Università degli Studi di Cagliari

PHD DEGREE

Economics and Business

Cycle XXXI

Mergers in Public Health and Education: A Policy Evaluation Design

Scientific Disciplinary Sector

SECS-P/03

PhD Student:	Sara Pau
Coordinator of the PhD Programme	Prof. Andrea Melis
Supervisor	Prof. Rinaldo Brau
Second Supervisor	Prof.ssa Adriana Di Liberto

Acknowledgments

I am very grateful to prof. Rinaldo Brau and Adriana Di Liberto for their guidance. I am also indebted to prof. Joan Llull, Yun Cao, Friedrich Breyer, Silvia Balia, Eugenio Zucchelli and Marco Nieddu for their comments and valuable advice.
Moreover, I am grateful to the School Principals prof. Rosanna Sardu and Antonietta Melas for our thorough conversations on their professional experiences, of which chapter three truly benefited.
Furthermore, I would like to acknowledge my friends and colleagues Christelle Traboulsi, Giulia Contu, Carola Casti, Gianfranco Pischedda, Salvatore Loriga, Farideh Tavazoee, Giulia Boi, Simone Zandara, Cristian Usala, You-jun Shin, Alberto Nonnis and Luca Naitana for their help, brainstorming sessions, technical assistance, and, most importantly, for their friendship.
Finally, I thank my family and my friends for always being by my side during my pursuit of this doctoral degree and for supporting me all the way.

Contents

List of Figures

List of Tables

Summary

Over the last decades, in Italy, the increase of life expectancy and the consequent population aging raised concerns on the dynamics of health spending in such a way that institutional settings have changed and policies mostly oriented at containing costs and increasing efficiency have been implemented. The mergers of Local Health Authorities (LHAs henceforth) represents a prominent example. Indeed, the Italian LHAs experienced a drastic reduction in their number, passing from 659 in 1992 to 101 in 2017.

The same pattern characterizes the educational system in Italy. New centralized criteria were introduced with the School Consolidation Process (*Dimensionamento Scolastico*), established by the D.P.R. 233/1998. The schools with less than 500 students (300 in mountain areas and small islands) would have been merged to another school or managed by a regent School Principal (SP hereafter), namely a SP "borrowed" from another school.

This doctoral thesis investigates the impact of such consolidation policies in health and education institutions on patients' and students' outcomes, respectively.

Chapter one provides an introductory overview on the new dynamics of life expectancy, GDP and education in Italy, and is chiefly aimed at mapping the longstanding geographical disparities among the NUTS 1 macroregions.

We start the analysis by presenting an extensive literature review on health capital accumulation and classical health determinants, namely income and education. We then shift attention to a

descriptive analysis of the most recent data on regional life expectancy, GDP and different levels of educational attainment. Findings suggest that the Italian economy is still characterized by a sharp dichotomy between the North and the South, confirming the existence of “two Italies”. If such discrepancies are not properly taken into account by central governments when designing public policies, further inequality may be created.

Chapter two starts investigating the effects of one merger policy that interested the National Health Service in Italy. More specifically, the chapter evaluates the impact of LHAs consolidation process on population mortality within Italian municipalities from 2002 to 2016. We apply a quasi-experimental research design using an event study framework to test for potential pre-existing trends and anticipation effect, as well as phase-in policy impact. Since each region set up the reform in different years, we use a staggered two-way fixed effect Difference in Difference as identification strategy, and distinguish between mountain and non-mountain municipalities to detect possible heterogeneous effects. Findings show that population mortality has increased in the years following the policy change, with the mountain municipalities being the most affected, and suggest that policy makers should carefully weight the efficiency gains with the impact on health care effectiveness, specifically on local communities, before allowing more mergers.

Chapter three analyses another merger policy that concerned the Italian educational system. More precisely, the chapter aims to assess the impact the consolidation process that involved the Italian schools on two measures of student dropouts: voluntary dropout rates and grade retention. With a focus on secondary schools, we exploit the fuzzy discontinuity in treatment regime and apply a continuity-based local polynomial approximation Fuzzy RDD.

Estimation results show that the program exposure may have had detrimental heterogeneous effects in upper secondary schools, especially for treated schools far away from the directional school (where the school principal is based) at least 10 km. As a result, the isolation of the merged schools appear as a causality channel that makes the principals' responsibilities substantially aggravated. We conclude that the school consolidation process appears as a cost-saving policy that may fail to consider the indirect costs arising in term of poorer school performance, which may later produce labor market inequalities due to losses in human capital accumulation.

Chapter 1

Mapping Regional Disparities in Health Outcomes and their Determinants: The Two Italies

Abstract

This introductory chapter provides an overview on the new dynamics of health outcomes in Italy, and is chiefly aimed at mapping the longstanding geographical disparities among the NUTS 1 macroregions. We start the analysis by presenting an extensive literature review on health capital accumulation and classical health determinants, namely income and education. We then shift attention to a descriptive analysis of the most recent available data on regional life expectancy, GDP and different levels of educational attainment. Findings suggest that the Italian economy is still characterized by a sharp dichotomy between the North and the South, confirming the existence of "two Italies". If such discrepancies are not properly taken into account by central governments when designing public policies, further inequality may be created.

1 Introduction

In this introductory chapter, we aim to map the new dynamics and spatial disparities on population health in the Italian context with the most recent available data. More precisely, this study explores the trends in health outcomes and socioeconomic conditions, concentrating on their geographical distributions among the Italian regions during the period 1990-2016.

We start by reviewing the existing literature on how health is produced at the individual level and what makes us healthier, with a specific focus on the classical determinants of health, namely income and education.

The issue turns out relevant since people value health highly. Thus, a clear understanding of the mechanism by which health is produced is a starting point for policies oriented at enhancing population well-being and at lessening inequalities in society.[1]

Given the wide diffusion of medical and health technologies in a country where the National Healthcare Service was established in 1978, and Essential Levels of Care (*Livelli Essenziali di Assistenza, LEA*) are set by the central state to assure free, equal and effective health care to all individuals throughout the country, one might expect the mortality being dissociated to the socioeconomic conditions. However, although health care is universally available, in Italy still persists an association between income distribution, education and population health.

Furthermore, we will pay attention to the most common methodologies and findings on health forecasting. In fact, understanding the path of health in the future, and consequently of longevity, is crucial to shape the social security system in a sustainable manner, according to the forthcoming levels of people's quantity of life.

This chapter will be split into three further sections. First, we examine the salient literature

[1] One example of such policies is PROGRESA (*Programa de Educación, Salud, y Alimenación*), implemented in 1997 by the government of Mexico as an integrated approach to poverty alleviation and health status improvement through the development of human capital (Schultz, 2004).

on health capital accumulation, socioeconomic determinants of health, and on the future health forecasting methodologies and findings. Second, we map the spatial disparities in the Italian regions during the period 1990-2016 with respect to different specifications of life expectancy, income and education. Finally, the chapter concludes.

2 Literature Review

2.1 The Determinants of Health

Since health is one of the most important aspects of human being, many scholars have focused on understanding what makes us healthy. More specifically, increasing attention has been paid by economists to health capital accumulation and to the relationship between health and socioeconomic conditions. The former can count on the celebrated paper by Grossman (1972) and its further extensions. The latter is the subject of a voluminous literature, with a main focus put on the classical determinants of health: income and education. Both theoretical and empirical contributions have been made and are reviewed in what follows.

2.1.1 Theoretical Studies

Health Capital Accumulation Theoretically, the standard economic model of health, developed by Grossman (1972), illustrates the mechanism underlying the health production function. In this model, health is seen as a durable capital stock that produces healthy time. Individuals inherit an initial stock of health that depreciates with age but can also be increased with investment. The intertemporal utility of the representative consumer is given by:

$$U = (\phi_0 H_{0,...,} \varphi_n H_n, Z_{0,...,} Z_n),$$

where H_0 is the inherited stock of health, H_i is the stock of health in the ith period, φ_i is the service flow per unit stock, Z_i is the consumption of a generic good in the ith period.

The net investment in the health stock is equal to:

$$H_{i+1} - H_i = I_i - \delta H_i,$$

where I_i is the gross investment and δ is the depreciation rate (which is assumed exogenous but it may vary with age of individuals).

The gross investment I_i is a function of medical care M_i, time inputs TH_i and human capital E_i:

$$I_i = I_i(M_i, TH_i, E_i).$$

The health capital at any given time is thus a result of a set of factors that comprises family genetic endowments, health in the previous time period, medical care received, the stock of knowledge about what improves health and a vector of other personal and environmental level factors.

An alternative approach is made by Dalgaard and Strulik (2014) which is based on health deficits rather than health capital. The authors note that, during the process of aging, humans develop an increasing number of deficits, for which Mitnitski *et al.* (2002) propose an index of "frailty". The law of increasing frailties or deficits accumulation equation can be expresses as the following:

$$\dot{D}(t) = \mu(D(t) - E),$$

where D is the proportion of deficits, μ is considered a physiological parameter that captures the force of aging and E is interpreted as the impact of nonbiological factors, namely the attempt of an individual to slow down the aging process by means of a deliberate health investment. When the number of deficits rises, the human body becomes more frail, leading, eventually, to death. That is, the authors define actual health as the maximum achievable health (i.e., the state of health of a 15 years old), minus all accumulated deficits:

$$H = \bar{H} - D,$$

where $\bar{H}$ is the maximum health and D is the accumulated health deficits. This leads to $\dot{H} = -\dot{D}$. Inserting the deficits accumulation equation into $\dot{H}$ and substituting D with $\bar{H} - H$, they obtain:

$$\dot{H} = -\mu(\bar{H} - H - E),$$

that can be rewritten as:

$$\dot{H}(t) = \mu E - \mu(\bar{H} - H(t)).$$

According to the above equation, an improvement in health status $H(t)$ makes the aging process slow down and is consistent with the fact that health losses are small for high levels of health stock and become bigger as the health status deteriorates.

Health and Economic Conditions Smith (1999) deeply analyses the nexus between health and wealth, which is the stock of disposable economic resources mainly accumulated through the flows of income.

Economic resources and health increasingly interact over people's life. Households savings and consumption are seen as a function of health. Presumably, based on current information about their health, people make projections about their future health states. The risk of future morbidity may affect positively saving behavior and alter households wealth profile eventually. The prospect of high medical expenses has been the main way in which health risks have been incorporated into life-cycle models (Lillard and Weiss, 1997). On the contrary, savings may decrease if health deteriorates because that reduces the amount of labor supplied or increases the purchase of health care goods.[2] Another reason why health may affect savings is that marginal utility of consumption could depend on health status. If the marginal utility of consumption declines with poorer health conditions, individuals will consume more when they are healthy, and savings will be higher when they are unhealthy. Since health shocks may permanently alter heath stock in future years, potential effects on future income, consumption, and medical expenses can be systematic. Wealth may be higher among individuals who did not experience negative health shock in the previous years, and more economic resources could protect individuals from the ravages of age so that their subsequent health is better. Interestingly, Smith notes that inequality raises levels of psycho-social stress which negatively affects endocrine and immunological processes. Across industrialized countries, average mortality is not associated to average income differences between countries, but rather to income inequality within countries. At least in industrial countries, the stress associated with being at the bottom end of an unequal society matters. Social cohesion enters then in the model as inequality antagonist and health-enhancing trait. Societies placing a strong value on caring for one another are supposed to be healthier ones.

Cutler *et al.*, (2006) contributed to the debate on the modern dynamics of health and wealth by noting that during the twentieth century in the United States and other high-income countries, growth in real incomes was associated with a rise of longevity by nearly 30 years. The decline in mortality came along with better health living conditions: people are not only living longer

[2] During retirement, new health shocks will not alter income since the pension is fixed and no labor is supplied. However, they can still deteriorate savings.

but also better and healthier than their forebears.

2.1.2 Empirical Studies

Understanding empirically the relationship between health and economic conditions is a challenging exercise. Difficulties arise for several reasons. One first issue is ontological and regards how to measure health and economic conditions. These latter are indeed multidimensional entities, and different proxies are adopted by researchers and further reviewed in this section.
Secondly, the most powerful analytical methodologies are not suitable in this context: randomized and controlled experiments involving exogenous manipulations of human beings' statuses are not ethically possible.
However, it is widely recognized that a variety of social factors, such as income and education, has an impact on health outcomes (Zimmerman *et al.*, 2015).

Health and Income Even though the mechanism underlying the relationship between health and income is the subject of voluminous literature in the past four decades, the direction of causality is still doubtful and often depending on the specific characteristics of social, economic, and political contexts (De Vogli *et al.*, 2005). The pathways of causation between income and health has been summarized as follows. High levels of income may lead to health improvements through a better lifestyle or diet, fewer monetary worries, better access to medical services and an improved living environment (Adler *et al.*, 1994; Smith, 1999).
In contrast, poverty affects health through different channels such as lack of resources held by individuals to buy medical treatments or sufficient food, lower access to prevention and precarious housing conditions. Systematic underinvestment may also be a result across a wide range of human, physical, health, and social infrastructures (for instance, education, health services,

transportation, environmental controls, availability of good quality food, quality of housing, and occupational health regulations) (Lynch *et al.*, 2000).

The relationship also goes in the other direction. Indeed, health enters in the accumulation of human capital function: people in good health are more likely to be economically productive, invest more in schooling and have higher incomes (see Curie and Madrian 1999, for a review of the evidence).[3] Other channels can be found: health affects how much an individual works, and how much an individual works determines the skills acquired on the job (learning-by-doing), then variations in health will influence human capital and, subsequently, future labor supply choices and income. Prior bad health conditions may alter past incomes which may affect current income. Periods of poor health in middle age reduce earnings and could have negative implications for pension and social security income during retirement (Hokayem and Ziliak 2014). Therefore, health and income are simultaneously determined and reverse causality problems emerge when dealing with finding their association.

Moreover, individual unobservable characteristics and third factors are likely to determine both income and health. Examples are genetics endowment, social background, discount rates (Frijters *et al.* 2005). Individuals with a low rate of time preference may undertake investments in human capital that enhance future earnings as well as engage in behaviors that improve future health (Barsky *et al.*, 1997).

We now review the empirical evidence on the association between health and income more in detail, concentrating on data used, methodology applied and obtained findings.

Preston (1975) was the first drawing the graph about the relationship of average health and income, that took the name of "Preston Curve", depicted in figure 1.1. In his seminal contribution, he presented a scatter diagram of the relationship between level of cross-country life expectancy (average, male and female) and per capita national income (expressed in 1963 U.S.

[3]A deeper attention will be paid to the association between health and education in the following paragraph.

dollars) in the 1900s, 1930s, and 1960s, showing that life-expectancy is smaller for countries with lower levels of per capita income.

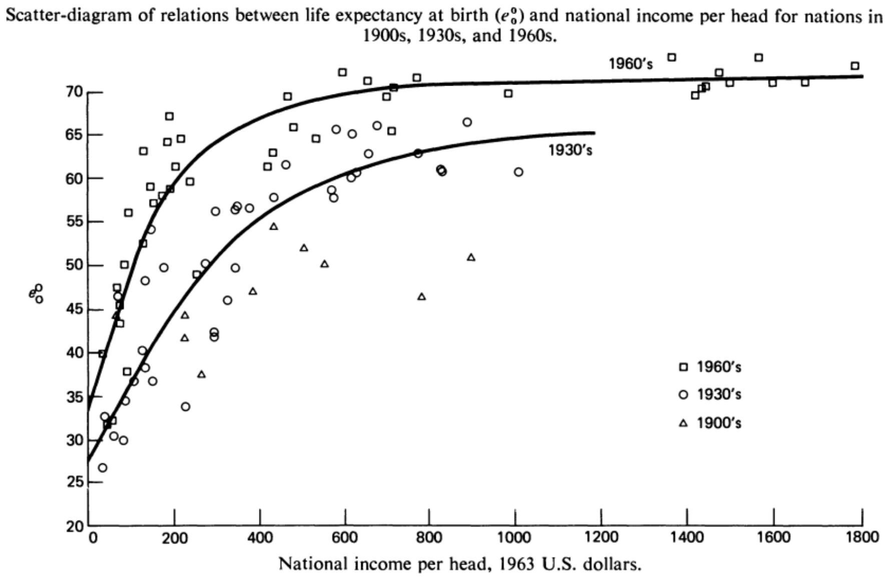

Figure 1.1: Preston Curve. Source: Preston (1975)

One of the first empirical work which aimed at estimating the structural effect of income on different measures of health was Ettner (1996). More specifically, the paper analyses the effect of family income on self-assessed health status, work and functional limitations, bed days, average daily consumption of alcohol, scales of depressive symptoms and alcoholic behavior.[4] Ettner applied an ordered probit instrumental variable approach where family income was instrumented with state unemployment rate, work experience, parental education and spouse characteristics, using individual-level data from the 1987 National Survey of Families and Households (NSFH), the 1986 and 1987 panels of the Survey of Income and Program Participation (SIPP), the Alcohol Supplement of the 1988 National Health Interview Survey (NHIS) and Bureau of Labor Statistics data on state unemployment rates. Based on the IV estimates, the results show a strong positive association between income and self-assessed health status and a negative asso-

[4]Number of days in the previous four months when the interviewed had to stay in bed more than half a day due to illness, including hospital stays.

ciation between income and depressive symptoms, work limitations, functional limitations and hospital bed days. As the author noted, the main shortcoming of the analysis approach is the exclusion restriction of the instruments used: state unemployment rate, parental education, work experience and spouse characteristics may be directly related to health.

Ruhm (2000), with longitudinal data of 50 states and District of Columbia over the 1972-1991 period and fixed-effect model, examines how economic conditions affect mortality rates for ten specific causes of death.[5] His study provides evidence that health improves as the macroeconomic conditions deteriorate. Specifically, state unemployment rates are negatively and significantly correlated to mortality and eight of the ten causes of death, with suicides representing an exception. One percentage point rise in the unemployment rate is associated with a 0.5 to 0.6 percent decrease in total mortality (in absolute terms, a reduction of around 11,000 fatalities annually). In fact, higher joblessness is associated with reduced smoking and obesity, increased physical activity and improved diet. Leisure time declines when the economy expands, making it more costly to undertake health-enhancing activities (such as exercise). Similarly, the time price of medical care will rise if individuals work more hours, finding it harder to schedule medical appointments. Hence, lifestyles should become less healthy during economic upturns. Moreover, during expansions, work-related accidents are more likely. Drinking and driving rise in good times, leading to higher motor vehicle fatality rates (Evans and Graham 1988; Ruhm 1995).

Further, Meer *et al.* (2003) deal with endogeneity in this debate using inheritance during the last five years as an instrumental variable for changes in wealth in a probit approach where the dependent variable was a dichotomous self-reported health status variable (healthy vs. not healthy). The authors use data from four waves of the Panel Study of Income Dynamics (PSID)

[5] In details: malignant neoplasms, major cardiovascular diseases, pneumonia or influenza, chronic liver disease and cirrhosis of the liver, motor vehicle accidents, other accidents and adverse effects, suicide, homicide and legal intervention, infant mortality (deaths within the first year) and 10) neonatal mortality (deaths within the first 28 days). These ten specific sources account for around 80 percent of all deaths.

during the time span 1984–1999.[6] They confirm the positive association between health and changes in wealth. However, this result turns out insignificant after instrumenting.

Frijters *et al.* (2005) investigate the causal relationship between households income and self-assessed health satisfaction and interestingly exploit the wall of Berlin fall in 1989 and the consequent German reunification as a natural experiment resulting in an exogenous variation of real income for Est Germans. Reunification was completely unanticipated and determined a rapid and exogenous rise in average real household incomes for East Germans. They used the German Socio-Economic Panel (GSOEP) between 1984 and 2002 and applied fixed-effects ordinal estimator to control for individual unobservable heterogeneity. The results suggest that increased income leads to improved health satisfaction even though the quantitative size of this effect is very small: one log point increase in income leads to a 0.083 (in a 10 points scale) improvement in health satisfaction for East men, 0.067 for West men and 0.088 for West women. In some robustness exercises - that did not change their results - they used different measures of income: a measure of "permanent" income averaged over a 3 or 5 years period, a household-equivalised income measure (household income adjusted taking into account the number of family members) and a measure of relative household poverty (*i.e.*, being at or below 60% of the median household income).

There is also a growing consensus that improving health can have equally large indirect payoffs through accelerating economic growth (see, e.g., Bloom and Canning 2005).

Murphy and Topel (2006) measured the value of falling mortality in the United States and the social benefits of better health living conditions during the twentieth century. Based on individuals' willingness to pay, they found cumulative gains in life expectancy after 1900 were worth over 1.2 million per person in 2000, whereas post-1970 gains added about 3.2 trillion per year to national wealth, equal to about half of GDP.

[6]The Panel Study of Income Dynamics (PSID) is a longitudinal survey, ongoing since 1968, of a representative sample of U.S. individuals (men, women, and children) and the families in which they reside.

Further, Acemoglu and Johnson (2007) investigates the effect of general health conditions, proxied by life expectancy at birth, on economic growth, exploiting the large improvements in life expectancy driven by the international epidemiological transition (namely the international health interventions, the more effective public health measures and the introduction of new chemicals and drugs starting in the 1940s, e.g., discovery and mass production of penicillin and streptomycin, or the discovery and widespread use of DDT against mosquito vectors). Using the country's initial mortality rate from various diseases around the world before the international epidemiological transition as an exogenous variation for increased life expectancy, they found no evidence that the large increase in life expectancy raised income per capita.

Hokayem and Ziliak (2014), using data from Survey Research Center subsample of the PSID from during the years 1999-2009, with a GMM estimation demonstrated a positive association between health conditions and future wages: one unit increase in standard deviation of health status, the future wages (t+1) increase by 1.3 percent. The used data describe a situation where, starting around age 35, unhealthy men have significantly lower life cycle profiles of wages and assets, which is preceded by an increase in obesity around age 30. The rise of obesity could be linked to different poor health conditions such as type II diabetes, high blood pressure, heart disease and some forms of cancer. Even though many obesity-related conditions are not permanently disabling and thus do not directly affect the labor productivity, they make individuals more susceptible to illnesses that may interrupt the human capital accumulation process.

Boyce and Oswald (2012) study the effect of promotion, namely an increase in the job status and wage, on health conditions (both subjective and mental health status) to test the hypothesis that individuals with high occupational status have good health and low rates of premature mortality. They used longitudinal data, from British Household Panel Survey (BHPS), on a sample of British workers and applied cross-sectional and difference-in-difference methods between promoted group non-promoted control group.[7] Little evidence is found that a person's health improves after promotion. In details, while the promoted group has significant better subjective health status across different time periods after promotion, the same individuals have significantly better mental health only at period T, suffering a relative worsening in their mental health from period T to T+3.

More recently, Chetty *et al.* (2016), measured the level, time trend, and geographic variability in the association between income and life expectancy in the United States from 2001 to 2014. They found that higher income was accompanied by a greater life expectancy, which differences across income groups increased over time. Also, they highlighted a substantial variation in the association between longevity and income across areas: differences across income groups decreased in some areas and increased in others and they were correlated with health behaviors and local area characteristics.

Health and Education The twentieth century experienced a dramatic increase in world-wide population life expectancy. According to the Human Mortality Database (HMD), life expectancy at birth in Italy rose from 41.85 in 1900 to 79.21 in 1999, becoming 82.95 in 2014. Oeppen and Vaupel (2002) suggest that the increase in life expectancy was mainly driven by improvements in education, as well as in sanitation and housing, causing a decline at a steady pace in early and mid-life mortality, which was mainly due to bacterial infections. In fact, the positive correlation between health and education is well-documented in the literature (Adams *et al.*, 2003 cite an exhaustive summary).

Broadly speaking, there are three possible explanations for the link between health and education. Cutler and Lleras-Muney (2006) provide an extensive review.

[7] The British Household Panel Survey has been ongoing since 1991, following the same representative sample of individuals and it is household-based, interviewing every adult member of sampled households.

One possibility is that poor health leads to low levels of schooling. In fact, sick students are more likely to miss school, less likely to learn while in school, and ultimately obtain fewer years of schooling (Case, Fertig and Paxson 2005).

Another possibility is that increasing education improves health. Indeed, more educated individuals may have "better" jobs that offer safer work environments, better access to information and improved critical thinking skills. As a result, highly educated people appear to trust science and to make use of new health related information first, which affects the adoption of health-enhancing behaviors such as improved diet, exercise, quitting smoking and drinking. Moreover, they are more likely to comply correctly with their medical treatments and even non-complex health behaviors, such as belt use, have a strong education gradient. Another causality channel that goes from education to health is related to the social support systems: the more educated individuals may have larger social networks that provide financial, physical and emotional support, which may have a causal effect on health (Berkman 1995).

Lastly there may be third factors, such as family background, genetic traits, abilities and time preferences, that increase both schooling and health. Smarter individuals may be more likely to obtain more schooling and also take better care of themselves. Another possibility is that individuals with lower discount rates are more likely to invest more in both education and health.

De Vogli *et al.* (2005) specifically calculate the Pearson correlation coefficients between health and education. [8] In their findings, educational attainment is positively correlated with life expectancy at birth ($r = 0.476; p < 0.001$). More precisely, regions with a higher proportion of 19 years old people with a high school diploma experience higher level of life expectancy compared with regions with a lower proportion of young adults of the same age holding the

[8] For the sake of completeness, they also calculate the Pearson correlation coefficients between health, income per capita and income inequality for all 20 Italian regions. They find that per capita income was positively correlated with life expectancy at birth ($r = 0.538, p < 0.001$), and income inequality, measured by Gini coefficient, has a strong negative correlation with health conditions, defined again in terms of life expectancy at birth ($r = -0.785; p < 0.001$). Regions whose income inequality was higher, such as Campania and Sicily, had a significantly lower life expectancy than regions where income inequality was relatively low (e.g., Marche and Umbria).

same quantity of educational attainment.

It is important for policy to understand the mechanisms underlying the link between health and education, and whether and to what extent the effect of education on health causal is causal. Subsidies for schooling or any other kind of policy supporting education would be effective in improving population health only if in fact education causes health.

Clark and Royer (2013) try to overcome the difficulty of establishing to what extent the correlation between education and health reflects a direct causal effect. The authors use the changes in British compulsory school laws in 1947 and 1972 as exogenous source of variation of education. The 1947 change established that children who turned 14 before 1 April 1947 (born before April 1, 1933) could leave school when they turned 14, whereas children born after this date could not leave until they turned 15. The 1972 change extended the compulsory schooling age and children born after September 1, 1957 could not leave until they turned 16. With different sources of data, the results show that compulsory school law changes affected the educational attainment of people born just days apart, who had more years of schooling. With a regression discontinuity method the authors suggest, however, that the health returns to this extra education are small and cohort-specific.

Kemptner*et al.* (2011) apply the same approach in the German case. They exploit the changes in compulsory schooling laws between 1949 and 1969 as an exogenous variation of education to investigate the causal effect of years of schooling on health and health-related behavior. Applying an instrumental variables approach, they find evidence for a strong and significant causal effect of years of schooling on long-term illness and work-disability about among men (while they found a non significant effect among woman). Although they find a weak evidence of a causal effect of education on health-related behavior, their estimates still suggest significant non-monetary returns to education on health outcomes.

2.2 Forecasting Mortality and Life Expectancy

As already mentioned in the previous paragraph, the dramatic increase in worldwide population life expectancy during the first half of the twentieth century was mainly due to a decline at a steady pace in early and mid-life mortality, driven by improvements in sanitation, housing, and education. The increasing trend in longevity continued with the development of vaccines and antibiotics. However, by the latter half of the twentieth century, there was little room for further contraction in early and mid-life mortality. The continuing increase in life expectancy is chiefly due to a new phenomenon: the decline in late-life mortality (Oeppen and Vaupel, 2002). Consequently, the share of elderly people has grown as the life expectancy rose higher and higher.

Although the rising longevity is a worldwide health objective to increase people's well-being, the aging of the population questions the sustainability of social security system, whose plan relies on previous lower life expectancy levels. Forecasting mortality rates and, consequently, population longevity, becomes increasingly important to reshape the public welfare accordingly. That is, different methods for mortality modeling and forecasting have been developed (for reviews see Pollard 1987; Tabeau 2001; Wong-Fupuy and Haberman 2004; Booth and Tickle 2008). They can be divided into three approaches: extrapolation, explanation, and expectation (Booth and Tickle 2008).

Extrapolative methods use the regularity typically found in both age patterns and trends in time.

The explanation approach use structural or epidemiological models of mortality from certain causes of death for which the key exogenous variables are known and can be measured.

The expectation approach is based on the subjective opinions of experts involving varying degrees of formality (Stoeldraijer *et al.*, 2013).

The most applied and developed methods fall into the category of extrapolative approaches, of which the Lee-Carter method plays a dominant role and has become a milestone of the literature. This method specifies mortality by age and period for a single population as an overall time trend, an age component, and the extent of change over time by age (Lee and Carter 1992), by means of the Singular Value Decomposition. More specifically, the original formulation of the model is the following:

$$ln(m_{x,t}) = \alpha_x + \beta_x k_t + \epsilon_{x,t}.$$

The log of a time series of age-specific death rates $m_{x,t}$ is seen as the sum of an age-specific term α_x, which represents the general mortality shape across age and is independent of time, and another component that is the product of a time-varying parameter k_t, representing the change in the level of mortality across time, and an age-specific component β_x, that represents the age specific response to variations in the time index. Lastly, an error term $\epsilon_{x,t} \sim N(0; \sigma^2)$.
Afterwards, the Lee-Carter method has been extended in several direction to include a cohort dimension (Renshaw and Haberman, 2006) and medical knowledge and information on behavioral and environmental changes, such as smoking and obesity (e.g. Pampel 2005; Bongaarts 2006; Janssen and Kunst 2007; Stewart, Cutler, and Rosen 2009; Wang and Preston 2009; King and Soneji 2011; Janssen, van Wissen, and Kunst 2013, Danesi *et al.*, 2015).
Haberman and Russolillo (2005) investigate how the Lee-Carter approach can be used to forecast Italian national mortality, using death rates from year 1950 to 2000 to forecast the level and of the age distribution of mortality for the next 25 years. They also used the standard Box and Jenkins ARIMA model to make a comparison. They find different results, as expected, with the Lee-Carter model projecting a life expectancy in 2025 of 86,58 for females and 79,98 for males while time-series-based forecasts of 89,19 for females and 82,48 for males.
The Lee-Carter model is very robust in situations where age-specific log mortality rates have

linear trends, namely when relative gains to life expectancy are fairly constant year after year (Booth *et al.*, 2006). The Lee-Carter model also assumes that all the age-specific death rates move up or down together, although not necessarily by the same amounts (Haberman and Russolillo, 2005). However, the Lee-Carter model has some limitations when applied to recent data: some countries have less linear trend, as pointed out by different studies, [9] and we are witnessing a growth in mortality rate for some specific sub-group of population (see Case and Deaton 2015, for midlife white non-Hispanic Americans, Cutler *et al.* 2006, for American blacks and English male manual workers).

3 Data Description

We collect data for all 20 Italian regions from 1990 to 2016 to map the spatial distribution and disparities of the main health outcomes and the their determinants.[10] Our main source of information is the latest available version (June 2018) of ISTAT-Health for All.

The Italian economy is characterized by a sharp dichotomy between the northern and the southern regions. Different rates of growth separate the North from the South since Italy unification in 1861, originating what is known as "Southern Question" (Davis, 2015). The data analyzed in what follows confirm the existence of "two Italies".

3.1 Health Outcomes

We focus on four specific health outcomes: life expectancy at birth for males and females, and life expectancy at 65 years old, again by gender. The former is the most used and reliable proxy for the general health status of the population of interest. The latter assumes a crucial

[9] See e.g. Booth, Maindonald, and Smith 2001, for Australia; Renshaw and Haberman 2006, for England and Wales; Janssen, Kunst, and Mackenbach 2007, for the Netherlands.

