Università degli Studi di Torino

Vilfredo Pareto Doctorate in Economics

Essays on Economics of Health and Education

Sarah Zaccagni

2016-2019

Università degli Studi di Torino

Dipartimento di Scienze Economico-Sociali e Matematico-Statistiche

Vilfredo Pareto Doctorate in Economics

XXXII Cycle

Essays on Economics of Health and Education

Supervisor:

Marina Di Giacomo

Co-Supervisor:

Massimiliano Piacenza

Candidate:

Sarah Zaccagni

Coordinator: Fabio Cesare Bagliano

2016-2019

Contents

List of Tables

List of Figures

Declaration of Co-Authorship

The second chapter of this thesis is a joint work with Marina Di Giacomo (University of Turin), Luca Pieroni (University of Perugia), and Luca Salmasi (Catholic University of the Sacred Heart).

The third chapter of this thesis is a joint work with Ainoa Aparicio Fenoll (University of Turin and Collegio Carlo Alberto), and Flavia Coda-Moscarola (CeRP-Collegio Carlo Alberto and Ufficio Studi Compagnia di San Paolo).

Introduction

This thesis collects three contributions: the first and the second mainly refer to the field of Health Economics, the third to that of Economics of Education.

In the first chapter we analyze how risk attitude changes when parents make decisions that affect their children's health rather than their own. In several situations individuals make health decisions on behalf of others: the case of parents and children is one of them. We use data from the German Socio-Economic Panel (SOEP) survey to test the theoretical prediction that parental risk aversion has a positive impact on the investment in child's health. As explanatory variable, we use the answers to a question asking people about their willingness to take risks in general. We estimate a model using children's Body Mass Index and incidence of obesity as dependent variables. We assume that these variables reflect family food habits and therefore express the result of parents' decisions regarding their children's health. We find that the answer of the parents to the question about risk taking in general is a good predictor of decisions that increase the risk of obesity for children. Being well-known in the literature the relationship between individual risk attitude and individual health behaviors, our contribution consists in the assessment of the stability of risk attitudes in health decisions made for oneself or on behalf of one's child.

The second chapter of this thesis focuses on hospital quality at regional level. Hospital quality measures are increasingly being used in the public domain to compare hospitals and constitute one of the main drivers of hospital choice at least for elective pathologies. Therefore, a rigorous investigation of hospital quality across regions is a useful tool to forecast patients' mobility and address the intervention of the policymaker. The aim of this paper is to revisit the two-step methodology proposed by Papanicolas and McGuire [2017] and apply it to the data from Piano Nazionale Esiti (PNE) for the years 2008-2016 to reach the following extents. The first is to extract regional fixed effects from yearly cross-sectional hospital-level data for three standard health quality measures for acute myocardial infarction (30-day mortality rate, 365-day readmission rate and the ratio of 30-day crude rate mortality for treated with Percutaneous Transluminal Coronary Angioplasty versus not treated). The second is to smooth the fixed effect obtained in the first step through a vector auto-regression in order to forecast the future quality of the regional healthcare systems in Italy.

The third chapter explores the theme of suitability of special programs for gifted students. The suitability of special programs for gifted students is highly controversial. In this chapter, we evaluate a mathematics camp for talented high-school students. The camp covers Mathematics topics outside the school curricula with "hands-on" learning style. We randomize participation in the camp and test the effects of such participation on problem-solving skills, self-concept, and career intentions. Potential participants are designated by their teachers. Results show that participants improve their problem-solving skills, specially in questions that regard logic as opposed to mechanical solutions. We also find positive short-run effects on students' self-concept. Students who benefit more from the program are those who have a lower math grade in the first semester. Finally, younger students who participate in the program are influenced in their choice to continue their studies after high-school.

CHAPTER 1

Parental Risk Preferences and Children's Health

1. Introduction

People are often in a position where either they make decisions on behalf of someone else or someone else is making decisions on their behalf. While there is an extensive literature which deals with the theme of individual decision-making, the literature on making decisions for other people is not large.

Situations in which people make decisions for others can be grouped into three types: those in which someone requests that the decision is made for them; those in which the decision is imposed on them; and those in which their condition requires that the decision is made for them. Health decisions belong to the last case and we refer to this as surrogate decision-making: in fact the person for whom the decision is made is unable to make the appropriate decision. This happens for example when doctors make treatment decisions on behalf of their patients, when parents make health decisions for children or even when adult children help elderly parents with healthcare decisions (Edwards and Elwyn [2009]).

The relationship between the individual who makes the decision and the one the decision has to be made for may be somehow assimilated to an agency problem. In the literature, few papers investigate the nature of this relationship and they mostly focus on the doctor-patient relationship. The aim of these papers is to understand what patients might want in terms of involvement in the decision-making process and to address the problem of asymmetric information between the two parties (Vick and Scott [1998]).Evidence shows that health-care consumers may experience disutility from being involved in the decisions about the appropriate treatment for their health problems; therefore non-involvement or asymmetric information may even be aimed rather than avoided like it happens in standard economics contexts (McKinstry [1992]).

When we refer to health problems, we must point out that health has many determinants. In industrialized countries, where morbidity and mortality are primarily related to chronic rather than infectious diseases, health behaviors - like smoking, drinking, diet, and physical activity - are particularly important Ezzati et al. [2004]. Among the determinants of individual health behaviors there are individual innate characteristics, such as attitude toward risk and time, and socio-cultural factors, such education, income, job position, etc. Therefore, individual health behaviors can be considered partially dependent on the intergenerational transmission of the innate component of risk and time preferences (Wickrama et al. [1999]), partially the result of habits molded by parents and partially influenced by the social, economic and cultural environment in which the individual lives. In the first part of one's life, health behaviors are determined by parents who behave as surrogate decision-makers for their children. For this reason, it is important to understand how parents make decisions for their children's health and to evaluate the results of these decisions in terms of children's health outcomes.

In this paper, we focus on the parent-child agency relationship. In particular, we investigate the effect of the decision-maker's willingness to take risks on her decisions regarding her child's health. Risk attitude is important because almost every decision involves risk: not only economic decisions but also decisions regarding education and health investment or occupational choice. One important open question concerns how to measure attitude toward risk. Experimental studies (Andersen et al. [2006], Andersen et al. [2008]), which measure risk-taking behavior with real money at stake, offer an incentive-compatible measure of risk attitudes. However, this technique is costly and difficult to perform with a large, representative sample, preventing large-scale studies. On the other hand, survey questions are not incentive compatible, and it is difficult to say whether self-reported personal attitudes and traits are behaviorally meaningful. Various factors in fact, including self-serving biases, inattention, and strategic motives could cause respondents to distort their reported risk attitudes.

We address this problem using data from the German Socio-Economic Panel (SOEP), which measures the risk attitudes of more than 22,000 individuals. In the SOEP, one question directly asks individuals to make a global assessment of their willingness to take risks ("How willing are you to take risks, in general?"). Respondents rate their willingness on a scale from 0 to 10. This measure has been validated by Dohmen et al. [2011b], who conduct a field experiment with a representative sample of 450 subjects from the SOEP sample. Participants who attend the experiment made choices in a real-stakes lottery. Comparing the answer to the SOEP direct question to the experimental result, Dohmen et al. [2011b] found that responses to the general risk question are a reliable predictor of actual risky behavior, even controlling for a large number of observables.

Relying on the behavioral validity of this question, we test the effect of risk attitude on investment on children's health. We present a model of utility maximization, in which children's utility function enters parent's value function according to a specific rate of altruism (Doepke and Zilibotti [2017]). From the model we obtain the prediction that a higher willingness to take risks results in a higher investment in children's health. As health outcome for the children we use their Body Mass Index (henceforth BMI): we assume in fact that BMI during childhood is determined by food habits, that parents impose on their children. In the choice of the health outcome, we follow the approach of Weller et al. [2008] and Chabris et al. [2008], who both focus on nutrition as health outcome potentially affected by individual preferences in adult subjects.

The organization of the paper is as follows. Section 2 is dedicated to a brief literature review of the three important literature streams to which this paper refers: the determinants of individual health decision making; decision-making on behalf of others in healthcare and the role of parenting in deciding for and transmitting preferences to children. Section 3 is devoted to outline the general model of parent's value maximization over the investment in health for their children.

In Section 4, after a description of the SOEP dataset and of the measures that we use, we present the empirical methodology. Section 5 shows the results obtained from the empirical analysis, including results from additional collateral analyses.

Section 6 discusses the implications of our results, highlighting some critical aspects, and then we move to the conclusions, in Section 7.

2. Literature Review

2.1. Determinants of individual health decisions. In industrialized countries, the role of health behavior as a determinant of the onset of chronic diseases in increasingly gaining attention. Several studies which analyze the deaths related to modifiable risk in US show that between 2000 and 2010 about 40 percent of deaths were related to smoking, diet, physical activity, and alcohol (Ezzati et al. [2004]). Even if these estimates are potentially biased by the incidence of other confounding factors, it is uncontested that modifiable behaviors represent an important determinant of both mortality and morbidity. It is interesting to look at the trends in health behaviors because the results are mixed. First of all, individual health behaviors are more relevant in modern industrialized economies, while in poorer nations infectious diseases and environmental risks play a greater role (Singh and Singh [2008]). Even if we restrict the view to high income countries from 1970 up to now, we see a trend toward healthier behaviors on some dimensions but not others. The fact that is not possible to define an unequivocal conclusion makes more urgent the study of health behaviors.

The foundation for economics research on health behaviors is the model of health capital developed by Grossman [1972]. In his model it is assumed that individuals inherit an initial stock of health that depreciates over time at an increasing rate and can be increased by investment. Death occurs when the stock falls below a certain level (i.e. lifetime is not fixed but is an endogenous variable and no source of uncertainty occurs). Gross investments in health capital are produced by household production functions whose direct inputs include the own time of the consumer and market goods such as medical care, diet, exercise, recreation, and housing (Becker [1978]). The last assumption is that more educated people are more efficient producers of health: schooling may improve health by enhancing allocative efficiency (participation in healthier behaviors) or productive efficiency (obtaining more health from the same set of inputs). Health is demanded by consumers for two reasons. As a consumption commodity, it directly enters their preference functions. As an investment commodity, it determines the total amount of time available for market and non-market activities. This model represents the first attempt to shape the demand for the good "good health" and explains variations in health among persons as variations in supply and demand for health capital.

To understand these variations in the demand for health, it is important to take into account individual time and risk preferences. One of the first economists to examine the relationship between rate of time preference and health behaviors was Fuchs [1980], who argues that the correlation of education with good health could reflect differences in rate of time preference. Fuchs finds that a more patient rate of time preference (elicited from questions about willingness to exchange a certain amount of money today for a larger amount in the future) is associated with greater schooling and usually also with healthier behaviors, although the point estimates are often small and not always statistically significant. A more sophisticated and recent approach is the one proposed by Becker and Mulligan [1997], who model time preference as endogenous. Before their attempt, rates of time preference were almost invariably taken as exogenous, with little discussion of what determines their level. Schooling plays a relevant role in their specification too. In fact it may provide a method of decreasing one's rate of time discount through several channels: the study of history and other subjects for schooling focuses students' attention on the future; schooling can communicate images of the situations and difficulties of adult life, which are the future of childhood and adolescence; in addition, through repeated practice at problem-solving, schooling helps children learn the art of scenario simulation. Thus, educated people should be more productive at reducing the remoteness of future pleasures and this is a possible mechanism through which education improves health.

More recently, experimental research has shown that risk and time preferences serve as a good predictor for field behavior, in particular in the health domain. Attitudes toward risk are likely to affect the purchase of health insurance, the use of preventive medical care, and the propensity to engage in behaviors that either increase or decrease mortality risk, such as cigarette smoking, seat belt use, alcohol consumption, exercising or nutrition (Chabris et al. [2008], Weller et al. [2008], Dohmen et al. [2011b], Barsky et al. [1997]). A few papers specifically examine the impact of risk preferences on Body Mass Index (BMI) and the incidence of overweight and obesity. Anderson and Mellor [2008], Sutter et al. [2013], and de Oliveira et al. [2016] find that individuals who are more risk tolerant are more likely to have a higher BMI and be obese.

In the empirical part of this paper we investigate the relationship between parental risk attitude and children's BMI and incidence of obesity. The aim is to check whether this positive relationship, stated by the existing literature, holds also when parents decide on behalf of their children.

2.2. Decision-making on behalf of others: the agency model in healthcare. Since in this paper we investigate how individuals make health decisions that do not affect directly their own health but the health status of someone else a relevant stream of literature is that on decision-making on behalf of others. In the health domain, this happens for example when doctors make treatment decisions on behalf of their patients, when adult children help elderly parents with healthcare decisions or when parents make health decisions for children. While in the standard agency model the role of an agent is to maximize the utility of the principal within available resources, in healthcare this role changes according to the subjects involved, the kind of relationship between them and the resources available. The main literature regarding the nature of the agency relationship in healthcare focus on the interaction doctor-patient (Vick and Scott [1998]). In analyzing this relationship it is possible to outline two main models (McKinstry [1992]) emphasizing the different understandings of the goals of the patient-doctor interaction, the guidelines to whom the doctor must be compliant, the role of patient values, the relevance given to patient autonomy and the efficiency of the information flow between the two parties.

The first one is the paternalistic model. In this model, the doctor makes the decision with the clear objective to maximize the patient's utility. The doctor does not have information about patient's utility function; she evaluates the patient's utility according to her utility function, ensuring that patients receive the interventions that best promote their health and well-being. The flow of information is asymmetric by the side of the doctor, who presents the patient with selected information that will encourage her to accept the decision the physician considers the most appropriate one. The doctor does not ask for information because she uses her skills to determine the patient's medical condition. This model is characterized by low patient's autonomy. As we will state later, in this paper we consider this model the appropriate framework to evaluate also the parent-child relationship.

The second is the informative model. In this relationship, the doctor does not see herself as being in charge of the decision-making process, but considers the patient to be the final arbiter. The objective of the interaction between the two is for the doctor to provide the patient with all relevant information and for the patient to select the medical interventions she wants. To this end, the doctor informs the patient of her disease state, the nature of possible diagnostic and therapeutic interventions, the nature and probability of risks and benefits associated with the interventions, and any uncertainties of knowledge. The flow of information goes from doctor to patient because the patient is the one who makes the choice. This model is characterized by a high patient's autonomy. There are specific contexts in which this model is not applicable because the agent is in a condition that does not allow her to make a choice.

Considering the criteria according to which the distinction between the two models is made, it is possible to identify intermediate models of interaction. In particular, when the decision is let to the patient, the role of the doctor can be interpreted in different ways. The doctor can provide the patient with information on the nature of the condition and the risks and benefits of possible alternatives, like in the informative model, and assists her in understanding her values and thus her utility function. In this model the doctor suggests the patient the most effective treatment to maximize patient's utility. The role of the doctor could also be more intrusive that is when she collects information from the patient but simultaneously tries to shape the patient's utility function according to the health-related values that the doctor herself considers more relevant.

Even if this literature is relevant to frame our problem well, we focus on the relationship between parent and child. The aim of this paper is to fill this gap in the literature since the existing literature has neglected the family context to favor the more strictly medical context.

2.3. Parenting. As we pointed out before, the objective of our research lies in the analysis of the relationship between parents and children. To analyze health decisions it is useful to make a brief review of how preferences are transmitted from parents to children and of how parents make decisions for their children.

Since risk and time preferences play an important role in individual decisions to invest in education, pensions and not least health, it becomes important to understand how preferences are formed. Individuals may be born with innate time and risk preferences or preferences may be learned. There is some evidence that at least at the beginning of one's life preferences are determined by the genetic makeup of the individual. However, evidence shows also that preferences vary over the lifecycle which suggests that preferences may be endogenous, as supposed and modeled by Becker and Mulligan [1997]. Therefore, individuals have the opportunity to invest resources to modify their preferences over time and risk. During childhood, parents can influence children's preferences by investing resources in teaching them to be more risk averse or to discount less future consumption. Under these assumptions, the transmission of preferences from parents to children may occur through both genetic inheritance and learning.

Despite of the channel through which parents transmit preferences to their children, there have been few attempts to examine empirically correlations in time and risk preferences between parents and their offspring and how these preferences affect their outcomes.

Webley and Nyhus [2006] choose future orientation as a proxy for time preferences and examine whether parents and children's' future orientation is correlated. Knowles and Postlewaite [2005] investigate correlations in time preference using saving residuals to measure it. Both the studies state a significant positive correlation, while Reynolds et al. [2009] and Kosse and Pfeiffer [2012] examining correlations in time preference still find a positive correlation but not significant.

A similar number of studies have examined correlations in risk preferences. Dohmen et al. [2011a] investigate the intergenerational transmission of risk preferences using a general question regarding willingness to take risk from the German Socio-Economic Panel. The results show that risk preferences of parents and their children are significantly correlated. Hryshko et al. [2011] and Charles and Hurst [2003], using similar analyses find that risk preferences, measured using a gamble with different levels of lifetime income, are correlated but at the more extreme end of the distribution only.

The positive significant relationship between parents and their young adult offspring risk and time preferences is confirmed by Brown and van der Pol [2015]. In addition to previous evidence, they take into account gender differences. Regarding risk attitude, risk seeking parents are more likely to have risk seeking offspring except for the father/daughter dyad. Daughters are more likely to be influenced by their mother's risk preferences, however, sons are equally influenced by both parents. Moving to time attitude, the association in parental and offspring time preference was larger for mothers than fathers. We see that evidence from several empirical studies suggests that parents affect their children's preferences.

The most relevant attempt to outline a theoretical model of how parents interfere with children's utility has been made by Doepke and Zilibotti [2014]. They develop a theory of parent/child relations that rationalizes the choice between alternative parenting styles. In their theory, according to different parenting styles, parents can affect their children's choices via two channels: either by influencing children's preferences or by imposing direct restrictions on their choice sets. The parenting styles outlined in their study are the following: permissive, authoritative and authoritarian. Permissive parents are those who allow children to make free choices according to their natural inclinations. This parenting style can be assimilated to the doctor-patient informative model we presented in the previous section: in this case children make decisions, as in the informative model patients make them. Authoritative parents are those who attempt to mold their children's preferences, with the aim of inducing choices that they view as conducive to success in life. To continue the parallelism with the doctor-patient interaction, this can be considered as the intermediate situation in which the doctor tries to shape the patient's utility function according to her health-related values. Lastly, authoritarian parents restrict children's choices with the objective to impose their will on the child. Authoritarian parents make decisions on behalf of their children, as paternalistic doctors make decisions on behalf of their patients.

This theory can be applied to different aspects of life where there is a conflict of interest between parents and children: one of the aspect is education but it is possible to apply the model to the analysis of health decisions, which is the objective of this research work.

3. Theoretical Model

In this section we present a model of parental decision-making with respect to health investment decisions. In this model a central role is played by individual risk preferences, and therefore it leads to a prediction of the optimal choice of investment based on the individual risk attitude. This model is shaped on the Doepke and Zilibotti [2014] general theory. While their framework considers three alternative parenting styles and is sufficiently flexible to be applied to many kinds of choices children make and that their parents may disagree with, our one is specific for the field of health and needs few important assumptions.

We assume that the model economy is populated by two overlapping generations of children and parents. The life span of an individual life is divided in childhood (period 1) and parenthood (period 2). Each parent has one child, for simplicity.

The period utility function is defined on a consumption vector c and a preference parameter a. Parents can put child-rearing effort to mold their children's preferences. The preference parameter a is acquired during childhood but it affects utility in both childhood and parenthood. The parameter $a \in A = [1, \bar{a}]$, where $\bar{a} > 1$, captures children's innate preference for instant gratification: the same parameter in fact is assumed to be equal to 1 for parents. The preference parameter of children is chosen by the parent, and the possibility of choosing $a < \bar{a}$ captures the option for parents to stifle the child's enjoyment of young-age consumption. Although cardinal utility is maximized by setting $a = \bar{a}$, the parent may choose a lower a in order to make the child more patient. However, we assume that the parent affects only the choice of health investment and not the preference parameter, thus $a_y = \bar{a}$.

Age also has an effect on preferences. For instance, a child may be intrinsically less patient or less risk averse than an adult. Thus, there are separate period utility functions for children $U_y(c_y|\bar{a})$ and for parents $U_p(c_p|1)$ [1]. We parameterize preferences by a Constant Relative Risk Aversion (CRRA) utility function. Therefore, the individual's utility is given by:

$$U_i(c_i|a_i) = a_i \frac{(c_i)^{1-\theta}}{1-\theta} \tag{1}$$

where i ={y, p} and $\theta \in (0, 1)$ measures the degree of relative risk aversion.

The stochastic level of consumption for both, parents and children, is implied by the investment in health, x. This variable can be interpreted as the effort a parent can put to make her child adhere to a healthy lifestyle. For simplicity, we assume that parental effort is costless.

The health investment choice is continuous, $x \in [0, 1]$, and together with the realization of an exogenous stochastic shock s, contributes to the accumulation of human capital. The return to health investment x is determined by a parameter R. We also assume that in every period an individual is equipped with a constant wealth endowment, w, which can be approximated with her own income in adulthood or parent's income in childhood. According to these assumptions, the laws of motion of the model are the following:

- $c_y = (1 - x)w$
- $c_p = (1 + Rx)w + s$

Against this backdrop, the parent's value function is defined as follows[2] :

$$V_p(a_p, x) = \max_x \left\{ U_p(c_p|a_p) + \delta\Omega(a_p, a_y, x) \right\} \tag{2}$$

where $a_p = 1$ and $\delta \in (0, 1)$ is the overall altruism of the individual, i. e. how much the individual discounts the utility derived from someone else with respect to her own utility. The value of δ determines the relevance of child's utility into the parent's one. The child's value function is the following:

[1] We use y-subscript for the young/child and p-subscript for the parent.

[2] See Appendix A for the complete derivation.

$$V_y(a_y, x) = \max_x \left\{ E[U_y(c_y|a_y)] + \beta V_p(a_p, x) \right\} \tag{3}$$

where $a_y = \bar{a}$.

Child's utility does not enter the parent's utility function in its entirety. In particular the utility a parent derives from her child's experiences is given by:

$$\Omega(a_p, a_y, x) = E_s[(1-\lambda)U_y(c_y|a_y)] + \lambda U_p(c_y|a_p) + \beta V_p(a_p, x)] \tag{4}$$

The function $\Omega(x, a_p, a_y)$ comprises both an altruistic and a paternalistic component, whose weights are respectively $1-\lambda$ and λ. Altruism is the standard enjoyment of the child's own utility as in Becker [1974], while paternalism is the evaluation of the child's actions through the lens of the parent's utility function. Paternalism applies only to childhood, and not to the child's felicity when the child has grown up. Hence, the child's adultage utility enters the parent's value function as $\beta V_p(a_p, x)$, where β is the discount factor between period 1 and period 2 and $V_p(a, x)$ is the value function of the child when she turns into a parent. Restricting paternalistic motives to childhood is broadly realistic because preferences change with age, implying that there is more scope for conflict with an adolescent than with a grown-up child.

When considering health investment, we assume that a parent deliberately acts in an authoritarian way, that is, directly forces her child to undertake the investment she desires. We consider this a reasonable assumption because the stakes involved in health decisions are too high to go for an alternative parenting style. Therefore, the theoretical problem becomes a problem of parent's value function maximization as follows:

$$V_p(\bar{a}, x) = \frac{((1+Rx)w)^{1-\theta}}{1-\theta} + \max_{x \in X} \left\{ \delta\Omega(a_p, a_y, x) \right\} \tag{5}$$

$$\Omega(\bar{a}, x) = (1-\lambda)\left[\bar{a}\frac{((1-x)w)^{1-\theta}}{1-\theta}\right] + \lambda\left[\frac{((1-x)w)^{1-\theta}}{1-\theta}\right] + \beta\left[\frac{((1+Rx)w)^{1-\theta}}{1-\theta}\right] \tag{6}$$

$$= w\left[\left(\lambda + (1-\lambda)\bar{a}\right)\left(\frac{(1-x)^{1-\theta}}{1-\theta}\right) + \beta\left[\frac{(1+Rx)^{1-\theta}}{1-\theta}\right]\right]$$

$$\max_{x \in X} \Omega(\bar{a}, x) \tag{7}$$

$$x^* = argmax\Omega(\bar{a}, x) = \frac{1 - \left(\frac{\lambda+(1-\lambda)\bar{a}}{\beta R}\right)^{\frac{1}{\theta}}}{1 + R\left(\frac{\lambda+(1-\lambda)\bar{a}}{\beta R}\right)^{\frac{1}{\theta}}} \tag{8}$$

According to the result obtained through the maximization, the parent's investment on her child's health depends positively on her altruism (λ), her risk aversion (θ) and on the return to the investment (R), and negatively on her discount rate (β).

In Figure 1 is displayed a simulation of the relationship between investment in health and risk aversion and between investment in health and parental altruism[3]. The second graph shows that for extremely high levels of parental altruism the gap between the investment in health for risk lovers and risk averse people disappears. This is consistent with the fact that completely altruistic parents invest in their children's health as much as they can.

In the following section, we empirically test the positive relationship between risk aversion and the investment in health.

4. Empirical Analysis

4.1. Data and descriptives. In the previous section we outlined a theoretical model for parents' choice of health investment for their children. The objective of the second part of this work is to test whether the prediction of the theoretical model that a higher risk aversion (θ) leads to higher parental investment in children health, x, is supported by evidence. In order to achieve this objective, we use data from the Socio-Economic Panel Survey (SOEP), a representative panel survey of the German population.

