SAPIENZA
UNIVERSITÀ DI ROMA

PhD programme in Economics and Finance
XXXII ciclo

Three Essays in Health Economics

Phd Candidate: Vanessa Cirulli
Matricola: 1680767

Coordinator: Rita Laura D'Ecclesia
Tutor: Marco Marini

Abstract

The present PhD Thesis titled "Three Assays on Health Economics" consists of three separates pieces of research on three different topics.

The first paper is entitled "*Do mergers affect hospital outputs and outcomes? Evidence from the English secondary care sector*". We analyze the impact of hospital mergers on several measures of hospital outputs and outcomes over the period 2000-2008 in order to understand whether English hospital Trusts that have merged are able to reconfigure their service offer more significantly than non-merged Trusts and consequently to achieve an advantage relative to non-merged Trusts. In order to answer this question, for the analysis on hospital outputs, we adopt an innovative flexible conditional difference-in-difference approach, developed by Dettman et al. (2020) able to capture mergers with varying start dates and varying treatment durations. Regarding the analysis on hospital outcomes, we adopt a fixed effect ordered logit model as developed by Dickerson et al. (2014). Our empirical analysis shows a negative impact of hospital mergers on both hospital outputs and outcomes.

The second paper is entitled "*Financial crisis, fiscal austerity, and health in Italy*". This paper aims to assess whether the economic crisis (2007-2008) and the Italian sovereign debt crisis (2010-2011) have both had any impact on the health of the Italian population, proxied by a wide set of indicators. Following previous papers by Kentikelenis (2015), Kentikelenis et al. (2014) and Karanikolos et al. (2013) we analyze the effect of regional bail-out plans adoption on a broad set of health status measures during the period 1999-2015 considering the bail-out effect on physical and psychological measures of health status, focus on social distress. We adopt a Variable Instrumental approach to address potential endogeneity issues associated with the choice of adopting bail-out plans. Our empirical results show a general increase in mortality rate, and also of the incidence of some infectious diseases. The adoption of bail-out plans affects mainly vulnerable people with psychological diseases.

The third paper is entitled "*Does co-payment exemption increase diagnostic care utilization?*

A causal approach for the Italian care system". The purpose of this paper is to analyze the effect of co-payment exemption on the diagnostic care utilization. Increased utilization of healthcare can be driven either by health needs or by opportunistic behavior. In this preliminary analysis we overcome the potential endogeneity associated to co-payment exemption by adopting an Instrumental Variable approach. We consider the Global Competitiveness Index at regional level as the proxy of bureaucracy and administrative slowness. Our findings reveal a weakness of instrument due to weak joint statistical independence. In order to estimate the possible effect of co-payment on diagnostic cares utilization, we adopt an alternative empirical method based on the estimation of intersection bound. Preliminary findings confirm that co-payment exemption increase the average number of diagnostic care and also, reveal potential opportunist behavior. Even if our preliminary results cannot allow to estimate the exact effect of co-payment exemption, the inference on intersection bounds permits to identify the possible dimension of the issue.

Acknowledgement

I would like to express my sincere gratitude to my supervisor Prof. Giorgia Marini for continuous support on my PhD studies, for her patience, motivation, and immense knowledge. Her supervision helped me in each moment of PhD studies and research, encouraging me professionally and personally.
She has never abandoned me. She has made my PhD experience as one the best period of my life.

I would like to thank my tutor Prof Marco Marini for believing in me and for his suggestions and encouragement. He always supported me with sincere advice, huge smiles and a lot of peace.

The last but not least thanks are for Giovanni Cerulli. He introduced me into the beauty of econometric studies and in the research thoughts sharing.

Do mergers affect hospital outputs and outcomes? Evidence from the English secondary care sector*

Vanessa Cirulli
Sapienza University of Rome
vanessa.cirulli@gmail.com

October 30, 2020

Abstract

We analyse the impact of merger on several measures of hospital outputs and outcomes among English Trusts during the period 2000-2008. We adopt an alternative difference-in-differences approach, called "flexible conditional DID" as developed by Dettman et al., 2020, to capture time-varying effect related to its starting time and duration. We find that merger policy has a negative effect on hospital outputs and outcomes mainly in the three years after merger. Given that merger achievement is enhancing efficiency and improve quality between hospitals, our findings suggest that merger policy in the short-run may be an inappropriate strategy of dealing with poorly performing hospitals.

JEL Codes: I11, I13, I18, L32

Keywords: Hospital mergers, Performance, Flexible conditional DID

1 Introduction

Over the past fifty years, there have been marked changes in organisational structures and budgetary arrangements in the English National Health Service (NHS) causing, among other things, a widespread merger activity in the hospital sector that has significantly reshaped the health system: several waves of hospital consolidation have dramatically reduced the number of providers operating in England from about 400 in 1960, serving an average population of 100,000 people, to about 150 in 2018, serving an average population of 450,000 people. Such reshaping of course has posed questions of quality of services provided, performance of hospital

*The views expressed are those of the authors and do not necessarily reflect those of the institutions to which they belong.

providers, efficiency in terms of economies of scale and scope.

In this context of reorganizational changes, we want to investigate whether merging activity has had any significant effect on hospital outputs and outcomes. In particular, we refer to horizontal mergers between neighbouring providers carrying out similar services to overlapping or contiguous population in order to rationalize the offered services (Collins, 2015). "Horizontal merger is one route through which a firm can acquire dominance over the supply of goods or services in a market" (Goddard and Ferguson, 1997, p. 15). As efficiency of a firm depends not only on the degree of competition of the market, but also on the degree of monopoly power within the firm, it is fair to ask to what extent mergers, as one way to acquire monopoly power, are a potential problem in a health care system. Based on the existing literature, the existence of monopoly power can give poor incentives for management to take action in order to operate efficiently as implementing cost-saving (i.e., efficient) mechanisms is likely to cause the hospital management considerable time and effort. Moreover, the pricing regime in the NHS, which allows Trusts to cover costs plus an allowance for rate of return on capital, provides further incentives for those in monopoly positions to put less effort into restricting costs. Finally, mergers involving reconfiguration of services onto centralised sites and the closure of others will have implications for patients in terms of ease and cost of access. Therefore, whether mergers can be expected to deliver benefits overall to patients depends largely on the incentives generated for improving efficiency.

In order to capture some evidence of merger hospitals in terms of efficiency and performance, we address a policy evaluation of merger effects on several measures of hospital activities, such as inpatient admissions, elective admissions, emergency admissions, outpatients, day cases and various combinations of them, and additionally on two measures of performance based on quality of services and use of resources. We analyse hospital mergers occuring at different years over the period 2000-2008.

We contribute to the existing literature on hospitals merger showing an alternative approach to dealing with Trust merging in different years and how the time-varying treatment and its different duration can impact on hospital outputs and outcomes. The lack of data on the merger date of decision to merge or the merger date of announcement do not allow to implement any event study, as in Gaynor et al. (2012). However, following the recent methodology strategy "flexible conditional DID", as proposed Dettmann et al. ((2020), we are able to estimate the average effect of merged Trusts (ATT) with time-varying treatment, by selecting appropriate controls in each year of merger occurred and having the same features of merged Trusts. The definition of pre-merger features and time are crucial into analysis as well as the definition of pre-merger time output development and its post-merger development. Moreover, this approach assumes conditional parallel trends, replacing the common parallel trends assumption as required by the canonical DID and event studies. Thus, it allows to capture not only the merger effect on treated Trusts related to the different starting of treatment, but also considering different duration between treated and controls. We test conditional parallel trend assumption by using Cerulli and Ventura (2019) methodology considering both time leads and time trends.

The remainder of this article is organized as follows. We introduce a brief institutional context in order to explore how NHS reforms influenced merger hospitals in Section 2. In section 3. we illustrate the Literature review representing our background analysis. We explain our Methodology strategy, presenting the flexible conditional DID in details, how we estimate the DID with fixed effects and also, why and how we demonstrate the conditional parallel trend assumption in section 4. In section 5 we present data used in our analysis, their descriptive statistics, and also, how we reorganized our data in order to taking into account time-varying treatment effects. In section 6 we present separate output and outcome results, showing in detail flexible conditional DID and fixed effect findings. Moreover, we show graphically and by testing both time leads and time trends variables the proof of conditional common parallel trends. Section 7 concludes. Additional graphical results are reported in the Appendix.

2 Institutional framework

Since its establishment (1948), the English National Health Services was been provided free of charge. During the following years, several reforms have been set out. The attention was put on definition of administrative responsibility in order to satisfy healthcare demand but also concerning on costs containment of public healthcare sector.

The problem of cost savings and better resources' allocation in the health public sector worsened during the 1970s also, as result of the international oil crisis (1973-1974). The Conservative Party, that won the national election in 1979, during the '80s decade started a process of transformation in each public sector introducing private sector economic mechanisms. The health sector goals of the Thatcher Governments in 1979-90, sought to quality improvement of health services by reducing the monopoly's power of public health sector and introducing competition. They set out several market-oriented changes, an "imposition of a 'managerialist' regime in Britain's public services" (Dorey, 2015) and also, the introduction of audits, analysis and monitoring to measure the performance of public health sector and its staff. Thus, the reorganizational process began in 1982 with the abolition of area health authorities (AHAs). The district health authorities (DHAs) were put under the control of the regional health authorities. Subsequently, in order to promote health-enhancing, the Governments announced in the "White Paper, Promoting better health" (1987) and in "White paper Working for Patients" (1989), improvements in patient choice, a broadening of services provided by pharmacists and nurses, quality and financial incentives to raise the processes of delivering care, with extra pay for undertaking health promotion, screening and other preventative actions. They introduced Trusts for hospitals with proven managerial ability, giving them more freedom (self-governing NHS trusts), and also, suggested the figure of the medical audit into hospitals and primary care.

Then, with their most important reform the "National Health Service and Community Care Act" (1990), they sought increasing efficiency gains by introducing competition in the public healthcare sectors. They introduced a state-financed internal market within the NHS and a split between purchasers and providers of care. This separation should have permitted to purchasers, mainly health authorities, to invest their budgets in obtaining services from different providers (such as acute, mental health, elderly or disabled people hospitals). Moreover, the reform established GP fund-holdings, which allowed GPs to make budgets, purchasing secondary care services from hospitals and other providers on behalf of their patients[1]. In the light of these reforms, providers began to merge their specialty or activities, both as a formal Trust merger process and as applying for NHS Trust status[2]. Thus, an increasing concentration of hospitals in some areas, lessened the competition's principle advocated by the reforms. The Governments pointed out on the importance of competition on the supply side with the purpose of achieving both as the efficiency gains derived from contracts between purchasers and as providers and the adoption of regulatory framework for public accountability ("The Operation of the NHS Internal Market: Local Freedoms, National Responsibilities",1994).

Several critiques focused on claims that the reforms led on establishing a two-tier health system. The Governments were accused of creating disparities between patients of GP fundholders and of non-fundholding GPs, in which the first ones would have had a more quickly access to health treatment than the latter. Until the Labour Party's victory (in 1997), the Thatcher Governments pursued its NHS reorganizational and rationalization goal by reducing the number of regional health authorities, lessening bed numbers, supporting changes in specialism mix and providing high technology services. In the late 1990s the Labour Party won elections. An element of its campaign was the abolition of the internal market, but the Labour Governments did not realise this point. Rather, they supported any sort of partnership between providers having a robust performance management. Specifically, the "NHS Plan (2000)" clarified the role of private in the public health sector, creating a more constructive and efficient collaboration between these two players. The competition introduction of Conservative Governments as leverage of enhancing efficiency and improving quality in the NHS sectors was replaced by the idea of "collaboration" between of private and public healthcare sectors. In particular, the effect of this policy was seen in public hospitals. Several hospitals were closed by with others geographically close. The Governments enhanced the involvement of the private and voluntary sectors to improve patients care in each stage of their clinical process. Thus, independent sector treatment centres (ISTCs) were established (managed both by the NHS and by the independent sector), in order to enhance healthcare services in terms of efficiency

[1] As not all GPs became fundholders, the local health authorities purchased services for them.
[2] A Trust could be conferred with the assets of multiple hospital sites.

and performance (i.e. fast, pre-booked surgery and diagnostic tests for NHS-funded patients by separating scheduled treatment from emergency care). Moreover, with the aim to reduce waiting lists and times, introduced NHS diagnostic and treatment centres as separate centres on NHS hospital sites specialising in routine diagnostics and operations that did not require hospital admission.

Also, Governments outlined the importance of modernized healthcare sectors through a set of investments and reforms. Their Agenda provided for investing in NHS facilities (extra beds in hospitals and intermediate care, new hospitals and GP premises, new one-stop primary care centres, scanners, modern IT systems in every hospital and GP surgery clean wards improvements and better hospital food) and staff (increasing the numbers of consultants, GPs, nurses, therapist, new medical school places). These large-scale investments should be coupled with widespread reforms "redesigned around the needs of the patient". In order to achieve these goals, they claimed the principles of subsidiarity and a new system of autonomy, where the centre would devolve the power to the local. Moreover, GPs fund-holdings were abolished and replaced with GPs primary care groups. These groups had indicative budgets and took control of the commissioning, previously managed from the health authorities. They also could become primary care trusts (PCTs) replacing health authorities as the principal commissioners of NHS services at local level.

In order to increase efficiency and improve quality of the healthcare sector it was introduced the "patient choice" in time and dates of hospital appointment (and subsequent of being treated by an alternative provider if they could not be treated within six months by the NHS), and also an activity-based payment system for hospitals known as Payment by Results (PbR). Since 2003, NHS trust with high performance were given the opportunity to apply for the status of NHS foundation trusts (FTs), gathering more financial and managerial freedoms remaining fully part of the NHS. In general, since the '70s and in each subsequent decade, the NHS organisational pattern was influenced by different aims and principles with the common intent to achieve financial savings in healthcare sector improving quality and performance. Thus, whether the definition of management functions featured the '80s in order to develop a management consensus since 1990s the main aim was the good governance based on accountability and integrity. Throughout different Governments, the phenomenon of hospital mergers activity was homogeneous and seemed to be the standard solution to achieve both Parties' intents.

3 Literature review

The industrial economics literature applied to the healthcare market offers several explanations about the driving forces behind merger activity and the impact of mergers on the performance of Trusts. In effect, Trusts may have used mergers as a strategic tool to improve their financial performance through price increases (made possible by increased market power) and/or cost reductions (made possible by either economies of scale and scope, monopsony power or favourable adjustments in the product mix), with important policy implications both in terms of services provided to patients and in terms of an efficient use of available resources. From a theoretical point of view, the aim of hospital mergers has been enhanced efficiency and quality benefits, such as production costs reduction, output increase, quality improvements and both operating and managerial efficiencies enhancements, reinforcement of financial sustainability, simplification of staff recruitment. Whilst, these benefits were not always achieved. Indeed, for instance, the management savings from NHS mergers have been highly variable and sometimes much lower than expected or they have been more likely to injury finances of trusts than improve them (Gaynor *et al.*,2012), furthermore the process of staff recruitment has not been made easier after the merging process (Fulop *et al.*, 2005). Thus, a wide empirical literature has investigated on possible consequences of hospital mergers activity in terms of economic and non-economic benefits. Some authors have attributed reconfigurations (either mergers or acquisitions) as possible devices to change the mix of services offered (Krishnan *et al.*,2004), or as tools to gain efficiency (Dranove, 1998; Preyra and Pink, 2006; Kjekshus and Hagen, 2007; Radach Spang *et al.*, 2009). Other researchers have highlighted the impact of mergers on prices (Radach Spang *et al.*, 2009)) or how the type of merging hospital have affected costs (Schmitt, 2017a[3]).

[3]According to Schmitt (2017), in fact, mergers between independent hospitals have a small and insignificant effects on costs, "while acquisitions by multihospital system can have larger and significant cost reductions".

However, ambiguous results have been gathered about merger process' effects on social welfare (Town *et al.*, Town, Feldman, and L. 2006) or quality of the services provided (Ho and Hamilton, 2000; *et al.*, 2004), although Bloom *et al.* (2005) suggest that higher management quality and improvement of performance have derived from higher competition results. An analysis of the impact of mergers between NHS hospitals on financial performance, productivity, waiting times and measures of clinical quality found little evidence of improvement in any of these areas and, on some measures, performance actually declined, producing little benefits in terms of patient welfare (Gaynor *et al.*,2012). As described by Collins 2015), the merge process is complex and characterised by a long and costly proposal process, that influence not only the realization of merger but also, the future life of new merged Trust.

Additionally, several evidences have highlighted how the impact of competition in markets with fixed prices has led to improvements in hospital performance (Gaynor, 2004 or on hospital quality and efficiency (Propper *et al.*, (2008 and 2012; Bloom *et al.*, 2005).

Previous studies report mixed results. In the UK, Propper *et al.* (2004) show that competition reduced quality (i.e. increased death rates), although waiting time has been reduced too (Propper *et al.*, (2008). The Choose and Book reform (improved user choice coupled with DRGs implying that money follows the patient) is shown to improve health care quality (Gaynor *et al.*,2012 and 2013). On the contrary, Cooper *et al.* (2011) show that increased competition between private and public hospitals decreased productivity among the public sector (as more complicated cases were selected there) while it increased in the private sector.

Despite the popularity of mergers, the evidence on the success of this policy in terms of delivering benefits it promised, or indeed, any benefits at all, is so far negative and, in general, evidence is indeed mixed. It may depend on the methodological strategy applied and on the appropriate selection of control group. Some studies adopted a difference-in differences approach to compare average changes on characteristics and performance variables for the pre-merger hospitals with the new post-merger entities as a group. For example, Schmitt (2017b) highlighted costs-savings findings in the acquired hospitals in the years following the acquisition, by using difference-in-differences models and considering a variety of different control strategies.

Cooper *et al.* (2011) applied DID estimator to test whether hospital quality improved after NHS introduction after the 2006.

However, some doubts in the appropriateness of canonical DID assumption of homogeneous effects are raised in the literature, especially treatment effect is dynamic. In this context the event study, thought as an extended DID, can be applied. For instance, Gaynor *et al.*,2012 adopted an event study strategy to measure merger effects on financial performance, productivity, waiting times and clinical quality. By matching procedure, they select an appropriate control group and compare "hospital performance before and after the merger and between hospitals that ever merged and those that never merged hospitals". Thus, they overcome possible anticipation effects. Moreover, they assume that the outcomes in the treated and control groups follow parallel trends in the absence of the treatment. However, the adoption of event studies required generalized form of parallel trends assumption, no anticipation effect of treatment and also, impose any variation on treatment effects across groups. However, the interpretation of heterogenous treatment effect can be difficult by using event studies. Recent literature (Athey and Imbens, 2018; Imai and Kim, 2019; Callaway and Sant'Anna, 2019; Abraham and Sun, 2018; Dettman et al, 2020) propose alternative DID specification to capture treatment effect heterogeneity. This is the reason why we also adopt this alternative DID approach.

4 Methodology

4.1 "Flexible conditional difference in difference" approach

We investigate possible differences in the level and in the composition of hospital activity as well as in the performance of hospital Trusts between merged and non-merged Trusts. We examine whether the merger activity, affecting different hospitals in different years, has produced any difference at all or whether indeed there are long-standing differences between these different types of organisations (in terms of hospitals' output and outcome), which have led some of them more likely to merge than others.

We consider nine different measures of hospital activity and two different measures of performance, and we compare these measures over time for merged Trusts and non-merged Trusts.

One of the main challenges in evaluating whether Trust mergers have any effect on hospital outputs and outcomes is the ability to draw firm conclusions based on the comparison between merged and non-merged Trusts, when the decision to merge takes place in different years among different hospital Trusts, and it is voluntary (likely due to poor performance and efficiency of one of the Trusts involved). Allowing for this potential selection bias is, therefore, a key component of our research, and we describe below our approach to this.

We use the "flexible conditional difference in difference" approach, following Dettmann et al., (2020), to consider the potential selection bias associated with heterogeneous treatment effects in a panel data structure.

The 'flexible conditional DID approach, as developed by Dettman et al. (2020), is a non-parametric conditional difference-in-differences approach (as introduced by Heckman et al., 1998) within the framework of the staggered adoption design, where units (they are treated once in the observation time) are considered as treated units from that date onwards [4].
The main idea is based on the combined propensity-score matching with DID to find adequate controls for the treated units. It is based on some assumptions of Callaway and Sant'Anna (2019) and Imai et al. (2019) approaches.

Like Callaway and Sant'Anna (2019) approach, they apply the staggered adoption framework and define time in relation to the treatment start. Moreover, the flexible conditional DID estimator can be considered as a special case of the group-time average treatment effects approach proposed by Callaway and Sant'Anna (2019[5]), but differently from them, each group is based on number of treated units weighted by specific group sizes and not based only on the first year of treatment. First, they select a set of groups which is composed of one single

[4] Dettmann et al.'s statement derived from Callaway and Sant'Anna (2019) who assume the *Irreversibility of Treatment*. They consider T periods and a particular time period t (with t=*1, ..., T*). In a canonical DID setup, $T = 2$, no one is treated in period $t = 1$, D_{it} is a binary variable equal to one if unit i is treated in period t; zero otherwise. Thus, the treatment process will assume $D_1 = 0$ almost surely. For $t = 2, \ldots, T$, $D_{t-1} = 1$ implies that $D_t = 1$ almost surely. Namely, this assumption states that no one is treated at time $t = 1$; and that once a unit becomes treated, that unit will remain treated in the next period.

[5] Callaway and Sant'Anna define group-time average treatment effect as a disaggregated causal parameter, namely the average treatment effect for group g at time t, where a "group" is defined by the time period when units are first treated.

treated unit having specific pre-treatment features in a specific pre-treatment time. We define them as “benchmark groups”. According to the benchmark groups features, they select treated units and a set of controls for each individual treated unit. The set of control observations are selected from other units in the same time period with identical pre-treatment features. The set of treated and controls groups is the sample matched. Also, they apply a post-treatment time for both treated and controls' output development. Each output development is related to different treatment starting time and treatment duration of each group. Due to heterogeneous treatment durations, the observed periods may be heterogeneous among the treated individuals. Finally, they compare individual outcome development (as proposed by Imai et al., 2019) propose a matching-DID estimator for time series cross sectional data. The estimator selects potential control observations for every treated unit at a specified time period by using propensity score matching and weighting schemes. First, they use the exact matching to align the treatment time for a specific time span creating the matched sets for treated units. Then, they refined these matched sets by using the caliper matching of pre-treatment-outcome and additional covariates. Finally, the apply a DID estimation as weighted average of individual differences, namely they compare the outcome development of the treated with the average outcome in each refined matched set. Moreover, as measure of matching quality, they compare first, the standardized mean difference between a treated and its matched control in each covariate at each pre-treatment time period, and then considering the aggregate measure of standardized mean difference among all treated observations for each covariate and each time period, and the average treatment effect for the treated is thus a weighted average of different observation periods.

Dettmann et al. (2020) define a flexible procedure in the definition of treatment start and treatment duration, including individual treatment time information from the panel into the matching process, removing potential calendar time effect. Also, it ensures that varying treatment phases can be properly accounted and that the point in time when an individual is compared to his 'statistical twin', can be exactly determined. Moreover, differently from other approaches (Callaway and Sant'Anna, 2019 and Imai et al., 2019) the statistical matching procedure is based on a textitweighted average of scale-specific distance function (as a combination of mean absolute difference for continuous variables and the generalized matching coefficient for categorial ones). This statistical distance function considers only the exact similarities and disparities regarding the individual covariates, so, the overall indicator reflects the comparability of the observations considering only the relevant similar or dissimilar covariates. Additionally, it helps to account potential anticipation of a treatment that may lead to a temporary change in the behaviour of the applicants (i.e. Ashenfelter's dip) and determines exactly the matching and the outcome observation time (concerning the treatment start).

As this approach is a combination of propensity-score matching and difference-in differences methodology, it assumes that unobservable individual characteristics must be in variant over time for units with the same observed characteristics. Thus, they adopt "*conditional parallel trend assumption*" (as proposed by Callaway and Sant'Anna, 2019) instead of the *conditional independence assumption* (CIA) required for matching and the *common trend assumption* necessary for DID. Following Callaway and Sant'Anna (2019) assumption, they assume that the anticipation of the treatment (anticipation effect), δ, is restricted to all "eventually treated"' groups[6]. Given that, conditioning on covariates (x), the average outcomes for the group first treated in period g and for the "never-treated" group would have followed parallel paths in the absence of treatment (*Conditional Parallel Trends based on a "Never-Treated"*). This can be particularly important in cases where there are covariate specific trends in outcomes over time and when the distribution of covariates is different across groups.

Furthermore, the *common support condition* must be satisfied as required by matching procedure. Additionally, following Imai et al. (2019) they assume that the potential carryover effects do not influence the matching variables at the matching time. Finally, the stable unit treatment value assumption for matching (no spillover effects) is reached.

Now, we explain in detail how the flexible conditional DID works. It is based on two processes: the pre-processing, where the original dataset is rearranged creating several groups composed by one treated unit, containing al pre-treatment features necessary for matching process (benchmark selection).

[6]When $\delta = 0$, it imposes a "no-anticipation" assumption.

In the step, we specify for all Trusts and years in the dataset, the treatment variable, and a set of pre-treatment characteristics necessary for matching process (called *matchvars* for matching and *tchvarsexact* for the exact matching, respectively). Moreover, it is defined a relative time specification, according to the treatment start, that defines the time of matching (called *matchtimerel*), namely the time when the matching process is conducted. The result of this pre-processing is a temporary dataset with information that is crucial for the use of the processing step.

Then, the matching process (flexpaneldid processing) selects a set of treated units considering only those observed just at the individual matching treatment start. The matching algorithm, by using a set of pre-treatment characteristics (included in matchvars and tchvarsexact option, respectively), selects one or more statistical twins among these pre-selected units. Finally, the observation time of matching variables and the outcomes are normalized such that they are measured related to the individual treatment start.

The second step (processing process) estimates the average treatment effect for the treated by using the matching procedure that allows eliminating any systematic differences after conditioning on observables. Such differences may arise, for example, because of selection based on unobservable characteristics, or because activity outcomes for merged and non-merged Trusts may be measured in different ways. Also, when data is extracted from different sources, the identification conditions required for matching may be violated. In general, the DID strategy overcomes this problem by allowing for temporally invariant differences in outcomes between treated and untreated. By using the flexible conditional DID framework, we allow for average treatment effects for Trusts in a DID setups considering multiple time periods and variation in treatment timing. The exact matching is required for this phase. As we describe above, the exact matching option (tchvarsexact) has already been required the pre-selection process to consider the time information. Unlike Dettmann et al. (2020), we consider only the mean absolute difference for continuous variables[7] calculated using the normalized absolute difference of the respective variables for a treated Trusts i and a non-treated Trusts j, as follow:

[7] They apply a matching procedure based on a combined statistical distance function as a weighted average of scale-specific distance functions taking into account the mean absolute difference for continuous and the generalized matching coefficient for categorical variables

$$ADC_{n,ij} = \frac{1}{N_c}\sum_{n=1}^{Nc}\frac{|x_{n1} - x_{nj}|}{diff_{max}(x_n)} \tag{1}$$

where N_c is the number of continous variables, $||$ denotes absolute values, and $diff_{max}(x_n)$ is the maximum observed difference of variable x_n

Finally, the average treatment effect for the treated (ATT) is estimated on matching process, considering the mean absolute difference for continuous variables selected (as above described), the relative time specification and the pre-treatment outcome (*outcometimerelstart* and *outcomedev* in Stata options, respectively). The first defines the end of the outcome development in relation to the treatment starts[8] , while the second sets out the outcome development, where the its start and end are defined on the period before the treatment starts[9]

Thus, the flexible conditional DID allows on the comparison of the mean of the individual differences in outcome development between the treated firms i and their respective controls j, in contrast of canonical DID model that compares the mean outcome in the treated and the control group.

We can define the estimator of the individual comparison as the following:

$$\begin{aligned}\delta(F,L) = E\{Y_{i,t+F}(X_{i,t} = 1, (X_{i,t-1} = 0, \{X_{i,t-l}\}_{l=F}^{L})-\\ E\{Y_{i,t+F}(X_{i,t} = 0, (X_{i,t-1} = 0, \{X_{i,t-l}\}_{l=F}^{L})|X_{i,t}, (X_{i,t-1}\}\end{aligned} \tag{2}$$

where F is outcome development (post-treatment time), L is the pre-treatment time, with $1 < l < F$. $X_{i,t-1}$ and $X_{i,t}$ are treated units and $Y_{i,t+F}(X_{i,t} = 1, (X_{i,t-1} = 0, \{X_{i,t-l}\}_{l=F}^{L})$ is the development output under treatment; $Y_{i,t+F}(X_{i,t} = 0, (X_{i,t-1} = 0, \{X_{i,t-l}\}_{l=F}^{L})$ is the potential outcome without treatment (so, $X_{i,t-1}$=$X_{i,t} = 0$), and $\{X_{i,t-l}\}_{l=F}^{L})$ are the pre-treatment characteristics.

[8]For example, if we define the "*outcometimerelstart(1)*", we will observe the outcome development from the individual treatment start to one year after the start of the treatment.

[9]For example, if we define "*outcomedev(-4 -2)*", we set out the outcome development in a span time from four to two years before the individual treatment starts.

The matching algorithm used in our analysis is the nearest neighbour matching with replacement.

The Stata comand *flexpaneldid* also runs a t-test with corrected standard errors to draw causal inference in the presence of non-random sampling[10]

The heterogeneity of treatment duration is due to the presence of different observed period among the treated individuals (Trusts). Hence, the average treatment effect for the treated is a weighted average of different observation periods.
We, therefore, compare the change in the level of activity and performance for merged Trusts before and after the merger with the change in the level of activity and performance for Trusts in a comparator group that are not undergoing the intervention, over different years. The DID method enables us to estimate the treatment effect of a merger on the level of activity and performance of merged Trusts.

4.2 Analysis on hospital outputs

We use a fixed effect model to identify the average effect of a merger in different years and to explore the robustness of our results. We investigate on the effect of merger on the level of hospital activities, by using the following model:

$$y_{it} = \beta_0 + \beta_1 M_i + \sum_{t=1}^{9} \beta_2 D_{it} + \sum_{t=1}^{9} \delta M_i D_{it} + \sum_{t=1}^{9}\sum_{k=1}^{13} \beta_3 X_{kit} + \sum_{t=1}^{9} \beta_4 Z_t + \mu_i + \epsilon_{it} \tag{3}$$

where y_{it} is the log of each output analysed for Trust i in year t; M_i is a dummy variable taking value equal to 1 if the trust is treated, 0 otherwise; D_{it} is a dummy variable equal to the relative difference time from the treatment start to $t+2$ (outcome development in the two years after merger) for each Trust i taking in the account pre-treatment features in the year $t-1$. For example, if we consider a Trusts merged in 2001, D_{t+2} will assume value equal to 1 in

[10]"*The correction terms are implemented by using the matching-based procedure of Abadie et al. (2004); Abadie and Imbens (2006, 2011). The number of matches is fixed to two (like the default setting in the "teffects nnmatch" comand in Stata.)*" (Dettmann et el., 2020)

2003; 2 in 2004,..., 6 in 2008; 0 otherwise. Specifically, the analysis considers all years after the exact year of the merger. The M_iD_{it} captures the policy effect for treaded Trusts over different years (defined as the relative time distance from the starting of treatment). Moreover, δM_iD_{it} captures the interaction between M_i and D_{it}, , X_{kit} is the k-th observable time-variant factor (inputs, controls, hospital characteristics) affecting our dependent variables for Trust i in year t. Z_t represents year fixed effect, μ_i represents hospital fixed effect while ϵ_{it} is the error term.

4.3 Analysis on hospital outcome

Hospital outcome is measured by a star system and it represents the hospital performance. We analyse two types of hospital performance rating the quality of services (*Performance rating – Type 1*) and for the use of resources (*Performance rating – Type 2*), both ranked from level 1, equals to the low level of performance, to the 4, that it is the high level. For both dependent variables, we apply a fixed effect ordered logit model (FE-OL) with blow up and cluster (BUC) estimator (Baetschmann et al., 2015).[11]

We investigate on the effect of merger on hospital performance by using the following model:

$$y_{it} = \beta_0 + \sum_{t=1}^{9} \delta M_i D_{it} + \sum_{t=1}^{9} \sum_{k=1}^{13} \beta_3 X_{kit} + \epsilon_{it} \tag{4}$$

where y_{it} is the log of outcome for Trust i in year t. As the fixed effect ordered logit model (FE-OL, Baetschmann et al., 2015) does not allow for factorial variables, the M_iD_{it} represents the overall interaction between M_i and D_{it}, where M_i is a dummy variable taking value equal to 1 if the trust is treated (onwards and subsequent merger), 0 otherwise; while D_{it} is a dummy variable equal to the relative difference time from the treatment start to $t+2$ (outcome development in the two years after merger) for each Trust i taking in the account pre-treatment features in the year $t-1$. X_{kit} is the k-th observable time-variant factor (inputs, controls, hospital characteristics) affecting our dependent variables for Trust i in year t and ϵ_{it} is the error term. As a change in the offer of services provided might have an impact on

[11] The BUC estimator ("Blow-Up and Cluster") is an alternative to the DvS estimator to avoid the problem of small sample sizes associated with some cut-off values. It involves estimating the model using $K-1$ cut-offs simultaneously, subject to the restriction that $\beta^2 = \beta^3 = ... = \beta^K$.

hospital's patient choice and on the quality of the services, our results might be helpful from a policy perspective.

4.3.1 Conditional parallel trend assumption

Our analysis is a policy evaluation of merger effect with time-varying treatment. Following Dettmann et al.'s (2020) work, we assume conditional parallel trends assumption. Thus, we assume that any possible anticipation of treatment is only related to potential treated units. Moreover, given the restricted anticipation effect, we posit that the average outcomes for the group first treated in period and for the "never-treated" group would have followed parallel paths in the absence of treatment is conditionated on a set of covariates.

To demonstrate our assumptions, we follow Cerulli and Ventura (2019) procedure or Stata command (*tvdiff*) that permits the estimation of the pre- and posttreatment average treatment effects with binary time-varying treatment, allowing also to test the parallel trend assumption. Furthermore, this command performs the common trend assumption by using both time leads and time-trend variables.

The binary treatment variable generated in the matching-DID process (*flexpaneldid*), captures the effect of merger related to its starting year on treated Trusts and his development on two years after merger. Moreover, we use the same set of covariates applied in the fixed model estimation to verify the parallel trend assumption for each hospital output.

Also, we set the time pre-treatment considering one year before merger, as average time value between relative time matching (one year before merger) and the pre-treatment development (from two to one year before merger). We align the post-treatment time to the outcome development, namely two years after merger. We align this time defintion according to flexpaneldid estimation.

We demostrate the following parallel trend assumption after conditioning on the treatment, outcome development, and covariate features

$$E[Y_{i,t+F}(X_{i,t}=0, X_{i,t-1}=0, \{X_{i,t-l}\}_{l=F}^{L}) - Y_{i,t-1}|X_{i,t}=1, X_{i,t-1}=0, \{X_{i,t-l}, X_{i,t-l}\}_{l=F}^{L}, \{Z_{i,t-l}\}_{l=0}^{L}] =$$
$$= E[Y_{i,t+F}(X_{i,t}=0, X_{i,t-1}=0, \{X_{i,t-l}\}_{l=F}^{L}) - Y_{i,t-1}|X_{i,t}=0, X_{i,t-1}=0, \{X_{i,t-l}, X_{i,t-l}\}_{l=F}^{L}, \{Z_{i,t-l}\}_{l=0}^{L}] \quad (5)$$

where F is outcome development (post-treatment time), L is the pre-treatment time, with $1 < l < F$. $X_{i,t-1}$ and $X_{i,t}$ are treated units and $Y_{i,t+F}(X_{i,t} = 1, (X_{i,t-1} = 0, \{X_{i,t-l}\}_{l=F}^{L})$ is the development output under treatment; $Y_{i,t+F}(X_{i,t} = 0, (X_{i,t-1} = 0, \{X_{i,t-l}\}_{l=F}^{L})$ is the potential outcome without treatment (so, $X_{i,t-1}$=$X_{i,t} = 0$), $\{X_{i,t-l}\}_{l=F}^{L})$ are the pre-treatment characteristics, and $\{Z_{i,t-l}\}_{l=0}^{L}$ are feature of covariates.

We use the fixed effect model (as expressed in equation 6) for each hospital output and outcome. In this model we verify the conditional parallel trend by using for each hospital output and outcome, their average development in the two years after merger (as applied in equation 3 and 4) starting, but here we constrain the analysis period considering one year before and two years after merger.

$$y_{it} = \beta_0 + \{\sum_{t=1}^{9} \delta D_{it}\}_{t-2}^{t+1} + \sum_{t=1}^{9}\sum_{k=1}^{13} \beta_3 X_{kit} + \sum_{t=1}^{9} \beta_4 Z_t + \mu_i + \epsilon_{it} \quad (6)$$

where y_{it} is the log of outputs and outcomes for Trust i in year t. D_{it} is is a dummy variable equal to 1 from the treatment start to t+2 (outcome development in the two years after merger) for each Trust i taking in the account pre-treatment features in the year $t-1$. δ_{-1} and δ_{-2} are the coefficient that measures the impact of the treatment one and two periods after its occurrence, respectively. δ_{+1} is the coefficient that measures the impact of the treatment one before its occurrence. X_{kit} is the k-th observable time-variant factor (inputs, controls, hospital characteristics) affecting our dependent variables for Trust i in year t. Z_t represents year fixed effect, μ_i represents hospital fixed effect while ϵ_{it} is the error term.

We verify conditional parallel trend assumption:

1) by using the "leads", we test the null hypotesis:

$$H0 : \delta_{+1} = 0 \tag{7}$$

Accepting the H0 the condition for parallel-trend is hold.

2) by using the "time-trend", we test the null hypotesis:

$$H0 : \delta_0 = 0 \tag{8}$$

Accepting the null $H0 : \delta_0 = 0$ implies accepting that the parallel-trend assumption is not violated, namely no "anticipation effects"

5 Data

5.1 Data Sources

Our data is longitudinal, available annually for a period of 9 years from 2000 to 2008. It contains information on all acute, specialist and teaching hospitals in England with a unique identifier for each hospital. Our unique dataset combines information from several data sources: administrative data providing information on performance, as well as hospital characteristics, extracted and/or derived from the Hospital Episode Statistics (HES); the Hospital Activity Statistics (HAS); the NHS Foundation Trust Directory; the Medical and Dental Workforce Census (Department of Health), and from individual hospitals' websites. Our data include 1-year pre-treatment policy (year 2000) and 8 years of data post policy. The dataset contains 1,581 observations for: 195 hospitals in year 2000, 186 in year 2001, 175 in 2002, 172 in years 2003, 2004 and 2005, 171 in 2006 and 169 hospitals in years 2007 and 2008. According to Table 1, the number of hospital providers in England has in fact decreased by 13%, from 195 acute and specialist Trusts in 2000 to 169 in 2008.

Table 1: Hospitals by their merging status over time. England, years 2000-2008

Year	Total number of Trusts	Number of merged Trusts	Number of merging Trusts	% merged Trusts	% merging Trusts
2000	195	0	19	0%	10%
2001	186	10	21	5%	11%
2002	175	10	6	6%	3%
2003	172	3	0	2%	0%
2004	172	0	0	0%	0%
2005	172	0	2	0%	1%
2006	171	1	3	1%	2%
2007	169	2	0	1%	0%
2008	169	0	0	0%	0%

5.2 Variable Definitions and Measurements

5.2.1 Dependent Variables

We consider different hospital outputs measured by several variables to account for various hospital services provided and their possible combinations. In particular, we consider the following hospital activities: number of inpatient spells, number of elective admissions, number of emergency admissions, number of patients attending the first outpatient appointment, number of patients attending first A&E, number of day cases (day hospital or day surgery).[12] The output analysis is completed with three more dependent variables built on a selection of the above variables: the proportion of elective and emergency admissions, the share of inpatients over outpatients, and the share of day cases over elective admissions. These extra variables will be used to assess if and how hospital mergers alter the combination of services provided. As Trusts differ mostly in the amount of services provided, rather than the decision to provide as service at all, we will focus here on the intensive margin of the degree of providing a service, which we will measure by a log-transformation of the dependent variables.

In order to assess the effect of hospital mergers on outcome, we also consider two different measure of hospital performance, built combining the star rating performance index with either the quality of services index or the use of resources index. The star rating is a composite index

[12] We exclude from the analysis both subsequent outpatient attendances and total outpatient attendances, in order to avoid patients' double counting. The same reasoning holds for A&E attendances, thus we exclude subsequent A&E attendances and total A&E attendances as well.

score that places Trusts into one of four categories of performance: from highest (awarded three stars) to poorest (awarded zero stars). The star rating is defined over 43 performance indicators (`http://ratings2004.healthcarecommission.org.uk`): 9 key targets including financial management; 12 clinical indicators (e.g., number of deaths within 30 days of a heart by-pass surgery); 15 patient focus indicators (e.g., waiting time for breast cancer treatment); and 7 indicators for capacity and capability (e.g., data quality). From 2005/06 the star rating system has been replaced with a new system, referred as the annual health check. The annual health check is a more sophisticated performance rating that places Trusts into one of four categories of performance, from highest (awarded three stars) to poorest (awarded zero stars), based on two aspects: efficient use of hospital resources and quality of the services provided but the assessment is now based solely on clinical indicators (e.g., waiting time for cancer treatment; clinical responsiveness in the provision of thrombolysis) and patient level indicators (e.g., smoking during pregnancy and breastfeeding initiation).

5.2.2 Policy variables

To assess the impact of the reorganizational change due to hospital merger on our output and outcome measures, we construct a dummy variable for hospital merger status. Specifically, the *merged_forward* equals to 1 in the year the new merged hospital starts its activity and subsequent years, and zero otherwise.