[10] The period of observation is shorter for some variables, given the smaller availability of data.

importance to shape the social security system.

Figure 1.2 provides the historical trend of life expectancy at birth, both for males and females,

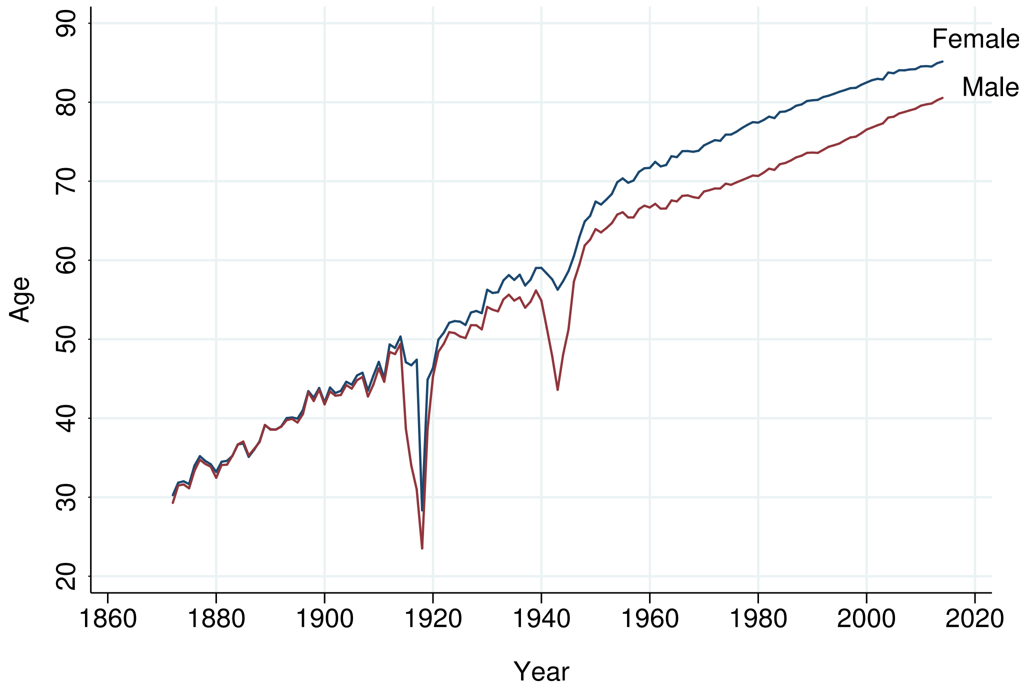

Figure 1.2: Italy - Life expectancy at birth 1872-2014. Source: Author graph based on HMD.

from 1872 to 2014 at the national level.[11] The figure allows to depict the striking twentieth century increase in population life expectancy. In fact, according to the HMD, life expectancy at birth in Italy rose from 41.85 in 1900 to 79.21 in 1999, becoming 82.95 in 2014. Interestingly, the figure suggest that life expectancy is not uniform by gender: a gap between male and female life expectancy started after the trough due to the World War I and became wider after the World War II.[12]

Cutler *et al.* (2006) explain this gender health inequality as a result of historical gender time-gap in health behaviors such as smoking. In fact, smoking may cause cardiovascular disease, with a relatively short lag, and lung cancer with a longer lag. Men registered an increase in

[11] The historical data disaggregated at the regional level are not available.

[12] In 1918-19 the peak also includes mortality due to the Spanish flu epidemic. For example, infant deaths from influenza rose from 972 in 1917 to 44.968 in 1918 (ISTAT, 2014).

smoking in the second quarter of the twentieth century and a current reduction in smoking, widespread throughout the rich countries. That contributed to the lesser mortality decline in the third quarter of the century. Contrariwise, women began to smoke later than men and women's smoking rates are still rising in some European countries, likewise their mortality rate from lung cancer. As a result, life expectancy by gender may converge in the near future.

Time and spatial variation of life expectancy at birth, aggregate for the 5 Italian macroareas and differentiated by gender, is depicted in figure 1.3 and 1.4.[13]

Figure 4 and 5 represent the time and spatial variation for life expectancy at 65 years old, again by genders and macroareas.

These graphs show two important features of the health dynamics: (1) women always register a higher longevity than men; (2) there is an evident and systematic ranking in the performance of the macroareas, with the North-East performing always better than everyone else in recent years. The North-West and the Center follow in this order and the South and the Island are ranked the lowest, showing also a bigger gap from the others.

Interestingly, while the ranking remains fairly constant along all the period of observation for female life expectancy, we observe an inversion of order at the edges of the time series for men. In fact, in 1990 North-West and North-East registered the worse performance in male life expectancy, probably due to the sharp differences in the labor market structure between genders and areas in that period.[14] The same geographical pattern is followed by the life expectancy at 65 years old for females and males, and depicted respectively in figure 1.5 and 1.6.

[13]These 5 macroareas, also known as macroregions, are the official NUTS 1 partition for Italy and they comprises homogeneous regions, either economically as well as along other dimensions. In details, they are North-West (which comprehends Valle d'Aosta, Piedmont, Lombardy, Liguria), North-East (Friuli-Venezia Giulia, Veneto, Trentino-Alto Adige, Emilia-Romagna), Center (Tuscany, Marche, Lazio, Umbria), South (Abruzzo, Molise, Campania, Basilicata, Puglia, Calabria) and Islands (Sicily and Sardinia).

[14]Indeed, the employment rate for men has always been higher than for females and, after a first reduction in work-related accidents during the post–World War II, Italy witnessed an increase in work injuries ever since, leading to the adoption of stricter occupational safety regulations in the following years. The National Institute for Insurance against Occupational Accidents (INAIL) confirm that the injuries to workers with permanent damage have increased significantly during from 1987 and 1995, exceeding one million cases. Therefore, the more industrialized North with higher male employment rates may have experienced an increase in mortality rates.

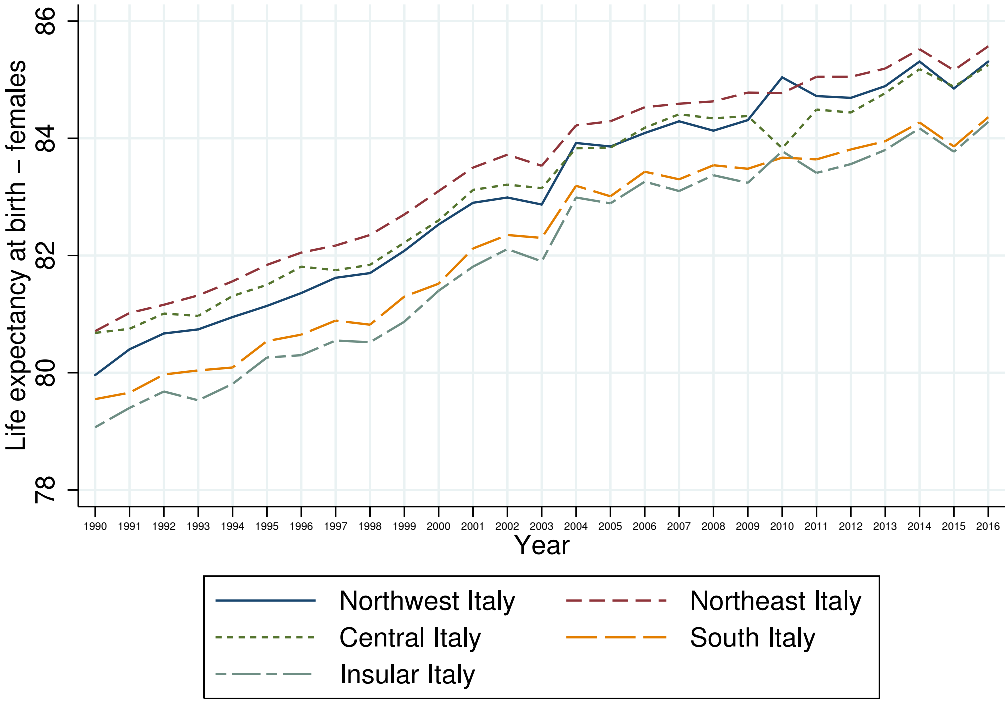

Figure 1.3: Life expectancy at birth - Females. Source: authors graph based on ISTAT data.

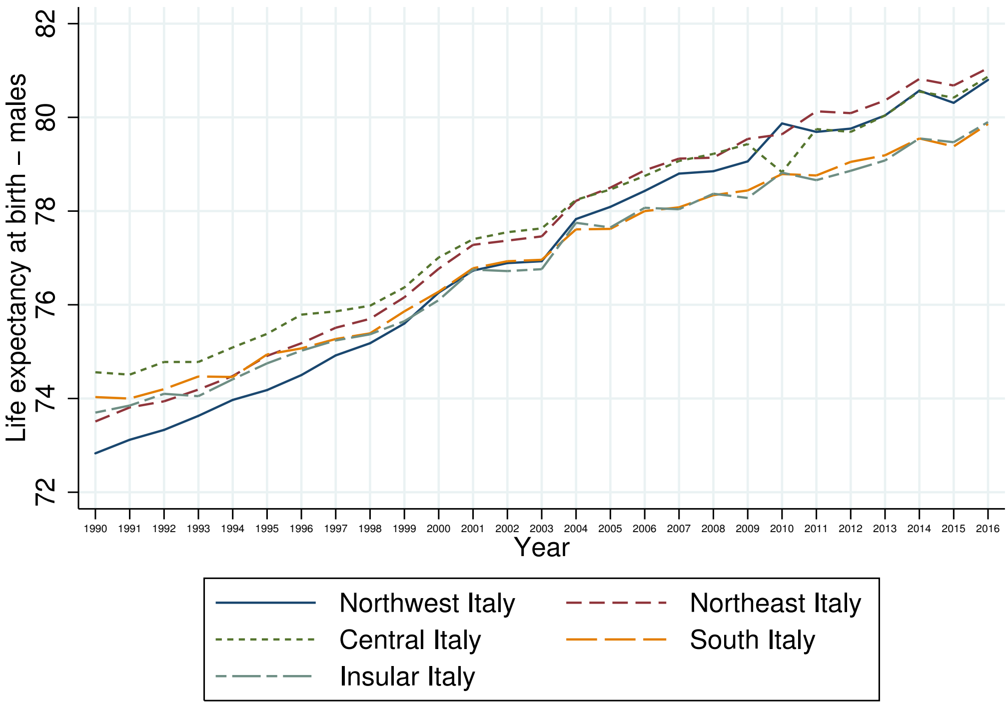

Figure 1.4: Life expectancy at birth - Males. Source: authors graph based on ISTAT data.

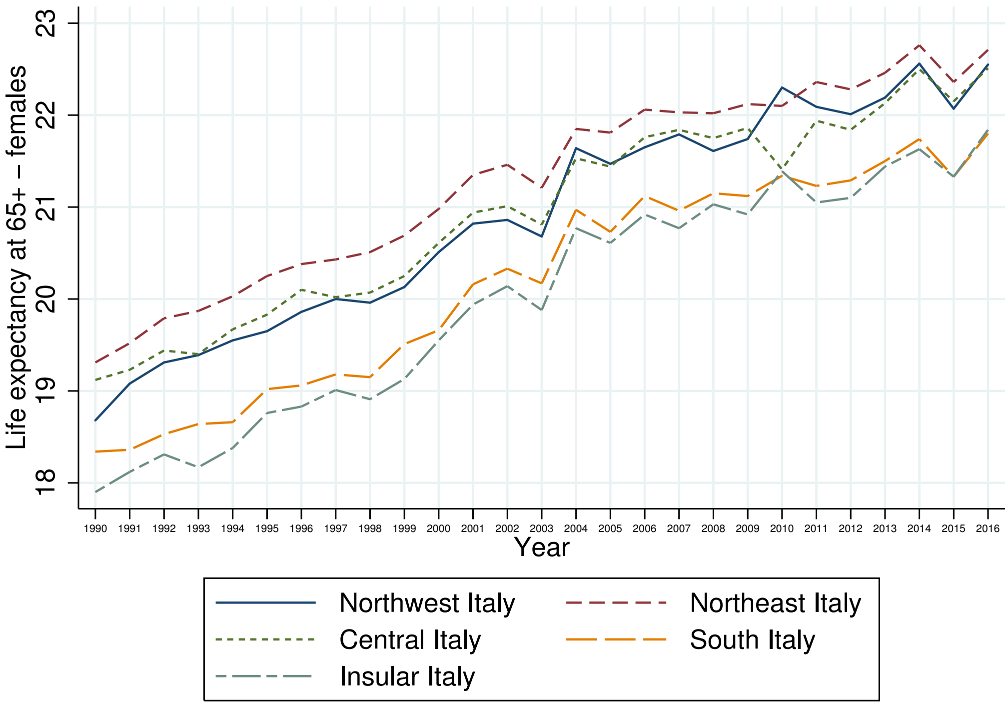

Figure 1.5: Life expectancy at 65 years old - Females. Source: authors graph based on ISTAT data.

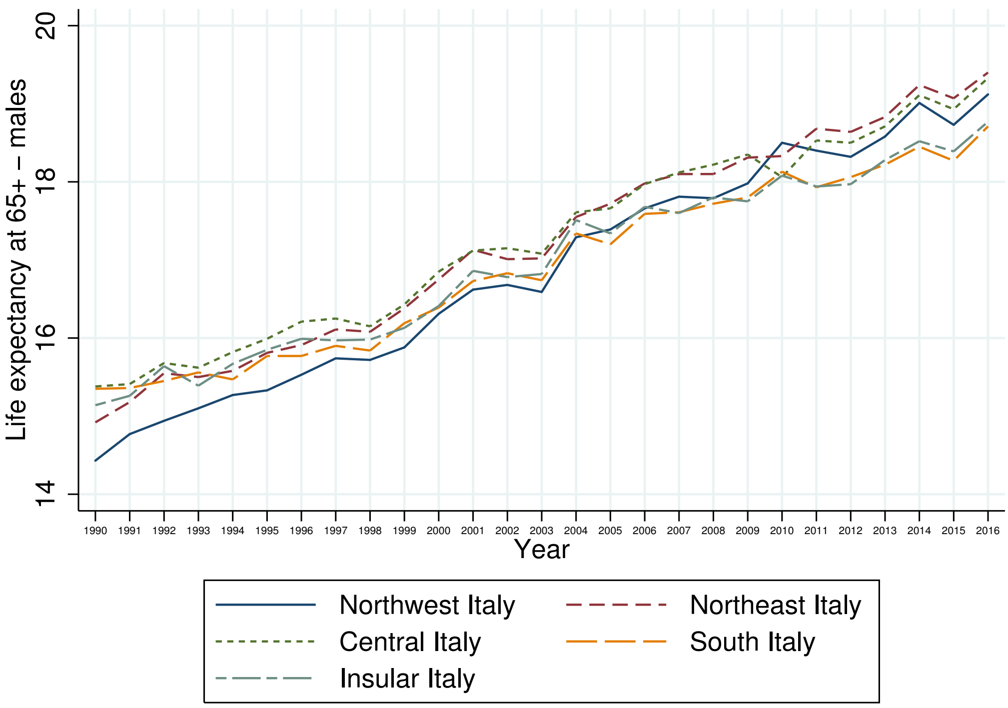

Figure 1.6: Life expectancy at 65 years old - Males. Source: authors graph based on ISTAT data.

Table 1.1 reports the main descriptive for the four health outcomes, divided by macroarea. In the best performing North-West region, the average life expectancy at birth during the period 1990-2016 was 77.496 years for men and 83.484 for women, with a standard deviation of 2.457 for men compared to 1.562 for women. The same gender discrepancy is observed for life expectancy at 65.

On the contrary, the worst performing Islands registered during the same time period an average life expectancy at birth of 76.851 years for men and 81.957 for women, with a higher standard deviation for men (1.992 vs 1.702 for women). Again, life expectancy at 65 years old follows the same pattern, with women living longer than men albeit the former register a slightly bigger standard deviation.

Table 1.1: Health outcomes: Summary statistics by macroarea

Variable	Mean	Std. Dev.	Min.	Max.	Growth rate	N
North-East						
Life expectancy 0 M	77.497	2.457	73.51	81.05	0.31***	27
Life expectancy 0 F	83.484	1.562	80.71	85.57	0.19***	27
Life expectancy 65 M	17.218	1.373	14.92	19.40	0.17***	27
Life expectancy 65 F	21.285	1.058	19.31	22.76	0.13***	27
North-West						
Life expectancy 0 M	77.043	2.603	72.83	80.8	0.32***	27
Life expectancy 0 F	83.012	1.717	79.96	85.31	0.21***	27
Life expectancy 65 M	16.870	1.457	14.43	19.12	0.18***	27
Life expectancy 65 F	20.894	1.186	18.68	22.56	0.15***	27
Center						
Life expectancy 0 M	77.63	2.068	74.51	80.87	0.26***	27
Life expectancy 0 F	83.101	1.500	80.68	85.25	0.18***	27
Life expectancy 65 M	17.268	1.242	15.38	19.33	0.15***	27
Life expectancy 65 F	20.932	1.074	19.12	22.51	0.13***	27
South						
Life expectancy 0 M	76.937	1.924	74	79.86	0.24***	27
Life expectancy 0 F	82.197	1.606	79.55	84.36	0.20***	27
Life expectancy 65 M	16.903	1.113	15.35	18.71	0.14***	27
Life expectancy 65 F	20.187	1.169	18.34	21.80	0.14***	27
Islands						
Life expectancy 0 M	76.851	1.992	73.70	79.90	0.25***	27
Life expectancy 0 F	81.957	1.702	79.07	84.28	0.21***	27
Life expectancy 65 M	16.945	1.121	15.14	18.77	0.14***	27
Life expectancy 65 F	19.994	1.260	17.90	21.84	0.15***	27

Source: Author elaboration on ISTAT-Health For All data. Growth rates for each of the 20 series are calculated using a regression of the respective variable with the time trend as the only exogenous variable. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3.2 Socioeconomic Conditions

Income As reviewed in section 2, one key variable associated to health is income. For this reason, our variable of interest is the regional per capita GDP (constant 2010 market prices). We focus on GDP because it is an indicator of living standards since it comprises the value of all final goods and services produced in a certain time.

Figure 1.7 shows the spatial and time variation of per capita GDP for the 5 macroreas, highlighting the big gap between the North and the South of the country, namely what is known as the Italian dichotomy. The standard ranking of the macroareas, which sees the north performing

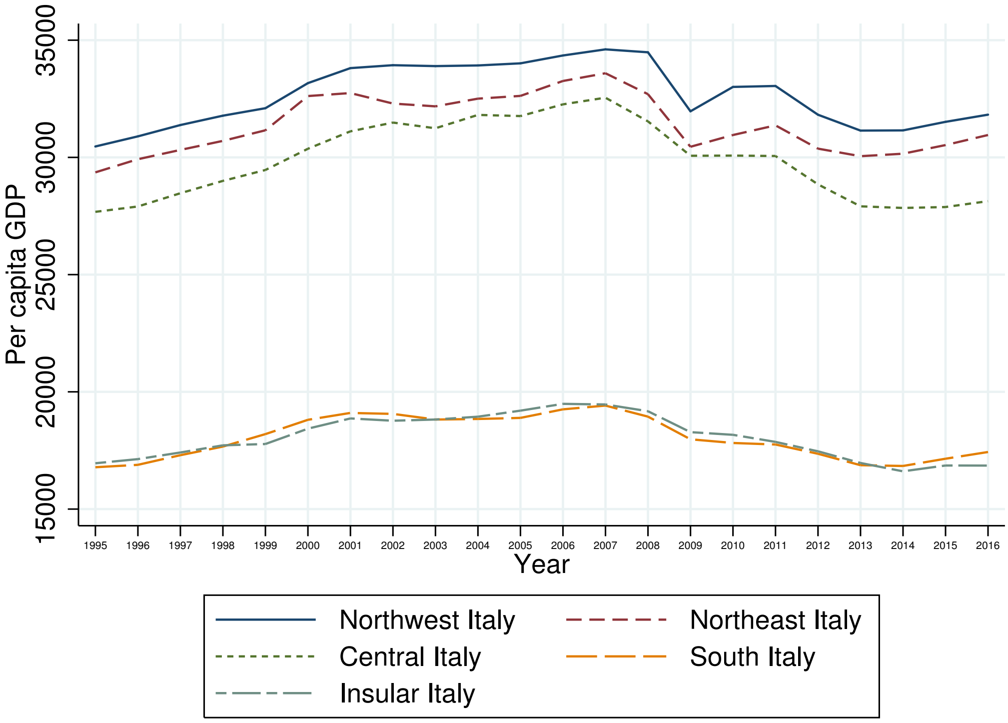

Figure 1.7: Per-capita GDP. Source: authors graph based on ISTAT data.

better than the central and the south holds.

Education Furthermore, we explore different levels of educational attainment, distinguishing once again between genders and among macroareas. More specifically, we consider the share of population with more than primary school diploma (calculated by using 1 minus the available data on the share of people with primary school diploma or no education at all), and the share of population with a university degree or a doctorate. The former is depicted in figure 1.8 (for men) and in figure 1.9 (for women). The latter follows in figure 1.10 and 1.11.

The graphs describe a country with two different groups of regions, those in the North and in the South, and two different subgroups of people, men and women, that grow at a very different pace. The share of population with more than primary school diploma is mostly increasing in our observed period but two big disparities arise. In recent years, the Italian "Mezzogiorno", which includes the South and the Islands, is experiencing the worst performance in terms of education

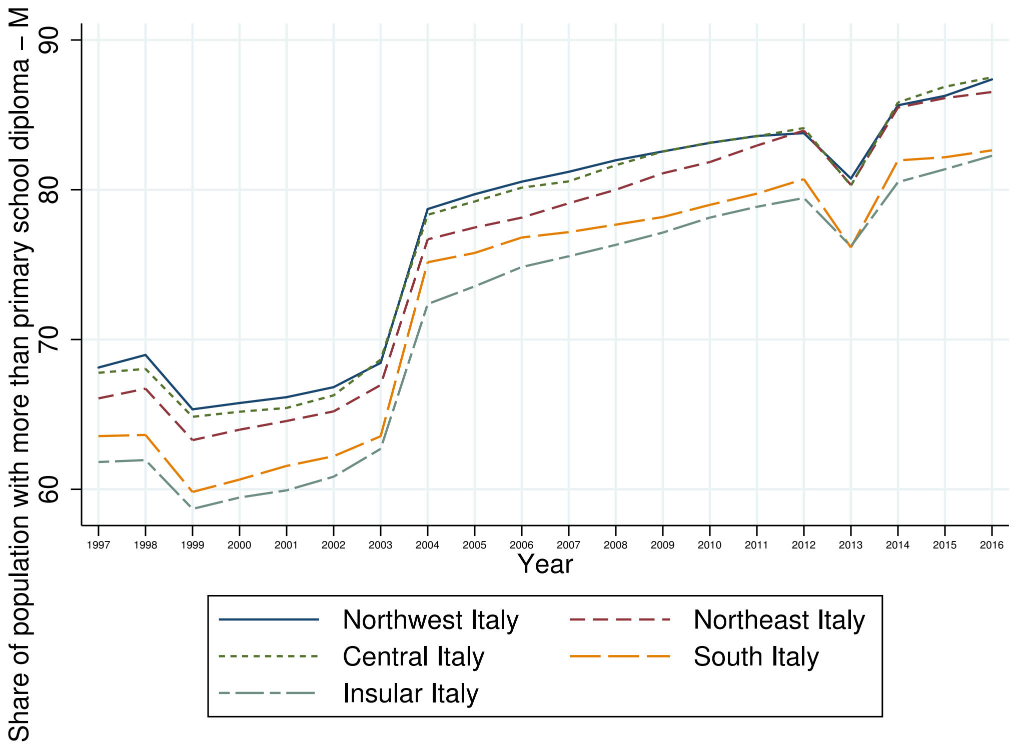

Figure 1.8: Share of male population with more than primary school diploma. Source: Author graph based on ISTAT data.

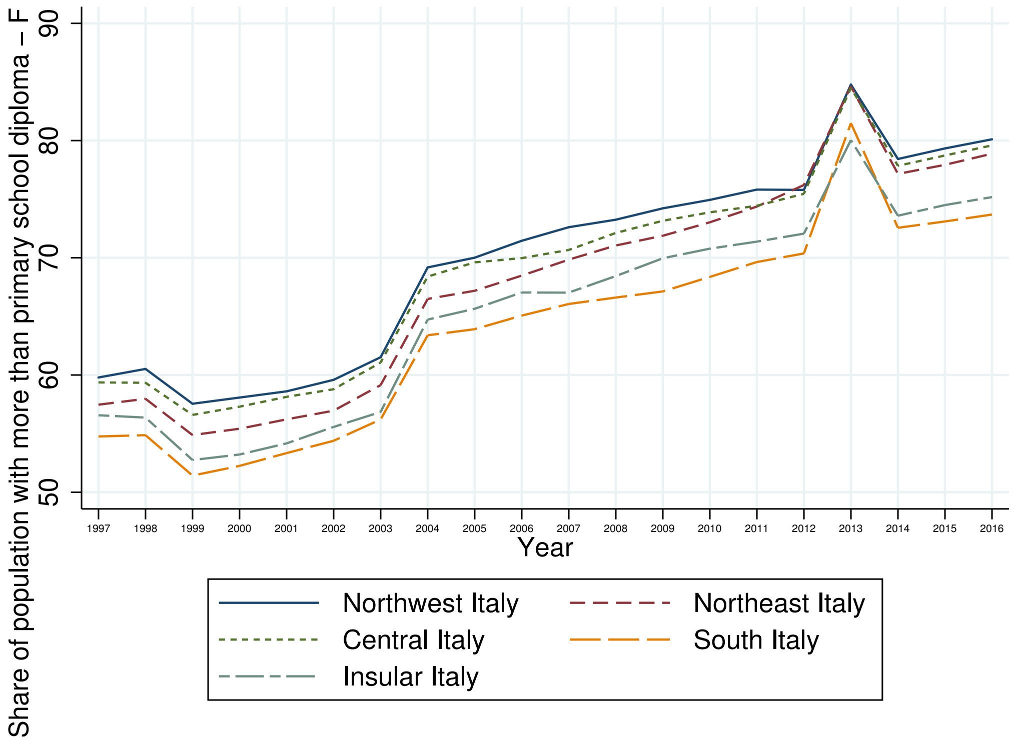

Figure 1.9: Share of female population with more than primary school diploma. Source: Author graph based on ISTAT data.

attainment, with roughly 5% less educated people compared the other regions. Interestingly, women appear as the worse-off in terms of level of education, with a gap compared to men that can reach almost 10% of difference.

Figures 1.10 and 1.11 show a similar scenario: the level of graduated people raised over time, although with a big variability among macroareas. In this case, the Central Italy experiences the best performance during the whole period of observation and women register the highest scores. The North of Italy immediately follows and the South and Island register, as usual, the lowest performance.

Table 1.2 reports the regional descriptive statistics for the abovementioned GDP and educational attainment levels, which confirm and details numerically what just described.

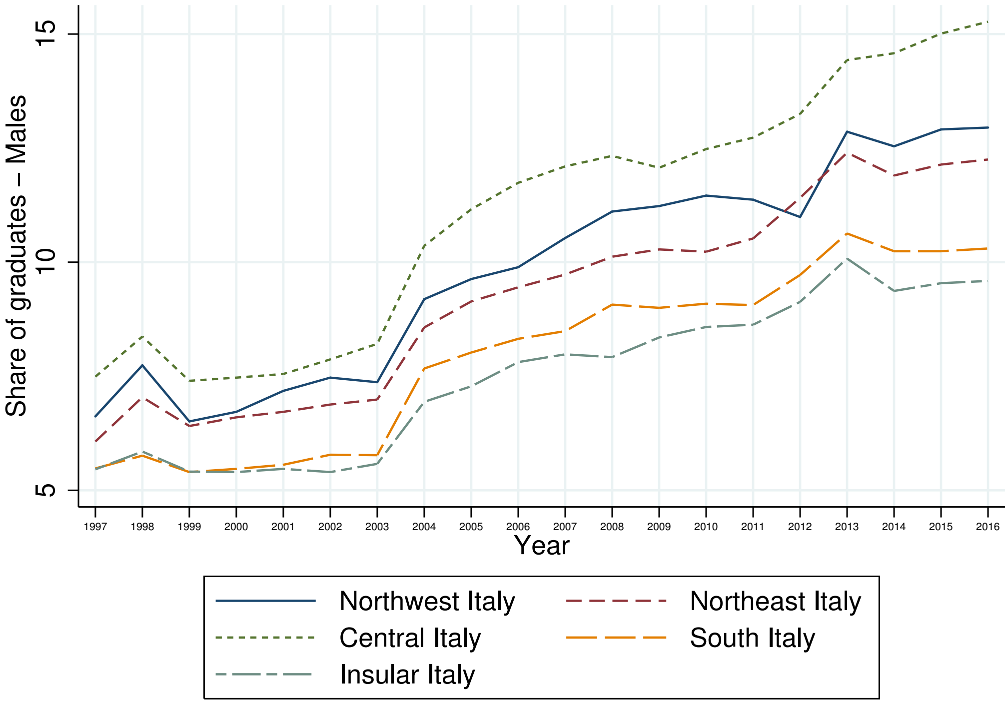

Figure 1.10: Share of male graduates. Source: authors graph based on ISTAT data.

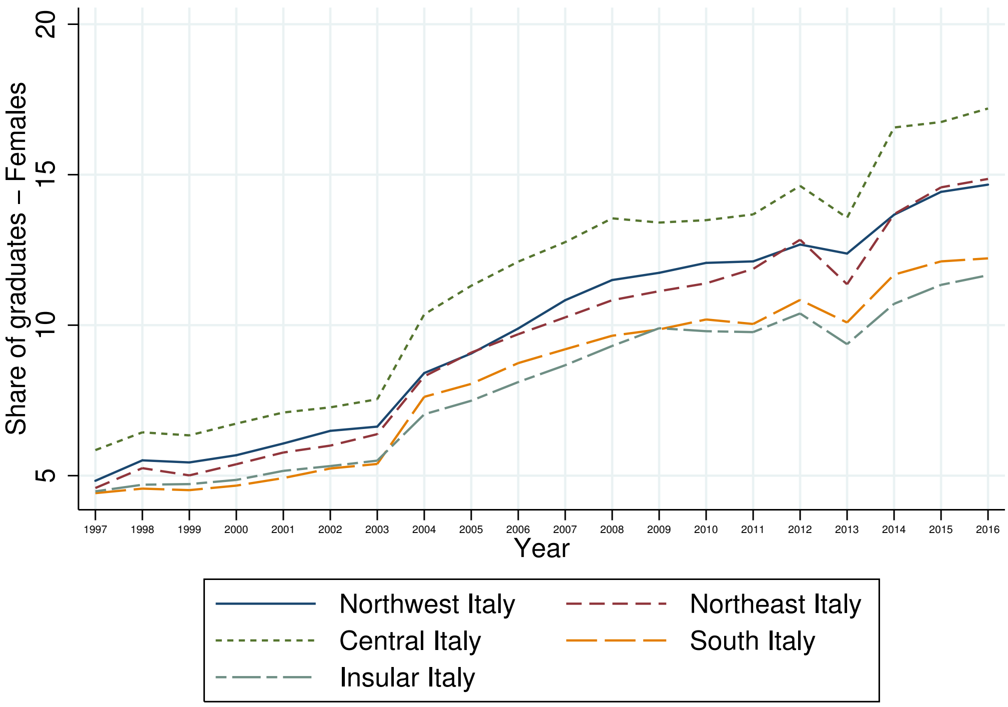

Figure 1.11: Share of female graduates. Source: authors graph based on ISTAT data.

