Ministry of Economy and Finance

Department of the Treasury

Analisi e Programmazione
Economico Finanziaria

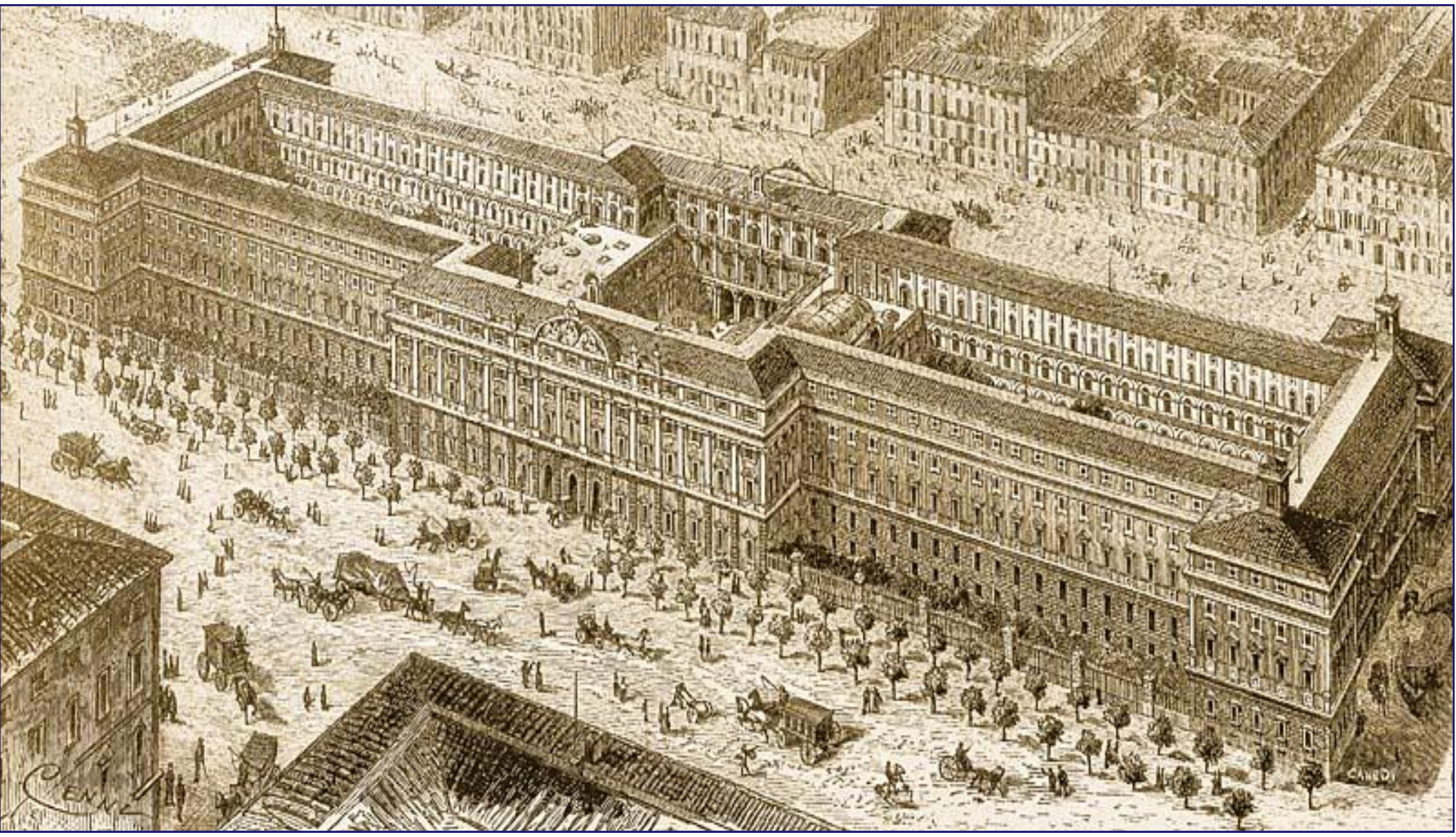
Working Papers

N°4 - February 2011

ISSN 1972-411X

Early-life environment, height and BMI of young adult males in Italy

Emilia Arcaleni, Franco Peracchi



Working Papers

The working paper series promotes the dissemination of economic research produced in the Department of the Treasury (DT) of the Italian Ministry of Economy and Finance (MEF) or presented by external economists on the occasion of seminars organised by MEF on topics of institutional interest to the DT, with the aim of stimulating comments and suggestions.
The views expressed in the working papers are those of the authors and do not necessarily reflect those of the MEF and the DT.

Early-life environment, height and BMI of young adult males in Italy[1]

Emilia Arcaleni (*), Franco Peracchi (**)

Abstract

This paper studies the relationship between the two main dimensions of early-life environment, namely disease burden (measured by infant mortality) and economic conditions (measured by income or consumption per capita), and height and body-mass index (BMI) of recent cohorts of young adult males in Italy. By combining high-quality micro-level data on height and weight with regional- and province-level information, we are able to link individual height and BMI at age 18 to regional and provincial averages of environmental variables in the year of birth.

Our results are consistent with the hypothesis that, in rich low-mortality setting, the scarring effects of childhood disease dominate selection. We also show that both income and disease matter, and their relative importance differs depending on the outcome considered and the available background information. In particular, we find that income matters more than disease for height, while the opposite is true for BMI. Finally, using detailed province-level information, we show that income per capita is a proxy for a variety of environmental indicators that are highly correlated with economic conditions.

JEL Classification: C31, C81, J13.
Keywords: Body height, BMI, obesity, income, infant mortality.

(*) Emilia Arcaleni , University of Rome "La Sapienza" and ISTAT.
(**) Franco Peracchi, Faculty of Economics, Tor Vergata University, via Columbia 2, 00133 Rome, Italy (franco.peracchi@uniroma2.it) and EIEF.

[1] We thank Luigi Cannari, Graziella Caselli, Istituto Guglielmo Tagliacarne, and the Italian National Statistical Institute (ISTAT) for the data, and Anne Case, Angus Deaton, Chris Paxson, Climent Quintana-Domeque and Alessandro Tarozzi for insightful discussions. Franco Peracchi also thanks the Center for Health and Wellbeing at Princeton University for hospitality during summer 2009.

CONTENTS

1 INTRODUCTION

Body height and body weight provide indicators of health and wellbeing that complement more conventional indicators, such as self-reported health, income or consumption. For this reason, they have long being of interest to demographers, nutritionists and public health researchers, and are now becoming increasingly popular also among economists.

Living conditions during the growing years, especially in early childhood, influence body height through their impact on net nutrition, namely the balance between the supply of nutrients and the demands of metabolism, physical exertion, and disease (Silventoinen 2003, Steckel 2009). Thus, adult height is a useful marker of the economic and disease environment in childhood. The relative abundance of data on adult height and the link between income and height have been exploited by economic historians to analyze the well-being of populations and historical periods for which other data sources are lacking (Fogel 1994). As argued by Deaton (2007), however, this link "is importantly contingent on the disease environment". Adult height is also a bridge between the present and the future, for it has been shown to be a strong predictor of earnings, cognitive function, and health outcomes at older ages, including longevity (Strauss and Thomas 2008, Case and Paxson 2008a, 2008b).