5.2.3 Control variables

To account for other variables that may be correlated with our output measures, as well as the key variable of interest (the policy variable), we control for various hospital characteristics. To account for other variables that may be correlated with our output measures, as well as the key variable of interest (the policy variable), we control for various hospital characteristics.

We include the Foundation Trust status, *FT*, that is equal to 1 in the year the hospital becomes a Foundation Trust and subsequent years, and zero otherwise[13]. The main reason why

[13] In 2003 the UK Parliament passed the HSC Act 2003, a bill that allowed some NHS Trusts to acquire a new legal status – Foundation Trust – and become non-profit public benefit corporations in charge of providing goods and services for the purposes of the NHS in England (HSC Act 2003, Part 1, section 1). Several hospitals have thus experienced an organizational change by acquiring this status. FTs have acquired a new set of freedoms in comparison to non-FTs. Specifically, FTs have a higher degree of independence from the Department of Health

we also account for this policy is related to the fact that many mergers were motivated to allow NHS trusts to acquire the FT status through the merger transaction (Collins, 2015). Similarly to other healthcare studies, we include inputs of a production function to control for overall hospital capacity (e.g., Wagstaff, 1989; Duggan, 2000; Horwitz and Nichols, 2009). Thus, we consider as measure of capital *total available beds*, *total available acute beds*, and total operating theatres as measures of capital. According to OECD (2020) definition the number of total available beds is a "*measure of the resources available for delivering services to inpatients in hospitals in terms of number of beds that are maintained, staffed and immediately available for use. Total hospital beds include curative care beds, rehabilitative care beds, long-term care beds and other beds in hospitals*", while the *acute beds hospitals* are definied as "*beds accommodating patients in a hospital or hospital department whose average length of stay is 30 days or less until the 1980s and 18 days or less after* (OECD, 2001). We can use the broad definition of total beds as a proxy of hospital size while the detailed explanation of acute beds can be a proxy of hospital efficiency. Additionally, we consider *proportion of medical staff* and the *proportion of non-medical staff* as measure of labor.

Second, to account for differences in the complexity of the patients among hospitals, similarly as other studies (e.g., Herr, 2008; Bloom et al., 2005), we include *ALOS* - average length of stay as more severe patients stay in hospital longer. ALOS is often used as a patient complexity measure since it allows capturing the variation of severity not only between, but also within diseases (Wagstaff, 1989). Moreover, we control for *median waiting time* to account for differences in the quality of the service provided and for the *number of tests* dispensed to account for overall hospital use of resources. We also account for differences in the population served by considering the *proportion of patients aged 0-14*, the *proportion of patients aged 60* and

and more freedom in their corporate governance decisions. For example, more control over appointing and rewarding staff, directors and board members; as well as more control over their long/short term strategies and the way services are managed and operated. More decentralization, managerial and governance flexibility also brings more financial freedoms. In particular, FTs can retain their surpluses, obtain faster access to capital by raising it from both the public and private sectors, invest in the best mix of services for their patients and thus develop business strategies that better coordinate their financial and operating structure with the needs of their local communities. Moreover, these freedoms should also facilitate outsourcing of both medical and non-medical services (e.g., laundry, cleaning, catering, lab analysis, etc.) allowing further increases in efficiency. As a result of all these organizational changes, one can expect that FTs would be encouraged to change their behaviour, and ultimately their performance (HSC Act 2003; Commission for Healthcare Audit and Inspection, 2005). In fact, FT policy advocates tend to argue that the new freedoms of FTs should lead to their better organizational performance, including lower costs and improved efficiency.

over, while the *proportion of female* patients represents the share of female population over the number of inpatient admission spells. Since we estimate semi-log specifications, we transform continuous variables into logarithms.

Moreover, we construct an Herfindahl Index (HHI) to capture market competition, using hospital market shares of bed days within a 30 miles radius for each hospital[14]. Given that the HHI already reflects percentages, in our estimations we include HHI in levels rather than the logged values. Finally, we include a dummy variable teaching that is equal to 1 if the Trust is a *teaching* hospital, and zero otherwise; and a categorial variable *Performance rating - type 2* ranging from 1 (the poorest level of composite performance) to 4 (the highest level of composite performance).

To account for other variables that may be correlated with our outcome performance, as well as the key variable of interest (the policy variable), we control for the several hospital characteristics, according to the type of performance rating. As the *Performance rating - type 1* measures hospital performance based on quality of services, we include in the analysis the *total available number of acute beds* and the *proportion of non-medica staff* as measure of capital and labor, respectively. Also, we consider the *ALOS*, and whether it is a teaching Trust (*teaching*), and a variable that captures the interaction between HHI on bed days (radius 30 miles) and teaching (*Interaction 1*).

To control for the performance rating based on use of resources (*Performance rating - type* 2), we consider some financial measures, such as *surplus* (that measures surplus or deficit for the financial year), expenditure for directors (*Directors' costs*) and *RCI* (including excess bed days). Moreover, we include the *proportion of non-medical staff* as measures of labor, *median waiting time*, *teaching* and also *Interaction 1* and *Interaction 2* (measuring the interaction between HHI on bed days - radius 30 miles and specialist and FT, respectively).

5.3 Characteristics of data structure

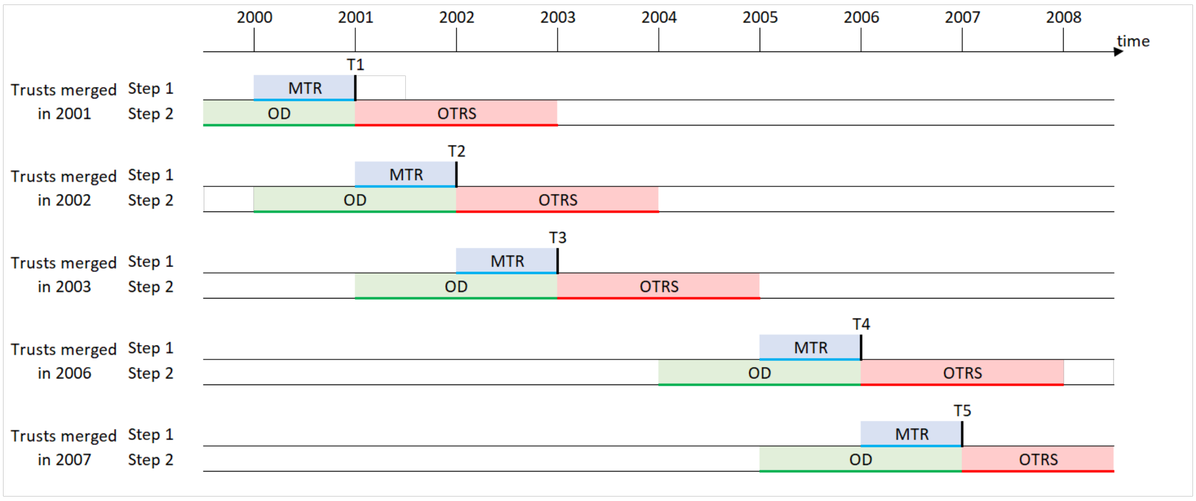
Our panel dataset collects all information on treated and untraded observations from 2000 to 2008 where the treatment starts in different years among different Trusts, creating an unbal-

[14]This is a plain measure of competition defined on the simple number of neighbour competitors and used to control for non-price competition (e.g., quality and/or demand competition), instead of price competition (e.g., technical efficiency). The value within 30 miles was chosen on sensitivity analysis' results.

anced panel. To restore the balance between treated and untreated, we need to capture the time-varying effect of treatment among different Trusts. Following Dettmann et al. 2020) approach, we adopt a flexible data structure, namely a 'staggered treatment adoption'. Figure 1 shows five years of merger Trusts (T1 in 2001, T2 in 2002, and so on, until the T5 in 2007). Each Trusts group has a set of pre-treatment, treatment, and post-treatment characteristics which can influence the whole setting analysed and each phase of flexible panel DID framework. The flexible conditional DID framework runs by two processes: the pre-processing and processing steps (called Step1 and Step 2 in Figure 1, respectively).

In the pre-processing (Step 1) we organise our dataset considering a set of selected pre-treatment characteristics (matchvars and exactmatchvar) and the relative time specification, based on the year of merger, that defines the time of matching (matchtimerel, MRT in Figure 1). We fix MRT one year before merger. Then, in the matching process (Step 2 in Figure 1), we estimate the average treatment effect for treated by using the pre-processing data time structure. Here, the matching process runs by considering the mean absolute difference for continuous variables selected and the pre-treatment outcome development and its relative time specification. As the pre-treatment outcome is a selected period of outcome development before the treatment starts, we define this span time considering from two to one year before merger (OD in Figure 1), and we assume that the outcome will develop over the two years after merger to have a more balanced time span with the pre-treatment outcome development (OTRS). The merger process is complex, and it is articulated in several phases. According to Fullup et al. (Fulop, King, and C. 2005), the "merger is a complex organisational phenomenon without clear boundaries, where the begins and the ends are not obvious". The articulated proposal merger phase influences it. Collins (pp. 23, 2015) estimates that providers "can take from one to two years to identify their preferred merger partners, and one to four years to gain approvals and complete the merger process". For those reasons, we identify our pre-treatment periods as the average time of these two proposal phases.

Figure 1: Data structure in the Flexible conditional Panel DID



5.4 Aggregate Data Patterns and Descriptive Statistics

Table 2 shows descriptive statistics for our sample overall. Our dataset is cleaned by the presence of outlier values. Among 1,581 observations in our sample about 2% of hospitals merge in a given year. Activities of hospitals are represented by several measures of efficiency and performance. In particular, hospitals are used to reduce admissions (lessening their high costs) in favour of outpatient activities. On average, outpatient first attendances are quite 10,000 units greater than inpatients spells. Also, data show that on average, hospitals tend to program admission activity to use more efficiently their resources. Indeed, average number of elective admissions (43430.35) are almost double than emergency admissions (24073.54). These data confirm a decreasing utilization of emergency care in favour of planned care, as also shown by the value of elective-emergency ratio. However, data highlights a higher utilization of emergency care, as shown by A&E first attendance mean (78748.18).

On average, the value of planned care (elective admissions) is double than planned care without overnight remaining is not so high (day cases is only 22944.84).

The ratio between planned and unplanned hospital activities (namely, elective-emergency ratio), shows that the average value is more than three hundred per cent. This is due to the presence in the sample of hospital trusts with a constant emergency activity over time. On average the proportion between impatient admissions and outpatient activities is around 94%, while the daycase-elective ratio is around 51%.

We analyse dimension of each hospital by considering capital inputs, such as number of total beds (on average almost 742) and acute beds (on average almost 582), number of available theatres (on average almost 16) and labour inputs, where the number of non-medical staff is higher than medical staff (non-medical staff is around 89% of total staff while the medical staff is only the 11%).

The median value of waiting time in days (around 51 days) and the ALOS in a hospital (between 5-6 days) are used as measure of efficiency and quality of hospital services. ALOS is calculated as the average number of days spent by each inpatient in hospital. This variable is included in the empirical specification to control for the outpatient variation among inpatients not captured by the number of admitted patients. The total number of tests, including CT, MRI, obstetric and non-obstetric ultra-sound tests, radio isotopes and radio-graph tests and fluoroscopy is on average around 179000. The CT and MRI around 7% and 4% of the total number of tests respectively, are used for the sensitivity analysis. Competition between Trusts is measured by the number of bed days within 30 miles[15] (around 48 km) range of each Trust. This is a plain measure of competition defined on the simple number of neighbour competitors and used to control for non-price competition (e.g., quality and/or demand competition), instead of price competition (e.g., technical efficiency).

We consider three groups to control for population's differences: the proportion of young people (about 14%), of elderly population (about 41%) and of females (51% of total population). Moreover, we capture both hospital characteristics before hospital mergers in the matching process and to better explain the effect of hospital mergers on performance by using several financial measures. As we know from literature on hospital mergers ((i.e. Fulop et al., 2005; Goddard and Ferguson, 1997; Gaynor et al., 2012; Propper et al., 2004; Collins, 2015), several economic, institutional, social, and political drivers led to enhance the process

[15]Also, we explore how different distances among competitors affect Trusts merged by using the HHI on bed days within 15 and 20 miles in a additional sensitivity analysis.

of merger. In particular, we consider hospitals' financial and economic conditions as the most relevant merger drivers. In line with Heckman et al. (1997, 1998), economic environment influences the performance and output of firms, so we should consider it when analysing treatments effects in order to overcome the 'calendar time effect'. Thus, it is suitable a synthetic measure of potential efficiency in the investment. In the general definition, the Return on Investment (ROI) is a performance measure used to evaluate the efficiency of an investment comparing the amount of return on a particular investment, relative to the investment's cost. In our definition, we use the proportion of surplus on the total amount of expenditure as measure of hospital performance that is around at 2.5%. Also, we consider the proportion of total director's and managers' cost on the total expenditure (around 3.3%) as useful indicator in the merger process in order to control the presence of high costs. Then, the total amount of expenditure and the retained surplus of Trusts can help us to take in the account the whole Trusts' financial condition. The total amount of expenditure (on average £190579) collects all financial and operating costs. The retained surplus is a highly volatile measure that shows whether the NHS trust has achieved a breakeven in the year. It is the surplus after paying all public dividend capital dividends and on average is around £ -140,09[16]).

We investigate the effect of hospitals merger on performance, considering three financial measures such as RCI (including excess bed days), surplus (or deficit) in a given financial year, and directors costs[17] (on average £100.46, £4661.76 and £735, respectively). The financial management measure, the RCI, can be used as a Trust's measure of efficiency related to the costs . The average measure of RCI above 100, shows Trusts with relative efficiency-hig costs. Additionally, we consider the surplus, that is the Trusts' net amount of operating and financial revenues and costs, and directors' costs, that represent a high cost item in the Trusts budget. The latter is around the 4% of total expenditure. All values of financial measure presented in

[16]The statuary duty establishes that breakeven is measured 'taking one year with another', but it is normally measured over three years, so a potential retained deficit for just one year does not mean a violation of the statutory duty (NHS trust accounts, Revised 2010 edition

[17]Efficiency Trusts is defined on RCI. RCI is an activity-weighted average of a Trust's Healthcare Resource Group (HRG) unit costs relative to the national average. With the introduction of HRG casemix funding under PbR, reference costs are used to set the national tariff on which Trusts are reimbursed (Audit Commission, 2004). Thus, whether a Trust has an RCI below 100 means that the Trust has the relative efficiency-low-cost (high efficiency), while with an RCI above 100, the Trust is not efficient because of high costs (Department of Health, 2006)

the analysis, are expressed in thousands of £).

Finally, we analyse hospital characteristics. We have around 34% of teaching hospitals and two measures of the interaction between level of Trusts' competition with hospital teaching (interaction 1) and with Foundation Trust status (Interaction 2). We include two types of quality variables to take in account performance. Both are ranking from 1 to 4 and their average score is around 2.9.

Table 2: Summary statistics - full dataset

Variable	N	Mean	Std. Dev.	Min.	Max.
Dependent variables					
Number of inpatient spells	1579	67475.69	38128.685	2264	232033
Number of elective admissions	1577	43430.348	25184.739	2119	154926
Number of emergency admission	1579	24073.538	14090.597	13	85135
Outpatient first attendances	1576	77768.218	44429.45	1006	257783
A&E: first attendances	1568	78749.177	45778.437	0	279532
Number of daycase	1574	22944.841	14652.876	0	82856
Elective-emergency ratio (%)	1577	320.788	1252.306	61.908	43969.23
Inpatient-outpatient ratio (%)	1574	93.873	47.654	6.746	818.52
Daycase-elective ratio (%)	1571	51.367	11.244	0	96.904
Policy variables					
merged forward	1581	0.111	0.314	0	1
FT	1581	0.166	0.372	0	1
Inputs					
Operating theatres	1567	15.836	9.168	0	57
Total available beds	1578	742.039	411.194	44	2838
Total available acute beds	1575	581.803	343.654	44	2142
Share of medical staff (% in ln)	1559	11.159	2.197	4.61	19.362
Share of non-medical staff (% in ln)	1559	88.841	2.197	80.638	95.39
Controls					
Average length of stay (ALOS)	1577	5.531	1.78	1	23
Median waiting time in days	1559	50.638	20.109	6	163
Patients aged 0-14 (%)	1553	13.999	12.778	0	94.988
Patients aged 60 and over (%)	1563	40.704	10.12	0	70.400
Female patients (%)	1578	50.914	6.662	29.909	112.241
Total tests	1567	178634.843	95586.141	6730	626807
CT	1571	13049.296	9953.686	0	82316
MRI	1565	6025.043	5054.216	0	34456
HHI on bd - radius 15 miles	1581	5822.067	3860.922	478.037	10000
HHI on bd - radius 20 miles	1581	4706.753	3710.223	380.523	10000
HHI on bd - radius 30 miles	1581	2984.791	2866.884	310.415	10000
RCI (including excess beds)	1386	100.463	10.703	69.092	162.489
Surplus	1298	4661.763	7106.681	-76901	73800
Retained surplus	1568	-149.087	6144.63	-84823	55990
Pseudo-ROI	1281	2.589	3.121	-19.908	13.696
Directors' costs	1282	735.274	265.352	198	2088
CEO (%)	1030	3.296	1.055	0.79	10.222
Total expenditure	1369	190579.218	126978.351	10434	845474
Hospital characteristics					
teaching	1581	0.341	0.474	0	1
Performance rating - type 1	1556	2.927	0.86	1	4
Performance rating - type 2	1556	2.867	1.008	1	4
Interaction 1	1581	218.916	916.351	0	8577.114
Interaction 2	1581	837.722	2005.164	0	10000

In the table 3, we compare the average differences of Trusts that merged in different year (considering the whole spam from year of merging to the end of our sample) with those that did not (i.e. for merged_forward=0 in all years). The mean differences for dependent variables (in the last column of Table 3) suggest that hospitals' mergers improve their efficiency. These data pattern support the expectations of many merging policy advocates, who claimed that merger processes were initially thought with the intention to create more efficient trusts (e.g., Eaton, 2005). All differences in mean are significant at 1%. In particular, the effect of merger is larger in the planned activities than in unplanned ones, as shown by elective and emergency admissions' average (31082.9 vs 18300.7, both significant at 1%). This is consistent with the fact that unplanned admissions include many unexpected treatments, such as emergencies, and if these occur, lifting inefficiency constraints and providing better allocation of resources as a result of hospital re-organization would be much more important than for treatments that can be planned in advance. However, merged hospitals do not provide significantly larger ratio in the combined outcomes of both elective-emergency and inpatient-emergency, which may further raise the question of how merger policy interacted with hospital efficiency in the short term.

Merged hospitals present a smaller average level of FT status than non-merging hospitals. The difference in mean confirms that FT status will not play a key role in the merging process in the long term. Moreover, considering capital and labor inputs, merged hospitals have significant differences in means for all inputs. However. the effect of merger decreases slightly the number of medical staff (the difference on average is only 0.5%). That is in line with general literature, merged hospitals decrease the number of medical staffs in favor of enhancing investments in capital inputs. The positive effect of merger policy is reinforced form the decreasing in the average level of average length of stays.

By analyzing financial measures, results show that the costs' reduction achievements did not reach completely through merger process. Indeed, even if the surplus increases on average in merged hospitals, at the same time, the retained surplus lessens on average. Moreover, the total expenditure and the directors' costs rise on average in the merger Trusts, even thought the proportion of managers and directors' costs on total expenditure lessen in merger Trusts. We suppose that it may be depended by a reduction of managers and, hence, by a lessening of this expenditure item in the budget. On average the level of efficiency, measure by the RCI, is lower in merged hospitals than those are not. In line with literature, our data confirm that the merger process increase the number of teaching hospitals.

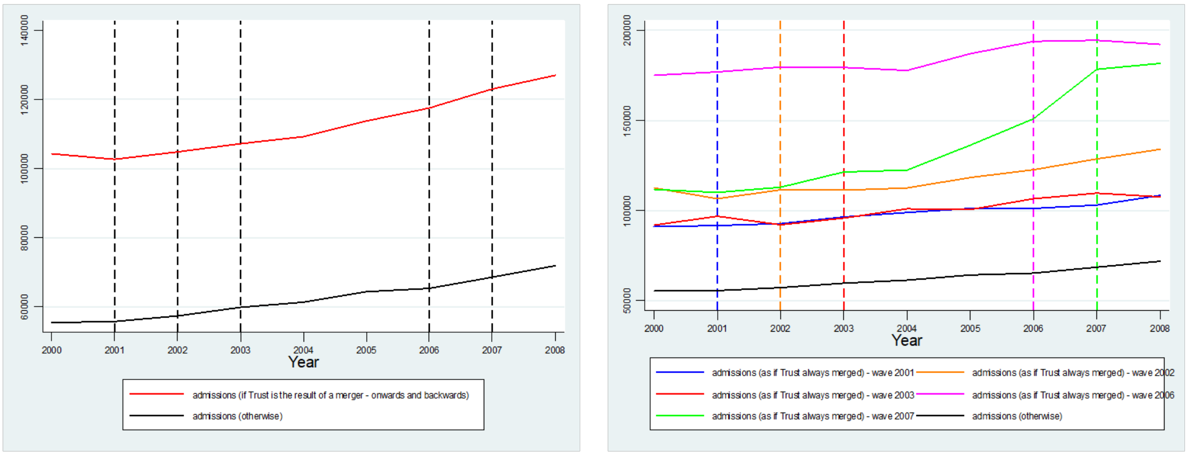
Though these correlations in aggregate data are appealing, between merged and non-merged hospitals, also suggest that hospital heterogeneity (e.g., teaching vs. non-teaching status), differences in financial conditions, but also merging in different years will play an important role when it comes to teasing out the impact of organizational change imposed by merger policy. The goal of our empirical analyses described below is to further explore these data patterns.

Table 3: Mean Comparison: Merged forward vs. Non-Merged forward (by the end of our sample period; 2000-08)

Variable	Merged forward (= 1 by the year of merger N	Mean	Merged forward (= 0 in all years N	Mean	Difference
Dependent variables					
Number of inpatient spells	175	111409.943	1404	61999.554	49410.389***
Number of elective admissions	175	71063.971	1402	39981.073	31082.899***
Number of emergency admission	175	40345.971	1404	22045.278	18300.693***
Outpatient first attendances	174	123366.845	1402	72109.045	51257.800***
A&E: first attendances	174	130871.971	1394	72243.175	58628.796***
Number of daycase	175	39328.966	1399	20895.362	18433.604***
Elective-emergency ratio (%)	175	177.626	1402	338.657	-161.031
Inpatient-outpatient ratio (%)	174	93.401	1400	93.932	-0.531
Daycase-elective ratio (%)	175	55.054	1396	50.905	4.149***
Policy variables					
FT	175	0.194	1406	0.162	0.032
Inputs					
Operating theatres	175	24.034	1392	14.805	9.229***
Total available beds	175	1184.863	1403	686.804	498.059***
Total available acute beds	175	952.714	1400	535.439	417.275***
Share of medical staff (% in ln)	175	10.709	1384	11.216	-0.507***
Share of non-medical staff (% in ln)	175	89.291	1384	88.784	0.507***
Controls					
Average length of stay (ALOS)	175	5.28	1402	5.562	-0.282**
Median waiting time in days	175	48.977	1384	50.848	-1.870
Patients aged 0-14 (%)	175	12.352	1378	14.209	-1.857*
Patients aged 60 and over (%)	175	42.489	1388	40.479	2.010**
Female patients (%)	175	50.833	1403	50.925	-0.091
Total tests	174	274983.356	1393	166599.925	108000000***
CT	175	19424.594	1396	12250.1	7174.494***
MRI	175	8915.606	1390	5661.124	3254.482***
HHI on bd - radius 15 miles	175	6906.655	1406	5687.072	1219.583***
HHI on bd - radius 20 miles	175	5096.359	1406	4658.26	438.099
HHI on bd - radius 30 miles	175	3367.654	1406	2937.137	430.517*
RCI (including excess beds)	175	102.102	1211	100.226	1.876**
Surplus	160	6480.331	1138	4406.077	2074.254***
Retained surplus	175	-338.383	1393	-125.306	-213.077
Pseudo-ROI	155	2.278	1126	2.632	-0.354
Directors' costs	155	913.497	1127	710.762	202.735***
CEO (%)	126	2.976	904	3.341	-0.364***
Total expenditure	170	287141.112	1199	176888.208	110000***
Hospital characteristics					
teaching	175	0.469	1406	0.325	0.144***
Performance rating - type 1	175	2.846	1381	2.938	-0.092
Performance rating - type 2	175	2.737	1381	2.883	-0.146*
Interaction 1	175	0	1406	770.519	-246.164***
Interaction 2	175	1377.647	1406	434.509	607.127***

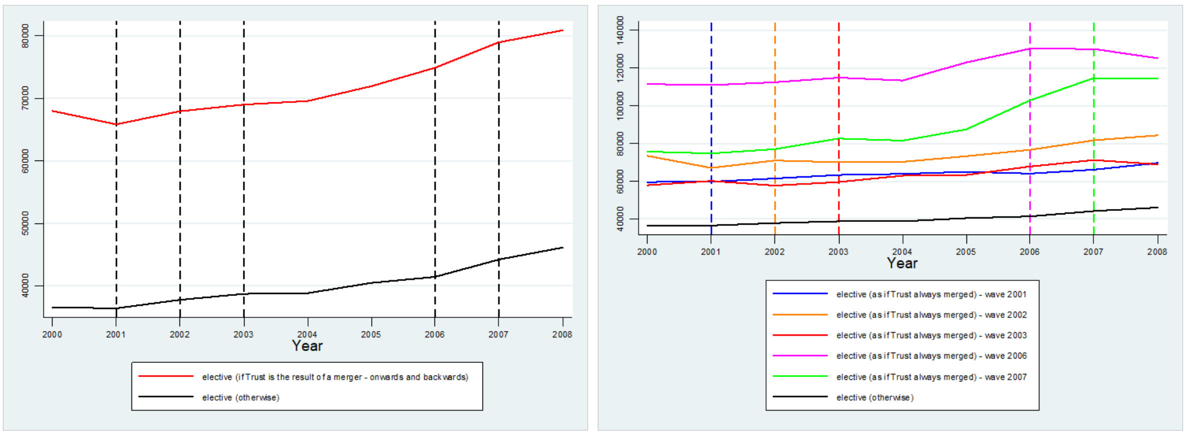
Figures 2-10 represent the level of each output and outcome for merged Trusts and for non-merged Trusts (comparator group). In the panel (a) we present for each of our dependent variables, their overall level considering all merged Trusts (red line) and comparator group (dark line). Whilst, the panel (b) shows dependent variables trends according to the year of merger (different coloured lines by year of merger) compared to non-merged Trusts (dark line). We create the level of each output and outcome over time considering the same features ex-ante and post-merger. Specifically, thought we consider a new Trusts born in 2001, called i.e. C, as the merger of two merging Trusts, i.e. A and B, we build its level of i.e. inpatient admission in the pre-merger period as the sum of the level of Trust A's inpatient admissions and those of Trust B. In that way we have one single value of inpatient admission for the new merger Trust C in the period 2000 to 2008. In general, we can observe a parallel trend if we consider the overall level of all merged Trusts and comparator group (panel a), but this condition is not always achieved if we consider the time-varying treatment effect (panel b). As we know form literature (Angrist and Pischke 2009, Krueger and Card, 2000), the DID estimate is the difference between the change in outcomes before and after treatment in the treated and controls and this estimate represent the interaction of treatment group dummy and a post-treatment period dummy. In terms of sample means, the regression to potential outcomes two-group/two-period (2x2) DID identifies the average treatment effect on the treated, under the parallel trends' assumption (namely, trend difference between treatment and comparison groups is equal to 0). Unlike, our analysis diverges from the canonical 2x2 DID set-up, as our treatment occurs at different time. Adopting the flexible conditional DID and the conditional parallel assumption, we can estimate the average treatment effect in the presence of time-varying treatment.

Figure 2: Inpatient admissions, 2000-2008



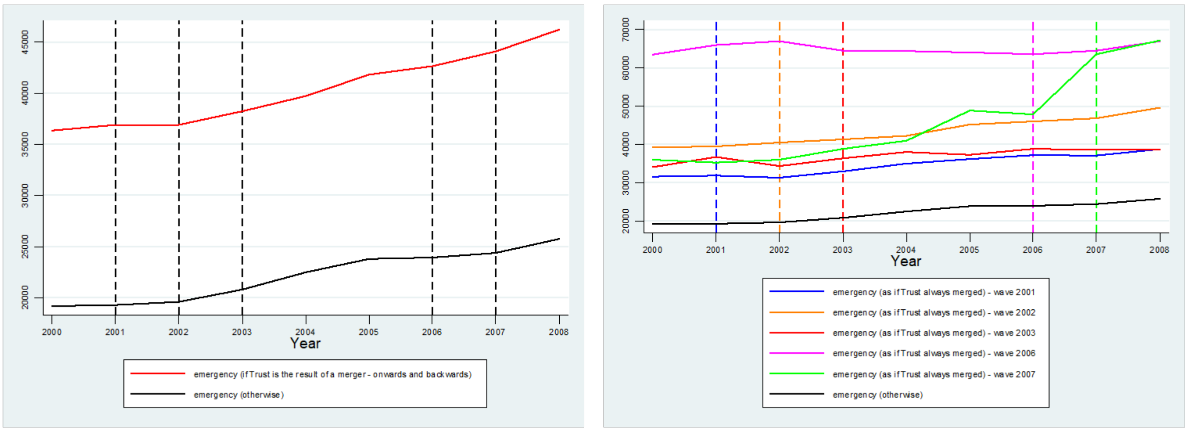
Panel (a) Panel (b)

Figure 3: Elective admissions, 2000-2008



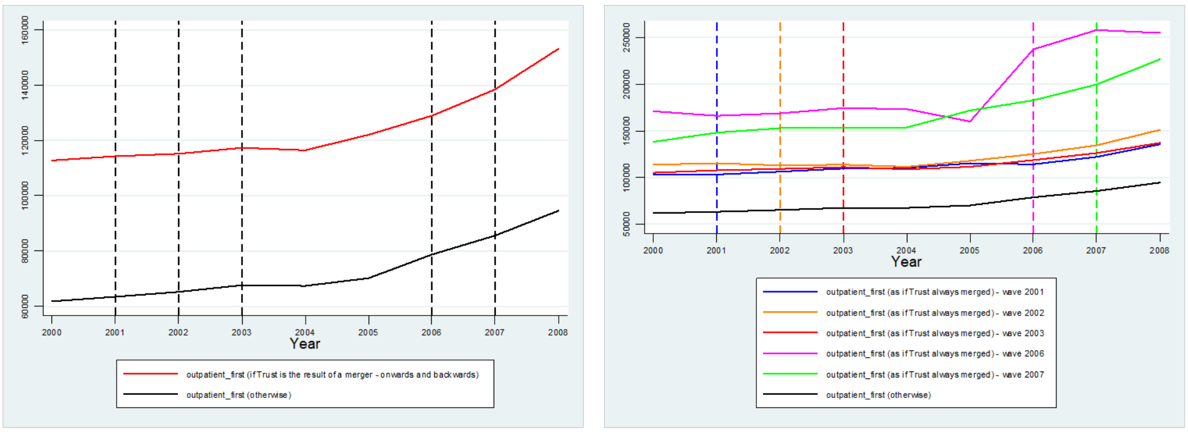
Panel (a) Panel (b)

Figure 4: Emergency admissions, 2000-2008



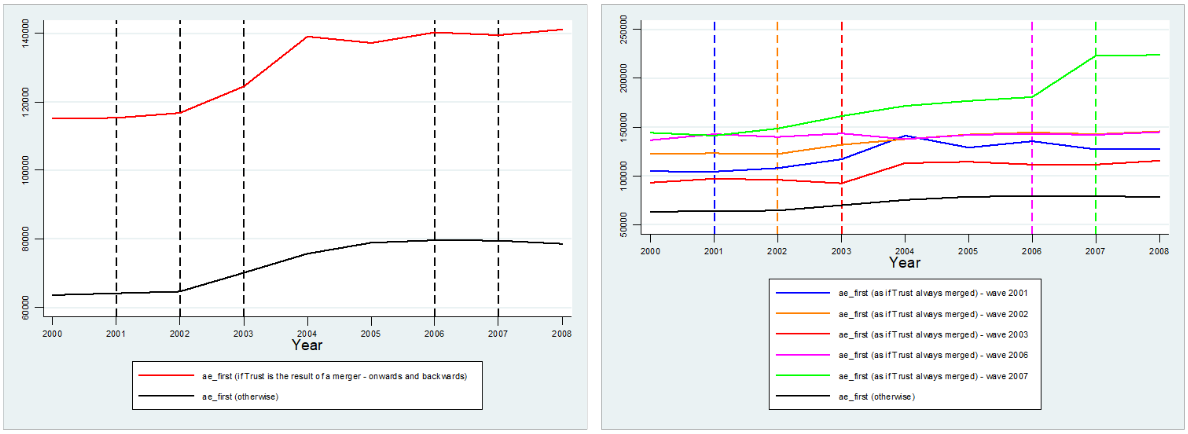
Panel (a) Panel (b)

Figure 5: Patients attending first outpatient appointment, 2000-2008



Panel (a) Panel (b)

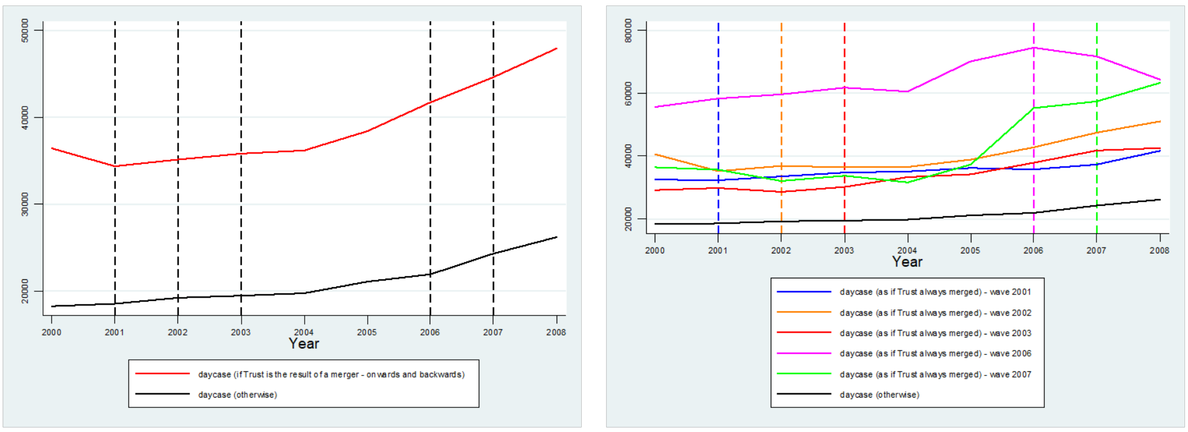
Figure 6: Patients first visiting A&E department, 2000-2008



Panel (a)

Panel (b)

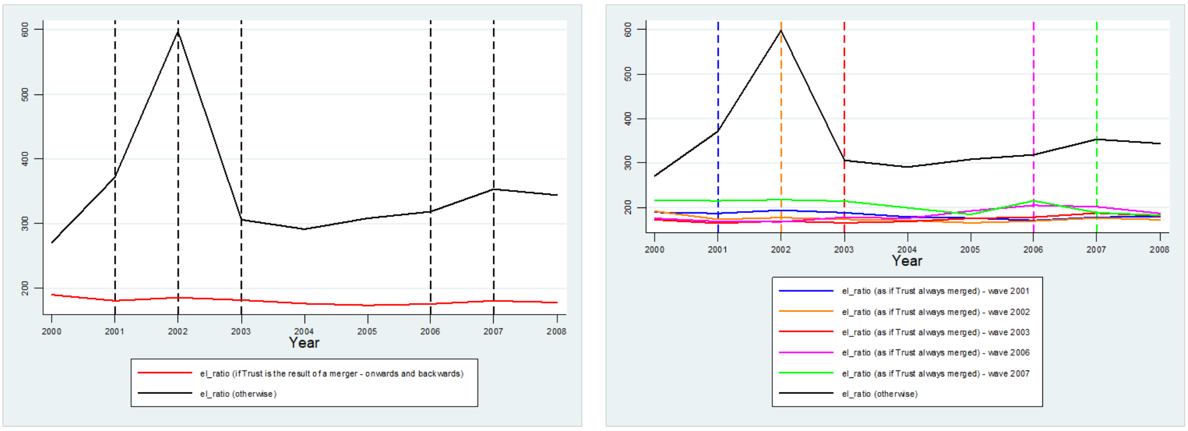
Figure 7: Inpatient admissions (with no overnight stay), 2000-2008



Panel (a)

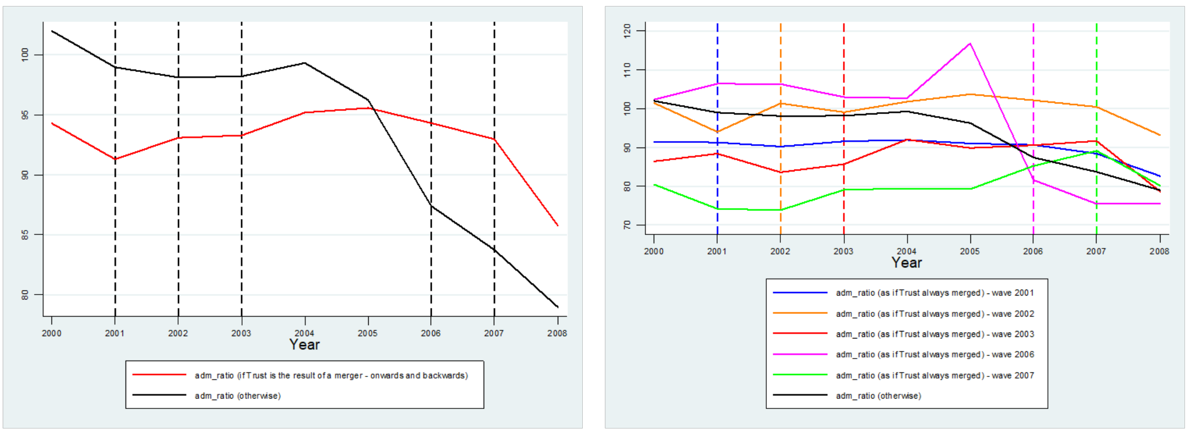
Panel (b)

Figure 8: Elective-emergency ratio, 2000-2008



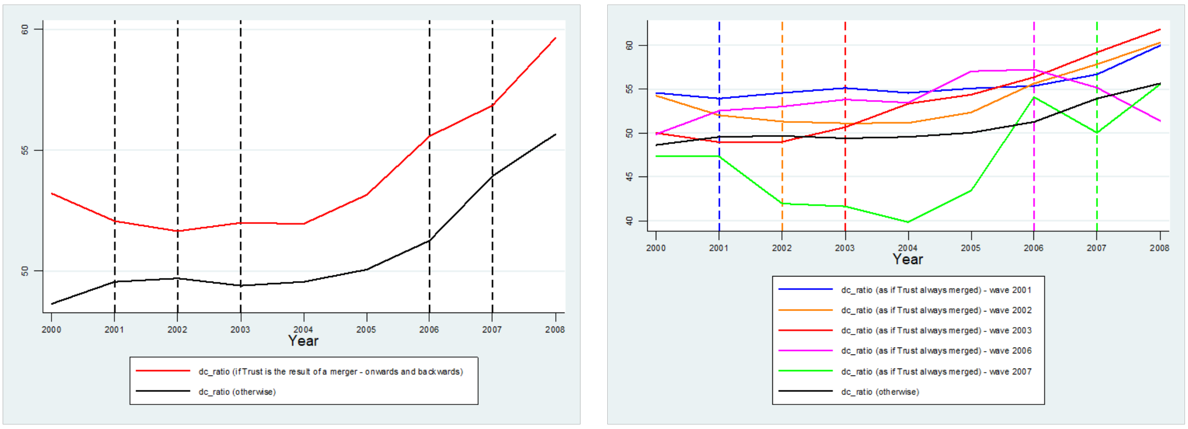
Panel (a) Panel (b)

Figure 9: Inpatient-outpatient ratio, 2000-2008



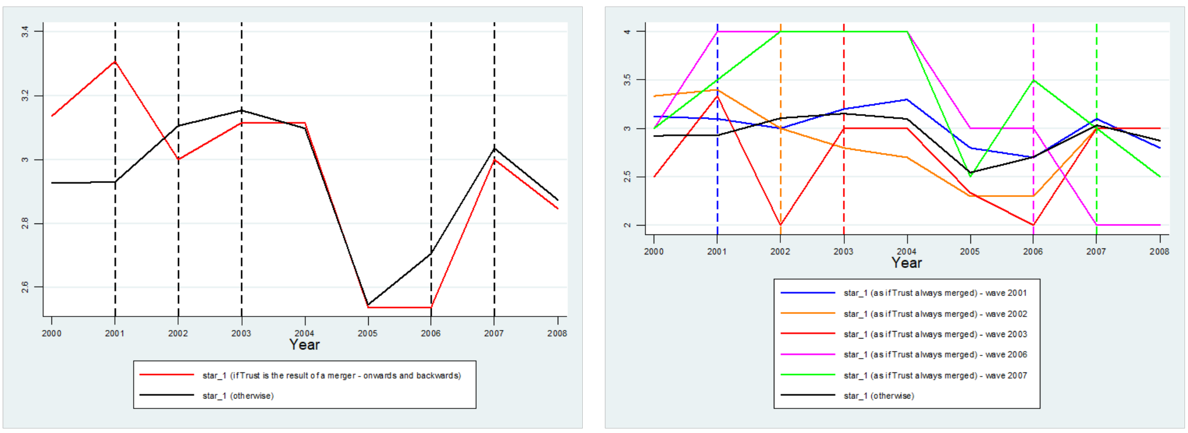
Panel (a) Panel (b)

Figure 10: Daycase-elective ratio, 2000-2008



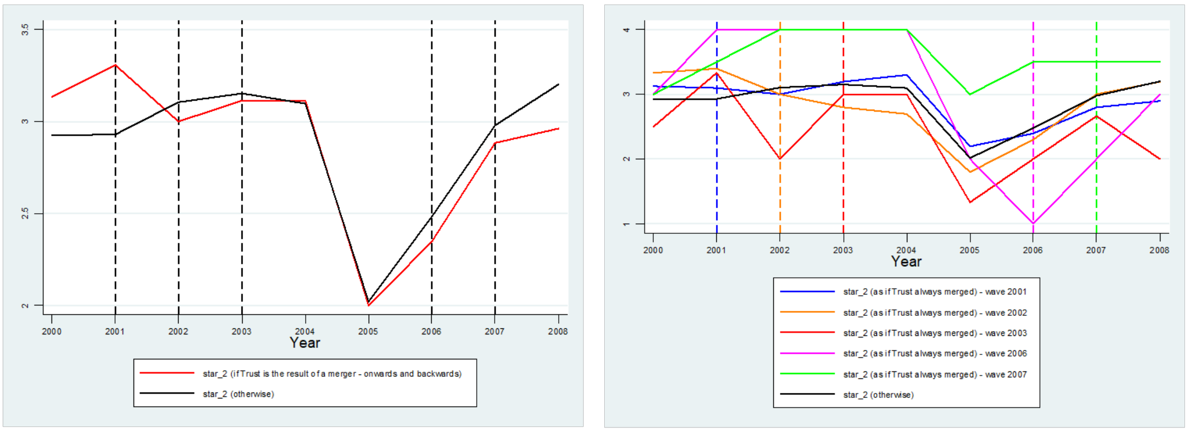
Panel (a) Panel (b)

Figure 11: Performance rating - type 1, 2000-2008



Panel (a) Panel (b)

Figure 12: Performance rating - type 2, 2000-2008



Panel (a) Panel (b)

6 Results

The aim of our analysis is the effect of merger on hospital outputs and outcomes. We use the Dettmann et al. (2020) approach (flexible conditional DID) to deal with time-varying treatment effect, as hospitals merged in different years.