Table 1.2: GDP and education - summary statistics by regions

Variable	Mean	Std. Dev.	Min.	Max.	Growth rate	N
North-East						
GDP	31400.627	1231.967	29361.119	33589.613	-7.93	22
% pop no education or primary M	24.175	8.418	13.47	36.72	-1.33***	20
% pop no education or primary F	32.241	9.237	15.38	45.12	-1.48***	20
% pop graduated M	9.242	2.192	6.07	12.4	0.36***	20
% pop graduated F	9.414	3.384	4.59	14.86	0.56***	20
North-West						
GDP	32648.618	1323.531	30466.08	34607.648	3.5	22
%pop no education or primary M	22.759	7.961	12.63	34.66	-1.23***	20
% pop no education or primary F	30.221	8.611	15.22	42.45	-1.38***	20
% pop graduated M	9.813	2.302	6.51	12.95	0.37***	20
% pop graduated F	9.705	3.322	4.83	14.67	0.55***	20
Center						
GDP	29885.543	1646.106	27678.031	32547.123	-25.00	22
% pop no education or primary M	23.002	8.212	12.49	35.16	-1.28***	20
% pop no education or primary F	31.051	8.625	15.51	43.4	-1.38***	20
% pop graduated M	11.093	2.793	7.4	15.27	0.46***	20
% pop graduated F	11.333	3.832	5.850	17.2	0.63***	20
South						
GDP	18053.698	902.311	16785.529	19415.785	-22.6	22
% pop no education or primary M	27.096	8.4	17.37	40.18	-1.28***	20
% pop no education or primary F	36.069	8.588	18.51	48.58	-1.36***	20
% pop graduated M	7.954	1.926	5.4	10.63	0.31***	20
% pop graduated F	8.201	2.8	4.42	12.22	0.46***	20
Islands						
GDP	18055.358	939.366	16610.656	19488.119	-27.4	22
% pop no education or primary M	28.401	8.538	17.73	41.31	-1.34***	20
% pop no education or primary F	34.706	8.486	19.96	47.25	-1.35***	20
% pop graduated M	7.489	1.674	5.4	10.08	0.27***	20
% pop graduated F	7.915	2.493	4.48	11.66	0.41***	20

Source: Author elaboration on ISTAT-Health For All data. Growth rates for each of the 25 series are calculated using a regression of the respective variable with the time trend as the only exogenous variable. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4 Final Remarks and Discussion

In this introductory chapter, we aimed to explore the new dynamics of life expectancy, GDP and education attainment in Italy, concentrating on how they are distributed among the NUTS 1 macroregions during the period 1990-2016.

We started by reviewing the existing literature on how health is produced at the individual level and what makes us healthier, with a specific focus on the classical determinants of health: income and education.

We can draw mainly three findings from our analysis: (1) health outcomes, measured by life expectancy at birth and at 65 years old, are registering an increasing improvement in the last 27 years, in any NUTS 1 macroregion, and both for men and women. Consequently, there is no evidence to date of a potential ceiling in life expectancy that makes the average number of years a person can expect to live stop rising. (2) Although health care and education are universally available in Italy, the longstanding dichotomy between the North and the South is still reflected in the levels of life expectancy, education attainment, and GDP, with the *Mezzogiorno* always being the worst-performing. The existence of "two Italies" is thus confirmed. (3) Women still live longer that men and, although they graduate more, the share of low educated females exceeds the same share of men.

A clear understanding of the association between health and socioeconomic conditions, along with their spatial distribution, is a starting point for a policy maker in order to identify the weaknesses to be addressed with policies oriented to shrink inequalities in a society. Inequalities that may actually be exacerbated if spatial disparities in health outcomes are not properly taken into account when age thresholds that drive the eligibility for social benefits are established. In fact, if people live longer, governments may be interested in changing the age of eligibility for social security on lockstep with life expectancy gains, setting up new threshold at a national level. But, if the quantity of life is not equally distributed throughout the country and depends on the place of residence, as in the Italian case, then only people living in certain places would get significant improvements in life expectancy. As a result, when a raise in the age of eligibility is set globally, without taking into consideration those discrepancies, more inequality may be created. Indeed, in some specific regions people die relative earlier than in other regions but share the same costs for social security system without benefiting for the same amount of time.

Bibliography

[1] Acemoglu, D., & Johnson, S. (2007). Disease and development: the effect of life expectancy on economic growth. Journal of political Economy, 115(6), 925-985.

[2] Adams, P., Hurd, M. D., McFadden, D., Merrill, A., & Ribeiro, T. (2003). Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status. Journal of Econometrics, 112(1), 3-56.

[3] Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Syme, S. L. (1994). Socioeconomic status and health: the challenge of the gradient. American psychologist, 49(1), 15.

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

[5] Berkman, L. F. (1995). The role of social relations in health promotion. Psychosomatic medicine, 57(3), 245-254.

[6] Bloom, D. E., & Canning, D. (2005). Schooling, health and economic growth: reconciling the micro and macro evidence. unpublished, Harvard School of Public Health (February 2005).

[7] Bloom, D. E., Canning, D., & Fink, G. (2014). Disease and development revisited. Journal of Political Economy, 122(6), 1355-1366.

[8] Bongaarts, J. (2006). How long will we live?. Population and Development Review, 32(4), 605-628.

[9] Booth, H., Hyndman, R. J., Tickle, L., & De Jong, P. (2006). Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions. Demographic Research, 15, 289-310.

[10] Booth, H., Maindonald, J. H., & Smith, L. (2001). Age-time interactions in mortality projection: Applying Lee-Carter to Australia.

[11] Booth, H., & Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. Annals of actuarial science, 3(1-2), 3-43.

[12] Boyce, C. J., & Oswald, A. J. (2012). Do people become healthier after being promoted?. Health economics, 21(5), 580-596.

[13] Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

[14] Case, A., & Deaton, A. (2015). Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. Proceedings of the National Academy of Sciences, 112(49), 15078-15083.

[15] Case, A., Fertig, A., & Paxson, C. (2005). The lasting impact of childhood health and circumstance. Journal of health economics, 24(2), 365-389.

[16] Chetty, R., Stepner, M., & Cutler, D. (2016). Relationships Between Income, Health Behaviors, and Life Expectancy—reply. Jama, 316(8), 880-881.

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

[18] Currie, J., & Madrian, B. C. (1999). Health, health insurance and the labor market. Handbook of labor economics, 3, 3309-3416.

[19] Cutler, D., Deaton, A., & Lleras-Muney, A. (2006). The determinants of mortality. Journal of economic perspectives, 20(3), 97-120.

[20] Cutler, D. M., & Lleras-Muney, A. (2006). Education and health: evaluating theories and evidence (No. w12352). National bureau of economic research.

[21] Dalgaard, C. J., & Strulik, H. (2014). Optimal aging and death: understanding the Preston curve. Journal of the European Economic Association, 12(3), 672-701.

[22] Danesi, I. L., Haberman, S. & Millossovich, P. (2015). Forecasting mortality in subpopulations using Lee-Carter type models: A comparison. Insurance: Mathematics and Economics, 62, pp. 151-161. doi: 10.1016/j.insmatheco.2015.03.010

[23] Davis, J. A. (2015). A Tale of Two Italys? The "Southern Question" Past and Present. In The Oxford Handbook of Italian Politics.

[24] De Vogli, R., Mistry, R., Gnesotto, R., & Cornia, G. A. (2005). Has the relation between income inequality and life expectancy disappeared? Evidence from Italy and top industrialised countries. Journal of Epidemiology & Community Health, 59(2), 158-162.

[25] Denton, F. T., Feaver, C. H., & Spencer, B. G. (2005). Time series analysis and stochastic forecasting: An econometric study of mortality and life expectancy. Journal of Population Economics, 18(2), 203-227.

[26] Ettner, S. L. (1996). New evidence on the relationship between income and health. Journal of health economics, 15(1), 67-85.

[27] Evans, W., & Graham, J. D. (1988). Traffic safety and the business cycle. Alcohol, drugs, and driving, 4(1), 31-38.

[28] Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189-1232.

[29] Frijters, P., Haisken-DeNew, J. P., & Shields, M. A. (2005). The causal effect of income on health: Evidence from German reunification. Journal of health economics, 24(5), 997-1017.

[30] Grossman, M. (1972). On the concept of health capital and the demand for health. Journal of Political economy, 80(2), 223-255.

[31] Haberman, S. & Russolillo, M. (2005). Lee Carter mortality forecasting: application to the Italian population (Report No. Actuarial Research Paper No. 167). London, UK: Faculty of Actuarial Science & Insurance, City University London.

[32] Hartog, J., & Oosterbeek, H. (1998). Health, wealth and happiness: Why pursue a higher education?. Economics of Education Review, 17(3), 245-256.

[33] Hokayem, C., & Ziliak, J. P. (2014). Health, human capital, and life cycle labor supply. American Economic Review, 104(5), 127-31.

[34] Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org (data downloaded on August 2018).

[35] INAIL - National Institute for Insurance against Occupational Accidents https://www.inail.it/cs/internet/attivita/dati-e-statistiche/statistiche-storiche/casi-denunciati.html.

[36] ISTAT - Health for All, version June 2018, https://www.istat.it/it/archivio/14562. Accessed August 2018.

[37] ISTAT - Statistiche Focus, 15 gennaio 2014. La mortalità dei bambini ieri e oggi in Italia. https://www.istat.it/it/files/2014/01/Mortalita_-sotto_i_5_anni-.pdf.

[38] Janssen, F., & Kunst, A. (2007). The choice among past trends as a basis for the prediction of future trends in old-age mortality. Population Studies, 61(3), 315-326.

[39] Janssen, F., Kunst, A., & Mackenbach, J. (2007). Variations in the pace of old-age mortality decline in seven European countries, 1950–1999: The role of smoking and other factors earlier in life. European Journal of Population/Revue europeenne de demographie, 23(2), 171-188.

[40] Janssen, F., van Wissen, L. J., & Kunst, A. E. (2013). Including the smoking epidemic in internationally coherent mortality projections. Demography, 50(4), 1341-1362.

[41] Kemptner, D., Jürges, H., & Reinhold, S. (2011). Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany. Journal of Health Economics, 30(2), 340-354.

[42] King, G., & Soneji, S. (2011). The future of death in America. Demographic research, 25, 1.

[43] Lee, R. D., & Carter, L. R. (1992). Modeling and forecasting US mortality. Journal of the American statistical association, 87(419), 659-671.

[44] Lynch, J. W., Smith, G. D., Kaplan, G. A., & House, J. S. (2000). Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions. Bmj, 320(7243), 1200-1204.

[45] Meer, J., Miller, D. L., & Rosen, H. S. (2003). Exploring the health–wealth nexus. Journal of health economics, 22(5), 713-730.

[46] Mitnitski, A. B., Mogilner, A. J., MacKnight, C., & Rockwood, K. (2002). The accumulation of deficits with age and possible invariants of aging. The Scientific World Journal, 2, 1816-1822.

[47] Murphy, K. M., & Topel, R. H. (2006). The value of health and longevity. Journal of political Economy, 114(5), 871-904.

[48] Oeppen, J., & Vaupel, J. W. (2002). Broken limits to life expectancy.

[49] Pampel, F. C. (2005). Diffusion, cohort change, and social patterns of smoking. Social science research, 34(1), 117-139.

[50] Pollard, J. H. (1987). Projection of age-specific mortality rates. Population Bulletin of the United Nations, (21/22), 55-69.

[51] Preston, S. H. (1975). The changing relation between mortality and level of economic development. Population studies, 29(2), 231-248.

[52] Renshaw, A. E., & Haberman, S. (2006). A cohort-based extension to the Lee–Carter model for mortality reduction factors. Insurance: Mathematics and economics, 38(3), 556-570.

[53] Ruhm, C. J. (1995). Economic conditions and alcohol problems. Journal of health economics, 14(5), 583-603.

[54] Ruhm, C. J. (2000). Are recessions good for your health?. The Quarterly journal of economics, 115(2), 617-650.

[55] Schultz, T. P. (2004). School subsidies for the poor: evaluating the Mexican Progresa poverty program. Journal of development Economics, 74(1), 199-250.

[56] Smith, J. P. (1999). Healthy bodies and thick wallets: the dual relation between health and economic status. Journal of Economic perspectives, 13(2), 145-166.

[57] Stewart, S. T., Cutler, D. M., & Rosen, A. B. (2009). Forecasting the effects of obesity and smoking on US life expectancy. New England Journal of Medicine, 361(23), 2252-2260.

[58] Stoeldraijer, L., van Duin, C., van Wissen, L. J. G., & Janssen, F. (2013). Impact of different mortality forecasting methods and explicit assumptions on projected future life expectancy: The case of the Netherlands. Demographic Research, 29, 323-354. [13].

[59] Tabeau, E. (2001). A review of demographic forecasting models for mortality. In Forecasting Mortality in Developed Countries (pp. 1-32). Springer, Dordrecht.

[60] Wang, H., & Preston, S. H. (2009). Forecasting United States mortality using cohort smoking histories. Proceedings of the National Academy of Sciences, 106(2), 393-398.

[61] Well, D. N. (2007). Accounting for the effect of health on economic growth. The quarterly journal of economics, 122(3), 1265-1306.

[62] Wong-Fupuy, C., & Haberman, S. (2004). Projecting mortality trends: recent developments in the United Kingdom and the United States. North American Actuarial Journal, 8(2), 56-83.

[63] Zimmerman, E. B., Woolf, S. H., & Haley, A. (2015). Understanding the relationship between education and health: a review of the evidence and an examination of community perspectives. Population health: behavioral and social science insights. Rockville (MD): Agency for Health-care Research and Quality, 347-84.

Chapter 2

One Plus One Makes Less than Two? Consolidation Process in Italian Local Health Authorities and Population Mortality

Abstract

This chapter evaluates the impact of Local Health Authorities consolidation process on population mortality within Italian municipalities from 2002 to 2016. We apply a quasi-experimental research design using an event study framework that enables us to test for potential pre-existing trends and anticipation effects, as well as phase-in policy impact. Since each region set up the reform in different years, we use a staggered two-way fixed effect Difference in Differences method as identification strategy. We distinguish between mountain and non-mountain municipalities to detect possible heterogeneous effects. Our findings show that population mortality has increased in the years following the policy change, with the mountain municipalities being the most affected, and suggest that policy makers should carefully weight the potential efficiency gains with the negative impact on health care effectiveness, specifically in local communities, before allowing more mergers.

1 Introduction

While the provision of public services has been increasingly devoted to decentralization, following the spreading of the subsidiarity principle, in health care management centralization seems to be the new rage. A merging process in health care started in the 1990s in the United States and the United Kingdom, leading to a radical change in the structures of their Health Systems (Gaynor *et al.* 2012, Goddard and Ferguson 1997). As pointed out by Schmid and Varkevisser (2016), other countries, as Germany, the Netherlands, and Canada, are moving in the same direction, such that this new rage can be defined as "merger mania". Italy is not an exception since mergers in health structures are still an ongoing process. Indeed, over the last decades, in Italy, the increase of life expectancy and the consequent population aging raised concerns on the dynamics of health spending in such a way that institutional settings have changed and policies mostly oriented at containing costs and increasing efficiency have been implemented. The consolidation of Italian Local Health Authorities (LHAs henceforth) represents a prominent example of this pattern and is the focus of this study. The Italian LHAs experienced a drastic reduction of their number, passing from 659 in 1992 to 101 in 2017 (Italian Ministry of Health, 2018).

This chapter seeks to answer the question: does the amalgamation of two or more LHAs affect the quality of health care?

The question of whether Italian LHAs amalgamation programs affect the quality of health care for patients has not been answered yet and would represent a further step in the international debate regarding the effects of mergers in the provision of public services. Thus, the motivation of this analysis relies on giving some evidence on the unintended effects that this "merging mania" may produce on the quality outcomes. In fact, the more straightforward direct effects are related to cost-saving benefits by cutting off the number of General Managers and administrative staff. Given this strength, from a theoretical perspective, the maximum benefit would be obtained by establishing a single LHA per region or a single LHA for the whole country. Nevertheless, this merging design may have some weaknesses and produce, eventually, a worsening of health care services. As a matter of fact, a limit might exist to the complexity that a single management body can reasonably supervise.

Using a staggered two-way fixed effect difference in differences approach, we are able to deal with the variation occurring from different timing of amalgamation reforms across the Italian regions. Moreover, we are able to control for potential anticipation effects and to disentangle the full dynamics of the treatment effect in the short and in the medium-long run.

Heterogeneous effects are also possible with respect to the municipality geographical characteristics since, as suggested by Moore and Carpenter (1999), the patterns of morbidity and mortality depend on geographical and environmental factors. As a result, in a "one size fits all" health care provision, the specific needs of some local communities could no longer be taken into proper account.

Our findings show that population mortality increased in the years following the reform, especially in mountain municipalities, reasonably the most isolated ones, and suggest that policy makers should carefully consider the impact on health care effectiveness and on accessibility to care before allowing more mergers.

This chapter is organized as follows. Section 2 gives the background of the amalgamation process in Italy in our observed period. Section 3 provides the literature framework to which this analysis belongs. In section 4 we discuss the data. Section 5 the methodology. Section 6 presents the results, section 7 the robustness checks, section 8 concludes.

2 Institutional Changes: the Consolidation Process

The Italian National Health Service (NHS henceforth) was established in 1978 with the objective of providing uniform and adequate health care throughout the country. Ever since, health care expenditure increased systematically over the years, while life expectancy considerably grew at the national level, as depicted in figure 2.1. The increase of life expectancy and the consequent aging of population, shown in figure 2.2, raised concerns on the dynamics of health spending.

In order to control such expenditure growth, various subsequent Italian Governments introduced policy reforms such that the institutional features of health care providers remarkably changed over time. With a constitutional reform in 2001, a major process of decentralization invested the NHS structures, shifting responsibilities from central state to regions, which exert their role in providing public health care through the LHAs. LHAs are public bodies responsible for organizing, coordinating and providing at least the Essential Levels of Care (*Livelli Essenziali*

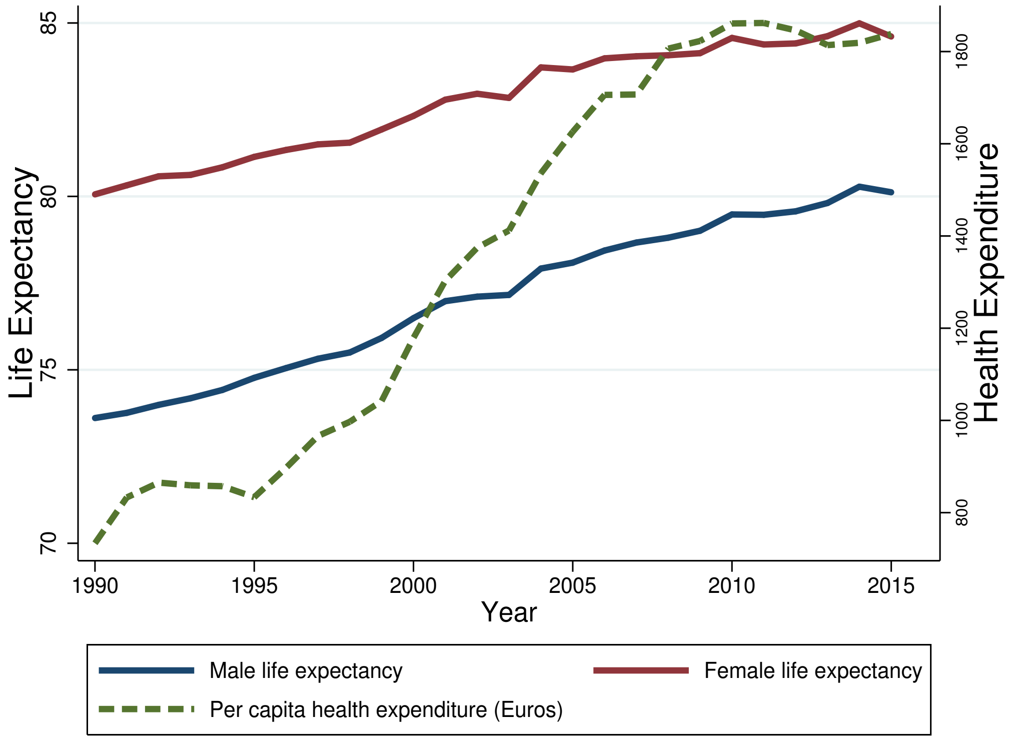

Figure 2.1: Life expectancy by gender and health expenditure in Italy from 1990.
Source: Author graph based on ISTAT data.

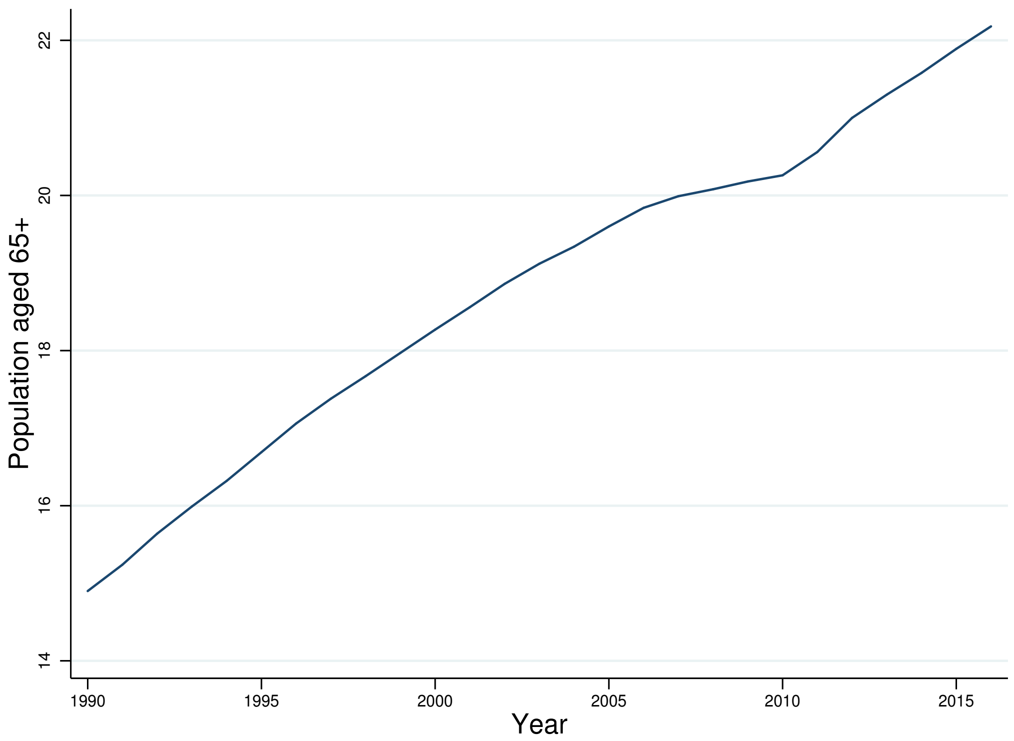

Figure 2.2: Share of population aged 65 and older in Italy from 1990.
Source: Author graph based on ISTAT data.

di Assistenza, LEA) in specific geographical areas within the regions.

The Italian LHAs experienced a drastic reduction of their number, passing from 659 in 1992 to 197 in 2001 (Di Novi *et al.*, 2018), resulting in 139 in 2015 and 101 in 2017 (Ministry of Health, 2018). They decreased by almost 85% from 1992 to 2017. As a trivial result, the declining number of LHAs over time provoked an increase in the average population size under the responsibility of one single LHA.

The merging of LHAs seems to belong to policies mostly oriented to contain costs and increase efficiency. Indeed, several studies have documented the presence of regional differentials in Italian public spending in health sector (Francese and Romanelli, 2014). Starting from 2007, the Central Government, in order to avoid financial failure in some regions where emerged significant management deficits in health expenditure, introduced a special regime called "Recovery Plan"(*Piani di Rientro*), aimed at restoring the economic-financial equilibrium of the regions concerned. The consolidation process seems to be nested in this cost saving-oriented policies: in the last wave of merging, from 2001 to 2015, 32 out of 58 LHAs merged are located in regions under Recovery Plan (Di Novi *et al.*, 2018). However, it is worth noting that the merging process does not regard only LHAs with deficit problems but appears a more comprehensive approach involving even the most diligent regions, including Emilia Romagna, Tuscany and Lombardy, such that gained the definition of "merging mania".

Figure 2.3 depicts the abovementioned Italian LHAs amalgamation process, showing its variation in time, and implicitly in space. The number of LHAs is diminishing over years because regions decide to consolidate their LHAs in different periods.

Figure 2.4 provides a geographical overview in the consolidation variation by Italian macroareas (NUTS 1) and figure 2.5 details the mergers at the regional level. These latter highlight that the consolidation process is a widespread phenomenon throughout the country, which has affected all regions, with the trivial exception of Valle d'Aosta, where the LHA has always been unique. Table 2.10 in appendix provides details on how LHAs have been involved in the mergers.

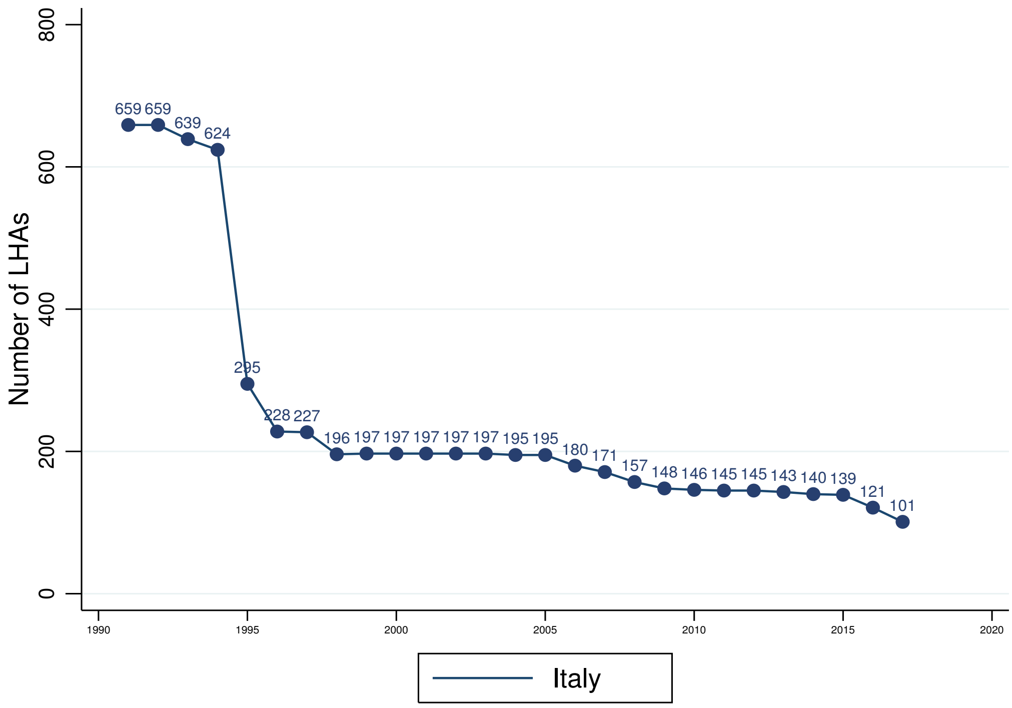

Figure 2.3: Number of LHAs in Italy from 1990 to 2017.
Source: Author elaboration on ISTAT- Health for all Data

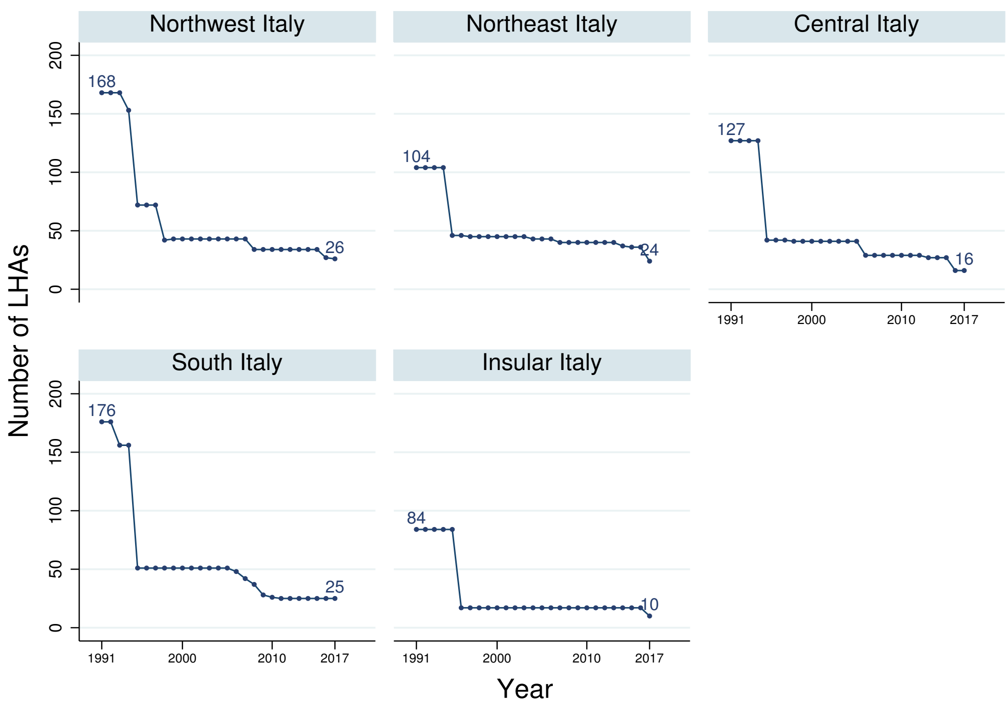

Figure 2.4: Number of LHAs in Italy from 1990 to 2017 by macroareas
Source: Author elaboration on ISTAT- Health for all Data

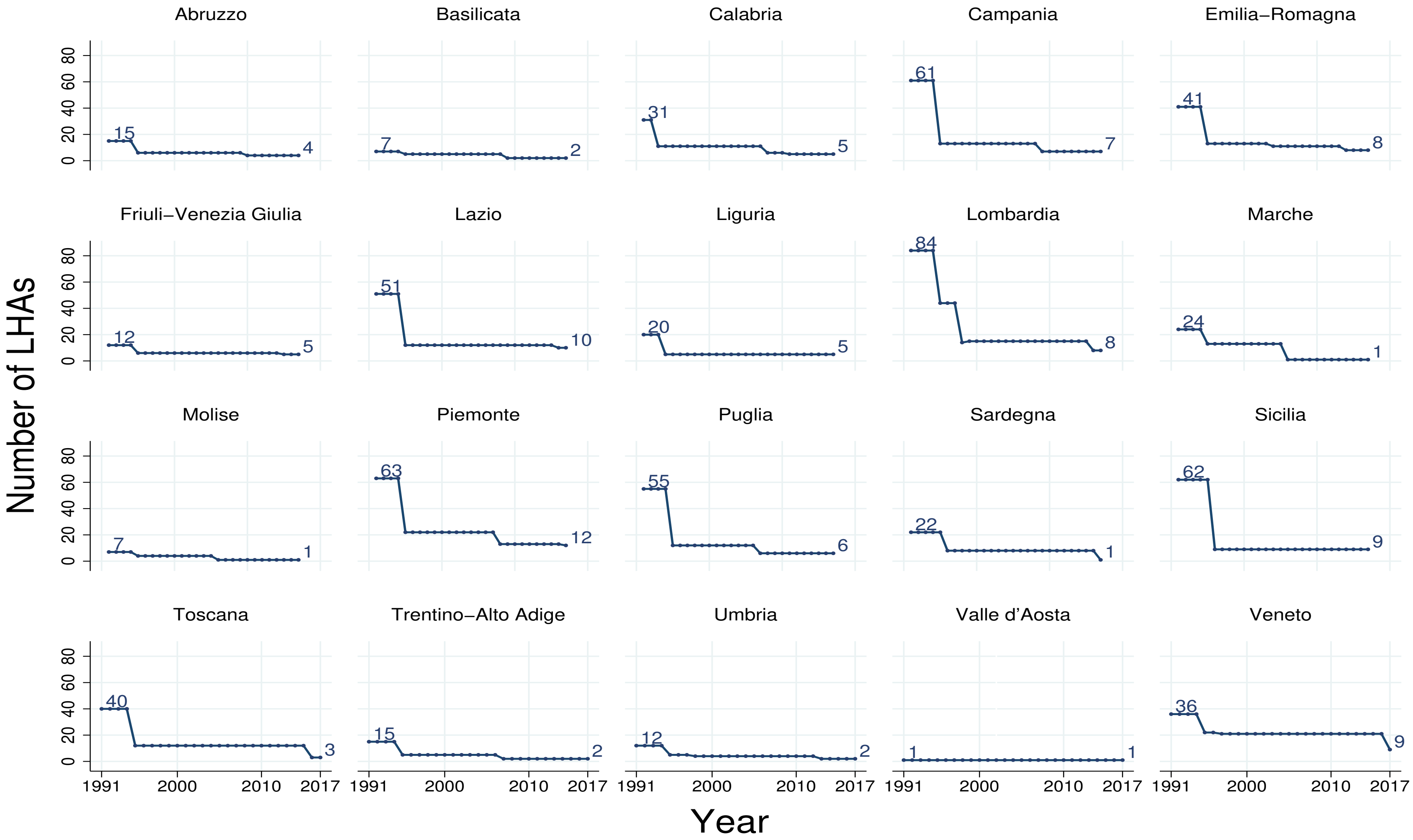

Figure 2.5: Number of LHAs in Italy from 1990 to 2017 by regions. Source: Author graph on ISTAT- Health for All data

3 Literature Review

This chapter relates to two different strands of the literature that, in this case, build ties one to another.