[3]In Figure 1, the left graph is drawn for a particular value of the paternalism parameter (λ= 0,4) and the right one for three particular values of the parameter θ ("Low risk aversion": θ=0,3 ; "Risk neutrality" θ=0,6 ; "High risk aversion" θ= 0,9).

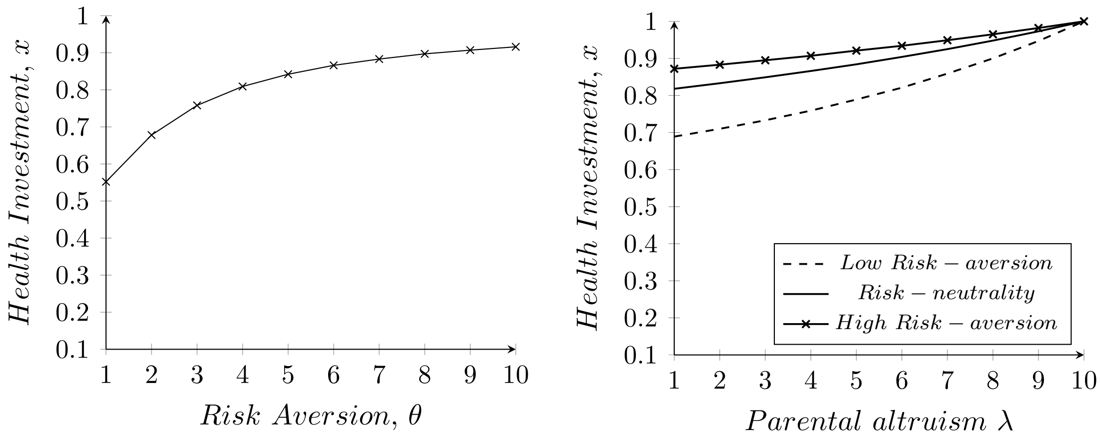


Figure 1. Left: Health investment as a function of parental risk aversion measured on a 1-10 scale. Right: Evolution of health investment with respect to parental altruism for three levels of risk attitude.

From 1984 until today, the SOEP surveys the head of each household in the sample, but also gives the full survey to all other household members over the age of 17. In each wave approximately 11.000 households and more than 20.000 persons are surveyed.

This survey fits the scope of our work for two main reasons.

The first is that from 2003[4], in addition to information regarding adult individuals and the whole household, in the families in which there are children up to 17 years old, are collected also complementary questionnaires (Mother-Child Questionnaire, Parents Questionnaire, Pre-teen Questionnaire and Early Youth Questionnaire) that cover information on children and pre-adolescents aged 2-14 years. Through the completion of these questionnaires, specific information on children's health are recorded.

The second is that the SOEP data provide unique measures of individual subjective attitude toward risk taking, which is what we are interested in as independent variable in our model. In fact from 2004[5], respondents were asked about their attitude toward risk in general[6]. The general question allows respondents to rate their willingness to

[4]The Mother-Child questionnaire was collected for the first time in 2003 for children aged 1, from 2005 for children aged 3, from 2008 for children aged 6 and from 2010 for children aged 1-10. Information related to 11-14 years old subjects are available from 2013.

[5]The general risk question is included in years 2004, 2006 and every year from 2008.

[6]The question directly asks individuals to make a global assessment of their willingness to take risks: "How willing are you to take risks in general?".

take risks (henceforth WTR) on a 11-point scale ranging from 0, indicating complete unwillingness, and 10, indicating a very high willingness. In 2004, 2009 and 2014 the survey includes five additional questions that use the same scale as the general risk question, but ask about risk taking in specific contexts: car driving, sports and leisure, career, health, financial matters. As robustness check, we replicated the same analysis using the health related question: we didn't find any significant difference with respect to using the general risk question.

A crucial concern is whether survey questions can be meaningfully interpreted in terms of actual risk-taking behavior. In this case, the predictive power of the measures, and in particular of the general question, was tested by Dohmen et al. [2011b]. They compared answers in the SOEP with the results of a complementary laboratory experiment with a representative subject pool. According to the results of their validation, all questions in the SOEP provide valid measures of risk attitudes.

As a first descriptive analysis, we display the distribution of general risk attitudes in our representative sample of 5,011 observations in Figure 2. Each bar indicates the fraction of individuals choosing a given number on the eleven-point risk scale. The figure reveals substantial heterogeneity in risk attitudes across the population: the modal response is 5, but risk attitudes vary widely over the entire scale, with mass distributed over the entire support. The fraction of respondents who chooses a value of 10, indicating that they are very willing to take risks, is lower than the fraction of individuals who chooses 0, indicating that they are not at all willing to take risks.

For the dependent variable we use data provided by the complementary questionnaires regarding children. Since we do not have a direct measure for a parent investment in her child health we use two health outcomes that reasonably represent it: child's BMI and child's incidence of overweight.

Child's BMI is obtained exploiting data on height and weight. With the metric system the formula for BMI is weight in kilograms divided by height in meters squared. Although the increasing debate about its appropriateness, BMI is still considered the most reliable indicator of body fatness. BMI does not measure directly body fat, but it is demonstrated that it correlates to direct measures of body fat such as underwater

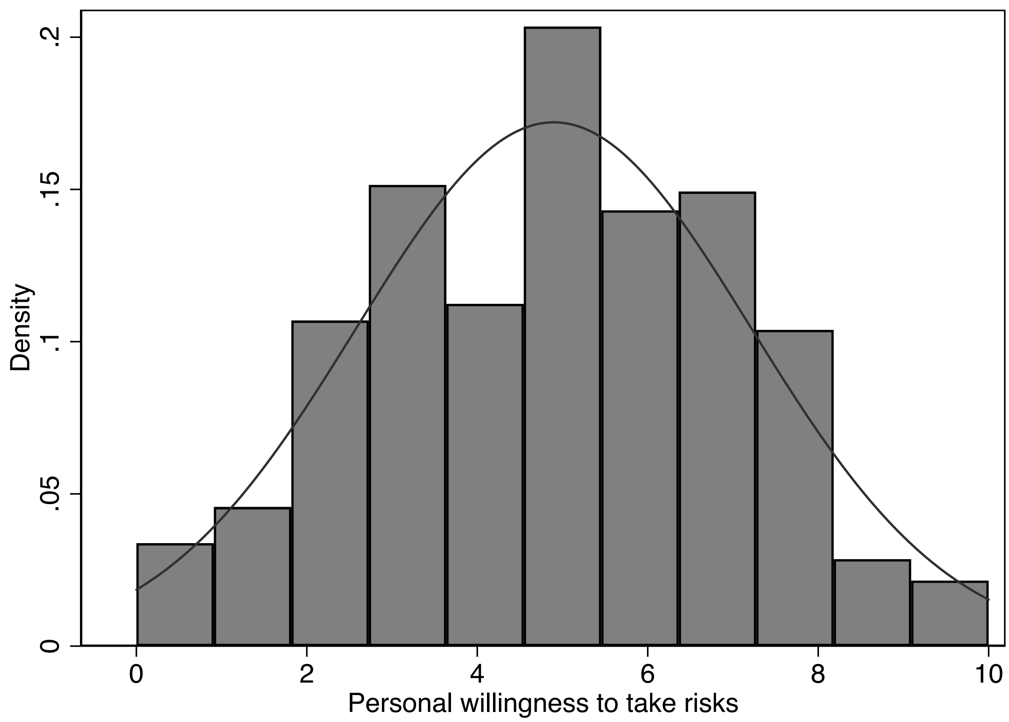


Figure 2. Histogram of responses to the question about willingness to take risk "in general" measured on an eleven-point scale (0= not at all willing; 10=very willing).

weighting and dual energy x-ray absorbptiometry (DXA). Our assumption is that a parent who cares less about health risks associated with higher BMI and obesity is also more likely to let her children adhere to an unhealthy lifestyle. In fact, among unhealthy behaviors, bad food habits play a major role and children during their childhood do not decide autonomously their meals. At the age 2-14, covered in our sample, parents are in charge of making nutritious choices on behalf of their children. According to this reasoning, the decisions regarding children's nutrition can be considered an investment in health made by the parents. It is well known in medicine literature that bad eating habits and practices may have harmful effects on children health. This category of habits, such for example skipping breakfast, having a diet rich in Trans fats or the consumption of sugar-sweetened beverages (primarily soda and juice), for the children leads to increase body weight, which becomes a risk factor for chronic diseases (Silveira et al. [2013]). To test the robustness of our reasoning, we also use, as an alternative dependent variable, the incidence of overweight among children, considering overweight those children whose BMI is higher than the 85th percentile.

Our attention is limited to children aged between 2 and 14 years and parents, father and mother, aged between 20 and 65 years. The final sample consists of 5,011 children, among which 2,527 are females, and the other 2,484 are males. The children are all coupled with a parent. The distribution of parents is 2,816 mothers, and 2,195 fathers. Table 1 and Table 2 provide descriptive statistics of the main estimation sample .

In Figure 3, we display the distribution of the BMI measure for the children in our sample.

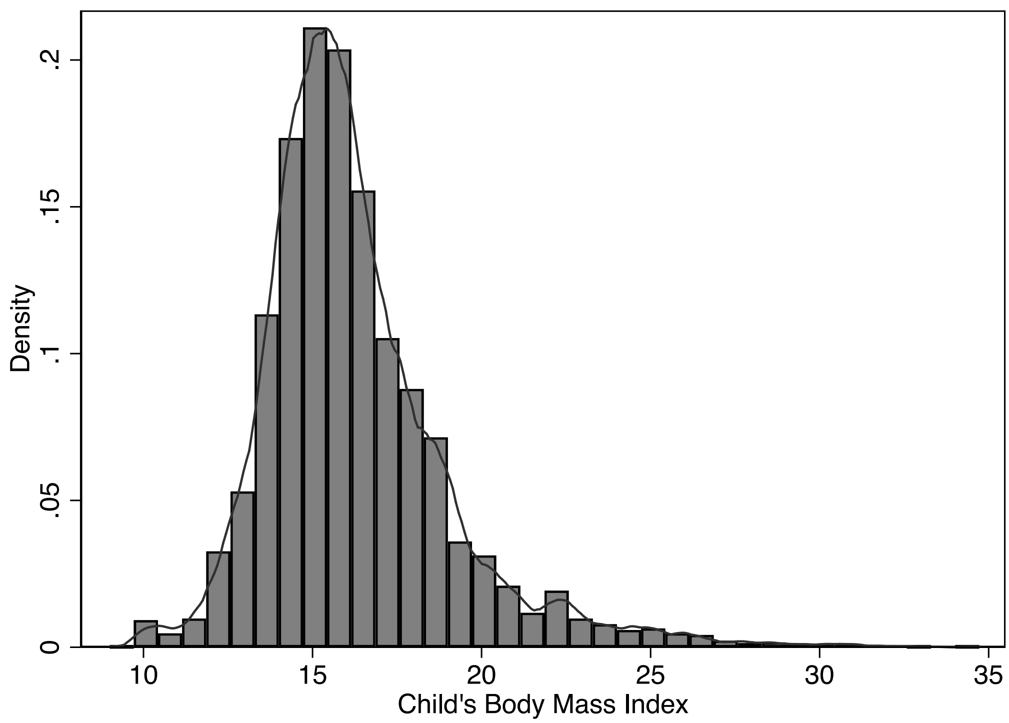


Figure 3. Distribution of children's Body Mass Index (BMI)

For children and teens, overweight and obesity are defined according to the relative BMI's percentile[7]. Due to the fact that weight and height change during growth and

[7]The SAS program for calculating percentile for body mass index-for-age is available at http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm

development, a child's BMI must be interpreted relative to other children of the same sex and age. In our sample the overall mean is 16.24, aligned with the value of 16 (for children 2-14) reported by the *"Reference percentiles for anthropometric measures and blood pressure based on the German Health Interview and Examination Survey for Children and Adolescents 2003-2006 (KiGGS)"*. Coherently with the literature, the mean for male children (16.36) is slightly higher than the mean for female children (16.11).

We also construct the BMI of the parent, using the information weight and height, and use it as additional control in the regression. Obesity in children is in fact heavily influenced by genetic transmission: early studies, such as Coate [1983], already reveal that the probability of an adolescent being obese increases by 20% if either of her parents is obese (Costa-Font and Gil [2013]). Parents, therefore, are responsible not only for for passing onto their children habits and attitudes but also for the genetic structure of their children's body. Figure 4 shows the distribution of parents' BMI in our sample.

For adults, BMI is interpreted using standard weight status categories that are the same for all ages and for both men and women. The standard weight status categories associated with BMI ranges for adults are:

- Underweight: BMI below 18.5;
- Normal weight: BMI between 18.5 and 24.9;
- Overweight: BMI between 25 and 29.9;
- Obese: BMI over 30.

In our sample the mean of parents' BMI is 25.83, which is classified as overweight. This evidence and the right skewness of the distribution, shown in Figure 4, are consistent with the increasing burden of obesity among adult population in Germany (International Association for the Study of Obesity).

4.2. Empirical Model. The first specification we apply to the data is the following:

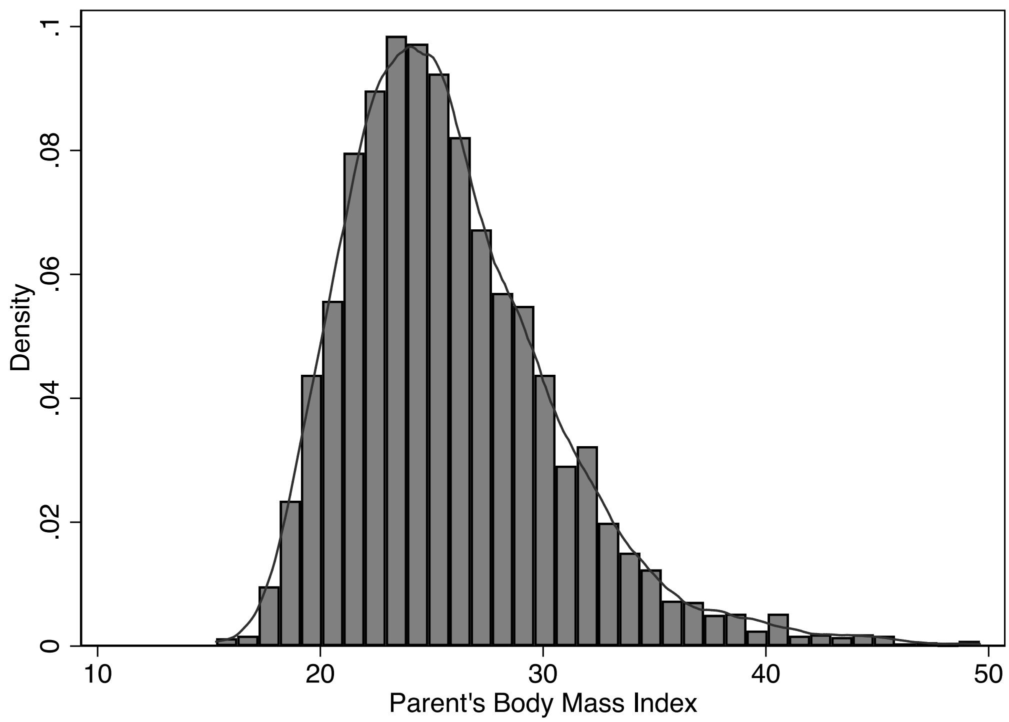


Figure 4. Distribution of parents' Body Mass Index (BMI)

$$BMI_y = \beta_0 + \beta_1 WTR_p + \beta_2 X_p + \beta_3 X_y + \epsilon \tag{9}$$

We estimate this specification by OLS because the dependent variable, Child's BMI, is continuous. Our key explanatory variable is the parent willingness to take risks (WTR). We choose to use the WTR Dummy instead of the continuous alternative. We include a set of controls for the parent (X_p) and a set of controls for the child (X_y). X_p includes income, education, employment status, gender and age. We estimate twice the same regression, using two alternative measures of income (See Section 5). X_y includes gender and age. We also consider cross-section fixed effects for the state of residence. errors are clusterized at household level.

The second specification is the following:

$$Overweight_y = \beta_0 + \beta_1 WTR_p + \beta_2 X_p + \beta_3 X_c + \epsilon \tag{10}$$

where the dependent variable (i.e. incidence of child overweight) is binary and in particular is equal to 1 if the child is overweight and 0 otherwise. For this reason we estimate this model by Probit.

In both specifications, the last term, ϵ, is the stochastic error, which we assume to be uncorrelated with the explanatory variables and the outcome observed. We can reasonably assume that reverse causality does not occur: child's BMI is not responsible for modifications of the parent's attitude toward risks. The willingness to take risks can be here considered as a genetic characteristics of a parent, which cannot easily be affected by external factors. The inclusion of income and education as controls exhausts the factors that can eventually affect both, the attitude toward risks and the child's BMI. We consider the measurement issue negligible due to the reliable validation of the WTR measure made by Dohmen et al. [2011b].

5. Results

Results of the OLS are reported in Table 3, while results of the Probit are reported in Table 4. The reported coefficients in Table 4 are Probit marginal effects estimates, evaluated at the mean of the independent variables [8].

As all columns show in both Tables, the willingness to take risk of the parent has a significant impact on her child's BMI (coefficients range in Table 3: 0.332 - 0.347; coefficients range in Table 4: 0.045 - 0.048), thus confirming the positive relationship stated by the theoretical model. The columns differ accordingly to the inclusion or not of specific controls in the related specification. In columns (3) - (6) fixed effects for the State of residence are included. We find a significant negative relation between income and education and children's BMI. Inverse associations between socio-economic status and adiposity in children are well established (Shrewsbury and Wardle [2008]). Our paper in particular agrees with the review of Cohen et al. [2013], who analyze 289 articles finding that the relationship between educational attainment and obesity depends on the country's level of development. For high-income countries, such Germany, 70% of

[8] The interpretation of the marginal effect of a dummy variable on a binary outcome is intended to be interpreted as follows: a 1% increase in the probability of being risk-lover affects the probability of that parent's child being overweight by 4.5%.

the studies assess an inverse association. As well as for education, the inverse association between family income and childhood obesity is consistent with the results of Jo [2014]. This relationship does not hold for very low-income families and generally increases as children age. In order to give evidence that the choice of a different income measure produces estimates consistent with those obtained in the main analysis, we estimate all the specifications by adopting two alternative definitions and measures of income. The first one is the annual household labor income ("HH Income"). We exclude from the sample the top and bottom 1 percent observations. Excluding observations in the two tails of the income distribution is standard in this literature. In our sample this excludes at the top 51 observations with an income between 219K and 713K euros, at the bottom 48 observations with an income less or equal 2,160 euros per year. The second one is the net income last month ("Income"). Even with this measure we proceed excluding the top 40 observation with a net labour income last month higher than 6K euros (to 10K) and the bottom 40 with a net labour income last month lower than 100 euros. We find a significant negative coefficient for the first measure while we don't find a significant coefficient for the second measure but regardless of which measure we use the results for the WTR coefficient hold. In columns (5) and (6) we include, as a control variable, the parental employment status. In the literature the relationship between childhood obesity and maternal employment has been investigated by Anderson et al. [2003] and Cawley and Liu [2012], who both find a positive association between the number of working hours of the mother and the incidence of obesity in the child. More hours worked by the mother tend to be negatively related to positive routines like eating meals as a family or at regular times, or having family rules about hours of television watched. Our results agree with these findings [9], even if the magnitude of the coefficient reflects a small impact on the outcome variable. The gender and age of the parent do not affect significantly the child's BMI, while we assess a significantly lower BMI for daughters with respect to sons. This finding can be explained by a number of reasons, including girls usually being more conscious about their physical appearance, and boys

[9] The variable "Employment status" is a 1-9 index, where 1 is the "Full employed" condition and 9 the "Not employed" condition.

being more brand loyal and therefore susceptible to the billions of pounds spent on marketing to children through brand characters and sports stars. We are analyzing gender differences more in details in Subsection 5.1.

5.1. Heterogeneous effects. We also compared the performance of the risk attitude in terms of prediction accuracy when splitting the sample. Results of these heterogeneous effects are reported in Table 5. All columns include fixed effects for the state of residence and the parent's employment status among the control variables. In column (1) we restrict the sample to German subjects[10] and we see that the coefficient for the WTR increases slightly (0.381). Keeping the sample restricted to Germans, we find that in the age range 4-8 years old the effect of the WTR is greater (0.468). Splitting the sample according to child's gender, we find a powerful result: the WTR coefficient for German females aged 4-8 doubles compared to the baseline (0.723), while if we consider German males of the same age the value of the coefficient is the same we find in the baseline specification but it loses its significance. Columns (5) - (6) report the results obtained by splitting the sample according to the parent's gender. We find that paternal willingness to take risks has a slightly higher impact on child's BMI than maternal WTR, except if we restrict the sample to daughter. For daughters in fact column (7) shows that having a mother who loves to takes risks increases significantly the daughter's BMI of 0.629. In columns (8) and (9) we follow the path that the last results suggest showing that daughters' BMI is more influenced by parental WTR when they grow into adolescence (>8 years old) and that the highest influence is exerted when the sample is restricted to mothers.

5.2. Robustness checks. The results of a second version of the Probit specification where we look at the obesity incidence instead of the overweight are provided in Table 6. We find that the probability of being overweight and obese is affected by having a risk lover parent at the same magnitude level (4-5 %). These results confirm the inverse association between socio-economic status and obesity, while we don't find significant differences according to age and gender of both, parents and children. In

[10]Both parents born in Germany.

our sample only 251 children out of 5,011 are obese, therefore we consider showing the results for the overweight incidence more relevant to the analysis. We keep the results related to obesity as robustness check.

As additional robustness check, we replicate the estimations including also parents' BMI as a control to check whether the results hold even taking into account the genetic contribution of a parent's BMI to her child's one. Here, we use the Parent's BMI Index (8-point scale index, built according to the WHO BMI reference tables for adults). The results we obtain are reported in Table 7. It is not surprising the significance of the genetic contribution of the parent's BMI to her child's one. The predictive value of parental BMI for an individual BMI has been assessed in several studies. Our results are consistent with the findings of Svensson et al. [2011]: the impact of parental BMI on the BMI in children is strengthened as the child grows into adolescence. We test for the stronger influence of the genetic factor the older the child gets by comparing the results in two subsamples: the first composed of children younger than 10 years and the second of children older than 10 years (to 14). The difference between the two estimates of the coefficient for the parent's BMI was on average 0.7 (0.2 vs 0.9). In our sample we do not find a significant difference between paternal and maternal BMI as predictors for child's BMI[11].

In Table 8 we report the results of the final robustness checks we compute for parents. In particular in the first and the second column are reported the results of OLS regressions, with the parent's BMI as dependent variable. The explanatory variable is the parent's WTR. All the controls are known from Tables 3-4. We find a positive but not significant relationship between the willingness to take risks and the individual BMI. The coefficient becomes significant in the Probit regression - columns (3) and (4) - indicating that being more risk lover increases the probability of being overweight for adult subjects. The coefficient of the variables Income and Education are negative and significant, confirming the findings of Table 3. The socio-economic status affects BMI for all the family components: parents and children. The age of the individual is positively associated with BMI: BMI increases as the individual grows old. The results

[11]Results of additional robustness checks specifications are not reported but available upon request.

related to "Gender" confirm the fact that on average males have a BMI higher than women. Finally, the results reported in (5) and (6) are obtained by a Probit estimation where the dependent variable "Smoke" is binary. The results show that a risk-lover individual is 7% more likely to be a smoker. Even in these last two specifications, the socio-economic status has a relevant impact on the choice of smoking. High educated and high income family components are more likely to choose to not smoke. These last estimates are aligned with the results presented by Dohmen et al. [2011b]. In all the six columns, cross-section fixed effects are included.

6. Discussion

The current study provides a snapshot of the processes through which parental characteristics, and in particular parental willingness to take risks, may put children at risk for obesity. The discussion will review the results, acknowledge the limitations, and offer suggestions for applying the evidences collected. While maximizing their utility function, parents face trade-offs between investing in their child's health (i.e. molding child's health choices or deciding on behalf of her) and devoting their wealth to consumption. In literature, parental investment theory has mostly been examined through the analysis of educational child outcomes, with less emphasis on the actions parents take to promote a better health condition. Here, we examine whether parent's characteristics (willingness to take risks, BMI, age, gender, education, income, employment status) and child factors (gender, age) influence parental investment in health-seeking behaviors. We measure the parental investment in child's health through the result of a preventative healthy lifestyle: the child's Body Mass Index. Obesity in fact presumably represents an outcome of prolonged poor habits. The differences in households' characteristics, parents' characteristics, and children's characteristics between obese and non-obese children revealed some interesting findings about parent's attitudes and economic well-being. First of all, higher parental education and wealth are associated with higher levels of health investment in children, thus lower child's BMI. However, considerable variation in parental health-seeking exists even between households with comparable levels of education and income. This implies that other factors can help explain investment biases within and between households: our study focuses on parental risk attitude. The results state that children who are obese are more likely to have parents with a higher willingness to take risks when compared to their non-obese age mates.

However, although the current study contributes to our understanding of how parental risk attitude affects the decisions regarding children health, caution must be taken when interpreting the results, due to limitations attributable to the information availability in the SOEP Survey.