To obtain a more deeply understanding of mergers effects on hospital activities and performance, we separate their results in two separate sections (results on hospital activities in 6.1 and on perfomance in 6.2).

6.1 Results on hospital outputs

6.1.1 Flexible conditional DID's results

As already explained in par. 4.1, the flexible conditional DID sets out two steps by running in Stata[18] the separate routines called "*flexpaneldid_preprocessing*" and "*flexpaneldid*[19]", respectively . As already known, several elements need to be considered when we want to investigate on the effect of merger starts in different years among different Trusts. First, we need to overcome all possible time calendar effects considering the general context and possible drivers

[18] We use both STATA 14 and 16.

[19] Further details about "flexpaneldid_preprocessing" and "flexpaneldid" are provide in Dettmann et al.'s paper (2020).

for mergers. Competition mechanisms (1991 NHS reform) and then collaborative purposes (1997 NHS reform) are used from policy makers to enhance the quality and reduce the costs of healthcare services. Trusts' economic and financial conditions were put "under the magnifying glass" of reformers with the intention to achieve more efficiency and quality. For that reason, financial variables as the pseudo-ROI, CEO, total expenditure and retained surplus can help us to capture the pre-treatment hospitals' characteristics.

In the pre-processing process, we organise dataset considering both pre-treatment hospital characteristics and the time before merger related to them. We use retained surplus for the exact matching (*matchvarsexact* option). Even thought the retain surplus is a high volatile financial measure, the choice is driven by the fact that it represents the total net surplus (or deficit) from all the years of the NHS trust's operation. When it is negative, it means that Trust has a cumulative net deficit. Moreover, that break-even was one of the key measures in the star rating performance regime. Thus, the use of retained surplus allow us to take in the account both Trusts' efficiency and quality. Additionally, we include a list of variables required for matching process such as the pseudo-ROI, that it is a measure of Trust's performance, CEO to understand the weight of managers and directors' costs on the total hospital expenditure and the total expenditure.

The exact matching option works like a filter and it is already relevant in this step, as it selects only units with identical values among variables defined in matching list. Hence, only non-treated Trusts with identical values are chosen as potential partners for every treated Trusts. All pre-treatment variables are in log-form. Moreover, the pre-processing process organise data selecting 26 treated units according to the relative time of matching, that we fix one year before merger.

Then, we run the matching process runs considering the data structure as built in the pre-processing step. As our purpose is to observe possible changes in the number of hospital outputs considering time-varying treatment effect, we need to define the textitoutcome development time span considering the pre-treatment characteristics and time before and after merger. Thus, the relative pre-treatment time (or pre-treatment development) is from two to one year before merger and its development post-treatment is from merger start to two years afterwards. Defining the outcome development time span, we include the size of the effect that may be depend on the length of exposure of it and, thus, we overcome the phenomenon called "dynamic treatment effect". The definition of relative time matching and both pre-treatment relative time and post-treatment outcome development are based on merger process Collins' assumptions (more details in par. 5.3).

After pre-processing, we execute the matching procedure taking in the account the individual (internal generated) indication of possible partners for each treated unit. In tables 4-12 are shown matching results as following explained. We illustrate a short information on the number of the matched treated units and the number of the control units used for matching. For each hospital output, the matching process select a specific number of matched treated units out of the 26 treated units generate in the pre-processing process and use the non-treated units as partners. The matching process runs by using the nearest neighbour matching with replacement estimator, so non-treated units are used as partner for more than one treated.

The matching procedure provides both propensity score test (as developed by Leuven and Sianesi, 2003) and a graphical for the balance of the variable distributions in the treated and the control group. Both tests are made at matching time as defined in the pre-processing procedure (one year before matching). For each of matching variables, we find the mean for treated and control group, a measure for the standardized percentage difference between the means in both groups *%bias* in the table), and the t-test check if the means in the control group are equal to those in the treated group. We conclude that matching variables' means are balanced. Graphical results are shown and commented in the Appendix A (figure 16 and 17).

Due to the matching process is based on statistical distance function, we present an additional scale-specific statistics test for the matching variables to verify the significance of each variable distributions' differences between treated and control group. The results are provided by the Kolmogorov-Smirnov test[20] (*Combined K-S* in the table), as the matching variables are continuous. Kolmogorov-Smirnov test results show for each matching variable, the distributions difference between treated and non-treated units (D), p-value (*P-value*) and its corrected estimate (*Corrected*). Considering matching variable's p-values estimates, our results show that there are non-significant differences in the variable distributions between treated and control group.

Table 4: Matching selection of the appropriate comparator group for Inpatient admissions

	Non-Treated		Treated					
All	143		26					
Matched sample	11		14					
	Mean			t-test		Combined K-S		
Variable	Treated	Control	%bias	t	p > |t|	D	P-value	Corrected
Pseudo-Roi	1.0674	1.1787	-25.5	-0.67	0.506	0.1429	0.999	0.997
CEO	1.1459	1.1176	10.7	0.28	0.778	0.1429	0.999	0.997
Total expenditure	19.238	19.077	34.2	0.91	0.373	0.2857	0.617	0.505
Outcome development	-.04132	-.01757	-29.2	-0.77	0.447	0.2143	0.905	0.847
Pseudo R2	LR chi	p > chi2	MeanBias	MedBias				
0.059	2.12	0.714	18.0	19.1				

Table 5: Matching selection of the appropriate comparator group for Eelective admissions

	Non-Treated		Treated					
All	143		26					
Matched sample	12		14					
	Mean			t-test		Combined K-S		
Variable	Treated	Control	%bias	t	p > |t|	D	P-value	Corrected
Pseudo-Roi	1.0674	1.2055	-30.6	-0.81	0.426	0.2143	0.905	0.847
CEO	1.1459	1.1061	14.9	0.39	0.696	0.1429	0.999	0.997
Total expenditure	19.238	19.16	19.1	0.51	0.617	0.2857	0.617	0.505
Outcome development	-.05827	-.02483	-29.1	-0.77	0.449	0.2143	0.905	0.847
Pseudo R2	LR chi	p > chi2	MeanBias	MedBias				
0.077	2.97	0.562	23.4	24.1				

[20]The Kolmogorov–Smirnov test (K–S test) is a nonparametric test for equality of continuous distribution functions. It compares two samples quantifying a between the empirical distribution functions of two samples. The null distribution of this statistic is calculated under the null hypothesis the samples are drawn from the same distribution.

Table 6: Matching selection of the appropriate comparator group for Emergency admissions

	Non-Treated		Treated					
All	143		26					
Matched sample	13		14					
	Mean			t-test			Combined K-S	
Variable	Treated	Control	%bias	t	p > |t|	D	P-value	Corrected
Pseudo-Roi	1.0674	1.1852	-27.0	-0.71	0.482	0.2143	0.905	0.847
CEO	1.1459	1.1087	14.3	0.38	0.708	0.2143	0.905	0.847
Total expenditure	19.238	19.156	18.2	0.48	0.635	0.2143	0.905	0.847
Outcome development	-.00711	.01026	-46.9	-1.24	0.226	0.3571	0.334	0.237
Pseudo R2	LR chi	p > chi2	MeanBias	MedBias				
0.063	2.45	0.654	26.6	22.6				

Table 7: Matching selection of the appropriate comparator group for Outpatient first attendances

	Non-Treated		Treated					
All	143		26					
Matched sample	11		14					
	Mean			t-test			Combined K-S	
Variable	Treated	Control	%bias	t	p > |t|	D	P-value	Corrected
Pseudo-Roi	1.0674	1.2061	-31.8	-0.84	0.408	0.2143	0.905	0.847
CEO	1.1459	1.1104	13.7	0.36	0.720	0.1429	0.999	0.997
Total expenditure	19.238	19.14	22.7	0.60	0.554	0.2857	0.617	0.505
Outcome development	.00749	.00148	17.9	0.47	0.641	0.3571	0.334	0.237
Pseudo R2	LR chi	p > chi2	MeanBias	MedBias				
0.043	1.66	0.798	21.5	20.3				

Table 8: Matching selection of the appropriate comparator group for A&E First attendances

	Non-Treated		Treated					
All	143		26					
Matched sample	11		14					
	Mean			t-test			Combined K-S	
Variable	Treated	Control	%bias	t	p > |t|	D	P-value	Corrected
Pseudo-Roi	1.0674	1.1803	-25.8	-0.68	0.500	0.1429	0.999	0.997
CEO	1.1459	1.1419	1.6	0.04	0.967	0.2143	0.905	0.847
Total expenditure	19.238	19.145	19.2	0.51	0.616	0.3571	0.334	0.237
Outcome development	.00109	-.00177	5.8	0.15	0.878	0.2857	0.617	0.505
Pseudo R2	LR chi	p > chi2	MeanBias	MedBias				
0.024	0.94	0.919	13.1	12.5				

Table 9: Matching selection of the appropriate comparator group for Daycases

	Non-Treated		Treated					
All	143		26					
Matched sample	13		14					
	Mean			t-test			Combined K-S	
Variable	Treated	Control	%bias	t	p > |t|	D	P-value	Corrected
Pseudo-Roi	1.0674	1.1686	-23.3	-0.62	0.543	0.1429	0.999	0.997
CEO	1.1459	1.0899	21.7	0.57	0.572	0.1429	0.999	0.997
Total expenditure	19.238	19.089	31.3	0.83	0.415	0.2857	0.617	0.505
Outcome development	-.08339	-.03583	-24.1	-0.64	0.530	0.2143	0.905	0.847
Pseudo R2	LR chi	p > chi2	MeanBias	MedBias				
0.065	2.51	0.643	25.1	23.7				

Table 10: Matching selection of the appropriate comparator group for Elective-emergency ratio

	Non-Treated		Treated					
All	143		26					
Matched sample	11		14					
	Mean			t-test		Combined K-S		
Variable	Treated	Control	%bias	t	p > |t|	D	P-value	Corrected
Pseudo-Roi	1.0674	1.208	-31.1	-0.82	0.417	0.2143	0.905	0.847
CEO	1.1459	1.1018	16.6	0.44	0.664	0.1429	0.999	0.997
Total expenditure	19.238	19.194	12.0	0.32	0.753	0.2857	0.617	0.505
Outcome development	-.05115	-.04772	-3.0	-0.08	0.938	0.2857	0.617	0.505
Pseudo R2	LR chi	p > chi2	MeanBias	MedBias				
0.035	1.37	0.849	15.7	14.3				

Table 11: Matching selection of the appropriate comparator group for Inpantient-outpatient ratio

	Non-Treated		Treated					
All	143		26					
Matched sample	12		14					
	Mean			t-test		Combined K-S		
Variable	Treated	Control	%bias	t	p > |t|	D	P-value	Corrected
Pseudo-Roi	1.0674	1.1764	-23.5	-0.62	0.540	0.1429	0.999	0.997
CEO	1.1459	1.2082	-26.5	-0.70	0.490	0.2857	0.617	0.505
Total expenditure	19.238	19.186	11.4	0.30	0.765	0.2857	0.617	0.505
Outcome development	-.04881	-.03329	-15.2	-0.40	0.691	0.1429	0.999	0.997
Pseudo R2	LR chi	p > chi2	MeanBias	MedBias				
0.067	2.60	0.627	19.1	19.3				

Table 12: Matching selection of the appropriate comparator group for Daycase-elective ratio

	Non-Treated		Treated					
All	143		26					
Matched sample	12		14					
	Mean			t-test		Combined K-S		
Variable	Treated	Control	%bias	t	p > |t|	D	P-value	Corrected
Pseudo-Roi	1.0674	1.1531	-19.6	-0.52	0.609	0.1429	0.999	0.997
CEO	1.1459	1.1113	13.8	0.36	0.718	0.1429	0.999	0.997
Total expenditure	19.238	19.104	29.2	0.77	0.446	0.2857	0.617	0.505
Outcome development	-.02512	-.01197	-15.0	-0.40	0.695	0.1429	0.999	0.997
Pseudo R2	LR chi	p > chi2	MeanBias	MedBias				
0.041	1.61	0.808	19.4	17.3				

Finally, we present flexible conditional DID results, defined as a comparison of individual differences between treated and their controls (as formally expressed in the equation 2). The table 13 shows the average treatment effect for the treated results[21].

In general, results show that merged Trusts decrease the development of their number of activities in the two years subsequent the merger. Specifically, for inpatients, elective and emergency admissions the mean differences are positive in both treated and control groups. Unlike, difference of their mean differences is negative, so we can conclude that merger effect decreases the development of the number of these hospital activities from the merger start until the two years afterward. Moreover, as expected the effect of merger is greater on the development of planned activities than unplanned. Indeed, the emergency admissions' mean difference is 0.0334 compared to the mean difference (0.0078) of elective admissions. This result is confirmed by the sign of elective-emergency ratio mean difference on treated and controls. Also, the merger has a negative impact on outpatient services development in the two years after its starting. Its mean difference on Trusts merged is negative as well as the difference between merged and non-merged Trusts' differences. The decreasing development of outpatient services' number in the two years subsequent the merger is weaker than the number of inpatients development, among Trusts merged. That is confirmed by the positive sign of inpatient-outpatient ratio on treated, even if the difference between treated and controls mean differences is negative. The development of number of accidents and emergency department activities as well as the number of inpatient admissions without overnight stay (daycase) in the two years subsequent the merger, is positive in both treated and controls groups and also, interestingly, their differences in mean are positive[22]. The increasing development of daycases numbers is greater than the lessening development of outpatient services' number. Thus, the sign of daycase-outpatient ratio mean difference is positive on Trusts merged but is negative in non-merged Trusts. Finally, we show the p-value estimates to assess the statistical significance of these differences. The p-value estimates indicate that the differences are not significant.

[21] The regression-based bias correction of Abadie and Imbens (2006, 2011) is applied to the Difference-in-Differences in order to adjust the start and the end values of the outcome development.

[22] Considering admissions as a proxy of discharges, our daycase results are in line with European Commission's analysis (2016), which denounces an increasing of number of all hospital discharges (including inpatient and day cases) in the period 2003-2012 approximately by 14% due to day cases discharges increased by 50% (while inpatient discharges stayed more or less constant).

Table 13: Conditional DID results

Outcome	mean Diff Treated	Control	DID*	AI robust S.E.	z	p > |z|
Inpatient admissions	0.0181	0.0727	-0.0545	0.0796	-0.6851	0.5053
Elective admissions	0.0078	0.0158	-0.0080	0.0663	-0.1203	0.9061
Emergency admissions	0.0334	0.1899	-0.1565	0.0951	-1.6466	0.1236
Patients attending first outpatient appointment	-0.0012	0.0894	-0.0906	0.0972	-0.9324	0.3681
Patients first attending A&E department	0.1525	0.0951	0.0574	0.1072	0.5355	0.6013
Inpatient admissions without overnight stay	0.0081	0.0008	0.0073	0.1216	0.0599	0.9532
Elective-emergency ratio	-0.0256	-0.1072	0.0816	0.0861	0.9475	0.3607
Inpatient-outpatient ratio	0.0193	0.0375	-0.0182	0.0788	-0.2311	0.8209
Daycase-outpatient ratio	0.0003	-0.0151	0.0154	0.0599	0.2571	0.8011

⋆ Consistent bias-corrected estimator as proposed in Abadie & Imbens (2006,2011).

6.1.2 Fixed effect results

Tables 14-15 estimate the mean treatment effect for the treated within the FE model for the time from the earliest treatment start by the end of our sample.

We highlight the effect of hospital mergers, considering the relative distance from the starting of treatment for each year. This effect is capture by the variable *post merged* from the year 2003 to 2008. Also, the *Merged*post merged* from the year 2003 to 2008 shows the policy effect for treated Trusts related to the distance from the merger year.

As shown in the par. 5.3 (Characteristics of our data structure), the first merger year was in 2001. Consequently, since the time policy effect is defined as the distance from merger year, we capture its effect starting from 2002, considering 2002 as the baseline.

In both tables, we show the effect with only policy variables and also including inputs, controls and hospital characteristics only to have a broader understanding of the overall effect of the merger over time. However, we focus our attention on the model specification with regressors.

In general, the policy effect is positive in all years after the merger (*post merged*) except for Patients first attending A&E department and inpatient admissions without an overnight stay. Also, the sign of both Elective-emergency ratio and Daycase-outpatient ratio is negative. The first result, in line with the previous results found, confirms that the increasing of unplanned activities are more remarkable than planned ones, while the second confirms that hospital mergers are cost-cutting policies that reduce the most expensive activities, such as day case, in favor of outpatient activities. Further, the effect of the merger over time is more significant in inpatient and emergency admissions, in inpatient-outpatient and daycase-outpatient ratio.

The merger policy effect tends to be negative on merged Trusts in each year after the merger (*Merged*post merged*). Thus, we can see that the merger decreases the number of planned and unplanned activities, except for outpatient services and A&E first attendances. Their number increases over time on merged trusts. These findings are confirmed by inpatient-outpatient ratio trend. Indeed, in line with the literature (Goddard and Ferguson, 1997), these findings show that hospital merger in order to achieve efficiency gains, reduce the most expensive activities such as inpatient admission in favour of outpatient services predominantly in the first years after the merger. Moreover, the positive sign of elective-emergency ratio over time in Trusts merged, confirms that the decreasing number of planned activities (inpatient admissions) is wider than the unplanned ones (emergency activities). The policy effect on merger Trusts over time is significant in inpatient and emergency admission.

We assessed a simple sensitivity analysis to test the robustness and better understanding, and also to reduce the uncertainty of our results, comparing several inputs, controls and hospital characteristics. First, we compared the model specification comparing two different inputs: the total available beds and the total available acute beds. The effect was similar on both *post merged* and *Merged*post merged* variables in terms of sign and significance. Any changes in outputs are not found comparing the total number of tests and their subcomponents, CT and MRI, respectively. Finally, we asserted our results comparing the performance rating – type 2 with the performance rating – type 1

Table 14: Estimantes on hospital outputs - part 1

	Inpatient admissions		Elective inpatients		Emergency admissions		Patients attending first outpatient appointment		Patients first attending A&E department		Inpatient admissions without overnight stay	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)
Policy variables												
post merged_year 2003 (β_{21})	0.044	0.041	0.023	0.020	0.075*	0.072*	0.017	0.015	0.031	0.027	-0.030	-0.005
	(0.031)	(0.030)	(0.032)	(0.030)	(0.040)	(0.039)	(0.032)	(0.032)	(0.097)	(0.102)	(0.037)	(0.044)
post merged_year 2004 (β_{22})	0.093***	0.087***	0.048	0.033	0.159***	0.167***	-0.013	-0.014	-0.002	-0.009	-0.025	-0.001
	(0.028)	(0.030)	(0.034)	(0.036)	(0.052)	(0.048)	(0.035)	(0.038)	(0.118)	(0.123)	(0.043)	(0.056)
post merged_year 2005 (β_{23})	0.093***	0.082**	0.039	0.021	0.166***	0.167***	-0.104**	-0.089**	-0.020	-0.024	-0.053	-0.040
	(0.033)	(0.031)	(0.050)	(0.045)	(0.057)	(0.049)	(0.048)	(0.042)	(0.144)	(0.147)	(0.076)	(0.080)
post merged_year 2006 (β_{24})	0.028	-0.008	-0.027	-0.063	0.108	0.072	-0.118	-0.111	-0.020	-0.027	-0.145	-0.157
	(0.072)	(0.073)	(0.087)	(0.081)	(0.080)	(0.082)	(0.103)	(0.118)	(0.170)	(0.174)	(0.101)	(0.099)
post merged_year 2007 (β_{25})	0.194**	0.138**	0.132	0.067	0.283***	0.243***	0.031	0.056	-0.025	-0.036	0.067	0.023
	(0.073)	(0.063)	(0.090)	(0.078)	(0.086)	(0.066)	(0.118)	(0.133)	(0.208)	(0.213)	(0.115)	(0.094)
post merged_year 2008 (β_{26})	0.202***	0.112*	0.167	0.051	0.247***	0.198***	0.021	0.020	-0.041	-0.045	0.151	0.070
	(0.070)	(0.058)	(0.103)	(0.087)	(0.081)	(0.060)	(0.108)	(0.123)	(0.240)	(0.242)	(0.150)	(0.122)
Merged*post merged_year 2003 (δ_1)	-0.023	-0.025	-0.009	-0.013	-0.048	-0.050	-0.013	0.006	-0.001	0.004	0.013	-0.027
	(0.036)	(0.036)	(0.042)	(0.039)	(0.034)	(0.042)	(0.038)	(0.034)	(0.030)	(0.039)	(0.057)	(0.053)
Merged*post merged_year 2004 (δ_2)	-0.077**	-0.077**	-0.042	-0.032	-0.133***	-0.153***	-0.032	0.003	-0.015	-0.003	-0.052	-0.079
	(0.032)	(0.030)	(0.044)	(0.044)	(0.033)	(0.037)	(0.039)	(0.040)	(0.029)	(0.039)	(0.073)	(0.063)
Merged*post merged_year 2005 (δ_3)	-0.082**	-0.076**	-0.049	-0.033	-0.124***	-0.140***	-0.015	0.008	-0.014	0.000	-0.038	-0.058
	(0.033)	(0.031)	(0.052)	(0.047)	(0.032)	(0.037)	(0.049)	(0.043)	(0.035)	(0.046)	(0.089)	(0.076)
Merged*post merged_year 2006 (δ_4)	-0.003	0.006	0.030	0.038	-0.052	-0.045	-0.060	-0.016	-0.007	0.005	0.110	0.073
	(0.066)	(0.065)	(0.083)	(0.075)	(0.047)	(0.058)	(0.118)	(0.115)	(0.037)	(0.047)	(0.106)	(0.091)
Merged*post merged_year 2007 (δ_5)	-0.077	-0.064	-0.034	-0.004	-0.135**	-0.155***	-0.116	-0.083	-0.015	0.002	0.011	-0.002
	(0.046)	(0.042)	(0.063)	(0.052)	(0.050)	(0.046)	(0.077)	(0.075)	(0.046)	(0.053)	(0.099)	(0.083)
Merged*post merged_year 2008 (δ_6)	-0.049	-0.020	-0.036	0.017	-0.056	-0.074*	-0.097	-0.049	-0.032	-0.019	-0.028	-0.011
	(0.056)	(0.048)	(0.081)	(0.067)	(0.053)	(0.042)	(0.082)	(0.077)	(0.059)	(0.062)	(0.123)	(0.105)
FT		0.044		0.074*		-0.006		0.081		0.002		0.086
		(0.028)		(0.036)		(0.028)		(0.052)		(0.026)		(0.059)
Inputs												
Operating theatres (in ln)		0.026		0.049		-0.025		0.004		-0.013		0.072
		(0.053)		(0.081)		(0.082)		(0.101)		(0.057)		(0.107))
Total available beds (in ln)		-0.024		-0.047		0.025		0.141		-0.001		-0.138
		(0.038)		(0.057)		(0.046)		(0.097)		(0.050)		(0.082)
Share of medical staff (% in ln)		0.023		0.021		0.033		-0.025		-0.094		0.031
		(0.088)		(0.103)		(0.084)		(0.133)		(0.087)		(0.136)
Controls												
Average length of stay (in ln)		-0.074		-0.083		-0.061		0.139		-0.024		-0.014
		(0.091)		(0.123)		(0.057)		(0.120)		(0.052)		(0.125)
Median waiting time in days (in ln)		-0.080**		-0.067		-0.102**		-0.038		-0.004		-0.174**
		(0.038)		(0.049)		(0.047)		(0.076)		(0.030)		(0.072)
Patients aged 0-14 (% in ln)		-0.232		-0.428*		0.085		-0.134		-0.111		-0.355
		(0.161)		(0.218)		(0.153)		(0.259)		(0.076)		(0.291)
Patients aged 60 and over (% in ln)		0.196		0.438		-0.235		-0.004		-0.371		1.959*
		(0.557)		(0.756)		(0.349)		(0.577)		(0.372)		(0.957
Female patients (% in ln)		0.356		0.939		-0.672*		0.112		0.009		1.879**
		(0.418)		(0.590)		(0.382)		(0.567)		(0.307)		(0.836)
Total tests (in ln)		0.049		0.030		0.062		0.005		-0.054		-0.035
		(0.069)		(0.090)		(0.069)		(0.109)		(0.081)		(0.148)
hh_bd_30		0.010		0.016**		0.002		0.010		-0.014		0.034***
		(0.006)		(0.007)		(0.008)		(0.015)		(0.015)		(0.011)
Hospital fixed effects												
teaching		0.028		-0.002		0.076**		-0.103		0.028		0.00
		(0.040)		(0.050)		(0.033)		(0.074)		(0.053)		(0.069)
Performance rating (star + use_resources)		0.007		0.006		0.009		0.008		-0.003		0.023
		(0.007)		(0.010)		(0.009)		(0.011)		(0.015)		(0.014
Constant (β_0)	11.439***	9.689**	11.012***	6.846	10.369***	13.177***	11.530***	10.343***	11.522***	14.158***	10.342***	-1.840
	(0.018)	(4.399)	(0.026)	(6.325)	(0.011)	(2.666)	(0.018)	(3.588)	(0.013)	(3.037)	(0.041)	(7.667)
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	234	225	234	225	234	225	233	224	234	225	234	225
R-squared	0.587	0.659	0.425	0.563	0.670	0.701	0.564	0.620	0.527	0.541	0.458	0.634
Number of id2	26	26	26	26	26	26	26	26	26	26	26	26

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

Table 15: Estimantes on hospital outputs - part 2

	Elective-emergency ratio		Inpatient-outpatient ratio		Daycase-outpatient ratio	
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)
Policy variables						
post merged_year 2003 (β_{21})	-0.052	-0.052	0.026	0.025	-0.053***	-0.025
	(0.036)	(0.033)	(0.042)	(0.044)	(0.019)	(0.025)
post merged_year 2004 (β_{22})	-0.112*	-0.134**	0.107**	0.101**	-0.073***	-0.034
	(0.062)	(0.056)	(0.041)	(0.042)	(0.024)	(0.033)
post merged_year 2005 (β_{23})	-0.128	-0.145*	0.198***	0.172***	-0.092**	-0.061
	(0.084)	(0.072)	(0.060)	(0.052)	(0.038)	(0.044)
post merged_year 2006 (β_{24})	-0.135	-0.135*	0.144*	0.100	-0.118**	-0.094**
	(0.089)	(0.071)	(0.082)	(0.082)	(0.053)	(0.043)
post merged_year 2007 (β_{25})	-0.151	-0.176**	0.163*	0.081	-0.065	-0.044
	(0.102)	(0.080)	(0.086)	(0.097)	(0.065)	(0.055)
post merged_year 2008 (β_{26})	-0.080	-0.147	0.182*	0.090	-0.016	0.019
	(0.136)	(0.111)	(0.094)	(0.108)	(0.082)	(0.074)
Merged*post merged_year 2003 (δ_1)	0.039	0.037	-0.010	-0.030	0.022	-0.014
	(0.037)	(0.037)	(0.055)	(0.053)	(0.032)	(0.035)
Merged*post merged_year 2004 (δ_2)	0.092*	0.121**	-0.045	-0.079	-0.010	-0.047
	(0.051)	(0.058)	(0.049)	(0.049)	(0.043)	(0.039)
Merged*post merged_year 2005 (δ_3)	0.075	0.107*	-0.076	-0.091*	0.012	-0.025
	(0.063)	(0.061)	(0.060)	(0.052)	(0.051)	(0.044)
Merged*post merged_year 2006 (δ_4)	0.081	0.083*	0.057	0.024	0.080	0.035
	(0.058)	(0.047)	(0.093)	(0.089)	(0.055)	(0.041)
Merged*post merged_year 2007 (δ_5)	0.101	0.151**	0.039	0.020	0.045	0.002
	(0.077)	(0.057)	(0.076)	(0.071)	(0.054)	(0.043)
Merged*post merged_year 2008 (δ_6)	0.019	0.091	0.048	0.030	0.009	-0.027
	(0.096)	(0.076)	(0.076)	(0.073)	(0.061)	(0.049)
FT		0.080*		-0.037		0.011
		(0.040)		(0.050)		(0.033)
Inputs						
Operating theatres (in ln)		0.073		0.023		0.023
		(0.130)		(0.099)		(0.066)
Total available beds (in ln)		-0.072		-0.162*		-0.091
		(0.078)		(0.093)		(0.074)
Share of medical staff (% in ln)		-0.012		0.052		0.010
		(0.076)		(0.144)		(0.083)
Controls						
Average length of stay (in ln)		-0.022		-0.213		0.070
		(0.113)		(0.151)		(0.054)
Median waiting time in days (in ln)		0.035		-0.043		-0.108**
		(0.060)		(0.082)		(0.044)
Patients aged 0-14 (% in ln)		-0.513**		-0.098		0.073
		(0.238)		(0.280)		(0.122)
Patients aged 60 and over (% in ln)		0.674		0.203		1.520***
		(0.692)		(0.560)		(0.311)
Female patients (% in ln)		1.611**		0.239		0.940**
		(0.682)		(0.579)		(0.429)
Total tests (in ln)		-0.031		0.051		-0.066
		(0.098)		(0.106)		(0.083)
hh_bd_30		0.014		-0.001		0.018***
		(0.009)		(0.015)		(0.006)
Hospital fixed effects						
teaching		-0.078*		0.133*		0.009
		(0.043)		(0.076)		(0.034)
Performance rating (star + use_resources)		-0.003		-0.001		0.018**
		(0.013)		(0.013)		(0.008)
Constant (β_0)	5.248***	-1.726	4.514***	3.844	3.936***	-4.081*
	(0.027)	(6.506)	(0.024)	(2.843)	(0.019)	(2.272)
Year fixed effect	YES	YES	YES	YES	YES	YES
Hospital fixed effect	YES	YES	YES	YES	YES	YES
Observations	234	225	233	224	234	225
R-squared	0.143	0.385	0.231	0.347	0.335	0.558
Number of id2	26	26	26	26	26	26

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

6.1.3 Conditional parallel trend assumption

Table 16 shows results on conditional parallel trend assumption by using both leads and time trend variable. This assumption is ensured among hospital outputs. The binary treatment $merged_t$ is defined as the average development of number of hospital outputs in the two years after merger, considering pre-treatment features in the year before merger. For each hospital outputs, the treatment is defined as the tendency of treated Trusts to increase in hospital output in a specific year compared with a baseline reference, measured as the average development of hospital output in the two years after merger.

Thus, Figure 13 show hospital outputs' path in the period t-1 to t+2. For inpatient admissions' ATE at time of treatment (merged at time t) is higher than its average development in the two years after merger. Moreover, the ATE of inpatient admissions in the period from t to t+2, is decreasing but higher than its average development in the two years after merger. Also, elective and outpatient services' trend is decreasing related to their average developments in the two years after merger. Considering emergency admissions, we can notice that its ATE at time t and in the following period (t+1, t+2) is higher than its average development in the two years after merger, reaching a positive peak in t+1. The ATE of number of impatient attending A&E department is higher than its average development in the two years after merger but decreasing below its average development in the two years after merger, at time t+1 and t+2. The ATE of number of inpatient admissions without overnight stay is less than its average development in the two years after merger from the period t to t+2, but with a tendentially increasing trend. Moreover, The ATE of elective-emergency ratio is below its development in the two years after merger in the period from t to t+2, reaching a negative peak in t+1. Inpatient-outpatient ratio's ATE is increasing in the period from t to t+2 and always higher than its development in the two years after merger. Finally, Ate of daycase-outpatient ratio is below its development in the two years after merger, its trend is increasing in period from t to t+2, reaching the equality with its development in the two years after merger in t+2.

Table 16: Conditional parallel trend results

Variables	Inpatient admissions (Coeff.\) (Std. Err.)	Elective inpatients (Coeff.\) (Std. Err.)	Emergency admissions (Coeff.\.) (Std. Err.)	Patients attending first outpatient appointment (Coeff.\) (Std. Err.)	Patients first attending A&E department (Coeff.\) (Std. Err.)	Inpatient admissions without overnight stay (Coeff.\) (Std. Err.)	Elective-emergency ratio (Coeff.\) (Std. Err.)	Inpatient-outpatient ratio (Coeff.\) (Std. Err.)	Daycase-outpatient ratio (Coeff.\) (Std. Err.)
$merged_{t-1}$	-0.029	-0.042	-0.018	-0.035	0.016	-0.037	-0.024	0.004	0.005
	(0.028)	(0.036)	(0.032)	(0.091)	(0.049)	(0.049)	(0.043)	(0.070)	(0.027)
$merged_{t}$	0.012	0.006	0.012	-0.023	0.051	-0.021	-0.007	0.036	-0.027
	(0.016)	(0.023)	(0.024)	(0.043)	(0.094)	(0.031)	(0.032)	(0.038)	(0.016)
$merged_{t+1}$	0.008	-0.015	0.037	-0.054	-0.023	-0.034	-0.052	0.063	-0.019
	(0.046)	(0.040)	(0.059)	(0.032)	(0.052)	(0.051)	(0.034)	(0.044)	(0.022)
$merged_{t+2}$	-0.004	-0.017	0.016	-0.098	-0.009	-0.015	-0.033	0.090	0.002
	(0.029)	(0.028)	(0.039)	(0.047)	(0.041)	(0.041)	(0.027)	(0.044)	(0.020)
Parallel trend using the "leads"									
F(1, 25) =	1.12	1.34	0.31	0.15	0.11	0.55	0.31	0.00	0.04
Prob > F	0.3007	0.2573	0.5844	0.7034	0.7402	0.4634	0.5835	0.9553	0.8486
Parallel-trend	passed	passed	passed	passed	passed	passed	passed	passed	passed
Parallel trend using the "time-trend"									
F(1, 25) =	2.17	1.17	1.93	0.06	1.57	0.85	0.21	3.08	0.21
Prob > F =	0.1531	0.2904	0.1766	0.8074	0.2217	0.3647	0.6495	0.0913	0.6504
Parallel-trend	passed	passed	passed	passed	passed	passed	passed	passed	passed
Observations	140	140	140	139	140	140	140	139	140
Id2	26	26	26	26	26	26	26	26	26

Figure 13: Conditional paralle trend of hospital outputs

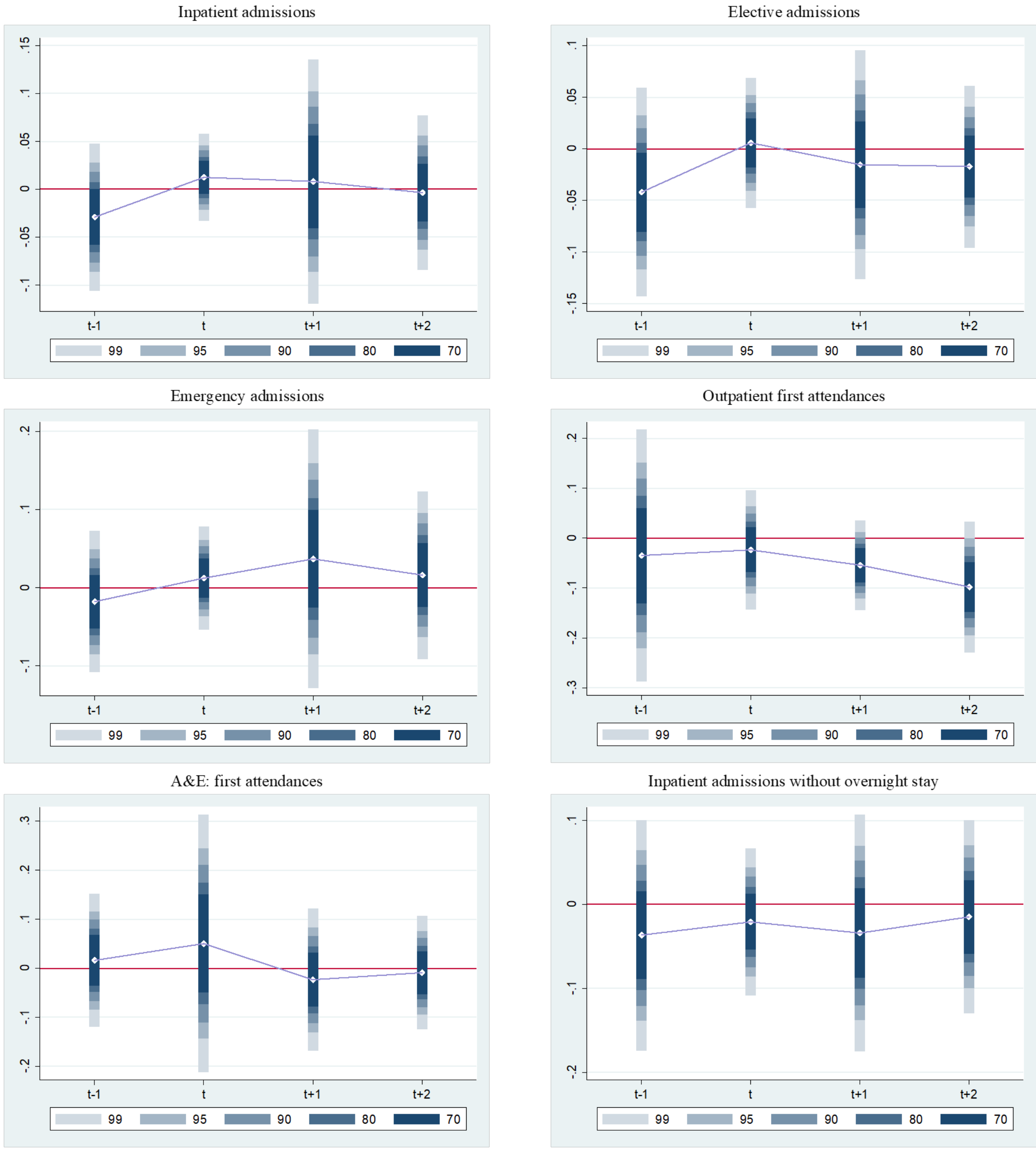
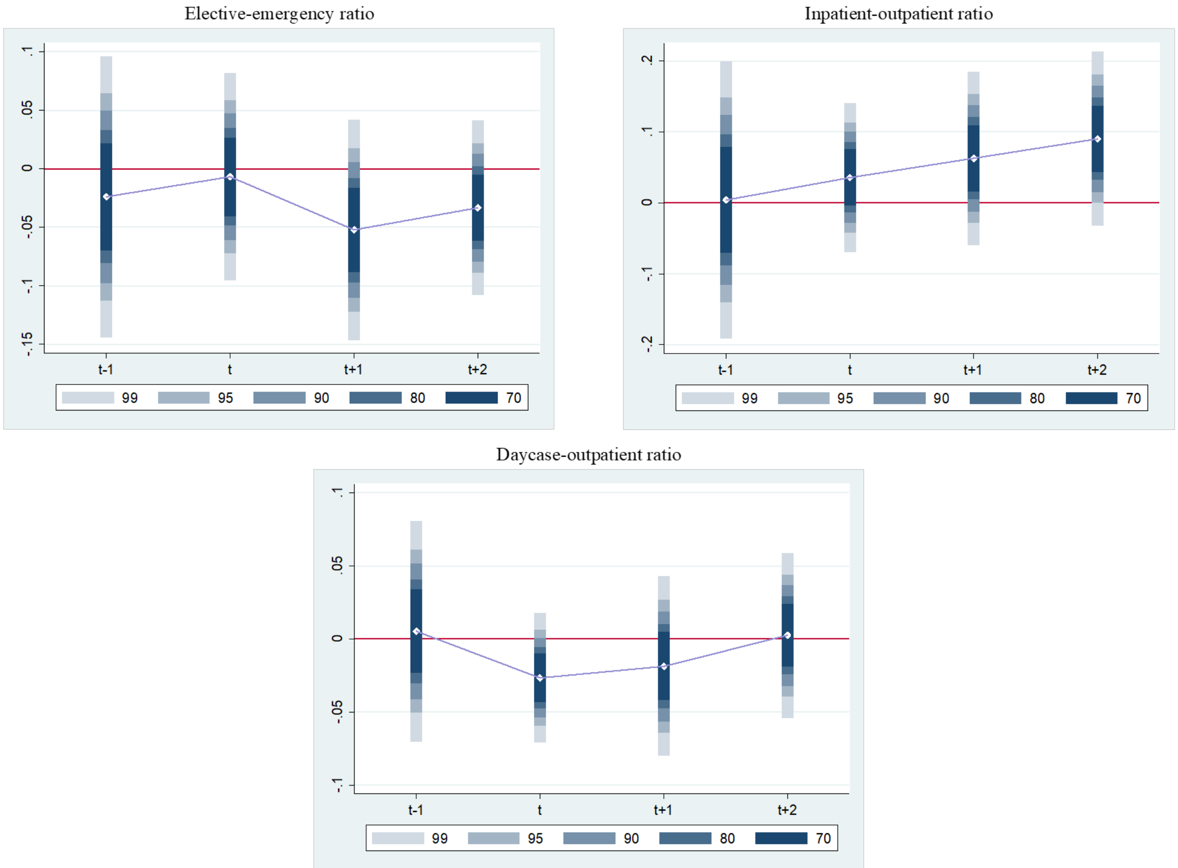


Figure 14: Conditional paralle trend of hospitals outputs (ratio)



6.2 Results on hospital outcomes

6.2.1 Flexible conditional DID's results

In tables 17-19 are shown matching and DID results on Performance rating - Type 1 and Performance rating - Type 2. We adopt the same assumptions in term selection of of pre-treatment characteristics, relative time of matching, pre-treatment outcome development and its development post-merger (see par. 6.1.1 for further details).