First, a branch related to the impact of health department and hospital mergers on the potential cost-minimization effect - by means of economies of scale and scope - and on quality of care - through knowledge spillover arising from the increased volume of treated cases. Second, the piece of the literature linked to the centralized versus decentralized provision of local public goods debate.

A large body of empirical literature documents the existence of a positive correlation between the volume of cases treated by health care providers and better health outcomes achieved by patients. Mortality appears lower in those hospitals that perform more times a given procedure. Although the consolidation of LHAs does not imply the merger of their controlled hospitals, a bigger LHA may benefit from the increased volume of cases under its responsibility that a consolidation causes. The leading explanation for this positive correlation is the "practice makes perfect" effect. More specifically, larger health care providers may be able to furnish better quality treatments and then to achieve better outcomes for their patients through learning-by-doing or quality enhancing scale economies (Luft *et al.*, 1979 and 1987, Gaynor *et al.*, 2005, Cavalieri *et al.*, 2013). This direction of causality from volume to better outcomes matters for policy, supporting the idea of centralization of procedures in a few hospitals and, eventually, the mergers policies. What is not trivial in this picture is to shed light on the temporal pattern that characterizes the positive correlation between volume and outcomes. In particular, Gaynor *et al.* (2005), examined the relationship between volume and outcome for a type of heart surgery, the coronary artery bypass graft (CABG), questioning whether the effect is only contemporaneous and due to static-scale-economies or could be inlaid in the framework of learning models, namely via lags with a more complicated dynamics. Their probit estimates suggest that the volume-outcome effect is mainly contemporaneous albeit there is not sufficient variation in their data to robustly identify the lags dynamics.

Ho and Hamilton (2000) study the impact of hospital mergers and acquisitions on the quality of care in California during the 1990s. They compare patients' outcomes - mortality for heart attack and stroke, 90-day readmission for heart attack and discharge within 48 hours for normal newborn babies - before and after merger and acquisitions that took place between 1992 and 1995. Their results indicate that mergers and acquisitions have a positive but not significant effect on inpatient mortality (although the standard errors are large due to a small sample size and, therefore, they cannot rule out the possibility that the real effect may be smaller than the standard error associated to the estimates). In addition, they find mergers and acquisitions to be associated with a higher readmission rates for heart attack and early discharge of normal newborns.

Our analysis is also akin under certain aspects to the study made by Gaynor *et al.* (2012). They contributed to the recent literature on the merging mania of public service providers examining the impact of a wave of general hospitals mergers in England between 1997 and 2006 on a large list of hospital performance outcomes including clinical quality, productivity, waiting times and financial performance. Since the policy shift occurred in different years across hospitals, they used an extended Difference in Difference approach, also called "event study", where the treatment date is the date of merger and the control group is selected through a matching procedure among those hospitals that never merged that are similar in terms of observable characteristics. They studied the impact on abovementioned outcomes two years prior to the year of the merging approval, controlling also for potential anticipation effects, and four years after, in order to disentangle the full dynamics of the impact both in the immediate post reform period and in the medium run. Their results are not reassuring: although hospital admissions and staff size fell after the merger, quality measures remained unchanged, productivity did not increase while financial deficits and waiting times did.

Hoornbeek *et al.* (2015) specifically focused on the impact of 20 local health departments (LHDs) consolidations on the total and administrative expenditures in Ohio from 2001 to 2011. They found that consolidating LHDs yielded financial benefits. Indeed, consolidated LHDs experienced a reduction in total expenditures of roughly 16% and no statistically significant change in administrative expenses. Interviewed senior local health officials also asserted that consolidations caused public health service improvements.

To the best of our knowledge, Di Novi, Rizzi and Zanette (2018) is the only study that focused on the Italian LHAs consolidation and studied the potential cost savings arising from the LHAs reconfigurations. They found the presence of economies of scale linked to the size of the LHAs population. Indeed, their estimation highlights an elasticity of administrative per capita costs with respect to population size of -2.50, that becomes -0.112 and -0.127 for goods and non-health-service costs, respectively. Hence, *ceteris paribus*, the smaller the LHA, the higher the per capita non-health-service costs.

Their analysis focused only on administrative costs, cost of goods and non-health-service costs, namely costs not directly involved in the provision of health care, under the assumption that the quantity and the quality of health care provided remained constant. Our study instead relax this assumption and aims at estimating the effect of mergers in terms of health care quality outcome. Indeed, as Di Novi *et al.* themselves suggested, savings may be an important driver in limiting growing cost of health care and may be used to improve the quality of treatments, for instance, in terms of making more R&D investments.

Furthermore, the classic problem in public finance regarding the trade-off between centralized and decentralized provision of local public goods have been of interest to public economists (see e.g., Boffa *et al.* 2015, Besley and Coate 2003, Enikolopov and Zhuravskaya 2007, Rodríguez-Pose and Ezcurra 2009).

The literature on decentralization is wide and mostly related to fiscal federalism. However, we believe the principles may be applied in the case object of this study, focused on the effect of a lesser decentralization in the provision of health care.

The standard approach, formalized by Oates (1972), assumes that the main drawback with a centralized system is that it produces a "one size fits all" provision of public goods which does not reflect, and eventually satisfy, local needs. Therefore, decentralization is more efficient than centralization when local areas are heterogeneous and there are no policy spillovers. Indeed, decentralization is believed to produce a large range of benefits. Bardhan (2002) reviews theoretically the main differences between the two governmental configurations. Fragmenting central authority and introducing intergovernmental branches is viewed as a way of reducing the role of central state and to make government more responsive and efficient. While centralization can exploit economies of scale and scope better and profit from policy coordinations - which becomes very important in presence of spillovers - local bureaucrats are believed to have more information on the local needs than the upper-tier governments and many of the public goods needed by citizens are often community- and site-specific. The crucial difference regards the political accountability. In case of centralization, the number of principals is very large while the number of agents is very little. In case of decentralization, the number of agents is larger

and the ratio principal-agent decreases. The higher the number of agents, the more straightforward the identification of who bears responsibilities for the welfare of a given locality. Moreover the intergovernmental competition stemming from decentralization may be a powerful tool to increase economic efficiency and to induce citizens to reveal their preferences, as first proposed by the well-known Tiebout model (1956). However, when the technical and administrative capacity is not homogeneous among local bureaucracies, an unequal quantity and quality of public goods provision may result, specifically in those services that require sophisticated expertise, like the case of health care management. As Bird (1994) states,"the central government may not know what to do; the local government may not know how to do it".
That is, the best configuration in the provision of health care, namely a public service that requires sophisticated expertise and may be community- and site-specific, should be determined empirically.

4 Data

This analysis draws data from different sources. The major sources are the Italian National Statistical Institute (ISTAT) for the demographic and mortality data, the Ministry of Health (MH) for LHAs information, and the Ministry of Economy and Finance (MEF) for income data.

4.1 Dependent Variable and Treatment Indicator

We collected data on population mortality for any Italian municipality from 2002 to 2016. The sources are the demographic balance of the resident population in the "Geo-Demo" section of ISTAT, which provides the results of the annual survey "Movement and calculation of the resident population" (ISTAT P.2 model) that ISTAT performs every year at the registry offices of the Italian municipalities.[1] Thus, our dependent variable y_{it} is the mortality rate of all people

[1]Following the consolidation rage, some municipalities experienced a merger to other municipalities in terms of administrative borders. Thus, we dropped from the sample 292 municipalities that experienced a merger or an acquisition, since their essential characteristics are not stable in our observation period and they may not be randomly selected into merger. We also rule out the municipality of Rome because it comprises more than one LHAs that share different treatment regimes, as well as LHA 205 in region Calabria because it was involved in a consolidation twice in the observed period. Similarly, we do not consider region Piedmont because the consolidation policy occurred both in 2006 and in 2017. Due to the lack of up-to-date data, we are not able to distinguish the treated LHAs in the last wave of consolidation.

residing in the municipality i at time t. Note that the health outcome of interest is the mortality rate of the population as a whole, not just those admitted to an hospital that received health care goods. We analyze this more comprehensive dimension of health outcome because the responsibility of a LHA is to guarantee public health with both primary and secondary care. Moreover, potential endogeneity problems are posed by the possibility that hospital choices may be driven by the perceived quality of care among patients. Indeed, the Italian Constitution establishes that patients are free to seek health care in any LHA throughout the country. As a result, patients may choose the better perceived hospitals, especially in the most severe clinical cases. Therefore, the potential worse off patient outcomes registered in the selected hospitals may depend on an increased complexity of treated cases rather than on a lower quality of care. Further, we accurately reconstructed the features of the public policy switch. LHAs reconfiguration was recorded by means of the yearly-based databases released by the Ministry of Health including the comprehensive list of the whole population of Italian hospitals and the LHAs to which they belong, and the LHAs-municipalities correspondence. Following the correspondence among any single hospital and its LHA over the years, we were able to precisely identify the LHAs mergers since our data-driven method showed a perfect aggregation of all hospitals belonging to separated LHAs to only one LHA. In other words, the available administrative data assure the absence of a chaotic amalgamation process where, for instance, only a share of hospitals belonging to a LHA ended up under the responsibility of the bigger one resulting after the merger. The consolidation was instead much simpler and could be summed up as the aggregation of all the hospital belonging to two or more LHAs into a bigger one, as table 2.10 details in the appendix. Once we gained a clear picture of the consolidation process, distinguishing between treated and untreated municipalities, namely municipalities belonging or not belonging to a merged LHA, was straightforward given the correspondence LHAs-Municipalities issued by the Ministry of Health.

4.2 Control Variables

To control for confounder factors which may be correlated to mortality rate, we collect data on population size, population aged 55 and older, per capita income, regions under Recovery Plan

(RP) and natural disasters.

Population size may be a key component of health needs in an area (Leone, 2010). Population density is also significantly positively associated with trends in life expectancy. Indeed, the most densely populated municipalities may have public policies that improve health (e.g., smoking bans) or greater funding for public services (Chetty *et al.*, 2016).

Population aged 55 and older aims at controlling for the frailer share of people living in a municipality. Although the aging process is only imperfectly captured by chronological age, it may be defined as "the intrinsic, cumulative, progressive, and deleterious loss of function that eventually culminates in death" (Dalgaard and Strulik, 2014 p.675). Following a similar approach as in Lichtenberg and Waldfogel (2009), we adopt the 55 age cutoff to control for the share of population that reasonably faces a higher risk of poor health but we also compute the shares with 65 and 75 cutoffs and use them as further robustness checks.

The association between *per capita income* and health conditions is the focus of a voluminous literature. Even though the evidence on the sign and direction of the association is mixed, many studies are suggestive of a key role of income in determining health outcomes (see e.g., Ettner 1996, Ruhm 2000, Frijters *et al.* 2005). We collected data on municipality per capita income thanks to the data "Open Data Dichiarazioni" issued by the MEF on a yearly basis, which include the pre-tax income at the municipality level.

Recovery Plan regimes are intended to control for peculiar financial constraints that may have interfered with the quality of health care provided. Indeed, starting from 2007, ten Italian regions had to implement the reorganization plans of the local health system, commonly called Recovery Plans, due to excessive financial deficits. Six of the ten regions (Abruzzo, Calabria, Campania, Lazio, Molise and Sicily) have been continuously under RP since the period 2007-09, two (Piedmont and Puglia) have started their RP regime in 2010, while two others (Liguria and Sardinia) were subject to a Plan only in the three-year period 2007-2009. From 2017 Piedmont is no longer under RP (Aimone Gigio *et al.*, 2018). Recovery Plan is thus a dummy variable with value 1 for the municipalities belonging to a region under RP in a given year, 0 otherwise.

Natural disaster refers to earthquakes, flooding and any other meteorological event that caused deaths. Data are available only at province level. We then constructed a dummy variable taking value 1 for all the municipalities whose province experienced a natural calamity in a given year, 0 otherwise.

4.3 Geographical Classification

Throughout the empirical analysis, we distinguish between mountain and non-mountain municipalities, following the ISTAT municipality classification of relevance, to detect possible heterogeneous effects. The motivation of this choice relies on the following two aspects. The distinction of the national territory between mountain and non-mountain municipalities dates back to the post-World War II years. In fact, the adopted ISTAT classification is based on the law n. 991 of 1952, with which the mountain areas were introduced for the first time. The indication of the criteria for the classification of mountain municipalities was intended to establish the possibility to access specific benefits as territories subject to development delay. Moreover, health differences are observed by urban, rural and mountain areas in the Italian case (Bertoncello *et al.*, 2007).

4.4 Summary Statistics

Table 2.1 recaps a brief description of the data, the sources and the main descriptive statistics for all municipalities during the entire observation period.

Table 2.1: Summary statistics - all municipalities

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**	**Source**
Mortality rate (%)	1.115	0.491	0	9.090	101311	ISTAT
Total deaths	74.711	308.606	0	26873	102701	ISTAT
Population aged 55 and older(%)	34.154	6.761	9.492	81.818	101576	ISTAT
Per capita income (thousands)	10.216	3.284	1.562	52.904	100989	MEF
Natural disasters	0.009	0.092	0	1	103187	European Union and Italian Civil Protection
Population (thousands)	7.636	28.084	0.031	1345.851	101576	ISTAT
Recovery Plan	0.24	0.427	0	1	101590	Our elaboration on MH data
9 plus years before the policy	0.143	0.35	0	1	103187	Our elaboration on MH data
7 and 8 years before the policy	0.056	0.23	0	1	103187	Our elaboration on MH data
5 and 6 years before the policy	0.07	0.255	0	1	103187	Our elaboration on MH data
3 and 4 years before the policy (the baseline)	0.08	0.271	0	1	103187	Our elaboration on MH data
1 and 2 years before the policy	0.081	0.272	0	1	103187	Our elaboration on MH data
1 and 2 years after the policy	0.05	0.219	0	1	103187	Our elaboration on MH data
3 and 4 years after the policy	0.037	0.189	0	1	103187	Our elaboration on MH data
5 and 6 years after the policy	0.035	0.183	0	1	103187	Our elaboration on MH data
7 and 8 years after the policy	0.033	0.177	0	1	103187	Our elaboration on MH data
9 plus years after the policy	0.034	0.182	0	1	103187	Our elaboration on MH data

Source: The author

Table 2.2 gives an insight on the difference between never treated municipalities and those treated before the reform. This allows us to identify pre-existing differences. The descriptive statistics show a higher mortality rate and exposure to Recovery Plan in the never treated municipalities.

Table 2.2: Summary statistics: Never treated vs treated

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Never treated					
Mortality rate (%)	1.159	0.523	0	9.090	39089
Total deaths	77.109	322.94	0	10767	39144
Population aged 55 and older(%)	34.713	7.133	9.492	81.818	39169
Per capita income (thousands)	10.431	3.129	2.222	52.904	38976
Recovery Plan	0.289	0.454	0	1	39030
Natural disasters	0.012	0.11	0	1	39318
Population (thousands)	7.650	30.486	0.062	1008.419	39169
Treated pre-reform					
Mortality rate (%)	1.053	0.446	0	7.761	44047
Total deaths	73.622	301.698	0	14417	44073
Population aged 55 and older(%)	33.284	6.248	12.382	78.832	44214
Per capita income (thousands)	10.303	3.427	1.562	45.608	43973
Recovery Plan	0.066	0.249	0	1	44367
Natural disaster	0.009	0.095	0	1	44367
Population (thousands)	7.659	28.442	0.031	1337.155	44214

Source: The author

Table 2.3 provides the main descriptive statistics for mountain municipalities while table 2.4 distinguishes between never treated and treated before the reform. Sources are unchanged. Again the mortality rate and the exposure to Recovery Plan are higher in the never treated municipalities, that also show a higher per-capita income and a minor population density.
Table 2.5 and 2.6 focus on non-mountain municipalities and confirm the previous custom about higher mortality rate and exposure to Recovery Plan for never treated municipalities.

Table 2.3: Summary statistics - Mountain municipalities

Variable	Mean	Std. Dev.	Min.	Max.	N
Mortality rate (%)	1.262	0.586	0	9.090	42829
Total deaths	30.574	49.407	0	1077	42870
Population (thousands)	2.818	5.199	0.031	117.317	42903
Population aged 55 and older(%)	36.669	7.387	12.382	81.818	42903
Per capita income (thousands)	9.681	3.074	1.562	52.904	42877
Natural disasters	0.01	0.099	0	1	43068
Recovery Plan	0.255	0.436	0	1	42966
9 plus years before the policy	0.114	0.318	0	1	43068
7 and 8 years before the policy	0.051	0.22	0	1	43068
5 and 6 years before the policy	0.068	0.251	0	1	43068
3 and 4 years before the policy	0.08	0.272	0	1	43068
1 and 2 years before the policy	0.081	0.273	0	1	43068
1 and 2 years after the policy	0.057	0.232	0	1	43068
3 and 4 years after the policy	0.047	0.212	0	1	43068
5 and 6 years after the policy	0.044	0.204	0	1	43068
7 and 8 years after the policy	0.04	0.196	0	1	43068
9 plus years after the policy	0.026	0.159	0	1	43068

Source: The author

Table 2.4: Mountain municipalities: Never treated vs treated

Variable	Mean	Std. Dev.	Min.	Max.	N
Never treated					
Mortality rate (%)	1.301	0.629	0	9.090	16811
Total deaths	28.491	47.598	0	1077	16845
Population aged 55 and older(%)	37.462	7.832	16.05	81.818	16812
Per capita income (thousands)	10.486	2.873	2.362	52.904	16812
Recovery Plan	0.256	0.436	0	1	16791
Natural disaster	0.01	0.099	0	1	16893
Population (thousands)	2.545	4.948	0.062	117.317	16812
Treated pre-reform					
Mortality rate (%)	1.197	0.526	0	7.761	16812
Total deaths	31.992	48.362	0	998	16819
Population aged 55 and older(%)	35.576	6.762	12.382	78.832	16885
Per capita income (thousands	8.991	3.043	1.562	24.027	16871
Recovery Plan	0.094	0.292	0	1	16969
Natural disaster	0.015	0.123	0	1	16969
Population (thousands)	3.051	5.279	0.031	98.657	16885

Source: The author

Table 2.5: Summary statistics - Non mountain municipalities

Variable	Mean	Std. Dev.	Min.	Max.	N
Mortality rate (%)	0.978	0.354	0	5.536	46826
Total deaths	106.865	394.85	0	14417	46854
Population (thousands)	11.313	37.077	0.056	1345.851	46865
Population aged 55 and older(%)	31.902	5.434	9.492	58.261	46865
Per capita income (thousands)	10.854	3.396	2.222	45.608	46667
Natural disasters	0.007	0.086	0	1	46961
Recovery Plan	0.207	0.405	0	1	46896
9 plus years before the policy	0.182	0.386	0	1	46961
7 and 8 years before the policy	0.064	0.245	0	1	46961
5 and 6 years before the policy	0.075	0.263	0	1	46961
3 and 4 years before the policy	0.083	0.276	0	1	46961
1 and 2 years before the policy	0.084	0.278	0	1	46961
1 and 2 years after the policy	0.046	0.209	0	1	46961
3 and 4 years after the policy	0.027	0.163	0	1	46961
5 and 6 years after the policy	0.026	0.16	0	1	46961
7 and 8 years after the policy	0.026	0.158	0	1	46961
9 plus years after the policy	0.018	0.133	0	1	46961

Source: The author

Table 2.6: Non mountain municipalities: Never treated vs treated

Variable	Mean	Std. Dev.	Min.	Max.	N
Never treated					
Mortality rate (%)	1.032	0.386	0	5.536	17286
Total deaths	114.638	424.631	0	10767	17301
Population aged 55 and older(%)	32.278	5.672	9.492	55.895	17286
Per capita income (thousands)	10.605	3.334	2.222	39.523	17266
Recovery Plan	0.293	0.455	0	1	17264
Natural disasters	0.014	0.118	0	1	17329
Population (thousands)	11703.406	39151.767	68	1008419	17286
Treated pre-reform					
Mortality rate (%)	0.935	0.337	0	4.918	22821
Total deaths	98.869	403.277	0	14417	22834
Population aged 55 and older(%)	31.467	5.235	12.516	58.261	22860
Per capita income (thousands)	11.358	3.385	2.989	45.608	22750
Recovery Plan	0.045	0.207	0	1	22913
Natural disasters	0.005	0.068	0	1	22913
Population (thousands)	10667.395	37993.362	56	1337155	22860

Source: The author

5 Identification Strategy

We apply a quasi-experimental research design using an event study framework that enables us to examine anticipation and phase-in policy effects. Indeed, since each region set up the policy in different years, we use a staggered two-way Difference in Differences (DID) as identification strategy to asses the policy impact. DID framework is one of the workhorse models used in public health policy research to learn about causal relationships when randomized controlled trials are not suitable in practice. The identifying assumption in the two-way fixed effects models is that there were no other variables - such as other public policies – that also affected the outcome of interest, that were coincident with the policy evaluated (Bitler and Carpenter, 2016). Recovery Plan regimes do not arise concern with this regard because they where not coincident in time and space with the adoption of the LHAs consolidation policy, as pointed out in section 2. Moreover, since we have panel data, we control for unobservable factors by including municipality fixed effect and time fixed effect. Municipality fixed effects eliminate any confounding factor, observed or unobserved, that is constant over time within each municipality. Year fixed effects eliminate any confounding factor, observed or unobserved, that is constant across all municipalities within each year, namely they trace out the common trend.

To properly identify a DID model, a control group able to act as a valid comparison for the treated group is needed. With this regard, the major concern is the policy endogeneity (Bellou and Bhatt, 2013). In fact, unobserved characteristics of municipalities belonging to specific LHAs may drive the decision to make the policy switch and they may be correlated to quality of care provided that affects eventually population mortality. Descriptive statistics provided in section 4 suggest the occurrence of pre-existing differences between the never treated and treated municipalities. Hence, following Raifman *et al.* (2017), we ran a "Granger-Type" causality test, examining if the current mortality rate is associated to the probability of adopting the consolidation policy in the future 3-4 years. We select this period to take into account the closest period to the policy shift which is reasonably not contaminated by potential anticipation effects. Results are reported in tables 2.14 and 2.15 in the appendix and the coefficients estimated do not reject the null hypothesis of no association. If the lead variable for implementing the consolidation policy in the future was associated with mortality rate, it would have indicated that our results may be owing to time trends in treated municipalities being systematically different from time trends in control municipalities.

Figure 2.6 in the appendix depicts the mortality trend by treated and never treated municipalities, suggestive of no violation of the common trend assumption typical of the standard DID approach.[2] We employ OLS to estimate the following baseline standard two-way fixed effect DID model for the general treatment effect with :[3]

$$y_{it} = \alpha_i + \delta_t + \gamma D_{it} + \phi X'_{it} + \epsilon_{it}, \tag{2.1}$$

where y_{it} is the mortality rate for municipality i at time t, α_i are the municipalities fixed effects and δ_t the year fixed effects, D_{it} is the treatment dummy assuming value 1 in any year after the reform for the treated municipality i, 0 otherwise, and X'_{it} is a set of the controls. We further extend the model to better fit the complexity of the consolidation reform. More precisely, given the availability of data, we standardized the time dimension as $m = -15$ periods before and $n = +12$ periods after the treatment. We then have a certain time window around the adoption of the policy $(-15, -14, ..., 0, ..., 11, 12)$, where 0 is the year of policy switch, that allows us to capture either the immediate effect of the policy, potential pre-existing trends, possible anticipation effects, and any additional effects that occur n periods after adoption. This event study design is similar to the approach undertaken by Stevenson and Wolfer (2006), Gippet *et al.*, (2015), Bellou & Bhatt (2013), Anderson *et al.* (2015), Bitler & Carpenter (2016), Simon (2016), Marcus & Siedler (2015) and Paik *et al.* (2013). Likewise, Brot-Goldberg *et al.* (2017) and Alpert (2016), specifically look for anticipatory effects while Kolstad & Kowalski (2012) consider periods before, during, and after treatment. That is, following Bellou and Bhatt (2013) we are able to capture and disentangle the full dynamic response of the mortality rate to the institutional change by evaluating the effect for each year following the adoption of the consolidation reform. Additionally, the dummies leading up to the policy shift test if there is evidence of pre-existing trends after controlling for other covariates. We combine the years (1-2, 3-4, etc.) instead of using one year increment because we expect the policy to affect the health outcomes

[2]More formally, the parallel trends assumption in DID frameworks means that the difference between the treatment and comparison group means remains constant in the absence of treatment (Alpert, 2016).

[3]Following many studies that deal with limited dependent variables (see e.g. Chen and Jin 2012, Bellou and Bhatt 2013, Carpenter *et al.* 2017, Stevenson and Wolfer, 2006) we prefer the specification of a linear model for the ease of interpretation and to avoid the incidental parameter problem that arises in the estimations of nonlinear panel models including individual specific fixed effects (first noted by Neyman and Scott 1948 for static models, and by Nickell 1981 for dynamic models). However, We ran fixed effect Poisson regression for the death count as well. Results are similar and are not reported here.

gradually. We set the baseline period as 4 and 3 year before. The terminal point is fixed at 9 plus year, both in the before and after periods, to avoid having a small number of observation in later years.

This approach not only gives us a deeper insight on the time-varying policy effects and their persistence over time, but also allows for potential anticipation effects by looking at the municipality mortality one and two years before the reform approval. In general, anticipation effects are possible whenever a policy includes a time gap between announcement and effective date (Wing *et al.*, 2018). This is actually the case of LHAs mergers where the decision is made by the regional government after long and lively political debates.

We estimate the following equation:

$$y_{it} = \alpha_i + \delta_t + \sum_{k=m}^{n} \gamma_k D_{i,t} + \phi X'_{it} + \epsilon_{it}, \tag{2.2}$$

where y_{it} is the mortality rate for municipality i at time t, α_i are the municipalities fixed effects and δ_t the year fixed effects, D_{it} is a vector of time effects relative to time from $m = -15$ to $n = +12$ and X'_{it} is a set of the controls.

The standard errors are clustered by municipality in any specification as our observations are non-independent and serial correlation may lead to serious overestimation of t-statistics (Bertrand *et al.*, 2004).[4]

6 Results

Table 2.7 reports the coefficients for the aggregated treatment effect for three different sample of municipalities. Columns 1 refers to the population mortality for all municipalities, regardless of their geographical characteristics. Columns 2 and 3 distinguish between mountain municipalities and non-mountain municipalities, respectively. The results are suggestive of an increase of mortality rate due to the policy shift. More specifically, the implementation of mergers yields

[4] The cluster-robust standard error estimator (CRSE) converges to the true standard error as the number of clusters approaches infinity. Kézdi (2004) shows that 50 clusters, with roughly equal cluster sizes, is close to infinity for accurate inference, and even in the absence of clustering, there is little to no cost of using the CRSE estimator, as long as the number of clusters is not less than 50. The number of municipalities included in this analysis is large enough to restrain concerns on the appropriateness of CRSE in this setting.
We also tried cluster-robust standard errors at the LHAs level. Results do not change.

to a higher mortality rate by 0.0263 percentage points. Given that the mortality rate may range between 0 and 100, with a sample pre-reform average of 1.053, the estimated coefficient can be readily interpreted in terms of increased mortality rate. It implies that, being in the post-reform years (a unit increase in the associated indicator), would increase the population mortality by approximately 2.5%.[5] Mountain municipalities register a roughly three times higher increase compared to non-mountain municipalities.

The aging of population seems to be positively correlated to mortality, as expected, likewise the recovery plan regime for all municipalities. The negative sign of the per capita income confirms the positive association between health and income suggested by the literature. Population density is also significant and negatively associated with mortality, confirming the positive association between population size and better health outcomes reported in section 4. The coefficient associated to Natural Disaster never turns out associated to mortality. This might be due to the weakness of the variable deputed to capture the increased level of mortality risk since the only available data is the province where the calamity occurred.

[5] 1.84 more deaths in absolute terms $(0.0263 * (73.622/1.053))$.

Table 2.7: Estimation results: Two-way fixed effect DID

	(1)	(2)	(3)
VARIABLES	All municipalities	Mountain	Non-mountain
Treatment	0.0263***	0.0393***	0.0117**
	(0.00450)	(0.00824)	(0.00472)
Controls			
Per capita income	-0.0156***	-0.0233***	-0.00961***
	(0.00263)	(0.00466)	(0.00284)
Population aged 55 and older	0.0352***	0.0348***	0.0366***
	(0.00163)	(0.00249)	(0.00214)
Natural Disaster	-0.00678	-0.0155	0.0146
	(0.0121)	(0.0211)	(0.0129)
Recovery Plan	0.0293***	0.0426***	0.0157***
	(0.00455)	(0.00874)	(0.00498)
Population	-0.00313***	-0.0243***	-0.000616
	(0.00114)	(0.00711)	(0.000675)
Municipality fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	100,721	42,777	46,606
R-squared	0.600	0.540	0.640

Notes: Robust standard errors in parentheses, clustered at municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variable is the municipality population mortality rate by year.