Body weight is a measure of mass. It is usually combined with other information, such as height. Unlike height, which remains stable during adult life until old age, body weight changes with the amount of fat and muscle in the body. The body-mass index (BMI), namely the ratio of weight to squared height (kg/m2), is just one way of combining weight and height into a single measure. Although not a direct measure of body fat, the BMI is considered a reliable proxy for total body fat in adults. The BMI is the most widely used diagnostic tool to identify obesity problems within a population, because it is easy to measure and calculate. As such, it has received a great deal of attention in the recent literature on the obesity epidemic and its economic and public health consequences (Cutler, Glaeser and Shapiro 2003, Philipson and Posner 2008, Brunello, Michaud and Sanz-de-Galdeano 2009). The relationship between early life-environment and weight and BMI has instead received less attention. Barker (1997) put forward the hypothesis that prenatal nutritional and disease experience is related to both small size and pathologies such as obesity and cardiovascular disease in later life. Power, Manor and Matthews (2003) find that gestation/early infancy for men and the period of "adiposity rebound" (5–7 years) for women are important periods during which socio-economic conditions may influence adult obesity. Kestil¨a et al. (2009) find that childhood circumstances, such as parental-education, parental unemployment and single-parent family, increase the risk of being overweight or obese in early adulthood (18–29 years of age). They also find that their effect is stronger on obesity than on overweight, and in women than in men.

Early-life environment not only affects height and possibly BMI, but also plays a prominent role in the development of cognitive and non-cognitive skills (Cuhna and Heckman 2007, Strauss and Thomas 2008, Currie 2009). Our paper does not consider problems of skill formation and focuses exclusively on physical development. It complements recent work by Akachi and Canning (2007), Bozzoli et al. (2009) and Bosch et al. (2009) who use country-cohort or region-cohort data to explore the relationship between adult height and the two main dimensions of early-life environment, namely income and disease, trying to establish their respective roles.

Akachi and Canning (2007) focus on mean cohort height of females born between 1945 and 1985 in Sub-Saharan Africa. They combine country-level averages of micro-data on height from the Demographic and Health Surveys (DHS) with country-level data on infant mortality, GDP per capita, and average protein and calorie consumption. Although country effects capture most of the observed variation in cohort height over time, they do find some role for the variables that describe the early-life environment.

Bozzoli et al. (2009) combine data on mean height of cohorts born between 1950 and 1980 in the USA (from the National Health Interview Survey) and 11 European countries (from the European Community Household Panel) with country-level data on infant mortality and GDP per capita. They find that postneonatal mortality (defined as death between one month and one year of age) is strongly negatively associated with height even after controlling for GDP per capita and country and year effects. They argue that "given that postneonatal mortality is a sensitive indicator of the disease environment in the first year of life, these results support accounts in which some form of 'scarring' in infancy negatively affects lifetime health, as marked by adult height. In [. . .] low mortality countries, the stunting effect of childhood disease dominates any possible height-based selective mortality in childhood that would induce a positive relationship between disease in early life and adult health". Based on these findings, they develop a formal model of selection and stunting in which the early life burden of nutrition and disease is not only responsible for mortality but leads to lower adult height and a higher risk of late-life disease. Their model predicts that at sufficiently high mortality levels selection can dominate scarring, leaving a taller population of survivors. Using data for developing countries from the DHS, they find evidence of this on female height in the poorest and highest mortality countries of the world.

Bosch et al. (2009) combine data on mean height of cohorts born between 1969 and 1999 in 17 Spanish regions (from the Spanish Health Survey) with region-level data on infant mortality and GDP per capita. They also find that infant mortality dominates the relationship between adult height and early-life environment. Focusing on a single country, as they do, has the important advantage of avoiding the problems associated with cross-country differences in the definition and measurement of the variables describing early-life environment, and the issue of what country effects stand for. Both these issues can hardly be ignored in cross-country studies.

Our paper examines the relationship between height and early-life environment for six annual cohorts of young Italian males born between 1973 and 1978. In addition to height, we also consider BMI and related measures, namely the fraction of people who are overweight (25 _ BMI < 30) or obese (BMI _ 30). Italy is an interesting case to consider because of its large and persistent regional differences in economic conditions, disease environment, access to health care, etc. These differences offer the possibility of separating the respective role of income and disease in a relatively homogeneous country.

Our height and weight data are drawn from the Italian military archives. These archives represent a rich and largely unexploited source of data that contains socio-demographic, anthropometric and health information on every young male who took the compulsory medical examination to ascertain fitness for military service. Compared to the survey data usually employed in the literature, our data offer several important advantages. First, due to the compulsory nature of military service in Italy during the period considered, our data provide a nearly complete coverage of a very large subset of the male population. Second, the size of the data is massive as we have, on average, over 300,000 individual records available each year. Third, in addition to height, we can examine weight and BMI. Fourth, anthropometric measures are not self-reported but are the outcome of a detailed physical inspection carried out by a medical team.

Our data are not without problems, however. One is potential selection effects due to the lack of data on the Navy draft. Other problems include missing records, rounding and heaping in measured height and weight, and editing errors in the available files. Most importantly, these are administrative data that lack the background information on a person typically available in household surveys. For example, we lack the information that would be necessary to estimate a second chapter height production function, nor we can follow individuals or cohorts overs time. Thus, our strategy consists of relating individual height and weight to the few available micro-level variables and to region- and province-level information on economic conditions and disease burden in the year of birth of each cohort.

The remainder of this paper is organized as follows: Section 2 describes the data, Section 3 presents some preliminary descriptive statistics, Section 4 presents the results of our regression analysis, and Section 5 concludes.

2 DATA

We use two types of data. For height and weight we use micro-data drawn from the military archives. These data contain information on all the Italian males who, between 1987 and 1996, took the compulsory medical examination that is part of the Army and Air Force draft process. Most of the information describing the early-life environment consists instead of aggregate data at the region- and province-level.

2.1 Height and weight

Our micro-data have been obtained by merging several electronic tapes produced by the Italian Ministry of Defense and stored at the Italian National Statistical Institute (Istat). These tapes record information on all the Italian males who were inspected, usually at age 18, to ascertain fitness for military service in the Army and Air Force. The tapes were produced in the late 1990s by a working group set up at Istat composed of Istat researchers, Army officials and university professors. The group dissolved after the sudden death of its leader, Floriano Pagnanelli. To our knowledge, this is the first time that these data have been used for scholarly research.

Our data offer several important advantages with respect to sample surveys. First, due to the compulsory nature of military service in Italy during the period considered, they provide nearly complete coverage of the population subject to the Army and Air Force draft.

Second, the size of the data is massive. On average we have more than 300,000 individualrecords each year, representing over 80 percent of each male cohort born between

1973 and 1978. Because of this large sample size, issues of statistical precision become of second order.

Third, in addition to height, we can examine weight, BMI and related measures, such as the prevalence of overweight and obesity. Further, the available measures are the outcomes of a detailede physical inspection by a medical team and do not suffer the distortions typically associated with self-reported measures. A general conclusion is that people tend to underreport their weight and overreport their height, resulting in underestimation of BMI and related measures (see e.g. Bolton-Smith et al. 2000).

Yet, our data are not perfect. First, because females were not subject to military conscription, we only have information on males. Second, age 18 is still a relatively early age and adult height may not have been reached by a non-negligible fraction of conscripts. This may be especially true for areas of the country lagging behind economically. Third, some features of the draft raise issues of selection bias.

During the period considered, the draft process was organized as follows. Each year, municipal authorities compiled a conscription list with the names of all resident male citizens who turned 17 that year. The following year, when draft operations started, the conscription list was updated, dropping people selected for the Navy draft or deceased, and adding people deferred from the previous draft or erroneously omitted, draft dodgers, and people who became Italian citizens in the meantime. Every young male on the list was expected to show up at the local conscription center (about one per province) for the medical inspection, or send a justification to avoid being declared a dodger. The inspection normally took place within three months from the 18th birthday. After the inspection, conscripts were assigned to three categories: fit for service, unfit for service, or deferred to the next draft. In a few special cases (e.g., residents abroad, people already enrolled in military bodies as volunteers, or people with a documented serious disability), the assignment was made without medical inspection.