As shown in tables 17-18, the matched sample is composed of 13 treated units out of the 26 treated units generate in the pre-processing process and use 12 non-treated units as partners for both hospital outcomes. As in the previous analysis, we adopt the nearest neighbour matching with replacement estimator, so non-treated units are used as partner for more than one treated.

For each of matching variables, we find the mean for treated and control group, a measure for the standardized percentage difference between the means in both groups (%bias in the table). Moreover, the t-test checks if the means in the control group are equal to those in the treated group. We conclude that matching variables' means are balanced. Graphical results are shown and commented in the Appendix A (figure 18) . Considering matching variable's p-values estimates, Kolmogorov-Smirnov test results show that there are non-significant differences in the variable distributions between treated and control group.

Finally, we present flexible conditional DID results, defined as a comparison of individual differences between treated and their controls (as formally expressed in the equation 2). Table 19 shows the average treatment effect for the treated results .

In general, results show that merged Trusts decrease the development of hospitals' performance in the two years subsequent the merger. Specifically, the mean differences are negative in both treated and control groups for each dependent variable, but the decreasing development on hospitals' performance is greater in on Performance rating - Type 1 rather than Performance rating - Type 2 in the two years subsequent the merger. In other words, the negative effect of merger is stronger in the quality of services than in the use of resources in the short-term. The p-value estimates show that the differences are not significant.

Table 17: Matching selection of the appropriate comparator group for Performance rating - Type 1

	Non-Treated		Treated					
All	143		26					
Matched sample	12		13					
	Mean			t-test		Combined K-S		
Variable	Treated	Control	%bias	t	p > |t|	D	P-value	Corrected
Pseudo-Roi	1.0344	1.1598	-29.3	-0.75	0.462	0.1538	0.998	0.995
CEO	1.1464	1.1404	2.2	0.06	0.956	0.2308	0.879	0.811
Total expenditure	19.247	19.084	31.6	0.81	0.428	0.3077	0.570	0.455
Outcome development	-.30769	-.38462	8.9	0.23	0.822	0.0769	1.000	1.000
Pseudo R2	LR chi	p > chi2	MeanBias	MedBias				
0.059	2.12	0.714	18.0	19.1				

Table 18: Matching selection of the appropriate comparator group for Performance rating - Type 2

	Non-Treated			Treated					
All	143			26					
Matched sample	12			13					
	Mean				t-test		Combined K-S		
Variable	Treated	Control	%bias		t	p > |t|	D	P-value	Corrected
Pseudo-Roi	1.0344	1.1598	-29.3		-0.75	0.462	0.1538	0.998	0.995
CEO	1.1464	1.1404	2.2		0.06	0.956	0.2308	0.879	0.811
Total expenditure	19.247	19.084	31.6		0.81	0.428	0.3077	0.570	0.455
Outcome development	-.38462	-.38462	0.0		0.00	1.000	0.0769	1.000	1.000
Pseudo R2	LR chi	p > chi2	MeanBias		MedBias				
0.049	1.76	0.780	15.8		15.8				

Table 19: Conditional DID results

Outcome	mean Diff		DID*	AI robust	z	p > |z|
	Treated	Control		S.E.		
Performance rating - Type-1	-0.3846	-0.0981	-0.2865	0.4639	-0.6176	0.5484
Performance rating - Type-2	-0.3846	-0.2036	-0.1810	0.4882	-0.3708	0.7172

6.2.2 Fixed effect results

Table 20 estimates the mean treatment effect for the treated within fixed effect ordered logit model (FE-OL) with blow up and cluster (BUC) estimator. This model allows only for policy effect of merged Trusts considering the overall effect post merged (*Merged*post merged*).

As we know the effect of policy over time, is defined from 2003 to 2008, considering 2002 as the baseline. Thus, we have only the overall effect of merger in the period 2003-2008. It may be a weakness of our analysis, but it can a good starting point for further investigations on testing categorial variable with time-varying treatment effect.

To analyse the effect of hospital merged on Performance rating – Type 1, we include elements that better portrays the quality of hospital services and hospital characteristics (number of acute beds, proportion of non-medical staff, teaching, ALOS and the interaction between the HHI and teaching, called Interaction 1). In line with the literature, we investigate the performance based on the use of resources (Performance rating – Type 2) by using financial variables (RCI and directors' costs), the proportion of non-medical staff to take into account hospital inputs and some hospital characteristics (teaching, ALOS, Interaction 1 and the interaction between HHI and FT status, called Interaction 2. In addition, in both models we include other policy interventions (FT).

In general, the effect of policy variable is negative on both outcomes even if it is significant at least 5% in the performance rating – type 2. Interestingly, FT status has a positive and significant effect on performance rating – type 2, due to the greater FTs' financial flexibility and freedom that allow Trusts, among other advantages, to retain financial surpluses and invest in buildings and new services. Also, a positive relationship is between surplus and use of resources. As expected, the increasing on inputs have a positive effect on both hospital activities. An enhancing of ALOS, decreases the quality of services but improve the use of resources.
Additionally, we assessed also for hospitals outcome a simple sensitivity analysis to test the robustness, better understanding, and also to reduce the uncertainty of our results, comparing our results by using total available beds or the total available acute beds on both outcomes. The results are similar in the two different model specification. Moreover, we confirmed our results the performance rating – type 2 comparing surplus and retained surplus.

Table 20: Estimantes on hospitals outcomes

	Performance rating Type-1 (star + quality of the services)		Performance rating - Type 2 (star + use of resources)	
	(1)	(2)	(3)	(4)
Variables	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)	(Coeff.\Std. Err.)
Policy variables				
$Merged * post_merged(\delta)$	-0.142	-0.052	-0.195***	-0.297**
	(0.107)	(0.137)	(0.057)	(0.145)
FT		-0.735		2.232**
		(0.544)		(1.103)
Inputs				
Total available acute beds (in ln)		1.343		
		(1.231)		
Share of non-medical staff (% in ln)		3.733		40.787**
		(12.814)		(18.743)
Controls				
Average length of stay (in ln)		-0.807		0.191
		(1.037)		(1.274)
Surplus (in ln)				1.148***
				(0.336)
RCI (including excess beds) (in ln)				-4.681
				(4.079)
Directors' costs (in ln)				-0.797
				(0.631)
Hospital characteristics				
teaching		0.397		-12.079***
		(2.033)		(2.980)
Interaction 1 HHI on bd (radius 30 miles) * teaching		-0.397		-0.533
		(0.438)		(1.153)
Interaction 2 HHI on bd (radius 30 miles) * FT				-0.030
				(0.192)
Observations	421	416	474	312

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

6.2.3 Conditional parallel trend assumption

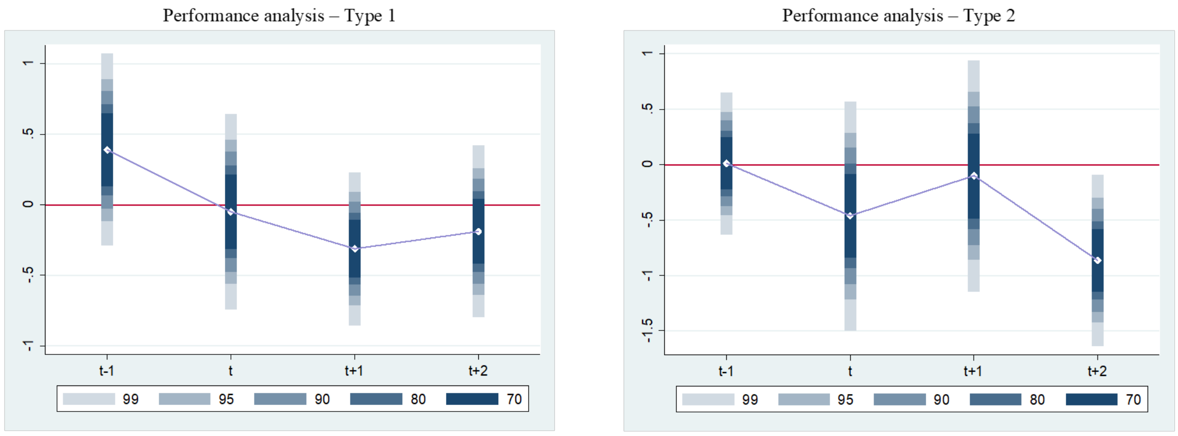
Table 21 shows results on conditional parallel trend assumption by using both leads and time trend variable. This assumption is ensured among hospital outcomes.
The binary treatment $merged_t$ is defined as the average development of number of hospital outcomes in the two years after merger, considering pre-treatment features in the year before merger. For each hospital outcomes, the treatment is defined as the tendency of treated Trusts to increase in hospital outcome in a specific year compared with a baseline reference, measured as the average development of hospital outcome in the two years after merger.

Thus, figure 15 show hospital outcomes path in the period t-1 to t+2. The ATE of performance rating – type 1 is just below its average development in the two years after merger, decreases in the period from t to t+2 reaching a negative peak in t+1. Whilst the ATE of performance rating – type 2 is always below its average development in the two years after merger in the period from t to t+2, having a bump on the path in t+1.

Table 21: Conditional parallel trend results

Variables	Performance rating - Type 1 (Coeff.\Std. Err.)	Performance rating - Type 2 (Coeff.\Std. Err.)
mergedt-1	0,390	0,011
	(0,244)	(0,220)
mergedt	-0,050	-0,462
	(0,248)	(0,354)
mergedt+1	-0,312	-0,101
	(0,194)	(0,357)
mergedt+2	-0,189	-0,863
	(0,217)	(0,265)
Parallel trend using the 'leads'		
F(1, 25) =	2.56	0.00
Prob > F =	0.1227	0.9595
Parallel-trend	passed	passed
Parallel trend using the 'time-trend'		
F(1, 25) =	0.94	0.10
Prob > F =	0.3424	0.7584
Parallel-trend	passed	passed

Figure 15: Conditional paralle trend of hospital outcomes



7 Conclusions

In this paper, we provide evidence on the effect of merger policy on several measures of hospital outputs and outcomes during the year 2000-2008. As the starting of merger differs between different hospitals, we consider different groups of treated Trusts. Also, as the merger effect is heterogeneous between groups, we adopt an alternative difference-in differences approach. Following the Dettmann et al (2020) approach, we adopt the "flexible conditional DID" to taking into account both time-varying treatment effect and its heterogeneity over time. In line with literature (Goddard and Ferguson, 1997; Propper et al. 2004; Fulop et al., 2005; Gaynor et al, 2012; Collins, 2015) we consider hospital mergers as complex processes mainly driven by economic and financial reasons. Thus, we reorganized our data structure consider financial hospital features one year before merger, identifying 26 groups of one single treated unit. They represent our "benchmark groups" imposed in the matching process. The definition of pre-treatment merger is an average time built considering different proposal phases, as described by Collins (2015, pp. 23).

The matching procedure selects for each treated unites the appropriate controls at the same time with the same features (as identified by the "benchmark groups"). As we need to capture the different development of outcome for our treated and control groups, in this step we define the time of pre-treatment and post-merged outcome. Finally, we estimate the average treatment effect of our dependent variables for treated Trusts (ATT). It is the difference of development's average differences between treated and controls for each hospital output and outcome. In gen-

eral, findings show that mean differences of hospital outputs development among treated Trust, is overall positive, but the differences between treated and controls show that the general effect of merger process decreases the development number of hospital admissions (inpatient, elective and emergency) and also outpatient services. Unlike development of both number of patients attending A&E department and patient admissions without overnight stay increase in the two years after merger. These findings may reveal two concerns: patients make use of higher A&E department services due to the lessening planned and unplanned hospital activities; and also merged hospital reduce expensive activities, such as admissions with overnight stay, in order to achieve more efficiency.
Findings on performance highlights a decreasing on average development on both quality of services and use of resources in two years after merger. Furthermore, we test a fixed model, including a set of inputs, controls, and hospital characteristics in order to estimate the effect of merger policy on hospital outputs and outcomes. In general, considering time span between 2000-2008, merged Trusts decrease the level of hospital activities and performance mainly in the three years after merger (in line with the previous findings). As, performance measures are categorical variables, we estimate the effect of merger on them by using the BUC estimator. The drawback of analysis is that it can capture only the overall effect without showing the effect of merger policy in each year. We implement this issue in future analysis.
As required from adopted approach, we test conditional parallel trends (Callaway and Sant'Anna, 2019) by using both time leads and time trends variables by using Cerulli and Ventura (2019) procedure. These results confirmed that heterogenous treated and control groups follow parallel trends in the absence of the treatment.

In conclusion, we contribute to the literature on hospital mergers by using an alternative methodological approach that allows to overtake the potential limits associated to important missing information (such as date of decision and announcement) fundamental to assess the effect of mergers. As the time-varying treatment effect is crucial for policy evaluation, we illustrate another methodological approach to address the effect of policy introduction in different moment, even in the lack of fundamental information as required by event studies. Moreover, by using the flexible conditional DID approach permit to evaluate the average treatment effect among different period without losing the clearness of results interpretation. Even thought our analysis provide short-run findings, they could be a good starting point for enrich our study evaluating the effect of merger policy in the long-term. We reserve in future studies the improvement of performance assessment.

Appendix A Quantile-Qualtile plot

Figure 14-15 show the quantile-quantile plots of matching variables for hospital outputs and performance. They compare the distributions in both groups by means of the plotted quantiles. The line represents identical distributions. Treated group are represented in the x-axis, while the control groups are in the y-axis. In general, we can see a small deviation from the line for all displayed variables, mostly at the tails of the distributions.

Figure 16: QQ-plot - matching variables at matching time for hospital outputs

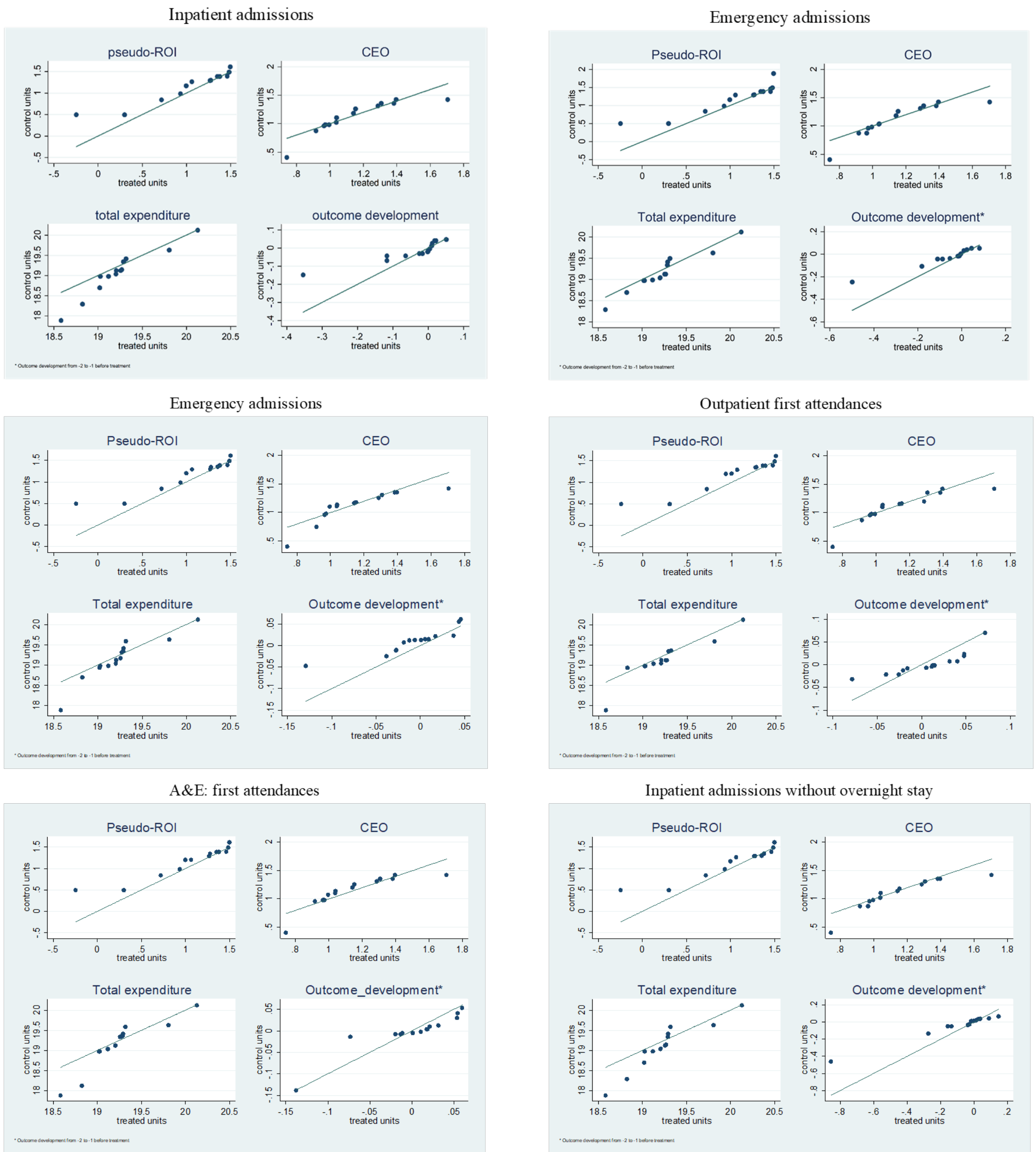
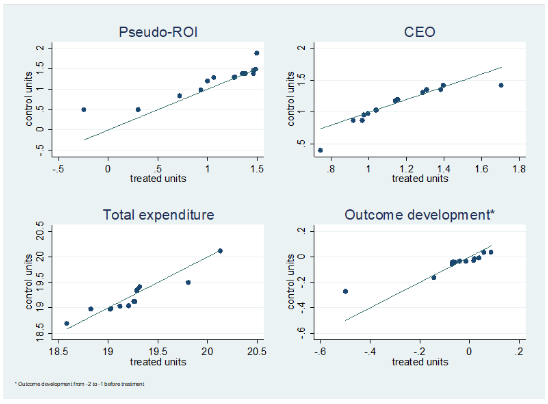
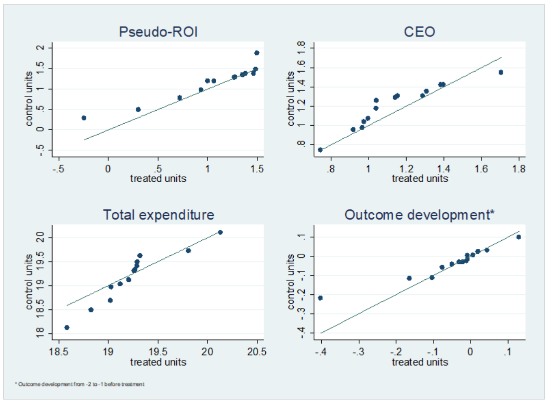


Figure 17: QQ-plot - matching variables at matching time for hospital outputs (ratio)

Elective-emergency ratio



Inpatient-outpatient ratio



Daycase-outpatient ratio

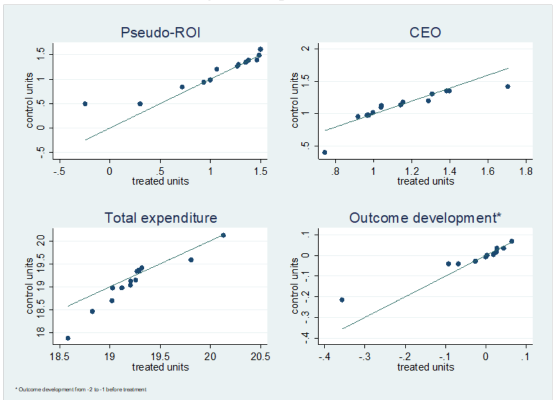
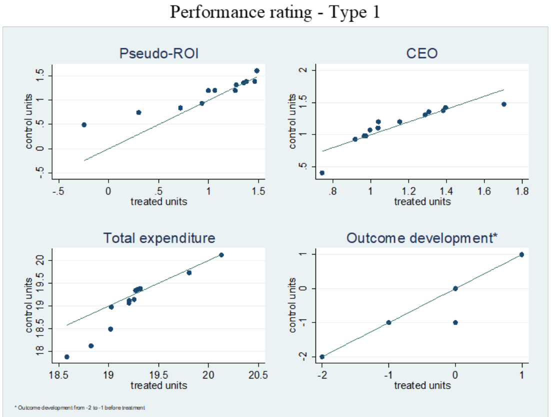
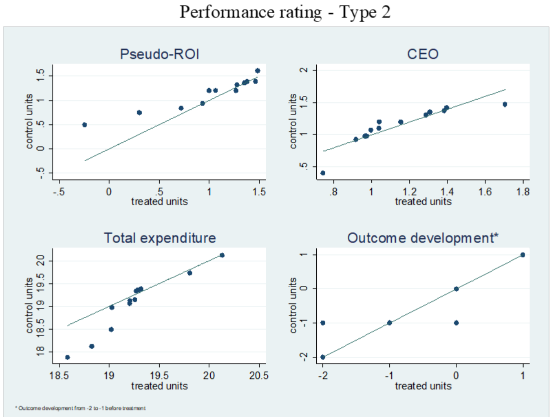


Figure 18: QQ-plot - matching variables at matching time for hospital outcomes

Performance rating - Type 1



Performance rating - Type 2



References

Abadie, A. and G. Imbens (2011). "Bias-Corrected Matching Estimators for Average Treatment Effects". In: *Journal of Business and Economic Statistics* 29.1, 1–11.

— (2006). "Large Sample Properties of Matching Estimators for Average Treatment Effects". In: *Econometrica* 74.1, pp. 235–267.

Abraham S. Sun, L. (Apr. 2018). "Estimating Dynamic Treatment Effects in Event Studies With Heterogeneous Treatment Effects". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3158747.

Athey S., Imbens G. (2018). *Design-based Analysis in Difference-In-Differences Settings with Staggered Adoption*. arXiv: 1808.05293 [econ.EM].

Audit, Commission and HFMA (Revised 2010 edition). "NHS Trus Account: A guide for non-executives". In.

Baetschmann G, Staub K. E. and Winkelmann R. (2015). "Consistent estimation of the fixed effects ordered logit model". In: *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 178.3, pp. 685–703.

Bloom N. A., Propper C., Seiler S., and Van Reenen J. (2005). "The Impact of Competition on Management Quality: Evidence from Public Hospitals". In: *The Review of Economic Studies* 82.Issue 2, pp. 457–489.

Callaway B, Sant'Anna P. H. C. (2019). "Difference-in-Differences with Multiple Time Periods". In: Available at SSRN: https://ssrn.com/abstract=3148250https://ssrn.com/abstract=3148250 or http://dx.doi.org/10.2139/ssrn.3148250.

Cerulli G., Ventura M. (2019). "Estimation of pre- and posttreatment average treatment effects with binary time-varying treatment using Stata". In: *Stata Journal* 19.3, pp. 551–565.

Collins, B. (2015). "Foundation trust and NHS trust mergers 2010 to 2015". In: *The King's Fund*. Available on line: http://www.kingsfund.org.uk/publications/foundation-trust-and-nhs-trust-mergers.

Cooper Z., Gibbons S., Jones S., and McGuire A. (2011). "Does Hospital Competition Save Lives? Evidence From The English NHS Patient Choice Reforms". In: *The Economic Journal* 121.554, F228–F260.

Dettmann, E., Giebler A., and Weyh A. (2020). *Flexpaneldid: A Stata toolbox for causal analysis with varying treatment time and duration.* IWH Discussion Papers 3/2020, Leibniz-Institut für Wirtschaftsforschung Halle (IWH), Halle (Saale.

Dorey, P. (2015). "The Legacy of Thatcherism - Public Sector Reform". In: *Observatoire de la société britannique* 17.

Dranove, D. (1998). "Economies of scale in non-revenue producing cost centers: implications for hospital mergers". In: *Journal of Health Economics* 17.1, 69–83.

Duggan M., G. (2000). "Hospital Ownership and Public Medical Spending". In: *The Quarterly Journal of Economics* 115.4, pp. 1343–1373.

Fulop N., Protopsaltis G., Allen P. Hutchings A. King A., and Normand C. (2005). "Changing organisations: a study of the context and processes of mergers of health care providers in England". In: *Social Science & Medicine* 60.1, pp. 119 –130.

Gaynor, M. (2004). *Competition and quality in hospital markets. What do we know? What don!t we know?* Vol. 12. 2, pp. 3–40.

Gaynor, Martin, Rodrigo Moreno-Serra, and Carol Propper (2013). "Death by Market Power: Reform, Competition, and Patient Outcomes in the National Health Service". In: *American Economic Journal: Economic Policy* 5.4, pp. 134–66. DOI: 10.1257/pol.5.4.134. URL: https://www.aeaweb.org/articles?id=10.1257/pol.5.4.134.

Gaynor M., Laudicella M. and Propper C. (2012). "Can governments do it better? Merger mania and hospital outcomes in the English NHS". In: *Journal of Health Economics* 31.3, pp. 528 –543.

Goddard, M. and Ferguson B. (1997). "Mergers in the NHS: Made in heaven or marriages of convenience?." In: *Nuffield Trust.* Nuffield Occasional Papers Health Economics Series.1.

Heckman, J. et al. (1998). *Characterizing Selection Bias Using Experimental Data.* NBER Working Papers 6699. National Bureau of Economic Research, Inc.

Heckman J Ichimura H, Todd PE. (1997). "Matching as an econometric evaluation estimator: evidence from evaluating a job training programme". In: *The Review of Economic Studies* 64.4, Special Issue: Evaluation of Training and Other Social Programmes, pp. 605–654.

Herr, A (Sept. 2008). "Cost and Technical Efficiency of German Hospitals: Does Ownership Matter?" In: *Health economics* 17, pp. 1057–71. DOI: 10.1002/hec.1388.

Ho, V. and Hamilton B.H. (2000). “Hospital mergers and acquisitions: does market consolidation harm patients?” In: *Journal of health economics* 19.5, 767—791. ISSN: 0167-6296. DOI: 10.1016/s0167-6296(00)00052-7.

Horwitz, J. R. and A. Nichols (2009). “Hospital ownership and medical services: Market mix, spillover effects, and nonprofit objectives”. In: *Journal of Health Economics* 28.5, pp. 924–937.

Imai K., Kim I. S. and E. Wang (2019). *Matching Methods for Causal Inference with Time-Series Cross-Sectional Data*. Working paper. Harvard University.

J., Angrist and Pischke J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist’s Companion*. 1st ed. Princeton University Press.

Kjekshus, L. and T. Hagen (2007). “Do hospital mergers increase hospital efficiency? Evidence from a National Health Service country”. In: *Journal of Health Services Research & Policy* 12.4, pp. 230–235. DOI: 10.1258/135581907782101561.

Krishnan R. A., Joshi S. and Krishnan H. (2004). “The influence of mergers on firms’ product-mix strategies”. In: *Strategic Management Journal* 25.6, pp. 587–611. DOI: 10.1002/smj.410.

Krueger A. B., Card D (2000). “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Reply”. In: *American Economic Review* 90.5, pp. 1397–1420.

Leuven, E. and B. Sianesi (2003). “PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing”. In: *Statistical Software Components, Boston College Department of Economics*. URL: URL: https://ideas.repec.org/c/boc/bocode/s432001.html.

Preyra, C. and Pink G. (2006). “Scale and scope efficiencies through hospital consolidations”. In: *Journal of Health Economics* 25.6, pp. 1049 –1068. ISSN: 0167-6296. DOI: https://doi.org/10.1016/j.jhealeco.2005.12.006.

Propper, C. (2012). “Competition, incentives and the English NHS”. In: *Health Economics* 21.1, pp. 33–40.

Propper C., Burgess S. and Gossage D. (2008). “Competition and Quality: Evidence from the NHS Internal Market 1991-9”. In: *Economic Journal* 118.525, pp. 138–170.

Propper C., Burgess S. and Green K. (2004). “Does competition between hospitals improve the quality of care?: Hospital death rates and the NHS internal market”. In: *Journal of Public Economics* 88.7-8, pp. 1247–1272.

Schmitt, M. (2017a). “Do hospital mergers reduce costs?” In: *Journal of Health Economics* 52, pp. 74 –94. ISSN: 0167-6296. DOI: https://doi.org/10.1016/j.jhealeco.2017.01.007.

— (2017b). “Do hospital mergers reduce costs?” In: *Journal of Health Economics* 52, pp. 74 –94. ISSN: 0167-6296. DOI: https://doi.org/10.1016/j.jhealeco.2017.01.007.

Spang H. Radach, Arnould R. J. and Bazzoli G. J. (2009). “The effect of non-rural hospital mergers and acquisitions: An examination of cost and price outcomes”. In: *The Quarterly Review of Economics and Finance* 49.2, pp. 323–342.

Town R., Wholey D., R. Feldman, and Burns L. (June 2006). “The Welfare Consequences of Hospital Mergers”. In: *NBER Working Paper; No. 12244*. DOI: 10.3386/w12244.

Wagstaff, A. (1989). “Econometric studies in health economics: A survey of the British literature”. In: *Journal of Health Economics* 8.1, pp. 1–51.

Financial crisis, fiscal austerity, and health in Italy*

Vanessa Cirulli
Sapienza University of Rome
vanessa.cirulli@gmail.com

December 10, 2020

Abstract

We analyse the effect of regional bail-out plans adoption on a broad set of health status measures during the period 1999-2015. Also, we focus on the impact of austerity policies on social distress, considering several dimensions of physical and psychological human diseases. We adopt an IV strategy to address the potential endogeneity of bail-out plans, by using the average percentage of people with low satisfaction on public transport in 2005 as a proxy of overall regional system inefficiency before the introduction of bail-out plans. In line with the existing literature, our results show a negative impact of bail-out plans on several dimensions of health.

JEL Codes: I11, I13, I18, L32

Keywords: Austerity measures, Recession, Bail-out plans, Mortality rate, Incidence of infectious diseases, Discharges rate, social distress; Physical and psychological human diseases.

1 Introduction

During 2007 the global banking crisis was unfolded, and consequently, many European countries were affected by a global economic crisis. Several authors argued that the macroeconomic crisis has weakened the fiscal sustainability of European welfare states and, in particular, their healthcare systems. Generally, healthcare costs have carried weight within public expense items in each European country. Indeed, the EU total health expenditure (in terms of % of GDP) has

*The views expressed are those of the authors and do not necessarily reflect those of the institutions to which they belong.

grown from 6.18 to 7.82 during the year 2000-2013. Consequently, the EU has required to their Member States reforms over the health system, has recommended them to introduce coercive and persuasive policies but always monitoring the possible negative effects over poor and more vulnerable population. Thus, most of the governments had to rethink and reform their health care systems in order to cope with such a crisis, and have introduced a set of policy adjustment with the aim to enhance the efficiency of the healthcare system and to better control their health expenditures (Appleby et al., 2015).

Most of them undertook a path of austerity, adopting several measures to deal with it, such as fiscal adjustments and budgeting cuts. The economic crisis was the pretext for several EU countries to introduce not only cuts to health spending but also, to social security spending related to unemployment. In this context, the growth of public health expenditure become one of the most relevant elements of the health policy debates across European economies, and the main concerns were about the effects of austerity measures on the European healthcare system. Several measures of cost cuts were introduced, in particular in the hospital and pharmaceutical sectors, for example, by switching to generic drugs, or shifting some healthcare costs from the State to the patient, and also reducing doctors' salaries or the number of non-medical staff (Arie, 2013). In response to the crisis, others decreased the extent of health coverage by instituting or increasing user charges for some health services. In the latter case, the financial burden was fallen on households increasing the inequality between the high and low-value care, especially for both people with low incomes and high-users of healthcare, even when user charges are low. Moreover, cuts in health spending have had negative effects also on the population health and well-being. For example, outbreaks of infectious illnesses were typical diseases during downturn periods. Besides, the constrained access to appropriate medical care and the removal of a set of welfare services has increased the cases of psycho-social distresses, such as alcohol and drugs abuse, psychological or mental disorders, and consequently the number of suicides (Karanikolos et al., 2013; Kentikelenis et al., 2014).

In Italy, as in the rest of Europe, the increasing of health public spending has always been one of the most important issues of public expenditure management. Over the past thirty years, the Italian Central Governments carried out several reforms of Regional Health Services (RHS) to improve both the efficiency and the quality of public healthcare sector. The increasing fiscal and organisational autonomy of Regions did not allow to contain their deficits. Thus, as a consequence of both European financial crises and the Italian sovereign debt crisis, Regions with a deficit level more than 7% (up to 2007, or more than 5% after 2010) of the total amount of total funding have to adopt an operational program called bail-out plan. We consider the bail-out plan as a measure of austerity because it is based on cuts of healthcare services, by the reduction of hospital rates, number of hospital beds and staff costs.

Several authors have analyzed the effect of austerity measures on the Italian healthcare sector. Recent papers have explored the effect of bail-out plans (BOP) on several dimensions of health. In particular, Atella et al. (2019) investigate how health policies adopted to contarst economic downturns have affected healthcare sector productivity. Piacenza and Turati (2014), instead, focus on the efficiency of the healthcare sector by assessing the effect of BOP on the average life expectancy (ALE) at different ages and the infant mortality rate (IMR), used as proxies for health status of regional populations. Finally, Depalo (2019) fouceses on the effects of BOP on hospitalization and mortality rates.

In this paper, we seek to provide evidence on whether the introduction of bail-out plans have negative effects on health status. We focus on a particular issue: the impact of mandatory operational programs (BOPs) and hence, the health regional system reconfigurations and healthcare cost-cuts may decrease the physical and psychological health status.

Following the literature (Karanikolos et al., 2013; Kentikelenis et al., 2014) on the impact of austerity measures on different health dimensions, we analyze the effect of regional bail-out plans adoption on a broad set of health measures. Furthermore, we focus on the impact of austerity policies on social distress, considering several dimensions of physical and psychological human diseases. We collect data over 17 years (from 1999 to 2015) on several dimension of health status (in line with WHO, 2015 guidelines). Our dataset is composed of data on mortality rate and its different causes (suicide rate murders rate and traffic roads rate), related to age (infant mortality rate, stillbirths rate, life expectancy rate and potential life lost – PYLL) and on mortality rate due to different diseases (alcohol, cancer, mental disorders, and also respiratory, heart and pneumonia diseases). Thus, by using these measures, we can capture the impact of austerity measures on the health status related to different specifications. Moreover, to address the effect of bail-out on morbidity, we collect data on the incidence of several infectious diseases (measles, meningococcus, parotitis, rubella, varicella, pneumococcus, aids, legionellosis and tuberculosis). These measures may provide the probability (or risk) of infectious in regions undergoing bail-out plans. Finally, we consider the effect of BOPs on healthcare services, but departing away from the "canonical" analysis on hospitalization rate, we focus our attention on the effect of BOPs on healthcare coverage related to vulnerable people and affected by social distress. Thus, we consider the HIV discharges rate, and the psychological diseases discharge rate. We adopt an Instrumental Variables approach, considering as instrument the average percentage of people with public transport low satisfaction in 2005, as a proxy of regional system inefficiency, before the adoption of bail-out plans.

In line with the literature we also measure and evalutate the effect of the same restrictive policy measures (BOP) on a wider set of indicators but we make a further assumption related to BOP, which make our analysis different from the others. In particular we ask ourselves why some regions are the different from the others and need to adopt BOP. It is not a geographic issue (regions either in the richer north or in the poorer south may have adopted a BOP), it is not a political issues (regions lead by iether center-letf or center-right may have adopted a BOP). Therefore, there must be some pre-existing (latent) characterstics that has pushed to the adoption. This point is crucial as it implies that BOP are endogenous and some further assumptions need to be done and that estimates either from a plain DID estimation model or a matched DID model would be biased.

The remainder of this article is organized as follows. We introduce a brief institutional context to explore both the devolution process of Italian Regional Health System and characteristics of regional bail-out plans in section 2.1 and 2.2, respectively. In section 3, we illustrate our Background analysis. Then, we explain our methodology strategy and Instrumental Variables assumptions in section 4. Moreover, in section 5, we present data used in our analysis and their descriptive statistics. In section 6, we present our results on health status dimensions and healthcare coverage and also their robustness check. In section 7 concludes. Additional tables on data sources and definitions are reported in the Appendix.

2 Institutional framework

2.1 The devolution process of Italian National Health System

The historical process and organizational structure of the Italian National Health System are important elements to analyze for a wide comprehension of causes and effects' bail-out plans on health outcome. As declared in the Italian Constitution, health is a fundamental right of the individual and collective interest essential healthcare services (art.32). In order to guarantee equal access and utilization of comprehensive and essential healthcare services, Italian National Healthcare System (NHS) is hierarchically structured from national to regional level, from regional to local level (Atella et al., 2019).
As several western European countries, Italy has undertaken several decentralization reforms to address a twofold goal: enhance efficiency of public health care sector and satisfy the emerging political and social desire for greater regional autonomy.

In line with Saltman et al (2007) definition [1], the Italian process of decentralization was characterized by the transfer of formal responsibility and power to make decisions from the Central Governments to "textitlower-level political authorities such as regions or municipalities", especially in the healthcare sector. According to this definition, we can define Italian NHS reorganization as a devolution process.

Since the first 90s, the process of devolution had led to the establishment of different Regional Health Systems (RHS). This process has involved in administrative, financial and fiscal aspects of the health care sector, implying more regional autonomy in programming, funding, organization and delivery of health care services in their geographical area. Furthermore,

[1] Saltman (pp.45-51, 2007) defines decentralization as "*The transfer of formal responsibility and power to make decisions regarding the management, production, distribution and/or financing of health services, usually from a smaller to a larger number of geographically or organizationally separate actors*". He identifies five decentralization mechanisms, that can be adopted combined together in various forms. Thus, he defines *devolution* (or political decentralization) as the transfer of formal responsibility and power to make decisions from central "*to lower-level political authorities such as regions or municipalities*"; deconcentration that can assume two forms: *vertical*, when decentralization is "*from a smaller number to a larger number of administrative actors within a formal administrative structure*"; or *horizontal* if it is "*from central management to other non-managerial groups such as health professionals*". Moreover, transfer of responsibility and power can occur through a mechanism of *bureaucratization* when it is "*from political levels to administrative levels*"; through *delegation and autonomization* that it is the "*transfer of selected functions to more or less autonomous public organizational management*"; or also by *privatization* when "*responsibility for particular functions is transferred from public to private actors either permanently, or for particular time periods*".

several incentives and competitive mechanisms[2] were introduced as levers for efficiency improvements and costs containment. Managerial and accountability principles were introduced in the public health sector. The Local Health Units were transformed into public entities (Local Health Authorities-ASL) and main hospitals into Agencies, managed by independent managers with the same principles and rules as private enterprises, both directly accountable to regions. Moreover, Regions in order to offer an equally delivering of healthcare services among their different territories, established the size, the managerial autonomy and organizational procedures for each their local health and hospital Agencies.

As the quasi-market and internal competition models between private and public health structures was not fully achieved[3], Regions adopted a collaboration model among health care providers and partnerships with local health authorities to enhance community care.

These reforms would have made Regions more autonomous and more responsible in terms of political, administrative and financial healthcare provisions. Thus, even though the central policies and strategies (such as the minimum amount of Essential Levels of Care - LEA) have been decided at the national level, Regions have been the main responsible for healthcare provisions and organizations (such as number of healthcare treatments, number of available hospital beds, number of general practitioners, medical and non-medical staff, and so on).