Table 2.8: Estimation results: Two-way fixed effect DID

VARIABLES	(1) All municipalities	(2) Mountain	(3) Non-mountain
9 plus years before the polic	0.00397	0.00676	0.00345
	(0.00659)	(0.0126)	(0.00750)
7 and 8 years before the policy	0.000506	-0.000131	-0.00287
	(0.00676)	(0.0135)	(0.00619)
5 and 6 years before the policy	-0.00499	-0.00743	-3.09e-05
	(0.00509)	(0.00978)	(0.00507)
3 and 4 years before the policy	-	OMITTED	-
1 and 2 years before the policy	0.00675	0.0126	0.00104
	(0.00507)	(0.00967)	(0.00516)
1 and 2 years after the policy	0.0172***	0.0289**	0.00989
	(0.00619)	(0.0114)	(0.00651)
3 and 4 years after the policy	0.0178**	0.0260**	0.00648
	(0.00711)	(0.0124)	(0.00745)
5 and 6 years after the policy	0.0370***	0.0512***	0.0108
	(0.00769)	(0.0137)	(0.00782)
7 and 8 years after the policy	0.0629***	0.102***	0.0168*
	(0.00862)	(0.0153)	(0.00905)
9 plus years after the policy	0.0720***	0.100***	0.0438***
	(0.0110)	(0.0192)	(0.0109)
Controls			
Per capita income	-0.0170***	-0.0259***	-0.00978***
	(0.00269)	(0.00495)	(0.00284)
Population aged 55 and older	0.0352***	0.0349***	0.0367***
	(0.00165)	(0.00252)	(0.00220)
Natural Disaster	-0.00611	-0.0116	0.0140
	(0.0121)	(0.0213)	(0.0129)
Recovery Plan	0.0221***	0.0323***	0.0125**
	(0.00512)	(0.00948)	(0.00592)
Population	-0.00291***	-0.0239***	-0.000556
	(0.00110)	(0.00693)	(0.000675)
Municipality fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	100,721	42,777	46,606
R-squared	0.600	0.540	0.640

Notes: Robust standard errors in parentheses, clustered at municipality level.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variable is the municipality population mortality rate by year. The baseline pre-reform dummy, obtained by aggregating the third and the fourth and year preceding the consolidation policy, is omitted and represents the reference category.

Table 2.8 records the estimation results of equation 2. Findings are coherent with the previous general specification. No evidence of pre-existing trends in mortality rates is found and there is statistically significant increase in municipality mortality in all samples after the policy change. The effect grows over time with respect to the baseline pre-reform years (3 and 4 years before). In fact, the coefficients for the post-reform dummies are positive and significant in any year following the reform. The effect for non-mountain municipalities turns on only after 7 years and remains quite small. Interestingly, the effect is always higher in mountain municipalities, whose coefficients are reported in column 2, shedding light on the fact that those municipalities may tow the average treatment effect on the treated.[6] The insignificance of the coefficient associated to the time dummy 2-1 years before the policy suggests that no anticipation effect ever occurred.

7 Robustness Checks

We conducted several robustness checks. Even though the "Granger-type" causality test and figure 2.6 are suggestive of strict policy exogeneity and parallel trends in mortality rates between treated and control groups, we test the model for the presence of municipality-specific linear trend. In other words, we test our model for the presence of unobserved factors that trend over time within a municipality and are correlated to population mortality. We augmented equation 2 with a municipality-specific linear trend by estimating the following DID regression:

$$y_{it} = \alpha_i + \delta_t + \beta_i(\alpha_i * t) + \sum_{k=m}^{n} \gamma_k D_{i,t} + \phi X'_{it} + \epsilon_{it}, \tag{2.3}$$

where the only difference from equation 2 is the term $\alpha_i * t$ that indicates the municipality-specific linear trends. We believe that in an analysis based on population mortality as outcome of interest, time dummies and time trends are not redundant. Indeed, there is a potential long-term linear trend of the outcome, due for instance to the changes in the population age composition over time, and it could be also subject to periodic macroeconomic shocks that are large enough to affect mortality.

[6] Indeed, according to Ryan *et al.* (2015), "DID estimates are typically considered to be average treatment effects on the treated, rather than average treatment effects. This is because DID estimates are generally thought of as applying to a particular group that was treated (rather than to a population that could have been treated)."

Results are reported in table 2.9. Findings show that the inclusion of individual linear trends cause the standard errors to increase while point estimates are roughly similar to, but slightly smaller than, those shown in table 2.8 for all municipalities and mountain mountain municipalities. Only for non-mountain municipalities even the point estimates increase along with the standard errors. However, the increase in standard errors produces results that are not precisely estimated enough to reject a null that the pattern of coefficients is the one recorded in table 2.8. In fact, underlying trends are better identified when the population size is big enough to bear random shocks in the deaths count. On the contrary, as pointed out by Jarner and Kryger (2011), the mortality evolution of small populations exhibits substantial variability and irregular patterns. The comparison between the descriptive statistics reported in table 2.3 and 2.5 for mountain and non-mountain municipalities indicates how the latter are mostly characterized by a small population size with a sample mean of less than 3 thousands people. Further, we examine the sensitivity of our results to the sample chosen by omitting in turn individual regions. Findings - which are not reported here for the sake of brevity - suggest that particular regions do not influence our main results.

Additionally, we substituted the variable *Population aged 55 and older* trying 65 and 75 age cutoffs without obtaining substantially different results. Finally we conduct falsification exercises by simulating the policy change in a numerous of various subsamples - balanced to being representative of all the NUTS 1 macroareas - of never treated municipalities and different years. The fake policy shift does not turn out significant.[7]

[7] Results are not reported here for the sake of briefness.

Table 2.9: Two-way fixed effect DID with municipality-specific linear trend

VARIABLES	(1) All municipalities	(2) Mountain	(3) Non-mountain
9 plus years before the policy	0.00152	-0.0175	0.00476
	(0.0124)	(0.0259)	(0.0130)
7 and 8 years before the policy	0.00551	0.00328	-0.00404
	(0.00874)	(0.0174)	(0.00890)
5 and 6 years before the policy	-0.00361	-0.00826	-0.00166
	(0.00588)	(0.0115)	(0.00603)
3 and 4 years before the policy	-	OMITTED	-
1 and 2 years before the policy	0.00605	0.0140	0.00232
	(0.00632)	(0.0122)	(0.00636)
1 and 2 years after the policy	0.0107	0.0240	0.0145
	(0.00976)	(0.0191)	(0.00979)
3 and 4 years after the policy	0.0283	0.00878	0.0183
	(0.0152)	(0.0287)	(0.0147)
5 and 6 years after the policy	0.0222	0.0244	0.0360*
	(0.0202)	(0.0384)	(0.0195)
7 and 8 years after the policy	0.0489*	0.0702	0.0534**
	(0.0264)	(0.0502)	(0.0249)
9 plus years after the policy	0.0514	0.0632	0.0746**
	(0.0329)	(0.0614)	(0.0321)
Controls			
Per capita income	-0.0313***	-0.0422***	-0.0216***
	(0.00437)	(0.00874)	(0.00492)
Population aged 55 and older	0.0699***	0.0757***	0.0654***
	(0.00333)	(0.00495)	(0.00442)
Natural Disaster	-0.0100	-0.0216	0.0146
	(0.0128)	(0.0224)	(0.0138)
Recovery Plan	-0.00107	-0.00172	0.00693
	(0.00708)	(0.0131)	(0.00791)
Population	0.00271*	0.131***	0.00123
	(0.00145)	(0.0236)	(0.000837)
Municipality fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Municipality-specific linear trend	Yes	Yes	Yes
Observations	100,721	42,777	46,606
R-squared	0.640	0.583	0.682

Notes: Robust standard errors in parentheses, clustered at municipality level.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variable is the municipality population mortality rate by year. The baseline pre-reform dummy, obtained by aggregating the third and the fourth and year preceding the consolidation policy, is omitted and represents the reference category.

7.1 Age-Adjusted Mortality Rate

The direct comparison of crude mortality rates may suffer from distortion caused by the different population age distributions among municipalities. In fact, an higher mortality rate in some municipalities may be a result of an older population rather than the exposure the policy change focus of this analysis. The coefficient associated with the variable *Population aged 55 and older* in the previous estimation results suggests, as expected, the role of an aging population in the crude mortality rates. We address this issue by calculating the age-adjusted mortality rates for any municipality included in our sample.[8]

The age-adjusted rates are those rates that would have existed if the observed population had the same age distribution as the population taken as "standard". Therefore, they are summary measures adjusted for differences in age distributions.

Following the example applied to the Italian LHAs provided by Consonni *et al.* (2012), we first calculated the age-standardized rates for any municipality using the Segi's world population as standard.[9] More specifically, the adjusted rate is defined as the weighted sum by age groups of the crude rates, where the weights are given by the standard population distribution (Rothman, 1986).[10] Table 2.10 reports the summary statistics for the resulting adjusted measure of mortality.

Table 2.10: Summary statistics

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Age-adjusted mortality rate (%)	0.329	0.183	0	13	91310

Source: The author.
Notes: The number of observation is now smaller because the age-specific mortality data are available only until 2015.

We reestimated equation 2.1, 2.2 and 2.3 with the age-adjusted mortality rate as dependent variable. Results from equation 2.1 estimation are reported in table 2.11. The treatment es-

[8] We are very grateful to Antonella Ciccarese from ISTAT that kindly provide us with the necessary databases "Decessi comunali per classi di età", which include the age-specific death count by municipality.

[9] The standard world Segi's population includes 18 age categories with decreasing weights for older categories. Table 2.14 in appendix provides the relevant details.

[10] Algebraically is defined by

$$Age\_Adj\_Rate = \frac{\sum_{i=1}^{18} W_i R_i}{\sum_{i=1}^{18} W_i} \tag{2.4}$$

where R_i is the stratum specific rate in age group i and w_i is the weight for stratum i given by the standard population.

timates confirm the previous results but are smaller in magnitude. In fact, being in the years following the reform is associated with a mortality increase of 0.00441 percentage points (1.34% at means).[11] Again, the effect is higher in mountain municipalities than in non-mountain ones but the difference between them is now smaller, suggesting that the previous results may have been driven by differences in the population age distribution by geographical municipality characteristics.

The new estimation results of equation 2.2 and 2.3 are reported in table 2.12 (columns 2, 4 and 6 include the municipality-specific linear trend). Overall, estimates are suggestive of an increase in the age-adjusted mortality rate after the policy switch. Therefore key findings are robust after adjusting for the age structure of the population in any given period. Interestingly, while a significance level of 0.10 indicates a weak evidence of some anticipation effects and preexisting trends on the two subsamples of municipalities and on the overall sample, respectively, estimates show a robust pattern of statistically different from zero policy effects in the years following the reform. More specifically, the effect appears always positive and increasing over time, and the differences between mountain and non-mountain municipalities have been smoothed out, although the mortality still remain slightly higher in mountain municipalities five years after the reform.

Table 2.11: Two-way fixed effect DID with age-adjusted mortality rate

VARIABLES	(1) All municipalities	(2) Mountain	(3) Non-mountain
Treatment	0.00441**	0.00755**	0.00585***
	(0.00207)	(0.00372)	(0.00195)
Controls			
Per capita income	-0.00582***	-0.00723***	-0.00327**
	(0.00124)	(0.00201)	(0.00162)
Natural Disaster	-0.000342	-0.00472	-0.00274
	(0.00563)	(0.00689)	(0.00503)
Recovery Plan	0.0136***	0.0145***	0.0102***
	(0.00202)	(0.00413)	(0.00223)
Population	0.000190	0.00285	-8.69e-05
	(0.000261)	(0.00184)	(0.000196)
Municipality fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	91,310	39,887	43,280
R-squared	0.192	0.172	0.256

Notes: Robust standard errors in parentheses, clustered at municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variable is the municipality population age-adjusted mortality rate by year.

[11] $(0.00441/0.329) * 100$

Table 2.12: Two-way fixed effect DID with age-adjusted mortality rate

VARIABLES	(1) All municipalities	(2) All municipalities	(3) Mountain	(4) Mountain	(5) Non mountain	(6) Non mountain
9 plus years before the policy	-0.00157	-0.00124	-0.00290	-0.0109	0.000663	0.00197
	(0.00307)	(0.00551)	(0.00640)	(0.0113)	(0.00300)	(0.00588)
7 and 8 years before the policy	0.00730*	0.00676	0.0141	0.0127	0.00158	-0.000447
	(0.00418)	(0.00499)	(0.00965)	(0.0116)	(0.00272)	(0.00391)
5 and 6 years before the policy	0.00243	0.00139	0.00323	0.00223	0.00176	-0.000367
	(0.00213)	(0.00261)	(0.00387)	(0.00495)	(0.00245)	(0.00295)
3 and 4 years before the policy	-	-	OMITTED	-	-	
1 and 2 years before the policy	0.00133	0.00268	0.00748*	0.00701	-0.00420*	0.000176
	(0.00219)	(0.00275)	(0.00408)	(0.00540)	(0.00232)	(0.00285)
1 and 2 years after the policy	0.00409	0.00242	0.00852*	0.00268	0.00333	0.00912*
	(0.00279)	(0.00490)	(0.00479)	(0.00929)	(0.00283)	(0.00487)
3 and 4 years after the policy	0.00336	0.000620	0.00954*	-0.00118	0.00357	0.0137*
	(0.00300)	(0.00718)	(0.00526)	(0.0137)	(0.00290)	(0.00717)
5 and 6 years after the policy	0.00916***	0.00511	0.0195***	0.00306	0.00192	0.0163*
	(0.00327)	(0.00967)	(0.00571)	(0.0183)	(0.00311)	(0.00963)
7 and 8 years after the policy	0.0144***	0.00936	0.0273***	0.00640	0.00859**	0.0263**
	(0.00361)	(0.0123)	(0.00635)	(0.0235)	(0.00395)	(0.0121)
9 plus years after the policy	0.0198***	0.0104	0.0391***	0.00935	0.00865*	0.0280*
	(0.00498)	(0.0156)	(0.00878)	(0.0293)	(0.00470)	(0.0153)
Controls						
Per capita income	-0.00629***	-0.00537***	-0.00802***	-0.00593**	-0.00349**	-0.00397
	(0.00124)	(0.00202)	(0.00199)	(0.00280)	(0.00163)	(0.00303)
Natural Disaster	0.000802	0.00107	-0.00110	-0.00181	-0.00236	-0.00271
	(0.00558)	(0.00581)	(0.00690)	(0.00780)	(0.00504)	(0.00590)
Recovery Plan	0.0145***	0.00560	0.0151***	0.00860	0.0105***	0.000180
	(0.00223)	(0.00346)	(0.00419)	(0.00684)	(0.00249)	(0.00403)
Population	0.000171	-0.000507*	0.00287	0.0146**	-9.02e-05	-0.000428*
	(0.000263)	(0.000259)	(0.00186)	(0.00704)	(0.000197)	(0.000239)
Municipality fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality-specific linear trend	No	Yes	No	Yes	No	Yes
Observations	91,310	91,310	39,887	39,887	43,280	43,280
R-squared	0.192	0.272	0.173	0.254	0.256	0.331

Notes: Robust standard errors in parentheses, clustered at municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variable is the municipality population age-adjusted mortality rate by year. The baseline pre-reform dummy, obtained by aggregating the third and the fourth and year preceding the consolidation policy, is omitted and represents the reference category.

8 Final Remarks and Discussion

The merging processes in health care, documented in the literature, started in the 1990s in the United States and the United Kingdom, leading to a radical change in the structures of their Health Systems. Other countries as well, as Germany, the Netherlands, and Canada, are moving in the same direction, such that this new rage can be defined as "merger mania". Italy is not an exception since mergers in health structures are still an ongoing process.

This chapter investigated the impact of LHAs consolidation process on population mortality within Italian municipalities from 2002 to 2016, aiming at providing some evidence on potential indirect effects produced by this recent "merging mania" that has characterized the Italian and international institutional dynamics. Since each Italian region set up the policy in different years, we used a staggered two-way fixed effect DID method as identification strategy to estimate the treatment effect, disentangling the full dynamics both in the short run and in the medium-long run. We estimated the impact of these policy shifts on population mortality adjusting for the population age structure and using the within municipality variation over time. We also distinguished between the geographical characteristics of municipalities (mountain Vs. non-mountain) and controlled for individual municipality linear trends. Key findings show that population mortality generally increased by roughly 1.34% (at means). The effect is persistent and increase over time. Mountain municipalities register a higher increase compared to non-mountain municipalities. We therefore conclude that policy makers, both in Italy and in the other countries where the mergers have been implemented, should carefully consider the impact on health care quality and effectiveness before allowing more mergers.

This analysis has some limitations. First, albeit we control for the age structure of the population, mortality rate is not risk adjusted for the presence of comorbidities, since data on the incidence of specific health conditions at municipality level are not available. Thus, the possibility of some municipalities having riskier patients was minimized but could not be completely ruled out.

Second, there is the possibility of uncontrolled geographical variation of some conditions that may ultimately lead to bad health and death such as pollution. In fact, they may vary over time and by municipality, without showing entirely some sort of time trend. As a result, municipality fixed effects, time fixed effects and municipality-specific linear trends are not able to pick up these potential confounders. To the best of our knowledge, data on pollution (for instance, the widely used levels of Particulate Matter 2.5 and 10) are available only for a few cities, specifically where the detection stations are located, and do not allow us to have a robust sample size. Third, given the availability of data, we were able to evaluate the consolidation policy only from 2002 to 2016. Previous mergers took place, as depicted in figure 2.3. However, most of the policy changes were implemented before 1995 and the total number of LHAs remained constant for many years preceding the wave of consolidation process analyzed here. Hence, our results can be interpreted as the effect of additional centralization in health care management.

Finally, our analysis does not allow to understand the mechanisms through which implementation of mergers increased population mortality.

Further research and more detailed data are needed to understand the association between centralization of health management and health status in the Italian case.

Bibliography

[1] Alpert, A. (2016). The anticipatory effects of Medicare Part D on drug utilization. Journal of health economics, 49, 28-45.

[2] Aimone Gigio, L., Alampi, D., Camussi, S. A., Ciaccio, G., Guaitini, P., Lozzi, M., Mancini, A. L., Panicara, E., & Paolicelli, M. (2018). La Sanità in Italia: Il Difficile Equilibrio Tra Vincoli Di Bilancio E Qualità Dei Servizi Nelle Regioni in Piano Di Rientro (The Italian Healthcare System: The Difficult Balance between Budget Sustainability and the Quality of Services in Regions Subject to a Financial Recovery Plan).

[3] Alpert, A. (2016). The anticipatory effects of Medicare Part D on drug utilization. Journal of health economics, 49, 28-45.

[4] Bardhan, P. (2002). Decentralization of governance and development. Journal of Economic perspectives, 16(4), 185-205.

[5] Bellou, A., & Bhatt, R. (2013). Reducing underage alcohol and tobacco use: Evidence from the introduction of vertical identification cards. Journal of health economics, 32(2), 353-366.

[6] Bertoncello, C., Cazzaro, R., Ferraresso, A., Mazzer, R., & Moretti, G. (2008). Prevalence of overweight and obesity among school-aged children in urban, rural and mountain areas of the Veneto Region, Italy. Public health nutrition, 11(9), 887-890.

[7] Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates?. The Quarterly journal of economics, 119(1), 249-275.

[8] Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates?. The Quarterly journal of economics, 119(1), 249-275.

[9] Besley, T., & Coate, S. (2003). Centralized versus decentralized provision of local public goods: a political economy approach. Journal of public economics, 87(12), 2611-2637.

[10] Bird, R. (1994). Decentralizing infrastructure: for good or ill? (No. 1258). The World Bank.

[11] Bitler, M. P., & Carpenter, C. S. (2016). Health insurance mandates, mammography, and breast cancer diagnoses. American Economic Journal: Economic Policy, 8(3), 39-68.

[12] Boffa, F., Piolatto, A., & Ponzetto, G. A. (2015). Political centralization and government accountability. The Quarterly Journal of Economics, 131(1), 381-422.

[13] Brot-Goldberg, Z. C., Chandra, A., Handel, B. R., & Kolstad, J. T. (2017). What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics. The Quarterly Journal of Economics, 132(3), 1261-1318.

[14] Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. The Review of Economics and Statistics, 90(3), 414-427.

[15] Carpenter, C. S., McClellan, C. B., & Rees, D. I. (2017). Economic conditions, illicit drug use, and substance use disorders in the United States. Journal of Health Economics, 52, 63-73.

[16] Cavalieri, Marina, Lara Gitto and Calogero Guccio. 2013. Reimbursement systems and quality of hospital care: An empirical analysis for Italy. Health Policy, 273– 289.

[17] Chen, Y., & Jin, G. Z. (2012). Does health insurance coverage lead to better health and educational outcomes? Evidence from rural China. Journal of health economics, 31(1), 1-14.

[18] Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., ... & Cutler, D. (2016). The association between income and life expectancy in the United States, 2001-2014. Jama, 315(16), 1750-1766.

[19] Consonni, D., Coviello, E., Buzzoni, C., & Mensi, C. (2012). A command to calculate age-standardized rates with efficient interval estimation. Stata Journal, 12(4), 688-701.

[20] Correia, S. (2017). reghdfe: Stata module for linear and instrumental-variable/gmm regression absorbing multiple levels of fixed effects. Statistical Software Components s457874, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s457874.html.

[21] Dalgaard, C. J., & Strulik, H. (2014). Optimal aging and death: understanding the Preston curve. Journal of the European Economic Association, 12(3), 672-701.

[22] Di Novi, C., Rizzi, D., Zanette, M. (2018). Scale Effects and Expected Savings from Consolidation Policies of Italian Local Healthcare Authorities. Applied health economics and health policy, 16(1), 107-122.

[23] Enikolopov, R., & Zhuravskaya, E. (2007). Decentralization and political institutions. Journal of public economics, 91(11-12), 2261-2290.

[24] Ettner, S. L. (1996). New evidence on the relationship between income and health. Journal of health economics, 15(1), 67-85.

[25] Francese, Maura, and Marzia Romanelli. "Is there Room for containing healthcare costs? An analysis of regional spending differentials in Italy." The European Journal of Health Economics 15.2 (2014): 117-132.

[26] Frijters, P., Haisken-DeNew, J. P., & Shields, M. A. (2005). The causal effect of income on health: Evidence from German reunification. Journal of health economics, 24(5), 997-1017.

[27] Fulop N, Protopsaltis G, King A, Allen P, Hutchings A, Normanda C. Changing organisations: a study of the context and processes of mergers of health care providers in England. Social Science & Medicine 2005; 60: 119–130.

[28] Fulop N, Protopsaltis G, Hutchings A, King A, Allen P, Normand C, Walters R. Process and impact of mergers of NHS trusts: multicentre case study and management cost analysis. BMJ 2002; 325(7358):246.

[29] Gaynor M, Seider H, Vogt WB. The volume-outcome effect, scale economies, and learning-by-doing. American Economic Review 2005;95(2):243–7.

[30] Gipper, Brandon, Christian Leuz, and Mark Maffett. Public audit oversight and reporting credibility: Evidence from the PCAOB inspection regime. No. w21530. National Bureau of Economic Research, 2015.

[31] Goddard, M., & Ferguson, B. (1997). Mergers in the NHS. London: The Nuffield Trust.

[32] Greene, W. (2004). Fixed effects and bias due to the incidental parameters problem in the Tobit model. Econometric reviews, 23(2), 125-147.

[33] Ho V, Hamilton BH. Hospital mergers and acquisitions: does market consolidation harm patients? Journal Health Economics, 2000.

[34] Hoornbeek, J., Morris, M. E., Stefanak, M., Filla, J., Prodhan, R., & Smith, S. A. (2015). The impacts of local health department consolidation on public health expenditures: Evidence from Ohio. American journal of public health, 105(S2), S174-S180.

[35] ISTAT - Health for All, version June 2018, https://www.istat.it/it/archivio/14562. Accessed August 2018.

[36] ISTAT - Geo-Demo. "Bilanci Demografici e Popolazione residente. Years 2002-2016. http://demo.istat.it/index.html. Accessed June 2018.

[37] Jarner, S. F., & Kryger, E. M. (2011). Modelling adult mortality in small populations: The SAINT model. ASTIN Bulletin: The Journal of the IAA, 41(2), 377-418.

[38] Kézdi, G. (2004). Robust Standard Error Estimation in Fixed-Effects Panel Models. Hungarian Statistical Review Special(9): 96-116.

[39] Kolstad, J. T., & Kowalski, A. E. (2012). The impact of health care reform on hospital and preventive care: evidence from Massachusetts. Journal of Public Economics, 96(11-12), 909-929.

[40] Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. Foundations and Trends® in Econometrics, 4(3), 165-224.

[41] Leone, T. (2010). How can demography inform health policy?. Health Economics, Policy and Law, 5(1), 1-11.

[42] Lichtenberg and Waldfogel (2009), Does Misery Love Company - Evidence from Pharmaceutical Markets before and after the Orphan Drug Act, 15 Mich. Telecomm. & Tech. L. Rev. 335.

[43] Luft, H. S., Bunker, J. P., & Enthoven, A. C. (1979). Should operations be regionalized? The empirical relation between surgical volume and mortality. New England Journal of Medicine, 301(25), 1364-1369.

[44] Luft, Harold S., Sandra S. Hunt, and Susan C. Maerki. "The volume-outcome relationship: practice-makes-perfect or selective-referral patterns?." Health services research 22.2 (1987): 157.

[45] Marcus, J., & Siedler, T. (2015). Reducing binge drinking? The effect of a ban on late-night off-premise alcohol sales on alcohol-related hospital stays in Germany. Journal of Public Economics, 123, 55-77.

[46] Mark Anderson, D., Hansen, B., & Rees, D. I. (2015). Medical marijuana laws and teen marijuana use. American Law and Economics Review, 17(2), 495-528.

[47] Ministry of Economy and Finance, Analisi statistiche - Open Data Dichiarazioni. http://www1.finanze.gov.it/finanze3/analisi_stat/index.php?search_class%5B0%5D=cCOMUNE&opendata=yes. Accessed June 2018.

[48] Ministry of Health, Corrispondenze ASL-Comuni e popolazione residente, 2017. http://www.salute.gov.it/portale/documentazione/p6_2_8_1_1.jsp?lingua=italiano&id=16. Accessed June 2018.

[49] Ministry of Health, Elenco Aziende sanitarie locali e Strutture di ricovero, 2017. http://www.salute.gov.it/portale/documentazione/p6_2_8_1_1.jsp?id\unhbox\voidb@x\bgroup\let\unhbox\voidb@x\setbox\@tempboxa\hbox{1\global\mathchardef\accent@spacefactor\spacefactor}\accent221\egroup\spacefactor\accent@spacefactor3. Accessed June 2018.

[50] Moore, D. A., & Carpenter, T. E. (1999). Spatial analytical methods and geographic information systems: use in health research and epidemiology. Epidemiologic reviews, 21(2), 143-161.

[51] Neyman, J., & Scott, E. L. (1948). Consistent estimates based on partially consistent observations. Econometrica: Journal of the Econometric Society, 1-32.

[52] Nickell, S. (1981). Biases in dynamic models with fixed effects. Econometrica: Journal of the Econometric Society, 1417-1426.

[53] Oates, W. E. (1972). Fiscal federalism. Books.

[54] OECD Health Statistics 2015, FOCUS on Health Spending.

[55] Paik, M., Black, B., & Hyman, D. (2013). The receding tide of medical malpractice litigation: part 2—effect of damage caps. Journal of Empirical Legal Studies, 10(4), 639-669.

[56] Rothman, K. (1986). Modern Epidemiology. Boston: Little, Brown.

[57] Ruhm, C. J. (2000). Are recessions good for your health?. The Quarterly journal of economics, 115(2), 617-650.

[58] Ryan, A. M., Burgess Jr, J. F., & Dimick, J. B. (2015). Why we should not be indifferent to specification choices for difference in differences. Health services research, 50(4), 1211-1235.

[59] Schmid, A., Varkevisser, M. (2016). Hospital merger control in Germany, the Netherlands and England: Experiences and challenges. Health Policy, 120(1), 16-25.

[60] Simon, D. (2016). Does early life exposure to cigarette smoke permanently harm childhood welfare? Evidence from cigarette tax hikes. American Economic Journal: Applied Economics, 8(4), 128-59.

[61] Stevenson, Betsey, and Justin Wolfers. "Bargaining in the shadow of the law: Divorce laws and family distress." The Quarterly Journal of Economics 121.1 (2006): 267-288.

[62] Raifman, J., Moscoe, E., Austin, S. B., & McConnell, M. (2017). Difference-in-differences analysis of the association between state same-sex marriage policies and adolescent suicide attempts. JAMA pediatrics, 171(4), 350-356.

[63] Rodríguez-Pose, A., & Ezcurra, R. (2009). Does decentralization matter for regional disparities? A cross-country analysis. Journal of Economic Geography, 10(5), 619-644.

[64] Tiebout, C. M. (1956). A pure theory of local expenditures. Journal of political economy, 64(5), 416-424.

[65] Wing, C., Simon, K., & Bello-Gomez, R. A. (2018). Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. Annual review of public health, 39, 453-469.

Appendix

Table 2.13: LHAs amalgamation from 2000 to 2017

Region	Year of Policy Shift	LHAs codes before and after the mergers
Piedmont	2008	101+102=201; 103+104=202; 105+106+110=203; 107+109=204; 115+116+117=210; 120+121+122=213
A. P. of Bolzano	2007	101 + 102 + 103 + 104 = 201
Emilia Romagna	2004	107+108+105=105
Marche	2006	101+102+103+104+105+106+ +107+108+109+110+111+112+113=201
Abruzzo	2010	101+104=201; 102+103=202
Molise	2006	101+102+103+104=201
Campania	2009	101+102=201; 104+105=203; 107+108=205; 109+110=206; 111+112+113=207
Apulia	2007	102+103+104+105=114; 107+108+109=115; 110+111=116
Basilicata	2009	101+102+103=201; 104+105=202
Calabria	2008	101+102+103+104=201; 106+107=203; 110+111=205
Calabria	2011	109+205=205
Umbria	2013	101+102=201; 103+204=202
Emilia Romagna	2014	110+111+112+113=114
Friuli V. Giulia	2015	102+105=202
Lombardy	2016	308+309+310+306=321; 314+303=322; 303+313+315=323; 324=305+311; 327=304+307
Tuscany	2016	201=103+104+110+111; 202=101+102+105+106+112; 203=108+107+109
Lazio	2016	201=105+101; 202=102+103
Sardinia	2017	101+102+103+104+105+106+107+108=201

Notes: The size of those not included remained unchanged. Some LHAs in Veneto and Piedmont experienced a merger in 2017 but, due to the lack of an up-to-date list of Italian hospitals by Ministry of Health for 2017, we are not able to reconstruct the merging design. Sardinia 2017 mergers are instead clear even without detailed information about single hospitals because all LHAs have been merged into only one.
Source: the author, on Ministry of Health data.

Table 2.14: "Granger-Type" Causality test

VARIABLES	(1) All municipalities	(2) All municipalities	(3) Mountain	(4) Mountain	(5) Non-mountain	(6) Non-mountain
Policy shift in 3-4 years	0.00502	0.00482	0.00378	0.00216	0.00709	0.00700
	(0.00479)	(0.00525)	(0.00888)	(0.0100)	(0.00479)	(0.00519)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality-specific linear trend	No	Yes	No	Yes	No	Yes
Observations	43,846	43,846	16,794	16,794	22,711	22,711
R-squared	0.612	0.664	0.551	0.614	0.631	0.681
Adjusted R-squared	0.571	0.585	0.498	0.512	0.595	0.613
F test	72.91	32.66	23.07	19.09	50.79	24.29
Within R-squared	0.0193	0.0239	0.0173	0.0271	0.0273	0.0282
Adjusted Within R-squared	0.0191	0.0237	0.0169	0.0267	0.0270	0.0279

Notes: Robust standard errors in parentheses, clustered at municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variable is the current municipality population mortality rate by year. The sample is restricted to treated municipalities before the reform.