Conscripts for the Navy were selected from coastal areas and among people with particular skills (shipyard work experience, fishing activities, or maritime studies). They were a nonnegligible fraction of each cohort (about 13‑15%), the fraction being higher in provinces with a maritime tradition. Taking the Navy draft into account, the size of our cohorts roughly matches Istat demographic data on cohort size by province. The available evidence suggests that excluding the Navy draft should not cause important biases, at least not for mean height. For example, using data for the 1928‑1929 birth cohorts, Arcaleni (1998) shows that Navy conscripts from Southern regions were slightly taller on average than the other conscripts (less than 2 mm) although, at the national level, there was no noticeable difference. These small regional differences may be attributed to a healthier diet (more fish) and better life conditions in coastal areas.

Fourth, these are administrative data and lack the rich personal and household-level information that is typically available in household surveys. In particular, they offer little information on earlylife environment. In addition, the nature and quality of the available micro-level information changes across cohorts. For those born between 1969 and 1972, the information is limited to place of birth (province and municipality, or foreign born), place of residence (province and municipality, or resident abroad), height (in cm), weight (in kg), and thoracic perimeter (in cm). For the cohorts born between 1973 and 1978 the available information also includes date of birth (day, month and year), marital status, number of children, educational attainment (highest degree attained and last class attended), type of job, reasons for cancelation from the conscription list, a synthetic health profile, and a measure of general intelligence.

Finally, there is evidence of rounding and heaping in individual measurements and there are missing data problems. In particular, we have a large fraction of missing values on height and weight for Italians residing abroad, a very large fraction of missing values on weight and thoracic perimeter and no information on age for the early cohorts (1969‐72), and a large fraction of missing data for conscripts from Sardinia born in 1974. Other minor issues include editing errors and a few duplicated records in the available data.

To reduce the impact of these problems, we confine attention to Italian residents born in Italy between 1973 and 1978, dropping individuals born abroad or foreign resident. We also use the information on the date of birth to select individuals aged 18 at the time of the medical inspection, thus avoiding spurious effects due to changes in the age composition of the sample (A' Hearn, Peracchi and Vecchi 2009). After discarding the few cases with missing values for age, place of birth, and place of residence, or with missing or implausible values for both height and BMI (namely height below 60 cm or above 250 cm, and BMI below 10 or above 70), we are left with a total sample consisting of over 300 thousands annual observations on the cohorts born between 1973 and 1978.

2.2 Early-life environment

To describe the early-life environment we use region- and province-level data on economic conditions and disease burden in the year of birth of the conscripts. These data refer to the Italian regions and provinces existing in the early 1970s. For consistency, we work with data broken down into 94 provinces, 18 regions and 5 macro-regions (or areas), namely North-West, North-East, Center, South, and Islands (Sicily and Sardinia). The average size of an Italian province is about 600 thousands inhabitans, the average size of an Italian region is about 5 times larger.

Our main indicator of economic conditions is income per capita in real terms. For regional incomes we use the series of real value added per capita at 1990 prices produced in the late 1990s by the Istituto per lo Sviluppo del Mezzogiorno (Svimez). For provincial incomes we use the series at 1999 prices produced in 2001 by the Istituto Guglielmo Tagliacarne (IGT). As a proxy of economic conditions we also consider private consumption per capita at 1990 prices, only available at the regional level from Svimez. All series use region-level prices to deflate nominal values. The main drawback of provincial data is that, being largely based on the decennial census, they are only available every ten years. The available years closest to the span of our height and weight data are 1971 and 1981. We use them to impute the provincial values for the years from 1973 to 1978 by interpolation, which may be justified by the low-frequency nature of the relationship linking height, weight and early-life environment.

Our main indicator of disease burden is infant mortality, namely the fraction of newborn males who die before their first birthday (per thousand) by region or province. Unfortunately, unlike Bozzoli et al. (2009), we cannot distinguish between neonatal and postneonatal mortality. Regionlevel data at the annual frequency have been obtained from Istat. Province-level data, taken from the database constructed by Graziella Caselli of the University of Rome "La

Sapienza", are only available at decennial intervals and are computed as averages over a 3-year period. Similarly to what was done for provincial incomes, we use the data for 1971‑73 and 1981‑83 to impute provincial values for the years 1973‑1978 by interpolation.

In addition to income per capita and infant mortality, we also consider province-level indicators of quality of the housing stock (the average number of people per room and the percentage of people living in houses without a toilet), ease of access to the health-care system (the number of hospital beds per 1,000 people), extra indicators of disease environment (the mortality of mothers, proxied by female mortality rates for the 26‑30 age group), and the number of cases of measles, mumps and rubella (MMR) and chickenpox, pertussis and scarlet fever (CPSF) per 1,000 inhabitants) and, following Thomas, Strauss and Henriques (1991), the educational attainments of mothers. Finally, as indicators of the educational attainments of mothers we use the fraction of women aged 6+ who hold at least a high school degree and the fraction of women aged 6+ who are illiterate, both drawn from the 1971 and 1981 censuses. Similarly to what was done for real incomes, we impute the values for the years 1973‑1978 by interpolation.

3 DESCRIPTIVE STATISTICS

The total number of available observations in our data falls steadily from about 372 thousands for the 1973 cohort to about 297 thousands for the 1978 cohort, a decline of about 20 percent. This trend is largely due to the shrinking size of the cohorts born after 1964, the year when the baby boom reached its peak in Italy. According to Istat demographic statistics, after decreasing slowly from 1964 to 1973, cohort size in Italy began falling swiftly from about 888 thousands life births in 1973 to about 721 thousands in 1978, a decrease of about 19 percent. This reflects the rapidly declining fertility of Italian women, a trend that during the period considered affects mainly the regions in the North and the Center, and is the result of the increasing share of women having at most two children. The delay in childbearing starts only in the late 1970s, and is not yet an issue for the period that we consider (Pinnelli, Hoffmann-Nowotny and Fux 2001).

Mean height at age 18 increases from 173.6 cm for the 1973 cohort to 173.9 cm for the 1975 cohort, and then falls slightly to 173.8 cm for the 1978 cohort. This pattern, which contrasts sharply with the rapid height growth experienced by the cohorts born in the 1950s and 1960s (Arcaleni 2006), mainly reflects three factors: a small increase of mean height in the Northern regions (equal to .25 cm in the North-West and .20 cm in the North-East) where conscripts are on average taller, a more sizeable increase of mean height in the Southern regions (equal to .38 cm in the South and .35 cm in the Islands) where conscripts are on average shorter, and a rising importance of the Southern regions, whose weight in the sample increased from 38.6 percent for the 1973 cohort to 42 percent for the 1978 cohort The fact that most of the height growth is observed for the cohorts born between 1973 and 1975 may be related to the high unemployment rates associated with the oil shock of 1973‑74, as babies conceived in times of high unemployment may have a better health on average and, consequently, may be taller and healthier in later life (Dehejia and Lleras-Muney 2004).

Mean weight at age 18 increases slightly more than mean height, from 68.3 kg for the 1973 cohort to 68.6 kg for the 1978 cohort. As a result, there is a small increase in mean BMI at age 18, from 22.6 for the 1973 cohort to 22.7 for the 1978 cohort. Again, this is mainly a compositional effect, due to the rising importance of the Southern regions, where conscripts have on average a higher BMI than in the North and the Center. What instead grows substantially is the variability of weight and BMI in our population, with the standard deviation of weight increasing from 10.5 to 11.5 kg, and the standard deviation of BMI increasing from 3.1 to 3.4. As a result of this increasing dispersion, the fraction of overweight conscripts increases from 15.4 percent for the 1973 cohort to 15.7 percent for the 1978 cohort, while the fraction of obese conscripts increases from 2.6 percent for the 1973 cohort to 3.8 percent for the 1978 cohort.