In the early 2000s, the National Health Fund[4] (NHF) was abolished and replaced by regional healthcare budget that gave more financial responsibilities and powers to each region. Differently from NHF, the regional healthcare budget was based on the capitation-based formula, adjusted for differences in healthcare needs (Decree Law 56/2000). Regional budget was allocated within the Local Health Authorities of each region whilst, regional grants were ensured through both regional taxation and healthcare co-payments (Nuti et al., 2016). Each region adopted a different level of taxation and healthcare co-payment schemes.

In general, the central Government decided the total amount of resources to be devoted to healthcare . This amount was redistributed among regions according to an appropriate formula.

[2] Private and public health structures could freely compete for the delivering of health care according to tariffs established with the diagnosis related group (DRG) international system by the Regions within maximum values adopted by the Ministry of Health.

[3] In many Regions, several healthcare services that could not be offered by public structures, were fully delivered by private sector.

[4] The National Health Fund was financed through general taxation.

In the presence of regions' deficit, the Central Government could decide ex-post the amount of deficit to cover through flexible bails out plans (Piacenza et al., 2014), without requiring any constrains or conditions on the achievement of specific level of economic results (Aimone et al, 2018).

As results of these reforms, each region ((21, considering the Autonomous Provinces of Trento and Bolzan) according to its health expenditure budget, could autonomously decide the amount of healthcare provisions and how to deliver public health, community health services and primary care through the Local Health Authorities, or the amount and type of secondary and specialist of care directly or through public hospitals or accredited private providers. Thus, Italian NHS was the result of 21 different healthcare systems.

Different studies reveal several consequences of devolution, such as the increasing disparities in healthcare provision and management among Italian Regions (Toth, 2014; Nuti et al, 2016) and also, the growing gap between northern and southern Regions, in terms of healthcare expenditure and quality (Neri, 2019, Toth, 2014). However, without any penalty mechanisms to prevent an increase of health expenditure, significant deficits persisted in some regions. In light of persisting deficits, the Central Government strengthened its control on regional healthcare expenditure by imposing a balanced budget assessment (Financial Stability Law L. 311/2004 and Financial Stability Law L. 296/2006).

Specifically, the Central Government imposed on Regions with significant deficits, the identification of deficit's causes and based on those, the development of a specific Operating Program to strengthen RHS reorganization, requalification. Through Region-State agreement (different in each Region), the Italian Government mandated a recovery process where Regions could have the right to access to bail-out plan to address pre-existing deficits and achieve balance budget, containing costs without prejudice to LEA's delivery and provision (Aimone et al., 2018).

Furthermore, the European economic crisis (2007-2008) and then the Italian sovereign debt crisis (2010-2011), worsened both national healthcare expenditure and the financial deficit of several regions . Thus, since 2007 some Regions have adopted restrictive policies (bail-out plans, BOPs) "in order to face with their serious structural and economic-financial unbalances in the regional health system" (MEF-RGS, 2014-2018).

2.2 Regional Bail-out Plans

The Bail-out Plan (BOP) is an industrial program, adopted in Italy to achieve a profound and structural reorganization of the Regional Health Service (SSR)[5]. It identifies inefficiency areas leading to economic unbalances and makes plans and implements appropriate measures to correct these problems. The BOP's adoption is compulsory if the level of regional deficit is more than 7% (threshold value up to 2010) or more than 5% (threshold value after 2010) of the total amount of total funding. The BOP duration is at least three years and each region can decide to extend it in the subsequent 3-years (or longer than the 3 years periods) until economic-financial unbalances in the regional health system will be removed (Aimone et al., 2018). Thus, Regions with a deficit higher than the threshold as fixed by the Central Government, had to adopt reorganizational plans in terms of structural endowments and staff. They were subjected to periodic monitoring aimed to control the effectiveness of commitments established concerning costs containment and LEA's provision and delivery[6]. During the first application of BOPs (2007-2009), the Central government monitored the implementation of both financial and economic balances and reorganization of health services in each RHS, assessing also the potential RHS sustainability after the end of the BOP. In the subsequent BOPs' adoptions, the monitoring activity assessed whether the economic results and LEA had been truly accomplished jointly.

Bail-out plans are characterized by the introduction of containment's measures on structural endowments and staff. Specifically, since 2005 was imposed a hospital beds number reduction up to 4.5 (per thousand inhabitants) to achieve by the 2007. During the years 2010-2012, the number of hospital beds was decreased to 4 (per thousand inhabitants) and in 2015, it was established equal to 3,7 (per thousand inhabitants), distinguishing this number between acute hospitals beds (3 beds per thousand inhabitants) and rehabilitation and long-term care (0,7 beds per thousand inhabitants). Moreover, it was established an hospital reorganization related to a reduction of hospitalization rate, namely reducing the number of hospital admissions in day-cases. Finally, it introduced staff costs containment and the block of employment turn-over.

[5] Also, in the 2016 was introduced Recovery Plans for healthcare Enterprises (Decree of Ministry of Health, 21 June 2016) that we are not going to deal with in our analysis.

[6] The share of NHF (premium quote) can be denied whether region does not achieve commitment established in BOP

In particular, it established that during the period 2007-2009, the staff expenditure should be less than the level of staff costs in the 2004 reduced by a fix percentage. This measure was extended in the subsequent periods. The Central Government established these measures to achieve a twofold goals: operating costs containment, (mainly by introducing containment measures of hospital beds and staff costs) and efficiency gains in terms of individuation of appropriate medical care and consequently choice of admission type.
Specifically, 10 out of 20 Italian Regions have carried out a bailout plan: Liguria, Abruzzo, Campania, Lazio, Molise, Sardinia and Sicily (since 2007), and Apulia, Calabria and Piedmont (since 2010). Some regions, such as Liguria and Sardinia, have successfully balanced their budget in few years (they have adopted a bail-out plan until the 2009 and 2010, respectively). Moreover, Lazio, Abruzzo, Molise, Campania and Calabria replaced their President of Region by an "ad acta commissioner" because of adoption of inadequate plans[7].

3 Literature review

The effect of the financial crisis on public expenditure is widely analyzed in the literature. The 2007-2008 economic crisis has worsened several European countries, and the main effect was the large cuts of public expenditures and consequently the reduction of public services, such as welfare protections, healthcare services, education. Several studies analyzed the effect of measures adopted to tackle the financial crisis (healthcare costs cuts) on health systems and also, in several health dimensions (Karanikolos et al., 2013; Kentikelenis et al., 2014; McKee et al., 2012; Appleby et al., 2015). In particular, Kentikelenis et al., (2014) and Karanikolos et al. (2013) investigated on the relationship between the economic downturns and some health dimensions, such as mortality, suicides and life expectancy. The largest cuts on public health budgets have an adverse effect on access care in time. The lack of obtaining timely quality medical care may increase the mortality rate and reduce the life expectancy at birth. Moreover, McKee et al. (2012) illustrated how different measures of austerity adopted, such as reduction on social protection, public health funding and healthcare coverage, can impact on several social and health dimension (poverty, depression, suicides, mental disorders).

[7]The Law states that the ad acta commissioner has to be the president of the region and regional taxes have to increase automatically up to a predefined value (Nuti et al., 2016).

In line with this literature, we investigate the effect of austerity measures adopted in Italy on several health dimensions and two measures of healthcare services during the period 1999-2015. In order to make Regions more autonomous and efficient, the Central Government undertook a process of regional decentralization and in particular, a devolution process of Regional Health systems. However, these goals have not truly achieved. The substantial differences in funding between regions (unevenly distributed tax base[8]) and the increasing deficit in several regions (mainly in the central and southern parts of the country) create disparities on healthcare provision and delivering among regions (Bordignon and Turati, 2009; Piacenza and Turati, 2014; Di Novi et al., 2019). In light of the European economic crisis and then the Italian sovereign debt crisis, the Italian the Central Government applied austerity measures on healthcare sector by introducing operating programs (bailout plans) for the reorganization, qualification or improvements of the regional health service, to restore the economic equilibrium, while respecting the essential levels of assistance (LEA).

Several authors have explored the effect of some institutional factors on health spending (Bordignon and Turati, 2009; Piacenza et al., 2014; Nuti et al., 2016; Atella et al., 2019). As one of the most pillars of bailout plans is the staff costs containment, Aimone et al. (2018) revealed a general decrease of medical and non-medical staff with an increase of flexible contractual forms instead of permanent ones the reduction of the number of beds. Moreover, Atella et al. (2019) analyzed the effect of bailout plans on healthcare sector productivity in the SSN, in comparison with the English NHS. They showed that the overall NHS productivity growth index increased by 10% over the period 2004-2011, at an average of 1.39% per year, while SSN productivity increased overall by 5%, at an average of 0.73% per year. They addressed the differences in productivity to the policy adopted in each country. The main aim in Italy was the efficiency in the provision of services achieved by a policy of cost containment and rationalized provision. Other authors analyzed the effect of bailout plans on both healthcare services and health outcomes. For instance, Depalo (2019) found a negative effect of bailout plans on hospitalization rate and mortality rate, without any increase in efficiency gains, while

[8]Regions may vary the tax rates through taxing corporations and a regional surcharge on income tax. However, as the financing system is based on unequal tax based distribution, in poorer regions where a room for manoeuvre is small, they may only increase tax rates (more than high-income regions), making disincentives for economic activities (Italy HiT, 2014)

Piacenza and Turati (2014) investigated how the bailout expectations can affect the expenditure for healthcare policies carried out by decentralized governments. They found that bailout expectations do not significantly affect people's (measured in terms of average life expectancy and infant mortality rate)

4 Methodology

In our analysis we estimate the effect of bail-out plans on several health dimension and two type of hospital discharges between regions adopting a bail-out plan and regions without it. In literature we found different approaches to address this scope. For instance, the Ministry for Health published a monitoring of LEA delivering comparing regions with and without BOPs by using as benchmark the observed outcome in regions without recovery plan and so they impose a state invariance restriction (SiVeAS, 2014).

Whilst Depalo (2019) use a methodological approach based on the identification of a set that estimates several bounds instead of single point (Manski, 1990)[9]

We estimate the causal effect of bail-out plans in a non-experiment design considering the problem on the selection on unobservable variables. As we know, the adoption of bail-out is compulsory for regions with level of deficit higher than a defined threshold, but also several fiscal and organizational reforms on health regional systems drove to significant differences on efficiency and quality of healthcare delivery among regions (Toth, 2014; Nuti et al., 2016; Neri, 2019; Cicchetti and Gasbarrini, 2016). These differences may also depend on unobservable features of the regional systems, leading some regions to be inefficient in the healthcare sectors (their deficits increase) and hence, undergo the program instead of others.

The presence of such latent, pre-existing characteristics not related to either geographical or political issues (Regions undergoing BOPs are located both in the north and in the south of the country, are led by centre-left or centre-right governors) implies that pre-existing (latent)

[9] Depalo (2019) verified a non-common trend between regions with BOP and regions without it, as required by using a Difference-in-Differences (DID) or a Synthetic Control (SC) method. He departs from considering a state invariance restriction or also a time invariant restriction (by using as benchmark for regions with BOP before their adoption) but assumes a possible set of admissible effects or bounds (Manski, 1990). He does not identify a specific true benchmark for regions which underwent the program, but its benchmark is bounded within a range. Then, he compares the actual outcome of the regions with the range of the benchmark, that provides a set of admissible parameters (lower and upper bound) for the treatment effects.

regional characteristics have pushed to the adoption of a BOP. This point is crucial as it implies that BOPs are endogenous and some further assumptions need to be done and that estimates either from a plain DID estimation model (or a matched DID model) would be biased.

In order to isolate the true effect of bail-out plans on health outcomes and discharges, and hence, overcome the selection bias into treatment (selection on unobservable variables), we adopt instrumental variable approach. In our analysis, we use the average percentage of people with negative judgement of public transport as instrument. Different studies (Bentivogli et al., 2008; Suguiy et al., 2013; EC Mobility and Transport, 2019) compare the transportations planning at country level or at different regional level (regions, municipalities) to measure the efficiency and performance of the public policy adopted. For instance, Bentivogli et al (2008) show a heterogeneity of public transport among Italian regions in terms of efficiency after the adoption of several reforms during the years '90s.

As, the public transports are essential to daily lives of citizens, we use the average percentage of people with negative judgment of public transport (buses, tram, trolleybuses) in each region as a proxy of inefficiency of public policies adopted by regions to meet the basic needs of the population.

Following the econometrics literature (Sargan, 1958; Angrist and Krueger, 1991; Abadie et al., 2002; Angrist and Imbens, 1995; Angrist and Pischke, 2009; Angrist, 1991; Angrist et al., 1996), we need to assume an instrumental variable z directly correlated with the selection process (D) but uncorrelated with the outcome Y in order to produce consistent estimation of average treatment effect (ATE).

We estimate a consistent ATE by adopting a direct two-stage least squares (Direct-2SLS). It is a simultaneous system of two OLS regressions where in first stage (equation 1), we calculate predicted values of the endogenous variable D_{it} assuming that z_{it} (our instrumental variable) is directly correlated with the selection process (D_{it}) and captures the effect of z_{it} on D_{it}, adjusting for covariates X_{it}.

$$D_{it} = \eta + \sum_{t=1}^{17} \delta z_{it} + \sum_{t=1}^{17} \sum_{k=1}^{n} \beta_1 X_{kit} + \nu_{it} \tag{1}$$

Then, we use these predictions as regressor in a second OLS regression to estimate the outcome Y_{it} (equation 2 expressed in reduced form) assuming that (X_{it}, z_{it}) are uncorrelated with the error term eit, thus (X_{it}, z_{it}) are exogenous and ϵ_{it} is is still correlated with D_{it}

$$Y_{it} = \beta_0 + \sum_{t=1}^{17} \beta_1 D_{it} + \sum_{t=1}^{17} \sum_{k=1}^{n} \beta_2 X_{kit} + \sum_{t=1}^{17} \beta_3 Z_t + \epsilon_{it} \tag{2}$$

where Y_{it} represent our different measure of health outcomes and discharges (the total mortality rate and measure of mortality rate by causes, by age and by diseases; incidence of infectous diseases and discharges rate for HIV and psychological diseases) for each region i in the year t; D_{it} is is a dummy variable taking value equal to 1 in the year when Regions adopt it and in the subsequent years (*bop*); X_{kit} is the k-th observable time-variant factor (socio-economic characteristics, inputs) affecting our dependent variables for region i in year t. The number of controls depends on the depedend variable of interest. For instance we use 9 different controls in mortality rates and discharges' analyses ($n = 9$) and 8 controls for incidence of infectious diseases estimate ($n = 8$). z_{it} represents year fixed effect while ϵ_{it} is the error term.

This approach allows to measure the effect of BOPs on several outcomes Y_{it} by assuming that the selection into the program may depend on the same factors affecting the outcome plus z_{it} (our instrument), namely through an indirect effect on Y_{it}. The relation between the endogenous variable D_{it} and the outcome Y_{it} exists, but only through an indirect link produced by the direct effect of z_{it} on D_{it}. In other word, we need to assume zit correlated with D_{it} and uncorrelated with any other determinants of our dependent variables state that z_{it}. If both requirements are achieved, our instrumental variable z_{it} satisfies the *exclusion restriction* property, as required for the identification of casual parameters of interest in the IV method (Angrist and Pischke, 2009).

Moreover, in line with empirical literature we detect for weakness of our instrument by applying several statistical tests. In our analysis we consider one-dimensional model on Y_{it}, with z_{it} is *n x 1* and D_{it} is *m x 1* and $E[\epsilon_{it}, z_{it}] = 0$ (i.i.d data $\{Y_{it}, D_{it}, z_{it}\}$). Assuming D_{it} as endogenous and z_{it} as exogenous ($E[\epsilon_{it}, z_{it}] = 0$), we also first assurance that our instrument is relevant, namely $E[z'_{it}, D_{it}]$ has rank n. Whether this condition is achieved, our model is identified.

Additionally, we analyse the Kleibergen-Paap rk LM statistic for the underidentification test, where the null hypothesis is Ho: matrix of reduced form coefficients has rank=m1-1 (underidentified). Rejecting the null hypotesis we found that the model is identified (matrix has rank=m1). This is a LM test that has a null distribution does not depend on μ^2 but on $\chi 2$.

Then, we compare the first-stage F statistic with a cut-off (Stock-Yogo, 2005: weak ID test critical values for endogenous regressor; Table 1).

Table 1: Stock-Yogo cut-off table

cut-off	critical values
10% maximal IV size	16.38
15% maximal IV size	8.96
20% maximal IV size	6.66
25% maximal IV size	5.53

Stock and Yogo (2005)[10] showed that the first-stage F-statistic is distributed as a non-central $\chi 2$ with a non-centrality parameter directly related to the concentration parameter μ. As a result, the first-stage F-statistic can be considered as an indicator of the value of μ. Specifically, in the first-stage F-statistic is the statistic for testing $\delta = 0$ in the first stage regression (equation 1). We know that $F = 1 + \mu^2/m$, so we can estimate μ^2/m as $F - 1$. Then we compare the obtained μ^2/m with cut-off values as reported in Stock-Yogo table and finally if accordin to the Rule of thumb, wheter $F < 10$ means that our instrument is weak.

As the first-stage using the F statistics reliesing heavily on the assumption of conditional homokedasticity, and we apply the robust options in the first stage, we also analyse the Kleibergen-Paap rk Wald statistic under the null hypothesis is Ho: equation is weakly identified.

All statistical tests are provided by running ivreg29 command in Stata 14.

[10] They provide critical values that depend on the number of endogenous regressors, the number of instruments, the maximum bias.

5 Data

5.1 Data Sources

We have collected longitudinal data, available annually, for a period of 17 years (from 1999 to 2015) for 19 Regions and 2 Autonomous Provinces (Trento and Bolzen) for a total of 357 observations. All data are collected by using several sources (Health for All database, HFA; I.Stat data warehouse; Italian Society of Hygiene, Preventive Medicine and Public Health, SITI; Italian National Institute of Health, ISS and Ministry of Health. Our data contain information on several measures of health outcome, proxy measures of service coverage, inputs variables and socio-economic characteristics. Data sources details are in Appendix (Table 21)

5.2 Variable Definitions and Measurements

5.2.1 Dependent Variables

In line with "100 core Health Indicators" (WHO, 2018), we consider several measures of health outcomes and two measure of health services coverage. As shown in Table 2, we consider several measures of mortality rate and incidence of infectious diseases (ISS) as measures of health outcomes. In line with the literature (Karanikolos et al., 2013; Kentikelenis et al., 2014; Ruhm, 2015), we consider all measures of health outcomes indirectly affected by austerity measures. Thus, we analyse the total mortality rate defined as the number of deaths per 10,000 inhabitants, for both males and females and all age groups. Also, we extend our investigation considering subcategories of external deaths such as suicides rate, murders rate as proxy measures of depression and mental disorders, and mortality due to traffic roads deaths as a proxy measure of general economic conditions (Ruhm, 2015) . Suicides rate is defined as the number of deaths by suicides per 10,000 inhabitants, for both males and females and all age groups; murders rate is the number of murders per 100,000 inhabitants, for both males and females and all age groups; while the traffic roads death rate is built as the percentage of the total killed in road accidents out of the total road accidents (killed and injuries) for both males and females and all age groups. Moreover, we investigate the effect of BOP on mortality by age considering infant mortality rate at birth, as the number of deaths at age 0 per 10,000 live births for both males and females; life expectancy at birth as the average number of years that each infant (aged 0) is likely to live for both males and females. In line with OECD definition, we built the Potential years of life lost (PYLL) as the sum of deaths (considering mortality rate) occurring at different age classes (0-14; 14-24;25-34; 35-44; 45-54; 55-64; 65-74; 75 and over) and multiplying this with the number of remaining years to live up to a selected age limit (age 70). Finally, to investigate the impact of BOPs on illness and vulnerable people, we consider a health status subgroup including mortality by several diseases, such as alcohol, cancer, cirrhosis, mental disorders, total hearth diseases and total pneumonia (for the last two diseases we consider people aged 65 and over too). Here, each mortality rate is the number of deaths due to each specific disease per 10,000 inhabitants, for both males and females and all age groups.

As several studies highlighted that economic downturns could have a negative impact not only in the transmission of infectious diseases but also in the control of them (Suhrcke et al., 2011, Kentikelenis et al., 2015, Quaglio et al., 2013; Paes-Sousa et al., 2019), we investigate the possible effect of BOPs on the incidence of several diseases. Our data on incidence of infectious diseases represent the rate of incidence per 100,000 inhabitants, for both males and females and all age groups.

Additionally, we analyse how hospitals containment measures can affect more vulnerable people. Thus, we consider measure of discharges (as proxy of admissions) for people affected by HIV and psychological diseases, detaching from recent literature carried out on hospitalization rate (Depalo, 2019). Both HIV and psychological diseases discharges rate are calculated per 10,000 inhabitants, for both males and females and all age groups. All dependent variables are expressed in log-transformation form.

Table 2: List of dependent variables

Health outcomes		Healthcare services covarage
Mortality rate measures	Incidence of Infectous diseases	
Mortality rate	Measles	HIV discharges rate
	Meningococcus	Psycological diseases discharges rate
by causes	Parotitis	
Suicides rate	Rubella	
Murders rate	Varicella	
Traffic road deaths rate	Pneumococcus	
	Aids	
by age	Legionellosis	
Infant mortality rate	Tuberculosis	
Life expectancy at birth		
Potential years of life lost (PYLL)		
by diseases		
Alcohol		
Cancer		
Cirrhosis		
Mental disorders		
Respiratory disease		
Heart diseases		
Heart diseases aged 65+		
Pneumonia		
Pneumonia aged 65+		

5.2.2 Policy variable

We analyse the effect of austerity measures on our dependent variables by using a dummy variable *bop* is equal to 1 in the exact year when the Region adopts the BOP and onwards. In our analysis, BOP is the combination of hospital rate reduction, hospital beds reorganization and staff costs containment. Specifically, we consider the adoption of BOP from 2007 onwards for Liguria, Abruzzo, Campania, Lazio, Molise, Sardinia and Sicily; and from 2010 onwards for Apulia, Calabria and Piedmont.

5.2.3 Instrumental variable

We use the average percentage of people with negative judgement of public buses transportation for each region (*neg_ transp*) in the year 2005. The variable is the result of percentage difference of people (aged 14 years old and over) using buses, tram and trolleybuses and with high and very high satisfaction for these public transports, as collected from I.Stat Datawarehouse. The satisfaction is measured considering several quality and efficiency aspects of public service (frequency and on-time buses, seats availability, cleanliness of buses, possible connections with other municipalities, ticket costs, buses travel time, suitable waiting stops and time schedules). We calculate the mean of all percentages of people for each aspect in each region in the year 2005. The difference by 100, is our average percentage of people (aged 14+) with negative judgement of public transport in the year 2005. We consider this variable as a proxy of inefficiency in each region. For that reason, we do not consider also the average percentage of people with negative judgement on train transportation because of lack on separate data on regional or inter-regional travel. Moreover, we consider the 2005 as the "baseline" year before bails-out introduction (in 2007 and in 2010).

5.2.4 Control variables

We control our different dimensions of health status and healthcare service coverage by using several measures, as shown in Table 3. Some socioeconomic variables are common for all dependent variables, such as the *total population*, the *proportion of people aged 85 and over*, *household poverty index* and *GDP per capita (at current prices)*. As bail-out plans are measures adopted also considering the population size of each region, we include the total population as control for all dependent variables. Moreover, we use the proportion of people aged 85 and over to control the relationship between several dimensions of health status and discharges and the elderly population. Additionally, we include household poverty index, as the measure of the incidence of poor households out of the total household in each region. It can be considered as a proxy of deprivation and vulnerability at the household level but also, on a broader definition, as a proxy of social protection measures (UN, 2012). Also, we include GDP per capita at current prices (expressed in € mln) as a measure of the region's wealth.

As proxy of social ties, we control all dimensions of health status by using the *average number of household components* (Boitsov and Samorodskaya, 2016). In line with Ruhm's works (2000, 2015), we control all mortality rate dimensions by using the *employment-population ratio for people aged 15 and over* and *young unemployment rate (people aged 15-24)*, both as proxy measures of economic downturn effects.

As after periods of economic downturns, it was found an increasing level of depression and anxiety disorders in young people (Paes-Sousa et al., 2019), we control discharges measures including the *proportion of young population (people aged 0-14)*. Whilst we consider the *proportion of population aged 15-64* as proxy of active population in the analyses of all dimensions of health status. Additionally, we include the *percentage of people with the primary education level* (or no education) and *with tertiary education* as proxy of low[11] and high income, respectively (Boitsov and Samorodskaya, 2016). We use the first control for incidence of infectious diseases and the latter as control for different mortality rates.

As infectious diseases may affect different population groups, we add the *proportion of pediatricians with more than 800 patients* when we analyze measles, meningococcus, parotitis, rubella varicella and pneumococcus, as typical child infectious diseases; while we consider the *proportion of general practitioners with more than 1000 patients* for aids, legionellosis and tuberculosis, that can have more incidence on adults. Both measures can be considered as proxy measures of hospital healthcare unmet. In line with the healthcare literature, we analyze the effect of bail-out plans on discharges measures considering inputs of a production function (e.g.,

[11]In general, people with low level of education have a low income and consequently, they are not always access to risk factors' information, symptoms of diseases, opportunities to receive medical care, etc.

Wagstaff, 1989; Duggan, 2000; Horwitz and Nichols, 2009). Thus, we consider the *total number of acute ordinary public beds* as measures of capital and the *rate of doctors* and *nurses* as measures of labor. We consider the acute ordinary public beds because of the reduction of acute hospital beds is one of the BOP pillars.

Table 3: Controls for each dependent variable group

	Dependent variables				
	Total MR and MR by causes	MR by age	MR by diseases	ISS	HIV and psycological diseases discharges
Total population (in ln)	X	X	X	X	X
Population aged 0-14 (in ln)					X
Population aged 15-64 (in ln)	X	X	X	X	
Population aged 85+ (% in ln)	X	X	X	X	X
Acute ordinary pubblic beds (rate in ln)					X
Doctors (rate in ln)					X
Nurses (rate in ln)					X
Pediatricians with 800+ patients (% in ln)				X(a)	
General practitioners with 1500+ patients (% in ln)				X(b)	
Employment-population ratio 15+ (in ln)	X	X	X		
Unemployed rate aged 15-24 (in ln)	X	X	X		
Primary/No education (% in ln)				X	X
Tertiary education (% in ln)	X	X	X		
Average number of household components	X	X	X	X	
Household poverty index (in ln)	X	X	X	X	X
GDP per capita at current prices (in ln)	X	X	X	X	X

Note:
(a) Only for measles, meningococcus, parotitis, rubella, varicella and pneumococcus
(b) Only for aids, legionellosis and tuberculosis
MR= mortality rate; ISS= Incidence of Infectous diseases

5.3 Aggregate Data Patterns and Descriptive Statistics

Table 4 shows descriptive statistics for our sample overall. Among 357 observations in our sample, about 3% of Regions adopt a BOP in a given year (bop=1 in the exact year when regions adopt a BOP). In general, average values on mortality rate and its causes on average are equal to 100 for the total mortality rate, 74 for suicide rate, the traffic roads rate is around 73 and the murders rate is around 84. On average, Italian people have a life expectancy at least 80 years old, while the potential years of life lost for people aged 0-70 is around 269 per 10.000 inhabitants. The rate of infant mortality is around 34% on average, while the stillbirth rate is 16% on average. Analysing the mortality rate by several diseases, we can see that on average the most relevant cause is heart diseases in people over 65 years old (around 58% versus 13% of total people affected by heart diseases), followed by people affected by cancer (around 29%). Also, we can notice that people with pneumonia diseases are on average around 2% while if we consider age class over 65 years old, the rate increases up to 7% on average. Moreover, respiratory diseases seem to be a significant cause of mortality (on average, around 7% of people die for this disease). Remaining pathologies look having a less impact on mortality rate (such as mortality by alcohol abuse that is on average, only 0,05% and cirrhosis around 1,5% on average). As incidence measures the proportion of people affected by a specific disease during a specific period (meant like new cases), we can consider incidence as the risk (namely the probability) of each individual to be affected by specific diseases. Our results show that during the period 1999-2015, people are more likely to be affected by varicella (average incidence equal to 146), followed by parotitis (14), pneumococcus (7) and measles (6). The average incidence to be affected by TBC is, on average equal to 4, while the incidence of both meningococcus and rubella is equal to 3. The incidence of both aids and legionellosis is less than two on average. Finally, the discharge rate for psychological diseases are greater than people affected by HIV on average (52 vs 4). Descriptive statistics of control variables are shown in table 5.

Table 4: Summary statistics - full dataset

Variable	Mean	Std. Dev.	Min.	Max.	N
Dependent variables					
Mortality rate by causes					
Mortality rate	100.775	13.737	74.510	144.73	349
Suicides rate	0.743	0.248	0.31	1.93	349
Traffic road deaths rate	2.731	1.066	0.692	7.516	315
Murders rate	0.835	0.606	0.124	4.083	259
Mortality rate by age					
Life expectancy at birth	81.19	1.203	77.58	83.69	357
Infant mortality rate	34.448	11.472	2.05	83.260	349
Still births rate	16.219	6.547	0	37.44	349
Potential years of life lost (PYLL)	268.689	44.714	171.15	409.386	349
Mortality rate by diseases					
alcohol	0.048	0.039	0	0.26	349
cancer	28.803	4.652	19.04	40.03	349
cirrhosis	1.461	0.472	0.62	3.06	349
mental disorders	2.288	1.08	0.5	7.59	349
respiratory disease	6.954	1.255	4.69	10.88	349
heart diseases	13.152	2.702	8.279	21.13	349
heart diseases aged 65+	58.537	9.947	34.2	88.210	349
pneumonia	1.558	0.690	0.37	3.76	349
pneumonia aged 65+	7.117	3.054	1.97	16.73	349
Incidence of Infectious diseases					
measles	6.274	26.025	0	383	322
meningococcus	3.157	3.118	0	24.9	323
parotitis	14.146	33.227	0	273.1	266
rubella	3.026	10.272	0	154.12	322
pneumococcus	6.833	9.381	0	53.355	323
varicella	146.282	120.38	0	605.58	266
aids	1.957	1.284	0	6.94	323
legionellosis	1.268	1.251	0	6.65	323
tuberculosis	4.43	4.121	0	19.46	323
Healthcare services coverage					
HIV discharges rate	4.007	4.68	0.16	39.45	357
Psycological diseases discharges rate	52.104	16.376	22.5	99.210	357
Policy variables					
bop (in the exact year of adoption)	0.028	0.165	0	1	357
bop	0.196	0.398	0	1	357
neg_transp	23.846	24.498	0	68.8	357
Controls					
Total population	2808148.221	2385475.045	118754	10005482	357
Population aged 0-14	395221.98	350063.512	14992	1421772	357
Population aged 15-64	1853275.857	1587353.158	81201	6490987	357
Population aged 85+ (%)	2.626	0.693	1.2	4.71	357
Avg n. of household components	2.51	0.212	2.05	3.12	336
Employment-population ratio (aged 15+)	44.539	6.963	30.35	58.66	357
Unemployed rate aged 15-24	27.913	15.168	2.28	66.25	357
General practitioners with 1500+ patients (%)	7.977	0.833	4.75	9.370	315
Pediatricians with 800+ patients (%)	63.126	13.211	11.76	87.17	315
Acute ordinary pubblic beds (rate)	33.207	6.318	17.61	50.84	355
Doctors (rate)	18.834	2.568	12.82	27.22	315
Nurses (rate)	48.1	8.395	31.85	63.89	315
Primary/No education (%)	28.717	8.34	13.79	46.93	357
Tertiary education (%)	9.295	2.856	3.65	18.17	357
Household poverty index	12.022	7.481	2.3	31	247
GDP per capita at current prices (in € mln)	25052.552	6751.440	12407.292	41445.832	357

Table 5 compares summary statistics between Regions with a BOP with those that do not adopt a BOP (i.e. for them bop = 0 in all years), considering the period from the exact year of BOP introduction by the end of our sample. The average differences' magnitude between two samples analysed, is not so large; but it is significant for several of our dependent variables, especially if we consider mortality related to diseases and measures of the incidence of different infectious diseases.

On average, the total mortality rate and also its specifications by causes are lower in regions adopting a BOP, i.e. on average, the difference in means for suicide rates equal to -0.190, even though it is not significant). The mortality rate due to traffic roads accidents is higher in Regions adopting a BOP, even though the difference in means is not significant. Overall, differences in means of mortality rates related to age, are negative and significant. Thus, people living in regions with BOP have a life expectancy slightly longer compared to people living in regions without BOP (on average almost 82 instead of 81 years old) and also, a narrowed potential year of life lost on average (on average the PYLL in regions with BOP is 256 instead of 272 in those without BOP). Furthermore, on average, the mortality rate among different diseases is less in regions with BOP than regions without it, except for mortality rate due to mental disorders. Even though it is higher in regions with BOP, the difference between regions with and without BOP is not significant (0.045). Likewise, on average, the incidence of infectious diseases is smaller in regions with BOP.

We can observe a lower discharge rate for psychological diseases in regions with BOP compared to regions without BOP (46% vs 54%, respectively) but during the same period analysed, the discharge rate of people affected by HIV is higher in regions adopting a BOP. In general, significant differences in means are revealed for controls variables between regions adopting a BOP and those without it. These findings suggest that these factors may play an important role when we consider the overall time span after the introduction of bail-out plans. On average total population and its groups by age classes, are higher in regions adopting a BOP rather than those without it. The opposite and significant sign of differences in means of both the employment-to-population rate and young unemployment rate show the same result. In regions adopting a BOP may exist work-related problems. Moreover, on average, both the percentage of general practitioners (with more than 1500 patients) and paediatricians (with more than

800 patients), are higher in the regions adopting a BOP. However, the difference in means' magnitude of the percentage of paediatricians is larger than the percentage of general practitioners (4.7% instead of 0.25%, respectively). Both differences may suggest a possible hospital healthcare services unmet in regions adopting a BOP. Analysing controls used as capital and labour inputs, we notice that on average, both rates of acute ordinary public beds and nurses are lower in regions with BOP, whilst doctors' rate is higher in regions adopting a BOP. These findings are in line with the main bails-out plans targets. Socio-economic variables reveal that regions adopting a Bop are more educated on average (the percentage of people with tertiary education is higher in regions with a BOP than those without it; 11% vs 9%, respectively). Then, on average household living in a region adopting a BOP are poorer than households living in regions without a BOP (household poverty index is around 16 while it is only 11 in regions without BOP). Finally, the GDP per capita at current prices is higher in regions without BOP, on average equal to 25.789 instead of per capita GDP of regions with BOP equal to 22.035 (both expressed in € mln).

Table 5: Mean Comparison: Regions with BOP vs. Regions without BOP (by the end of our sample period; 1999-2015)

	bop (= 1 if regions adopt a BOP		bop (= 0 if regions do not adopt a BOP		Difference
Variable	N	Mean	N	Mean	
Dependent variables					
Mortality rate by causes					
Mortality rate	70	100.139	279	100.935	-0.795
Suicides rate	70	0.592	279	0.781	-0.190***
Traffic road deaths rate	70	2.603	245	2.768	-0.164
Murders rate	69	0.906	190	0.809	0.097
Mortality rate by age					
Life expectancy at birth	70	81.548	287	81.103	0.445***
Infant mortality rate	70	35.483	279	34.188	1.295
Still births rate	70	16.22	279	16.219	0.001
Potential years of life lost (PYLL)	70	255.667	279	271.957	-16.290***
Mortality rate by diseases					
alcohol	70	0.038	279	0.051	-0.013**
cancer	70	27.149	279	29.218	-2.069***
cirrhosis	70	1.376	279	1.482	-0.106*
mental disorders	70	2.324	279	2.278	0.045
respiratory disease	70	6.613	279	7.039	-0.427**
heart diseases	70	12.684	279	13.269	-0.584
heart diseases aged 65+	70	56.993	279	58.924	-1.931
pneumonia	70	1.099	279	1.673	-0.574***
pneumonia aged 65+	70	4.89	279	7.676	-2.786***
Incidence of Infectious diseases					
measles	69	3.724	253	6.969	-3.245
meningococcus	70	1.807	253	3.53	-1.723***
parotitis	46	0.669	220	16.963	-16.295***
rubella	69	0.726	253	3.653	-2.927**
pneumococcus	70	6.045	253	7.052	-1.007
varicella	46	44.91	220	167.478	-122.568***
aids	70	1.371	253	2.119	-0.747***
legionellosis	70	0.905	253	1.369	-0.463***
tuberculosis	70	1.755	253	5.171	-3.415***
Healthcare services coverage					
HIV discharges rate	70	4.699	287	3.838	0.861
Psycological diseases discharges rate	70	46.043	287	53.582	-7.539***
neg_transp	70	56.046	287	15.993	40.053***
Controls					
Total population	70	3399776.7	287	2663848.592	736000**
Population aged 0-14	70	490160.271	287	372066.3	118000**
Population aged 15-64	70	2242643.371	287	1758308.171	484000**
Population aged 85+ (%)	70	2.768	287	2.591	0.176*
Avg n. of household components	70	2.52	266	2.507	0.013
Employment-population ratio (aged 15+)	70	38.924	287	45.908	-6.984***
Unemployed rate aged 15-24	70	39.222	287	25.155	14.067***
General practitioners with 1500+ patients (%)	54	8.18	261	7.935	0.245**
Pediatricians with 800+ patients (%)	54	66.996	261	62.325	4.670**
Acute ordinary pubblic beds (rate)	70	27.955	285	34.496	-6.541***
Doctors (rate)	54	19.038	261	18.792	0.246
Nurses (rate)	54	42.12	261	49.337	-7.217***
Primary/No education (%)	70	23.57	287	29.973	-6.403***
Tertiary education (%)	70	11.234	287	8.822	2.412***
Household poverty index	62	15.669	185	10.8	4.869***
GDP per capita at current prices (in € mln)	70	22035.119	287	25788.512	-3753.392***

To have a more comprehensive representation of dependent variables over time, we graphically illustrate their trends in level. Figures 1-5 represent the level of each health outcomes (mortality rate, mortality rate by causes, by age and by diseases; the incidence of infectious diseases) and discharges for regions adopting BOP in 2007 (blue line) and in 2010 (red line) and regions without BOP (comparator group). Each figure shows dependent variables trends (from 1999 to 2015) according to the year of BOP adoption compared to regions do not adopt a bail-out plan (dark line).

We present briefly trends of each dependent variable in regions adopting a BOP compared with regions without BOPs. The level of total mortality rate increases after the introduction of BOPs but in comparison with regions without a BOP, it is higher in regions adopting a BOP in 2007 while it is tendentially equal to the level of comparator group for regions adopting a BOP in 2010 (especially after 2014). The level of suicides rates in regions with BOP is lesser than regions without BOPs but their level increase after the adoption of a BOP. The levels of both murders rate and traffic roads deaths rate are tendentially decreasing over time, but they are higher in regions with BOP. The level of life expectancy at birth (LEB) is tendentially rising over time in all groups, but people living in regions with BOP have a LEB's level lesser than regions without BOP. Conversely, PYLL decreases over time in all groups, but its level is higher in regions with BOP. In general, the levels of both infant mortality rate and stillbirth rate diminishing over time among the three groups, but they are higher in regions adopting a BOP. The overall level of level of mortality rate by diseases increases over time, especially for mortality due to mental disorders, respiratory diseases, cancer and pneumonia, even though their level of mortality rate is lesser in regions with BOPs than regions without them. Interestingly, the levels of both mortality rate due to cirrhosis and hearth diseases decrease over time, but the level of mortality rate due to cirrhosis is higher in regions adopting a BOP. The level of mortality rate due to alcohol present a non-linear trend pattern among the three groups, but its level in regions with BOPs is below the level of regions without BOPs.

In general, the level of incidence of infectious diseases is decreasing over time, except for people affected by pneumococcus and legionellosis. The level of these incidences in regions with

BOPs is almost below than level of regions without BOPs. Finally, the level of both discharge rates analyzed decrease over time, even though the level of HIV discharge rate in regions with BOPs is higher than the level of regions without them; whilst only the level of discharge rate for psychological diseases in regions adopting a BOP in 2007 is higher than both regions adopting BOPs in 2010 and without BOPs.