Table 2.15: "Granger-Type" Causality test with age-adjusted mortality rate

VARIABLES	(1) All municipalities	(2) All municipalities	(3) Mountain	(4) Mountain	(5) Non-mountain	(6) Non-mountain
Policy shift in 3-4 years	0.00179	0.00157	0.00155	0.00332	0.00328	0.00224
	(0.00247)	(0.00261)	(0.00475)	(0.00524)	(0.00274)	(0.00275)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality-specific linear trend	No	Yes	No	Yes	No	Yes
Observations	41,710	41,710	16,529	16,529	22,071	22,071
R-squared	0.231	0.335	0.212	0.316	0.273	0.358
Adjusted R-squared	0.148	0.174	0.118	0.132	0.201	0.218
F test	2.222	0.107	1.724	0.720	1.709	0.546
Within R-squared	0.000138	1.13e-05	0.000125	6.04e-05	0.000319	9.54e-05
Adjusted Within R-squared	3.18e-05	-0.000108	-0.000146	-0.000247	0.000120	-0.000125

Notes: Robust standard errors in parentheses, clustered at municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variable is the current municipality population age-adjusted mortality rate by year. The sample is restricted to treated municipalities before the reform.

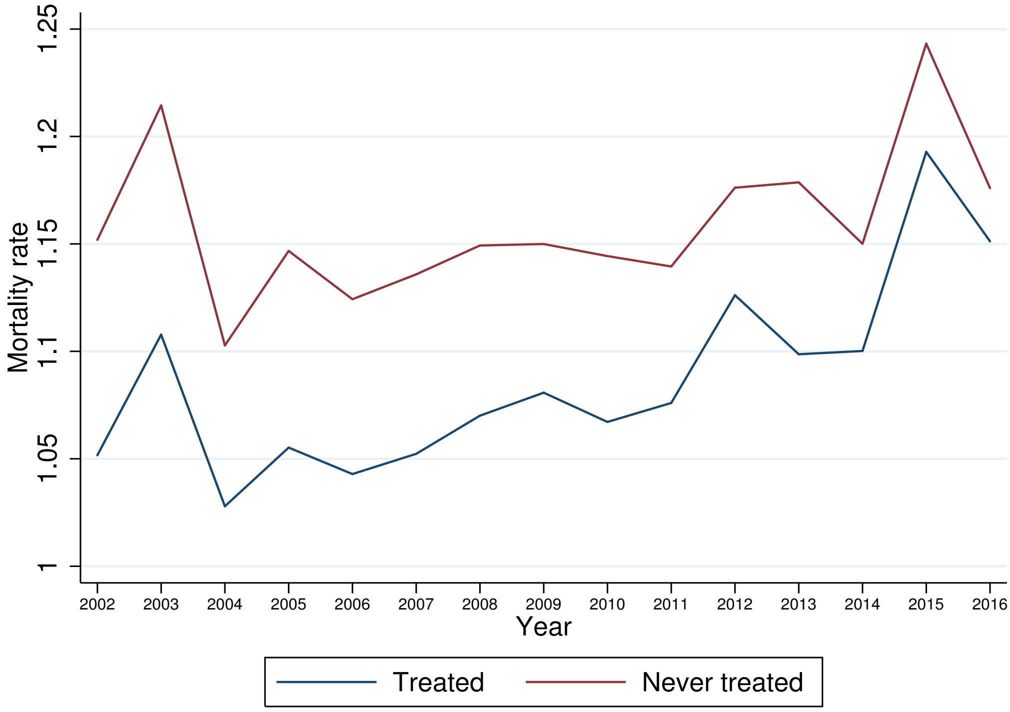

Figure 2.6: Mortality trends by treated and controls.
Source: Author elaboration on ISTAT data

Table 2.16: Segi's world standard population

Age group (years)	Population
0-4	12.00
5-9	10.00
10-14	9.00
15-19	9.00
20-24	8.00
25-29	8.00
30-34	6.00
35-39	6.00
40-44	6.00
45-49	6.00
50-54	5.00
55-59	4.00
60-64	4.00
65-69	3.00
70-74	2.00
75-79	1.00
80-84	0.50
85+	0.50

Chapter 3

Does a Part-Time School Principal Harm Students? Italian School Consolidation Process and Student Outcomes

Abstract

Schools are the main producers of human capital, which plays a key role in the process of economic growth. In this chapter we aim to assess the impact of the Italian secondary schools consolidation process on student outcomes. We exploit the fuzzy discontinuity in treatment regime and apply a continuity-based local polynomial approximation Fuzzy RDD. Estimation results show that the program exposure has had detrimental heterogeneous effects in upper secondary schools, especially for treated schools far away from the directional school, where the school principal is based. As a result, the isolation of the merged schools appears as a causality channel that makes the principals' responsibilities substantially aggravated. We conclude that the school consolidation process appears as a cost-saving policy that may fail to consider the indirect costs arising in term of poorer school performance, which may subsequently produce labor market inequalities due to losses in human capital accumulation.

1 Introduction

Human capital plays a crucial role in the process of economic growth (Lucas 1988, Barro 2001, Romer 1990), which has been shown to be closely related to the quality of schooling (Hanushek and Kimko, 2000). The underlying mechanism sees economies with more human capital innovating at a higher rate than those with less human capital, implying that societies with larger human capital in their workforce register more productivity gains and, as a result, a path of continued economic growth.

Since schools are the main "producers" of human capital, understanding the factors linked to variations in school performance becomes a key point to forward-thinking policy makers interested to trace economic development far into the future. In fact, many countries around the world seek to improve their schools aiming to enhance the skills and employability of their youth or to reduce inequalities in economic outcomes within their population (Hanushek and Woessmann, 2010).

Since 1990s, Italian institutional settings have changed and policies mostly oriented at containing costs have been the new rage. The school consolidation process (*Dimensionamento Scolastico*) belongs to this framework (Fontana, 2008). Indeed, new centralized criteria for the optimal schools size (both primary and secondary schools) have been introduced with the D.P.R. 233/1998, aiming at ensuring the optimal use of professional and instrumental resources. More specifically, to acquire or maintain legal personality, educational institutions must normally have a population between 500 and 900 students. Some exceptions are provided. In small islands, mountain municipalities, as well as in geographic areas characterized by ethnic or linguistic specificity, the school population can be reduced to 300 pupils. If schools do not reach the above mentioned thresholds, horizontal unification is planned, with other schools of the same grade included in the same territorial area, or vertical in comprehensive schools, according to the educational needs of the local area. Moreover, in some specific situations, a school where the position of the School Principal (SP hereafter) is left vacant, the Regional School Office can appoint a regent. What, in any case, the two schemes - unification and regency - have in common is the absence of a SP who is constantly based in the school. For this reason we adopted in this analysis the definition of "Part-Time School Principal", namely a SP that is in charge to guarantee simultaneously the functioning of more schools: the "directional school" where he or she was already based, plus one or more merged complexes. Hence, from a single merged school point of view, this policy design results in having a SP "borrowed" from the directional school that has to share his working time and effort among more than one schools. During the school year 2010/2011, 57.6% of total Italian schools were involved in the consolidation process, having then a shared SP.

The administrative cost reduction arising from the consolidation appear as a trivial result, meeting at least one of the policy change objectives related to ensure the optimal use of instrumental resources. What is neither trivial in this picture nor already assessed in the existing literature, is a clear understanding whether the use of professional resources is also optimal. More specifically, is the centralization of school management a win-win policy?

Since there is a growing attention regarding the role played by the SP in running the school (Bloom *et al.*, 2015), which is increasingly considered crucial, this study is aimed at investigating the impact of school consolidation process on student dropout behaviour.

In fact, Hanushek *et al.* (2008) investigated the underlying causes of dropping out of school and found that a student is much less likely to remain in school if attending a low quality school rather than a high quality school. Yet, mass-media and involved actor groups have voiced concerns that these mergers substantially aggravate the SP responsibilities and may have negative implications for the school performance, resulting in poorer student outcomes and, ultimately, less economic growth.

Due to the availability of data, we focus on Lower Secondary Schools (LSS henceforth) and Upper Secondary Schools (USS in what follows) during the school year 2010/2011. The student outcomes of interest are two different measures of dropout behaviour. They include voluntary dropouts, namely students that spontaneously stop attending classes during the year, and grade retention, measured by those students that are not admitted to the next class. They represent the "two sides of the same coin" since both of them hinder the expected school completion levels and give reason to concern about.

The analysis is performed both at school and cohort level.

The chapter is structured as follows. Section 2 provides details on the consolidation policy. Section 3 reviews the literature framework to which this analysis belongs. In section 4 we discuss the data. In section 5 the methodology. Section 6 presents the results and some robustness checks, section 7 concludes.

2 Institutional Settings

During the school year 2010/2011, the optimal size of educational institutions was governed by the D.P.R. 233/1998. Article 2 states: "The administrative, organizational, didactic and educational research and planning autonomy is recognized to the educational institutions of every order, including those already endowed with juridical personality, that reach dimensions suitable to guarantee the optimal balance between the demand and the organization of the education supply."

Further, the parameters are specified: "[...] To acquire or maintain legal personality, educational institutions must normally have a population, consolidated and predictably stable for at least five years, between 500 and 900 students; these indexes are taken as terms of reference to ensure the optimal use of professional and instrumental resources." Some exceptions are also provided. In small islands, mountain municipalities, as well as in geographic areas characterized by ethnic or linguistic specificities, the reference "[...] indexes can be reduced to 300 pupils for institutions including pre-school, primary and secondary schools, or upper secondary education which include courses or sections of different order or type [...]; in the above mentioned locations that are in conditions of particular isolation can also be constituted institutions inclusive of schools of every order and degree. The maximum index - 900 pupils - can be exceeded in areas of high population density, with particular regard to secondary educational institutions for training purposes that require structural assets, laboratories and workshops of high artistic or technological value." Moreover, in order to preserve the peculiarity of some North-Est local areas, schools with Slovenian teaching language will maintain the autonomy even in the absence of the minimum parameters of 300 pupils.

If schools do not reach the aforementioned reference indexes, horizontal unification is planned, with schools of the same grade included in the same territorial area, or vertical in comprehensive schools, according to the educational needs of the territory and local area planning.

To ensure the permanence in the territorial areas of schools that do not reach, alone or combined with schools of the same grade, optimal dimensions, educational institutions including kindergarten, primary and secondary schools are constituted, assuming the name of *Istituti Omnicomprensivi*. With the same purpose and to ensure the necessary variety of educational programs proposed by each institute and to meet the demand for education expressed by the school population, the unification of institutions of different order or type that do not reach, separately, the optimal dimensions is practicable. These institutions assume the name of *Istituti di Istruzione Superiore*.

The educational institutions merger plans are defined in provincial conferences for the organization of the school network, in compliance with the planning guidelines and general criteria previously adopted by the regions. The regions approve the regional merger plan, based on the provincial plans, by December 31 of the school year preceding the one to which the plan refers.

2.1 Regencies and Unifications

In some specific situations, a school where the position of the SP is left vacant, the Regional School Office can appoint a regent, that is a SP of another educational institution who is in charge to guarantee simultaneously the functioning of two or more schools.

The SP position may be vacant for several reasons:

- prolonged illness, leave or acceptance of other positions by the previous SP.
- Merger: if the number of students does not reach the thresholds set by the D.P.R. 233/98 and it is not possible to proceed to unification with other plexuses (e.g. because the total number of students after the unification would exceed the maximum threshold).[1]
- Retirement of the previous SP without the appointment of the new SP, regardless the number of students enrolled in the school.

The regency of the SP differs from the horizontal and vertical unifications for various reasons and concern the collaborators of the SP and the administrative apparatus. The regent SP may have a vicar both in the school where he is effective and in the school where he is a regent, the SP of an unified school has only one vicar.

A SP can however appoint collaborators up to 10% of the teaching staff, and often, in the unified plexuses, there is the figure of the "Director of Plexus".

The schools under regent maintain their own administrative apparatus, the unified schools do

[1] Only in this case, the Director of General and Administrative Services (DGAS) will be a regent as well.

not but, as already specified, only in case of regency for merger, the DGAS will also be a regent "borrowed" from another school. In all other cases of regency, the school will have its own secretarial offices.

Conversely, in the case of unification, a single administrative apparatus is envisaged with competence both on directional school and on unified plexuses. What, in any case, the two schemes - unification and regency - have in common is the absence of a SP who is constantly based in the directional school. For this reason we adopted in this analysis the definition of "part-time School Principal".

3 Literature Review

Our analysis contributes to two strands on the existing literature, one linked to the role of SP on school performance and pupils' outcomes and one related to the debate regarding the centralized versus decentralized provision of public services.

A large number of studies link high-quality leadership with positive school outcomes, such as student achievement (see Hallinger and Heck, 1998).

SPs may affect pupils outcomes through a variety of channels. In the Italian context, as school leaders, they are in charge to set up the school-level policy decisions. Although they have little influence on the composition of the school workforce, they still maintain considerable power on teacher supervision and retention, student discipline and student allocation to teachers and classes. They can also potentially have an impact on introducing new curricula and teaching techniques (Coelli and Green, 2012) and are responsible for monitoring the quality of instruction delivered by teachers (Clark *et al.*, 2009). Moreover, since teachers quality matters to student achievements, as pointed out by a voluminous literature (see, among others, Rivkin *et al.* 2005, Aaronson *et al.*2007, Konstantanopoulos 2007, Kane *et al.* 2008, Koedel 2008 and Leigh 2010), the role of SP becomes crucial as long as he or she can have an impact on teachers composition and performance.

Further, Grissom (2011) finds that SP effectiveness is associated with greater teacher satisfaction and a lower probability that the teacher leaves the school within a year. The impact of principal skills on retaining teachers is significant since high rates of teacher turnover negatively impact school performance through several channels: high turnover means greater school instability, disruption of curricular cohesiveness and a continual need to hire inexperienced teachers, who typically are less effective, as replacements for teachers who leave. Moreover, the positive impacts of principal effectiveness on teacher satisfaction and lower turnover is even greater in disadvantaged schools. These findings support those policies focused on getting the best principals into the most challenging school environments.

Bloom *et al.* (2015) answer the question whether the managerial practices are related to meaningful educational outcomes by means of a cross-sectional analysis and survey methodology over 1,800 schools randomly sampled across eight countries. They constructed a management quality index starting from information on management practices adoption collected by double-blind telephone interviews with school principals. More in details, the survey investigated the adoption of 20 basic management practices, where the level of adoption was evaluated through a scoring grid that ranged from one, indicating a "worst practice", to five, signaling a "best practice".The index was then a result of an average across those 20 management practice measures in four areas of management: operations, monitoring, target setting and people. They find that one standard deviation increase in the abovementioned managerial index is associated with a 0.232 to 0.425 standard deviation increase in students' outcomes, namely educational achievements involving standardize and non-standardize examination results.

Moreover, higher principal turnover is also found to occur in low achieving schools and the switching SP are more likely to move to higher performing schools (Cullen and Mazzeo, 2008, Branch *et al.*, 2012 and Miller, 2009).

Further, Clark *et al.* (2009), present evidence on the relationship between principal characteristics and school performance using data from New York City Department of Education. They find that schools perform better when they are led by experienced principals, particularly for math test scores and student absences. Thus they confirm the well known learning by doing principle: workers performing more tasks will become more productive with experience at the task, especially when the task is as demanding as running a school. They also warn about the high potential cost in term of policy implications of having experienced principals leave their jobs, and provide evidence on the benefits associated with retaining experienced principals. As a result, their findings alert district administrators to the distributional consequences arising from higher rates of principal turnover in disadvantaged schools.

Grissom and Loeb (2011) try to add a step forward on the literature that recognizes that SPs affect school outcomes, by shedding light on how principals influence these outcomes, namely the specific skills that principals need to achieve school success. They use principals' self-assessments on 42 tasks to distinguish five effectiveness dimensions (Instruction Management, Internal Relations, Organization Management, Administration, and External Relations).[2] In their findings, only the principals' Organization Management skills, consistently predicts students' math and reading achievement gains.[3]

Further, Coelli and Green (2012), using the method that Rivkin *et al.* (2005) used to analyze teacher impacts, estimated the effect of individual SPs on student high school graduation probabilities and grades from the Canadian province of British Columbia. More specifically, they are able to isolate the effect of principals from the effect of schools thanks to the SP rotation across schools occurred by districts. They find that a principal who is one standard deviation better in the principal effects distribution is associated with higher graduation rates by 2.6 percentage points. The authors also hypothesize that the effect of principals may not be immediately evident but it may take numerous years for them to exert a significant impact on student outcomes. In particular, in a school where the principal turns over every year, he or she may not be able to have a full sizable effect on the school, finding that English exam scores will be higher by 2.5 percentage points if principals are given time to fully " leave their mark" on the school.

Being a school principal is a stressful job, and many school districts find difficult to attract quality applicants and to keep successful principals in their jobs. Coelli and Green results confirm that public policies should make an effort to retain good school principals.

Furthermore, the longstanding problem in public finance regarding the trade-off between centralized and decentralized provision of public services has been of interest to public economists (see e.g., Boffa *et al.* 2015, Besley and Coate 2003), and it has been applied to the educational field as well (Galiani *et al.*, 2008). Since it has already been the focus of a deep review in the second chapter of thesis, we refer the reader to section 3 of chapter 2 for a broader analysis.

[2] Instruction Management is related to the set of tasks principals conduct to support and improve the implementation of curricular programs. Internal Relations is linked to the capability of principals to building strong interpersonal relationships within the school. Organization Management refers to the set of tasks principals exert to oversee the functioning of the school and to pursue the school's medium and long-term goals. Administration regards more routine administrative duties and tasks executed to comply with state or federal regulations. External Relations addresses the capability to work with stakeholders outside the school.

[3] Math and reading scores from the Florida Comprehensive Assessment Test (FCAT) during the years 2007 and 2008 at the student-level.

However, to facilitate a clear reading of the present analysis, we provide here a brief recap. The standard approach, formalized by Oates (1972), assumes that the main drawback with a centralized system is that it produces a “one size fits all” provision of public goods and services which does not reflect, and eventually satisfy, local needs. While centralization can exploit economies of scale and scope better, and profit from policy coordination - which becomes very important in presence of spillovers - local bureaucrats are believed to have more information on the local specific needs than the upper-tier governments and many of the public goods and services needed by citizens are often community- and site-specific. This argument can be easily translated into educational environment where, on one hand, the needs of students may be school-specific and educational strategies require to be tailored accordingly; on the other hand, an upper-tier school management can have economies of scale in designing curricula and in prescribing and enforcing minimum quality standards (Bardhan, 2002).

4 Data

We focus on Italian Secondary Schools, both lower and higher order. We link multiple sources of administrative micro-data at the school level mainly issued by the Ministry of Education and available for the school year 2010/2011, through the project “Scuola in Chiaro”. Our dependent variables are proxies for student failure and include student dropouts, namely students that voluntarily stop attending classes during the year, and students that are not admitted to the next class.[4] The analysis is made at school level with regard to both dependent variables. Additionally, student dropouts data are available at cohort level as well, and we therefore conduct the analysis with this greater level of disaggregation only for dropouts. We also allow for heterogeneous effects by distinguishing between the geographical characteristics of the schools, driving distance from the directional school (where the SP is based), and USS typology. The available data on the Register of Italian Schools contain the unique national identification code (*Codice Meccanografico*) and the address for all school complexes with also the indication of

[4]Eurostat uses the definition of “early school leavers” to indicate “[...]the population aged 18 to 24 having attained at most lower secondary education and not being involved in further education or training”. Likewise, a piece of American literature employs the term “dropouts” to designate “young people who leave school without gaining a high school diploma” (Lamb and Markussen, 2011).

their directional school address. By comparing the addressed of the plexus and the directional school, we are able to identify treated (different address, SP not in place) to controls (directional schools, SP in place). Further, we bring together the information regarding the regencies by means of Regional decrees of Regency Assignment. [5]

Moreover, given the availability of the addresses for both the directional schools and their merged complexes, we requested Google Maps Platform, through its Directions API, to calculate the shortest driving distance in Kilometers between each pair of directional schools and any of their merged complexes.

Our dataset includes a number of schools characteristics for 11,402 Italian secondary schools. We do not include in the sample those located in the regions of Valle d'Aosta and Molise, and provinces Trento and Bolzano.[6] We also exclude 489 schools that have been exposed to the policy change in the school year after (2011/2012) to the one observed (2010/2011) and knew about their next policy change by December 2010, as pointed out in section 2. This allows us to get rid of potential anticipation effects that may occur when there is a time gap between the policy announcement and the actual policy implementation. Table 3.1 details the variables with a brief description and the data sources.

4.1 Descriptive Statistics

The following figures aim at depicting the students dropouts phenomenon in the observed school year. Figure 3.1 and 3.2 give an overview on the dropouts, at the national average, per grade on lower and upper secondary schools. As expected, dropouts are higher in USS, which show a wider variability by grade. Indeed, while in the LSS the educational offers are homogeneous and the choice of school mostly depends on the geographical proximity to the school buildings, a numerous of different choices of USS is instead available to students. A poor orientation system may lead students to enroll in a school not suitable for their preferences or capabilities and then change school type during the first or second year, corresponding to the 9th and 10th grade.

The higher level of dropouts during the 4th year of USS, corresponding to the 12th grade, may

[5] The detailed list of normative references is not reported here for the sake of brevity but is available from the author.

[6] Valle d'Aosta, Trento and Bolzano share a status of autonomous provinces and their data were not included in the national database. We were not able to find the Regional Decrees of Regency Assignment for the region Molise and, due to the impossibility to distinguish between under regent and not under regent schools, it is excluded from the sample.

Table 3.1: Data Description

Dependent Variables	Brief description	Source
Dropouts grade 6	Percentage of early leavers in the 6 grade (lower secondary school)	Miur- Scuola in Chiaro
Dropouts grade 7	Percentage of early leavers in the 7 grade (lower secondary school)	Miur- Scuola in Chiaro
Dropouts grade 8	Percentage of early leavers in the 8 grade (lower secondary school)	Miur- Scuola in Chiaro
Dropouts grade 9	Percentage of early leavers in the 9 grade (upper secondary school)	Miur- Scuola in Chiaro
Dropouts grade 10	Percentage of early leavers in the 10 grade (upper secondary school)	Miur- Scuola in Chiaro
Dropouts grade 11	Percentage of early leavers in the 11 grade (upper secondary school)	Miur- Scuola in Chiaro
Dropouts grade 12	Percentage of early leavers in the 12 grade (upper secondary school)	Miur- Scuola in Chiaro
Dropouts grade 13	Percentage of early leavers in the 13 grade (upper secondary school)	Miur- Scuola in Chiaro
Total dropouts per school	Sum of overall early leavers over total number of students	Our elaboration on MIUR data
Not admitted to the next class	Percentage of students not admitted to the next class per school	MIUR - Scuola in Chiaro
Treatment indicator		
Part - time SP (merged and/or under regent)	1 if merged and/or under regent, 0 othewise	Our elaboration on various sources
Running variable		
Total students per school	Sum of overall students per school in the school year 2010/2011	MIUR - Scuola in Chiaro
Other schools characteristics		
Number of complexes	Number of merged complexes per school	Our elaboration on MIUR data
School level	1 if upper secondary school, 0 if lower secondary school	MIUR - Scuola in Chiaro
Driving distance	Driving distance in Km of the shortest road linking merged schools	Google Direction API
USS School typology	1 if Lyceum, 2 if Technical Institute, 3 if Professional Institute	Our elaboration on MIUR data
Mountain municipality or small island	1 if the school is placed in mountain municipalities or small islands, 0 otherwise	Our elaboration on MIUR data

Source: The author

depends on the fact that, in Italy, school attendance is compulsory until the age of 16 years old. As a result, since students on the 12th grade should be at least 17 years old, no law obliges them to stay in the classroom any longer.

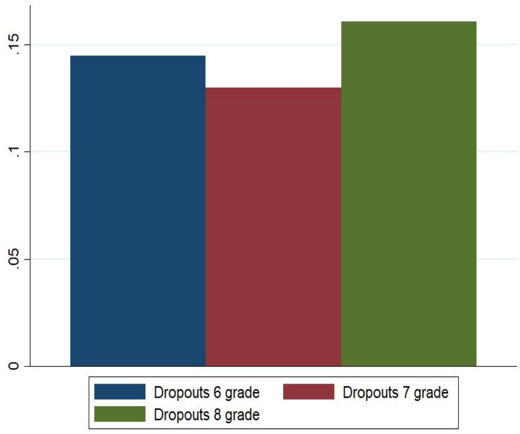

Figure 3.1: Dropouts (average) per grade in LSS. Source: Author graph on MIUR - "Scuola in Chiaro" data

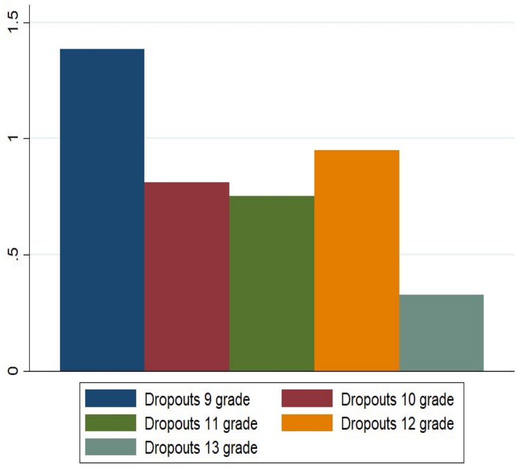

Figure 3.2: Dropouts (average) per grade in USS. Source: Author graph on MIUR - "Scuola in Chiaro" data

Figure 3.3 plots the level of dropouts by treatment status. At a first glance, dropouts appear higher in treated schools.

Figure 3.4 details the level of dropouts by treatment status, taking into account the intensity of treatment. More specifically, the dropouts levels are plotted against the number of school merged complexes. In fact, we expect the more the school complexes that one single SP has to manage, the higher the complexity of his responsibilities. The picture actually shows higher levels of dropouts as the number of complexes increases but this does not apply on the right tale of the complexes distribution, where, however, we observe only a small number of observations that are characterized by such high number of merged complexes.

Figure 3.5 and 3.6 give information of the relationship between three variables: the level of dropouts, the treatment status and the forcing variable, namely the number of overall students enrolled in the given school. Both of them suggest that dropouts look higher in small and treated schools.

Table 3.2 reports the main descriptives for the LSS. Table 3.3 outlines the summary statistics for the USS. Both of them distinguish between treated and untreated (control) schools. We perform t tests on the equality of means for each dependent variable by treatment status. Results are

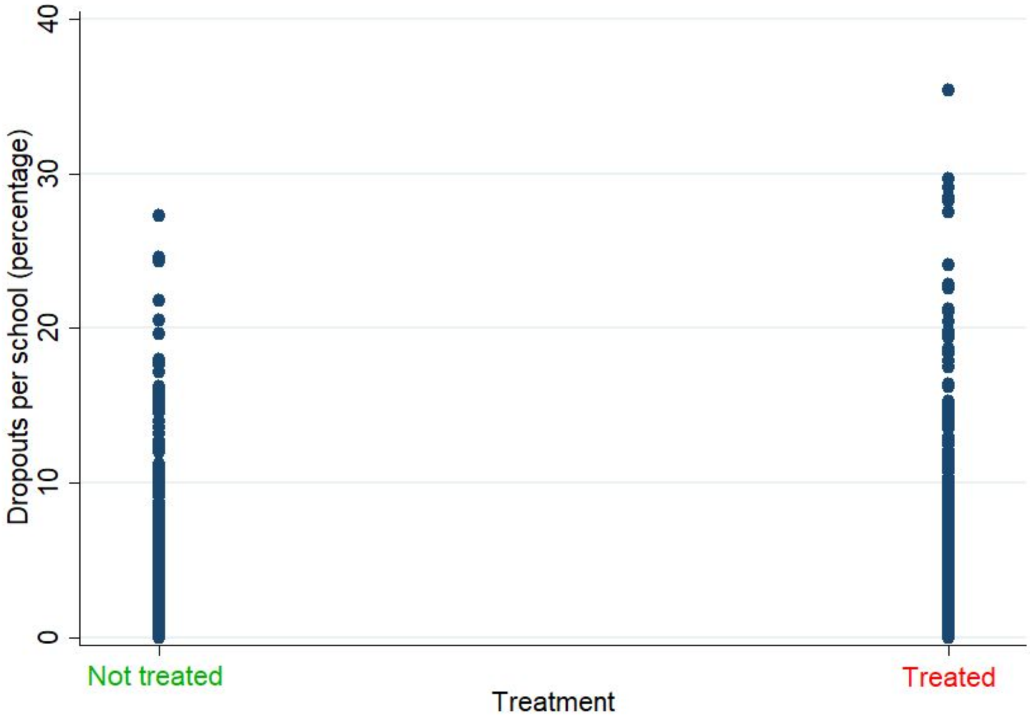

Figure 3.3: School dropouts Vs treatment (merged and/or under regent). Source: Author graph

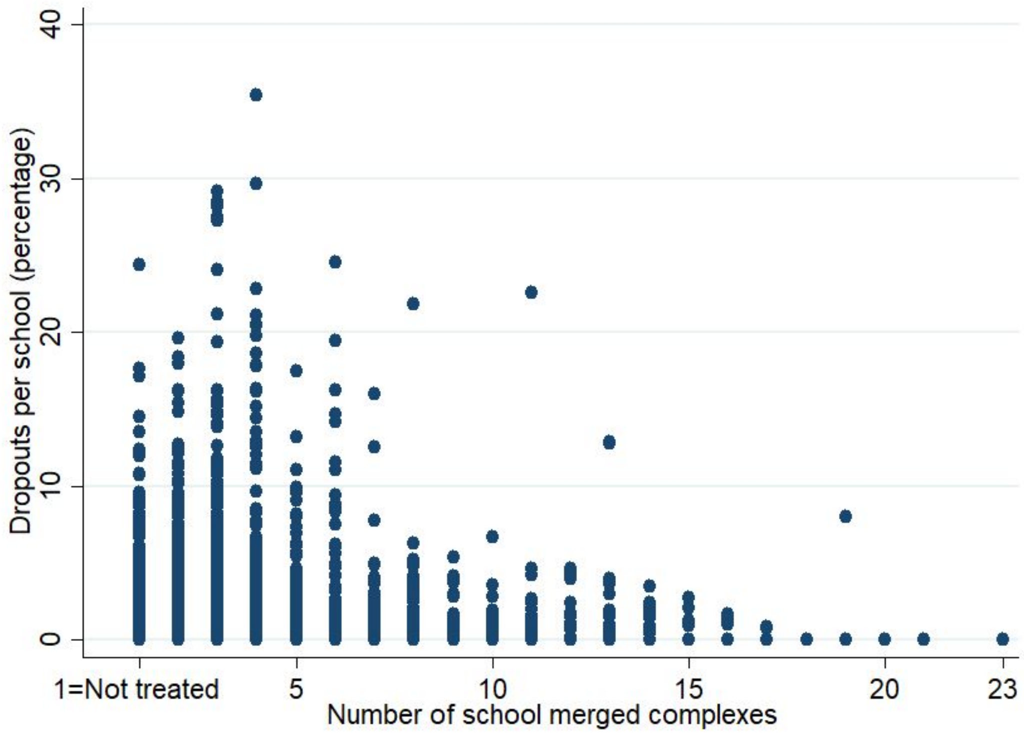

Figure 3.4: School dropouts Vs number of merged school complexes. Source: Author graph

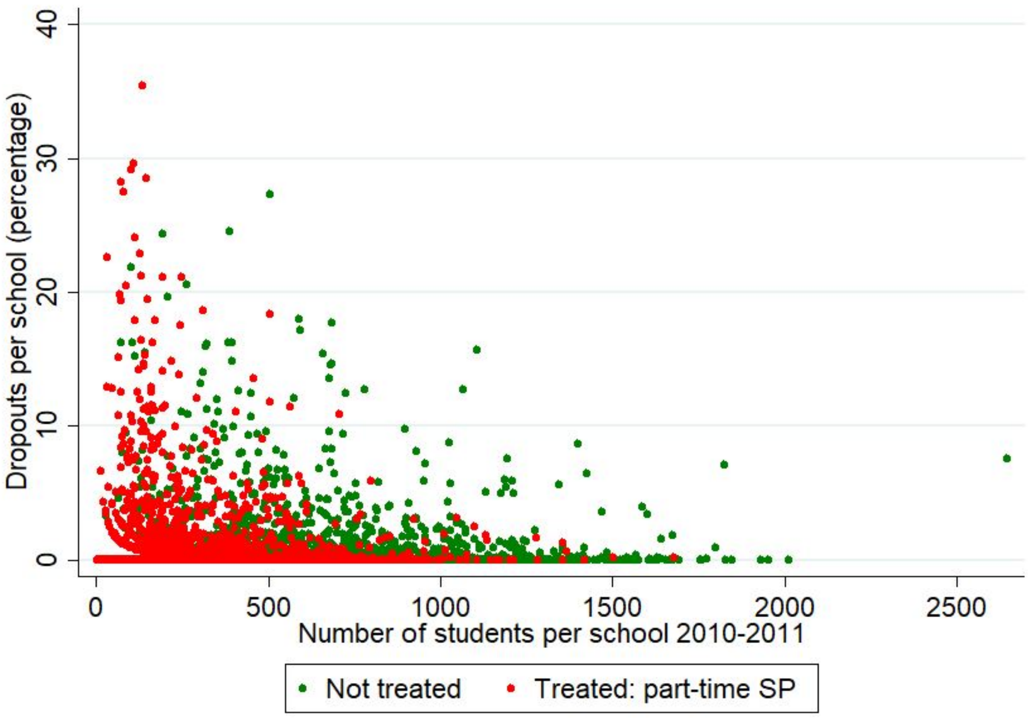

Figure 3.5: School dropouts Vs. number of students and treatment. Source: Author graph

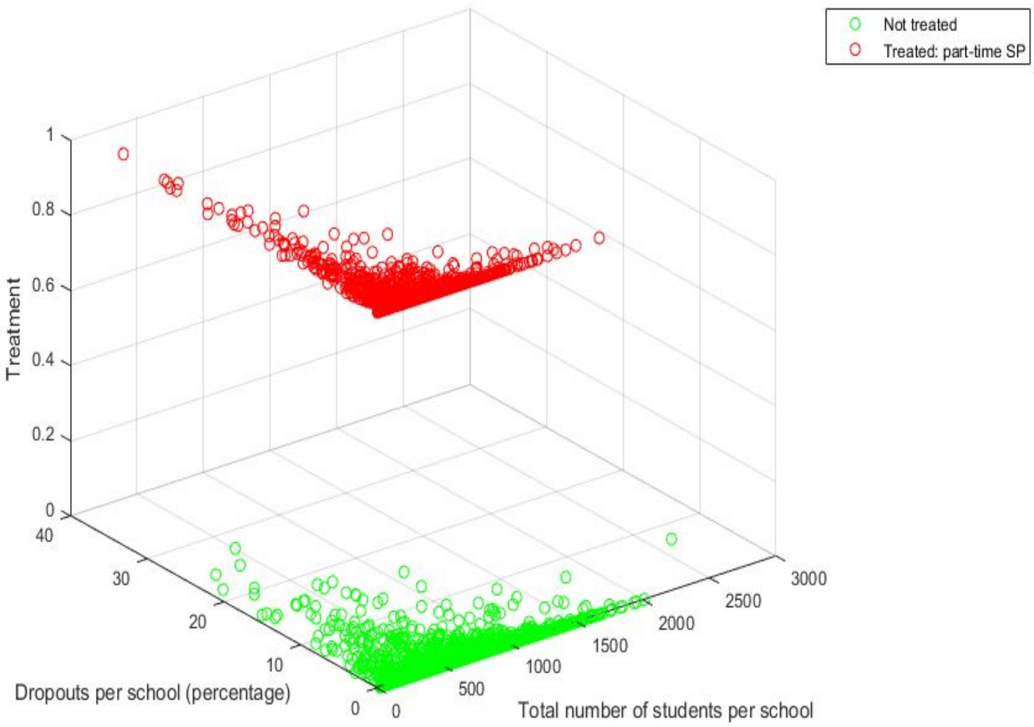

Figure 3.6: School dropouts Vs. number of students and treatment: 3D scatterplot. Source: Author graph

never statistically significant and do not allow to reject the null hypothesis of equality.