Table 1 - Mean and standard deviation of region-level variables describing early-life environment by birth cohort (infant mortality per thousand, income and consumption per capita in million Liras at 1990 prices)

	Infant mortality		Income per capita		Consumption per capita	
Cohort	Mean	Std.	Mean	Std.	Mean	Std.
1973	28.73	6.33	15.08	3.40	8.66	1.61
1974	25.26	5.45	14.75	3.15	8.70	1.65
1975	24.40	5.11	15.56	3.56	9.10	1.76
1976	22.51	3.89	16.00	3.58	9.43	1.83
1977	21.23	3.89	16.54	3.81	9.72	1.91
1978	19.43	3.61	17.49	4.11	10.39	2.02

Table 1 presents the (unweighted) means and standard deviations of our region-level indicators of early-life environment. Two opposite trends were at work for the cohorts considered. In the case of infant mortality, both the mean level and the variability between regions fell, as infant mortality dropped more in the South and the Islands where it was initially higher. For infant mortality, the mean fell by one third between 1973 and 1978 (from 28.7 to 19.4 deaths per thousand), whereas the between-region variation, measured by the standard deviation, fell by more than 40 percent (from 6.3 to 3.6 deaths per thousand). In the case of income and consumption per capita, instead, both the mean and the variability between regions increased, with the variability increasing faster than the mean. For example, the mean of real income per capita increased by 16 percent (from 15.1 to 17.5 million Liras at 1990 prices), whereas its standard deviation increased by 21 percent (from 3.4 to 4.1 million Liras). The pattern for consumption per capita, which represents about 60 percent of income per capita, is very similar. Thus, while Italian regions became more equal in terms of infant mortality, they became less equal in terms of income and consumption per capita.

Table 2 describes the correlation pattern of our variables at the region level. Height at age 18 is negatively correlated with BMI at age 18 and infant mortality, and strongly positively correlated with income and consumption per capita. BMI at age 18, on the other hand, is weakly positively correlated with infant mortality and negatively correlated with income and consumption per capita. There is also a substantial amount of correlation between the variables

Table 2 - Correlation patterns at the region level

	Height	BMI	Infant mortality	Consumption per capita
BMI	-.3450			
Infant mortality	-.4868	.2862		
Income pc	.7308	-.5070	-.6684	
Consumption pc	.7917	-.4641	-.6920	.9597

describing early-life environment. Income and consumption per capita are very strongly positively correlated (their correlation coefficient is above 90 percent). Further, infant mortality is negatively correlated with income and consumption per capita.

Turning to the province level, Figures 1 and 2 show the scatterplots of average provincial height at age 18 for the 1973 and 1978 cohorts against, respectively, infant mortality and real

Figure 1: Infant mortality at birth and mean height at age 18 by province

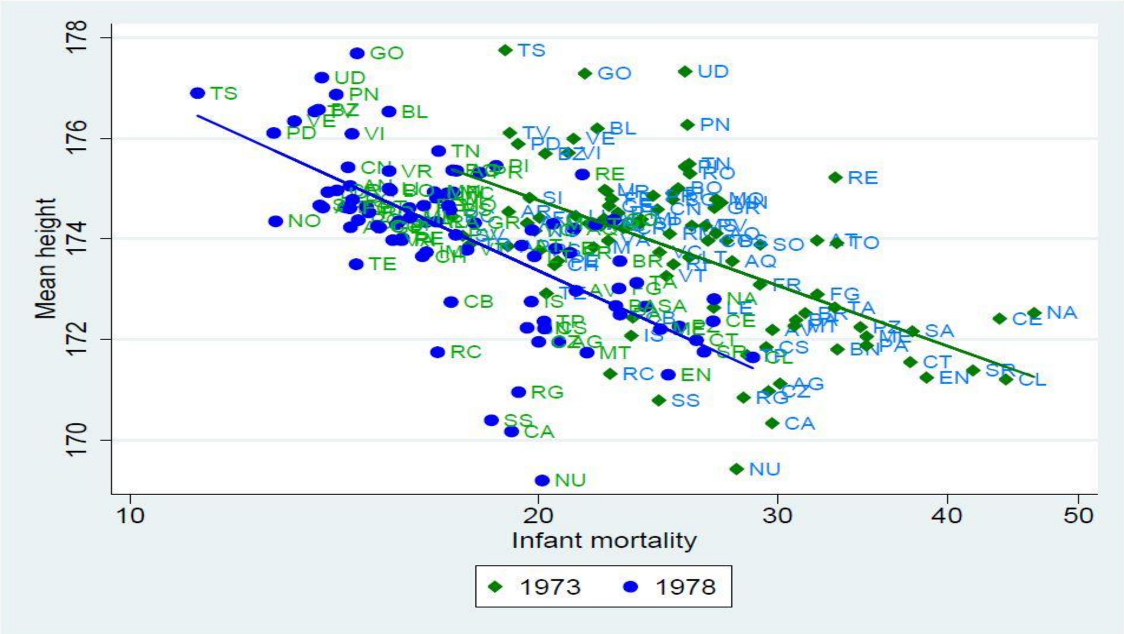


Figure 2: Income per capita at birth and mean height at age 18 by province

income per capita income (both on a log scale). Figures 3 and 4 show instead the scatterplots of average provincial BMI at age 18 against our two main environmental indicators. The solid lines are the fits from bivariate OLS regressions using province-level data. These figures show that only the bivariate cross-sectional relations of mean height with infant mortality and income per capita are relatively stable across cohorts. The figures also reveal the presence of strong differences between the Northern and the Southern provinces: the North has taller conscripts, lower BMI, lower infant mortality, and higher incomes; the South has shorter conscripts, higher BMI, higher infant mortality, and lower incomes.