The first critical aspect of the study is the choice of the dependent variable. According to us, child's BMI represents the best available proxy of the variable "Investment in child's health". If in the dataset we would have found a more appropriate information to measure it, perhaps we would have had the opportunity to compare the results obtained by the two alternatives.

Another potential issue is posed by the use of parent reports to gather relevant information about both, parents themselves and their children. A number of alternative research methodologies could be used in future studies to reduce any method bias associated with the use of self-reported measures. For instance, there are specific datasets in which the anthropometric measurements of the child are taken directly. On the other hand, we are less concerned about the self-reported measure of willingness to take risks, since we rely on the validation procedure of Dohmen et al. [2011b] (See Section 1 and 4.1 for further information).

7. Conclusions

This paper provides evidence of the behavioral validity of the willingness to take risk not only as a predictor of health behaviors regarding the self but also for decisions made on behalf of one's children. What clearly emerges from the results is the association between the parent's attitude toward risks and children's obesity, controlling for socio-economic characteristics.

Based on the findings of this study, the most viable pathways to increase parental investment in child's health seems to be making parents more aware of the risks related to being obese from childhood. Childhood obesity is in fact associated with an increased risk for other diseases not only during youth but also later in life, including diabetes, arterial hypertension, coronary artery disease, and fatty liver disease. Furthermore, obesity accelerates atherosclerosis progression already in children and young adults (Barton [2012]). In addition to giving parents more information regarding the risks related to obesity, the challenge is to improve their ability to understand this health-related information, that means increase their health literacy. Health literacy is the ability to obtain, read, understand, and use healthcare information in order to make appropriate health decisions. In this study, a higher health literacy could improve parental ability to evaluate the return on the investment in child's health. According to our theoretical framework, an underestimated return on such investment reduces the optimal investment level in the solution of the maximization problem. In addition, having high educated parents who are more able to understand health-related risks and low educated ones who are not will potentially increase the socio-economic gradient of obesity which is already a social concern. Future interventions should go in the direction of reducing the overall obesity but also closing this socio-economic gap.

Although the current study emphasizes family and parent contributions to children's BMI, children's role in their own physical health and well-being can not be discounted. Toward that end, consistent efforts need to be made to work directly with children on changing their habits and increasing their awareness of consequences of their own choices on their health. For the specific outcome we look at in this study, i.e. children obesity, the results underline the importance of helping children develop lifelong habits for regular physical activity and healthy nutrition. Parents have an irreplaceable role in encouraging children to connect with the physical and mental health benefits of conducing a healthy lifestyle from childhood.

Tables

Table 1. Description of the Variables

Variable	Description
Dependent variables	
Child's Overweight	Overweight defined according to the WHO BMI reference tables (percentile 85). Binary variable: 0 if Child's BMI<85 percentile, 1 if Child's BMI>85 percentile
Child's Obesity	Obesity defined according to the WHO BMI reference tables (percentile 95). Binary variable: 0 if Child's BMI<95 percentile, 1 if Child's BMI>95 percentile
Child's BMI	Child Body Mass Index, built as the person's weight in kilograms divided by the square of height in meters
Explanatory variables	
Parent's Willingness to Take Risks (WTR)	WTR measured on a 11-point scale ranging from 0 (complete unwillingness) and 10 (very high willingness)
Parent's Willingness to Take Risks Dummy (WTR Dummy)	0 if the parent is risk-averse (WTR<8), 1 if the parent is risk-seeker (i.e. WTR≥8)
HH Income	Annual household labor income, in €
Income	Net income last month, in €
Parental education	Education With Respect to High School (1 less than HS, 3 more than HS)
Employment status	10-point scale ranging from 0 (Full-time employed) to 9 (Not employed)
Controls for parents	
Gender	1 if male, 2 if female
Smoke	0 if the subject currently smokes , 1 if not
Age	Parent's age in years
Parent's BMI	BMI of the household adult parent (father o mother)
Parent's BMI Index	1 to 7 (underweight to obese) according to the WHO BMI adult references
Parent's Overweight	Overweight defined according to the WHO BMI reference tables (>25)
Child's characteristics	
Gender	1 if male, 2 if female
Age	Child's age in months
Eating habits during week	1 if the child does not eat at home or eats alone
Eating habits during week-end	1 if the child does not eat at home or eats alone

Table 2. Summary of Descriptive Statistics

Variable	Mean	Std. Dev	Q1	Q2	Q3	N
Dependent variables						
Child's Overweight	0.19	0.39	0	0	0	5,011
Child's Obesity	0.05	0.21	0	0	0	5,011
Child's BMI	16.25	2.77	14.57	15.70	17.35	5,011
Explanatory variables						
Parent's Willingness to Take Risks (WTR)	4.89	2.31	3	5	7	5,011
Parent's Willingness to Take Risks Dummy (WTR Dummy)	0.14	0.34	0	0	0	5,011
HH Income	62,946.85	34,414.8	39,600	58,400	78,970	5,011
Income	1,834.48	1,171.64	930	1,600	2,400	5,011
Parental education	2.18	0.61	2	2	3	5,011
Employment status	2.58	2.66	1	1	2	4,986
Controls for parents						
Gender	1.49	0.50	1	1	2	5,011
Smoke	0.68	0.46	0	1	1	5,011
Age	41.97	6.69	37	42	47	5,011
Parent's BMI	25.85	4.76	22.53	25.09	28.40	5,011
Parent's BMI Index	3.71	0.92	3	4	4	5,011
Parent's Overweight	0.50	0.49	0	1	1	5,011
Child's characteristics						
Gender	1.50	0.50	1	2	2	5,011
Age	69.52	43.52	34	66	75	5,011
Eating habits during week	0.53	0.49	0	1	1	821
Eating habits during week-end	0.13	0.33	0	0	0	836

Table 3. Child's Body Mass Index OLS Results

VARIABLES	Child's Body Mass Index					
	(1)	(2)	(3)	(4)	(5)	(6)
WTR	0.332**	0.352**	0.334**	0.353**	0.325**	0.347**
	(0.149)	(0.150)	(0.148)	(0.149)	(0.148)	(0.149)
HH Income	-0.184***		-0.181***		-0.196***	
	(0.054)		(0.0559)		(0.0561)	
Income		-0.0314		-0.0258		-0.0289
		(0.0459)		(0.0452)		(0.0452)
Education	-0.204*	-0.307***	-0.219**	-0.325***	-0.221**	-0.332***
	(0.108)	(0.110)	(0.104)	(0.107)	(0.103)	(0.106)
Employment status					-0.0379***	-0.0279*
					(0.0145)	(0.0145)
Parent's characteristics						
Gender	0.052	0.0562	0.0622	0.0661	0.0630	0.0667
	(0.094)	(0.0925)	(0.0960)	(0.0948)	(0.0958)	(0.0946)
Age	-0.032	-0.0678	-0.0390	-0.0716	-0.0387	-0.0733
	(0.078)	(0.0760)	(0.0780)	(0.0769)	(0.0782)	(0.0769)
Child's characteristics						
Gender	-0.305***	-0.298***	-0.317***	-0.309***	-0.310***	-0.304***
	(0.116)	(0.115)	(0.115)	(0.114)	(0.115)	(0.114)
Age	-0.059***	-0.0583***	-0.0587***	-0.0585***	-0.0590***	-0.0587***
	(0.007)	(0.00682)	(0.00684)	(0.00683)	(0.00683)	(0.00682)
Constant	18.798***	19.50***	19.20***	19.87***	19.33***	20.01***
	(1.602)	(1.566)	(1.629)	(1.608)	(1.631)	(1.609)
State FE	No	No	Yes	Yes	Yes	Yes
Observations	5,011	5,011	5,011	5,011	5,011	5,011
R-squared	0.226	0.222	0.231	0.227	0.232	0.228

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Dependent variable in all columns is a continuous variable. Each coefficient estimate is based on a separate regression of the dependent variable on the willingness to take risks dummy and a set of controls, whose coefficient estimates are all reported except for the dummies for the state of residence (refer to FE indication) and the age quadratic trends. All specifications include a constant. Robust standard errors are reported in parentheses below the coefficient estimates.

Table 4. Child's Overweight PROBIT Results

VARIABLES	Child's Overweight Incidence					
	(1)	(2)	(3)	(4)	(5)	(6)
WTR	0.045***	0.0476***	0.0466***	0.0490***	0.0453***	0.0482***
	(0.015)	(0.0155)	(0.0154)	(0.0155)	(0.0154)	(0.0155)
HH Income	-0.024***		-0.0225***		-0.0243***	
	(0.005)		(0.00513)		(0.00520)	
Income		-0.00848*		-0.00791*		-0.00827*
		(0.00466)		(0.00466)		(0.00467)
Education	-0.032***	-0.0434***	-0.0335***	-0.0448***	-0.0339***	-0.0459***
	(0.009)	(0.00849)	(0.00892)	(0.00858)	(0.00891)	(0.00858)
Employment status					-0.00426**	-0.00309
					(0.00197)	(0.00195)
Parent's characteristics						
Gender	0.004	0.00429	0.00402	0.00420	0.00423	0.00437
	(0.010)	(0.0102)	(0.0101)	(0.0102)	(0.0101)	(0.0102)
Age	0.006	0.00213	0.00559	0.00215	0.00562	0.00194
	(0.007)	(0.00698)	(0.00703)	(0.00708)	(0.00701)	(0.00706)
Child's characteristics						
Gender	-0.031***	-0.0305***	-0.0325***	-0.0319***	-0.0317***	-0.0313***
	(0.010)	(0.00993)	(0.00984)	(0.00987)	(0.00984)	(0.00987)
Age	-0.003***	-0.00325***	-0.00341***	-0.00334***	-0.00343***	-0.00336***
	(0.001)	(0.000645)	(0.000641)	(0.000644)	(0.000641)	(0.000644)
State FE	No	No	Yes	Yes	Yes	Yes
Observations	5,011	5,011	5,011	5,011	5,011	5,011

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Dependent variable in all columns is a binary variable. Reported coefficients are Probit marginal effects estimates, evaluated at the means of independent variables. Each coefficient estimate is based on a separate regression of the dependent variable on the willingness to take risks dummy and a set of controls, whose coefficient estimates are all reported except for the dummies for the state of residence (refer to FE indication) and the age quadratic trends.

Table 5. Heterogeneous Effects of Parent's WTR on Child's Body Mass Index

VARIABLES	Child's Body Mass Index								
	(German)	(German 4-8)	(German female 4-8)	(German male 4-8)	(German mothers)	(German fathers)	(Females/ mothers)	(Females>8)	(Females>8/ mothers)
WTR	0.381**	0.468**	0.723**	0.294	0.344*	0.387**	0.629**	1.020**	1.358**
	(0.153)	(0.230)	(0.306)	(0.335)	(0.183)	(0.180)	(0.254)	(0.499)	(0.685)
HH Income	-0.180***	-0.221**	-0.304***	-0.130	-0.163**	-0.215***	-0.262***	-0.490**	-0.598**
	(0.058)	(0.0910)	(0.114)	(0.128)	(0.0694)	(0.0718)	(0.0924)	(0.225)	(0.282)
Education	-0.251**	-0.271	-0.428**	-0.175	-0.230	-0.270**	-0.378***	-0.184	-0.257
	(0.107)	(0.181)	(0.174)	(0.274)	(0.157)	(0.125)	(0.145)	(0.381)	(0.441)
Employment status	-0.030**	0.0237	-0.0390	0.0958**	-0.0209	-0.0429**	-0.0394	-0.112**	-0.152*
	(0.015)	(0.0252)	(0.0334)	(0.0384)	(0.0194)	(0.0180)	(0.0240)	(0.0565)	(0.0798)
Parent's characteristics									
Gender	0.061	0.00842	0.261	-0.216				-0.134	
	(0.100)	(0.172)	(0.167)	(0.246)				(0.369)	
Age	-0.057	-0.0860	0.0419	-0.158	-0.0372	-0.107	0.109	0.118	0.0827
	(0.081)	(0.160)	(0.222)	(0.198)	(0.0998)	(0.111)	(0.160)	(0.351)	(0.638)
Child's characteristics									
Gender	-0.281**	-0.287			-0.326**	-0.240*			
	(0.117)	(0.181)			(0.141)	(0.142)			
Age	-0.058***	0.147	0.166	0.199	-0.0585***	-0.0576***	-0.0562***	-1.965	0.0980
	(0.007)	(0.684)	(0.997)	(0.851)	(0.00834)	(0.00860)	(0.0105)	(1.222)	(1.283)
Constant	19.684***	12.15	9.687	11.32	19.41***	20.80***	16.08***	161.7*	8.003
	(1.691)	(24.86)	(35.38)	(31.07)	(2.010)	(2.411)	(3.188)	(91.63)	(97.99)
Observations	4,766	1,596	800	796	2,676	2,090	1,364	509	292
R-squared	0.236	0.051	0.079	0.072	0.237	0.251	0.291	0.170	0.207

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Dependent variable in all columns is a continuous variable. Each coefficient estimate is based on a separate regression of the dependent variable on the willingness to take risks dummy and a set of controls, whose coefficient estimates are all reported except for the dummies for the state of residence (refer to FE indication) and the age quadratic trends. All specifications include a constant. Robust standard errors are reported in parentheses below the coefficient estimates.

Table 6. Child's Obesity PROBIT Results

VARIABLES	Child's Obesity Incidence					
	(1)	(2)	(3)	(4)	(5)	(6)
WTR	0.042***	0.0441***	0.0432***	0.0446***	0.0423***	0.0440***
	(0.010)	(0.00981)	(0.00966)	(0.00992)	(0.00960)	(0.00990)
HH Income	-0.011***		-0.00959***		-0.0105***	
	(0.003)		(0.00296)		(0.00300)	
Income		-0.00799***		-0.00751***		-0.00776***
		(0.00276)		(0.00278)		(0.00279)
Education	-0.017***	-0.0206***	-0.0187***	-0.0220***	-0.0188***	-0.0226***
	(0.005)	(0.00462)	(0.00485)	(0.00468)	(0.00487)	(0.00470)
Employment status					-0.00227**	-0.00185
					(0.00116)	(0.00118)
Parent's characteristics						
Gender	0.006	0.00680	0.00629	0.00640	0.00660	0.00674
	(0.006)	(0.00573)	(0.00577)	(0.00579)	(0.00574)	(0.00576)
Age	-0.006*	-0.00745**	-0.00576	-0.00674*	-0.00581*	-0.00695**
	(0.003)	(0.00343)	(0.00351)	(0.00350)	(0.00350)	(0.00348)
Child's characteristics						
Gender	-0.001	-0.000762	-0.00203	-0.00175	-0.00156	-0.00138
	(0.006)	(0.00557)	(0.00560)	(0.00560)	(0.00561)	(0.00562)
Age	-0.000	-0.000280	-0.000317	-0.000264	-0.000321	-0.000264
	(0.000)	(0.000397)	(0.000398)	(0.000399)	(0.000398)	(0.000399)
State FE	No	No	Yes	Yes	Yes	Yes
Observations	5,011	5,011	5,011	5,011	5,011	5,011

*** p<0.01, ** p<0.05, * p<0.1

Dependent variable in all columns is a binary variable. Reported coefficients are Probit marginal effects estimates, evaluated at the means of independent variables. Each coefficient estimate is based on a separate regression of the dependent variable on the willingness to take risks dummy and a set of controls, whose coefficient estimates are all reported except for the dummies for the state of residence (refer to FE indication) and the age quadratic trends.

Table 7. Child's Body Mass Index OLS Results with parents' BMI

VARIABLES	Child's Body Mass Index					
	(1)	(2)	(3)	(4)	(5)	(6)
WTR	0.301**	0.316**	0.308**	0.322**	0.301**	0.318**
	(0.149)	(0.150)	(0.148)	(0.149)	(0.148)	(0.149)
HH Income	-0.155***		-0.152***		-0.164***	
	(0.053)		(0.0548)		(0.0551)	
Income		-0.0350		-0.0304		-0.0325
		(0.0445)		(0.0440)		(0.0440)
Education	-0.157	-0.238**	-0.177*	-0.260**	-0.179*	-0.266**
	(0.104)	(0.107)	(0.100)	(0.104)	(0.100)	(0.104)
Employment status					-0.0287**	-0.0203
					(0.0146)	(0.0146)
Parent's characteristics						
Parent's BMI	0.086***	0.0887***	0.0855***	0.0877***	0.0845***	0.0871***
	(0.011)	(0.0111)	(0.0113)	(0.0112)	(0.0113)	(0.0111)
Gender	0.081	0.0841	0.0880	0.0906	0.0883	0.0909
	(0.089)	(0.0877)	(0.0911)	(0.0900)	(0.0909)	(0.0899)
Age	-0.064	-0.0940	-0.0686	-0.0957	-0.0681	-0.0968
	(0.076)	(0.0738)	(0.0761)	(0.0750)	(0.0762)	(0.0750)
Child's characteristics						
Gender	-0.317***	-0.311***	-0.328***	-0.323***	-0.323***	-0.319***
	(0.111)	(0.111)	(0.111)	(0.110)	(0.111)	(0.110)
Age	-0.060***	-0.0602***	-0.0604***	-0.0603***	-0.0607***	-0.0604***
	(0.007)	(0.00671)	(0.00674)	(0.00673)	(0.00673)	(0.00672)
Constant	17.169***	17.71***	17.42***	17.93***	17.54***	18.04***
	(1.555)	(1.531)	(1.588)	(1.576)	(1.588)	(1.576)
State FE	No	No	Yes	Yes	Yes	Yes
Observations	5,011	5,011	5,011	5,011	5,011	5,011
R-squared	0.226	0.222	0.231	0.227	0.232	0.228

*** p<0.01, ** p<0.05, * p<0.1

Dependent variable in all columns is a continuous variable. Each coefficient estimate is based on a separate regression of the dependent variable on the willingness to take risks dummy and a set of controls, whose coefficient estimates are all reported except for the dummies for the state of residence (refer to FE indication) and the age quadratic trends. All specifications include a constant.

Table 8. Parents' Check Regressions

VARIABLES	——BMI——		——Overweight——		——Smoke——	
	(1)	(2)	(3)	(4)	(5)	(6)
WTR	0.291	0.342	0.0615***	0.0658***	0.0732***	0.0777***
	(0.243)	(0.243)	(0.0203)	(0.0203)	(0.0191)	(0.0192)
HH Income	-0.386***		-0.0292***		-0.0421***	
	(0.103)		(0.00733)		(0.00666)	
Income		0.0426		0.00903		-0.00748
		(0.0800)		(0.00656)		(0.00596)
Education	-0.508***	-0.768***	-0.0377***	-0.0601***	-0.0650***	-0.0882***
	(0.180)	(0.174)	(0.0127)	(0.0121)	(0.0114)	(0.0109)
Employment status	-0.109***	-0.0871***	-0.0132***	-0.0115***	-0.00580**	-0.00373
	(0.033)	(0.0320)	(0.00265)	(0.00263)	(0.00249)	(0.00246)
Gender	-0.299*	-0.278*	-0.0354**	-0.0328**	0.00189	0.00250
	(0.165)	(0.163)	(0.0144)	(0.0145)	(0.0131)	(0.0132)
Age	0.369***	0.290**	0.0248**	0.0182*	-0.0343***	-0.0421***
	(0.123)	(0.117)	(0.00991)	(0.00987)	(0.00868)	(0.00868)
Constant	21.454***	22.87***				
	(2.633)	(2.558)				
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,011	5,011	5,011	5,011	5,011	5,011
R-squared	0.033	0.027				

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix A

$$\max_{x^{h*}\in X} \Omega(x^{h*},a) \tag{11}$$

$$F.O.C.: \frac{\partial\Omega(x^{h*},a)}{\partial x^h} = 0 \tag{12}$$

$$\frac{\bigg(\lambda+(1-\lambda)\bar{a}\bigg)(\theta-1)}{(1-\theta)}(1-x^h)^{-\theta} - \frac{\beta R(\theta-1)}{(1-\theta)}(1+Rx^h)^{-\theta} = 0$$

$$-\bigg(\lambda+(1-\lambda)\bar{a}\bigg)(1-x^h)^{-\theta} - \beta R(1+Rx^h)^{-\theta} = 0$$

$$-\bigg(\lambda+(1-\lambda)\bar{a}\bigg)(1-x^h)^{-\theta} = \beta R(1+Rx^h)^{-\theta}$$

$$-\frac{\bigg(\lambda+(1-\lambda)\bar{a}\bigg)}{\beta R} = \frac{(1+Rx^h)^{-\theta}}{(1-x^h)^{-\theta}}$$

$$-\bigg(\frac{\lambda+(1-\lambda)\bar{a}}{\beta R}\bigg)^{\frac{1}{\theta}} = \frac{(1-x^h)}{(1+Rx^h)}$$

$$-(1+Rx^h)\bigg(\frac{\lambda+(1-\lambda)\bar{a}}{\beta R}\bigg)^{\frac{1}{\theta}} = (1-x^h)$$

$$x^h\Bigg[1+R\bigg(\frac{\lambda+(1-\lambda)\bar{a}}{\beta R}\bigg)^{\frac{1}{\theta}}\Bigg] = 1-\bigg(\frac{\lambda+(1-\lambda)\bar{a}}{\beta R}\bigg)^{\frac{1}{\theta}}$$

$$x^{h*} = argmax\Omega(x^{h*},a) = \frac{1-\bigg(\frac{\lambda+(1-\lambda)\bar{a}}{\beta R}\bigg)^{\frac{1}{\theta}}}{1+R\bigg(\frac{\lambda+(1-\lambda)\bar{a}}{\beta R}\bigg)^{\frac{1}{\theta}}} \tag{13}$$

CHAPTER 2

Exploiting Persistence to Extract a Signal of Hospital Quality for Italian Regions

Joint work with Marina Di Giacomo, Luca Pieroni and Luca Salmasi

1. Introduction

An instructive and widespread aphorism is the following: "A data becomes information when it is able of changing the probability of decisions". This is true in several contexts but it applies very well to the case of healthcare. The lack of good information on performance or quality, in fact, is a core problem in many areas of public policy and policy evaluation. The difficulty of developing reliable information on the quality of healthcare providers for guiding public policies and individual choices is perhaps the most striking example. Availability of reliable measures of hospital quality (henceforth: HQMs) is important because they affect how individuals make health decisions (i.e. hospital choice). In fact hospital quality drives individuals' hospital choice together with other factors like distance (time and cost of transportation), waiting time, volumes, etc. In Italy, where the hospital quality measures are not widespread, in order to understand the individuals' hospital choice it could be useful to analyze the quality of the national health system at an aggregate level. The goal of this research is to consider a range of health outcome measures related to heart disease over time to develop and implement statistical techniques to extract a signal of health quality for Italian regions, smoothing standard health quality measures across hospitals and across years. At first, we extract regional fixed effects from data at hospital level and we highlight regional differences in the healthcare provision for the set of HQMs considered in the analysis. Then, we adapt vector auto-regression (VAR) methods for panel data to estimate the systematic relationship across outcomes and over time at hospital and regional level. Our scope is that these adjusted signals of quality would enrich the awareness of the differences in the quality of healthcare across regions and help explaining patients' mobility.

The chapter proceeds in the following way. In Section 2 we introduce the Italian National Health Service in detail and in separate subsections we analyze its regional structure and financial sources focusing on regional differences and patients' mobility. Section 3 describes our dataset, the role of indicators and the hospital quality measures included in our analysis. Section 4 presents the two-step multilevel methodology.

Section 5 establishes the main results for both the fixed effects and the vector auto-regression estimation . Finally, Section 6 concludes.

2. The Italian National Health Service

The Wold Health Organization (WHO) rated the Italian healthcare system as one of the best in the world. Italy's life expectancy is the 4^{th} highest among OECD countries with a per capita healthcare spending well below the average of other high-income countries (OECD (2017)). Despite this success, there are significant regional differences in the quality of healthcare and, as a consequence, in the health status of the Italian citizens. For example, average life expectancy is 82.3 years, but this value ranges from 83.5 years (81.2 for men and 85.8 for women) in Trento (North of Italy) to 80.5 years in Campania (78.3 for men and 82.8 for women) (South of Italy). A similar trend is observed for the reduction in mortality over the last 15 years: 27% in the North; 22% in the Centre and only 20%in the South (Osservasalute (2016)).

To understand the reasons behind these regional differences it is important to describe the characteristics of the Italian national health system from its origins and through the reforms that have contributed to making it as it is now.
The National Health Service (NHS) was established in 1978 (Art.1, L. 833/1978) and is composed by a system of institutions, both public and private, highly complex. The main principles on which the law of healthcare reform is founded are the following:

- universality of the right of health assistance for all people and all kinds of illness, without any discrimination;
- management of the supply of services committed to the U.S.L.s/A.S.L.s;
- equality of citizens and uniformity of the treatment in all the country.

The National Health Service was deeply reformed in 1999 (Art., Decree Law 229/1999) when it was transformed in a regional system giving a new role and responsibility to the regions. The central government has the responsibility of legislating, establishing the essential assistance levels (LEA), programming the healthcare policy and partially financing the national health service possibly intervening in the case of excessive deficit of the regions.

The regions, according to the Italian Constitution (Art. 117), have competence on the healthcare assistance on their territories, the responsibility of legislating within the framework of fundamental principles established by the central government, the faculty of collecting the regional taxes using them to finance the local health districts (which correspond to the ASLs, "Aziende Sanitarie Locali"). Regarding the organization of the National Health Service, the local health districts and the hospitals operate at a local level. The ASLs have the responsibility to supply healthcare services to the population of a given territory (district), directly with their own structures or through private suppliers, to organize and program the development of services and to allocate resources to the latter. On the other side, the hospitals are large structures, financed by ASLs, whose primary functions consist in making the activities of recovery of patients and offering specialized cares.