Figure 1: Mortality rate and its causes 1999-2015

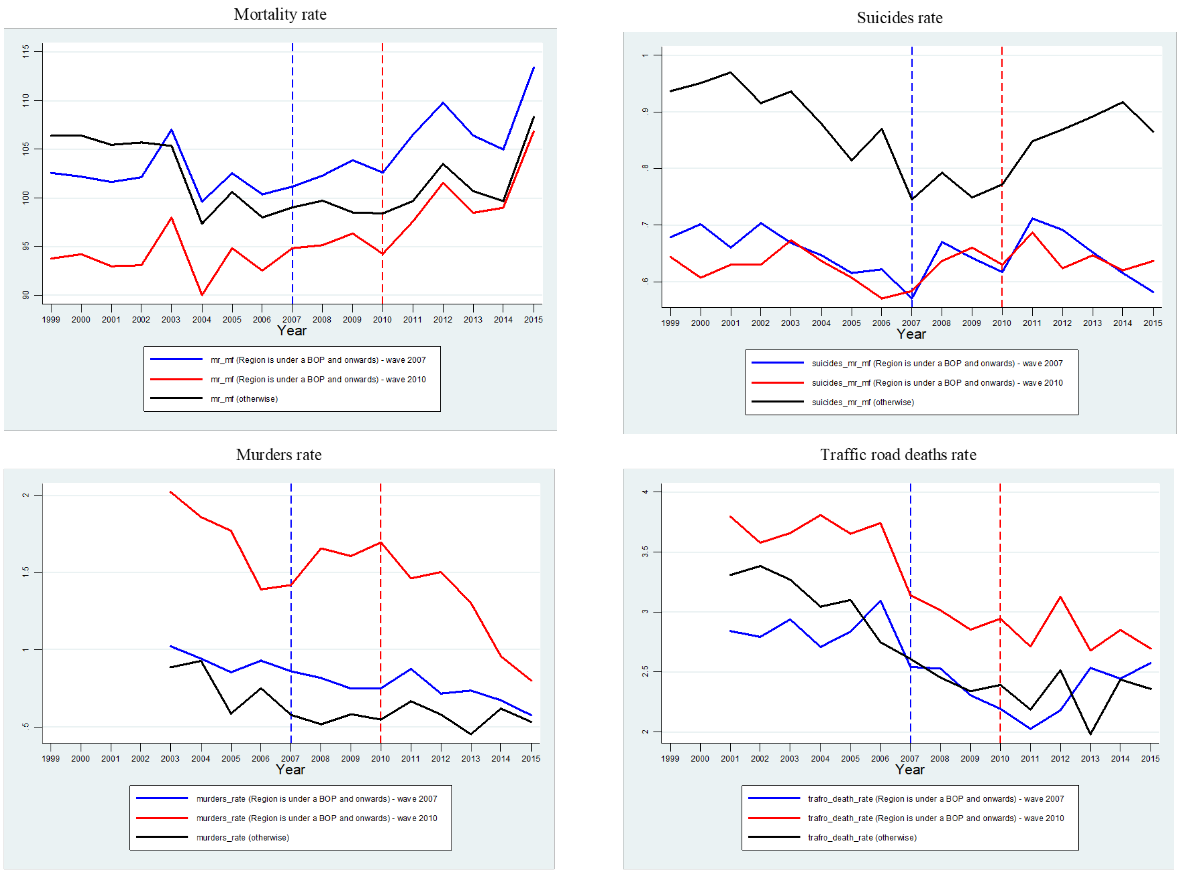


Figure 2: Mortality rate by age 1999-2015

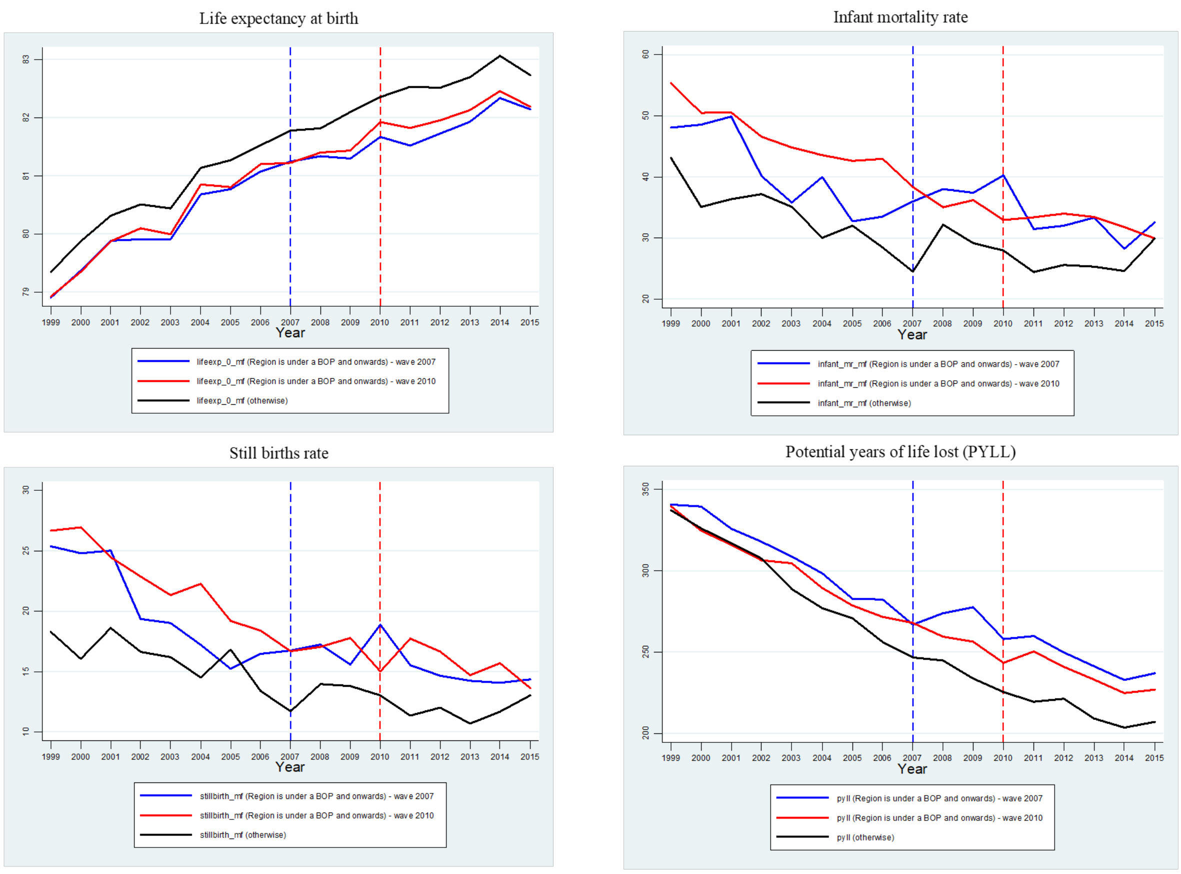
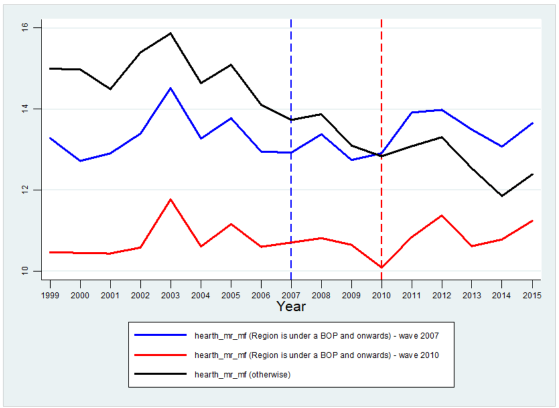
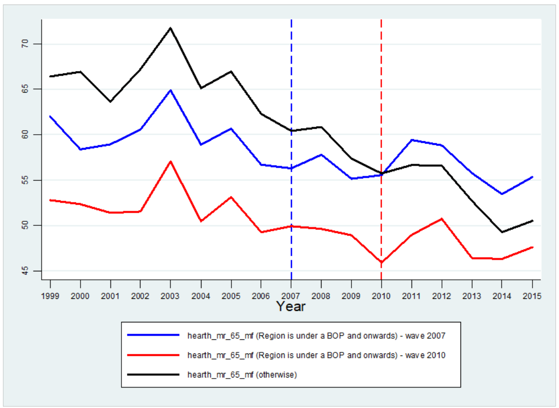


Figure 3: Mortality rate by diseases 1999-2015

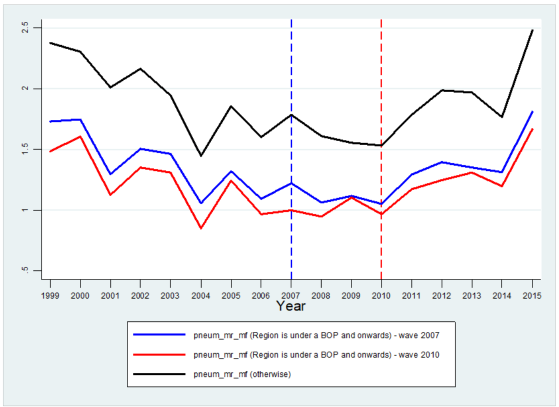
Heart diseases



Heart diseases aged 65+



Pneumonia



Pneumonia aged 65+

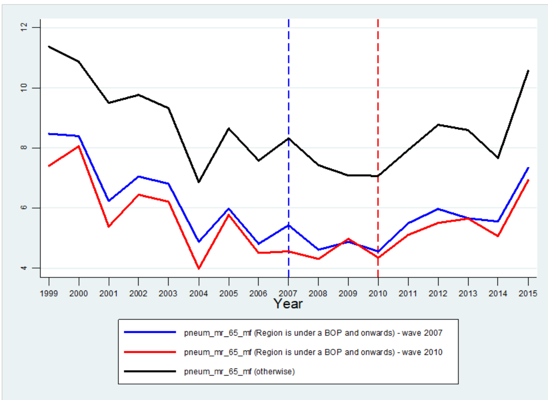
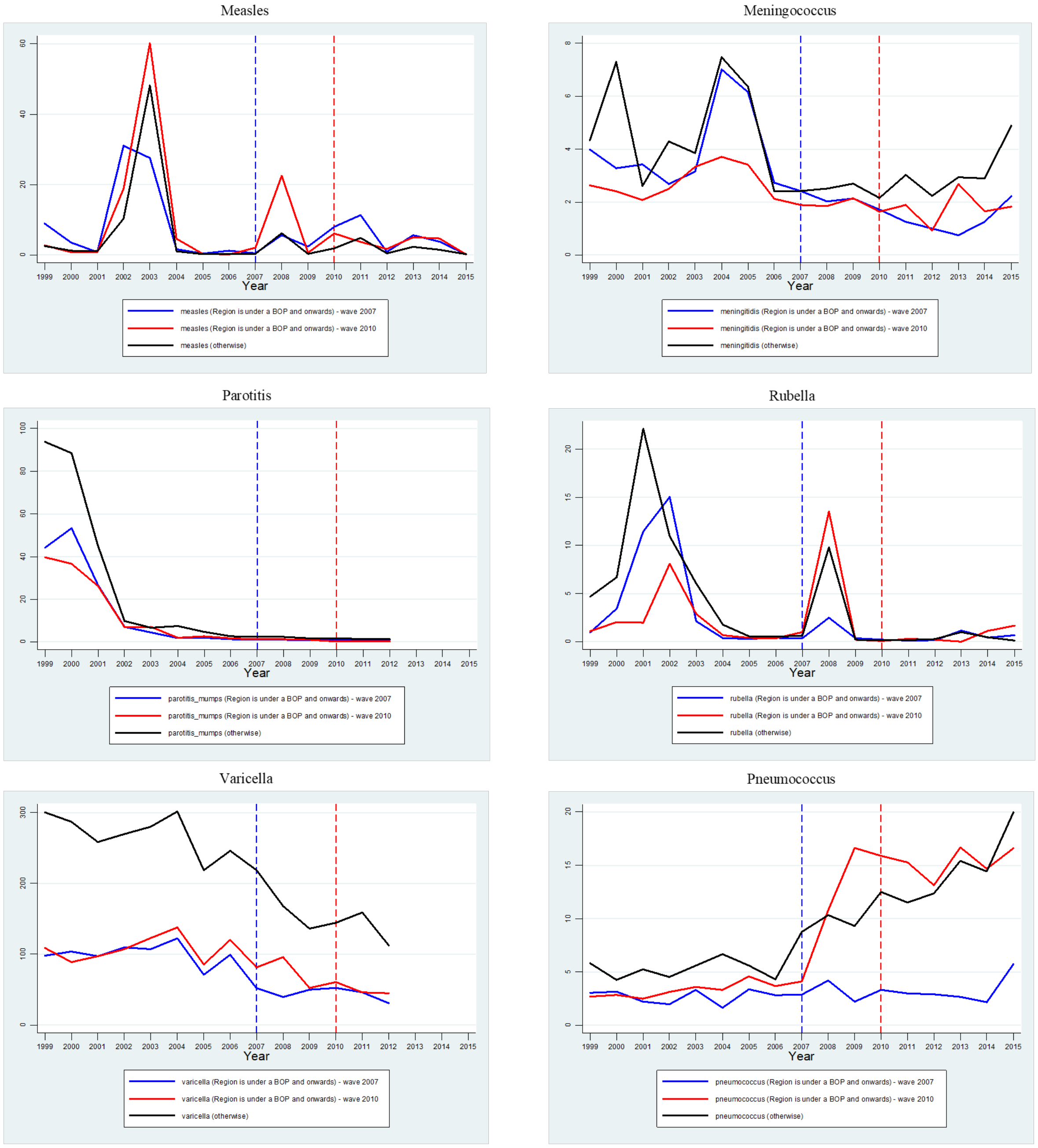
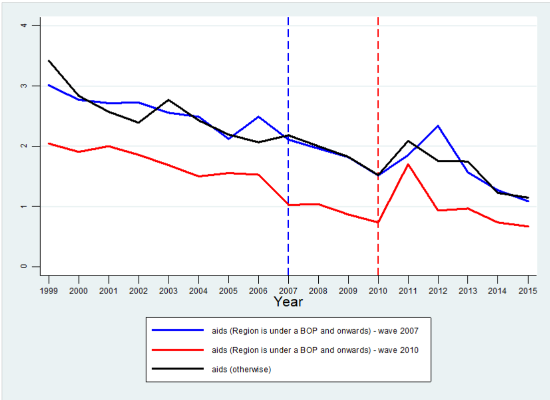


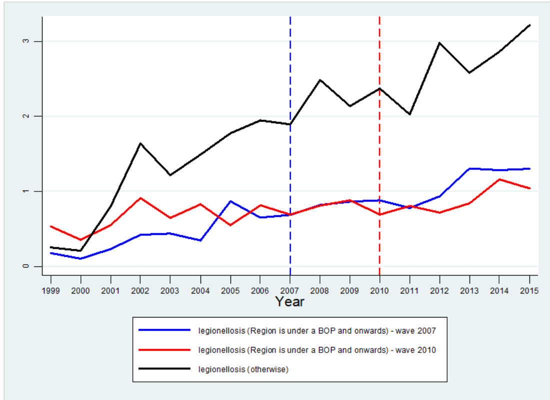
Figure 4: Incidence of Infectious diseases 1999-2015



Aids



Legionellosis



Tuberculosis

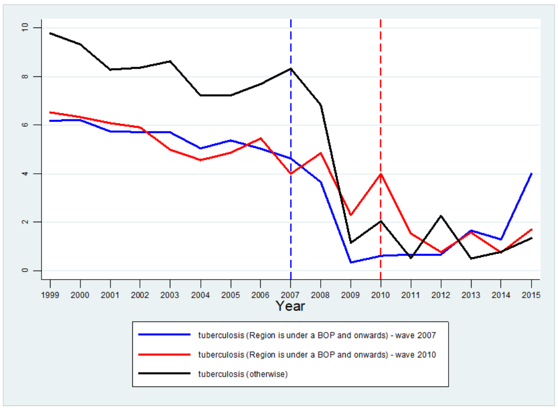
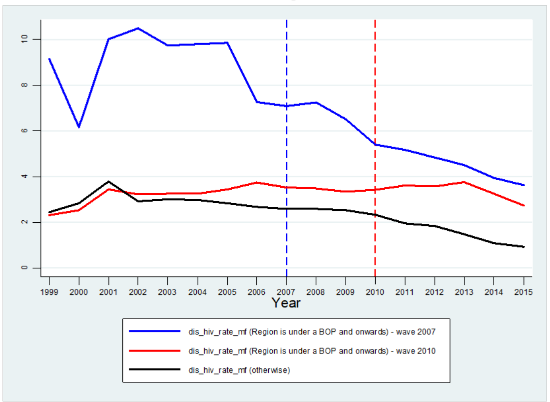
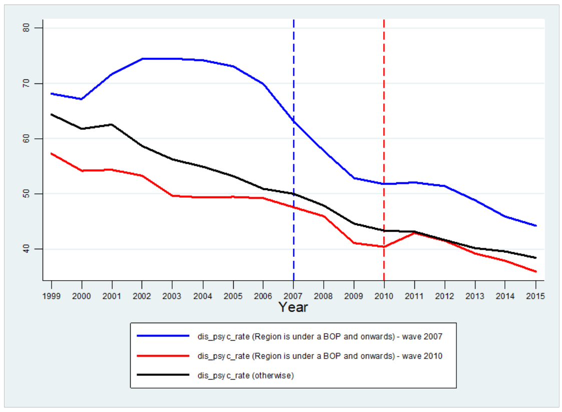


Figure 5: Discharges 1999-2015

HIV discharges rate



Psychological diseases discharges rate



6 Results

We present our results distinguishing among several groups of health dimension and discharges. In section 6.1 are represented results on total mortality rate, mortality rate due to different causes, to age and due to several diseases. Findings on the incidence on infectious diseases are presented in section 6.2. Morever, results on on discharges rate for HIV and psychological diseases are in section 6.3. Finally, we present robustness check on our results in section 6.4
In each section we show the second stage OLS regression estimates. Moreover, we present in additional tables, the estimate of our instrumental variable as provided by the first stage and several statistical tests to verified exclusion restriction, to detect on weakness of our instrument and weak under-identification.

6.1 Results on mortality rate

6.1.1 Results on mortality rate by causes

Table 6 shows the results of the second stage OLS regression for total mortality rate and several causes of mortality rate. In general, they increase after the introduction of a bail-out, but the effect is only significant for total mortality rate and suicides rate. The relationship between the total population and the total mortality rate is significant and negative, while it is positive but not significant for suicides rate and murders rate. Whilst if we analyze population by age classes, we notice that the total mortality rate increase but suicide rate and murders rate decreases in the regions with Bops. These results are in line with Boitsov and Samorodskaya (2016). Moreover, the average number of household components decrease all rates, and it is significant for the total mortality rate and the suicide rate in regions adopting a BOP (-0.494 and -1.521, respectively). We can interpreter these findings as a positive effect of social ties on several mortality rate dimensions. As expected, an increasing of employment-to-population ratio has a positive effect on mortality rate dimensions, especially on suicides rate and murders rate (respectively, 1.440 and 5.581 both significant at 1%). The sign of the coefficient for the unemployment rate is negative but not significant on both mortality rate and deaths due to traffic roads accidents. These findings are in line with Ruhm (2015). He considers the unemployment rate as a variation of economic recession, so the total mortality is "weakly related or unrelated to macroeconomic condition". Moreover, the lack of job reduces several bad habits and risky behaviours due to their economic weakness, people consume lower quantity alcoholic drinks, smokeless, reduce the utilization of car, decreasing accidents and invest their more leisure time in more physical activities and spend more time on socializing. However, an increase in the unemployment rate harms suicides rate and murders rate. The relationship between the percentage of people with tertiary education and all dimension of mortality rate is negative, except for the traffic road deaths rate. Whether we consider this variable as a proxy of income level, we can notice that the high-income level increase mortality rate. As expected, the level of poverty, as measured at the household level by the poverty index, has a negative effect on all dimensions of mortality rate except for the traffic road deaths rate. The utilization of car may be considered as an aspect of poverty dimension, so an increase in the poverty level reduces care utilization and consequently deaths due to traffic roads. In general, a rising of per capita GDP (at current prices) increase all mortality rates except for both deaths rate due to traffic roads and murders rate in regions with BOPs. Additionally, we present three statistical tests in order to verify the identification model and weakness of our instrument on the second stage. Both the Kleibergen–Paap rk LM statistic and the Hansen J statistics confirm that the model is identified in the second stage. The Kleibergen-Paap Wald rk F statistic confirms that the instrument is adequate to identify the equation.

In table 7, we present instrument result and several statistical tests in order to have a broad comprehension of our instrument. The positive and significant sign of *neg_ transp* shows that the regional inefficiency system increases the probability to adopt an operational program in the health sector (bail-out plans). We can consider the second stage as a standard OLS regression, while the first stage as a linear probability model (Cerulli, 2015). According to the rule of thumb, the value of the F-test statistic shows that our instrument is not weak. This result is confirmed by the value of Angrist-Prischke multivariate F test. (Both statistics are greater than 10). The model is well identified, as shown by both AP and Kleibergen-Paap rk LM statistics. Moreover, as the value of and Kleibergen-Paap rk F statistics is greater than the critical values obtained by Stock-Yogo cut-off (Table 1), we can reject the null hypothesis of weak model identification.

Table 6: Mortality rate by causes - Fixed effect results (second stage)

		Mortality rate by causes		
	Mortality rate	Suicides rate	Murders rate	Traffic road deaths rate
Variables	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.
Policy variable				
bop	0.040***	0.277***	0.329	0.010
	(0.015)	(0.085)	(0.207)	(0.087)
Controls				
Total population (in ln)	-0.700***	1.850	1.464	-0.077
	(0.227)	(1.719)	(6.107)	(2.106)
Population aged 15-64 (in ln)	0.285	-2.359	-5.840	0.095
	(0.240)	(1.682)	(6.048)	(1.945)
Population aged 85+ (% in ln)	0.231***	-0.555	-0.667	0.127
	(0.059)	(0.487)	(1.507)	(0.639)
Avg number of household components (in ln)	-0.494***	-1.521*	-1.377	-0.111
	(0.135)	(0.858)	(2.331)	(0.922)
Employment-to population ratio 15+ (in ln)	0.073	1.440***	5.581***	0.491
	(0.097)	(0.558)	(1.550)	(0.512)
Unemployment rate aged 15-24 (in ln)	-0.001	0.013	0.174	-0.036
	(0.010)	(0.078)	(0.202)	(0.074)
Tertiary education (% in ln)	0.010	0.183	0.897	-0.338
	(0.033)	(0.279)	(0.649)	(0.259)
Household poverty index (in ln)	0.008	0.018	0.001	-0.081*
	(0.007)	(0.064)	(0.136)	(0.049)
GDP per capita at current prices (in ln)	0.013	0.761	-1.619	-0.747
	(0.072)	(0.776)	(1.426)	(0.657)
Constant	10.290***	-5.889	52.771	7.466
	(1.881)	(18.636)	(38.696)	(14.359)
Regional fixed effect	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES
Observations	228	228	200	228
R-squared	0.979	0.823	0.697	0.903
Underidentification test				
K-P rk LM statistic Chi-sq(1)	34.018	34.018	25.136	34.018
P-val	0.0000	0.0000	0.0000	0.0000
Weak identification test				
K-P rk Wald F statistic	50.556	50.556	38.979	50.556
Overidentification test of all instruments				
Hansen J statistic	0.000	0.000	0.000	0.000

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P= Kleibergen-Paap

Table 7: First stage statistics tests on Mortality rate by causes

	Fisrt stage results on			
	Mortality rate	Suicides rate	Murders rate	Traffic road deaths rate
	bop	bop	bop	bop
Instrumental variable	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.
neg_transp	0.0217***	0.0217***	0.0227***	0.0217***
	(0.003)	(0.003)	(0.004)	(0.003)
F-test of excluded instruments				
F(1, 188)	50.56	50.56	38.98	50.56
P-val	0.0000	0.0000	0.0000	0.0000
Angrist-Pischke multivariate F test of excluded instruments				
AP F(1, 188)	50.56	50.56	38.98	50.56
P-val	0.0000	0.0000	0.0000	0.0000
Underidentification test				
AP Chi-sq(1)	61.31	61.31	48.42	61.31
P-val	0.0000	0.0000	0.0000	0.0000
K-P rk LM statistic				
Chi-sq(1)	34.02	34.02	25.14	34.02
P-val	0.0000	0.0000	0.0000	0.0000
Weak identification test				
K-P Wald rk F statistic	50.56	50.56	38.98	50.56
Nr. of observations	228	228	200	228
Nr. of regressors	40	40	39	40
Nr. of endogenous regressors	1	1	1	1
Nr. of instruments	40	40	39	40
Nr. of excluded instruments	1	1	1	1

Note: In the murders analysis the Anderson-Rubin Wald test is with F(1,161)
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P: Kleibergen-Paap

6.1.2 Results on mortality rate by age

Table 8 shows the results of second stage OLS regression for several reasons of mortality related to the age. In general, the adoption of bail-out increase the infant mortality rate, the stillbirths rate and the potential years of life lost, and consequently, we can see a reduction of life expectancy at birth. Demographic variables show heterogeneity finding on our dependent variable related to the age classes considered. For instance, the total population decreases the infant mortality rate (mainly due to the dimension of total population greater than infant population). However, if we consider only people aged between 15 and 64 years old, the effect is negative (namely the infant mortality increases) due to the size of this aged class population; conversely, the infant mortality decreases if we consider only the percentage of people aged more than 85 years old. The effect of demographic variables on our dependent variables is related to the size of population (or age classes) that we consider (in line with Boitsov and Samorodskaya, 2016). Interestingly, the social ties (as measured by the average number of household components) decrease both infant mortality and stillbirth rates and also the PYLL, while they increase the rate of life expectancy at birth in regions adopting a BOP. The effect of the employment-to-population ratio is significant for the potential years of life lost in regions with BOPs. The positive sign can be attributed to the quality of work life. As expected, the percentage of people with high education level decrease both infants and stillbirth rates (both signs are negative and significant at 1%). The household poverty index has a negative effect but not significant on all dependent variables. The per capita GDP has a positive effect on all dependent variables (expected of stillbirths rate); indeed an increase of per capita GDP reduces both infant mortality rate and PYLL and increase the rate of life expectancy at birth. The model is well identified, and the instrument is adequate for identification of the model in the second stage of the OLS regression, as confirmed by the statistical tests.

In table 9, we present instrument result and several statistical tests in order to have a broad comprehension of our instrument. Our instrument is positive and significant in the first stage of each equation model. Moreover, the value of the F-test statistic shows that our instrument is not weak. This result is confirmed by the value of Angrist-Prischke multivariate F test. (Both statistics are greater than 10). The model is well identified, as shown by both AP and

Kleibergen-Paap rk LM statistics. Moreover, as the value of and Kleibergen-Paap rk F statistics is greater than the critical values obtained by Stock-Yogo cut-off (Table 1), we can reject the null hypothesis of weak model identification. Our tests verify both exclusion restriction and identification model assumptions.

Table 8: Mortality rate by age - Fixed effect results (second stage)

Variables	Infant mortality rate Coeff.\Std. Err.	Still births rate Coeff.\Std. Err.	Life expectancy at birth Coeff.\Std. Err.	PYLL Coeff.\Std. Err.
Policy variable				
bop	0.088	0.289	-0.042	0.079***
	(0.172)	(0.228)	(0.119)	(0.029)
Controls				
Total population (in ln)	-3.592	7.503*	11.527***	-1.730***
	(3.066)	(3.953)	(2.222)	(0.485)
Population aged 15-64 (in ln)	1.886	-8.725*	-10.755***	1.238**
	(3.141)	(4.734)	(2.410)	(0.504)
Population aged 85+ (% in ln)	-0.217	-0.289	0.976	0.052
	(1.140)	(1.346)	(0.658)	(0.131)
Avg number of household components	-1.145	-2.627	2.010*	-0.928***
	(1.635)	(2.112)	(1.124)	(0.276)
Employment-to-population ratio 15+ (in ln)	1.701	-0.306	-0.781	0.459**
	(1.061)	(1.376)	(0.628)	(0.198)
Unemployed rate aged 15-24 (in ln)	0.012	-0.012	-0.023	0.021
	(0.135)	(0.187)	(0.090)	(0.021)
Tertiary education (% in ln)	-1.190*	-1.808*	0.138	-0.094
	(0.646)	(1.004)	(0.261)	(0.078)
Household poverty index (in ln)	0.059	0.129	0.081	-0.003
	(0.144)	(0.145)	(0.071)	(0.022)
GDP per capita at current prices (in ln)	-0.068	0.407	0.445	-0.091
	(0.852)	(1.379)	(0.602)	(0.185)
Constant	26.793	20.127	420.753***	12.972***
	(22.739)	(38.254)	(15.371)	(4.523)
Regional fixed effect	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES
Observations	228	228	228	228
R-squared	0.421	0.399	0.982	0.910
Underidentification test				
K-P rk LM statistic Chi-sq(1)	34.018	34.018	34.018	34.018
P-val	0.0000	0.0000	0.0000	0.0000
Weak identification test				
K-P rk Wald F statistic	50.556	50.556	50.556	50.556
Overidentification test of all instruments				
Hansen J statistic	0.000	0.000	0.000	0.000

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P= Kleibergen-Paap

Table 9: First stage statistics tests on Mortality rate by age

	Fisrt stage results on			
	Infant mortality rate	Still births rate	Life expectancy at birth	PYLL
	bop	bop	bop	bop
Instrumental variable	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.
neg_transp	0.0217***	0.0217***	0.0217***	0.0217***
	(0.003)	(0.003)	(0.003)	(0.003)
F-test of excluded instruments				
F(1, 188)	50.56	50.56	50.56	50.56
P-val	0.0000	0.0000	0.0000	0.0000
Angrist-Pischke multivariate F test of excluded instruments				
AP F(1, 188)	50.56	50.56	50.56	50.56
P-val	0.0000	0.0000	0.0000	0.0000
Underidentification test				
AP Chi-sq(1)	61.31	61.31	61.31	61.31
P-val	0.0000	0.0000	0.0000	0.0000
K-P rk LM statistic				
Chi-sq(1)	34.02	34.02	34.02	34.02
P-val	0.0000	0.0000	0.0000	0.0000
Weak identification test				
K-P Wald rk F statistic	50.56	50.56	50.56	50.56
Nr. of observations	228	228	228	228
Nr. of regressors	40	40	40	40
Nr. of endogenous regressors	1	1	1	1
Nr. of instruments	40	40	40	40
Nr. of excluded instruments	1	1	1	1

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star\, p < 0.05$, $\star p < 0.1$
K-P: Kleibergen-Paap

6.1.3 Results on mortality rate by diseases

The introduction of bail-out among regions has different effects on mortality rate related to the diseases analysed (Table 10). In particular, in regions with BOPs, the mortality rate increases among people affected by cancer and heart diseases (for both total population and people aged 65+ with heart diseases). The mortality rate due to other diseases decreases in region adopting a BOP even though the coefficient is not significant (except for people aged 65+ affected by pneumonia). The positive sign of mortality rate due to alcohol problems may indicate a decreasing wellbeing level among people living in the regions that undergo a bail-out plan. The effect of demographic variables on mortality rate is heterogeneous and depends on the size of age class and the diseases analysed. The social ties have a positive effect on the reduction of mortality rate due to cancer and heart diseases. In general, the effect of the employment-to-population rate is negative on the mortality rate due to different diseases, but interestingly its effect is positive on the mortality rate due to mental disorders. The effect of the other socio-economic variables shows that poor conditions increase overall the mortality rate due to different diseases (even though the sign is not significant). The model is well identified, and the instrument is adequate for identification of the model in the second stage of the OLS regression, as confirmed by the statistical tests.

As the findings provided in the previous analysis, in table 11, we summarize several statistical tests of the first stage tests in order to have a broad comprehension of the instrument in our model. Our instrument is positive and significant in the first stage of each equation model. Moreover, the value of the F-test statistic shows that our instrument is not weak. This result is confirmed by the value of Angrist-Prischke multivariate F test. (Both statistics are greater than 10). The model is well identified, as shown by both AP and Kleibergen-Paap rk LM statistics. Moreover, as the value of and Kleibergen-Paap rk F statistics is greater than the critical values obtained by Stock-Yogo cut-off (Table 1), we can reject the null hypothesis of weak model identification. Our tests verify both exclusion restriction and identification model assumptions.

Table 10: Mortality rate by diseases - Fixed effect results (second stage)

Variables	alcohol Coeff.\Std. Err.	cancer Coeff.\Std. Err.	cirrhosis Coeff.\Std. Err.	mental disorders Coeff.\Std. Err.	respiratory disease Coeff.\Std. Err.	heart diseases Coeff.\Std. Err.	heart diseases aged 65+ Coeff.\Std. Err.	pneumonia Coeff.\Std. Err.	pneumonia aged 65+ Coeff.\Std. Err.
Policy variable									
bop	0.316	0.065***	-0.033	-0.004	-0.023	0.114***	0.067**	-0.067	-0.127*
	(0.410)	(0.020)	(0.082)	(0.058)	(0.036)	(0.030)	(0.028)	(0.072)	(0.072)
Controls									
Total population (in ln)	19.824	-0.095	-4.287***	-0.118	-0.904	-3.425***	-5.642***	-0.779	-2.732*
	(12.094)	(0.377)	(0.974)	(1.095)	(0.675)	(0.573)	(0.609)	(1.557)	(1.649)
Population aged 15-64 (in ln)	-22.361**	-0.646	3.843***	-1.039	0.844	1.941***	4.617***	3.598**	5.908***
	(10.363)	(0.431)	(1.048)	(1.114)	(0.726)	(0.637)	(0.670)	(1.538)	(1.631)
Population aged 85+ (% in ln)	1.177	0.288***	0.044	1.142***	0.171	-0.036	-0.318**	0.357	0.012
	(2.673)	(0.088)	(0.298)	(0.300)	(0.167)	(0.148)	(0.158)	(0.459)	(0.460)
Avg number of household components	2.012	-0.789***	0.184	-0.813	0.200	-1.027***	-0.567**	0.399	0.834
	(4.783)	(0.184)	(0.791)	(0.570)	(0.355)	(0.309)	(0.287)	(0.695)	(0.687)
Employment-to-population ratio 15+ (in ln)	-2.261	-0.013	-0.265	-1.236***	0.577**	0.114	0.239	0.355	0.465
	(2.905)	(0.141)	(0.573)	(0.309)	(0.227)	(0.201)	(0.188)	(0.475)	(0.502)
Unemployed rate aged 15-24 (in ln)	0.438	-0.025	-0.004	-0.018	0.005	0.017	0.033	-0.077	-0.072
	(0.339)	(0.015)	(0.055)	(0.044)	(0.027)	(0.025)	(0.023)	(0.064)	(0.066)
Tertiary education (% in ln)	0.050	0.029	-0.062	0.248*	0.130	0.102	0.079	0.014	-0.042
	(0.911)	(0.052)	(0.204)	(0.131)	(0.092)	(0.087)	(0.082)	(0.207)	(0.216)
Household poverty index (in ln)	-0.324	0.008	-0.057*	-0.022	-0.002	-0.013	-0.026	-0.044	-0.048
	(0.283)	(0.014)	(0.031)	(0.030)	(0.029)	(0.019)	(0.018)	(0.049)	(0.056)
GDP per capita at current prices (in ln)	-0.403	0.027	0.140	-0.458	-0.449*	-0.287	-0.339*	-0.168	-0.334
	(3.444)	(0.120)	(0.356)	(0.312)	(0.238)	(0.183)	(0.181)	(0.483)	(0.513)
Constant	30.923	13.510***	7.740	25.058***	4.867	27.442***	23.641***	-38.003***	-39.082***
	(83.624)	(3.119)	(8.124)	(7.939)	(5.856)	(4.908)	(4.876)	(12.624)	(13.283)
Regional fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	228	228	228	228	228	228	228	228	228
R-squared	0.533	0.971	0.914	0.977	0.921	0.953	0.942	0.957	0.940
Underidentification test									
K-P rk LM statistic Chi-sq(1)	34.018	34.018	34.018	34.018	34.018	34.018	34.018	34.018	34.018
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Weak identification test									
K-P rk Wald F statistic	50.556	50.556	50.556	50.556	50.556	50.556	50.556	50.556	50.556
Overidentification test of all instruments									
Hansen J statistic	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P= Kleibergen-Paap

Table 11: First stage statistics tests on Mortality rate by diseases

	Fisrt stage results on								
	alcohol	cancer	cirrhosis	mental disorders	respiratory disease	heart diseases	heart diseases aged 65+	pneumonia	pneumonia aged 65+
	bop	bop	bop	bop	bop	bop	bop	bop	bop
Variables	Coeff.\ Std. Err.	Coeff.\ Std. Err.	Coeff.\ Std. Err.	Coeff.\ Std. Err.	Coeff.\ Std. Err.	Coeff.\ Std. Err.	Coeff.\ Std. Err.	Coeff.\ Std. Err.	Coeff.\ Std. Err.
neg_transp	0.0217***	0.0217***	0.0217***	0.0217***	0.0217***	0.0217***	0.0217***	0.0217***	0.0217***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
F-test of excluded instruments									
F(1, 188)	50.56	50.56	50.56	50.56	50.56	50.56	50.56	50.56	50.56
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Angrist-Pischke multivariate F test of excluded instruments									
AP F(1, 188)	50.56	50.56	50.56	50.56	50.56	50.56	50.56	50.56	50.56
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Underidentification test									
AP Chi-sq(1)	61.31	61.31	61.31	61.31	61.31	61.31	61.31	61.31	61.31
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
K-P rk LM statistic									
Chi-sq(1)	34.02	34.02	34.02	34.02	34.02	34.02	34.02	34.02	34.02
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Weak identification test									
K-P Wald rk F statistic	50.56	50.56	50.56	50.56	50.56	50.56	50.56	50.56	50.56
Nr. of observations	228	228	228	228	228	228	228	228	228
Nr. of regressors	40	40	40	40	40	40	40	40	40
Nr. of endogenous regressors	1	1	1	1	1	1	1	1	1
Nr. of instruments	40	40	40	40	40	40	40	40	40
Nr. of excluded instruments	1	1	1	1	1	1	1	1	1

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P: Kleibergen-Paap

6.2 Results on Infectious diseases (ISS)

The incidence of infectious diseases is the probability (or risk) to be affected by an infectious disease. As shown in Table 12, considering a broad set of different diseases, the risk to be affected by them is increased in regions adopting a bail-out plan. To have a clearer comprehension of this effect, we need to consider also that the immunisation was compulsory for some of these infectious diseases, and the mandatory was introduced in different years among children. For instance, the mandatory immunisation for measles, rubella and parotitis was introduced by 2001 for children between 1 and 6 years old, only for those born after 2000. Whilst the mandatory immunisation for meningococcus was introduced in 2012 for children of 2 years old, only for those born after 2011[12] . In particular, the incidence of measles and varicella increase in regions adopting a BOP, while the incidence of AIDS decreases in these regions. An increasing of the total population has a positive effect on the incidence of meningococcus, AIDS and TBC, while we have an opposite effect on these infectious diseases if we consider the population aged between 15-65 years old. As expected, we have a positive effect of the percentage of people aged 85+ on the incidence of some infectious diseases. It may depend on the nature of infectious diseases (children are mainly affected by measles, varicella, and pneumococcus). Also, the incidence of TBS is decreasing among the percentage of people aged 85+, as a consequence of social life reduction. The negative sign of the percentage of paediatricians shows that the healthcare services in regions with BOP, are mainly provided by the paediatricians for some infectious (such as measles, varicella, parotitis and pneumococcus). However, the positive and significant sign of coefficient in the incidence of rubella may reveal healthcare coverage unmet due to the larger number of patients (more than 800). The same consideration can be done analysing the effect of GPs with more than 1500 patients. The effect of the average number of household components is negative on the incidence of several infectious diseases mainly due to the high probability of transmission in large families. The negative sign of percentage of people with a low level of education may reveal a potential lack of interest of health conditions by these people rather than a potential positive effect of low

[12]Information are provided in the Ministry of Health website (`http://www.salute.gov.it/portale/vaccinazioni/dettaglioContenutiVaccinazioni.jsp?lingua=italiano&id=4829&area=vaccinazioni&menu=vuoto`)

education on incidence of infectious diseases in regions adopting a BOP. Same consideration ca be done analysing the effect of household poverty index on the incidence of infectious diseases. We can consider the positive sign of per capita GDP on the incidence of different infectious diseases in Regions with BOPs, as a consequence of increasing chances of transmissions due to growing social life. The model is well identified, and the instrument is adequate for identification of the model in the second stage of the OLS regression, as confirmed by the statistical tests.

Results in Table 13 are in line with previous findings. Specifically, the instrument is positive and significant in the first stage of each equation model. Moreover, the value of the F-test statistic shows that our instrument is not weak. This result is confirmed by the value of Angrist-Prischke multivariate F test. (Both statistics are greater than 10). The model is well identified, as shown by both AP and Kleibergen-Paap rk LM statistics. Moreover, as the value of and Kleibergen-Paap rk F statistics is greater than the critical values obtained by Stock-Yogo cut-off (Table 1), we can reject the null hypothesis of weak model identification. Our tests verify both exclusion restriction and identification model assumptions.