Table 3.2: Summary statistics: Untreated vs treated LSS

Variable	Mean	Std. Dev.	Min.	Max.	N
Untreated					
Dropouts grade 6th	0.084	0.421	0	6.100	1710
Dropouts grade 7th	0.083	0.5	0	13.6	1710
Dropouts grade 8th	0.093	0.438	0	6	1712
Total dropouts per school	0.086	0.281	0	4.603	1707
Not admitted to the next class per school	4.799	4.357	0	68.3	1714
Treatment: part-time SP	0	0	0	0	1714
Total number of students enrolled	412.579	236.343	12	1149	1714
Driving distance from directional school	0	0	0	0	1714
Mountain municipality or small island	0.153	0.361	0	1	1714
Treated					
Dropouts grade 6th	0.078	0.735	0	27.3	4916
Dropouts grade 7th	0.069	0.598	0	25	4900
Dropouts grade 8th	0.108	0.727	0	18.2	4927
Total dropouts per school	0.087	0.453	0	12.903	4862
Not admitted to the next class per school	4.538	4.313	0	40	4952
Treatment: part-time SP	1	0	1	1	4952
Total number of students enrolled	175.203	137.492	2	910	4951
Driving distance from directional school	7.456	48.844	0	1233.191	4491
Mountain municipality or small island	0.299	0.458	0	1	4952

Source: Author elaboration

Table 3.3: Summary statistics: untreated vs treated USS

Variable	Mean	Std. Dev.	Min.	Max.	N
Untreated					
Dropouts grade 9th	1.18	4.052	0	47.4	2376
Dropouts grade 10th	0.701	2.529	0	43.3	2382
Dropouts grade 11th	0.675	2.494	0	41.2	2380
Dropouts grade 12th	0.773	2.924	0	53.8	2373
Dropouts grade 13th	0.31	1.313	0	17.6	2364
Total dropouts per school	0.776	2.379	0	27.262	2329
Not admitted to the next class per school	10.915	7.76	0	60	2406
Treatment: part-time SP	0	0	0	0	2407
Total number of students enrolled	653.129	355.433	10	2648	2407
Driving distance to directional school	0	0	0	0	2407
Mountain municipality or small island	0.128	0.334	0	1	2407
Treated					
Dropouts grade 9th	1.463	5.266	0	59.1	2230
Dropouts grade 10th	0.909	3.727	0	52	2208
Dropouts grade 11th	0.818	3.32	0	68.8	2204
Dropouts grade 12th	1.102	4.418	0	56.3	2163
Dropouts grade 13th	0.35	1.961	0	36.8	2146
Total dropouts per school	0.985	3.106	0	35.419	2009
Not admitted to the next class per school	12.743	9.004	0	73.7	2328
Treatment: part-time SP	1	0	1	1	2329
Total number of students enrolled	281.445	214.854	9	1679	2328
Driving distance to directional school	8.041	44.491	0	1016.51	2059
Mountain municipality or small island	0.216	0.412	0	1	2329

Source: Author elaboration.

4.2 Graphical Analysis at the Cutoff

The following figures provide a graphical analysis aimed at showing intuitively and transparently any jumps across the threshold in the running variable.

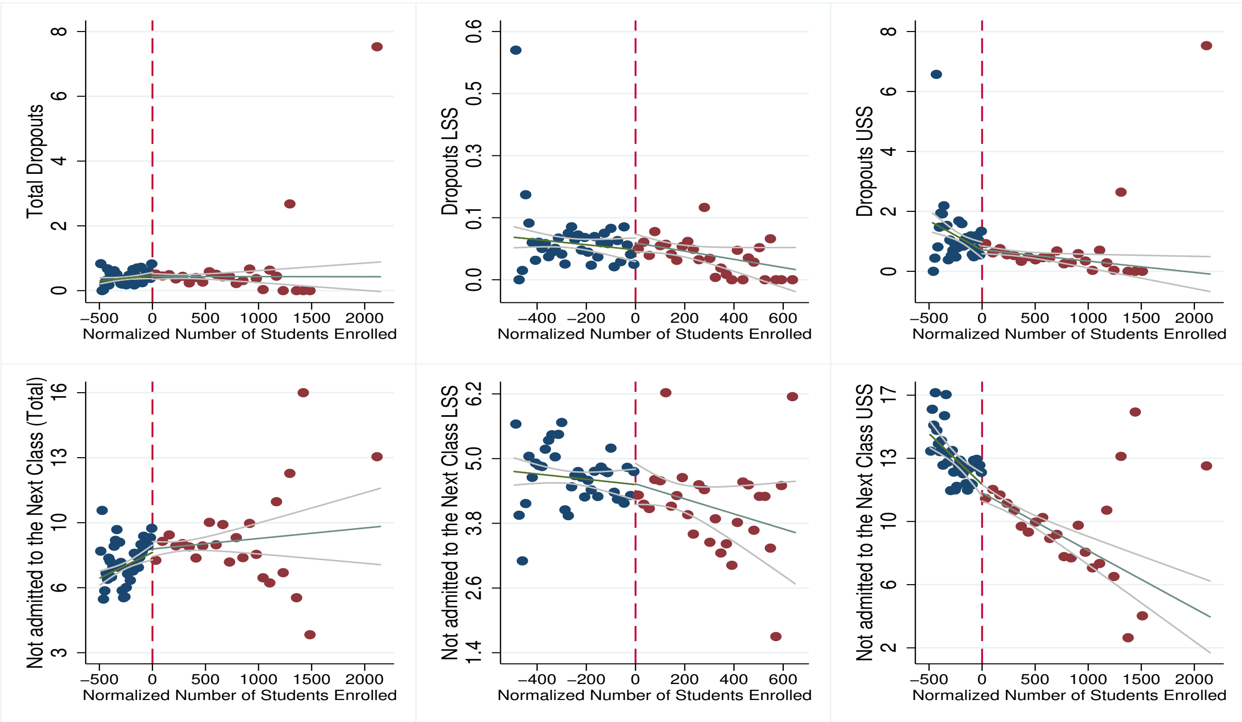

Figure 3.7: Dropouts and Not Admitted to the Next Class. Source: Author graph

Figure 3.7 explores the discontinuity around the threshold for the two outcomes of interest, both at the aggregate level for all schools and by school level. Figure 3.8 disentangles the volumes of dropouts by grade. Figure 3.9 studies the compliance of the treatment to the merging rule, showing a clear discontinuity at the cutoff. Although the level of compliance to the rule look high, the figure provide evidence of the occurrence of some fuzziness in the treatment compliance. Figure 3.10 studies the discontinuity with respect to the driving distance in kilometers from the directional school to the merged complexes, confirming the evidence on the great level of compliance to the merging rule.

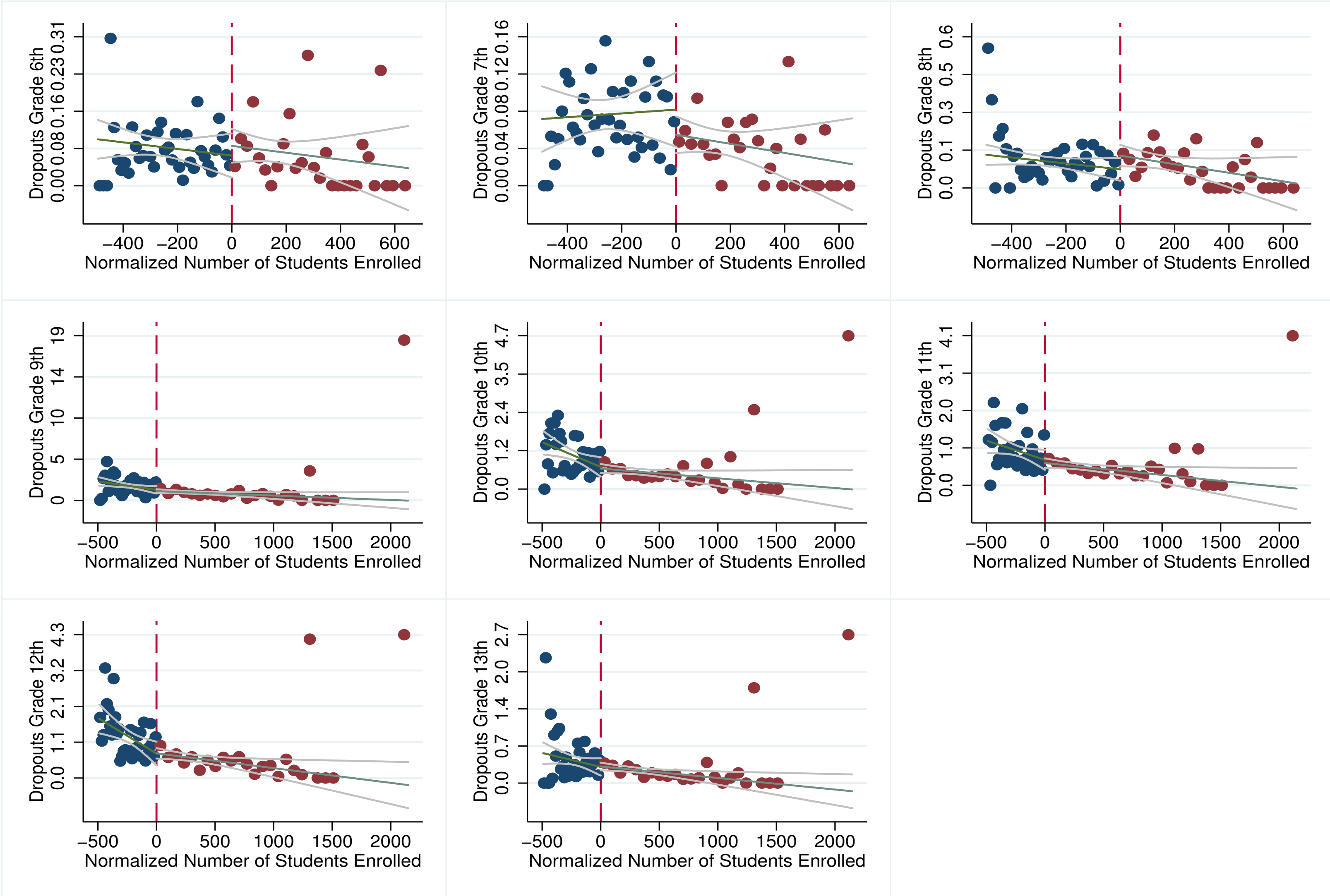

Figure 3.8: Dropouts per Grade. Source: Author graph

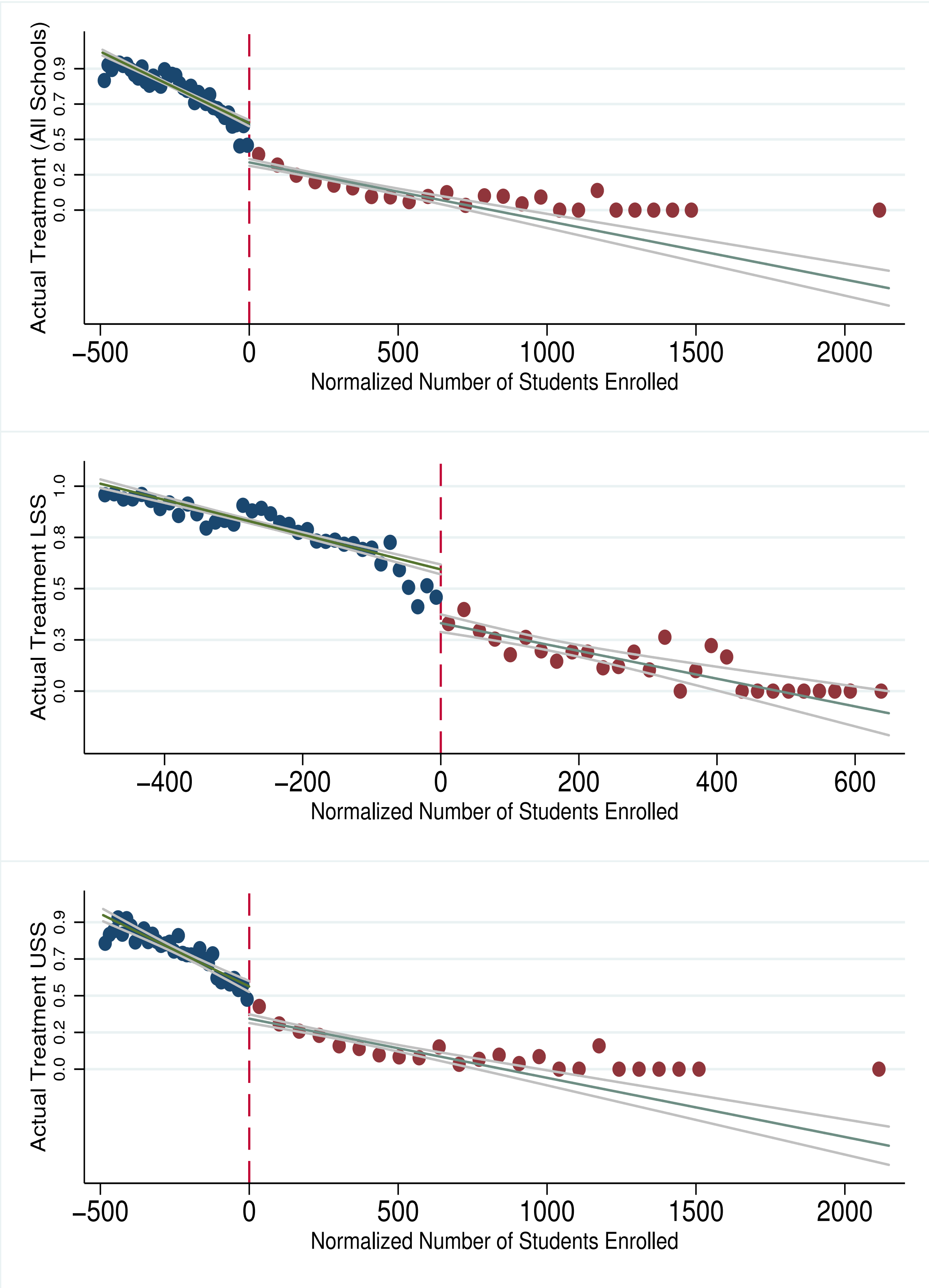

Figure 3.9: Treatment Compliance. Source: Author graph

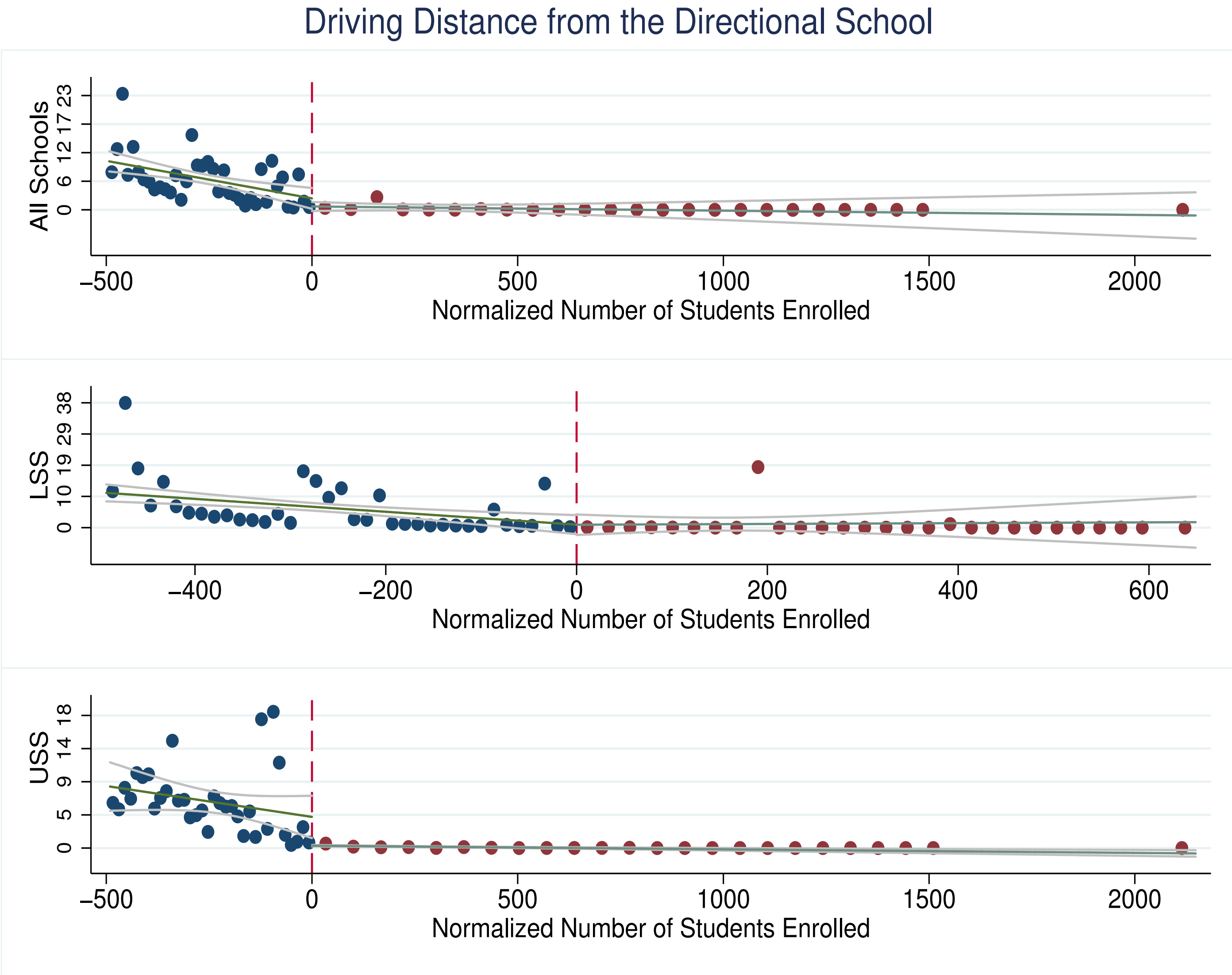

Figure 3.10: Driving Distance (Km) from the Directional School to the Merged Complexes. Source: Author graph

5 Identification Strategy

To assess the consolidation policy impact on student outcomes, we implement a continuity-based local polynomial approximation Fuzzy Regression Discontinuity Design (fuzzy RDD hereafter). In the general RDD framework, all units in the sample receive a score, also known as running or forcing variable, and a treatment is assigned to those units whose score is above a known cutoff, while those units whose score is below the cutoff are used as the counterfactual (Hahn *et al.*, 2001).

Fuzzy RDD is a quasi-experimental method in which the probability of receiving treatment changes discontinuously across the threshold as a function of a running variable (Dai *et al.* 2018), and the compliance with treatment assignment is no longer perfect, as in the Sharp design. [7]

Since the eligibility rules were not perfectly enforced in the Italian school case, and some ineligible schools received the program while some eligible schools failed to receive it, the program eligibility assigned discontinuously based on a score and a cutoff, and the imperfect compliance with eligibility status, makes this policy evaluation design suitable for a Fuzzy RDD where the running variable is the number of students enrolled in a given school, and the treatment is being a merged and/or under regent complex.

Treated and control schools may be different. As suggested by Cattaneo *et al.* (2017), if some of the characteristics on which they differ are correlated with the outcomes of interest, a simple comparison of the two groups will be misleading. For example, it seems reasonable to assume the occurrence of a selection process into the treated group: the "worst schools" may not be chosen by new students and therefore experience a decrease in the number of enrolled, reaching ultimately the merger program. The "worst schools" therefore have a greater risk of ending up in the treated group and a possible increase in their dropouts levels would not only depend on the policy implementation but on the pre-existing limited quality of the schools.

In such situations, a Regression Discontinuity (RD) strategy can be used to isolate a treatment effect of interest from all other systematic differences between treated and control groups. In other words, the observations just above and just below the cutoff will tend to be comparable in

[7] In contrast to Sharp RD, in a Fuzzy RD design, the change in the probability of being treated at the cutoff is always less than one.

terms of all characteristics, and point estimates get rid of systematic observed and unobserved differences between the groups.

As suggested by Calonico *et al.* (2018), modern RD application employ a low-order polynomial approximation (usually linear or quadratic) only around the cutoff, ruling out all other observations further away.[8]

Local polynomial methods estimate the chosen polynomials using only observations near the cutoff, separately for control and treatment observations. More formally, this approach uses only observations in the neighborhood of the threshold, essentially between $\bar{x}_{-}$ h and $\bar{x}_{+}$ h, where h is the given bandwidth. Moreover, within this bandwidth, a kernel function k(.) assigns a positive increasing weight to the observations depending on their proximity to the cutoff $\bar{x}$.[9]

Following Calonico *et al.* (2014, 2017), we choose the *p-th* order local polynomial estimator with the *q-th* order local polynomial bias correction, with triangular kernel function that assigns a proportionally higher weight to the observations closer to the cutoff. We also compute the optimal fuzzy RDD data-driven bandwidth for our sample.[10]

Yet, this approach can be viewed formally as a nonparametic local polynomial method, in which the fit is taken as an approximation of the unknown underlying regression functions within the bandwidth used, with a potential misspecification of the functional form of the regression function.

In the policy change focus of this chapter, as common in practice, the treatment is assigned using different cutoff values for different subgroups of observations. In fact, as pointed out in section 2, while the cutoff is normally 500, it is reduced at 300 students in schools located in mountain areas and small islands. Given this multi-cutoff RD design, we follow the commonly

[8] Gelman and Imbens (2018) give some cautionary advices against using higher-order polynomial in RD designs since they might inappropriately put large weights on observations further away from the cutoff. Higher-order polynomials tend to produce over-fitting of the data and thus unreliable results near boundary points. Local constant fits ($p = 0$) exhibit some undesirable theoretical features and usually under-fit the data. In practice, the recommended choices are usually $p = 1$ or $p = 2$, though theory and practice considers other polynomial orders as well (Cattaneo *et al.*, 2017).

[9] A kernel function may be triangular (assigning a positive weight to all observation within the bandwidth h but declining symmetrically and linearly as the value of the forcing variable gets farther from the cutoff, becoming zero outside h), uniform (which gives zero weight to all observations with score outside the bandwidth but equal weight to all the observations within h) and Epanechnikov (which gives a quadratic decaying weighting to observations within h and zero weight to the rest)(Cattaneo *et al.*, 2017).

[10] We implement this two-step procedure with the the Stata routine `rdrobust` which reports the correct standard errors. We run bandwidth selection with the Stata command `rdbwselect`, with the option `msetwo` for MSE - optimal point estimation using a different bandwidth on the two sides of the cutoff.

used approach and normalize (or center) the running variable at zero, such that all schools face the same common cutoff value, and then apply the standard Fuzzy RDD routines to the normalized score. However, we also consider the two different group of schools, and then each cutoff, separately, to detect potential heterogeneous effects.

In practice we estimate the following two equations. In the first-stage equation the outcome variable is a dummy that equals value 1 if the school is actually treated, 0 otherwise:

$$ActualTreatment_i = \delta_0 + \delta_1 PredictedTreatment_i + \delta_2(a) + \epsilon_i, \tag{3.1}$$

where *Predicted Treatment* denotes the treatment eligibility established by the D.P.R. 233/1998 and reviewed in section 2 ($PredictedTreatment = 1(a < 0)$). $\delta_2(a)$ is a polynomial of the running variable in the continuous measure of number of students enrolled in a given school.[11]

In the second equation we estimate:

$$y_{ji} = \gamma_0 + \gamma_1 ActualTreatment_i + \gamma_2(a) + u_i, \tag{3.2}$$

where y_{ji} are the set of j dependent variables outlined in section 4 for the school i, $\gamma_2(a)$ is again a continuous function of the running variable and the $ActualTreatment_i$ is the treatment status. Estimation results follow in the next session.

One potential concern on the validity of the chosen method is the continuity of the forcing variable at the cutoff. In our setting, where the running variable is the number of students in a given school, this means that schools cannot systematically manipulate the number of students enrolled. Following Cattaneo *et al.* (2017), we test the continuity of the normalized forcing variable by means of a number density tests that are reported in table 3.11 in appendix, along with the estimated density plots and histograms. The value of the statistic used to test is 0.04913 and the associated p-value is 0.6232, for all schools of the sample. This suggests that, under the continuity-based approach, we fail to reject the null hypothesis of no difference in the density of treated and control schools at the cutoff, consistent with the identification assumption

[11] We do not include covariates because the available data on other schools characteristics would lead to a decrease in the sample size. However, despite the inclusion of covariates is intended to increase the precision of the RD estimates, Calonico *et al.* (2018) note that there is no existing justification for using additional covariates for identification, estimation, or inference purposes, employing only continuity/smoothness conditions at the cutoff. They also highlight the fact that this has led, in empirical practice, to the proliferation of *ad hoc* covariate adjustments that may reduce the transparency of the estimation and result in noncomparable, or even inconsistent, estimators.

that treatment assignment around the discontinuity was "as good as random" (Amarante *et al.*, 2016). Figure 3.11 and 3.12 provide a graphical representation of the continuity in density tests, exhibiting both the estimated density (fig. 3.11) and the histogram of the data (fig. 3.12). We run the same tests distinguishing between the school levels. We find no discontinuity in the density of treated and control USS at the cutoff, while results for the subsample of LSS provide evidence of a sort of manipulation at 10% level of confidence.

6 Results

Obtained estimation results show that the treatment status hardly affected students outcomes. In fact, table 3.4 and 3.5 suggest that program exposure is scarcely associated to statistically significant variation in students dropouts and students that fail to pass to the next class. In fact, only the percentage of students not admitted to the next class looks higher by 24.94 percentage points in the treated USS schools, although this results is not robust to different choices of local polynomial order.[12]

Table 3.6 and 3.7 explore the impact of the treatment exposure on student dropouts by cohort and school type. Only in the 8th grade the merger policy seems to play a role in diminishing the dropouts phenomenon. However, this result is not robust to different choice of local polynomial order and may be biased by the evidence of manipulation given by the density test of the running variable for the LSS reported in table 3.11 in appendix.

6.1 Heterogeneous Treatment Effects

A common question of interest in a policy evaluation exercise is "which groups are the most affected by the treatment exposure?" A standard way to address this inquiry is to look at heterogeneity in treatment effects with respect to the baseline sample characteristics.

We first look at potential differences with regard to geographical school characteristics, which imply different cutoffs on the running variable as well. We therefore conduct the estimation exercise only for schools in mountain municipalities and small islands, whose cutoff determining

[12]Even though results are not reported here, uniform and epanechnikov kernel functions have been tried too. Results do not change.

Table 3.4: Student dropouts by school type: LSS and USS. Fuzzy RD estimates.

	(1)	(2)	(3)
VARIABLES	Dropouts LSS	Dropouts LSS	Dropouts LSS
RD Estimate	-0.0642	-0.547	-0.499
Robust 95% CI	[-.538 -.192]	[-3.439 -1.483]	[-2.911 - 1.99]
Kernel Type	Triangular	Triangular	Triangular
Observations	6569	6569	6569
Conventional Std. Error	0.142	1.133	1.147
Conventional p-value	0.650	0.629	0.663
Robust p-value	0.353	0.436	0.712
Order Loc. Poly. (p)	0.000	1.000	2.000
Order Bias (q)	1.000	2.000	3.000
	(1)	(2)	(3)
VARIABLES	Dropouts USS	Dropouts USS	Dropouts USS
RD Estimate	1.532	7.942	33.79
Robust 95% CI	[-2.1 - 6.541]	[-20.234 - 47.991]	[-494.97 - 636.006]
Kernel Type	Triangular	Triangular	Triangular
Observations	4338	4338	4338
Conventional Std. Error	1.759	15.225	259.022
Conventional p-value	0.384	0.602	0.896
Robust p-value	0.314	0.425	0.807
Order Loc. Poly. (p)	0.000	1.000	2.000
Order Bias (q)	1.000	2.000	3.000

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors clustered at the province level.