2.1. Regional healthcare systems. Since the ASLs are the institutions in charge of providing healthcare services to citizens, an important decision that has to be made by regional governments is in how many ASLs to divide their territory. Some regions prefer to have many small ASLs: this is, for example, the choice of Veneto, which currently has about twenty ASLs, with an average population of 235,000 inhabitants. Other regions have instead decided to have few larger ASLs: in Campania, for example, the seven local health authorities have an average population of around 837,000 inhabitants. The Marche has even set up a single regional healthcare company, which takes care of over 1,500,000 residents. This is also the direction of recent reforms in Umbria, which in 2012 has reduced the number of ASLs from 4 to 2, or in Tuscany, which has reduced the number of ASLs from 12 to 3. This decision is often driven by the aim of exploiting scale economies to improve efficiency.

In addition to the dimensions of the ASLs, the regions can choose whether to leave the hospitals under the management of the ASLs or transform them into autonomous hospital companies. In this regard, two models can be distinguished: the integrated one and the separate one. In the integrated model, hospitals remain under the control of the ASLs: this should favor the coordination between hospital and territorial care.

In the separate model, hospitals are instead separated from their respective ASLs and transformed into autonomous companies: this model, stimulating competition between the different hospital structures, reflects more the theoretical model of the "internal market". In our country, only one region has adopted with conviction the separate model: Lombardy. All other regions have either an integrated system or at most a mixed system (as in Piedmont). In this case the discussion underlying the decision of adopting one system rather than the other is on the positive or negative effect of enhancing the competition in the healthcare sector (Propper et al. [2004], Gaynor et al. [2016]).

A further strategic choice that regional governments must make concerns the involvement of private health. Each region is in fact free to decide which part of the services to provide with its own facilities and with its own staff, and which to outsource to private suppliers (nursing homes, private clinics, staff not employed by the NHS). The regions that mostly use private suppliers are Lazio, Lombardy and Puglia (they outsource more than 40% of regional health expenditure). Instead, the province of Bolzano, Valle d'Aosta, Friuli-Venezia Giulia, Umbria and Tuscany have an eminently public supply structure (less than a quarter of public health expenditure is outsourced). In general, the southern regions make more use of private suppliers than the central-northern regions do.

The regions enjoy wide discretion in reference to many other relevant issues, not only concerning the organization but also the financial management of the healthcare system. For example, each region is free to fix the tariffs through which to repay the suppliers (both public and private) independently. Tickets also change from region to region. Consider for example the ticket on drugs: in some regions it is not provided, in others it consists of a fixed quota, in others it is modulated based on income. So for the same pharmaceutical prescription you can pay 8 euros in Tuscany, 4 euros in Lombardy, 2 euros in Calabria, a 1 euro in Trento; in Friuli, Valle d'Aosta, Marche and Sardinia you pay nothing. In relation to the autonomy of economic management we will dedicate the next Section to describe in detail the financial resources of the NHS and of the regional healthcare systems.

2.2. The financial sources of the Italian healthcare system. The Italian law determines annually the overall level of the resources of the National Health Service (NHS) funded by the State. The standard national health requirement is determined, by agreement, in line with the overall macroeconomic situation and in compliance with public finance constraints and the obligations assumed by Italy in the European Union, coherently with the needs deriving from the determination of the essential assistance levels (LEA) provided in conditions of efficiency and appropriateness. The amount allocated to the ordinary statute regions and the quotas destined to institutions other than the regions are distinguished. The financing of the NHS was designed by Legislative Decree 56/2000 which provided for a system of financing of the NHS based on regional fiscal capacity, even if corrected by appropriate equalizing measures, establishing three sources of revenues: IRAP, the regional supplement to IRPEF (the personal income tax) and a quota of VAT.

The needs for the healthcare system are therefore financed by the following sources:

- revenue of the NHS bodies (tickets and revenues deriving from the intra-moenia of their employees);
- general taxation of the regions: IRAP (in the revenue component destined for health) and regional additional tax for IRPEF. General taxation, in its distinct IRAP and additional personal income tax components, passes through the Treasury accounts. The resources relating to the two taxes are paid to the regions each month in full (decree law 112/2008, article 77-quater);
- co-participation of the regions with special statutes and of the autonomous province of Trento and Bolzano: these institutions participate in health financing up to the requirements not met by the sources mentioned in the previous points, except for the Sicilian region, for which the co-participation rate it has been set since 2009 to the extent of 49.11% of its health needs (law n. 296/2006 article 1, paragraph 830);
- National budget: it finances the health needs not covered by other sources of funding essentially through the sharing of value added tax - VAT (destined for regions with ordinary statutes), and through the National Health Fund (a

quota is allocated to the Sicilian region , while the rest also finances health costs linked to certain objectives).

The fact that the health system is financed by regional revenues is one of the factors that can contribute to determining and explaining the gap between regions (Dirindin and Eva [2001]).

2.3. Regional and macro area performance gap: differences between healthcare systems. Each region has its own rules and its own organizational structure. The analysis of the relationship between alternative regional approaches and the healthcare quality provided is crucial in order to understand if there is a model - at least on paper - better than the others. If we look only at the organizational model, and not at the results it produces, it is not easy to answer that question. For example, consider the question of the size of ASLs: it is not clear whether it is better to have small ASLs or large ASLs. There is no agreement among the experts on the optimal size of a healthcare company, and there is no solid empirical research that sheds light on the issue. But if we investigate the results produced by each healthcare system is possible to make comparisons between them. Indeed, individual regional health systems perform very differently. And a clear gap emerges between the north and south of the country.

The most critical aspects regard the financial balance, a source of constant friction between the national government and regional administrations. In fact, few regions have been able to keep their accounts in order; among these are Friuli-Venezia Giulia, Lombardy, Veneto, Emilia-Romagna and Umbria. Other regions have systematically breached the budget. From 2001 onwards the regional health systems have accumulated a total deficit of almost 38 billion euros. However, the central and northern regions were responsible for just 13% of these losses. Southern regions, and some of the central ones including Lazio, have accumulated over 87% of the deficit. The central-northern regions, from 2001 to 2014, have on average accumulated a deficit of 139 euros per capita. The deficit of the southern regions - always per capita, and always in the same time frame

- was instead of well 1,235 euros: almost nine times higher than that of the central-northern regions. The absolutely most unruly regions were Campania and Lazio, which - alone - are responsible for 57% of the total health debt.

However, it is reductive to evaluate the goodness of a health system in economic terms; the focus of this research is on the state of health of the population and the quality of the services offered. But it is precisely here that the clinical picture becomes more complicated. The Ministry of Health, for example, monitors each year the extent to which the regions are able to provide the essential levels of assistance (LEA) through the National Outcomes Program (Piano Nazionale Esiti, PNE) (See Section 3). The latter constitute the package of care we are all entitled to and which should be provided uniformly throughout the entire national territory. From the monitoring of the ministry it emerges how the Central-Northern regions manage well (from a couple of years the first positions of the ranking are occupied - in order - by Tuscany, Emilia-Romagna and Piedmont); those of the South occupy the last positions of the ranking.

A recent study by the Center for the Economic Research applied to Health (Centro per la ricerca economica applicata in sanità , C.R.E.A. Sanità) provides a multi-dimensional assessment of the performance of individual regional health systems. Also here: the central-northern regions are all in the upper part of the ranking, the southern ones in the lower part. Many other indicators of efficiency or appropriateness could be considered. For example, the National Outcomes Program edited by AGENAS (National Agency for Regional Health Services), or the "targets" of Sant 'Anna di Pisa. The situation unfortunately does not change: all the rankings agree that in Italy we have higher quality health services in the Center-North and a lower quality healthcare, sometimes much lower, in the South.

2.4. Healthcare quality and patients' mobility. Citizens are aware of this gap in performance between regional healthcare systems. As emerges from the last Censis Report on the social situation of the country, 83% of the inhabitants in the South considers their regional health service "not adequate". This percentage is much lower in the northern regions: in the North-East the dissatisfied are in fact 35%, in the North-West less than 30%.

It is therefore not surprising that residents in the southern regions - when they can - go and seek treatment elsewhere. This is the well-known phenomenon of healthcare mobility: every year about half a million patients are hospitalized in a region other than that of their own residence. Also on this front, the North-South imbalance is evident. For every patient residing in the Center-North who is admitted to a hospital in the South, there are six who make the reverse journey, or rather from the southern regions go to be treated in hospitals in the Center-North. Lombardy, Emilia-Romagna and Tuscany are particularly attractive. From other regions, patients tend to flee: this is especially the case in Calabria, Campania and Sicily. All southern regions, with the exception of Molise, have a negative health mobility balance.

A deeper understanding of regional disparities - based not only on a mere comparison between indicators but by splitting the temporal component - can provide useful policy indications to manage the mobility flows of patients with requests for elective services and push the Ministry of Health to invest extraordinary resources in the health systems of the regions with a worse quality, also exercising a more accurate control on the use of the allocated resources. In the next Sections we will describe the dataset and the empirical model used to isolate the persistence effect and evaluate regional performance with more accuracy.

3. Dataset

The National Outcomes Program (Piano Nazionale Esiti, PNE) is a tool[1] for measuring, analyzing, evaluating and monitoring the clinical-welfare performance of health facilities available to the regions, companies and professionals for continuous improvement of our NHS. The results of PNE are published annually on the dedicated website. The indicators used to analyze the results of treatments, scientifically validated at international level, are aimed at achieving the following objectives:

- Continuous improvement of efficacy and appropriateness of care;

[1] Data collection for PNE is managed by AGENAS on behalf of the Ministry of Health.

- Greater equity of access to services of proven efficacy throughout the national territory, regardless of the area of residence;
- Transparency and empowerment of citizens and associations, with the dissemination of clear and scientifically validated information;
- Internal and external audit to identify possible critical issues in the quality of the data and in the clinical and/or organizational processes.

The data sources used for calculating the indicators are the Health Information Systems (Sistemi Informativi Sanitari, SIS) and the administrative sources for their accessibility and for their ability to provide information on the totality of healthcare provided by health facilities operating within the SSN. Furthermore, the use of these information sources allows systematic monitoring over time of the indicators included in the PNE. Currently, the hospital information system (Sistema Informativo Ospedaliero, SIO) is used, which collects information on all hospital admissions (acute and post-acute) for each patient discharged from public and private institutions throughout the national territory validated through the linkage with data from the hospital tax register (Anagrafe Tributaria, AT). Data derived from electronic archives are integrated through record linkage techniques with the aim of integrating the information present in different archives or in the same archive in different periods.

3.1. The role of indicators. The process that leads to the definition of an outcome indicator starts with a systematic review phase of the scientific medical literature related to the treatment or therapeutic diagnostic path to be evaluated. The information derived from this first revision phase allows defining a first version of the protocol to be used to carry out the preliminary analyzes that will allow to verify the validity of the indicator. The indicators are documented by protocols with explicit definition of the outcome in the study, of the selection criteria of the cases, of the follow-up times, of the data sources and of the factors used for the risk-adjustment. The indicator protocol and the results of the preliminary analyzes are subject to evaluation by representatives of the scientific societies of reference, of panels of expert clinicians and further discussed within the PNE Committee. There are different kinds of indicators. Among them the most important are outcome indicators and process indicators. The outcome indicators measure the outcome of a care process in terms of clinical outcomes (i.e. mortality, morbidity, hospitalization). Their relationship with the measured phenomenon is influenced by various determinants that are not directly correlated with the quality of the care process (risk markers, environmental factors, socio-economic variables) and which must be considered and possibly corrected during the calculation of the indicator . The robustness of the outcome indicators also depends on the time between the measurement and the actual delivery of the health service. Process indicators measure the degree of adherence of the care process to the reference standards of the best clinical practice based on evidence. For this reason they are considered proxies for the outcomes of assistance and their robustness, understood as predictive of clinical outcomes, depends on the strength of the clinical recommendation and the degree of evidence on which they were built. In addition to the outcome and process indicators, indicators that report volumes are calculated, for health interventions for which scientific evidence of the effectiveness of association between activity volumes and treatment outcomes is available.

3.2. Health Quality Measures in the analysis. In recent years PNE has constantly increased the number of indicators, assessed and selected. In particular, they increased in the angiological, orthopedic and pediatric area, going from 146 in 2015, to 165 in 2016, up to 166 indicators in 2017 (67 outcome and process, 70 volumes of activity and 29 hospitalization indicators). For this analysis, we exploited the following cardiovascular indicators belonging to the PNE:

- Acute Myocardial Infarction: 30 days mortality
- Acute Myocardial Infarction without PTCA: 30 days mortality
- Acute Myocardial Infarction with PTCA within 2 days: 30 days mortality
- Acute Myocardial Infarction with PTCA later than 2 days from recovery: 30 days mortality
- Acute Myocardial Infarction: percentage of treated with PTCA within 2 days
- Acute Myocardial Infarction: 365 days readmission rate (MACCE)

- Acute Myocardial Infarction: percentage of treated with PTCA within 7 days

In PNE, risk adjustment techniques are used which consist in the construction of a measure of gravity that describes the "clinical complexity" of the patient, based on the characteristics of the patient, the severity of the pathology in study and the concomitant pathologies of the patients, and in the use of such measure of gravity to obtain relatively risk-adjusted indicators, which allow a valid comparison between hospitals. Italian hospital performance is considered over the period 2008-2016. This gave an average of 82.421 cases each year for 283 hospitals involved in treating AMI patients and reporting results to PNE. We decided to exclude from the analysis hospitals located in regions with special statutes, assuming that due to the different financial funding (see Section 2.2), their regional healthcare systems are not comparable to those of the other regions. After this restriction, we have 276 hospitals for 80,479 cases each year.

The descriptive statistics are reported in Table 3 and Table 4[2]. Table 3 shows that in the sample we have 13 regions and 74 districts. The region with the highest number of hospitals(50), districts (12) and cases (15,248) is Lombardy, while the smallest one is Marche. we decided to consider Abruzzo and Molise as a unique region, due to the dimension of Molise and their geographic proximity. Table 4 shows a significant higher concentration of cases and hospitals in the North compared to the South but the reader must consider that these are not the true numbers but those reported from regions to PNE. The fact that Southern regions underreport their parameters is signaling again the poor quality of the healthcare system from the administrative perspective. Table 2 shows the 30-day mortality, the 365-day readmission and the ratio of 30-day CRM for treated w/PTCA versus not treated for the sample of hospitals across all years (for the first two HQMs reported per hundred deaths or readmissions). One can see from Table 2 that the trend for both, the 30-day mortality rate and the 365-day readmission rate, is decreasing over the time span considered. On the other hand, the trend of the ratio of 30-day CRM for patients treated with PTCA over CMR of not treated is increasing,

[2]For the same statistics with the inclusion of the regions with special statutes see Table 10 and Table 11 in Appendix A.

probably due to the decrease of the general mortality rate, thus confirming the positive direction that the Italian healthcare system,in its entirety, is pursuing.

4. The Empirical Model

The objective of this study is to return a reliable regional health quality measure (HQM) to investigate heterogeneity in healthcare provision in Italy at three levels of analysis:

- Hospital level
- Regional level
- Macro Area level

Since the health quality cannot be assessed directly, the valuable information on this measure can be returned through a two-step smoothing procedure proposed by Papanicolas and McGuire [2017], following the approach of McClellan and Staiger [1999]. According to the data availability, at hospital level only the second step of the smoothing procedure has been performed, while at regional and macro area level we proceed following the complete procedure.

4.1. First step: Fixed effects estimation. The first step of our analysis uses three unadjusted health outcome measures (30-day mortality, 365-day readmission rate and treatment with PTCA within 2 days) for the acute myocardial infarction (AMI) at hospital level and adjusts them for risk through linear regression against hospital's and district's characteristics.

The following first-stage, risk adjustment regression equation is run on each of the hospital's HQMs for each year of the analysis through ordinary least squares separately:

$$Y_{hr} = \mu_r + V_{hr} + X_{dr} + u_{hr} \tag{14}$$

where Y represents the HQM, h indexes the hospital and r the region to which the hospital belongs, V is a control for the volumes treated in the hospital and X represents a set of controls at the district level, such as the population over 55 years, the average income, the temperature and the employment rate.

Note that the μ_r in equation 14, estimated through the incorporation of dummy variables, are of greatest interest as they return a regional fixed effect, which can be considered a proxy measure of regional hospital quality. These μ_r are, therefore, estimates of regional health quality for each of the three HQMs gained through risk adjustment for hospital's and district's characteristics. As noted, equation 14 is run separately for each of the 9 years and for each of the three HQMs.

The regional fixed effects, returned from each of the yearly regressions, are used to construct a new vector, Q_r, of risk-adjusted regional quality, for each of the three HQMs.

Assuming T yearly time periods and K measures of quality, the hospital quality vector Q_r, which is constructed from each of the yearly regressions, has dimensions K. In our case with 3 measures of quality and 9 years of observations over the period 2008-2016, the vector Q_r has dimension 1×27 .

The vector Q_r is then assumed to represent the following relationship to the true regional health quality:

$$Q_r = q_r + \epsilon_r \tag{15}$$

where q_r represents the $1 \times TK$ vector of the 'true' underlying quality for region r, and ϵ_r is the estimation error (which is assumed to have mean 0 and to be uncorrelated with q_r).

Thus, equation 15 assumes that the estimated risk-adjusted regional fixed effects Q_r are suitable predictors of true quality, and anything that is not captured by these estimates is incorporated in the error term ϵ_r.

It is the removal of the error term ϵ_r from the estimated regional fixed effects Q_r which allows further improvement in the measures. The error term ϵ_r is related to the hospital level regressions (equation 14), in particular, to the variance-covariance of the regression estimates Q_r, i.e.

$$E[\epsilon_r' \epsilon_r] = S_r \tag{16}$$

where S_r represents the variance-covariance matrix of the regional effects estimates for region r for each year obtained by ordinary least squares estimation of equation 14 with the normal assumption that the off-diagonal elements of the covariance of the disturbances are all equal to 0.

Since the true latent regional quality measure q_r is not directly observable, adapting the McClellan and Staiger [1999] method, it is possible to create a linear combination of each regional observed risk-adjusted measures of quality for each year, in such a way that it minimizes the mean-squared error of the predictions. This could be conceptualized as running the following (hypothetical) regression for each year:

$$q_r = Q_r \beta_r + \omega_r \tag{17}$$

They noted, however, that equation 17 cannot be estimated directly, precisely because q_r represents the true unobserved quality for the defined outcomes, in each region r for each year. Assuming K measures of quality and T years, note that Q_r is a $1 \times TK$ vector and the optimal β for each quality measure k varies by hospital and year, given equation 15. The measurement challenge is to return the true hospital quality q_r for each quality measure k in each year, from the noisy estimate Q_r. We exploit McClellan and Staiger [1999] and Jones and Spiegelhalter [2012]'s insight that, although equation 17 cannot be estimated directly as q_r is not observed, the parameters of the hypothetical ordinary least squares regression represented by equation 15 and equation 17 can be retrieved from the existing data on the basis of the specified relationship between unobserved latent quality, observed quality and the error terms. In particular, the minimum least squared estimate, for each of the k quality measures over each of the t time periods, can be given by:

$$E(q_r|Q_r) = Q_r \beta \tag{18}$$

where:

$$\beta = E(Q_r' Q_r)^{-1} E(Q_r' q_r) \tag{19}$$

This best linear estimate can be returned by using the definitions:

$$E(Q_r'Q_r) = E(q_r'q_r) + S_r \tag{20}$$

$$E(Q_r'q_r) = E(q_r'q_r) \tag{21}$$

where $E(Q_r'Q_r)$ is the expected value of the products and cross-products of the regional fixed effects, which is gained from the first-stage hospital level regressions, and where $S_r = E(\epsilon_r'\epsilon_r)$ is the variance-covariance matrix of the disturbances that are associated with these fixed effects, which again is constructed from the first-stage hospital level regressions. $E(q_r'q_r)$ can be estimated by rearranging equation 20 such that $E(Q_r'Q_r - S_r) = E(Q_r'q_r)$. Subsequently equation 21 becomes $E(Q_r'Q_r - S_r) = E(q_r'q_r)$. Using these estimates and equalities, and inserting the relevant estimates into equation 19 allows derivation of the desired least squares parameters in equation 17 . The q_r can then be easily estimated by region for each year by using observed values.

$$\hat{q}_r = Q_r E(Q_r'Q_r)^{-1} E(Q_r'q_r) = Q_r\{E(q_r'q_r) + S_r\}^{-1} E(q_r'q_r) \tag{22}$$

4.2. Second step: the Vector Auto-regression. A further step in this smoothing procedure is represented by equation 23. This further step utilizes the information across the different time periods to improve these risk-adjusted, filtered quality outcome measures, $\hat{q}_r$, additionally. We consider this method a form of bidirectional smoothing estimation, because the measures reduce noise within regions, and across time periods. This is undertaken by using a vector auto-regression (VAR) model, with further structure imposed on the filtered quality estimates by assuming that each quality measure is reflective of its past performance, plus a contemporaneous shock that may be correlated across the different outcome measures. Noting that we have K measures of quality, which are interrelated and contain signals from past performance, and T years a second-order VAR model is specified to return the estimate $\hat{q}_{rt}$ for each of the k measures of quality. The estimate $\hat{q}_{rt}$ is a $1 \times K$ vector derived from estimating the auto-regressive process:

$$\hat{q}_{rt} = \hat{q}_{r,t-1}\Phi + \hat{q}_{r,t-2}\Phi' + v_{rt} \tag{23}$$

where Φ is the $K \times K$ matrix containing the estimates of the first lag coefficients of each of the HQMs and Φ' is the $K \times K$ matrix containing the estimates of the second lag coefficients of each of the HQMs.

Using the parameters that are estimated from the VAR model, we can estimate equation 23 to return non-stochastic smoothed estimates of quality for each region, incorporating the times series data; in this way the smoothing process has been applied to our measures bidirectionally, across waves and across hospitals. McClellan and Staiger (1999) referred to the results obtained through this two-step smoothing procedure as "predicted" estimates, while Papanicolas and McGuire (2016) adopted the term (bidirectional) smoothed estimators.

In addition to the implementation of the VAR procedure to the fixed effects obtained as result of the first step of the analysis, we tested the same specification at hospital level. The specification obtained is the following:

$$q_{ht} = q_{h,t-1}\Phi_h + q_{h,t-2}\Phi'_h + v_{ht} \tag{24}$$

The relevant value of this second step of the procedure, which consists in the application of the VAR, is that it allows the further production of forecast measures of healthcare quality. This is is of main interest to researchers and policy makers in particular at the regional level because it can offer insights about future mobility across hospitals and not only across regions. Forecast measures at both hospital and regional level can give useful policy indications. We will discuss these aspects related to policy implications in Section 6.

5. Results

We applied the two-step methodology described above to data on the three Health Quality Measures (30-day mortality rate, 365-day readmission rate and the ratio of

30-day crude rate mortality for treated with PTCA versus not treated) focusing on the treatment of AMI. We applied the first step of the methodology to the hospital data to extract a signal of quality for each region and each macro area. In the second step of the methodology, we applied the vector auto-regressive model to data at three levels: at hospital level, considering each hospital as an individual record; at regional level, considering for each region the fixed effect obtained in the first step and its temporal lags; at macro area level, considering for each macro area the fixed effect obtained in the first step and its temporal lags.

5.1. Results of the first step.

5.1.1. *Regional fixed effects.* The first step of the methodology has been used to extrapolate a regional and a macro area fixed effect from hospital data for each of the Health Quality Measures. Results of the regional fixed effects are reported in Table 5. As can be read in Table 5 and also seen from Figure 1, the regions which show a lower 30-day mortality rate - colored in dark blue - are: Emilia Romagna, Tuscany, Marche, Lombardy and Piedmont. A 30-day mortality rate higher than 10%, associated with a worse health quality, characterizes three regions of the South of Italy: Abruzzo, Molise and Puglia.

Looking at the second Health Quality Measure - the 365-day readmission rate - the regions with the lowest level of readmissions are Lazio and Piedmont, whereas the ones with highest rate of readmission are Veneto and Lombardy. The regional distribution of the readmission rate is represented in Figure 1, where regions with a lower readmission rate are colored in dark blue while in red there are those with the higher levels. Although readmission rates are a popular health quality measure, with higher emergency readmissions in particular thought to be indicative of worse quality, it cannot always be attributed to the overall quality of care delivered by the hospital (Fischer et al. [2011]). McClellan and Staiger [1999] and Laudicella et al. [2013] noted that high readmissions may not signal poor quality when hospital treatment is lowering mortality rates and more severely ill patients are surviving initial disease episodes. Under such circumstances higher readmission rates might be expected. Moreover, readmissions may reflect poor quality care in other parts of the healthcare system (e.g. the primary care sector), or individual behavioral factors beyond hospital control (e.g. poor adherence to medicines). Benbassat and Taragin [2000] concluded that readmission indicators are not good measures of quality of care for most conditions, as there is large variation in the percentage of this indicator that can be attributed to poor quality care. Their own study, using different readmission indicators for a range of conditions, estimated the variation for readmissions associated with improved quality of care to be between 9% and 50%. They did note that readmissions for specific conditions, such as childbirth, coronary artery bypass grafting and acute coronary disease, as well as approaches that ensure closer adherence to evidence-based guidelines, may provide more appropriate measures of quality. However, after initial use in the USA, there are now a growing number of European countries which employ readmission rates as a health service outcome measure (Fischer et al. [2011]) and attach financial incentives to them (Tunçalp et al. [2015]). For this reason we decided to include this health quality measure that is intended to be evaluated together with the trend of the crude mortality rate.