Figure 3: Infant mortality at birth and mean BMI at age 18 by province

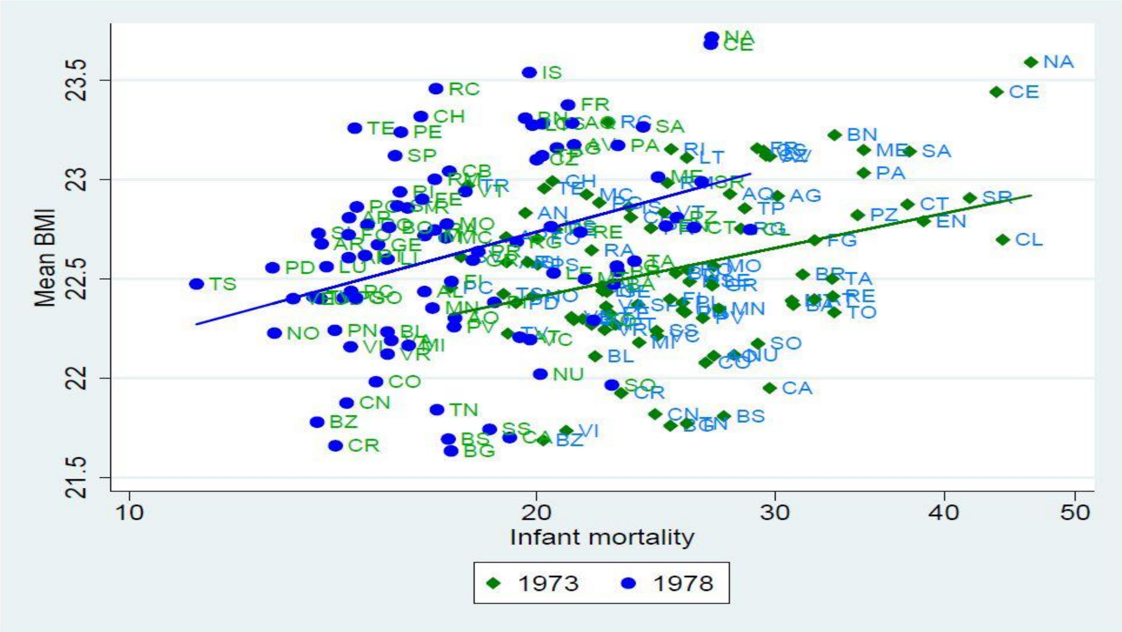
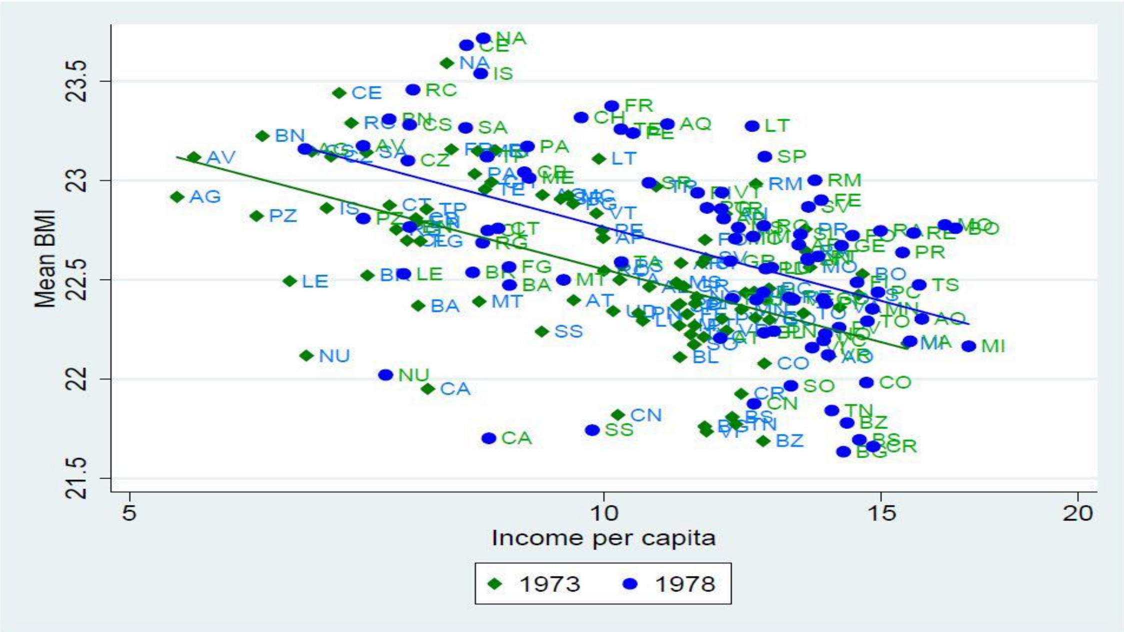


Figure 4: Income per capita at birth and mean BMI at age 18 by province



Finally, Table 3 presents the mean and standard deviation in 1971 and 1981 of the environmental variables at the province level and their contemporaneous correlation with income per capita. The main changes observed between 1971 and 1981 are a significant increase in incomes per capita (an increase of about 34 percent) and in the fraction of women with a high-school or a university degree, and a sharp decline in infant mortality, female mortality for the 26‒30 age group, the fraction of illiterate women, the fraction of people living in houses without a toilet, and the prevalence of infectious disease such as measles, mumps and rubella, but not chickenpox, pertussis and scarlet fever. The average number of hospital beds and the number of people per room change instead only slightly. At the same time we observe a substantial reduction of the between-province variation for most indicators except income per capita, the fraction of women with a high school degree, and the prevalence of chickenpox, pertussis and scarlet fever. As for the correlation between income per capita and the other environmental variables, this is highest (always above 50 percent in absolute value) for the number of people per room, the fraction of illiterate women, and the prevalence of chickenpox, pertussis and scarlet fever.

Table 3: Mean and standard deviation of environmental variables at the province level and their correlation with income per capita

	1971			1981		
	Mean	Std.	Corr.	Mean	Std.	Corr.
Income pc	9.65	2.30	1.0000	12.93	3.07	1.0000
Infant mortality (x1,000)	28.34	6.87	-.5293	13.50	3.00	-.4116
People per room	.94	.16	-.5810	.95	.14	-.6695
No toilet (%)	5.94	5.54	-.6011	1.91	1.54	-.4838
Hospital beds (x1,000)	10.69	4.42	.5055	9.90	2.89	.4659
Female mort. 26–30 (x1,000)	3.89	.83	-.3391	2.62	.62	-.0443
Prevalence of MMR (x1,000)	2.40	1.81	.4935	1.22	.95	.3774
Prevalence of CPSF (x1,000)	1.03	.90	.5908	1.55	1.27	.5618
Female HS grad. (%)	5.76	1.08	.2925	10.30	1.52	.3507
Female illiteracy (%)	7.10	5.69	-.7652	4.31	3.64	-.7544

4 REGRESSION ANALYSIS

In this section we model the mean of four outcomes‒individual height (in cm) and BMI at age 18, and indicators for being overweight or obese at that age‒conditional on information on economic conditions and disease burden in the year of birth of each cohort. Looking at the probability of being overweight or obese provides information on the conditional distribution of BMI beyond that contained in the conditional mean. We present results separately depending on whether environmental variables are measured at the region (Section 4.1) or province level (Section 4.2).

Our model for height and BMI is the linear regression

$$Y_{icr} = \alpha + \beta^{\top} X_{cr} + \gamma^{\top} W_{cr} + \delta^{\top} Z_{cr} + \tau_c + \mu_r + U_{icr}$$

where Y_{icr} is the outcome of interest for individual i of cohort c in region or province r, W_{cr} is the vector of focus regressors, namely the logarithms of infant mortality and income or consumption per capita in the year of birth of the individual, W_{cr} is a vector of additional region- or province-specific regressors, such as indicators of quality of the housing stock and ease of access to the health care system, Z_{cr} is a vector of individual specific regressors, such as indicators for the month of birth and the migration pattern, τ_c is a cohort trend, μ_r is an area or region fixed effect, U_{icr} is a random error assumed to be mean-independent of all variables on the right-hand side of the regression, and α, β, and γ are unknown parameters to be estimated by OLS. When the outcome is the indicator for being overweight or obese, we use a logit specification instead of linear regression, and we estimate the model parameters by maximum likelihood instead of OLS. All standard errors (reported in parentheses under the point estimates) are robust to heteroskedasticity of unknown form and are clustered geographically by municipality.

4.1 Region-level environmental variables

Tables 4-7 show the results for the case when the environmental variables are measured at the region-level. Results are presented for four specifications. The baseline specification (1) includes as regressors only the constant term and the logarithms of infant mortality and real income per capita in the year of birth. Specification (2) instead uses the logarithm of private consumption per capita as a measure of economic conditions. Specifications (3) and (4) are analogs of (1) and (2) with added dummies for year and month of birth, area of birth, birth in a maritime province (to partly control for the Navy draft), and migration pattern. Our set of dummies for migration pattern distinguishes between changes of province within the same region, changes of region within the same area, and changes from one area to another. Ignoring the dummies for month of birth and migration pattern, the models for height and BMI may be interpreted as regressing the deviation of an individual outcome from its area-level average on the deviations of region-level predictors on their area-level averages. Dividing by 10 the coefficients in the tables, gives an estimate of the difference in the mean value (or in the log-odds) of an outcome associated with a 10 percent change in one of the focus regressors, holding all the other regressors constant.