Table 3.5: Non admitted to the next class by school type: LSS and USS. Fuzzy RD estimates.

	(1)	(2)	(3)
VARIABLES	Not-admitted LSS	Not-admitted LSS	Not-admitted LSS
RD Estimate	1.738	10.77	-4.789
Robust 95% CI	[-4.782 - 11.273]	[-36.254 - 78.236]	[-37.693 - 20.125]
Kernel Type	Triangular	Triangular	Triangular
Observations	6665	6665	6665
Conventional Std. Error	3.526	26.433	13.554
Conventional p-value	0.622	0.684	0.724
Robust p-value	0.428	0.472	0.551
Order Loc. Poly. (p)	0.000	1.000	2.000
Order Bias (q)	1.000	2.000	3.000
	(1)	(2)	(3)
VARIABLES	Not-admitted USS	Not-admitted USS	Not-admitted USS
RD Estimate	24.94**	84.06	164.9
Robust 95% CI	[4.358 - 80.545]	[-238.347 - 451.248]	[-1174.776 - 1429.352]
Kernel Type	Triangular	Triangular	Triangular
Observations	4733	4733	4733
Conventional Std. Error	17.416	153.117	596.655
Conventional p-value	0.152	0.583	0.782
Robust p-value	0.029	0.545	0.848
Order Loc. Poly. (p)	0.000	1.000	2.000
Order Bias (q)	1.000	2.000	3.000

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors clustered at the province level.

Table 3.6: Student dropouts by cohort: LSS

	(1)	(2)	(3)
VARIABLES	Dropouts 6th grade	Dropouts 6th grade	Dropouts 6th grade
RD Estimate	0.146	0.788	0.756
Robust 95% CI	[-.248 - .798]	[-1.706 - 4.114]	[-1.715 - 3.131]
Observations	6625	6625	6625
Conventional Std. Error	0.219	1.404	1.156
Conventional p-value	0.505	0.574	0.513
Robust p-value	0.302	0.417	0.567
VARIABLES	Dropouts 7th grade	Dropouts 7th grade	Dropouts 7th grade
RD Estimate	0.0630	-0.0189	0.130
Robust 95% CI	[-.335 - .475]	[-2.313 - 2.001]	[-3.808 - 3.943]
Observations	6609	6609	6609
Conventional Std. Error	0.158	0.933	1.764
Conventional p-value	0.690	0.984	0.941
Robust p-value	0.736	0.888	0.973
VARIABLES	Dropouts grade 8th	Dropouts grade 8th	Dropouts grade 8th
RD Estimate	-0.501**	-2.527	-2.327
Robust 95% CI	[-1.846 - -.054]	[-17.147 - 8.71]	[-9.94 - 4.819]
Observations	6638	6638	6638
Conventional Std. Error	0.352	6.025	3.537
Conventional p-value	0.154	0.675	0.511
Robust p-value	0.038	0.523	0.496
Kernel Type	Triangular	Triangular	Triangular
Order Loc. Poly. (p)	0.000	1.000	2.000
Order Bias (q)	1.000	2.000	3.000

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors clustered at the province level.

Table 3.7: Student dropouts by cohort: USS

	(1)	(2)	(3)
VARIABLES	Dropouts grade 9th	Dropouts grade 9th	Dropouts grade 9th
RD Estimate	1.776	9.453	32.13
Robust 95% CI	[-4.591 - 9.532]	[-19.058 - 49.953]	[-212.83 - 300.722]
Observations	4605	4605	4605
Conventional Std. Error	2.824	15.657	117.212
Conventional p-value	0.529	0.546	0.784
Robust p-value	0.493	0.380	0.737
VARIABLES	Dropouts grade 10th	Dropouts grade 10th	Dropouts grade 10th
RD Estimate	0.843	0.643	14.70
Robust 95% CI	[-3.219 - 4.61]	[-13.575 - 15.95]	[-216.493 - 287.126]
Observations	4589	4589	4589
Conventional Std. Error	1.523	6.337	114.043
Conventional p-value	0.580	0.919	0.897
Robust p-value	0.728	0.875	0.783
VARIABLES	Dropouts grade 11th	Dropouts grade 11th	Dropouts grade 11th
RD Estimate	1.420	11.95	43.51
Robust 95% CI	[-1.982 - 6.242]	[-21.952 - 59.04]	[-327.06 - 483.219]
Observations	4583	4583	4583
Conventional Std. Error	1.665	18.196	184.557
Conventional p-value	0.394	0.511	0.814
Robust p-value	0.310	0.369	0.706
VARIABLES	Dropouts grade 12th	Dropouts grade 12th	Dropouts grade 12th
RD_Estimate	1.192	-0.113	1.109
Robust 95% CI	[-3.957 - 6.074]	[-23.68 - 19.393]	[-64.4 - 65.42]
Observations	4535	4535	4535
Conventional Std. Error	1.958	9.148	28.783
Conventional p-value	0.543	0.990	0.969
Robust p-value	0.679	0.845	0.988
VARIABLES	Dropouts grade 13th	Dropouts grade 13th	Dropouts grade 13th
RD Estimate	0.0821	-0.823	-2.498
Robust 95% CI	[-2.07 - 1.701]	[-7.997 - 4.778]	[-37.862 - 28.769]
Observations	4509	4509	4509
Conventional Std. Error	0.720	2.645	14.745
Conventional p-value	0.909	0.756	0.865
Robust p-value	0.848	0.621	0.789
Kernel Type	Triangular	Triangular	Triangular
Order Loc. Poly. (p)	0.000	1.000	2.000
Order Bias (q)	1.000	2.000	3.000

Note: *** p¡0.01, ** p¡0.05, * p¡0.10. Standard errors clustered at the province level.

the treatment status is set at 300 students. Table 3.8 shows the results for the two dependent variables of interest and by school level, which are suggestive of a statistically insignificant effect of the program on student outcomes.

Furthermore, we restrict the sample of treated school to those far away from the directional school at least 10 kilometers. Results are presented in table 3.9. Even though findings are not robust to the chosen local polynomial order, the local estimates show a negative impact of the program exposure on both dependent variables: student dropouts and students that fail to pass to the next class seem higher in treated and far away schools.[13]

Finally, we focus on the different categories of USS. In fact in the Italian system, after the LSS path, a student faces different school choices, which vary from Lyceums to Technical and Professional Institutes, Music Conservatory, Academy of Arts and National Dance Academy. Given the available data, a robust sample size is possible only for Lyceums, Technical and Professional Institutes. Results reported in table 3.10 are not suggestive of any significant heterogeneous effect with respect to the USS school typology.

[13]We conducted the estimation disentangling by school level and the USS seem the ones that tow the overall effect.

Table 3.8: Treatment effect in mountain municipalities and small islands

Lower Secondary Schools						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Dropouts	Dropouts	Dropouts	Not admitted	Not admitted	Not admitted
RD Estimate	-0.645	-0.266	-0.262	-5.815	-6.331	-2.298
Robust 95% CI	[-2.179 - .526]	[-.654 - .18]	[-.661 - .166]	[-34.033 - 10.188]	[-20.788 - 5.436]	[-8.948 - 5.027]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Observations	1697	1697	1697	1744	1744	1744
Conventional Std. Error	0.548	0.194	0.195	9.013	5.740	3.321
Conventional p-value	0.239	0.170	0.178	0.519	0.270	0.489
Robust p-value	0.231	0.265	0.240	0.291	0.251	0.582
Order Loc. Poly. (p)	0.000	1.000	2.000	0.000	1.000	2.000
Order Bias (q)	1.000	2.000	3.000	1.000	2.000	3.000
Upper Secondary Schools						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Dropouts	Dropouts	Dropouts	Not admitted	Not admitted	Not admitted
RD Estimate	5.169	-1.323	-1.568	-110.9	-15.99	-10.20
Robust 95% CI	[-7.313 - 39.221]	[-12.357 - 11.221]	[-12.204 - 9.831]	[-1477.815 - 2583.697]	[-70.381 - 51.669]	[-46.377 - 27.578]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Observations	707	707	707	810	810	810
Conventional Std. Error	9.526	5.115	5.015	867.551	26.681	16.278
Conventional p-value	0.587	0.796	0.754	0.898	0.549	0.531
Robust p-value	0.179	0.925	0.833	0.594	0.764	0.618
Order Loc. Poly. (p)	0.000	1.000	2.000	0.000	1.000	2.000
Order Bias (q)	1.000	2.000	3.000	1.000	2.000	3.000

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors clustered at the province level.

Table 3.9: Treated school far away from the directional school at least 10 Km

VARIABLES	(1) Dropouts	(2) Dropouts	(3) Dropouts
RD Estimate	2.305*	7.921	28.92
Robust 95% CI	[-.416 - 8.628]	[-16.61 - 46.485]	[-251.717 - 316.137]
Observations	7412	7412	7412
Conventional Std. Error	1.952	14.203	129.387
Conventional p-value	0.238	0.577	0.823
Robust p-value	0.075	0.353	0.824
VARIABLES	Not admitted	Not admitted	Not admitted
RD Estimate	27.36**	212.0	255.4
Robust 95% CI	[3.647 - 97.288]	[-2342.111 - 3216.67]	[-3753.949 - 4080.582]
Observations	7722	7722	7722
Conventional Std. Error	21.343	1237.5	1776.4
Conventional p-value	0.200	0.864	0.886
Robust p-value	0.035	0.758	0.935
Kernel Type	Triangular	Triangular	Triangular
Order Loc. Poly. (p)	0	1	2
Order Bias (q)	1	2	3

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors clustered at the province level.

Table 3.10: Estimation results by USS typology

Lyceums						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Dropouts	Dropouts	Dropouts	Not admitted	Not admitted	Not admitted
RD Estimate	0.163	5.026	3.090	5.376	15.95	24.5
Robust 95% CI	[-2.015 - 2.022]	[-60.614 - 80.449]	[-64.082 - 87.624]	[-2.044 - 22.458]	[-26.849 - 80.272]	[-110.395 - 163.75]
Observations	1539	1539	1539	1635	1635	1635
Conventional Std. Error	0.744	31.062	33.759	5.46	24.05	62.098
Conventional p-value	0.827	0.871	0.927	0.325	0.507	0.693
Robust p-value	0.997	0.783	0.761	0.102	0.328	0.703
Professional Institutes						
RD Estimate	0.407	15.43	5.997	16.3	-22.22	-8.891
Robust 95% CI	[-21.314 - 20.425]	[-325.557 - 315.894]	[-69.996 - 88.214]	[-57.932 - 159.23]	[-297.223 - 268.948]	[-99.509 - 90.06]
Observations	1079	1079	1079	1227	1227	1227
Conventional Std. Error	8.224	140.143	36.166	47.08	124.721	42.584
Conventional p-value	0.96	0.912	0.868	0.729	0.859	0.835
Robust p-value	0.967	0.976	0.821	0.361	0.922	0.922
Technical Institutes						
RD_Estimate	-0.157	0.453	2.481	6.882	32.12	10.05
Robust 95% CI	[-4.256 - 2.948]	[-9.95 - 14.033]	[-15.224 - 22.854]	[-6.599 - 34.134]	[-122.505 - 219.794]	[-62.35 - 75.618]
Observation	1516	1516	1516	1635	1635	1635
Conventional Std. Error	1.441	5.247	8.639	9.035	72.763	31.203
Conventional p-value	0.913	0.931	0.774	0.446	0.659	0.747
Robust p-value	0.722	0.739	0.695	0.185	0.577	0.851
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Order Loc. Poly. (p)	0	1	2	0	1	2
Order Bias (q)	1	2	3	1	2	3

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors clustered at the province level.

7 Final Remarks and Discussion

Starting from the crucial role of the human capital in the process of economic growth, we aimed to assess the impact of the consolidation process that involved the Italian Secondary Schools - the main "producers" of human capital. We exploited the fuzzy discontinuity in treatment regime induced by the number of students enrolled in a given school to evaluate the impact of mergers on two categories of student dropout behaviour: voluntary dropouts and grade retention.

Estimation results of a continuity-based local polynomial approximation Fuzzy RDD show that the program exposure is scarcely associated to statistically significant variation in students dropouts and failure to progress. In fact, only the percentage of students not admitted to the next class looks higher by 24.94 percentage points in the treated USS schools. We also find evidence of heterogeneous effects with respect to the driving distance from directional schools to merged complexes. Indeed, the student dropout behaviour seems higher in treated schools far away at least 10 Km from the directional schools. This provides an insight on the potential causality channels that may drive the treatment effect. Indeed, the SP responsibilities appear substantially aggravated with the program exposure and the geographical distance from the merged schools he or she is in charge to run appears as an additional cost for the SP that contribute to isolate the merged school complexes, whose school-specific needs are not satisfied. Thus, the school consolidation process seems as a cost-saving policies from a public finance point of view that may fail in considering the indirect costs arising in term of poorer school performance, which may later produce labor market inequalities due to losses in human capital accumulation.

School merger programs are not limited to the Italian case but have been embraced by many other countries, such as China (Mo *et al.*, 2012), United States (Logan, 2018), Denmark (Beuchert *et al.*, 2018), and Netherlands (De Haan *et al.*, 2016). Thus, we added to the relevant literature by providing some evidence on the impact of such consolidation policies on student achievement in an additional educational framework.

The analysis presents some limitations. First, the effect of these policy changes on student performance may depend on the years of exposure to the consolidation program. Since we do not observe multiple periods, we are not able to explore this possibility. In addition, we do not have any information on SP characteristics, such as tenure or management skills, and student cognitive and non-cognitive skills, that the existing literature claims as a determinant of school performance.

Finally, the applied methodology produces estimations of treatment effect that are very local in nature and tend to have a little external validity. However, compared to more conventional quasi-experimental methods, such as the Differences in Differences approach, the Local Average Treatment Effect, also known as LATE, resulting from the RDD estimation, has substantially a higher internal validity and is less sensitive to boundary-related problems and outliers.

Bibliography

[1] Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. Journal of labor Economics, 25(1), 95-135.

[2] Amarante, V., Manacorda, M., Miguel, E., & Vigorito, A. (2016). Do cash transfers improve birth outcomes? Evidence from matched vital statistics, program, and social security data. American Economic Journal: Economic Policy, 8(2), 1-43.

[3] Angrist, J. D., & Pischke, J.-S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton: Princeton University Press.

[4] Barro, R. J. (2001). Human capital and growth. American economic review, 91(2), 12-17.

[5] Beuchert, L., Humlum, M. K., Nielsen, H. S., & Smith, N. (2018). The short-term effects of school consolidation on student achievement: Evidence of disruption?. Economics of Education Review.

[6] Bloom, N., Lemos, R., Sadun, R., & Van Reenen, J. (2015). Does management matter in schools?. The Economic Journal, 125(584), 647-674.

[7] Besley, T., & Coate, S. (2003). Centralized versus decentralized provision of local public goods: a political economy approach. Journal of public economics, 87(12), 2611-2637.

[8] Branch, G. F., Hanushek, E. A., & Rivkin, S. G. (2012). Estimating the effect of leaders on public sector productivity: The case of school principals (No. w17803). National Bureau of Economic Research.

[9] Boffa, F., Piolatto, A., & Ponzetto, G. A. (2015). Political centralization and government accountability. The Quarterly Journal of Economics, 131(1), 381-422.

[10] Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. Econometrica, 82(6), 2295–2326.

[11] Calonico, Cattaneo, Farrell & Titiunik (2018): Regression Discontinuity Designs Using Covariates, Review of Economics and Statistics, forthcoming.

[12] Calonico, S., M.D. Cattaneo, Farrell, M. H., & Titiunik, R. (2017). rdrobust: software for regression-discontinuity designs. Stata Journal, 17, 372-404.

[13] Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2017). A Practical Introduction to Regression Discontinuity Designs: Part I.

[14] Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2018). A Practical Introduction to Regression Discontinuity Designs: Volume II.

[15] Cattaneo, M. D., Titiunik, R., & Vazquez-Bare, G. (2017). Comparing inference approaches for RD designs: A reexamination of the effect of head start on child mortality. Journal of Policy Analysis and Management, 36(3), 643-681.

[16] Clark, D., Martorell, P., & Rockoff, J. (2009). School Principals and School Performance. Working Paper 38. National Center for Analysis of longitudinal data in Education research.

[17] Coelli, M., & Green, D. A. (2012). Leadership effects: School principals and student outcomes. Economics of Education Review, 31(1), 92-109.

[18] Cullen, J. B., & Mazzeo, M. J. (2008). Implicit performance awards: An empirical analysis of the labor market for public school administrators. University of California, San Diego (December).

[19] Dai, F., Cai, F., & Zhu, Y. (2018). Returns to Higher Education in China: Evidence from the 1999 Higher Education Expansion Using Fuzzy Regression Discontinuity (No. 11735). Institute for the Study of Labor (IZA).

[20] De Haan, M., Leuven, E., & Oosterbeek, H. (2016). School consolidation and student achievement. The Journal of Law, Economics, and Organization, 32(4), 816-839.

[21] EUROSTAT - Statistic Explained https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Early_leaver_from_education_and_training. Accessed October 2018.

[22] Fontana, A. (2008). La rete scolastica: un modello per il numero di studenti per classe nella scuola primaria e secondaria di I grado. In: Paper presentato alla XX Conferenza della Società italiana di economia pubblica. Pavia, settembre.

[23] Galiani, S., Gertler, P., & Schargrodsky, E. (2008). School decentralization: Helping the good get better, but leaving the poor behind. Journal of public economics, 92(10-11), 2106-2120.

[24] Gelman, A., & Imbens, G. (2018). Why high-order polynomials should not be used in regression discontinuity designs. Journal of Business & Economic Statistics, 1-10.

[25] Google Maps Platform - Directions API https://developers.google.com/maps/documentation/javascript/directions.

[26] Grissom, J. A. (2011). Can good principals keep teachers in disadvantaged schools? Linking principal effectiveness to teacher satisfaction and turnover in hard-to-staff environments. Teachers College Record, 113(11), 2552-2585.

[27] Grissom, J. A., Kalogrides, D., & Loeb, S. (2015). Using student test scores to measure principal performance. Educational Evaluation and Policy Analysis, 37(1), 3-28.

[28] Grissom, J. A., & Loeb, S. (2011). Triangulating principal effectiveness: How perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills. American Educational Research Journal, 48(5), 1091-1123.

[29] Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica, 69(1), 201-209.

[30] Hallinger, P., & Heck, R. H. (1998). Exploring the principal's contribution to school effectiveness: 1980-1995. School effectiveness and school improvement, 9(2), 157-191.

[31] Hanushek, Eric A., and Dennis D. Kimko (2000). Schooling, labor force quality, and the growth of nations. American Economic Review 90,no.5 (December):1184-1208.

[32] Hanushek, E. A., Lavy, V., & Hitomi, K. (2008). Do students care about school quality? Determinants of dropout behavior in developing countries. Journal of Human Capital, 2(1), 69-105.

[33] Hanushek, E. A., & Woessmann, L. (2010). The high cost of low educational performance: The long-run economic impact of improving PISA outcomes. OECD Publishing. 2, rue Andre Pascal, F-75775 Paris Cedex 16, France.

[34] Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. Economics of Education review, 27(6), 615-631.

[35] Koedel, Cory (2008). Teacher quality and dropout outcomes in a large, urban school district. Journal of Urban Economics, 64 (3), 560-572.

[36] Konstantopoulos, Spyros (2007) : How long do teacher effects persist?, IZA Discussion Papers, No. 2893, Institute for the Study of Labor (IZA), Bonn.

[37] Lamb, S., & Markussen, E. (2011). School dropout and completion: An international perspective. In School dropout and completion (pp. 1-18). Springer, Dordrecht.

[38] Leigh, A. (2010). Estimating teacher effectiveness from two-year changes in students' test scores. Economics of Education Review, 29(3), 480-488.

[39] Logan, A. J. (2018). How Teachers Experience Change: A Case Study of the Merger Between Two Catholic Schools.

[40] Lucas Jr, R. E. (1988). On the mechanics of economic development. Journal of monetary economics, 22(1), 3-42.

[41] McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. Journal of econometrics, 142(2), 698-714.

[42] Miller, A. (2009). Principal turnover, student achievement and teacher retention. Princeton University: NJ.

[43] MIUR - Elenco Comuni di Montagna. https://archivio.pubblica.istruzione.it/reclutamento/allegati/elenco_comuni_montani_05.pdf.

[44] MIUR - Scuola in Chiaro. http://cercalatuascuola.istruzione.it/cercalatuascuola/

[45] MIUR - Elenco Scuole Piccole Isole. http://archivio.pubblica.istruzione.it:80/mobilita/bollettino/2010/scuole_isolane10.pdf.

[46] Mo, D., Yi, H., Zhang, L., Shi, Y., Rozelle, S., & Medina, A. (2012). Transfer paths and academic performance: The primary school merger program in China. International Journal of Educational Development, 32(3), 423-431.

[47] Oates, W. E. (1972). Fiscal federalism. Books.

[48] Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. Econometrica, 73(2), 417-458.

[49] Romer, Paul (1990). Endogenous technological change. Journal of Political Economy 99, n.5, pt.II: S71-S102.

Appendix

Table 3.11: Density tests of the running variable

Sample	T-statistic	p-value	Obs
All schools	0.4913	0.6232	11400
LSS	1.7581	0.0787	6665
USS	1.3421	0.1796	4735

Source: Author elaboration

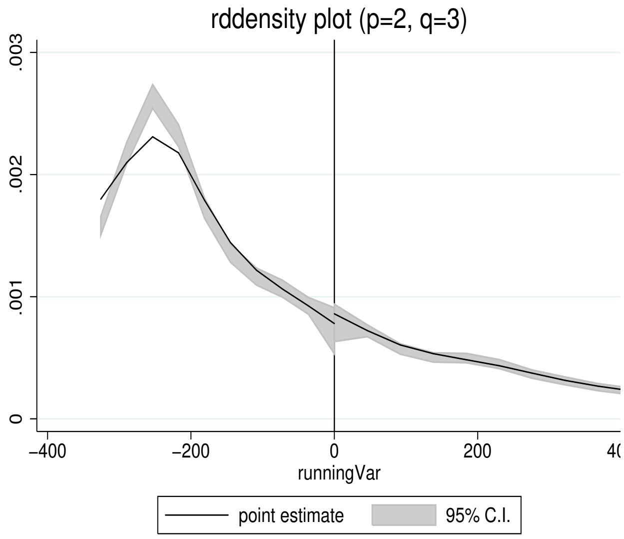

Figure 3.11: Estimated Density - All schools. Source: The author

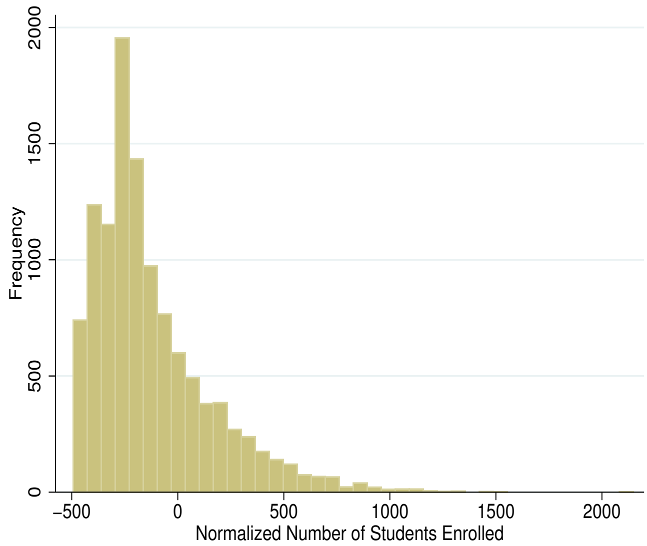

Figure 3.12: Histogram - All schools. Source: The author

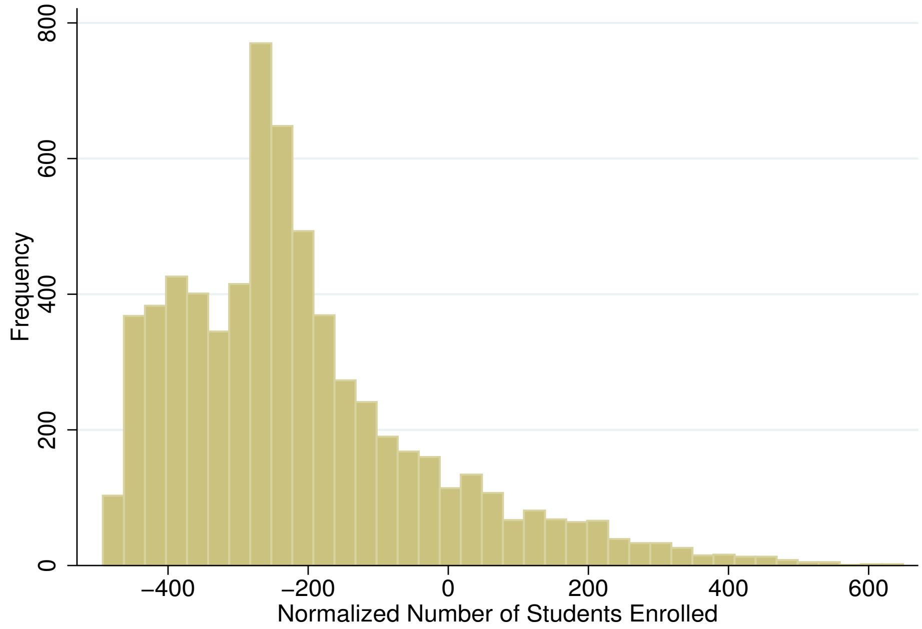

Figure 3.13: Histogram - LSS. Source: The author

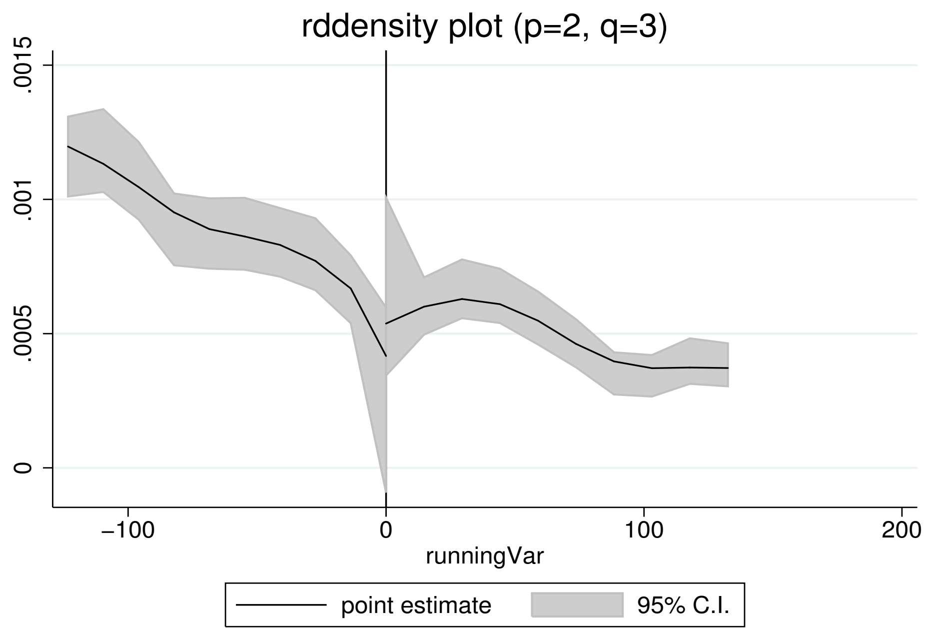

Figure 3.14: Estimated Density - LSS. Source: The author

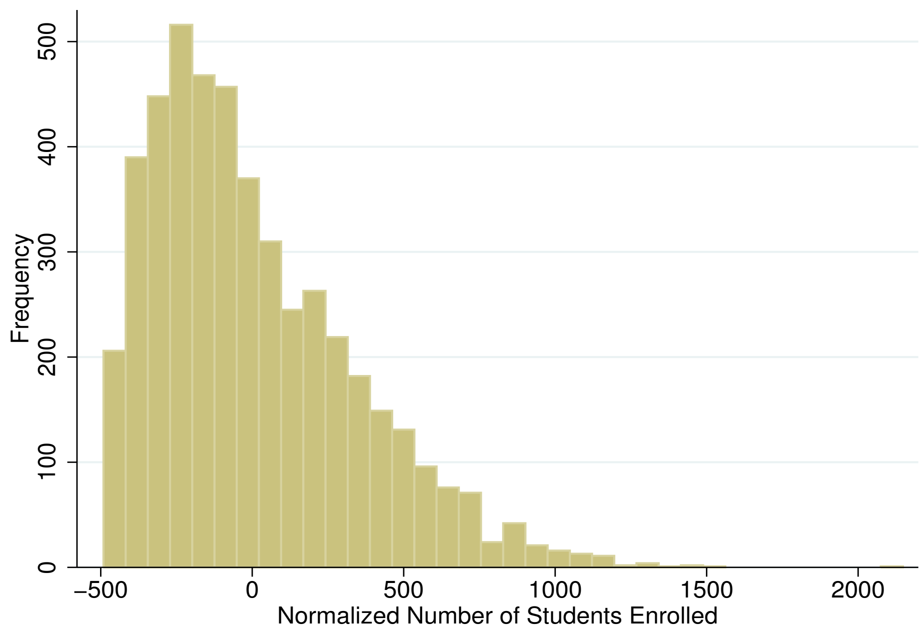

Figure 3.15: Histogram - USS. Source: The author

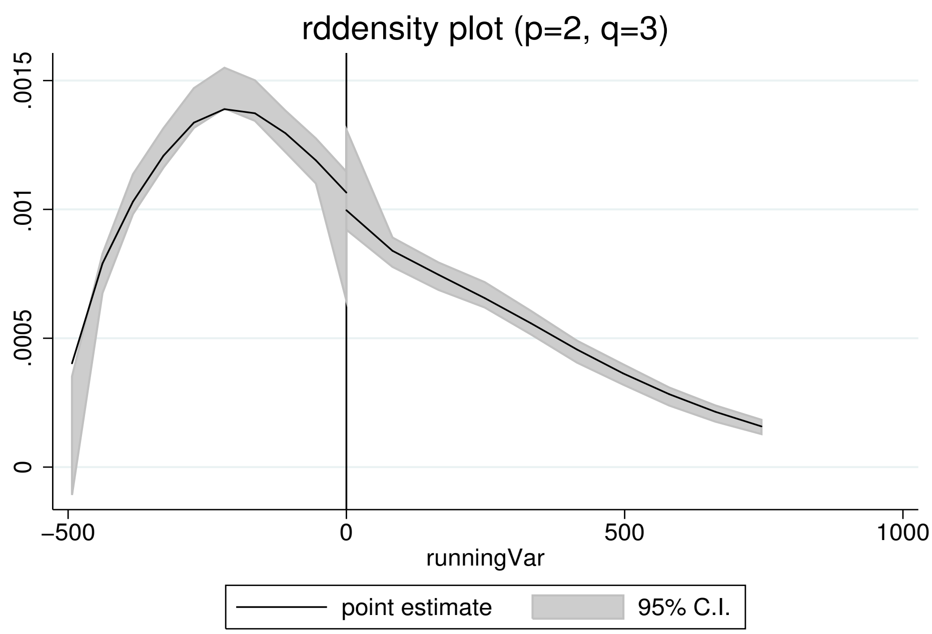

Figure 3.16: Estimated Density - USS. Source: The author