The third health quality measure analyzed is supposed to be an indicator of the quality of care for patients with AMI that are treated with PTCA. The precise definition of the measure is the ratio of the 30-day Crude Mortality Rate of patients treated with PTCA over the general Crude Mortality Rate (i.e. the first Health Quality Measure in Table 5). As one can see in the third column of Table 5, as well as from the third map of Figure 1, Lazio, Abruzzo and Lombardy are the regions with a higher 30-day mortality for patients treated with PTCA. This indicator has an ambiguous value because it can capture a high mortality associated with the administration of PTCA, which is a bad signal, but also a low general mortality rate, which is good. Unlike the other two measures, the results do not show significant regional differences. The homogeneity of incidence of mortality in patients treated with PTCA indicates a general widespread appropriateness of the administration of this practice to cases that require it.

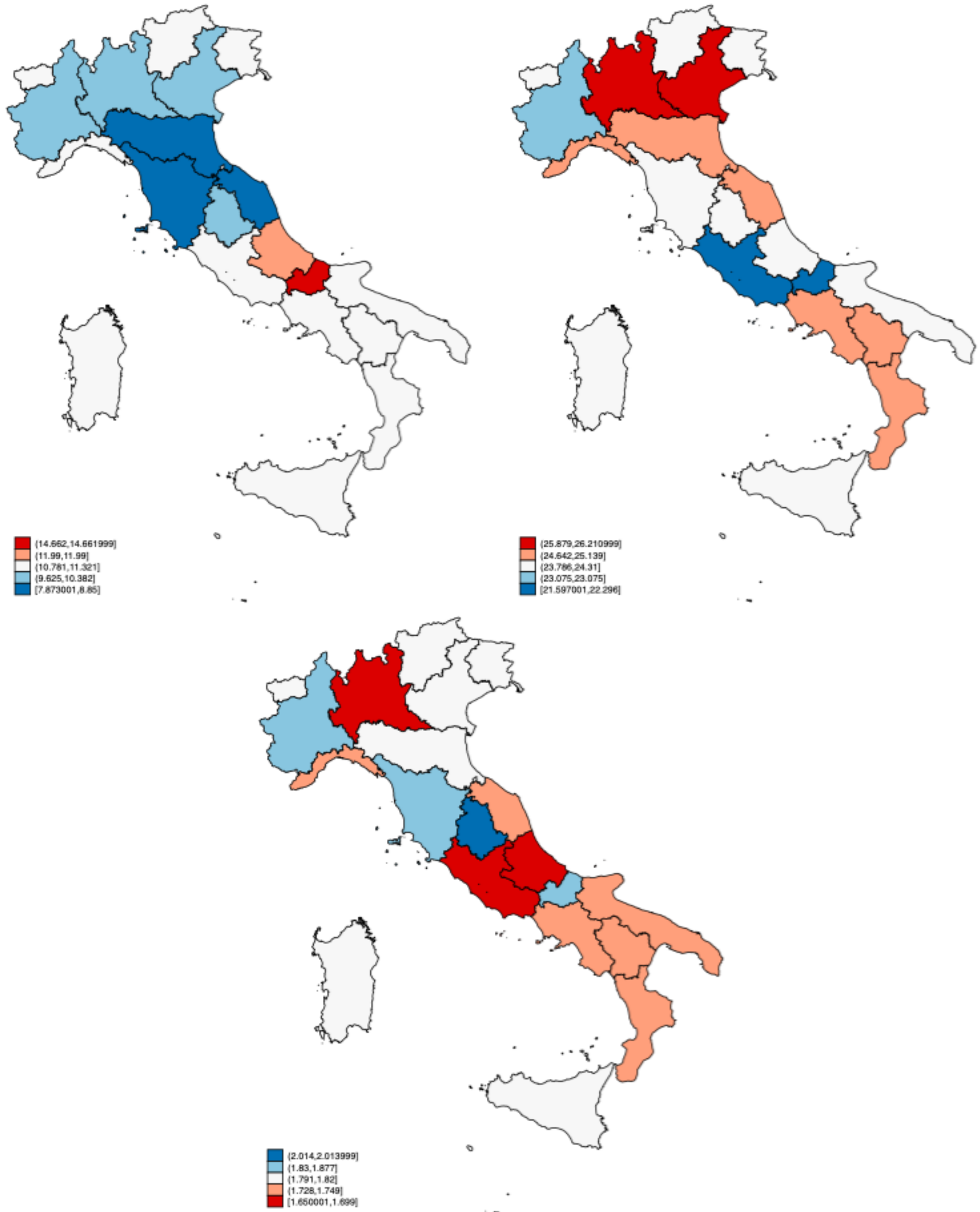


Figure 1. Regional distribution of 30-day mortality rate, 365-day readmission rate and Ratio of 30-day CRM for treated w/PTCA versus not treated

5.1.2. *Macro Area fixed effects.* We replicate the first step of the analysis for macro areas in order to get a broader perspective of geographical differences in the healthcare systems. Results of the macro area fixed effects are reported in Table 6. The results for the three health quality measures (30-day mortality rate, 365-day readmission rate and the ratio of 30-day crude mortality rate for treated with PTCA versus not treated) are respectively in the first, second and third column of the Table.

Starting looking at the 30-day crude mortality rate, as one can see there is a significant difference between North and South, with the South area showing a CMR 17% higher than the North. This confirms the expectations we had from the results obtained in the regional analysis. Results for the second HQM are controversial (see 5.1.1): the North shows a higher readmission rate at 365 days with respect to both South and Center. As we pointed out in the previous Section, this finding is not surprising. The combination of a high readmission rate with a low mortality rate can be a signal of good healthcare quality: people die less but as a consequence the readmission rate in the long run increases. Secondly, we need to take into account the interregional mobility in the case of readmissions due to non-emergency complications. This mobility will be driven by the desire to seek treatment in hospitals with better quality, whose regions will therefore register a higher readmission rate.

The third HQM analyzed is the ratio of 30-day crude mortality rate for treated with PTCA versus not treated. From Table 6 is possible to notice that North area and South area show the same performance, whereas Center area has a higher mortality rate of patients treated with PTCA over not treated patients, even if the difference is not consistent. A consistent and significant higher value would have suggested inappropriateness in the administration of PTCA treatment. The geographical distribution per area of each of the three HQMs is shown in Figure 2.

In both, the regional analysis and the analysis per macro area, we included control variables for district and hospital characteristics: the population over 55 years, which is exposed to a higher risk for AMI and cardiovascular diseases, and the volumes in the hospital.

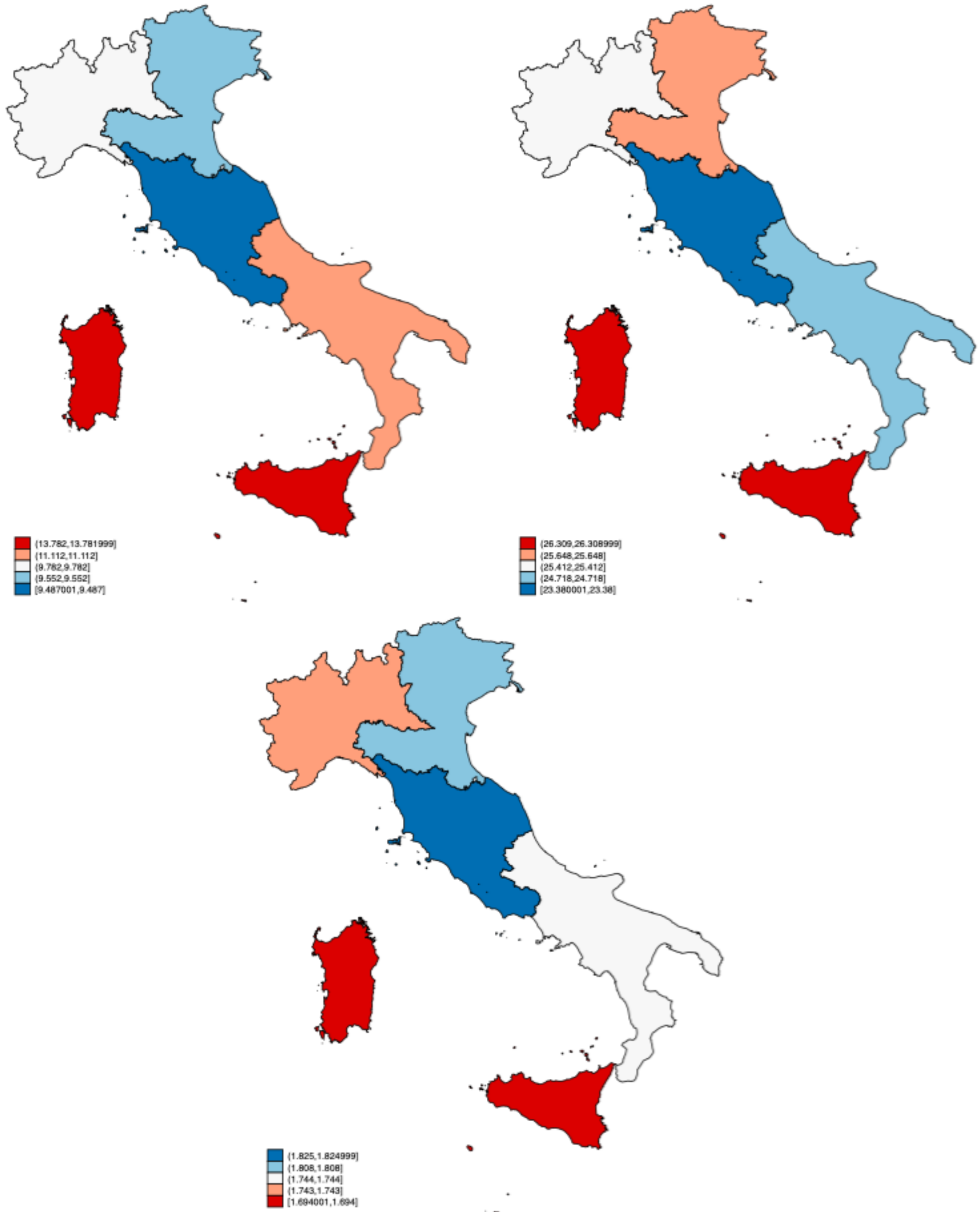


Figure 2. Distribution of 30-day mortality rate, 365-day readmission rate and Ratio of 30-day CRM for treated w/PTCA versus not treated per Area

5.2. Results of the second step. In the second step of the methodology, we use the time-varying effects obtained in the first step and also the full hospital-level data to implement the vector auto-regressive (VAR) analysis and thus assess the persistence of healthcare quality. The VAR parameters are estimated by using the information on the all three HQMs. While at hospital level we use the measures as they are provided in the dataset, at regional level we consider for each region the fixed effect obtained in the first step and its temporal lags. We don't perform the analysis at macro area level because we do not have enough information to obtain consistent estimates.

5.2.1. *Vector Auto-regression Model: hospital level.* The VAR analysis at individual level has been conducted by using the data at hospital level available in the PNE. Panel VAR analysis is predicated upon choosing the optimal lag order. Based on the model selection criteria by Andrews and Lu [2001] and the over-all coefficient of determination, second-order panel VAR is the preferred model[3], since this shows the smallest M-Bayesian Information Criterion, M-Akaike Information Criterion (Abrigo and Love [2016]).

The specification is the one in equation 4.2, although other specifications, with different lag lengths and with the exclusion of one of the three HQMs, were tested. The hospital smoothed parameter estimates are reported in Tables 7 and indicate the effect that past values of each of the three outcome measures have on their own current performance. The results suggest that the two HQMs "30-day mortality rate" and "365-day readmission rate" are quite persistent, thus seeing in particular the first lag of both having a significant impact on the same measure one year later. The coefficient of the first lag of the "30-day mortality rate" on determining the "30-day mortality rate" in year t is 0.361, whereas the coefficient of the first lag of the "365-day readmission rate" on determining the "365-day readmission rate" in year t is 0.655. Readmission rate one-year before shows higher impact on today readmission rate at hospital level. While the results suggest that the third HQM - "Ratio of 30-day CRM for treated w/PTCA versus not treated " - is not persistent. The inclusion of the third lag yielded

[3]STATA command *pvarsoc* uses the estimation sample with the highest lag order used, for all models that would be estimated by the program.

similar scores for all the three HQMs. The exclusion of the "Ratio of 30-day CRM for treated w/PTCA versus not treated " in the Vector Auto-regression of the "30-day mortality rate" strengthens the impact of the first lag of the same measure as well as the impact of the first two lags of the "365-day readmission rate". The same happens if we exclude the same HQM in the VAR of the "365-day readmission rate"[4].

5.2.2. *Vector Auto-regression Model: regional level.* To check the robustness of the persistence signal obtained in the Vector Auto-regressive model applied to hospital data, we ran the the same VAR model specification at regional level. The VAR analysis at regional level has been conducted by using the results obtained in the first step of the analysis (i.e. regional fixed effects). We chose to use a lag-length of two periods to obtain results comparable to those obtained in Section 5.2.1. Although other specifications, with different lag lengths and with the exclusion of one of the three HQMs, were tested. The regional smoothed parameter estimates are reported in Tables 8 and 9[5] and indicate the effect that past values of each of the three outcome measures have on their own performance. At regional level, the "30-day mortality rate" is slightly more persistent than at the hospital level: the value of the coefficient of the first lag is 0.462 in Table 8 and 0.884 in Table 9. The first lag of the "30-day mortality rate" is the only one which gives a significant contribution to the current value of the "30-day mortality rate". If we look at the readmission rate - second column of Table 8 - we see two significant coefficients: the second lag of the "30-day mortality rate" and the second lag of the "365-day readmission rate". However only the former has a considerable positive effect on the outcome (0.390). No significant persistence is observed for the third HQM ("Ratio of 30-day CRM for treated w/PTCA versus not treated "). Overall one can see that only the "30-day mortality rate" (in its first or second lag) can be used to predict the future values of the three HQMs.

[4]Results not reported, available upon request.

[5]Table 9 reports the results obtained running the VAR on the regional specific effects instead of the fixed effects obtained in the first step of the analysis.

6. Conclusions

The Italian national healthcare system is public, with universal coverage and financed with financial resources from general taxation. The system is jointly regulated by the state, which is responsible for defining the package of benefits and mandatory resources, and by the regions that organize health services in their area of responsibility, which are managed by healthcare companies. In the 1990s, to address the issue of financial risk competences, the State decided to transfer part of the tax leverage to the regions, thus aligning the autonomy of expenditure with that of financing.

Because of this marked regional autonomy in the management of health services, it is important to analyze the performance of Italian regions using shared reference indicators. This is possible thanks to the availability of the data collected by the Ministry of Health which converge in the annual publication of the Piano Nazionale Esiti (PNE). In this study we used a two-step methodology, proposed by Papanicolas and McGuire [2017], first to extract a quality signal from regional health systems and secondly to obtain information on the persistence of this quality signal. In their study, Papanicolas and McGuire [2017], following the approach of McClellan and Staiger [1999], used English, patient-level data for individuals suffering from heart disease (AMI) or who have undergone a HIP replacement surgery to create quality indicators at the hospital level. They suggested that their method could tackle some of the main limitations that are inherent in hospital quality measurement, allowing them to create indicators which reduce noise both within individual hospitals and across time, as well as integrate different dimensions of quality within a single estimator. In this paper we apply their method to Italian data at region level related to AMI to extract indicators of regional healthcare quality. We exploit the availability of risk-adjusted measures of quality for 283 Italian hospitals.

With this methodological exercise we prove the opportunity to extend Papanicolas and McGuire [2017] methodology to analyze the quality of the healthcare system at an aggregate level. In order to check the robustness of this intuition we applied this method also to extract a signal of quality of healthcare for three Italian macro areas: North, Center and South. Our application of this method to an extremely aggregate level raises issues, such the sample size, that one must deal with. To outline the approach in the first step we use the hospital level data to extract a regional fixed effect and a macro area fixed effect. In the second step of the methodology we relied on a VAR(2) specification to investigate the degree of persistence of hospital quality at hospital level and regional level.

Overall the estimates obtained through the two-step analysis based on three quality measures for an emergency condition, like the Acute Myocardial Infarction (AMI), seem to reasonably reflect differences in the underlying hospital quality across Italian regions. Furthermore, a multidimensional interpretation of the signal of persistence obtained in the second step of the analysis can help in predicting patients' mobility across regions and can address policy interventions to be implemented at national level to rebalance this gap.

Obviously, the results would be much more robust if the number of quality measures could be extended to indicators of efficiency and process of the hospital, without limiting the analysis to indicators of outcome. A second important extension would be to compare regional differences for emergencies and elective diseases. For the latter, data would show higher variability since patients have a greater chance of seeking treatments in hospitals of higher quality. Including these in the analysis would make it possible to obtain more reliable estimates for predicting interregional mobility.

Tables

Table 1. Summary Statistics of the Sample

Condition	Codes	Years analyzed	Mean cases per year	Number of hospitals
AMI	Italian classification code: ICD-9-CM 410.xx	2008-2016	80,479	276

Table 2. Hospital Quality Measures

Condition and year	30-day mortality rate per 100 (σ)	365-day readmission rate per 100 (σ)	Ratio of 30-day CRM for treated w/PTCA versus not treated (σ)
AMI 2008	11,64 (6,86)	26,54 (5,28)	1,74 (.31)
AMI 2009	10,51 (3,38)	25,48 (5,08)	1,79 (.34)
AMI 2010	10,86 (3,91)	25,23 (5,94)	1,82 (.32)
AMI 2011	10,62 (5,16)	24,45 (5,40)	1,84 (.34)
AMI 2012	10,06 (5,05)	23,32 (5,86)	1,86 (.37)
AMI 2013	9,12 (3,21)	22,88 (5,54)	1,83 (.36)
AMI 2014	9,11 (3,27)	-	1,89 (.34)
AMI 2015	8,97 (3,42)	21,52 (5,71)	1,90 (.42)
AMI 2016	8,28 (3,27)	-	1,95 (.44)

Table 3. Regional Characteristics (excluded "Regioni a statuto speciale")

Region	Number of hospitals	Number of districts	Mean cases per year
Abruzzo & Molise	9	5	2,486
Campania	36	5	9,209
Emilia Romagna	25	9	10,365
Lazio	28	5	8,671
Liguria	9	4	3.138
Lombardy	50	12	15,248
Marche	2	1	640
Piedmont	24	8	7.626
Puglia	26	6	5,628
Tuscany	22	10	7,952
Umbria	7	2	1.946
Veneto	38	7	7,181
Total	276	74	80,479

Table 4. Macro Area Characteristics (excluded "Regioni a statuto speciale")

Macro Area	Number of hospitals	Number of districts	Number of regions	Mean cases per year
North	146	44	5	45,500
Center	59	18	4	19,209
South	71	17	8	17,712
Total	276	74	17	80,479

Table 5. Regional Fixed Effects Results (excluded "Regioni a statuto speciale", with additional controls)

Region	30-day mortality rate per 100 (σ)	365-day readmission rate per 100 (σ)	Ratio of 30-day CRM for treated w/PTCA versus not treated (σ)
Abruzzo & Molise	12.283***	21.894***	1.638***
	(1.159)	(2.004)	(0.141)
Campania	11.217***	22.935	1.875***
	(0.625)	(0.545)	(0.058)
Emilia Romagna	8.591***	22.285	1.845***
	(0.529)	(0.874)	(0.046)
Lazio	10.49***	21.074	1.66
	(0.615)	(1.021)	(0.046)
Liguria	10.618***	22.111	1.774***
	(0.619)	(0.961)	(0.052)
Lombardy	9.723***	23.988*	1.74**
	(0.531)	(0.884)	(0.046)
Marche	8.766***	22.773	1.842***
	(0.841)	(1.218)	(0.069)
Piedmont	10.24***	18.294***	1.832***
	(0.589)	(0.905)	(0.053)
Puglia	11.867	22.057	1.873***
	(0.701)	(0.974)	(0.056)
Tuscany	7.689***	20.987	1.922***
	(0.533)	(0.883)	(0.048)
Umbria	10.761	20.202	1.966***
	(2.064)	(1.059)	(0.107)
Veneto	10.411***	20.791	1.724*
	(0.616)	(1.030)	(0.052)
Volumes	-0.001***	-0.001*	0.000***
	(0.000)	(0.000)	(0.000)
Pop over 55	-1.89	-3.85e-07	-2.87e-07***
	(1.00e-06)	(1.53e-06)	(1.02e-07)
Income	1.03	-2.43e-07	2.63e-07***
	(8.29e-07)	(1.29e-06)	(8.26e-08)
Employment rate	0.001	0.035	0.000
	(0.017)	(0.031)	(0.002)
Temperature	0.003	0.004	0.001
	(0.008)	(0.015)	(0.000)
Observations	1,918	1,918	1,439
R-squared	0.069	0.061	0.073

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6. Macro Area Fixed Effects Results (excluded "Regioni a statuto speciale", with additional controls)

Area	30-day mortality rate per 100 (σ)	365-day readmission rate per 100 (σ)	Ratio of 30-day CRM for treated w/PTCA versus not treated (σ)
North	11.990***	20.398***	1.656***
	(1.161)	(.383)	(.125)
Center	11.118*	18.859***	1.818*
	(.267)	(.356)	(.039)
South	13.677***	19.694	1.737
	(.228)	(.571)	(.022)
Volumes	-0.0007312	-.0015773	.0002657***
	(.0006548)	(.0006492)	(.0000468)
Pop over 55	-3.50e-06***	2.51e-06	-1.12e-07
	(9.19e-07)	(1.06e-06)	(8.02e-08)
Income	3.02e-06 ***	-2.34e-06	3.77e-08
	(7.60e-07)	(8.13e-07)	(6.26e-08)
Employment rate	-.0258322	.042	.0017599
	(.017)	(.024)	(.0017719)
Temperature	-.0079023	0.043	-.0003211
	(.007)	(.011)	(.0007492)

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7. Estimates of multivariate VAR(2) parameters at hospital level

Lag	30-day mortality rate per 100 (σ)	365-day readmission rate per 100 (σ)	Ratio of 30-day CRM for treated w/PTCA versus not treated (σ)
30-day mortality rate per 100 (σ) (t-1)	.361***	.141*	-.019***
	(.072)	(.085)	(.007)
30-day mortality rate per 100 (σ) (t-2)	.137**	.049	-.008*
	(.052)	(.041)	(.004)
365-day readmission rate per 100 (σ) (t-1)	.240***	.655***	-.012
	(.051)	(.064)	(.005)
365-day readmission rate per 100 (σ) (t-2)	.175***	.241***	-.009*
	(.035)	(.047)	(.003)
Ratio of 30-day CRM for treated w/PTCA versus not treated (σ) (t-1)	2.682*	1.542	-.056
	(1.259)	(1.441)	(.154)
Ratio of 30-day CRM for treated w/PTCA versus not treated (σ) (t-2)	1.505*	1.077	-.089
	(.876)	(1.076)	(.107)

Notes: No. of obs = 930; No. of panels = 167; Ave. no. of T = 5.569

Table 8. Estimates of multivariate VAR(2) parameters for regional fixed effects

Lag	30-day mortality rate per 100 (σ)	365-day readmission rate per 100 (σ)	Ratio of 30-day CRM for treated w/PTCA versus not treated (σ)
30-day mortality rate per 100 (σ) (t-1)	0.462***	-0.350	-0.021***
	(.197)	(.226)	(.016)
30-day mortality rate per 100 (σ) (t-2)	0.090	0.390***	-0.011
	(.121)	(.087)	(.011)
365-day readmission rate per 100 (σ) (t-1)	-0.305	-0.012	0.000
	(.215)	(.241)	(.016)
365-day readmission rate per 100 (σ) (t-2)	0.060*	-0.036***	0.017
	(.184)	(.216)	(.017)
Ratio of 30-day CRM for treated w/PTCA versus not treated (σ) (t-1)	-1.881	-1.676*	-0.462
	(3.674)	(3.809)	(.356)
Ratio of 30-day CRM for treated w/PTCA versus not treated (σ) (t-2)	-5.401*	-.807	-.127*
	(3.085)	(3.275)	(.239)

Notes: No. of obs = 60; No. of panels = 10; Ave. no. of T = 6.000

Table 9. Estimates of multivariate VAR(2) parameters for regional-specific effects

Region	30-day mortality rate per 100 (σ)	365-day readmission rate per 100 (σ)	Ratio of 30-day CRM for treated w/PTCA versus not treated (σ)
30-day mortality rate per 100 (σ) (t-1)	0.884***	3.649***	-0.027***
	(.083)	(.408)	(.007)
30-day mortality rate per 100 (σ) (t-2)	0.317***	2.110***	-0.004
	(.057)	(.356)	(.006)
365-day readmission rate per 100 (σ) (t-1)	-0.056***	-0.462***	0.000
	(.009)	(.052)	(.000)
365-day readmission rate per 100 (σ) (t-2)	0.018*	-0.431***	0.002***
	(.008)	(.044)	(.000)
Ratio of 30-day CRM for treated w/PTCA versus not treated (σ) (t-1)	-1.053	5.020*	-0.153
	(.956)	(5.588)	(.136)
Ratio of 30-day CRM for treated w/PTCA versus not treated (σ) (t-2)	-1.236*	-14.378***	-.126*
	(.520)	(2.555)	(.053)