Table 12: Incidence of Infectious diseases - Fixed effect results (second stage)

Variables	Measles Coeff.\Std. Err.	Meningococcus Coeff.\Std. Err.	Parotitis Coeff.\Std. Err.	Rubella Coeff.\Std. Err.	Varicella Coeff.\Std. Err.	Pneumococcus Coeff.\Std. Err.	Aids Coeff.\Std. Err.	Legionellosis Coeff.\Std. Err.	Tuberculosis Coeff.\Std. Err.
Policy variable									
bop	1.233*	0.084	0.115	0.212	1.060**	0.020	-0.297*	-0.304	0.247
	(0.643)	(0.285)	(0.260)	(0.501)	(0.527)	(0.369)	(0.173)	(0.382)	(0.252)
Controls									
Total population (in ln)	16.139	-14.979**	-0.138	-18.126	11.384	12.087	-7.710*	-0.103	-37.599***
	(14.838)	(6.687)	(8.839)	(12.587)	(16.793)	(8.206)	(4.647)	(6.779)	(8.150)
Population aged 15-64 (in ln)	-16.602	15.642*	-10.596	10.498	-4.032	-4.513	3.912	2.665	35.295***
	(17.038)	(8.162)	(10.613)	(15.085)	(19.426)	(9.134)	(5.136)	(8.538)	(8.849)
Population aged 85+ (% in ln)	-10.256***	0.141	1.139	1.571	-10.478***	-5.047**	1.575	2.903	-4.370**
	(3.755)	(1.838)	(1.862)	(2.740)	(3.525)	(2.245)	(1.076)	(2.031)	(1.848)
Pediatricians with 800+ patients (% in ln)	-2.609**	0.127	-0.061	2.561**	-0.790	-0.986			
	(1.294)	(0.504)	(0.672)	(1.125)	(1.104)	(0.694)			
GP with 1500+ patients (% in ln)							0.064	-0.221	0.597**
							(0.149)	(0.231)	(0.303)
Avg number of household components	-13.268*	1.315	3.961	-9.900*	5.498	1.302	5.011**	-3.371	-3.012
	(6.895)	(3.418)	(3.611)	(5.858)	(7.769)	(3.985)	(2.026)	(3.821)	(3.831)
Primary/No education (% in ln)	-5.828*	-2.367**	-1.096	-1.909	-3.877*	-1.410	-1.921***	-2.614***	-0.495
	(3.289)	(1.150)	(1.102)	(1.579)	(2.066)	(1.256)	(0.664)	(1.007)	(0.927)
Household poverty index (in ln)	0.120	-0.579**	0.161	-0.333	-0.287	-0.071	0.083	0.057	0.339
	(0.485)	(0.277)	(0.195)	(0.381)	(0.424)	(0.237)	(0.141)	(0.190)	(0.210)
GDP per capita at current prices (in ln)	0.400	1.688	-1.600	1.856	0.900	-0.159	2.679**	-0.189	-2.663
	(4.401)	(1.813)	(2.689)	(3.346)	(5.923)	(2.342)	(1.354)	(2.043)	(1.851)
Constant	48.166	-12.397	160.278**	97.164	-89.028	-92.467	27.727	-24.853	81.392
	(123.586)	(55.864)	(70.925)	(95.093)	(119.830)	(74.037)	(37.999)	(58.185)	(53.104)
Regional fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	208	209	190	208	190	209	209	209	209
R-squared	0.622	0.648	0.807	0.690	0.629	0.780	0.712	0.765	0.760
Underidentification test									
K-P rk LM statistic Chi-sq(1)	40.918	40.826	40.320	40.918	40.320	40.826	41.697	41.697	41.697
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Weak identification test									
K-P rk Wald F statistic	96.459	95.681	90.495	96.459	90.495	95.681	97.169	97.169	97.169
Overidentification test of all instruments									
Hansen J statistic	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note:

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

K-P= Kleibergen-Paap

Table 13: First stage statistics tests on Incidence of Infectious diseases

	Fisrt stage results on								
	Measles	Meningococcus	Parotitis	Rubella	Varicella	Pneumococcus	Aids	Legionellosis	Tuberculosis
	bop	bop	bop	bop	bop	bop	bop	bop	bop
Variables	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.
neg_transp	0.0218***	0.0218***	0.0218***	0.0218***	0.0218***	0.0218***	0.0220***	0.0220***	0.0220***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
F-test of excluded instruments									
F(1, 188)	96.46	95.68	90.49	96.46	90.49	95.68	97.17	97.17	97.17
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Angrist-Pischke multivariate F test of excluded instruments									
AP F(1, 188)	96.46	95.68	90.49	96.46	90.49	95.68	97.17	97.17	97.17
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Underidentification test									
AP Chi-sq(1)	118.02	116.94	112.38	118.02	112.38	116.94	118.76	118.76	118.76
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
K-P rk LM statistic									
Chi-sq(1)	40.92	40.83	40.32	40.92	40.83	40.83	41.70	41.70	41.70
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Weak identification test									
K-P Wald rk F statistic	96.46	95.68	90.49	96.46	90.49	95.68	97.17	97.17	97.17
Nr. of observations	208	209	190	208	190	209	209	209	209
Nr. of regressors	38	38	37	38	37	38	38	38	38
Nr. of endogenous regressors	1	1	1	1	1	1	1	1	1
Nr. of instruments	38	38	37	38	37	38	38	38	38
Nr. of excluded instruments	1	1	1	1	1	1	1	1	1

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P: Kleibergen-Paap

6.3 Results on healthcare services coverage

Our analyses focus on the impact of austerity measures adopting in several regions on several health dimensions of the population, in terms of psychological and physical conditions. By using both HIV and psychological diseases discharges rate, we would like to capture the effect of austerity measures on healthcare coverage considering people with high social distress. In line with the literature, countries adopting measures of austerity decrease the healthcare provisions with negative effects on population health ((Karanikolos et al., 2013; Kentikelenis et al., 2014). Findings in Table 14 show two potential consequences in the regions adopting bail-out plans. Regions cut healthcare service for people affected by HIV (this result is in line with the effect of bail-out introduction on the incidence of AIDS) but in the meantime, the positive sign of bail-out plans on psychological diseases discharges rate may depend on the increasing number of people with mental disorders. In general, our findings show that the adoption of austerity measure reduce healthcare services but increase social distress. The total population has an opposite effect among two types of discharges rate. Considering young people (aged 0-14), we can see that the sign is negative on HIV discharges due to a low HIV transmission while the sign is positive on psychological diseases. This result may reveal an increase in social distress among young people in regions with BOPs. An increasing of the percentage of elderly population increase both discharge rates. In general, an increase in hospital inputs has a positive effect on both discharges. The index of household poverty has a positive effect on HIV discharge, maybe because poor people have a low level of trust in HIV medical care. In regions adopting a BOP, the per capita GDP has a positive effect on discharges rate due to psychological diseases. The model is well identified, and the instrument is adequate for identification of the model in the second stage of the OLS regression, as confirmed by the statistical tests.

Results in Table 15 are in line with previous findings. Specifically, the instrument is positive and significant in the first stage of each equation model. Moreover, the value of the F-test statistic shows that our instrument is not weak. This result is confirmed by the value of Angrist-Prischke multivariate F test. (Both statistics are greater than 10). The model is well identified, as shown by both AP and Kleibergen-Paap rk LM statistics. Moreover, as the value of and Kleibergen-Paap rk F statistics is greater than the critical values obtained by Stock-

Yogo cut-off (Table 1), we can reject the null hypothesis of weak model identification. Our tests verify both exclusion restriction and identification model assumptions.

Table 14: Service covarage: Discharge rate - Fixed effect results (second stage)

Variables	HIV discharge rate Coeff.\Std. Err.	Psycological diseases discharge rate Coeff.\Std. Err.
Policy variable		
bop	-0.591***	0.126**
	(0.219)	(0.054)
Controls		
Total population (in ln)	5.407	-8.001***
	(5.194)	(1.419)
Population aged 0-14 (in ln)	-4.163*	2.944***
	(2.150)	(0.602)
Population aged 85+ (% in ln)	2.157	1.987***
	(1.729)	(0.428)
Acute ordinary pubblic beds (rate in ln)	-0.320	0.236**
	(0.378)	(0.118)
Doctors rate (in ln)	0.410	-0.013
	(0.693)	(0.172)
Nurses rate (in ln)	0.111	0.541***
	(0.717)	(0.156)
Primary/No education (% in ln)	0.939	-0.203
	(0.663)	(0.136)
Household poverty index (in ln)	-0.344**	0.040
	(0.138)	(0.037)
GDP per capita at current prices (in ln)	0.479	-1.723***
	(1.592)	(0.371)
Constant	-34.727	93.683***
	(51.721)	(13.710)
Regional fixed effect	YES	YES
Year fixed effect	YES	YES
Observations	228	228
R-squared	0.881	0.930
Underidentification test		
K-P rk LM statistic Chi-sq(1)	31.377	31.377
P-val	0.0000	0.0000
Weak identification test		
K-P rk Wald F statistic	39.368	39.368
Overidentification test of all instruments		
Hansen J statistic	0.000	0.000

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P= Kleibergen-Paap

Table 15: First stage statistics tests on Service covarage: Discharge rate

	Fisrt stage results on	
	HIV discharge rate	Psycological diseases discharge rate
	bop	bop
Variables	Coeff.\Std. Err.	Coeff.\Std. Err.
neg_transp	0.0173***	0.0173***
	(0.003)	(0.003)
F-test of excluded instruments		
F(1, 188)	39.37	39.37
P-val	0.0000	0.0000
Angrist-Pischke multivariate F test of excluded instruments		
AP F(1, 188)	39.37	39.37
P-val	0.0000	0.0000
Underidentification test		
AP Chi-sq(1)	47.74	47.74
P-val	0.0000	0.0000
K-P rk LM statistic		
Chi-sq(1)	31.38	31.38
P-val	0.0000	0.0000
Weak identification test		
K-P Wald rk F statistic	39.37	39.37
Nr. of observations	228	228
Nr. of regressors	40	40
Nr. of endogenous regressors	1	1
Nr. of instruments	40	40
Nr. of excluded instruments	1	1

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P: Kleibergen-Paap

6.4 Robustness check

In this section, we briefly present our robustness check on the previous results. We consider only 18 regions instead of 21, dropping from our analysis Aosta Valley, Trento and Bolzan). The value of their operating results should address them into a bail-out plan, but they did not undergo it. We investigate the robustness of our findings or if they depend on the number of regions without a bail-out plan. In Tables 16-20, we present only the second stage OLS estimations, confirming that the exclusion restrictions and identification model assumptions are satisfied in the first stage[13]. We discuss here only the effect of bail-out on several measures of health and discharges. As shown in Table 16, the effect of BOP on mortality rate and several its causes is always positive and significant for the total mortality rate and suicide rate. It is completely in line with findings considering 21 regions (Table 6) Comparing results on mortality rate by age between regions with 18 (Table 17) and 21 (Table 8), the effect of BOP is identical. The effect of bail-out introduction on mortality rate by several diseases is the same for all diseases in both samples with 18 and 21, except on mental disorders. The introduction of bail-out reduces the mortality rate due to mental disorders if we consider 21 regions but increases it if we analyse the effect on 18 regions, but in both, the coefficient is not significant. (Comparison of Table 18 and Table 10). Moreover, the effect of BOP on the incidence of several diseases considering 18 regions (Table 19), confirms results were provided with 21 regions (Table 12), except for meningococcus and pneumococcus. In the analysis with 18 regions the adoption of BOP reduces the incidence of these infectious diseases even though the coefficient is not significant. Finally, the effect on both discharges rate is identical if we consider the sample composed by 21 (Table 14) and 18 regions (Table 20). We can conclude that our results are robust. Moreover, to check additionally about the robustness of our findings, we replicate the same model estimations for each dependent variable considering Trento and Bolzano as one unique region[14].

[13]Results are provided upon request

[14]We do not present results here, but we can provide upon request.

Table 16: Mortality rate by causes on 18 Regions - Fixed effect results (second stage)

		Mortality rate by causes		
	Mortality rate	Suicides rate	Murders rate	Traffic road deaths rate
Variables	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.	Coeff.\Std. Err.
Policy variable				
bop	0.016*	0.101*	0.271	0.086
	(0.010)	(0.061)	(0.201)	(0.066)
Controls				
Total population (in ln)	-0.601***	1.085	0.590	0.153
	(0.186)	(1.453)	(6.112)	(1.935)
Population aged 15-64 (in ln)	0.282	-1.467	-5.159	-0.123
	(0.205)	(1.414)	(6.036)	(1.811)
Population aged 85+ (% in ln)	0.302***	-0.046	-0.552	-0.628
	(0.051)	(0.403)	(1.506)	(0.488)
Avg number of household components (in ln)	-0.302***	0.035	-1.180	-0.867
	(0.096)	(0.633)	(2.291)	(0.784)
Employment-to population ratio 15+ (in ln)	-0.078	0.882*	5.510***	0.342
	(0.058)	(0.478)	(1.584)	(0.457)
Unemployed rate aged 15-24 (in ln)	-0.002	0.016	0.198	-0.100
	(0.008)	(0.064)	(0.204)	(0.065)
Tertiary education (% in ln)	0.054**	-0.065	0.774	-0.172
	(0.022)	(0.202)	(0.672)	(0.190)
Household poverty index (in ln)	0.005	0.016	0.042	-0.028
	(0.006)	(0.059)	(0.147)	(0.043)
GDP per capita at current prices (in ln)	0.088*	1.071	-1.815	-1.233**
	(0.051)	(0.704)	(1.447)	(0.569)
Constant	8.413***	-9.609	57.848	13.851
	(1.335)	(16.455)	(38.939)	(12.918)
Regional fixed effect	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES
Observations	216	216	195	216
R-squared	0.989	0.859	0.699	0.929
Underidentification test				
K-P rk LM statistic Chi-sq(1)	30.819	30.819	25.136	30.819
P-val	0.0000	0.0000	0.0000	0.0000
Weak identification test				
K-P rk Wald F statistic	47.756	47.756	41.097	47.756
Overidentification test of all instruments				
Hansen J statistic	0.000	0.000	0.000	0.000

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P= Kleibergen-Paap

Table 17: Mortality rate by age on 18 Regions - Fixed effect results (second stage)

Variables	Infant mortality rate Coeff.\Std. Err.	Still births rate Coeff.\Std. Err.	Life expectancy at birth Coeff.\Std. Err.	PYLL Coeff.\Std. Err.
Policy variable				
bop	0.067	0.146	-0.102	0.041**
	(0.128)	(0.183)	(0.123)	(0.018)
Controls				
Total population (in ln)	-1.769	8.486**	11.217***	-1.557***
	(2.756)	(3.594)	(2.205)	(0.452)
Population aged 15-64 (in ln)	0.117	-9.491**	-10.789***	1.258***
	(2.848)	(4.428)	(2.402)	(0.447)
Population aged 85+ (% in ln)	-0.260	0.083	1.360**	0.110
	(0.975)	(1.218)	(0.685)	(0.100)
Avg number of household components	-0.308	-1.271	2.863**	-0.554***
	(1.359)	(1.807)	(1.148)	(0.199)
Employment-to-population ratio 15+ (in ln)	0.677	-1.038	-0.591	0.189
	(0.920)	(1.155)	(0.684)	(0.151)
Unemployed rate aged 15-24 (in ln)	-0.006	-0.026	-0.008	0.007
	(0.118)	(0.178)	(0.092)	(0.017)
Tertiary education (% in ln)	-0.902	-1.163*	-0.176	-0.041
	(0.636)	(0.692)	(0.273)	(0.061)
Household poverty index (in ln)	-0.028	0.032	0.120	-0.013
	(0.078)	(0.131)	(0.076)	(0.019)
GDP per capita at current prices (in ln)	-0.215	-0.199	0.476	0.027
	(0.761)	(1.120)	(0.623)	(0.180)
Constant	29.393	22.533	424.029***	9.645**
	(20.906)	(33.178)	(15.706)	(4.357)
Regiona fixed effect	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES
Observations	216	216	216	216
R-squared	0.464	0.466	0.982	0.937
Underidentification test				
K-P rk LM statistic Chi-sq(1)	30.819	30.819	30.819	30.819
P-val	0.0000	0.0000	0.0000	0.0000
Weak identification test				
K-P rk Wald F statistic	47.756	47.756	47.756	47.756
Overidentification test of all instruments				
Hansen J statistic	0.000	0.000	0.000	0.000

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P= Kleibergen-Paap

Table 18: Mortality rate by diseases on 18 Regions - Fixed effect results (second stage)

Variables	alcohol Coeff.\Std. Err.	cancer Coeff.\Std. Err.	cirrhosis Coeff.\Std. Err.	mental disorders Coeff.\Std. Err.	respiratory disease Coeff.\Std. Err.	heart diseases Coeff.\Std. Err.	heart diseases aged 65+ Coeff.\Std. Err.	pneumonia Coeff.\Std. Err.	pneumonia aged 65+ Coeff.\Std. Err.
Policy variable									
bop	0.799**	0.050***	-0.052	0.067	-0.063**	0.061***	0.015	-0.092	-0.148**
	(0.404)	(0.014)	(0.037)	(0.052)	(0.029)	(0.022)	(0.022)	(0.058)	(0.063)
Controls									
Total population (in ln)	18.027	-0.052	-4.127***	-0.098	-0.720	-3.409***	-5.682***	0.212	-1.804
	(12.163)	(0.338)	(0.881)	(1.115)	(0.624)	(0.541)	(0.599)	(1.438)	(1.556)
Population aged 15-64 (in ln)	-21.700**	-0.615	3.918***	-0.980	0.852	1.963***	4.696***	2.725**	5.077***
	(10.504)	(0.393)	(0.917)	(1.115)	(0.690)	(0.628)	(0.688)	(1.376)	(1.507)
Population aged 85+ (% in ln)	-1.444	0.314***	-0.001	0.768***	0.370***	0.146	-0.119	0.707**	0.398
	(2.771)	(0.079)	(0.217)	(0.295)	(0.141)	(0.123)	(0.138)	(0.337)	(0.371)
Avg number of household components	-3.422	-0.642***	0.028	-1.496***	0.657**	-0.567**	-0.117	0.946	1.360**
	(5.110)	(0.140)	(0.439)	(0.544)	(0.304)	(0.253)	(0.256)	(0.621)	(0.637)
Employment-to-population ratio 15+ (in ln)	-0.296	-0.064	-0.438	-1.087***	0.358*	-0.119	0.046	0.101	0.249
	(3.087)	(0.099)	(0.280)	(0.326)	(0.190)	(0.148)	(0.155)	(0.421)	(0.470)
Unemployed rate aged 15-24 (in ln)	0.369	-0.033**	-0.031	-0.039	0.006	0.020	0.038*	-0.064	-0.048
	(0.343)	(0.013)	(0.041)	(0.045)	(0.028)	(0.020)	(0.020)	(0.056)	(0.060)
Tertiary education (% in ln)	-0.016	0.028	-0.006	0.272**	0.125	0.106	0.075	0.136	0.062
	(1.024)	(0.042)	(0.125)	(0.130)	(0.077)	(0.080)	(0.078)	(0.170)	(0.191)
Household poverty index (in ln)	-0.066	0.023**	-0.029	0.013	-0.031	-0.018	-0.035**	-0.061	-0.083*
	(0.290)	(0.011)	(0.029)	(0.027)	(0.020)	(0.018)	(0.018)	(0.046)	(0.050)
GDP per capita at current prices (in ln)	-2.385	0.090	0.179	-0.442	-0.199	-0.225	-0.262	0.114	-0.037
	(3.396)	(0.106)	(0.325)	(0.337)	(0.171)	(0.177)	(0.183)	(0.446)	(0.473)
Constant	67.347	11.891***	4.824	24.215***	-0.087	26.552***	22.454***	-43.137***	-44.169***
	(83.064)	(2.713)	(6.937)	(8.228)	(4.574)	(4.641)	(4.914)	(11.797)	(12.448)
Regiona fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	216	216	216	216	216	216	216	216	216
R-squared	0.433	0.981	0.951	0.978	0.939	0.966	0.950	0.970	0.954
Underidentification test									
K-P rk LM statistic Chi-sq(1)	30.819	30.819	30.819	30.819	30.819	30.819	30.819	30.819	30.819
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Weak identification test									
K-P rk Wald F statistic	47.756	47.756	47.756	47.756	47.756	47.756	47.756	47.756	47.756
Overidentification test of all instruments									
Hansen J statistic	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P= Kleibergen-Paap

Table 19: Incidence of Infectious diseases on 18 Regions - Fixed effect results (second stage)

Variables	Measles Coeff.\Std. Err.	Meningococcus Coeff.\Std. Err.	Parotitis Coeff.\Std. Err.	Rubella Coeff.\Std. Err.	Varicella Coeff.\Std. Err.	Pneumococcus Coeff.\Std. Err.	Aids Coeff.\Std. Err.	Legionellosis Coeff.\Std. Err.	Tuberculosis Coeff.\Std. Err.
Policy variable									
bop	0.632	-0.025	0.303	0.402	0.541	-0.087	-0.140	-0.509	0.203
	(0.602)	(0.274)	(0.271)	(0.517)	(0.512)	(0.322)	(0.153)	(0.409)	(0.249)
Controls									
Total population (in ln)	18.095	-15.465**	-0.851	-17.485	13.101	9.934	-6.316	2.361	-38.679***
	(14.103)	(6.751)	(9.233)	(12.875)	(16.942)	(8.228)	(4.637)	(6.681)	(8.023)
Population aged 15-64 (in ln)	-17.943	16.534**	-10.040	9.785	-5.571	-3.549	2.294	0.825	36.503***
	(16.340)	(8.427)	(11.049)	(15.239)	(19.605)	(9.115)	(4.969)	(8.331)	(8.691)
Population aged 85+ (% in ln)	-7.941**	0.304	0.510	0.981	-8.748**	-5.117**	1.156	3.964*	-4.379**
	(3.665)	(1.819)	(2.001)	(2.861)	(3.561)	(2.180)	(1.044)	(2.086)	(1.859)
Pediatricians with 800+ patients (% in ln)	-2.563*	-0.225	-0.016	2.513*	-1.091	-0.687			
	(1.523)	(0.544)	(0.824)	(1.337)	(1.217)	(0.712)			
GP with 1500+ patients (% in ln)							0.021	-0.235	0.742**
							(0.159)	(0.250)	(0.324)
Avg number of household components	-6.519	1.944	2.041	-13.164**	11.246	2.863	3.548*	-0.730	-3.735
	(6.992)	(3.367)	(3.986)	(6.346)	(7.799)	(3.518)	(2.042)	(4.181)	(3.993)
Primary/No education (% in ln)	-4.787	-2.381**	-1.361	-2.303	-3.088	-1.108	-2.072***	-2.267**	-0.130
	(3.199)	(1.159)	(1.193)	(1.658)	(2.008)	(1.157)	(0.654)	(1.017)	(0.922)
Household poverty index (in ln)	-0.069	-0.337	0.199	-0.326	-0.515	-0.015	0.069	0.031	0.467**
	(0.529)	(0.239)	(0.226)	(0.416)	(0.435)	(0.246)	(0.152)	(0.221)	(0.230)
GDP per capita at current prices (in ln)	1.622	2.213	-2.174	0.995	2.640	-1.538	2.136	0.915	-2.687
	(4.455)	(1.761)	(3.099)	(3.672)	(6.241)	(2.295)	(1.351)	(2.119)	(1.927)
Constant	15.225	-22.844	171.415**	111.471	-117.276	-65.057	37.980	-50.005	79.126
	(122.268)	(55.273)	(79.069)	(100.006)	(120.417)	(72.800)	(38.292)	(59.754)	(55.414)
Regional fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	197	198	180	197	180	198	198	198	198
R-squared	0.630	0.665	0.796	0.682	0.662	0.806	0.741	0.761	0.766
Underidentification test									
K-P rk LM statistic Chi-sq(1)	36.959	36.779	37.300	36.959	37.300	36.779	36.879	36.879	36.879
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Weak identification test									
K-P rk Wald F statistic	96.648	95.540	93.483	93.483	93.483	95.540	94.133	94.133	94.133
Overidentification test of all instruments									
Hansen J statistic	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P= Kleibergen-Paap

Table 20: Service covarage: Discharge rate on 18 Regions - Fixed effect results (second stage)

Variables	HIV discharge rate Coeff.\Std. Err.	Psycological diseases discharge rate Coeff.\Std. Err.
Policy variable		
bop	-0.422**	0.068
	(0.168)	(0.045)
Controls		
Total population (in ln)	1.530	-6.688***
	(4.334)	(1.273)
Population aged 0-14 (in ln)	-2.897	2.463***
	(1.766)	(0.533)
Population aged 85+ (% in ln)	2.376	1.802***
	(1.623)	(0.408)
Acute ordinary pubblic beds (rate in ln)	-0.364	0.253**
	(0.353)	(0.111)
Doctors rate (in ln)	-0.218	0.181
	(0.611)	(0.178)
Nurses rate (in ln)	0.725	0.361**
	(0.624)	(0.159)
Primary/No education (% in ln)	0.940	-0.151
	(0.600)	(0.138)
Household poverty index (in ln)	-0.261*	0.011
	(0.137)	(0.037)
GDP per capita at current prices (in ln)	-0.653	-1.661***
	(1.401)	(0.357)
Constant	14.984	80.625***
	(43.205)	(12.453)
Regional fixed effect	YES	YES
Year fixed effect	YES	YES
Observations	216	216
R-squared	0.899	0.941
Underidentification test		
K-P rk LM statistic Chi-sq(1)	35.536	35.536
P-val	0.0000	0.0000
Weak identification test		
K-P rk Wald F statistic	47.844	47.844
Overidentification test of all instruments		
Hansen J statistic	0.000	0.000

Note:

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

K-P= Kleibergen-Paap

7 Conclusions

In this paper, we seek to provide evidence on the effect of mandatory bail-out plans, adopted by some Italians regions, on several measures of health status and two measures of discharges rate during the years 1999-2015. The bail-out plans are operational programs imposed by the Central Governments to restore regional budget balances, trough reorganizational and cost cuts of regional healthcare service but guaranteeing the LEA. Following the literature (Karanikolos et al., 2013; Kentikelenis et al., 2014) on the effect of austerity measures on the different health dimensions (in terms of physical and psychological diseases), we analyze the effect of regional bail-out plans adoption on a broad set of health status measures. Also, we focus on the impact of austerity policies on social distress, considering several dimensions of physical and psychological human diseases.

We adopt an IV strategy to handle with potential endogeneity of bail-out plan's adoption. As only some regions with a high level of deficit were imposed to adopt an operational recovery program, we state that some unobservable regional features beforehand influenced the selection on the bail-out plan. As seven regions adopted a bail-out plan in 2007 and three regions in 2010, we use the average percentage of people with low satisfaction on public transport in 2005 as a proxy of overall regional system inefficiency before the introduction of BOPs. This instrument is correlated with the adoption of bail-out plans but uncorrelated with each health status dimensions analysed. In general, our findings show a negative impact of bail-out plans on health status dimensions. These results are in line with previous studies that analyzed both the effect of economic downturns on health outcomes (Atella et al., 2018) and the impact of bail-out plans on citizens' well-being (Piacenza and Turati, 2014) or also, on mortality rate and hospitalization rate (Depalo, 2019). In particular, collecting data on several measures of mortality rate, the incidence of infectious diseases and discharges rates, we could capture both the physical and psychological dimension of health status. Thus, our findings show that the adoption of bail-out plans increase not only the total mortality rate but also the suicide, murders, and traffic road deaths rates. Whether the total mortality rate can be considered as a proxy of health status, the suicides, murders and traffic road deaths rates allow understanding the dimension of social distress of people during the economic downturn. Moreover, we investigate the effect of BOPs on the mortality rate related to age. Our findings show that the infant mortality and stillbirth rates increased in regions with bail-out plans and the life expectancy at birth decrease. In order to capture the effect on the potential life duration, we build the potential years' life lost indicator (PYLL). In the regions with bail-out plans, the PYLL increases significantly. In general, results show an increase in mortality rate due to several diseases in the regions with bail-out plans, especially for people affected by chronic diseases (such as cancer and heart attack). These findings may reveal that basic health needs are unmet.

Additionally, our analyses focus on the incidence of infectious diseases in the regions with bail-out plans. As the infectious diseases are a major cause of human suffering in terms of both morbidity and mortality, we investigate on the effect of reorganizational and costs cuts of healthcare services on health status in regions with BOPs, considering the probability (or risk) of infectious diseases. Overall results show an increasing incidence of infectious diseases, especially for measles and varicella. These findings may reveal as costs-cutting on healthcare services harm the immunization of people living in the regions adopting a BOPs.

Finally, as our main aim is the effect of austerity measures on health status by focusing on social distress, we use the discharges rate as a measure of healthcare coverage but considering only discharges of more vulnerable people and with affected by psychological disorders. Thus, our findings show the potential twofold effect of BOPs adoption. First, a reduction of HIV discharges, that may reveal a reduction of hospital resources for people affected by HIV; and an increase of psychological diseases discharges rate that leads to believe that austerity measures increase social distress.

In conclusion, by considering a broad set of physical and phycological health measures. We provide additional evidence with respect to the previous studies on the effect of bail-out plans in the Italian regions. We seek to capture the effect of BOPs on health status, focus on several dimensions of social distress. Moreover, by using an IV strategy, we address the bail-out plans endogeneity due to possible unobservable regional system features, that led some region to recover a bail-out plan and not others. In this perspective, we may also contribute to the literature on regional healthcare sectors disparities.

Appendix A Data descriptions and sources

Table 21: Data sources and descriptions

Variable name	Description	Source	Code in the original datawarehouse	Period of data availability
Dependent variables				
Mortality rate by causes				
Mortality rate	Mortality rate M+F	Health for ALL (HFA)	0270	1990-2015
Suicides rate	Mortality rate - suicides M+F	Health for ALL (HFA)	1940 T	1990-2015
Murders rate	Murders rate (per 100000 inab.)	built as murders/pop* 100000 inhab.		2003-2017
Traffic road	Traffic road deaths	I.stat Datawarehouse Datawarehouse - (morti in		2001-2018
deaths rate	rate M+F	incidenti stradali - rispetto al totale degli incidenti (on 100 inhab.)		2001-2018
Mortality rate by age				
Infant mortality rate	Infant Mortality rate M+F	Health for ALL (HFA)	0300	1990-2014
Still births rate	Still birth - Neonatal mortality rate 1-29 days M+F	Health for ALL (HFA)	0330	1990-2014
Life expectancy	Life expectancy rate -	Health for ALL	6090 and 6100	
at birth	aged 0+ (M+F)	(HFA)		1980-2017
Potential years of life lost (PYLL)	Potential years of life lost (PYLL)	built following OECD guidalines.		
Mortality by diseases				
alcohol	Mortality rate - alcohol M+F	Health for ALL (HFA)	5660	1990-2015
cancer	Mortality rate - cancer M+F	Health for ALL (HFA)	1090	1990-2015
cirrhosis	Mortality rate - cirrhosis M+F	Health for ALL (HFA)	1720 M+F	
mental disorders	Mortality rate - mental disorders M+F	Health for ALL (HFA)	1450	1990-2015
respiratory disease	Mortality rate - respiratory disease M+F	Health for ALL (HFA)	1600	1990-2015
heart diseases	Mortality rate - hearth diseases M+F	Health for ALL (HFA)	1540	1990-2015
heart diseases aged 65+	Mortality rate - heart diseases aged 65+ M+F	Health for ALL (HFA)	1543	1990-2015
pneumonia	Mortality rate - pneumonia M+F	Health for ALL (HFA)	1630	1990-2015
pneumonia aged 65+	Mortality rate - pneumonia aged 65+ M+F	Health for ALL (HFA)	1633	1990-2015
Incidence of infectious	Measles, Meningococcus, Parotitis, Rubella, Varicella,	Italian Society of Hygiene, Preventive Medicine and Public Health, SITI;		1999-2015
diseases	Pneumococcus, Aids, Legionellosis, Tuberculosis	Italian National Institute of Health, ISS and Ministry of Health		
Discharges rate				
HIV discharges rate	HIV discharges M+F (rate)	Health for ALL (HFA)	8060	1999-2016
Psycological diseases discharges rate	Psycological diseases discharges M+F (rate)	Health for ALL (HFA)	5150	1999-2016
Instrumental varibale				
bad_transp	avg % of people aged 14+ with good judgement of public transport	I.stat Datawarehouse - Aspetti della vita quotidiana	2001-2018	
Controls				
Total population	Total population M+F	Health for ALL (HFA)	0000	1982-2017
Population aged 0-14	Population aged 0-14	Health for ALL (HFA)	0001;0002	1982-2017
Population aged 15-64	Population aged 15-64	Health for ALL (HFA)	0003-0007	1982-2017
Population aged 85+	People aged 85+ (%)	Health for ALL (HFA)	0046	1982-2017
Acute ordinary pubblic beds	Acute ordinary pubblic beds (rate)	Health for ALL (HFA)	7224	1996-2015
Doctors rate	Doctors (rate)	Health for ALL (HFA)	9111	1994-2013
Nurses rate	Nurses (rate)	Health for ALL (HFA)	9112	1994-2013
Pediatricians with 800+ patients	Pediatricians with 800+ patients (%)	Health for ALL (HFA)	7019	1994-2013
General practitioners with 1500+ patients	General practitioners with 1500+ patients (%)	Health for ALL (HFA)	7014	1994-2013
Employment-to-population ratio 15+	Employment-to-population ratio 15+ M+F	Health for ALL (HFA)	0410	1993-2017
Unemployed rate aged 15-24	Unemployed rate aged 15-24 M+F	Health for ALL (HFA)	0455	1993-2017
Primary/No education	People witn primary/No education M+F (%)	Health for ALL (HFA)	0351	1997-2017
Tertiary education	People with tertiary education M+F (%)	Health for ALL (HFA)	0354	1997-2017
Avg number of household components	Average number of household components	Health for ALL (HFA)	0150	1994-2003,2005-2017
Household poverty index	Household poverty index	Health for ALL (HFA)	0481	2002-2017
GDP per capita at current prices	Per capita GDP at current prices (in € mln)	I.stat Datawarehouse		1995-2017

References

Abadie A., J. Angrist and Imbens G. (2002). "Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings". In: *Econometrica* 70.1, pp. 91–117. DOI: 10.1111/1468-0262.00270.

Aimone Gigio L., Alampi D. et al. (2018). *The Italian healthcare system: the difficult balance between budget sustainability and the quality of services in Regions subject to Financial Recovery Plan.* Questioni di Economia e Finanza (Occasional Papers) 427. Bank of Italy, Economic Research and International Relations Area. URL: https://ideas.repec.org/p/bdi/opques/qef_427_18.html.

Angrist J., G. Imbens and Rubin D. (1996). "Identification of Causal Effects Using Instrumental Variables". In: *Journal of the American Statistical Association* 91.434, pp. 444–455. DOI: 10.2307/2291629.

Angrist, J. and Krueger A. (1991). "Does Compulsory School Attendance Affect Schooling and Earnings?" In: *The Quarterly Journal of Economics* 106.4, pp. 979–1014.

Angrist, J. and Imbens G. (1995). "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity". In: *Journal of the American Statistical Association* 90.430, pp. 431–442. DOI: 10.1080/01621459.1995.10476535.

Appleby J., JK. Helderman and Gregory S. (2015). "The global financial crisis, health and health care". In: *Health Econ Policy Law* 10.1, pp. 1–6. DOI: doi:10.1017/S1744133114000206.

Arie, Sophie (2013). "Has austerity brought Europe to the brink of a health disaster?" In: *BMJ* 346. DOI: 10.1136/bmj.f3773. eprint: https://www.bmj.com/content/346/bmj.f3773.full.pdf. URL: https://www.bmj.com/content/346/bmj.f3773.

Atella V., Belotti F. et al. (2019). "How health policy shapes healthcare sector productivity? Evidence from Italy and UK". In: *Health Policy* 123.1, pp. 27 –36. ISSN: 0168-8510. DOI: https://doi.org/10.1016/j.healthpol.2018.10.016.

Atella, V. et al. (2018). *Economic Crisis, Mortality and Health Status. A New Perspective.* CEIS Research Paper 425. Tor Vergata University, CEIS.

Bentivogli C., R. Cullino and Del Colle D.M. (Sept. 2008). *Regulation and efficiency in Italian local public transport: the regional differences.* Questioni di Economia e Finanza (Occasional Papers) 20. Bank of Italy, Economic Research and International Relations Area. URL: https://ideas.repec.org/p/bdi/opques/qef_20_08.html.

Boitsov, S.A. and Samorodskaya I.V. (2016). "Determinants of mortality". In: *Her. Russ. Acad. Sci.* 86, 473–480. DOI: https://doi.org/10.1134/S1019331616060010.

Bordignon, M. and Turati G. (2009). "Bailing out expectations and public health expenditure". In: *Journal of Health Economics* 28.2, pp. 305 –321. ISSN: 0167-6296. DOI: https://doi.org/10.1016/j.jhealeco.2008.12.008.

Cerulli, G. (2015). *Econometric Evaluation of Socio-Economic Programs. Theory and Applications.* 1st ed. Springer-Verlag Berlin Heidelberg.

Cicchetti, A. and Gasbarrini A. (2016). "The healthcare service in Italy: regional variability". In: *Eur Rev Med Pharmacol Sci* 20.1 Suppl, pp. 1–3.

Commission, European (2019). *Transport in the European Union. Current Trends and Issues.* European Commission, Mobility and Transport.

Depalo, D. (2019). "The side effects on health of a recovery plan in Italy: A nonparametric bounding approach". In: *Regional Science and Urban Economics* 78.C. DOI: 10.1016/j.regsciurbeco.20. URL: https://ideas.repec.org/a/eee/regeco/v78y2019ics0166046218304836.html.

Duggan, M.G. (Nov. 2000). "Hospital Ownership and Public Medical Spending". In: *The Quarterly Journal of Economics* 115.4, pp. 1343–1373. ISSN: 0033-5533. DOI: 10.1162/003355300555097.

Ferré F., de Belvis AG Longhi S. Lazzari A. et al. (2014). "Italy Health system review". In: *Health Systems in Transition* 16.4. URL: https://www.euro.who.int/en/about-us/partners/observatory/publications/health-system-reviews-hits/full-list-of-country-hits/italy-hit-2014.

Horwitz, J. R. and Austin N. (2009). "Hospital ownership and medical services: Market mix, spillover effects, and nonprofit objectives". In: *Journal of Health Economics* 28.5, pp. 924–937.

J., Angrist and Pischke J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* 1st ed. Princeton University Press.

Karanikolos M., Mladovsky P., Cylus J., and et al. (2013). "Financial crisis, austerity, and health in Europe". In: *Lancet* 381.9874, pp. 1323–1331. DOI: 10.1016/S0140-6736(13)60102-6.

Kentikelenis A., Karanikolos M., McKee A., and Stuckler D. (2014). "Greece's health crisis: from austerity to denialism". In: *Lancet* 383.9918, pp. 748–753. DOI: 10.1016/S0140-6736(13)62291-6.

Kentikelenis A., Karanikolos M., Williams G., and et al. (2015). "How do economic crises affect migrants' risk of infectious disease? A systematic-narrative review". In: *Eur J Public Health* 25.6, pp. 937–944. DOI: doi:10.1093/eurpub/ckv151.

Manski, C. (1990). "Nonparametric Bounds on Treatment Effects". In: *The American Economic Review* 80.2, pp. 319–323.

McKee M., Karanikolos M. P. Belcher and Stuckler D. (2012). "Austerity: a failed experiment on the people of Europe". In: *Clin Med (Lond)* 12.4, pp. 346–350. DOI: doi:10.7861/clinmedicine.12-4-346.

Neri, S. (Sept. 2019). "Economic Crisis, Decentralisation and Health Inequalities: The Case of Italy". In: DOI: 10.5772/intechopen.89006.

Novi, Cinzia Di et al. (July 2019). *Does fiscal decentralization affect regional disparities in health? Quasi-experimental evidence from Italy.* DISCE - Working Papers del Dipartimento di Economia e Finanza def083. Università Cattolica del Sacro Cuore, Dipartimenti e Istituti di Scienze Economiche (DISCE). URL: https://ideas.repec.org/p/ctc/serie1/def083.html.

Nuti S., Vola F. A. Bonini and Vainieri M. (2016). "Making governance work in the health care sector: evidence from a "natural experiment" in Italy". In: *Health Econ Policy Law* 11.1, pp. 17–38. DOI: 10.1017/S1744133115000067.

Paes-Sousa R., JMA Schramm and Mendes LVP (2019). "Fiscal austerity and the health sector: the cost of adjustments". In: *Ciênc Saúde Colet* 24, pp. 4375–84.

Piacenza, M. and Turati G. (2014). "Does Fiscal Discipline Towards Subnational Governments Affect Citizens' Well-Being? Evidence On Health". In: *Health Economics* 23.2, pp. 199–224. DOI: 10.1002/hec.2910.

Quaglio G., Karapiperis T. L. Van Woensel, Arnold E., and McDaid D. (2013). "Austerity and health in Europe". In: *Health Economics* 113.1-2, pp. 13–19. DOI: 10.1016/j.healthpol.2013.09.005.

Ruhm, C. J. (May 2000). "Are Recessions Good for Your Health?*". In: *The Quarterly Journal of Economics* 115.2, pp. 617–650. ISSN: 0033-5533. DOI: 10.1162/003355300554872. URL: https://doi.org/10.1162/003355300554872.

— (2015). "Recessions, healthy no more?" In: *Journal of Health Economics* 42, pp. 17 –28. ISSN: 0167-6296. DOI: https://doi.org/10.1016/j.jhealeco.2015.03.004.

Saltman R., V. Bankauskaite and Vrangbabaek K. (2007). *Decentralization in health care: Strategies and outcomes.* WHO European Observatory.

Sargan, J. (1958). "The Estimation of Economic Relationships using Instrumental Variables". In: *Econometrica* 26, pp. 393–415. ISSN: 0167-6296. DOI: doi:10.2307/1907619.

SiVeAS (2014). "DG della Programmazione Sanitaria. Ministero della Salute". In.

Suguiy, T., M.F. Carvalho, and D.L. Nithack e Silva (2013). "Efficiency of Urban Public Transport: Case Study of Brazilian Cities". In: *IFAC Proceedings Volumes* 46.24. 6th IFAC Conference on Management and Control of Production and Logistics, pp. 379 –384. ISSN: 1474-6670. DOI: https://doi.org/10.3182/20130911-3-BR-3021.00019.

Suhrcke, M. et al. (June 2011). "The Impact of Economic Crises on Communicable Disease Transmission and Control: A Systematic Review of the Evidence". In: *PLOS ONE* 6.6, pp. 1–12. DOI: 10.1371/journal.pone.0020724. URL: https://doi.org/10.1371/journal.pone.0020724.

Toth, F. (2014). "How health care regionalisation in Italy is widening the North–South gap". In: *Health Economics, Policy and Law* 9.3, 231–249. DOI: 10.1017/S1744133114000012.

UN-DESA (2012). "Family-Oriented Policies for Poverty Reduction, Work-Family Balance and Intergenerational Solidarity". In.

Wagstaff, A. (1989). "Econometric studies in health economics: A survey of the British literature". In: *Journal of Health Economics* 8.1, pp. 1–51. URL: https://EconPapers.repec.org/RePEc:eee:jhecon:v:8:y:1989:i:1:p:1-51.

WHO (2018). *Global Reference List of 100 Core Health Indicators, 2015.* World Health Organization.