For the height and BMI regressions, the statistics reported at the bottom of each table are the sample size, the number of regressors excluding the constant term in each model, the adjusted, R^2 the root mean squared error (RMSE), and the F-statistic for the significance of the regression.

For the logit models, the last three statistics are replaced by the pseudo R^2 the maximized log likelihood, and the chi-square statistic for the significance of the regression.

Deaton (2007) and Bozzoli et al. (2009) argue that in richer low-mortality settings, such as contemporary Italy, the effect of infant mortality on height is likely to be negative as the stunting effect dominates the selection effect. The results in column (1) of Table 4 support both claims as they indicate that a 10 percent reduction in infant mortality, other things being equal, is associated with an increase of about .10 cm in mean height. However, this effect is small compared to the increase of about .40 cm associated with a 10 percent increase in income per capita. Measuring economic conditions by consumption rather than income per capita, as is done in column (2), further weakens the effect of infant mortality and increases the effect of economic conditions. When we include our full set of dummies (columns (3) and (4)), the coefficient on consumption remains positive and sizeable, while the coefficient on infant mortality remains negative and statistically significant, no matter what measure of economic conditions is used (although a 10 percent reduction in infant mortality now translates into only a .04 cm increase in mean height), but the coefficient on income becomes negative and looses statistical significance.

To save space, we do not present the coefficients on the other variables in the specifications (3)-(4) and only briefly report on their pattern. People born in a maritime province are slightly taller on average. The coefficients on the area dummies are small and not statistically significant for the North-East and the North-West, but are large and negative for the South and the Islands. Monthof-birth effects agree with the evidence reviewed by Bogin (1999, pp. 295-296), that is, they are highest for the months of May and June, lower for the fall months, and even negative for the winter months. These effects are not well understood yet, but may be related to sunlight influencing the patterns of human growth during the late fetal and early postnatal period. They may also reflect variation in mother and child diet, especially with respect to mineral and vitamin-rich fruits and vegetables that yield micronutrients essential to successful early childhood development. Relative to stayers, those who moved from low- to high-stature areas (e.g. from the South to the North) are on average taller, while those who moved from high- to low-stature areas are on average shorter.

This may reflect both area-specific effects and features of within-country migration during the period considered, especially the fact that the 1970s and 1980s are characterized by the end of the migration flows from the South to the North and by return migration from the North to the South.

Turning to the BMI regressions (Table 5), the relative importance of economic conditions and disease burden changes substantially depending on whether or not we include area, cohort, monthof-birth and migration dummies. When we exclude these dummies (specifications (1) and (2)), the coefficients on infant mortality are positive but small and not statistically significant, whereas the coefficients on economic conditions are large, negative and strongly statistically significant. When we instead include the dummies (specifications (3) and (4)), the relative importance of the two focus regressors is much more similar. For example, specification (3) implies that a 10 percent decrease in infant mortality has about the same effect as a 10 percent increase in income per capita, namely a reduction of the BMI by about .12-.13 points. The relative importance of disease burden is even more important in specification (4), where the coefficient on consumption per capita remains negative but is much smaller than in specification (2) and no longer statistically significant. As for the coefficients on the other variables in specifications (3) and (4), people born in maritime provinces have on average a slightly higher BMI. The coefficients on the area dummies show that, after controlling for all other things, the BMI is on average lower in the North and the Islands. Cohort effects are positive and strongly statistically significant, whereas month-of-birth effects show no clear pattern. Relative to stayers, those who moved from low- to high-BMI areas (e.g. from the North to either the Center or the South) have on average a higher BMI, while those who moved from high- to low-BMI areas have on average a lower BMI. As already mentioned, this may reflect both environmental effects and feature of within-country migration during the period considered.

Table 4: Height regressions with region-level environmental variables
(observed p-values: * p < :10; ** p < :05; *** p < :01)

	(1)	(2)	(3)	(4)
Log infant mortality	-.994 ***	-.288	-.422 *	-.384 *
	(.245)	(.267)	(.235)	(.226)
Log income pc	3.704 ***		-.647	
	(.185)		(.583)	
Log consumption pc		5.997 ***		1.737 **
		(.330)		(.693)
Area dummies	No	No	Yes	Yes
Cohort dummies	No	No	Yes	Yes
Month-of-birth dummies	No	No	Yes	Yes
Migration dummies	No	No	Yes	Yes
Sample size	1974286	1974286	1974286	1974286
No. obs.	2	2	45	45
Adj. R^2	.0259	.0289	.0382	.0382
RMSE	6.67	6.66	6.63	6.63
F-stat.	266	291	134	133

Table 5: BMI regressions with region-level environmental variables
(observed p-values: * p < :10; ** p < :05; *** p < :01)

	(1)	(2)	(3)	(4)
Log infant mortality	.160	.188	1.338 ***	1.401 ***
	(.133)	(.161)	(.184)	(.185)
Log income pc	-1.311 ***		-1.217 ***	
	(.118)		(.231)	
Log consumption pc		-1.621 ***		-.502
		(.159)		(.373)
Area dummies	No	No	Yes	Yes
Cohort dummies	No	No	Yes	Yes
Month-of-birth dummies	No	No	Yes	Yes
Migration dummies	No	No	Yes	Yes
Sample size	1746694	1746694	1746694	1746694
No. obs.	2	2	45	45
Adj. R^2	.0123	.0104	.0199	.0195
RMSE	3.22	3.22	3.21	3.21
F-stat.	62.5	53	41.6	41.4

The results from the logit models for being overweight (Table 6) or obese (Table 7) are consistent with the findings from the BMI regressions. Even after controlling for our full set of dummies, the coefficients on infant mortality are positive and statistically significant, while the coefficients on income and consumption are negative but are statistically significant only in the case of income. Finally, we again observe strongly positive and statistically significant cohort trends, especially for the probability of being obese.

Table 6: Logit models for the probability of being overweight with region-level environmental variables (observed p-values: * p < :10; ** p < :05; *** p < :01)

	(1)	(2)	(3)	(4)
Log infant mortality	.154 **	.172 **	.792 ***	.823 ***
	(.073)	(.087)	(.093)	(.093)
Log income pc	-.732 ***		-.791 ***	
	(.071)		(.148)	
Log consumption pc		-.910 ***		-.279
		(.091)		(.227)
Area dummies	No	No	Yes	Yes
Cohort dummies	No	No	Yes	Yes
Month-of-birth dummies	No	No	Yes	Yes
Migration dummies	No	No	Yes	Yes
Sample size	1746694	1746694	1746694	1746694
No. obs.	2	2	45	45
Pseudo R^2	.00692	.00603	.0109	.0106
Max. log lik.	-748449	-749123	-745484	-745711
χ^2-stat.	113	115	1603	1600

Table 7: Logit models for the probability of being obese with region-level environmental variables (observed p-values: * p < :10; ** p < :05; *** p < :01)

	(1)	(2)	(3)	(4)
Log infant mortality	.146 *	.188 *	1.097 ***	1.149 ***
	(.081)	(.101)	(.100)	(.103)
Log income pc	-.867 ***		-1.185 ***	
	(.069)		(.251)	
Log consumption pc		-1.041 ***		-.324
		(.108)		(.331)
Area dummies	No	No	Yes	Yes
Cohort dummies	No	No	Yes	Yes
Month-of-birth dummies	No	No	Yes	Yes
Migration dummies	No	No	Yes	Yes
Sample size	1746694	1746694	1746694	1746694
No. obs.	2	2	45	45
Pseudo R^2	.00677	.0057	.0128	.0123
Max. log lik.	-253091	-253363	-251547	-251670
χ^2-stat.	161	116	1703	1738