Notes: No. of obs = 60; No. of panels = 10; Ave. no. of T = 6.000

Appendix A

Table 10. Regional Characteristics - Complete sample

Region	Number of hospitals	Number of districts	Mean cases per year
Abruzzo	8	4	2,640
Campania	36	5	9,209
Emilia Romagna	25	9	10,365
Friuli Venezia Giulia	1	1	119
Lazio	28	5	8,671
Liguria	9	4	3.138
Lombardy	50	12	15,248
Marche	2	1	640
Molise	1	1	146
Piedmont	24	8	7.626
Prov. Auton. Bolzano	2	1	681
Prov. Auton. Trento	2	1	852
Puglia	26	6	5,628
Sicily	1	1	79
Tuscany	22	10	7,952
Umbria	7	2	1.946
Val D'Aosta	1	1	290
Veneto	38	7	7,181
Total	283	79	82,421

Table 11. Macro Area Characteristics - Complete sample

Macro Area	Number of hospitals	Number of districts	Number of regions	Mean cases per year
North	152	44	8	45,500
Center	59	18	4	19,209
South	72	17	8	17,712
Total	283	79	20	82,421

Table 12. Regional Fixed Effects Results (excluded "Regioni a statuto speciale")

Region	30-day mortality rate per 100 (σ)	365-day readmission rate per 100 (σ)	Ratio of 30-day CRM for treated w/PTCA versus not treated (σ)
Abruzzo & Molise	11.994***	22.777***	2.640***
	(.480)	(.727)	(.0386)
Campania	9.749***	21.974***	1.477***
	(0.344)	(0.545)	(0.044)
Emilia Romagna	6.873***	21.250***	1.464***
	(0.332)	(0.564)	(0.055)
Lazio	9.216***	18.699***	1.366***
	(0.408)	(0.690)	(0.051)
Liguria	9.302***	21.404***	1.453***
	(0.392)	(0.633)	(0.056)
Lombardy	8.033***	22.847***	1.403***
	(0.264)	(0.480)	(0.040)
Marche	7.492***	21.839***	1.529***
	(0.652)	(0.825)	(0.059)
Piedmont	8.474***	19.949***	1.568***
	(0.332)	(0.558)	(0.047)
Puglia	10.123***	21.273***	1.548***
	(0.532)	(0.455)	(0.040)
Tuscany	6.160***	20.265***	1.599***
	(0.290)	(0.461)	(0.046)
Umbria	9.021***	20.723***	1.722***
	(1.589)	(0.725)	(0.079)
Veneto	8.947***	22.078***	1.542***
	(0.246)	(0.408)	(0.044)
Volumes	0.004***	0.006***	0.001***
	(0.001)	(0.001)	(0.000)
Pop over 55	-0.000	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Observations	2,447	2,447	1,756
R-squared	0.802	0.918	0.934

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 13. Macro Area Fixed Effects Results (excluded "Regioni a statuto speciale")

Region	30-day mortality rate per 100 (σ)	365-day readmission rate per 100 (σ)	Ratio of 30-day CRM for treated w/PTCA versus not treated (σ)
North	9.254***	24.064***	1.763***
	(.231)	(.383)	(.021)
Center	9.024	22.236***	1.818*
	(.267)	(24.064)	(.021)
South	10.838***	23.477*	1.737
	(.228)	(.305)	(.022)
Volumes	.0001532	-.0006758	.0002952***
	(.0005304)	(.0005744)	(.0000443)
Pop over 55	2.63e-07	5.51e-07	-6.12e-08**
	(2.28e-07)	(2.87e-07)	(2.27e-08)

CHAPTER 3

Mathematics Camps: A Gift for Gifted Students?

Joint work with Ainoa Aparicio Fenoll and Flavia Coda-Moscarola

1. Introduction

Gifted students are likely to hold leading positions in future society. However, little research is devoted to them.[1] Two important questions for policy makers should be how to treat top-performing students to allow them to better express their skills and, if it is deemed effective to promote special programs dedicated to them, how to select them. As for the former, in the United States, the National Association for Gifted Students regrets the absence of a uniform federal policy for "gifted services". This lack of regulation results in a variety of State policies which go from "Accommodations in the regular classroom" to "Full-time grouping with students of similar abilities"[2]. As for the problem of an appropriate selection, the debate over whether education of top-performing children segregates them on the basis of pre-existing privilege rather than cognitive ability is by no means recent. The debate is certainly more heated in the United States where this theme has always been linked to that of the inequality of opportunities between whites and blacks. But it is important to study the phenomenon also in Europe where the theme of equality of opportunities becomes more and more urgent as the composition of the population becomes more heterogeneous. Adopted policies may, in fact, have relevant consequences in terms of both: talented students' future outcomes and increasing or decreasing inequality between students. In this paper, we study the effects of a mathematics camp targeting exclusively gifted students for problem-solving skills, psychological traits, and career intentions.

The mathematics camp is addressed to high-school students in grades one to four (ages fourteen to eighteen). Mathematics teachers select the top two performing students in each class to attend this intensive three days of mathematics camp at the end of the academic year. During the camp, students are randomly assigned to groups and proposed mathematics problems that are unrelated to the school curricula and must be solved in collaboration with their peers and using math manipulatives. Hence, the camp is characterized by peer-to-peer learning, "inquiry-oriented" activities, and a "hands-on"

[1]Some exceptions are Griffith and Rask [2007] and Horstschräer [2012] studies of talented children's school choice and Figlio and Lucas [2004] analysis of the impact of high grading standards on high ability students.

[2]See https://www.nagc.org/ for more detailed information.

learning style.

We evaluate the impact of participation in the math camp using a randomized control trial. Teachers typically choose a small number of students per class (median: 2) according to a subjective criterion, which includes their school math grade as well as their passion for scientific subjects and abilities shown in class. We asked teachers to select one additional student per class: in each class, they selected an additional student, the first one who would have been chosen if one of the two original students was not available. In other words, teachers select those students they would have selected in the absence of our program evaluation and the first student on the waiting list. We then randomly selected two out of the three signaled students to participate in the math camp. We then compare the answers of treated and control students in a questionnaire including demographic, psychological and intentions questions, and mathematics problems. Moreover, we test whether a selection mechanism based uniquely on the school grade could be meaningful. This result could help define a homogeneous selection criterion for all teachers and which also has the advantage of being transparent.

Our paper contributes to the literature on the impact of special educational programs which target talented students on their academic performance and self-concept. This impact has been found to be positive in some contexts (Cortes et al. [2015]; Aughinbaugh [2012]) but negative in others (Clotfelter et al. [2015]). Given the findings that self-esteem boosts performance (Ferkany [2008]), we are also interested in the impact of the math camp on self-concept. The effects of the camp on self-concept may be ambiguous as talented students surrounded by other talented students may update their beliefs about their position in the ability distribution upwards or downwards. For this reason, we test empirically the impact of participation in the camp on answers to the Big Five questionnaire, which captures five aspects of personality (i.e. Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism). Moreover, we test whether the result that mathematics courses increase students' propensity to enroll in BA in Mathematics (Maestri [2013]) also applies to our setup. Finally, the heterogeneous effects found for students with different capacities (Cortes et al. [2015]) motivates us to explore whether the identification of gifted students by teachers, instead of an objective

evaluation based on the math grade, provides an efficient selection procedure.
We find that students participating in the mathematics camp improve their problem-solving skills. The improvement is higher in problems that require logic skills rather than problems that require formal mathematics knowledge (formulas, standard solving methods, etc.). This confirms the findings in previous literature that effective mathematics programs are characterized by "inquiry-oriented" instruction (Blazar [2015]), frequent teacher feedback, the use of data to guide instruction, "high-dosage" tutoring, increased instructional time, and high expectations (Dobbie and Fryer Jr [2013]). The estimated positive effect is heterogeneous: students who experience the highest improvement in problem-solving skills after participating in the program are fourteen ot eighteen years old, lower-ranked, and with high-educated parents. Regarding personality traits, the camp leads to improvements in self-concept: it reduces the incidence of neuroticism and fosters the perception of being talented. Students who particularly benefit from the math camp in terms of psychological traits are younger, with higher school grades, and with highly educated mothers. Finally, only students in the earliest grade modify their career intentions after participation in the camp: they are found to be more likely to declare that they will enroll in university and that they will enroll in a scientific major.
The remainder of this paper is organized as follows. In Section 2 we introduce the Mathematics Camp in detail. Section 3 presents the design of our randomized control trial. Section 4 is devoted to the description of our data at the individual level, for both the treatment and the control group. In section 5 we present the randomization and the empirical strategy. Section 6 establishes the main results on the Mathematics Camp treatment effects, as well as checks for its robustness exploiting heterogeneous effects. Finally section 7 concludes. The Appendixes present any supplementary material.

1.1. Related Literature. The importance of studying Maths is well documented in the existing literature. A more rigorous high-school math curriculum is associated with a higher probability of attending college and of attending a 4-year college (Aughinbaugh [2012]). The household fixed effect results imply that students who take an advanced academic math curriculum in high-school (algebra II or precalculus, trigonometry, or calculus) are about 17 % more likely to go to college and 20 % more likely to start college at a 4-year school by age 21 compared to those students whose highest math class was algebra I or geometry. Identifying the types of mathematics content knowledge that are most predictive of students' long-term learning is essential for improving both theories of mathematical development and mathematics education.

A stream of this literature focuses on the predictors of mathematical achievement: inquiry-oriented instruction positively predicts student achievement (Blazar [2015]). Content errors and imprecisions are negatively related to achievement, though these estimates are sensitive to the set of covariates included in the model, while classroom emotional support and classroom organization are not related to achievement.

Looking at inquiry-oriented education programs it is shown that an index of five-policies - namely frequent feedback, use of data to guide instruction, high dosage tutoring, increased instructional time and high expectations about academic achievement - explains about 45% of the variation in school effectiveness (Dobbie and Fryer Jr [2013]). Class size, teacher certification, teacher training are not correlated with school effectiveness. The quality of school and teachers are particularly important for gifted students (Ellison and Swanson [2016]).

An additional interesting outcome is the choice of the university program and future career. Extra-curricular activities for secondary school students in Chemistry, Physics, Math and Materials Science increase the probability of enrolling in a scientific track by 3% for males, but have no effect on females (Maestri [2013]).

Regarding the effectiveness of extra-curricular courses studies find that HSTW (High-school that work - programs, professional development, curricula and technical support available for the elementary grades through community and technical colleges) has no effect on progression in mathematics and science pipelines for the average student.

On the other hand, there is some evidence of an increased gap in course taking between more advantaged and disadvantaged students. Up to our best knowledge, there has not been any academic work assessing whether the combination of knowledge, problem-solving skills, and test preparation at the AMC (the math Olympic games in the US) is predictive of success in college and beyond.

Transition Mathematics is a curriculum that uses the University of Chicago School Mathematics Project (UCSMP) textbook. The sequence of the topics intends to assist the transition from arithmetic to algebra and geometry. Transition Mathematics aims to increase applied arithmetic, pre-algebra, and pre-geometry skills in students in grades 7-12. This 1-year curriculum also addresses general application to different wordings of problems, types of numbers, and contexts for problems and aims to promote mathematical reading skills. This program was found to have mixed effects on mathematics achievement.

Peer effects are analyzed by Fuchs et al. [1997] who reported statistically significant effects of small magnitude of Peer-Assisted Learning Strategies on mathematics achievement based on the items of the Stanford Achievement Test (SAT) aligned with Peer-Assisted Learning Strategies and no significant effects of the program on mathematics achievement based on the items of the SAT unaligned with Peer-Assisted Learning Strategies. The same finding is reached by Abdulkadiroglu et al. (2013) who found no causal effect of peer effects on students performance.

The role of the gender gap in mathematics performance is crucial in explaining the gender gap in careers and wages. Justman and Méndez [2018] find that female students require stronger prior signals of mathematical ability to choose male-dominated subjects, and when choosing these subjects earn higher average scores than males, suggesting a possible loss of efficiency.

Previous research has shown that socio-economic disadvantage adversely affects boys more than girls. Moreover, students with a language background other than English choose STEM fields with greater frequency than other students, reflecting their comparative advantage, while exhibiting more markedly gendered subject choices, indicating a role for cultural factors. Finally, there is a significantly less gender streaming in STEM subjects among female students in all-girl schools than in co-educational schools, but there isn't such difference for male students. Gender gap widens at very high performance levels (99th percentiles) (Ellison and Swanson [2010]). High achieving girls may be less likely than extreme high achieving boys to decide to compete up (Niederle and

Vesterlund [2007]). The gender gap is usually small in countries with greater gender equity (Guiso et al. [2008]).

2. Background

The Mathesis Mathematics Camp is an intensive "three days" of mathematics that involves students from more than forty high-schools in Turin (Italy). It is organized by Mathesis, the association of math high-school teachers in the Italian region of Piedmont. The summer camp has taken place yearly since 1995. The initiative is supported by the financial contribution of Compagnia di San Paolo[3], which pays half of the fee. The remaining part of the fee is paid by the students, with few exceptions in which schools provide students with scholarships.

It takes place before the end of the academic year (last days of May and early June). For the three days, participants work away from classrooms, in a residence in Bardonecchia, a mountain location near Turin. The aim of the Mathematics Camp is to enhance excellence in mathematics.

In each edition about 1500 students from the first to the fourth grades of high-schools participated, followed by 120 high-school professors, 6 professors from the Department of Mathematics of the University of Turin, 20 undergraduate students in STEM (with didactic specialization), and 8 recent graduates in STEM[4]. Due to location capacity constraints, the students are divided into 4 rounds of three days each.

During the mathematics camp, students work mostly in open spaces. The spaces are allocated one for each grade, with tables for each group and space in between so that teachers can circulate and supervise activities. In a few specific moments, they go outside where they can examine what is proposed to them through concrete experiences. Participants have the opportunity to exploit continuous interactions not only with other

[3]Compagnia di San Paolo is a foundation, based in Turin, that supports a range of charitable activities in the fields of health, education and social welfare

[4]Most of them undergraduate and graduate students in Mathematics

students and teachers from different schools but also with undergraduates and graduates in STEM who have a leading role in illustrating the most complex concepts.
The peculiarities of the camp are the learning mode and the unusual contents. Important and complicated mathematics concepts are presented to students using modern tools and applications or themes of great actuality. The working style consists of "inquiry-oriented" activities and a "hands-on" learning style: support for the reasoning is given by manipulatives for a more effective and convincing vision of the theoretical concepts. Everyday problems and mathematical games of a certain difficulty are also proposed in order to urge the students to present particular and original solutions and problem-solving strategies in a climate of playful competition.
At the end of the mathematics camp students are involved in a "Treasure hunt" on the issues addressed during the camp itself with the dual purpose of verifying the acquired competencies carried out and concluding in a fun way also enhancing with an awarding of all those (individuals or groups) who have distinguished themselves in the various stages of the mathematics camp.

3. Randomized Experiment Design

The study was conducted between November and June. To carry out the evaluation we requested to implement some changes to the ordinary organization. On a regular year class teachers decided which students to send to the camp based on a discretionary evaluation of their mathematical skills and their propensity and passion for the subject. By January 2019 teachers of the high-schools involved in the program provide the list of participants. If in a regular edition of the camp, teachers would have chosen N_i students in each class, for the evaluation teachers were asked to select $N_i + 1$ students (add the first to the waiting list): according to this request, 2124 students were selected. We decided to restrict the population to $\Sigma^C(N_i + 1)$ individuals to obtain a treatment group as similar as possible to the one selected in a regular year. We considered a priority to respect the homogeneity of the treatment group with respect to having a larger control group. In this preliminary phase, teachers indicate a preference for the paper or electronic test. In the latter case the teacher must make sure that she had the opportunity to complete the test for students in a computer room or in class on a tablet.In February 2019 the pre-camp questionnaires were administered to all the students selected in the list of potential participants in the mathematics camp. The test consists of seventy-seven questions, divided into three sections: thirty socio-demographic questions, three mathematics problems and five questions related to the mathematical reasoning behind and thirty-nine psychological questions. By March 2019, based on the lists of candidates provided by the teachers, we randomly selected N_i from $N_i + 1$ to attend the mathematics camp. We used stratified randomization by class (a process that guarantees that each class is represented in the final sample by the usual number of students). For schools that send less than 12 students, we proceed to randomization by grade (i.e. putting together all the first-grades, all the second, ...). To foster collaboration in filling the questionnaires, the first, second or third-grade students assigned to the control group were guaranteed the opportunity to participate in the internship in the next edition. To fourth-grade students a summer school at the Collegio Carlo Alberto was offered. After randomization, we randomized 1479 students to participate in the camp and 645 to be part of the control group.

Upon their arrival at the residence, students are divided into teams and during the working hours each team sits in a table. The composition was decided according to the following criteria: obtaining groups with different gender compositions and avoiding groups with students belonging to the same class. Each team was composed of 6 students (for logistical reasons some exceptions are allowed in the 5-7 student range). The students' activity during the mathematics camp is evaluated in at least five distinct stages. The number of moments of intermediate evaluation varies according to the grade. Teachers are responsible for the evaluation: they report it on a board that students can consult at any time during the three days. Teachers assign a score to the group based on whether the answer is correct but also on the originality in the execution of the work, on the speed and on any other dimension the teachers consider interesting in order to return an exhaustive overview of what happens in the classroom. The last activity consists of a “treasure hunt” lasting two hours at the end of the last day of the mathematics camp. The final score deciding the winning team is calculated as the weighted average in which the sum of the scores in the intermediate stages has weight 0.2 and the “treasure hunt” 0.8.

In a week after the end of the mathematics camp, teachers administered a post-camp questionnaire to all the students (both the ones who participated in the internship and the ones belonging to the control group). The post-camp questionnaire consists of eighty-one questions of three categories: thirty socio-demographic questions, five mathematics problems e three questions related to the mathematical reasoning behind and forty-three psychological questions.

We checked that the randomization produced homogenous groups in terms of students’ characteristics. We then estimate the impact of the camp on students’ outcomes by comparing the performance of treated and control students based on their answers to the post-camp questionnaire. The effects we look at are mainly on problem-solving skills, psychological traits, and career intentions.

4. Data and Descriptive Statistics

In the analysis we use the information collected from the two questionnaires: the pre-stage questionnaire administered in February 2019 and a post-stage questionnaire administered one week after the stage. We use the first to make sure that the groups of treated and control students are comparable ex-ante and we use the answers to the second questionnaire to measure the differences between the two groups as a consequence of the stage. In this section, we describe the main characteristics at the individual level, for both the treatment and the control groups. In terms of demographics, at the individual level, we can observe participants’ gender, year of birth, the number of siblings, average grade obtained in the first semester in Math, Italian and the average of all subjects and finally father and mother’s education level. Demographic information is collected in both the pre and post-camp questionnaires to reduce as much as

possible missing values in the dataset. After merging the answers to the pre and post-camp questionnaires and cleaning the dataset from students who were substituted after the randomization, the final sample consists of 1346 students, among which 967 in the treatment group and 379 in the control group.
Table 1 reports the demographic characteristics of all students involved in the trial. Table 2 reports the demographic characteristics of students belonging to the treatment group (i.e. participants to the camp). Table 3 reports the demographic characteristics of students belonging to the control group (i.e. those not selected to participate in the camp).

The whole sample and both the treatment and the control group are balanced in gender composition (the percentage of male is 53% in the treatment group, 54% in the control group and 54% overall). Students involved in the study are aged 14-19, and their average age is 16 years and 4 months, as indicated by the year of birth (average year of birth is 2002.64). First-grade students are the majority in the sample, around 30%: they are slightly underrepresented in the treatment group with respect to the control. The less represented grades are the third and fourth, with each covering the 23% of the population. Selected students have good grades in both Math and Italian, and therefore also their average grade is good (8 over 10). In line with the context of the mathematics camp, their performance is slightly better in Math (8.3 over 10) than in Italian (7.7 over 10). No significant differences can be observed between treated and controls. The level of education of the mother is equally distributed among the two groups: mothers who have compulsory education only are around 7.6%, mothers with a high-school diploma around 45% and mothers with a university degree around 46%. Fathers are on average less educated than mothers: those who attained only compulsory schooling are 13-14% versus 7-8% of the mothers. The control group is characterized by slightly higher fathers' education than the other group.

In Table 4 we show the answers to the Big Five questions. We show the same answers for the complete sample, the treatment group and the control group. The

Big Five model, also known as the OCEAN (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism) model, classifies personality traits in five categories (Rothmann and Coetzer [2003]). The Big Five score is obtained applying the principal component analysis to the five questions. The personality trait that better describes the selected students' attitude is open-mindedness: to the corresponding question, the average answer is 8.5. Beyond being open-minded, students seem to be, in this order, conscientious and responsible, friendly, extrovert and sociable and finally neurotic. Students in the treatment group appear friendlier, more extrovert and more conscientious and responsible than students in the control group. The latter shows a slightly higher level of neuroticism. No relevant differences in their open-mindedness.

The last Table reports the answers to the five mathematics problems administered in the post-camp questionnaire. We present the figures separately for the complete sample, treatment and control group. Among the five questions, the first three replicate the three mathematic problems of the pre-camp questionnaire, the fourth and the fifth aim to explore the problem-solving skills of the students. We show also the Math Score, whose range is 0-5, and is computed simply as the sum of the scores obtained in the five questions.

5. Randomization and Econometric Strategy

In the randomization, we use stratification to improve the efficiency of the design. We partition the population into C strata, which in our case are classes. Formally, if the covariate space is $\mathbb{X}$, we partition $\mathbb{X}$ into $\mathbb{X}_1, ..., \mathbb{X}_C$ so that $\cup_c \mathbb{X}_c = \mathbb{X}$ and $\mathbb{X}_c \cap \mathbb{X}_{c'} =$. Let C_{ic} be an indicator for unit i belonging to stratum c, so that $C_{ic} = 1_{x_i \in \mathbb{X}_c}$. Let N_c be the number of units in stratum c, i.e. number of students in class c. Then we fix the number of treated students in each class as $N_{t,c}$, so that the total number of students in the treatment group is $N_t = \Sigma_{c-1}^{C} N_{t,c}$ (Athey and Imbens [2017]).

In Table 6 descriptive statistics for the randomized groups included in the study are shown. To test whether randomization produced homogenous groups, we considered not only the socio-demographic variables but also the answers to the mathematics problems administered in the pre-camp questionnaire. P-values in the third column show that there are no significant differences between the two groups.

From Table 6 we conclude that the randomization has produced two *ex-ante* comparable groups. We estimate the impact of participating in the stage on different outcomes using the following:

$$Y_i = \beta_1 T_i + \beta_2 \mathbf{X}_i + \delta_g + \epsilon_i \tag{25}$$

where T_i is the dummy equal to 1 if student i is randomly assigned to the treatment group and 0 otherwise; $\mathbf{X}_i$ is a vector of controls at student level (gender, age in months, class dummies, ranking indicators, number of siblings).

Y_i is the outcome variable. We consider three main sets of outcomes. The first outcome considered is the math score, which is a categorical variable in the range 0-5. The second outcome is the big five score, which reflects the psychological attitude of the individual, and is a continuous variable. The third set of outcomes outlines the individual intentions to go to university, to study a STEM subject at university, or to study maths at university. All the three are categorical variables in the range 1-5.

δ_s are the group fixed effects and account for time invariant heterogeneity in the group, determined by the school, the section, the class and the school type (scientific, linguistic, etc.). Finally, ϵ_i is the error term.

6. Results

In this section we show the main results for the three outcomes: mathematics and problem-solving skills, self-concept, academic intentions for the future. Table 7 presents the main estimates of the effect of participating in the mathematics camp on the mathematics and problem-solving skills of the students, measured through their answers to the post-camp questionnaire. The outcome of Column 1 of the table is the Math Score (in column 2 standardized) which is the number of correct answers given to the five questions. Columns 3-7 display the results for each of the five questions. The first answer implies the solution of a system of equations; the second the identification of a second degree polynomial with the graph of a parabola; the third is a geometry problem with the shape of a trapezoid. The fourth and the fifth problems don't require any specific math competency to be solved and therefore identify more directly the student's problem-solving skills: we label them "Logic I" and "Logic II" (See Appendix 7 - Post-camp questionnaire - questions 7-15).

Columns 1 and 2 show that the treatment has a significant positive effect on the Math Score (0.139) and on the standardized math score (0.178)[5]. The estimated treatment effect on the five questions ranges between 0.008 and 0.046. But, as expected from our initial assumptions, the effect is significant and bigger for the two logic questions (columns 6-7). These results show that the "hands-on" learning style, which characterizes the mathematics camp, has the desired effect of helping students in developing problem-solving skills.

In Table 8 we study for which students the camp was most useful. In Column 1 we interact the treatment dummy with the grade attended by the student; in Column 2 with the male dummy; in Column 3 we explore the potential heterogeneity of the effects by school grade. In the last two columns (4-5), we explore the role of the treatment for different levels of parental education.

The coefficient of the interactions of the treatment effect with the four grade dummies varies between 0.037 for the third grade and 0.291 for the first grade. The effect is strongest for students in first and fourth grades: the youngest and the oldest students benefit more from the initiative. Students who benefit more are also those with a lower math grade in the first semester, as suggested by results in column 3. The coefficient of the interaction between treatment and math grade is in fact -.083 and significant. Other significant coefficients are found for parental education, both of the mother and of the father: students whose parents went to university seem to enjoy to a greater extent the positive effects of treatment.

Tables 9 and 10 focus on the second set of outcomes: personality traits and psychological attitude. In Table 9 we measure the effect of the treatment separately for each

[5]It is standardized by grade to have zero mean and standard deviation equal to one. We standardized by grade because different grades have different proportions of correct answers

of the five personality traits defined within the Big Five model. The results suggest that participating in the mathematics camp reduces the chance of having mood swings (i.e. being neurotic) by 3.5 percent. Being part of the treatment group does not affect students' conscientiousness and openness. The positive effect on the perception of being agreeable and extrovert is not significant.

As we did for mathematical skills, we have summarized the five personality traits in a single indicator, obtained with the Principal Components analysis, called Big Five score. The effect of the treatment on this Big Five score, reported in Column 1 of Table 10, is 0.117 but it is not significant Columns 2-6 of Table 10 show possible sources of heterogeneity in the effect of the treatment on the Big Five score. As it happened for the math score, we do not find significant differences in the effect of the treatment between males and females and between those with a high and those with a low math grade. It is interesting to note that there is a reinforcement of the positive effect in the second grade: the students who seem to benefit most of the mathematics camp from a psychological point of view, are different from those who show better improvements in the problem-solving skills. The students whose self-concept is most affected by the treatment are those who a low-educated father. We explore the heterogeneity of the effect of attending the mathematics camp on each of the five personality traits in Tables from 13 to 17 in Appendix 7.

Regarding the psychological sphere, we also asked the students to rank the importance of the following factors that mainly affect the achievement of their academic results: effort, talent, luck. As one can see in Table 11, we find a positive significant coefficient for the outcome talent, meaning that students who participate in the mathematics camp have the perception that their school performance is mostly determined by talent respect to effort and luck compared to students in the control group.

The third set of outcomes we analyze is the intention to continue studies after high-school and the program preference. Columns 1 and 2 show that attending the mathematics camp does not have any effect on the probability of going to university neither on the probability of choosing a STEM program. The math camp positively impacts intentions to go to university only for first-grade students: the coefficient shown in Column 3 is 0.11. We do not find any relevance of the gender on the choice to go to university neither on the willingness to choose STEM subjects.

7. Conclusions

A growing literature evaluates early childhood interventions and establishes that they could substantially impact later life outcomes (e.g., Knudsen et al. [2006], Borghans et al. [2008], and Currie and Almond [2011]). The existing literature mainly focuses on programs that targeted disadvantaged and low IQ subjects. Less is known about interventions on adolescents and in particular on gifted ones.

This paper uses a randomized control trial to evaluate the impact of participation in a mathematics camp targeting gifted high-school students. The camp takes place every year in May/June and involves more than 40 high-schools in Piedmont (Italy). Our research seeks to understand whether participating to an intensive "three days" camp characterized by interactive teaching and "hands-on" learning style : (i) improves mathematical knowledge and reasoning skills; (ii) influences students' self-perception; (iii) has an impact on the chances of continuing the studies by going to university and choosing a scientific subject.

To perform the randomized control trial we asked teachers to select in each class an additional student (the first student in the waiting list). We then randomly selected students to participate in the math camp and students assigned to the control group. After the camp took place, the performance of treated and control students was compared using a questionnaire administered one week after the camp.

Our findings show that students' math skills are affected by the treatment only through the effect on problem-solving abilities: this confirms the effectiveness of peer-to-peer learning, "inquiry-oriented" activities, and a "hands-on" learning style to enhance mathematical reasoning. Introducing these interactive methodologies in class may be useful to improve students' performance in math. Our results show that the treatment is particularly effective for students of the first and fourth-grade students. We also find positive spillover effects on students' personality in the short run. Students in the treatment group, after experiencing the mathematics camp, declare to be less neurotic and consider their school performance mostly determined by talent instead of luck and effort. We can consider team-working and peer-to-peer learning determinant factors of the positive effect of the treatment on students' self-perception. These short-run effects could vanish after some time because associated with the recent exposure to the treatment, or they could consolidate and students who participated in the camp would have an easier time in following standard math classes afterward. Our last relevant finding is that students of the first grade who participate in the camp are more likely to choose to continue their studies going to university after high-school. However, the choice of the subject is not affected by the treatment.

In conclusion, based on our findings, we confirm our hypothesis that a mathematics camp based on a "hands-on" learning style would enhance problem-solving abilites, mathematics knowledge and positively affect self-concept in the short run. Furthermore, according to our results. policy makers willing to invest in mathematics programs should pay special attention to student in grades I and IV.

Tables

Table 1. Demographics - Complete sample

Variable	Mean	Std. Dev.	Min.	Max.	N
Male	0.536	0.499	0	1	1346
Year of birth	2002.64	1.142	2000	2005	1299
N siblings	1.077	0.804	0	7	1346
Class==I	0.296	0.457	0	1	1346
Class==II	0.253	0.435	0	1	1346
Class==III	0.226	0.418	0	1	1346
Class==IV	0.225	0.418	0	1	1346
Math grade	8.295	0.998	5	10	1332
Italian grade	7.724	0.891	5	10	1329
Average grade	8.006	0.700	6	10	1328
Mother below high-school	0.075	0.264	0	1	1346
Mother high-school	0.444	0.497	0	1	1346
Mother university	0.464	0.499	0	1	1346
Father below high-school	0.139	0.346	0	1	1346
Father high-school	0.415	0.493	0	1	1346
Father university	0.434	0.496	0	1	1346

Table 2. Demographics - Treatment group

Variable	Mean	Std. Dev.	Min.	Max.	N
Male	0.534	0.499	0	1	967
Year of birth	2002.614	1.142	2000	2005	925
N siblings	1.071	0.783	0	5	967
Class==I	0.291	0.454	0	1	967
Class==II	0.249	0.433	0	1	967
Class==III	0.233	0.423	0	1	967
Class==IV	0.228	0.419	0	1	967
Math grade	8.291	1.01	5	10	954
talian grade	7.74	0.901	5	10	951
Average grade	8.022	0.712	6	9.6	951
Mother below high-school	0.077	0.266	0	1	967
Mother high-school	0.446	0.497	0	1	967
Mother university	0.465	0.499	0	1	967
Father below high-school	0.146	0.353	0	1	967
Father high-school	0.42	0.494	0	1	967
Father university	0.425	0.495	0	1	967

Table 3. Demographics - Control group

Variable	Mean	Std. Dev.	Min.	Max.	N
Male	0.541	0.499	0	1	379
Year of birth	2002.703	1.139	2000	2005	374
N siblings	1.09	0.856	0	7	379
Class==I	0.309	0.463	0	1	379
Class==II	0.264	0.441	0	1	379
Class==III	0.208	0.407	0	1	379
Class==IV	0.219	0.414	0	1	379
Math grade	8.305	0.97	5.5	10	378
Italian grade	7.684	0.865	6	10	378
Average grade	7.966	0.671	6	10	377
Mother below high- school	0.071	0.258	0	1	379
Mother high-school	0.441	0.497	0	1	379
Mother university	0.462	0.499	0	1	379
Father below high-school	0.121	0.327	0	1	379
Father high-school	0.404	0.491	0	1	379
Father university	0.456	0.499	0	1	379

Table 4. Big Five Statistics

Variable	Mean	Std. Dev.	Min.	Max.
Complete sample:				
Big Five Score	0.049	1.386	-6.853	2.449
Do you consider yourself friendly?	7.995	1.511	1	10
Do you consider yourself neurotic?	6.235	2.534	1	10
Do you consider yourself conscientious and responsible?	8.298	1.435	1	10
Do you consider yourself extrovert and sociable?	7.276	1.995	1	10
Do you consider yourself open-minded?	8.583	1.449	1	10
N		1346		
Treatment group:				
Big Five Score	0.097	1.371	-6.853	2.449
Do you consider yourself friendly?	8.039	1.489	1	10
Do you consider yourself neurotic?	6.139	2.503	1	10
Do you consider yourself conscientious and responsible?	8.309	1.439	1	10
Do you consider yourself extrovert and sociable?	7.371	1.94	1	10
Do you consider yourself open-minded?	8.59	1.439	1	10
N		967		
Control group:				
Big Five Score	-0.075	1.417	-6.387	2.436
Do you consider yourself friendly?	7.881	1.563	1	10
Do you consider yourself neurotic?	6.48	2.599	1	10
Do you consider yourself conscientious and responsible?	8.272	1.426	1	10
Do you consider yourself extrovert and sociable?	7.034	2.112	1	10
Do you consider yourself open-minded?	8.565	1.476	3	10
N		379		

Table 5. Math statistics

Variable	Mean	Std. Dev.	Min.	Max.
Math score	4.561	0.738	1	5
Correct system of equations	0.969	0.174	0	1
Correct parabola	0.888	0.316	0	1
Correct trapezio	0.956	0.205	0	1
Correct Logic I	0.849	0.358	0	1
Correct Logic II	0.899	0.301	0	1
N		1346		
Treatment group:				
Math score	4.612	0.688	1	5
Correct system of equations	0.975	0.156	0	1
Correct parabola	0.9	0.301	0	1
Correct trapezio	0.962	0.192	0	1
Correct Logic I	0.865	0.342	0	1
Correct Logic II	0.911	0.285	0	1
N		967		
Control group:				
Math score	4.43	0.84	1	5
Correct system of equations	0.953	0.213	0	1
Correct parabola	0.858	0.35	0	1
Correct trapezio	0.942	0.234	0	1
Correct Logic I	0.810	0.393	0	1
Correct Logic II	0.868	0.339	0	1
N		379		

Table 6. Randomization Test

Variable	Mean treated	Mean control	p-value
Male	0.534	0.541	0.809
Year of Birth	2002.614	2002.703	0.202
N siblings	1.071	1.089	0.706
Class==I	0.290	0.308	0.512
Class==II	0.249	0.263	0.579
Class==III	0.233	0.208	0.339
Class==IV	0.227	0.219	0.737
Math grade	8.214	8.259	0.445
Italian grade	7.559	7.480	0.156
Average grade	7.945	7.930	0.636
Mother below high-school	0.076	0.077	0.741
Mother high-school	0.45	0.444	0.866
Mother university	0.461	0.452	0.905
Father below high-school	0.149	0.114	0.244
Father high-school	0.433	0.42	0.589
Father university	0.409	0.447	0.296
Pre-test			
Correct system of equations	0.969	0.971	0.848
Answered through logic	0.250	0.240	0.706
Answered through system of eq.	0.640	0.665	0.392
Answered through attempts	0.054	0.057	0.892
Correct parabola	0.872	0.839	0.117
Answered through formula	0.810	0.836	0.302
Answered through attempts	0.054	0.057	0.892
Correct rectangle	0.984	0.979	0.476
Answered through formula	0.335	0.314	0.459
Attempts drawing	0.133	0.108	0.210
Attempts without drawing	0.467	0.497	0.302

Table 7. Effect of the Treatment on Problem-Solving Abilities

	Math score	Std. score	Sys. of eq.	Parabola	Trapezio	Logic I	Logic II
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	0.139	0.178	0.017	0.027	0.008	0.046	0.041
	(0.043)***	(0.059)***	(0.013)	(0.021)	(0.012)	(0.022)**	(0.019)**
Obs.	1346	1346	1346	1346	1346	1346	1346
R^2	0.035	0.034	0.01	0.022	0.013	0.027	0.016

Table 8. Heterogeneity in the Effect of the Treatment on Math Score

	Math Score				
	(1)	(2)	(3)	(4)	(5)
Treatment		0.143	0.825		
		(0.066)**	(0.368)**		
Treatment in grade I	0.291				
	(0.105)***				
Treatment in grade II	0.078				
	(0.101)				
Treatment in grade III	0.037				
	(0.065)				
Treatment in grade IV	0.101				
	(0.06)*				
Treatment male		-.009			
		(0.086)			
Treatment math grade			-.083		
			(0.043)*		
Treatment mother<HS				0.13	
				(0.173)	
Treatment mother=HS				0.1	
				(0.059)*	
Treatment mother>HS				0.208	
				(0.07)***	
Treatment father<HS					-.003
					(0.12)
Treatment father=HS					0.156
					(0.069)**
Treatment father>HS					0.197
					(0.065)***
Obs.	1346	1346	1346	1346	1346
R^2	0.04	0.035	0.038	0.039	0.039

Table 9. Effect of the Treatment on Self-Perception

	Agreeableness	Neuroticism	Conscientiousness	Extroversion	Openess
	(1)	(2)	(3)	(4)	(5)
Treatment	0.116	-.352	0.014	0.206	0.037
	(0.108)	(0.149)**	(0.084)	(0.138)	(0.104)
Obs.	1346	1346	1346	1346	1346
R^2	0.022	0.089	0.032	0.02	0.037

Table 10. Heterogeneity in the Effect of Treatment on Big Five Score

	Big Five Score					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.091		-.065	0.23		
	(0.101)		(0.154)	(1.022)		
Treatment in grade I		0.061				
		(0.147)				
Treatment in grade II		0.531				
		(0.244)**				
Treatment in grade III		-.252				
		(0.195)				
Treatment in grade IV		-.023				
		(0.208)				
Treatment male			0.303			
			(0.22)			
Treatment math grade				-.017		
				(0.124)		
Treatment mother<HS					0.003	
					(0.375)	
Treatment mother=HS					0.036	
					(0.137)	
Treatment mother>HS					0.144	
					(0.143)	
Treatment father<HS						0.577
						(0.263)**
Treatment father=HS						-.160
						(0.14)
Treatment father>HS						0.171
						(0.17)
Obs.	1346	1346	1346	1346	1346	1346
R^2	0.032	0.039	0.034	0.032	0.032	0.038

Table 11. Effect of the Treatment on Determinants of School Performance

	Effort	Talent	Luck
	(1)	(2)	(3)
Treatment	0.189	0.287	0.114
	(0.131)	(0.145)**	(0.154)
Obs.	1339	1339	1329
R^2	0.046	0.038	0.023

Table 12. Heterogeneity in the Effect of the Treatment on the Academic Intentions

	University	STEM	University	STEM	University	STEM
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.036	0.037			0.033	-.0009
	(0.029)	(0.028)			(0.036)	(0.042)
Treatment in grade I			0.11	0.018		
			(0.055)**	(0.043)		
Treatment in grade II			-.012	-.031		
			(0.063)	(0.06)		
Treatment in grade III			-.012	0.071		
			(0.054)	(0.064)		
Treatment in grade IV			0.034	0.102		
			(0.038)	(0.063)		
Treatment male					0.006	0.074
					(0.056)	(0.061)
Obs.	1344	1341	1344	1341	1344	1341
R^2	0.036	0.034	0.039	0.037	0.036	0.035

Appendix A: Post-camp questionnaire

QUESTIONNAIRE FOR STUDENTS

Instructions

Dear Student, dear Student, our survey on the Bardonecchia stage of the Mathesis Association is turning to term and we need your last help!
In the test you will find five math problems and a series of simple and quick questions you can ask yourself we ask again to answer independently and in all sincerity. The compilation interested offers you no more than 50 minutes.
The test will consist of two parts.

- The first part must be completed by all in paper format, photographed or scanned and sent to the address: NDA
- The second part, when closed, can be completed: or online in the computer lab or or in paper format, photographed or scanned and sent to the email address: NDA

If for any reason the photo or the scan were not practicable, the questionnaires must be placed in cardboard boxes of transportable format (for example those containing the reams of paper for the photocopier). A researcher / research assistant will proceed to pick them up in the days of the week immediately following those of the test. Remember that because the sending went very well click on the "send" button that you find in the last one page of the questionnaire.
Thanks a lot for the collaboration!!

BASIC INFORMATION

(1) **Name**

. . . .

(2) **Surname**

. . . .

(3) **Class**

. . . .

(4) **Section**

. . . .

(5) **Type of school**

. . . .

(6) **Name of school**

. . . .

MATHEMATICS QUESTIONS

Here are five math problems. In the time available, 35 minutes:

– put your phone aside and get a sheet,
– start each page with your first and last name, section and school,
– for each problem, specify the steps that led you to the solution.

When you're done, take a picture on the test pages and send it to mathesis2@carloalberto.org. In the subject you specify your name and surname, class, section and school.

(7) **Did you send the email with your exercises by email?**

☐ Yes
☐ No

(8) **Problem 1**

The age of the father is 15 years higher than the age of the child. Knowing that the sum of the age of the father and child is 57 years, the age of the child is:

☐ 12
☐ 13
☐ 14
☐ 16
☐ Other:

(9) **I arrived at the solution in the following way:**

☐ System of equations
☐ Using deduction
☐ Trying alternatives
☐ Other:

(10) **Problem 2**

In one exam a student found the following equation for a parabola:

$$y - 6 = \frac{1}{2}(x-2)^2 \tag{26}$$

For the same parabola another student finds:

$$y = \frac{1}{2}x^2 - 2x + 8 \tag{27}$$

They can both be right?

- □ Yes
- □ No
- □ I don't know

(11) **I arrived at the solution in the following way:**

- □ Using the formula
- □ Trying alternatives of x and y
- □ Drawing the parabola
- □ Other:

(12) **Problem 3**

Consider an isosceles trapezoid with a base of more than 7 cm, a minor base 3 cm and a height of 4 cm. What can be the size of a rectangle with exactly its area?

- □ I didn't find a solution
- □ I found the right solution (area 20 cm^2)
- □ I found the wrong solution (area different from 20 cm^2)
- □ Other:

(13) **I arrived at the solution in the following way:**

- □ One example by drawing

- ☐ More than one example by drawing
- ☐ One example without drawing
- ☐ More than one example without drawing
- ☐ Writing the formula for the generic case
- ☐ Other:

(14) **Problem 4**

In a cage there were five parakeets. Their average price was 60 euros. One day the most beautiful flies away. The average price of the remaining is 50 euros. What was the price of the one who ran away?

. . . .

(15) **Problem 5**

In the same month, three Sundays fell on even days. What day of the week was the 20th of that month?

. . . .

SOCIO-DEMOGRAPHIC QUESTIONS

Thank you for your answers! Now we ask you to answer to a few socio-demograohic and attitudinal questions. Remember to answer 0 if you don't agree at all with the statement of the question and the maximum value vice-versa.

(16) **After high school are you going to enroll at university?**

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5

(17) **If yes, are you are going to enroll in a STEM program?**

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5

(18) **If yes, are you are going to enroll in a Math program?**

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5

(19) **Which is your Math grade in the current semester?**

. . . .

(20) **Which is your italian grade in the current semester?**

. . . .

(21) **Which is your average grade in the current semester?**

. . . .

Your grades are the result of:

Indicate on a 0-10 scale the weight of each component on obtaining your grades. Remember that the sum must be 10.

(22) **Effort**

. . . .

(23) **Talent**

. . . .

(24) **Luck**

. . . .

(25) **If you are around 100 people your age, how many do you usually consider more intelligent than you?**

. . . .

(26) **If you are around 100 people your age, how many do you usually consider better than you in math?**

. . . .

(27) **If you can't solve a mathematical problem that the professor gave you as a task (more than one answer possible):**

- ☐ Forget it and the next day you tell the professor that he couldn't solve it
- ☐ Ask an adult for help
- ☐ Ask a mate for help
- ☐ Look on books
- ☐ Other:

(28) **If you find it difficult to solve a mathematical problem, for how many minutes you try before adopting one of the solutions mentioned earlier?**

. . . .

Now we ask you to instinctively report how much you agree with the following statements (from 1 to 5):

(29) **I get discouraged easily**

. . . .

(30) **I don't speak in the presence of strangers**

. . . .

(31) **I easily make new friends**

. . . .

(32) **I find alternative solutions to problems**

. . . .

(33) **I flike to work with other people**

. . . .

(34) **IWhen I work with other people I am able to listen to their ideas**

. . . .

(35) **I know when it's time to talk about my personal problems to others**

. . . .

(36) **When I encounter obstacles, I remember the situations in which I encountered obstacles similar and I managed to overcome them**

. . . .

(37) **I expect to do well in most of the things I do**

. . . .

(38) **People easily confide in me**

. . . .

(39) **I struggle to understand the non-verbal messages of others**

. . . .

(40) **Some important episodes of my life have led me to rethink what is important and what is not**

. . . .

(41) **When my mood changes, I see new possibilities**

. . . .

(42) **Emotions are among the things that make life worth living**

. . . .

(43) **I am aware of my emotions**

. . . .

(44) **I have an optimistic attitude**

. . . .

(45) **I like to share my emotions with other people**

. . . .

(46) **When I experience a positive emotion, I know how to make it last**

. . . .

(47) **I organize events that others like**

. . . .

(48) **I try to do things that make me happy**

. . . .

(49) **I am aware of the non-verbal messages I send to others**

. . . .

(50) **I introduce myself so as to offer a good impression to others**

. . . .

(51) **When I'm in a good mood, it's easy to solve problems**

. . . .

(52) **Based on the facial expression, I can recognize the emotions felt by others**

. . . .

(53) **I know why my emotions change**

. . . .

(54) **When I am happy I am more like to have good ideas**

. . . .

(55) **I am able to control my emotions**

. . . .

(56) **I am able to easily recognize my emotions**

. . . .

(57) **I find motivation by imagining positive results for the tasks I face**

. . . .

(58) **I congratulate others when they do something well**

. . . .

(59) **I am aware of the non-verbal messages that people send**

. . . .

(60) **When someone tells me about an important event in his life, it almost seems to me to have it personally lived**

. . . .

(61) **When I feel an emotional change, I tend to produce new ideas**

. . . .

(62) **When I face a challenge, I give up thinking I won't make it**

. . . .

(63) **I just need to look at people to see how they feel**

. . . .

(64) **I help people feel better when they are feeling a little down**

. . . .

(65) **I use good humor to face obstacles**

. . . .

(66) **I understand how people feel by listening to their tone of voice**

. . . .

(67) **I find it hard to understand why people feel a certain way**

. . . .

BIG FIVE QUESTIONS (scale 0-10)

(68) **Do you consider yourself an extrovert and sociable person?**

. . . .

(69) **Do you consider yourself a friendly person?**

. . . .

(70) **Do you consider yourself a sociable and conscientious person?**

. . . .

(71) **Do you consider yourself a neurotic person?**

. . . .

(72) **Do you consider yourself an open-mind person?**

. . . .

Let us briefly make the point: to avoid losing pieces and not being able to adequately exploit all the information there you have provided. As part of the evaluation of the Mathematics Stage, this should be the second questionnaire you fill out. Think about it ...

(73) **Did you fulfill the Pre-stage questionnaire?**

- □ Yes
- □ No
- □ I son't remember

If you answered no or I don't know, the test continues with some quick ones socio-demographic and attitudinal questions.

(74) **Write your date of birth**

. . . .

(75) **Are you male or female?**

- □ M
- □ F
- □ I don't want to answer

(76) **Which is the Postal Code (CAP) of your home?**

. . . .

(77) **Indicate your mother's education**

- □ Graduate or Post-graduate
- □ High-School
- □ Compulsory school
- □ Nothing

(78) **Indicate your father's education**

- □ Graduate or Post-graduate
- □ High-School
- □ Compulsory school
- □ Nothing

(79) **How many brothers/sisters do you have?**

. . . .

(80) **Are your brothers/sisters younger or older than you?**

Brother/sister 1

- □ Older
- □ Younger

Brother/sister 2

- ☐ Older
- ☐ Younger

Brother/sister 3

- ☐ Older
- ☐ Younger

Brother/sister 4

- ☐ Older
- ☐ Younger

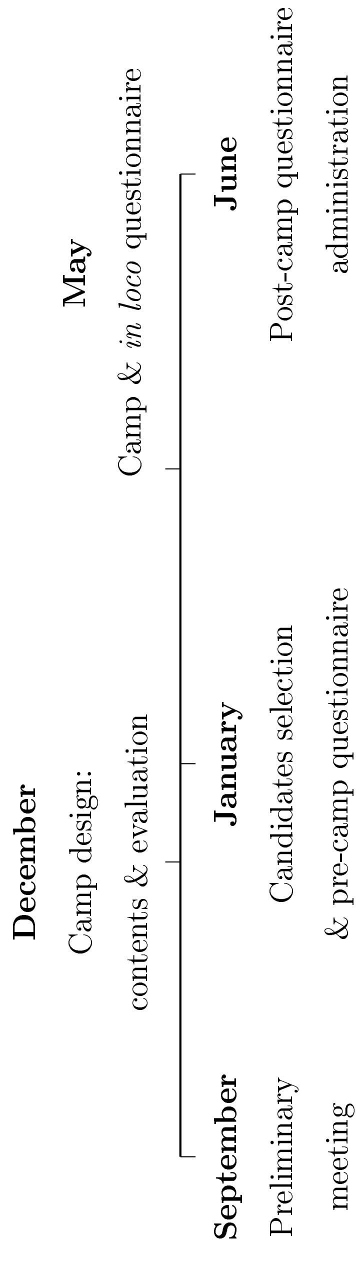
(81) **Write name and surname of your math teacher**

. . . .