4.2 Province-level environmental variables

To better understand the role of our two focus regressors, namely infant mortality and income per capita, we now consider regressions with environmental variables measured at the finer province level. Tables 8–11 show the results obtained for five specifications. The baseline specification in column (1), which is directly comparable to specification (1) in the previous section, includes as predictors only the constant term and our focus regressors. Specification (2) adds to (1) a set of dummies for the region of birth, the year and month of birth, and the migration pattern, and is directly comparable to specification (3) in the previous section. Specifications (3)‑(5) add to (2) an increasing set of variables that are intended to describe in more detail the characteristics of the early-life environment at the province level. Thus, specification (3) adds to (2) indicators for the quality of the housing stock, namely the average number of people per room and the percentage of people living in houses without a toilet. Specification (4) adds to (3) the number of hospital beds and, as additional indicators of disease environment, female mortality in the 26‑30 age group, and the number of cases of measles, mumps and rubella (MMR) and chickenpox, pertussis and scarlet fever (CPSF) for 1,000 inhabitants. Finally, specification (5) adds to (4) indicators for the educational attainments of mothers, proxied by the fraction of women aged 6+ who are illiterate nd the fraction of women aged 6+ who hold a high-school or a university degree. In the case of height and BMI, the specifications in columns (2)‑(5) may be interpreted as regressing the deviation of an individual outcome from its region-level average on the deviations of province-level predictors on their region-level averages.

Table 8: Height regressions with province-level environmental variables
(observed p-values: * p < :10; ** p < :05; *** p < :01)

	(1)	(2)	(3)	(4)	(5)
Log infant mortality	-1.319 ***	-.356	-.357	-.339	-.215
	(.244)	(.223)	(.237)	(.237)	(.230)
Log income pc	3.082 ***	.741 **	.301	.392	-.119
	(.245)	(.299)	(.371)	(.357)	(.345)
No toilet			-.038 **	-.043 **	-.032 *
			(.017)	(.017)	(.019)
People per room			.013	.132	.112
			(.421)	(.463)	(.494)
Hospital beds				0.013	0.007
				(0.012)	(0.012)
Prevalence of MMR				-.003	.010
				(.041)	(.040)
Prevalence of CPSF				-.130 **	-.144 **
				(.063)	(.064)
Female mort. 26–30				-.103 **	-.091 *
				(.051)	(.049)
Female HS grad.					.049 **
					(.023)
Female illiteracy					-.009
					(.016)
Region dummies	No	Yes	Yes	Yes	Yes
Cohort dummies	No	Yes	Yes	Yes	Yes
Month-of-birth dummies	No	Yes	Yes	Yes	Yes
Migration dummies	No	Yes	Yes	Yes	Yes
Sample size	1974286	1974286	1974286	1974286	1974286
No. obs.	2	60	62	66	68
Adj. R^2	.0262	.0425	.0426	.0427	.0428
RMSE	6.67	6.62	6.62	6.62	6.62
F-stat.	173	161	163	152	151

For all outcomes considered, the results from specifications (1) tend to agree in sign, magnitude and statistical significance with those from specifications (1) in the previous section, so we focus on the results from the other four specifications.

In the case of height, specification (2) in Table 8 produces very similar results to specification (3) in Table 4. Adding indicators for the quality of the housing stock, the number of hospital beds and extra indicators of the disease environment does not change much the size and statistical significance of the coefficients on infant mortality but essentially "kills" the role of income per capita. This result is important because it shows that income per capita is really a proxy for a variety of environmental indicators that are highly correlated with economic conditions. In particular, we find a strong and significant negative association between height and the percentage of people without a toilet at home, the prevalence of chickenpox, pertussis and scarlet fever, and female mortality in the 26–30 age group. Consistently with the results in Thomas, Strauss and Henriques (1991), we also find that height is positively associated with the fraction of women with a high-school degree and negatively associated with the fraction of women who are illiterate.

Table 9: BMI regressions with province-level environmental variables (observed p-values: * p < :10; ** p < :05; *** p < :01)

	(1)	(2)	(3)	(4)	(5)
Log infant mortality	.347 **	.113	.163	.198 **	.207 **
	(.158)	(.089)	(.099)	(.092)	(.088)
Log income pc	-.968 ***	.373 ***	.344 **	.249	.206
	(.095)	(.131)	(.164)	(.153)	(.152)
No toilet			-.005	-.007	-.011 *
			(.006)	(.007)	(.006)
People per room			-.174	-.191	-.224
			(.194)	(.208)	(.218)
Hospital beds				-0.003	-0.001
				(0.004)	(0.004)
Prevalence of MMR				-.035 **	-.029 **
				(.016)	(.015)
Prevalence of CPSF				.060 **	.058 **
				(.025)	(.025)
Female mort. 26–30				-.012	-.014
				(.023)	(.023)
Female HS grad.					.006
					(.010)
Female illiteracy					.007
					(.007)
Region dummies	No	Yes	Yes	Yes	Yes
Cohort dummies	No	Yes	Yes	Yes	Yes
Month-of-birth dummies	No	Yes	Yes	Yes	Yes
Migration dummies	No	Yes	Yes	Yes	Yes
Sample size	1746694	1746694	1746694	1746694	1746694
No. obs.	2	60	62	66	68
Adj. R^2	.0106	.0253	.0253	.0254	.0254
RMSE	3.22	3.2	3.2	3.2	3.2
F-stat.	54.1	116	114	114	106

In the case of BMI (Table 9), while the results for specification (1) agree with those in Table 5, those for specification (2) (which replaces area dummies with region dummies) do not. They main difference is the fact that the coefficient on income per capita remains statistically significant but becomes much smaller and switches from negative to positive. Introducing additional indicators of early-life environment (the last three specifications) does not change the sign (positive) and the size of the coefficient on infant mortality, which actually becomes statistically significant in specifications (4) and (5), but again "kills" the role of income per capita. Results that appear to be robust to alternative specifications are the positive association between average BMI and the prevalence of chickenpox, pertussis and scarlet fever, and the negative association between BMI and the average number of people per room and the prevalence of MMR. We find similar patterns in the case of the logit models for being overweight (Table 10) or obese (Table 11). In both cases, replacing area dummies with region dummies leads to substantial changes in the sign and magnitude of the estimated coefficients on the focus regressors relative to the corresponding specifications in the previous section. In particular, comparing the various specifications in Tables 10 and 11 shows that income is never statistically significant in the specifications that include our full set of regressors.

Table 10: Logit models for the probability of being overweight with province-level environmental variables (observed p-values: * p < :10; ** p < :05; *** p < :01)

	(1)	(2)	(3)	(4)	(5)
Log infant mortality	.271 ***	.089 *	.117 *	.124 **	.145 ***
	(.084)	(.050)	(.060)	(.057)	(.055)
Log income pc	-.529 ***	.214 ***	.226 **	.184 **	.105
	(.059)	(.071)	(.089)	(.085)	(.087)
No toilet			-.000	.000	-.000
			(.003)	(.003)	(.003)
People per room			-.086	-.079	-.097
			(.109)	(.118)	(.121)
Hospital beds				-0.000	-0.000
				(0.002)	(0.003)
Prevalence of MMR				-.017 *	-.013
				(.009)	(.009)
Prevalence of CPSF				.036 **	.034 **
				(.014)	(.014)
Female mort. 26–30				-.000	.001
				(.013)	(.013)
Female HS grad.					.009
					(.005)
Female illiteracy					.002
					(.003)
Region dummies	No	Yes	Yes	Yes	Yes
Cohort dummies	No	Yes	Yes	Yes	Yes
Month-of-birth dummies	No	Yes	Yes	Yes	Yes
Migration dummies	No	Yes	Yes	Yes	Yes
Sample size	1746694	1746694	1746694	1746694	1746694
No. obs.	2	60	62	66	68
Pseudo R^2	.0061	.0137	.0137	.0137	.0138
Max. log lik.	-746216	-740497	-740494	-740480	-740471
χ^2-stat.	98.8	4974	5057	5211	5217

Table 11: Logit models for the probability of being obese with province-level environmental variables (observed p-values: * p < :10; ** p < :05; *** p < :01)

	(1)	(2)	(3)	(4)	(5)
Log infant mortality	.231 **	.105	.201 **	.251 ***	.203 **
	(.102)	(.081)	(.087)	(.087)	(.086)
Log income pc	-.662 ***	.133	.101	-.026	.132
	(.061)	(.123)	(.150)	(.148)	(.163)
No toilet			-.005	-.007	-.008
			(.005)	(.005)	(.005)
People per room			-.279 *	-.293 *	-.265
			(.158)	(.159)	(.167)
Hospital beds				-0.007	-0.006
				(0.005)	(0.005)
Prevalence of MMR				-.042 **	-.050 ***
				(.018)	(.019)
Prevalence of CPSF				.091 ***	.095 ***
				(.026)	(.027)
Female mort. 26–30				-.017	-.020
				(.020)	(.020)
Female HS grad.					-.016
					(.012)
Female illiteracy					-.001
					(.005)
Region dummies	No	Yes	Yes	Yes	Yes
Cohort dummies	No	Yes	Yes	Yes	Yes
Month-of-birth dummies	No	Yes	Yes	Yes	Yes
Migration dummies	No	Yes	Yes	Yes	Yes
Sample size	1746694	1746694	1746694	1746694	1746694
No. obs.	2	60	62	66	68
Pseudo R^2	.00581	.0164	.0165	.0166	.0166
Max. log lik.	-253328	-250628	-250616	-250586	-250578
χ^2-stat.	136	4749	4665	4713	4689

5 CONCLUSIONS

In this paper we studied the relative importance of two dimensions of early-life environment, namely disease burden (measured by infant mortality) and economic conditions (measured by income or consumption per capita), for various health outcomes, namely height, BMI and related measures, using data on young adults in Italy.

Our results are consistent with the evidence in Bozzoli et al. (2009) that scarring dominates selection in rich low-mortality settings. We also show that both disease burden and economic conditions matter, and their relative importance differs depending on the outcome considered and the available background information. In particular, economic conditions appear to matter more than disease burden for height, while the opposite is true for BMI and the probability of being overweight or obese. Finally, using detailed province-level information, we show that income per capita is a proxy for a variety of environmental indicators that are highly correlated with economic conditions. Among these, particularly important appear to be the quality of the housing stock and the incidence of infectious diseses such as chickenpox, pertussis and scarlet fever.

REFERENCES

A'Hearn B., Peracchi F., and Vecchi G. (2009), "Height and the normal distribution: evidence from Italian military data." Demography, 46: 1-22.

Akachi Y., and Canning D. (2007), "The height of women in Sub-saharan Africa: The role for health, nutrition, and income in childhood." Annals of Human Biology, 34: 397-410.

Arcaleni E. (1998), "La statura dei coscritti italiani delle generazioni 1854-1976." Bollettino di Demogra_a Storica, 29: 23-59.

Arcaleni E. (2006), "Secular trend and regional differences in the stature of Italians, 1854-1980." Economics & Human Biology, 44: 24-38.

Barker D.J.P. (1997), "Maternal nutrition, fetal nutrition, and disease in later life." Nutrition, 13: 807-813.

Bogin B. (1999), Patterns of Human Growth (2nd ed.), Cambridge University Press, Cambridge, UK.

Bolton-Smith C., Woodward M., Tunstall-Pedoe H., and Morrison C. (2000), "Accuracy of the estimated prevalence of obesity from self reported height and weight in an adult Scottish population." Journal of Epidemiology and Community Health, 54: 143-148.

Bosch M., Bozzoli C., and Quintana-Domeque C. (2009), "Infant mortality, income and adult stature in Spain." Department of Economics, University of Alicante, mimeo.

Bozzoli C., Deaton A., and Quintana-Domeque C. (2009), "Adult height and childhood disease." Demography, 46: 647-669.

Brunello G., Michaud P.-C., and Sanz-de-Galdeano A. (2009), "The rise of obesity in Europe: An economic perspective (with discussion)." Economic Policy, 24: 551-596.

Case A., and Paxson C.H. (2008a), "Height, health and economic development." American Economic Review, 98: 463-467.

Case A., and Paxson C.H. (2008b), "Stature and status: Height, ability and labor market outcomes." Journal of Political Economy, 116: 499-532.

Cunha F., and Heckman J. (2007), "The technology of skill formation." American Economic Review, 97: 31-47.

Currie J. (2009), "Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development." Journal of Economic Literature, 47: 87-122.

Cutler D.M., Glaeser E.L., and Shapiro J.M. (2003), "Why have Americans become so obese." Journal of Economic Perspectives, 17: 93-118.

Deaton A. (2007), "Height, health, and development." Proceedings of the National Academy of Sciences, 104: 13232-13237.

Dehejia, R., and Lleras-Muney A. (2004), "Booms, busts, and babies' health." Quarterly Journal of Economics, 119: 1091-1130.

Fogel, R. (1994), "Economic growth, population theory, and physiology: The bearing of long-term processes on the making of economic policy." American Economic Review, 84: 369-395.

García Villar J., and Quintana-Domeque C. (2009), "Income and body mass index in Europe." Economics & Human Biology, 7: 73-83.

Kestilä L., Rahkonen O., Martelin T., Lahti-Koski M., and Koskinen S. (2009), "Do childhood social circumstances affect overweight and obesity in early adulthood?" Scandinavian Journal of Public Health, 37: 206-219.

Philipson T., and Posner R. (2008), "Is the obesity epidemic a public health problem? A decade of research in the economics of obesity." NBER Working Paper 14010.

Power C., Manor O., and Matthews S. (2003), "Child to adulthood socioeconomic conditions and obesity in a national cohort." International Journal of Obesity, 27: 1081-1086.

Pinnelli A., Hoffmann-Nowotny H.J., and Fux B. (2001), "Fertility and new types of households and family formation in Europe." Population Studies No. 35, Council of Europe, Strasbourg.

Silventoinen K. (2003), "Determinants of variation in adult body height." Journal of Biosocial Science, 35: 263-285.

Steckel R.H. (2009), "Heights and human welfare: Recent developments and new directions." Explorations in Economic History, 46: 1-23.

Strauss J.A., and Thomas D. (2008), "Health over the life course." In Schultz T.P. and Strauss J.A. (eds.), Handbook of Development Economics, Elsevier, 3375-3474.

Thomas D., Strauss J.A., and Henriques M.-H. (1991), "How does mother's education affect child height?" Journal of Human Resources, 26: 183-211.

Ministry of Economy and Finance

Department of the Treasury

Directorate I: Economic and Financial Analysis

Address:
Via XX Settembre, 97
00187 - Rome

Websites:
www.mef.gov.it
www.dt.tesoro.it

e-mail:
dt.segreteria.direzione1@tesoro.it

Telephone:
+39 06 47614202
+39 06 47614197

Fax:
+39 06 47821886