Information for scientific research (articles 1314 of the EU Reg 2016/679)

The test is finished, thanks for your help! Remember to click "submit / submit" before closing the page

Appendix B: Timing of the Study



Appendix C: Heterogeneity in Personality Traits

Table 13. Heterogeneity in the Effect of the Treatment on Agreeableness

	Agreeableness				
	(1)	(2)	(3)	(4)	(5)
Treatment		-.065 (0.175)	0.409 (0.992)		
Treatment in grade I	0.317 (0.216)				
Treatment in grade II	0.524 (0.223)**				
Treatment in grade III	-.360 (0.187)*				
Treatment in grade IV	-.141 (0.23)				
Treatment male		0.353 (0.236)			
Treatment math grade			-.035 (0.12)		
Treatment mother<HS				-.538 (0.376)	
Treatment mother=HS				0.069 (0.149)	
Treatment mother>HS				0.243 (0.152)	
Treatment father<HS					0.009 (0.311)
Treatment father=HS					-.139 (0.144)
Treatment father>HS					0.362 (0.158)**
Obs.	1346	1346	1346	1346	1346
R^2	0.034	0.025	0.022	0.026	0.027

Table 14. Heterogeneity in the Effect of the Treatment on Neuroticism

	Neuroticism				
	(1)	(2)	(3)	(4)	(5)
Treatment		-.700 (0.217)***	0.029 (1.502)		
Treatment in grade I	-.278 (0.27)				
Treatment in grade II	-.724 (0.28)***				
Treatment in grade III	-.115 (0.387)				
Treatment in grade IV	-.271 (0.249)				
Treatment male		0.677 (0.321)**			
Treatment math grade			-.046 (0.178)		
Treatment mother<HS				-.067 (0.529)	
Treatment mother=HS				-.487 (0.243)**	
Treatment mother>HS				-.209 (0.23)	
Treatment father<HS					-.799 (0.451)*
Treatment father=HS					-.213 (0.245)
Treatment father>HS					-.322 (0.261)
Obs.	1346	1346	1346	1346	1346
R^2	0.09	0.092	0.089	0.089	0.09

Table 15. Heterogeneity in the Effect of the Treatment on Extroversion

	Extroversion				
	(1)	(2)	(3)	(4)	(5)
Treatment		-.054 (0.22)	-1.298 (1.283)		
Treatment in grade I	0.589 (0.236)**				
Treatment in grade II	0.469 (0.27)*				
Treatment in grade III	-.332 (0.292)				
Treatment in grade IV	-.074 (0.27)				
Treatment male		0.506 (0.309)			
Treatment math grade			0.182 (0.158)		
Treatment mother<HS				-.902 (0.459)**	
Treatment mother=HS				0.21 (0.192)	
Treatment mother>HS				0.356 (0.191)*	
Treatment father<HS					-.077 (0.423)
Treatment father=HS					0.0001 (0.193)
Treatment father>HS					0.486 (0.176)***
Obs.	1346	1346	1346	1346	1346
R^2	0.028	0.023	0.021	0.026	0.023

Table 16. Heterogeneity in the Effect of the Treatment on Openness

	Openness				
	(1)	(2)	(3)	(4)	(5)
Treatment		-.004 (0.139)	0.841 (0.862)		
Treatment in grade I	0.311 (0.178)*				
Treatment in grade II	0.122 (0.197)				
Treatment in grade III	-.376 (0.203)*				
Treatment in grade IV	-.026 (0.183)				
Treatment male		0.08 (0.197)			
Treatment math grade			-.097 (0.104)		
Treatment mother<HS				0.184 (0.286)	
Treatment mother=HS				-.022 (0.159)	
Treatment mother>HS				0.05 (0.132)	
Treatment father<HS					0.026 (0.257)
Treatment father=HS					-.100 (0.159)
Treatment father>HS					0.129 (0.137)
Obs.	1346	1346	1346	1346	1346
R^2	0.044	0.038	0.038	0.038	0.039

Table 17. Heterogeneity in the Effect of the Treatment on Conscientiousness

	Conscientiousness				
	(1)	(2)	(3)	(4)	(5)
Treatment		-.060 (0.128)	0.219 (0.925)		
Treatment in grade I	-.298 (0.142)**				
Treatment in grade II	0.382 (0.209)*				
Treatment in grade III	-.010 (0.155)				
Treatment in grade IV	0.049 (0.154)				
Treatment male		0.145 (0.185)			
Treatment math grade			-.025 (0.113)		
Treatment mother<HS				0.394 (0.337)	
Treatment mother=HS				-.029 (0.119)	
Treatment mother>HS				-.003 (0.123)	
Treatment father<HS					0.83 (0.27)***
Treatment father=HS					-.143 (0.136)
Treatment father>HS					-.073 (0.16)
Obs.	1346	1346	1346	1346	1346
R^2	0.039	0.033	0.032	0.034	0.044

Bibliography

A. Abdulkadiroglu. School choice. *The handbook of market design*, pages 138–169, 2013.

M. R. Abrigo and I. Love. Estimation of panel vector autoregression in stata. *The Stata Journal*, 16(3):778–804, 2016.

M. Ali, R. Salehnejad, and M. Mansur. Hospital heterogeneity: what drives the quality of health care. *The European Journal of Health Economics*, 19(3):385–408, 2018.

S. Andersen, G. W. Harrison, M. I. Lau, and E. E. Rutström. Elicitation using multiple price list formats. *Experimental Economics*, 9(4):383–405, 2006.

S. Andersen, G. W. Harrison, M. I. Lau, and E. E. Rutström. Eliciting risk and time preferences. *Econometrica*, 76(3):583–618, 2008.

L. R. Anderson and J. M. Mellor. Predicting health behaviors with an experimental measure of risk preference. *Journal of health economics*, 27(5):1260–1274, 2008.

P. M. Anderson, K. F. Butcher, and P. B. Levine. Maternal employment and overweight children. *Journal of health economics*, 22(3):477–504, 2003.

J. Andreoni and C. Sprenger. Estimating time preferences from convex budgets. *American Economic Review*, 102(7):3333–56, 2012.

D. W. Andrews and B. Lu. Consistent model and moment selection procedures for gmm estimation with application to dynamic panel data models. *Journal of Econometrics*, 101(1):123–164, 2001.

F. J. Anscombe and R. J. Aumann. A definition of subjective probability. *The annals of mathematical statistics*, 34(1):199–205, 1963.

S. Athey and G. W. Imbens. The econometrics of randomized experiments. In *Handbook of Economic Field Experiments*, volume 1, pages 73–140. Elsevier, 2017.

A. Aughinbaugh. The effects of high school math curriculum on college attendance: Evidence from the nlsy97. *Economics of Education Review*, 31(6):861–870, 2012.

M. Awais, M. F. Laber, N. Rasheed, and A. Khursheed. Impact of financial literacy and investment experience on risk tolerance and investment decisions: Empirical evidence from pakistan. *International Journal of Economics and Financial Issues*, 6(1):73–79, 2016.

K. Baicker, A. Chandra, J. S. Skinner, and J. E. Wennberg. Who you are and where you live: How race and geography affect the treatment of medicare beneficiaries: There is no simple story that explains the regional patterns of racial disparities in health care. *Health affairs*, 23(Suppl2):VAR–33, 2004.

A. E. Barnato, F. L. Lucas, D. Staiger, D. E. Wennberg, and A. Chandra. Hospital-level racial disparities in acute myocardial infarction treatment and outcomes. *Medical care*, 43(4):308, 2005.

R. B. Barsky, F. T. Juster, M. S. Kimball, and M. D. Shapiro. Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics*, 112(2):537–579, 1997.

M. Barton. Childhood obesity: a life-long health risk. *Acta Pharmacologica Sinica*, 33 (2):189, 2012.

G. Becker. An economic approach to human behavior. 4, 09 1978.

G. M. Becker, M. H. DeGroot, and J. Marschak. Measuring utility by a single-response sequential method. *Systems Research and Behavioral Science*, 9(3):226–232, 1964.

G. S. Becker. A theory of social interactions. *Journal of political economy*, 82(6): 1063–1093, 1974.

G. S. Becker and C. B. Mulligan. The endogenous determination of time preference. *The Quarterly Journal of Economics*, 112(3):729–758, 1997.

G. S. Becker and K. M. Murphy. A theory of rational addiction. *Journal of political Economy*, 96(4):675–700, 1988.

J. Benbassat and M. Taragin. Hospital readmissions as a measure of quality of health care: advantages and limitations. *Archives of internal medicine*, 160(8):1074–1081, 2000.

H. P. Binswanger. Attitudes toward risk: Experimental measurement in rural india. *American journal of agricultural economics*, 62(3):395–407, 1980.

D. Blazar. Effective teaching in elementary mathematics: Identifying classroom practices that support student achievement. *Economics of Education Review*, 48:16–29, 2015.

L. Borghans, A. L. Duckworth, J. J. Heckman, and B. Ter Weel. The economics and psychology of personality traits. *Journal of human Resources*, 43(4):972–1059, 2008.

H. Brown and M. van der Pol. Intergenerational transfer of time and risk preferences. *Journal of Economic Psychology*, 49:187–204, 2015.

J. Cawley and F. Liu. Maternal employment and childhood obesity: A search for mechanisms in time use data. *Economics & Human Biology*, 10(4):352–364, 2012.

C. F. Chabris, D. Laibson, C. L. Morris, J. P. Schuldt, and D. Taubinsky. Individual laboratory-measured discount rates predict field behavior. *Journal of risk and uncertainty*, 37(2-3):237, 2008.

F. Chaloupka. Rational addictive behavior and cigarette smoking. *Journal of political Economy*, 99(4):722–742, 1991.

A. Chandra, J. Gruber, and R. McKnight. Patient cost-sharing and hospitalization offsets in the elderly. *American Economic Review*, 100(1):193–213, 2010.

K. K. Charles and E. Hurst. The correlation of wealth across generations. *Journal of political Economy*, 111(6):1155–1182, 2003.

L. M. Chen, A. K. Jha, S. Guterman, A. B. Ridgway, E. J. Orav, and A. M. Epstein. Hospital cost of care, quality of care, and readmission rates: penny wise and pound foolish? *Archives of internal medicine*, 170(4):340–346, 2010.

S.-Y. Chou, M. E. Deily, S. Li, and Y. Lu. Competition and the impact of online hospital report cards. *Journal of Health Economics*, 34:42–58, 2014.

C. T. Clotfelter, H. F. Ladd, and J. L. Vigdor. The aftermath of accelerating algebra evidence from district policy initiatives. *Journal of Human Resources*, 50(1):159–188, 2015.

D. Coate. The relationship between diet, parent" s fatness, and obesity in children and adolescents, 1983.

A. K. Cohen, M. Rai, D. H. Rehkopf, and B. Abrams. Educational attainment and obesity: a systematic review. *Obesity Reviews*, 14(12):989–1005, 2013.

K. E. Cortes, J. S. Goodman, and T. Nomi. Intensive math instruction and educational attainment long-run impacts of double-dose algebra. *Journal of Human Resources*, 50(1):108–158, 2015.

J. Costa-Font and J. Gil. Intergenerational and socioeconomic gradients of child obesity. *Social science & medicine*, 93:29–37, 2013.

J. Currie and D. Almond. Human capital development before age five. In *Handbook of labor economics*, volume 4, pages 1315–1486. Elsevier, 2011.

A. C. de Oliveira, T. C. Leonard, K. Shuval, C. S. Skinner, C. Eckel, and J. C. Murdoch. Economic preferences and obesity among a low-income african american community. *Journal of economic behavior & organization*, 131:196–208, 2016.

J. B. Dimick, H. G. Welch, and J. D. Birkmeyer. Surgical mortality as an indicator of hospital quality: the problem with small sample size. *Jama*, 292(7):847–851, 2004.

N. Dirindin and P. Eva. *Governare il federalismo. Le sfide per la sanità*. Il Pensiero Scientifico, 2001.

W. Dobbie and R. G. Fryer Jr. Getting beneath the veil of effective schools: Evidence from new york city. *American Economic Journal: Applied Economics*, 5(4):28–60, 2013.

M. Doepke and F. Zilibotti. Culture, entrepreneurship, and growth. In *Handbook of economic growth*, volume 2, pages 1–48. Elsevier, 2014.

M. Doepke and F. Zilibotti. Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica*, 85(5):1331–1371, 2017.

T. Dohmen, A. Falk, D. Huffman, and U. Sunde. Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(3):1238–60, 2010.

T. Dohmen, A. Falk, D. Huffman, and U. Sunde. The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies*, 79(2):645–677, 2011a.

T. Dohmen, A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550, 2011b.

A. C. Drichoutis and J. L. Lusk. What can multiple price lists really tell us about risk preferences? *Journal of Risk and Uncertainty*, 53(2-3):89–106, 2016.

C. C. Eckel and P. J. Grossman. Are women less selfish than men?: Evidence from dictator experiments. *The economic journal*, 108(448):726–735, 1998.

C. C. Eckel and P. J. Grossman. Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization*, 68(1):1–17, 2008.

A. Edwards and G. Elwyn. *Shared decision-making in health care: Achieving evidence-based patient choice*. Oxford University Press, 2009.

G. Ellison and A. Swanson. The gender gap in secondary school mathematics at high achievement levels: Evidence from the american mathematics competitions. *Journal of Economic Perspectives*, 24(2):109–28, 2010.

G. Ellison and A. Swanson. Do schools matter for high math achievement? evidence from the american mathematics competitions. *American Economic Review*, 106(6):1244–77, 2016.

J. Engle-Warnick, J. Escobal, and S. Laszlo. How do additional alternatives affect individual choice under uncertainty? *Canadian Journal of Economics/Revue canadienne d'économique*, 42(1):113–140, 2009.

L. G. Epstein. A definition of uncertainty aversion. *The Review of Economic Studies*, 66(3):579–608, 1999.

M. Ezzati, A. D. Lopez, A. A. Rodgers, and C. J. Murray. *Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors*. World Health Organization, 2004.

M. Ferkany. The educational importance of self-esteem. *Journal of Philosophy of Education*, 42(1):119–132, 2008.

D. N. Figlio and M. E. Lucas. Do high grading standards affect student performance? *Journal of Public Economics*, 88(9-10):1815–1834, 2004.

C. Fischer, H. A. Anema, and N. S. Klazinga. The validity of indicators for assessing quality of care: a review of the european literature on hospital readmission rate. *The European Journal of Public Health*, 22(4):484–491, 2011.

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

M. Friedman and L. J. Savage. The utility analysis of choices involving risk. *Journal of political Economy*, 56(4):279–304, 1948.

D. Fuchs, L. S. Fuchs, P. G. Mathes, and D. C. Simmons. Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34(1):174–206, 1997.

V. R. Fuchs. Time preference and health: an exploratory study, 1980.

S. Gable and S. Lutz. Household, parent, and child contributions to childhood obesity. *Family relations*, 49(3):293–300, 2000.

M. Gaynor, C. Propper, and S. Seiler. Free to choose? reform, choice, and consideration sets in the english national health service. *American Economic Review*, 106(11):3521–57, 2016.

I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of mathematical economics*, 18(2):141–153, 1989.

G. Gowrisankaran and R. J. Town. Estimating the quality of care in hospitals using instrumental variables. *Journal of health economics*, 18(6):747–767, 1999.

H. Gravelle and P. Sivey. Imperfect information in a quality-competitive hospital market. *Journal of Health Economics*, 29(4):524–535, 2010.

A. Griffith and K. Rask. The influence of the us news and world report collegiate rankings on the matriculation decision of high-ability students: 1995–2004. *Economics of Education Review*, 26(2):244–255, 2007.

M. Grossman. On the concept of health capital and the demand for health. *Journal of Political economy*, 80(2):223–255, 1972.

L. Guiso, F. Monte, P. Sapienza, and L. Zingales. Culture, gender, and math. *Science*, 320(5880):1164–1165, 2008.

N. Gutacker, L. Siciliani, G. Moscelli, and H. Gravelle. Choice of hospital: Which type of quality matters? *Journal of health economics*, 50:230–246, 2016.

C. A. Holt and S. K. Laury. Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655, 2002.

J. Horstschräer. University rankings in action? the importance of rankings and an excellence competition for university choice of high-ability students. *Economics of*

Education Review, 31(6):1162–1176, 2012.

H. S. Houthakker and L. D. Taylor. Consumer demand in the united states. 1970.

D. Hryshko, M. J. Luengo-Prado, and B. E. Sørensen. Childhood determinants of risk aversion: The long shadow of compulsory education. *Quantitative Economics*, 2(1): 37–72, 2011.

Y. Jo. What money can buy: family income and childhood obesity. *Economics & Human Biology*, 15:1–12, 2014.

H. E. Jones and D. J. Spiegelhalter. Improved probabilistic prediction of healthcare performance indicators using bidirectional smoothing models. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(3):729–747, 2012.

M. Justman and S. J. Méndez. Gendered choices of stem subjects for matriculation are not driven by prior differences in mathematical achievement. *Economics of Education Review*, 64:282–297, 2018.

K. Kesteloot and N. Voet. Incentives for cooperation in quality improvement among hospitals?the impact of the reimbursement system. *Journal of health economics*, 17 (6):701–728, 1998.

J. Knowles and A. Postlewaite. Do children learn to save from their parents (parc working paper series# 05–06). *Philadelphia, PA: University of Pennsylvania*, 2005.

E. I. Knudsen, J. J. Heckman, J. L. Cameron, and J. P. Shonkoff. Economic, neurobiological, and behavioral perspectives on building america's future workforce. *Proceedings of the National Academy of Sciences*, 103(27):10155–10162, 2006.

F. Kosse and F. Pfeiffer. Impatience among preschool children and their mothers. *Economics Letters*, 115(3):493–495, 2012.

D. Laibson. Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2):443–478, 1997.

M. Laudicella, P. L. Donni, and P. C. Smith. Hospital readmission rates: signal of failure or success? *Journal of health economics*, 32(5):909–921, 2013.

D. M. Levy. The fragile politics of addiction. *Public Choice*, 81(3-4):263–275, 1994.

G. Loewenstein and D. Prelec. Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, 107(2):573–597, 1992.

V. Maestri. Promoting scientific faculties: Does it work? evidence from italy. *Economics of Education Review*, 32:168–180, 2013.

M. H. Mahmoud and A. S. Taha. The association between eating habits and body mass index among nursing students.

M. McClellan and D. Staiger. The quality of health care providers. Technical report, National Bureau of Economic Research, 1999.

E. A. McGlynn, S. M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E. A. Kerr. The quality of health care delivered to adults in the united states. *New England journal of medicine*, 348(26):2635–2645, 2003.

B. McKinstry. Paternalism and the doctor-patient relationship in general practice. *Br J Gen Pract*, 42(361):340–342, 1992.

L. Miller, D. E. Meyer, and J. T. Lanzetta. Choice among equal expected value alternatives: Sequential effects of winning probability level on risk preferences. *Journal of Experimental Psychology*, 79(3p1):419, 1969.

P. Morosini. Indicatori in valutazione e miglioramento della qualità professionale: Rapporti istisan. 2004.

J. K. Murnighan, A. E. Roth, and F. Schoumaker. Risk aversion in bargaining: An experimental study. *Journal of Risk and Uncertainty*, 1(1):101–124, 1988.

M. Niederle and L. Vesterlund. Do women shy away from competition? do men compete too much? *The quarterly journal of economics*, 122(3):1067–1101, 2007.

T. O'Donoghue and M. Rabin. Doing it now or later. *American Economic Review*, 89 (1):103–124, 1999.

T. O'Donoghue and M. Rabin. The economics of immediate gratification. *Journal of Behavioral Decision Making*, 13(2):233, 2000.

A. Orphanides and D. Zervos. Rational addiction with learning and regret. *Journal of Political Economy*, 103(4):739–758, 1995.

I. Papanicolas and A. McGuire. Using a vector autoregression framework to measure the quality of english nhs hospitals. 2011.

I. Papanicolas and A. McGuire. Measuring and forecasting quality in english hospitals. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180(2):409–432, 2017.

D. W. Pitches, M. A. Mohammed, and R. J. Lilford. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? a systematic review of the literature. *BMC health services research*, 7(1):91, 2007.

C. Propper, S. Burgess, and K. Green. Does competition between hospitals improve the quality of care?: Hospital death rates and the nhs internal market. *Journal of Public Economics*, 88(7-8):1247–1272, 2004.

C. Propper, S. Burgess, and D. Gossage. Competition and quality: evidence from the nhs internal market 1991–9. *The Economic Journal*, 118(525):138–170, 2007.

F. P. Ramsey. A mathematical theory of saving. *The economic journal*, 38(152):543–559, 1928.

B. Reynolds, K. Leraas, C. Collins, and S. Melanko. Delay discounting by the children of smokers and nonsmokers. *Drug & Alcohol Dependence*, 99(1):350–353, 2009.

A. E. Roth, J. K. Murnighan, and F. Schoumaker. The deadline effect in bargaining: Some experimental evidence. *The American Economic Review*, 78(4):806–823, 1988.

S. Rothmann and E. P. Coetzer. The big five personality dimensions and job performance. *SA Journal of Industrial Psychology*, 29(1):68–74, 2003.

P. A. Samuelson. A note on measurement of utility. *The review of economic studies*, 4 (2):155–161, 1937.

L. J. Savage. *The foundations of statistics*. Courier Corporation, 1972.

R. Schubert, M. Brown, M. Gysler, and H. W. Brachinger. Financial decision-making: are women really more risk-averse? *American Economic Review*, 89(2):381–385, 1999.

Y.-C. Shen. The effect of financial pressure on the quality of care in hospitals. *Journal of health economics*, 22(2):243–269, 2003.

V. Shrewsbury and J. Wardle. Socioeconomic status and adiposity in childhood: a systematic review of cross-sectional studies 1990–2005. *Obesity*, 16(2):275–284, 2008.

B. M. Silveira, N. Kliemann, D. P. Silva, C. F. Colussi, and R. P. d. C. Proença. Availability and price of food products with and without trans fatty acids in food stores around elementary schools in low-and medium-income neighborhoods. *Ecology of food and nutrition*, 52(1):63–75, 2013.

A. R. Singh and S. A. Singh. Diseases of poverty and lifestyle, well-being and human development. *Mens sana monographs*, 6(1):187, 2008.

H. Smith, C. Currie, P. Chaiwuttisak, and A. Kyprianou. Patient choice modelling: how do patients choose their hospitals? *Health care management science*, 21(2): 259–268, 2018.

R. H. Strotz. Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3):165–180, 1955. doi: 10.2307/2295722. URL `http://dx.doi.org/10.2307/2295722`.

M. Sutter, M. G. Kocher, D. Glätzle-Rützler, and S. T. Trautmann. Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *American Economic Review*, 103(1):510–31, 2013.

V. Svensson, J. A. Jacobsson, R. Fredriksson, P. Danielsson, T. Sobko, H. B. Schiöth, and C. Marcus. Associations between severity of obesity in childhood and adolescence, obesity onset and parental bmi: a longitudinal cohort study. *International journal of obesity*, 35(1):46, 2011.

P. A. Taheri, D. A. Butz, and L. J. Greenfield. Length of stay has minimal impact on the cost of hospital admission. *Journal of the American College of Surgeons*, 191(2): 123–130, 2000.

O. Tiemann, J. Schreyögg, and R. Busse. Hospital ownership and efficiency: a review of studies with particular focus on germany. *Health Policy*, 104(2):163–171, 2012.

. Tunçalp, W. Were, C. MacLennan, O. Oladapo, A. Gülmezoglu, R. Bahl, B. Daelmans, M. Mathai, L. Say, F. Kristensen, et al. Quality of care for pregnant women and newborns?the who vision. *BJOG: an international journal of obstetrics & gynaecology*, 122(8):1045–1049, 2015.

M. Van der Pol. Health, education and time preference. *Health economics*, 20(8): 917–929, 2011.

E. Van Doorslaer, A. Wagstaff, H. Van der Burg, T. Christiansen, D. De Graeve, I. Duchesne, U.-G. Gerdtham, M. Gerfin, J. Geurts, L. Gross, et al. Equity in the delivery of health care in europe and the us. *Journal of health economics*, 19(5): 553–583, 2000.

M. Varkevisser, S. A. van der Geest, and F. T. Schut. Do patients choose hospitals with high quality ratings? empirical evidence from the market for angioplasty in the netherlands. *Journal of Health Economics*, 31(2):371–378, 2012.

S. Vick and A. Scott. Agency in health care. examining patients' preferences for attributes of the doctor–patient relationship. *Journal of health economics*, 17(5):587–605, 1998.

J. Von Neumann and O. Morgenstern. Theory of games and economic behavior. *Bull. Amer. Math. Soc*, 51(7):498–504, 1945.

R. Wang. The optimal consumption and the quitting of harmful addictive goods. *The BE Journal of Economic Analysis & Policy*, 7(1), 2007.

P. Webley and E. K. Nyhus. Parents' influence on children's future orientation and saving. *Journal of Economic Psychology*, 27(1):140–164, 2006.

R. E. Weller, E. W. Cook III, K. B. Avsar, and J. E. Cox. Obese women show greater delay discounting than healthy-weight women. *Appetite*, 51(3):563–569, 2008.

K. A. Wickrama, R. D. Conger, L. E. Wallace, and G. H. Elder Jr. The intergenerational transmission of health-risk behaviors: Adolescent lifestyles and gender moderating effects. *Journal of health and social behavior*, pages 258–272, 1999.

P. Wise and P. Foolish. Hospital cost of care, quality of care, and readmission rates. *Arch Intern Med*, 170(4):340–346, 2010.