Does exemption increase diagnostic cares? A causal approach for the Italian care system*

Vanessa Cirulli
Sapienza University of Rome
vanessa.cirulli@gmail.com

December 10, 2020

Abstract

The purpose of this paper is to analyze the effect of co-payment exemption on diagnostic care utilization. Increased utilization of healthcare can be driven either by health needs or by opportunistic behaviour. We address co-payment endogeneity by adopting one of the most advanced techniques available in econometrics to sidestep the difficulty in finding valid instrument, the inference on intersection bounds as developed by Chernozhukov et al. (2013, 2015) that allows identifying the identify the parameter of interest, providing an estimate of an interval within which it lays, rather than point identifying it. Findings show potential opportunistic behaviour in the utilization of diagnostic care.

Keywords: intersection bounds, diagnostic care, co-payment exemption.

1 Introduction

In Italy, healthcare services are guaranteed by universal principle and equitable access and utilization of healthcare services. People with the same health needs should have equal access to health care services (principle of horizontal equity on health). In this context, equity in healthcare utilization can be considered an element of achieving equity in health. However, since the 90s, the regional devolution processes led to establishing different Regional Health Systems (RHS) to restore regional budget balance and quality improvement in the healthcare sector. Each region acquired the making decision power to establish their co-payment scheme

*The views expressed are those of the authors and do not necessarily reflect those of the institutions to which they belong.

in order to raise regional revenues. Hence, the co-payment, meant as a financial contribution for healthcare services and goods consumed, is an instrument to contain healthcare costs, and in the meantime, it is a lever to enhance consumers' responsibility to demand only the optimal quantities of healthcare goods and services. However, regionals healthcare expenditure increased over time and during the 2000s, worsened by European economic crisis (2007-2008), so several healthcare containment costs were applied in Regions, especially in Regions with a high level of the budget deficit. The utilization of the healthcare services was potentially limited by introducing additional fees on co-payment that differs in each region. In this context, the most relevant effects of healthcare co-payments were among people with low economic resources (Atella, 2015 and 2014). Unmet needs for medical care and examinations increased, coupled with inequities in the utilization of specialist care. In particular, diagnostic care utilization was strongly penalized (Atella e al. 2015). Thus, the possibility (opportunity) in the utilization of healthcare services can be limited. However, the right of co-payment exemption for individuals with specific characteristic (age, presence of chronic diseases, low income, invalidity status and pregnancy) should guarantee the utilization and access to healthcare services among fragile population. Considering the co-payment exemption as a sort of healthcare insurance, a broad literature (starting from Arrow, 1963 until nowadays) investigate the impact of health insurance on healthcare expenditure, showing healthcare services' overconsumption among insured people ("moral hazard"). That implies that healthcare services' price is sensitive to the healthcare demand, and insured people may assume opportunistic behaviours.

Several studies analyzed this relationship. For instance, Manning et al., 1987 showed that health care demand is sensitive to the price (with an elasticity of about –0.2) and that individuals' health does not improve when healthcare services increase. Fireman and Swain (1996) found a reduction of 15% in the use of the emergency department after introducing a co-payment among some employees, considering as control group employees for whom the co-payment did not increase. Moreover, Tamblyn et al. (2001) found an increase of hospitalization for the elderly after introducing cost-sharing. Atella and Kopinska (2014) analyzed the causal effect between the introduction of co-payment on drug compliance, finding that in the presence of inefficient healthcare provision, co-payments are harmful to drug compliance, and this is especially true for patients who are originally good compliers. In Italy, the most relevant co-payment effect is the risk of impoverishment among fragile people (Costa et al. 2012). In particular, additional fees on cost-sharing were introduced on specialistic visits diagnostic cares and drugs prescriptions. Thus, in our work, we seek to evaluate the potential effect of co-payment exemption on diagnostic care utilization. As diagnostic cares are healthcare services, help on diseases prevention, and general health status controls for claimed people with diseases, this healthcare utilization analysis may be relevant to have a broad comprehension of the basic need unmet. Moreover, we investigate the effect of co-payment exemption on diagnostic care utilizations to understand if co-payment exemption is a lever for overcoming the increasing costs of healthcare services in real people in needs or there are some kind of opportunistic behaviours in people with exemption demanding diagnostic care.

The estimation of co-payment exemption on medical and healthcare services utilization is challenging since individuals' co-payment exemption is not randomly assigned. People can obtain the "right" of co-payment exemption according to specific features, such as people aged 65+, the presence of chronic diseases, low income, invalidity status or pregnancy. Moreover, observable and unobservable characteristics may also affect the demand for health services, influencing people health outcomes. For instance, Ponzo and Scoppa, 2016 applied a Regression Discontinuity Design, based on age, to access the exemption from cost-sharing in the Italian National Health System to estimate the effects of prescription drug consumption, specialist visits and diagnostic checks cost-sharing. We address potential co-payment endogeneity by adopting an instrumental variable approach by using the Regional Global Competitiveness Index as a proxy of the regional administrative and bureaucracy slowness. Regional healthcare services are provided by the Regional Health System (RHS) that differs in each region. Differences in healthcare delivery and infrastructures (including personnel and information technology) among regions may depend on regional systems' latent differences that can affect healthcare services utilization (also in the presence of co-payment exemption). Moreover, the co-payment exemption is a "right" conferred subsequently to a medical examination and administrative applications' submission. For that reason, we suppose that the slowness of regional administrative and bureaucracy can affect the "right" of co-payment exemption.

However, our estimates show that the instrument is weak, as confirmed by the joint statistical independence estimation. Specifically, a valid instrument needs to satisfy Late Independence (LI) and LATE monotonicity (LM). In other words, the instrument has to be independent of all potential outcomes and potential treatments; in the meantime, it must have any effect on the observed outcome beyond its effect on the observed treatment (LI) and also, a valid instrument variable has to ensure that the IV estimate identifies the average treatment effect, but it cannot be achieved in the presence of defiers (Imbens and Angrist, 1994) Thus, in line with Mourifié and Wan (2017), we adopt an intersection bounds strategy as developed by Chernozhukov, et al. (2013, 2015. We identify different confidence interval's widths related to the average number of diagnostic care utilization among people with co-payment exemption, conditioning to the average level of Regional Global Competitiveness Index with mean equal to zero.

Our estimates are based on the number of diagnostic care utilization by considering both the whole and female population. The distinction between the two populations is a consequence of several gender disparities due to differences in health and social conditions (Ministry of Health, 2016). At this work stage, we cannot identify the exact effect of co-payment exemption on the number of diagnostic cares utilization, but we may be able to better understand the phenomenon dimension. Moreover, in order to acquire more relevant information on the possible co-payment exemption effects among different individual characteristics, we built eight different groups related to the scarce self-perceived household economic resources, chronic disease (cancer), and geographical areas (north and south). These preliminary findings show potential opportunistic behaviour. Indeed, we use the scarce self-perceived household economic resources as a proxy of income (income information is not available in the dataset) and verify whether, given the same health disease, we can have some differences in the average number of diagnostic cares utilization. The increasing size (width) of confidence intervals in people with low economic resources may confirm our potential opportunistic behaviour hypothesis.

The remainder of this article is organized as follows. We introduce a brief institutional context to present co-payment exemption typologies, in Section 2. In Section 3.1, we explain our first assumption and concerns on Instrumental Variables approach, whilst in Section 3.2, we present the intersection bounds estimation methodology. Data sources and variables definition are in Section 4. In Section 5, we present our results by adopting the CRL strategy. Our conclusions are in Section 6.

2 Institutional framework

2.1 Copayment

The Italian National Health System aims to guarantee equality in healthcare access and utilization and establish the co-payment exemption for specific characteristics. It is a sort of insurance for Health. The co-payment exemption is based on the combination of age characteristics, income declared, and health status. We can distinguish between total and partial co-payment exemption. The principle of the requirement for the total co-payment exemption is mainly based on income and age. Thus, we have that children belonging to a household with total income less than €36151.98 can require the co-payment exemption. The same income threshold is applied to people with age more than 65 years old. Also, unemployed people with a specific household income threshold or individual income can ask for the co-payment exemption. Moreover, foreign people with special political asylum status and the residence permits can apply for it. Here are also included, people in several categories of invalidity. The partial co-payment exemption is based on the principle of illness. People affected by chronic diseases can obtain a co-payment exemption only for healthcare services related to their diseases. Pregnant women are fully guaranteed healthcare services for free charges during the period of pregnancy.

The SSN guarantees several screening services to prevent diseases, such as mammography for women aged 50 and over, pap smear, colorectal screening. People can participate in these programs prevention care for free of charges. The co-payment exemption is established at the central level (Central Government), whilst each region can autonomously fix the price for healthcare services, distinguishing between drugs prescription, specialistic visits and diagnostic cares. Thus, the analysis of different regional co-payment schemes and co-payment exemption can help on a broader understanding of geographical co-payment effects. Different studies reveal several consequences of the different regional fees' application on healthcare services, with an increasing gap in the access and utilization of health services between northern and southern Regions (Neri, 2019, Toth, 2014).

3 Methodology

3.1 Preliminary Assumption

Our analysis seeks to estimate the effect of co-payment exemption on diagnostic cares utilization, considering the whole and women population. We estimate the causal effect in a non-experiment (or quasi-experimental, as we deal with survey data) design considering the problem on the selection on unobservable variables. In the presence of co-payment exemption, the utilization of healthcare services, in our case diagnostic care, may increase more than individuals would have to pay the full price ("overconsumption"). As we know from health literature, the empirical estimation of the co-payment exemption (or also insurance) on healthcare service is a challenging issue, since the co-payment exemption is not randomly assigned. The healthcare co-payment exemption is a right guaranteed by the Italian Health System (SSN). However, the request for this right is left to the individual, subsequent the assessment of specific health condition diseases, as well as socio-economic conditions. Regional healthcare services are provided by the Regional Health System (RHS) that differs in each region. Differences in healthcare delivery but also infrastructures (including personnel and information technology) among regions may depend on latent unobservable differences in the regional systems that can affect the utilization of healthcare services also in the presence of co-payment exemption.

To isolate the actual effect of co-payment exemption on diagnostic care utilization, and hence, overcome the selection bias into treatment (selection on unobservable variables), we adopt the Instrumental Variable (IV) approach by using as instrument, the Regional Global Competitiveness Index. We consider it as a proxy of the regional administrative and bureaucracy slowness.

Thus, we estimate a consistent ATE by adopting a direct two-stage least squares (Direct-2SLS). It is a simultaneous system of two OLS regressions where in first stage (equation 1), we calculate predicted values of the endogenous variable D_i assuming that z_i (our instrumental variable) is directly correlated with the selection process (D_i) and captures the effect of z_i (Regional Global Competitveness Index, RGCI) on D_i, adjusting for covariates X_i.

$$D_i = \eta + \delta z_i + \sum_{k=1}^{14} \beta_1 X_{ki} + \nu_i \tag{1}$$

Then, we use these predictions as regressor in a second OLS regression to estimate the outcome Y_i (equation 2 expressed in reduced form) assuming that $(X_i,\ z_i)$ are uncorrelated with the error term ϵ_i, thus $(X_i,\ z_i)$ are exogenous and ϵ_i is is still correlated with D_i

$$Y_i = \beta_0 + \beta_1 D_i + \sum_{k=1}^{14} \beta_2 X_{ki} + \epsilon_i \tag{2}$$

where Y_i represent our number of total diagnost care utilization for both total population and female population for each individual i; D_i is is a dummy variable taking value equal to 1 if the individual have the co-payment exemption (*exemption*); X_{ki} individual socio-economic and geographical characteristics affecting the number of total diagnostic care utilization. ϵ_{it} is the error term.

In line with econometrics literature Sargan, 1958; Angrist and Krueger, 1991; Abadie et al., 2002; Angrist and Imbens, 1995; Angrist and Pischke, 2009; Angrist et al., 1996), an instrument is a random variable that is independent of certain unobserved latent terms, thereby facilitating the identification of the causal effect of an endogenous treatment on a particular outcome, by invoking different types of statistical independence. We consider the following potential outcome

$$Y = Y1(D) + Y0(1 - D) \tag{3}$$

where Y is the observed outcome, the number of diagnostic care underwent, $D = (0, 1)$ is the treatment, i.e. co-payment exemption for each individual; whilst the Y1(D) is the number of diagnostic care undergone by people with co-payment exemption, Y0(D) is the number of diagnostic care undergone by people without co-payment exemption).

We need to assume exclusion restriction, namely, the instrumental variable z is directly correlated with the selection process (D) but uncorrelated with the outcome Y in order to produce consistent estimation of average treatment effect (ATE). Moreover the instrument variable should be statistically independent of the vector of potential outcomes and treatments $Z \perp \{Y1, Y0, (Dz : z \in Z)\}$ (random assignment). Also, we need to verify that the instrumental variable should be statistically independent of the vector of potential outcomes only, and does not assume independence, or existence, of counterfactuals for D ($Z \perp (Y1, Y0)$, called joint statistical independence. However, Imbens and Angrist (1994)) state that a valid instrument need to be independent of all potential outcomes and potential treatments; in the meantime, it must have any effect on the observed outcome beyond its effect on the observed treatment (Late independence). When the treatment effect is heterogeneous, the valid instrument variable does not ensure that the IV estimate identifies the average treatment effect (ATE), so we need to assume the LATE monotonicity (LM, also known as the "no defiers" assumption), namely the instrument affects the treatment decision in the same direction for every individual. If both assumptions hold, the IV estimates could identify the ATE for the subpopulation of compliers, namely, the LATE. Although the results of Imbens and Angrist (1994) have been widely influential in the applied economics literature, there are still concerns about the validity of the key assumptions. Mourifié and Wan (2017) solve the sharp characterization of LI and LM by using a set of conditional moment inequalities. The feature of this conditional moment inequality representation is that the outcome variable goes into the inequalities as a conditioning variable. Thus, we can easily incorporate additional covariates into the moment inequalities as additional conditioning variables. They test the implications of both LI and LM assumptions by using the intersection bounds framework of Chernozhukov,et al. (2013, 2015).

As our IV approach reveals a potential weak instrument and weak instrument joint inference we adopt the intersection bounds strategy (CRL) as developed by Chernozhukov,et al. (2013, 2015).

3.2 Methodolgy on estimation by intersection bounds

We adopt the intersection bounds strategy as developed by Chernozhukov, et al. (2013, 2015), to identify a possible causal effect of co-payment exemption on diagnostic care utilization. So, we observe for each individual, the "right" of co-payment exemption (treatment effect), the standardized level of regional administrative and bureaucracy slowness (RGCI) with mean 0 and variance 1; and the number of diagnostic care underwent by individuals (for all and only women).

We estimate and make statistical inference on parameters restricted by using intersection bounds. Thus, we consider the parameter of interest (or true parameter value) $\theta^{\star}$ lies within the bounds $\theta^{l}(z)$ and $\theta^{u}(z)$

$$\theta^{l}(z) \leq \theta^{\star} \geq \theta^{u}(z) \tag{4}$$

Following Manski and Pepper (2000) we assume that $E[Y_i(D_i)|Z_i = z]$ is the conditional expectation of Y_i (number of diagnostic care utilization) is an increasing function of the regional bureaucracy and administrative slowness Z_i (*Monotone IV assumption*) and rises in the presence of co-payment exemption $(D)_i$ (*Monotone treatment response*). As our Instrumental Variable (IV) approach reveals weak instrument, weak instrument joint inference, violation of monotonicity condition (or presence o defiers), we cannot estimate a point identification of the average number of diagnostic care utilization. However, we are able to estimate the lower and upper bounds in which the average number of diagnostic care can be utilized by people with the co-payment exemption and living in a region with RGCI with media equal to zero.

Moreover, we know the "outcome space" of analysis, built as the total number of diagnostic care utilization subsequent to at least one specialistic visit. As in Italy, the general practitioners or specialist doctors may require diagnostic checks; we need to consider this sub-sample. Anymore assumption is made regarding the distribution of counterfactual outcomes. Then, we condition on average of RGCI ($Z_i = 0$). We built lower and upper bounds on the expected number of total diagnostic care at a given level of exemption (D_i) conditional on RGCI (Z_i) as follow:

$$LB_{u \leq z} E[Y_i^l | Z_i = u] \leq E[Y_i(D_i)|Z_i = z] \leq UB_{s \geq z} E[Y_i^u | Z_i = s] \tag{5}$$

where $\theta^{l}(z) = E[Y_i^l | Z_i = u]$ and $\theta^{u}(z) = E[Y_i^u | Z_i = s]$

Thus, we can find a set of bounds for different intervals of confidence (50%, 90%, 95% and 99%) of possibile increasing of average number of diagnostic care utilization under the null hypotesis H0: $\theta^{\star} = E[Y_i(D_i)|Z_i = z]$ at $z = 0$.

To understand if co-payment exemption can be a lever for overcoming the increasing costs of healthcare services in people with real needs or there are some opportunistic behaviours in people with co-payment exemption demanding diagnostic care, we built eight different population groups. These groups are based on the self-perceived household economic resources, distinguished in low and high economic resources; people affected by cancer, and geographical area (people living in the North or South area). For each group, we estimate intervals confidence at 90%, 95% and 90%. Their size can show if the average number of diagnostic care utilization increase among people with co-payment exemption and living in a region with an RGCI with mean equal to zero.

4 Data

4.1 Data Sources

We collect cross-sectional data from the Italian National Institute of Statistics Multipurpose Survey "Health conditions and use of health services", conducted by the Italian National Institute of Statistics in 2011-2012 (ISTAT, 2013). This survey is conducted every five years to evaluate the prevalence of chronic health conditions and healthcare services in the Italian population. In 2011-2012, 119,073 individuals in 49,811households were surveyed, aged between zero and 100 years. The survey is carried out on a sample of about 60000 households, distributed in 1456 Italian municipalities with different demographic size. According to a sampling strategy, each household is randomly selected from the municipal registry lists, creating a statistically representative sample of the resident population. The members of the extracted family are interviewed. It establishes a time frame for the interview. If one or more household members are not available for the interview during that period, or a new appointment can be arranged or will be interview another member of the household in place of the absent person.

4.2 Variable Definitions and Measurements

4.2.1 Dependent Variables

We consider two different diagnostic care measures to distinguish the effect of co-payment exemption on the diagnostic checks utilization of the whole population (*Number of total diagnostic cares- total population*), from women utilization (*Number of total diagnostic care – female population*). In line with the WHO goals (2015), we introduce the latter measure to investigate potential gender differences in health services utilization.

Information on the total number of diagnostic care utilization is obtained through a detailed questionnaire in which respondents are asked whether they have consumed, in the last four weeks before the interview diagnostic care and, in case of a positive reply, the number of tests they undergo during this period. We consider the total number of diagnostic tests as the sum of both blood and urine tests and specialist diagnostic tests, such as ultrasound, X-ray, computerized tomography (CT), magnetic resonance, and mammography, and pap-smear. We do not consider diagnostic check during hospital admissions and day hospitals. The *number of total diagnostic cares – female population* is obtained, selecting the total number of diagnostic care for women.

A high percentage of zero values characterizes the total number of diagnostic care utilization in the whole population, almost 80%, that means that only the 20% of people interviewed have undergone at least one diagnostic exam.

4.2.2 Policy variable

We analyse the effect of co-payment exemption on the number of diagnostic cares utilisation by using a dummy variable *exemption* that it is equal to 1 when people have a total or partial co-payment exemption[1]. A potential drawback of this variable is in its definition. It represents the co-payment exemption for both diagnostic care utilisation and specialistic visits. Unlike, we cannot distinguish between the two healthcare services due to the lack of clear information in the dataset, as provided by the Italian National Statistic Institute (ISTAT). Indeed, respondents declare of having the co-payment exemption and undergo both diagnostic tests and specialistic

[1] We consider both definitions as one unique right of co-payment exemption.

visits. Moreover, the only information available is the distinction between partial and total co-payment. Also, the information about the reason (age, invalidity status, chronic diseases, pregnancy) is not available, so we prefer to consider both specifications as one unique.

4.2.3 Instrumental variable

In line with World Bank definition (`https://govdata360.worldbank.org/indicators/gci`), the Global Competitiveness Index (GCI) is an aggregate index to measure the competitiveness of a country, considering several indicators on different dimensions. As it is based on several economic and institutional dimensions, we consider it a proxy of the regional administrative and bureaucracy slowness. We built the *Regional Global Competitiveness Index - RGCI*, considering both regional per capita GDP at current prices and each region's total population. It is a weighted index considering the regional GDP at current prices and regional population

It varies between 0 and 100 (that represent the high level of competitiveness).

4.2.4 Control variables

We control our dependent variables by considering both socio-economic measures and geographic areas. All controls are dummy variables taking value equal to 1 if the respondent has the specific characteristic. We consider the foreign citizen status (foreign) in order to control for potential cultural and linguistic barriers affecting the utilization of diagnostic checks. In line with the existing literature (Barbadoro et al., 2018; Reichard et al., 2015), we include several working-age adults and the elderly population (population aged 35-54; population aged 55-64 and people aged 65+). Moreover, we include a set of variables on health status. As the "presence of chronic disease remains the most important reason to use the healthcare system" (Barbadoro et al., 2018), we selected people affected by some relevant chronic diseases, such as cancer, heart attack and diabetes. Moreover, because of the self-perceived general health status may affect the decision to undergo diagnostic care, we include the variable bad health. We built this variable considering people that declare a "bad" and "very bad" self-perceived general health status. Furthermore, we include two-levels of education. In order to control whether the level of education can influence the utilization of diagnostic checks, we include the variable primary education (people having the primary education and unschooled people) and people with bachelor degrees, master or PhD (tertiary education). We include economic conditions, by analyzing the presence of unemployed people (unemployment) and also, the self-perceived household economic resources (low economic resource) as a proxy of income[2]. Finally, we consider the North and South areas for controlling potential geographical heterogeneity on diagnostic care utilisation.

4.3 Aggregate Data Patterns and Descriptive Statistics

As shown in Table 1, we consider overall survey respondents. Among 119069 people interviewed, almost 33% on average have a co-payment exemption in the year between 2011-2012. In the last four weeks before the interview, almost 35% of people underwent diagnostic care, of which almost 40% by women. The Regional Competitiveness Index is on average 54, about in half in the rank between 0 and 100. On average foreign people are almost 6% of the Italian population. In comparison, about 13% of the population is aged between 55 and 64 years old. People aged between 35 and 54 years old are on average almost 30% while older people (aged 65+) are almost 23%. These data show that more than one-third of the Italian population is composed of people with age more than 55 years old. On average, almost 4% of people are affected by cancer, 2% of heart attack and almost 6% of diabetes. On average, more than two-thirds of the population are unschooled, or with only primary education, and only 10% of Italian population have a title of tertiary education. On average, almost 40% of people interviewed perceived household economic resources as scarce. The average of unemployed people is around 7%. The population is almost homogenous between North and South areas (around 40% on average).

Table 2 compares summary statistics between people with co-payment exemption (exemption = 1) and people without it (exemption = 0). The average differences between the two samples analysed are significant for overall variables. In general, the average number of people with co-payment exemption undergoing diagnostic care is higher than people without it (almost 61% vs 22%. This average result is confirmed in women population. Interestingly, the index of RGCI is nearly the same in both populations, on average. Moreover, the average number of foreign people, Italian people aged 35-54, and 55-65 with co-payment exemption is lower than the population without it. People aged 65 and over with exemption is higher than people

[2]Income information is not available in Istat survey.

Table 1: Summary statistics - full dataset

Variable	Mean	Std. Dev.	Min.	Max.	N
Depedent variables					
Nr. of diagnostic care utilization - Total population	0.349	1.216	0	34	119069
Nr. of diagnostic care utilization- Female population	0.397	1.309	0	34	61635
Policy variable					
exemption (=1; 0 otherwise)	0.332	0.471	0	1	119069
Instrumental variable					
Global Competitiveness Index	54.039	20.377	12	93	119069
Controls					
Foreign (=1; 0 otherwise)	0.059	0.235	0	1	119069
Population aged 35-54 (=1; 0 otherwise)	0.303	0.459	0	1	119069
Population aged 55- 65 (=1; 0 otherwise)	0.133	0.339	0	1	119069
Population aged 65+ (=1; 0 otherwise)	0.227	0.419	0	1	119069
Cancer (=1; 0 otherwise)	0.038	0.19	0	1	119069
Heart Attack (=1; 0 otherwise)	0.02	0.14	0	1	119069
Diabetes (=1; 0 otherwise)	0.06	0.238	0	1	119069
Bad health (=1; 0 otherwise)	0.069	0.254	0	1	119069
Primary education (=1; 0 otherwise)	0.629	0.483	0	1	113048
Secondary education (=1; 0 otherwise)	0.271	0.444	0	1	113048
Tertiary education (=1; 0 otherwise)	0.1	0.3	0	1	113048
Low economic resourses (=1; 0 otherwise)	0.393	0.488	0	1	119069
Unemployment (=1; 0 otherwise)	0.072	0.259	0	1	102949
North area (=1; 0 otherwise)	0.421	0.494	0	1	119069
South area (=1; 0 otherwise)	0.401	0.49	0	1	119069

without co-payment exemption (people aged 65+ without co-payment exemption is only 8% instead of 52% of people with exemption in the same age class). As expected, the average number of people with the exemption, affected by chronic diseases, is higher than people without it. Also, the average number feeling bad health status and with co-payment exemption is greater than those who do not have it (17% vs 2%). Moreover, the average number of people with co-payment exemption and with tertiary education is lower than the population with the same level of education but without exemption. Whilst the population with a low level of education and co-payment exemption is greater than those with the same level of education but without co-payment exemption. We have the same comparison result in the population with scarce self-perceived household economic resources. We can see negative, but exceedingly small, average differences comparing people with and without co-payment exemption considering unemployed people and people living in the North Italian area. Even the average difference of people living in southern Italy is very small but is slightly more the average population with co-payment exemption.

Table 2: Mean Comparison: People with co-payment exemption vs. People without co-payment exemption

	exemption = 1		exemption = 0		Difference
Variable	N	Mean	N	Mean	
Depedent variables					
Nr. of diagnostic care utilization - Total population	39550	0.615	79519	0.216	0.398***
Nr. of diagnostic care utilization- Female population	22133	0.648	39502	0.257	0.390***
Instrumental variable					
Regional Global Competitiveness Index	39550	53.422	79519	54.346	-0.924***
Controls					
Foreign (=1; 0 otherwise)	39550	0.035	79519	0.071	-0.035***
Population aged 35-54 (=1; 0 otherwise)	39550	0.167	79519	0.371	-0.204***
Population aged 55- 65 (=1; 0 otherwise)	39550	0.13	79519	0.134	-0.004**
Population aged 65+ (=1; 0 otherwise)	39550	0.52	79519	0.081	0.440***
Cancer (=1; 0 otherwise)	39550	0.093	79519	0.01	0.083***
Heart Attack (=1; 0 otherwise)	39550	0.05	79519	0.005	0.045***
Diabetes (=1; 0 otherwise)	39550	0.144	79519	0.018	0.126***
Bad health (=1; 0 otherwise)	39550	0.168	79519	0.02	0.149***
Primary education (=1; 0 otherwise)	37113	0.79	75935	0.55	0.240***
Tertiary education (=1; 0 otherwise)	37113	0.051	75935	0.124	-0.074***
Low economic resourses (=1; 0 otherwise)	39550	0.461	79519	0.359	0.102***
Unemployment (=1; 0 otherwise)	35472	0.066	67477	0.076	-0.009***
North area (=1; 0 otherwise)	39550	0.415	79519	0.424	-0.009***
South area (=1; 0 otherwise)	39550	0.417	79519	0.393	0.024***

5 Results

We present our results on the effect of co-payment exemption on diagnostic care utilization for both total and female population. In section 5.1, we present estimates adopting an instrumental variable approach. We show results applying the CRL methodology, in section 5.2. Finally, in order to capture some information on the utilization of diagnostic care related to individual characteristics, in section 5.3 we present results by using the CRL methodology considering eight different population groups. These groups are based on the self-perceived household economic resources, distinguished in low and high economic resources; people affected by cancer, and geographical area (people living in the North or South area).

5.1 Instrumental variable results

Table 3 shows second stage OLS regression results for the number of diagnostic care utilization in the total and female population. In both analyses, the coefficient of our treatment variable (exemption) is negative and significant only in the women utilization of diagnostic care. These results may reveal that even though, we expected an increase of diagnostic care utilization in people with co-payment exemption, the inappropriateness of infrastructure and technologies due to regional administrative and bureaucracy slowness negatively influence their utilization. Moreover, the significant coefficient of the number of diagnostic care utilization in the female population show potential needs unmet.

However, Table 4 shows that in the first stage, the coefficient of RGCI is positive and significant, but analysing statistical test, we can claim that the instrument is weak. According to the rule of thumb, the F-test is greater than 10. Also, if we consider the Anderson-Rubin Wald test (F and Chi-squared) and Stock-Wright LM S statistic, we can conclude that our instrument is weak. Exclusion restriction and Joint statistical independence are not met. We provide an alternative empirical strategy to estimate the average treatment effect by using an inference by intersection bounds.

Table 3: Results on diagnostic care utilization - Fixed effect results (second stage)

Variables	Nr. of diagnostic care - total population Coeff.\Std. Err.	Nr. of diagnostic care- women popualtion Coeff.\Std. Err.
Depedent variables		
exemption	-0.622	-1.042*
	(0.515)	(0.559)
Controls		
Foreign	-0.065**	-0.080***
	(0.026)	(0.028)
Population aged 35-54	0.070***	0.088***
	(0.020)	(0.020)
Population aged 55-65	0.183**	0.232***
	(0.073)	(0.079)
Population aged 65+	0.446*	0.624**
	(0.264)	(0.286)
Cancer	0.429***	0.531***
	(0.134)	(0.151)
Heart Attack	0.207***	0.197***
	(0.071)	(0.063)
Diabetes	0.234**	0.294***
	(0.094)	(0.088)
Bad health	0.395***	0.436***
	(0.089)	(0.086)
Primary education	0.025	0.058
	(0.042)	(0.047)
Tertiary education	0.016	0.021
	(0.016)	(0.016)
Low economic resourses	0.035	0.060*
	(0.030)	(0.034)
Unemployment	0.080	0.111**
	(0.058)	(0.049)
North area	0.001	0.015
	(0.009)	(0.013)
South area	-0.021	-0.010
	(0.013)	(0.016)
Constant	0.167***	0.235***
	(0.025)	(0.039)
Observations	102,949	53,865
R-squared	-0.175	-0.499
Underidentification test		
K-P rk LM statistic Chi-sq(1)	11.262	12.826
P-val	0.0008	0.0003
Weak identification test		
K-P rk Wald F statistic	11.259	12.822
Overidentification test of all instruments		
Hansen J statistic	0.000	0.000

Note:

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

K-P= Kleibergen-Paap

Table 4: First stage statistics tests on Diagnostic care utilization

Variables	Fisrt stage results on Nr. of diagnostic care - total population exemption Coeff.\Std. Err.	Nr. of diagnostic care - women popualtion exemption Coeff.\Std. Err.
Regional Competitiveness Index	.0003***	.0004***
	(.00008)	(.0001)
F-test of excluded instruments		
F(1, 102933)	11.26	12.82
P-val	0.0008	0.0003
Angrist-Pischke multivariate F test of excluded instruments		
AP F(1, 102933)	11.26	12.82
P-val	0.0008	0.0003
Underidentification test		
AP Chi-sq(1)	11.26	12.83
P-val	0.0008	0.0003
K-P rk LM statistic		
Chi-sq(1)	11.26	12.83
P-val	0.0008	0.0003
Weak identification test		
K-P Wald rk F statistic	11.26	12.82
Weak instrument robust inference		
AR Wald test		
F(1,102933)	1.80	5.35
P-val	0.1793	0.0208
AR Wald test		
Chi-sq(1)	1.80	5.35
P-val	0.1793	0.0207
SW LM S statistic		
Chi-sq(1)	1.80	5.35
P-val	0.1793	0.0207
Nr. of observations	102949	53865
Nr. of regressors	16	16
Nr. of endogenous regressors	1	1
Nr. of instruments	16	16
Nr. of excluded instruments	1	1

Note:
Robust standard errors in parentheses $\star\star\star p < 0.01$, $\star\star p < 0.05$, $\star p < 0.1$
K-P: Kleibergen-Paap; AR: Anderson-Rubin; SW: Stock-Wright

5.2 Intersection bounds estimation

Results in Table 5 show different intervals of confidence combining upper and lower bounds for the average total number of diagnostic care utilization in both the total and in the female population. In this analysis, we seek to estimate the upper and lower bounds assuming that the co-payment exemption increases the average number of diagnostic care utilization, conditional to the average standardized Regional Global Competitiveness. In general, findings reveal a positive effect of co-payment exemption on the utilization of diagnostic cares for both populations. Specifically, people living in a region with an average level of GCI, increase the number of diagnostic care utilization within a confidence interval with a minimum of almost 0.05 up to a maximum of almost 0.56 (at 90%, 95% and 99% level). Whilst the confidence of intervals (at 90%, 95% and 99%) of the increasing number of diagnostic care utilization, vary from a minimum of almost 0.27 up to almost 0.75 whether we only consider women living in a region with an average level of GCI.

Table 5: Results on estimation intersection bounds

	Nr. diagnostic care utilization - total population		Nr. of diagnostic care utilization - female population	
	LB Value	UB Value	LB Value	UB Value
90% two-sided C.I.	[0.049,	0.557]	[0.269,	0.747]
95% two-sided C.I.	[0.049,	0.559]	[0.267,	0.749]
99% two-sided C.I.	[0.047,	0.563]	[0.264,	0.752]

Note: AIS(adaptive inequality selection) is applied

5.3 Estimate of intersection bounds by several individual characteristics

In Table 6 we show results on inference bound estimation on the averahe number of diagnostic cares in the total population. We consider eight groups combining the following individual characteristics: individuals with scarce self-perceived household economic resources, affected by cancer and living in the North or South area. In line with Manning et al. 1987, we consider the scarce self-perceived economic resources (at the household level) as a proxy of income to investigate some possible opportunistic behaviours and whether the utilization of diagnostic care is sensitive to the price. For simplicity, we comment results for the confidence interval at 90%, as slight differences appear in comparison with the other level of interval confidence.

In general, the average utilization of diagnostic cares increases among people with co-payment exemption. However, distinguish between people affected by chronic diseases, such as cancer, and people with a scarce economic resource (denominated "poor" for simplicity) we can notice some relevant differences. First, we consider people living in northern Italy. Confidence interval at 90%, in comparison between poor and wealthy individuals affected by cancer, show that in both groups, the presence of co-payment exemption increases the average number of diagnostic care utilization. However, the interval for poor people is broader than wealthy individuals. Thus, given individuals with a chronic disease, the "income" differences reveal potential opportunistic behaviour. In line with previous results, we observe that if we consider people do not affect by cancer, the interval of confidence at 90% is wider in the group of poor people than in the wealthy group. As we are analyzing the effect of co-payment, this group is composed of people having exemption due to other chronic diseases or different co-payment exemption type (age, income, invalidity status, pregnancy). Now, we consider people living in southern Italy. We consider people affected by cancer. The interval of confidence at 90% is wider in the group of people with scarce household economic resources. This result confirms that income differences drive an increase on the average number of diagnostic care utilization. Results on people do not affect by cancer and living in the southern regions, confirm that exemption increased the average number of diagnostic cares among people with low economic resources. Confidence intervals are not similar in the two geographical areas. They are wider in the southern regions. This result reveals a relevant geographical heterogeneity in the average number of diagnostic care utilization among people with co-payment exemption.

Table 6: Intersection bounds estimation by the self-perceived economic resources, people affected by cancer and geographical areas - Results for average number of diagnostic care utilization in the total population

	poor, cancer, N		non-poor, cancer, N	
	LB Value	UB Value	LB Value	UB Value
90% two-sided C.I.	[-0.011,	2.442]	[0.133,	2.355]
95% two-sided C.I.	[-0.016,	2.44]	[0.107,	2.411]
99% two-sided C.I.	[-0.028,	2.651]	[0.060,	2.517]
	poor, no cancer, N		non-poor, no cancer, N	
	LB Value	UB Value	LB Value	UB Value
90% two-sided C.I.	[0.416,	0.988]	[0.512	0.711]
95% two-sided C.I.	[0.404,	1.005]	[0.503	0.721]
99% two-sided C.I.	[0.381,	1.038]	[0.485,	0.741]
	poor, cancer, S		non-poor, cancer, S	
	LB Value	UB Value	LB Value	UB Value
90% two-sided C.I.	[0.086,	3.505]	[0.373	2.620]
95% two-sided C.I.	[0.051,	3.595]	[0.298	2.687]
99% two-sided C.I.	[-0.010,	3.761]	[0.209,	2.815]
	poor, no cancer, S		non-poor, no cancer, S	
	LB Value	UB Value	LB Value	UB Value
90% two-sided C.I.	[0.447,	1.181]	[0.674	0.989]
95% two-sided C.I.	[0.436,	1.202]	[0.650	1.004]
99% two-sided C.I.	[0.414,	1.241]	[0.605,	1.033]

Note: AIS(adaptive inequality selection) is applied
Poor: people declaring scarce self-perceived household economic resources
Cancer: people affected by cancer
N: North area; S: South area

Table 7 shows testing results for the verification of the null hypotesis by using CRL techniques as developed by Chernozhukov,et al. (2013, 2015) in the total population. These results show that the testing value is not rejected in the 90%, 95%, 99% confidence interval. In other words, the null hypothesis is NOT rejected at the 10%, 5%, 1% level. It it verifed that people with co-payment exemption and living in a region with an average level of RGCI, increase their utilization of diagnostic care on average.

Table 7: Testing results for GCI in the total population.

	P, C, N	NP, C, N	P, NC, N	NP, NC, N	P, C, S	NP, C, S	P, NC, S	NP, NC, S
Subgroup ID	1	2	3	4	5	6	7	8
Observations	234	439	1887	3890	231	230	2212	2536
10%	NR	NR	NR	NR	NR	NR	NR	NR
5%	NR	NR	NR	NR	NR	NR	NR	NR
1%	NR	NR	NR	NR	NR	NR	NR	NR

P: people declaring scarce self-perceived household economic resources
NP: people declaring good self-perceived household economic resources
C: people affected by cancer
NC: people do not affected by cancer
N: North area; S: South area

In Table 8, we show the confidence of interval for the average number of diagnostic care utilization by considering only the female population. As in the previous analysis, we consider eight groups combining the following individual characteristics: individuals with scarce self-perceived household economic resources, affected by cancer and living in the North or South area. Findings confirm that also women having a co-payment undergo on average more than people without it. Among women with co-payment exemption and affected by cancer, the most relevant driver of the different level of diagnostic care utilization, on average, seems to be the "income". Indeed, analyzing poor and wealthy women affected by cancer and living in the northern regions, we notice that the size of confidence interval is wider among poor women. We obtain the same results if we compare the same group of women with the same health condition and economic resources but living in the South regions. Moreover, the average utilization of diagnostic care is heterogeneous among geographical areas. In line with previous results, the interval of confidence is wider in the groups of women living in the south area. Interestingly, the size of interval confidence for women do not affect by cancer, is smaller than the size of those estimated in the female population affected by cancer. This result may reveal that cancer is a relevant determinant of diagnostic care utilization in the female population.

Table 8: Intersection bounds estimation by the self-perceived economic resources, people affected by cancer and geographical areas - Results for average number of diagnostic care utilization in the female population

	poor, cancer, N		non-poor, cancer, N	
	LB Value	UB Value	LB Value	UB Value
90% two-sided C.I.	[-0.006,	0.952]	[0.029,	0.755]
95% two-sided C.I.	[-0.009,	0.988]	[0.017,	0.781]
99% two-sided C.I.	[-0.009,	1.054]	[-0.003,	0.830]
	poor, no cancer, N		non-poor, no cancer, N	
	LB Value	UB Value	LB Value	UB Value
90% two-sided C.I.	[0.056,	0.138]	[0.064	0.082]
95% two-sided C.I.	[0.053,	0.141]	[0.062	0.085]
99% two-sided C.I.	[0.050,	0.147]	[0.059,	0.091]
	poor, cancer, S		non-poor, cancer, S	
	LB Value	UB Value	LB Value	UB Value
90% two-sided C.I.	[-0.012,	1.190]	[0.071	0.966]
95% two-sided C.I.	[-0.028,	1.244]	[0.054	1.001]
99% two-sided C.I.	[-0.062,	1.333]	[0.022,	1.069]
	poor, no cancer, S		non-poor, no cancer, S	
	LB Value	UB Value	LB Value	UB Value
90% two-sided C.I.	[0.047,	0.130]	[0.066	0.112]
95% two-sided C.I.	0.045,	0.135]	[0.064	0.114]
99% two-sided C.I.	[0.042,	0.141]	[0.060,	0.119]

Note: AIS(adaptive inequality selection) is applied
Poor: people declaring scarce self-perceived household economic resources
Cancer: people affected by cancer
N: North area; S: South area

Table 9 shows testing results for the verification of the null hypotesis by using CRL techniques as developed by Chernozhukov,et al. (2013, 2015), by considering the female population. These results show that the testing value is not rejected in the 90%, 95%, 99% confidence interval. In other words, the null hypothesis is NOT rejected at the 10%, 5%, 1% level. It it verifed that women with co-payment exemption and living ina region with average level of RGCI, increase their utilization of diagnostic care on average.

Table 9: Testing results for GCI in the female population.

	P, C, N	NP, C, N	P, NC, N	NP, NC, N	P, C, S	NP, C, S	P, NC, S	NP, NC, S
Subgroup ID	1	2	3	4	5	6	7	8
Observations	403	843	8488	18216	402	391	11868	13522
10%	NR	NR	NR	NR	NR	NR	NR	NR
5%	NR	NR	NR	NR	NR	NR	NR	NR
1%	NR	NR	NR	NR	NR	NR	NR	NR

P: people declaring scarce self-perceived household economic resources
NP: people declaring good self-perceived household economic resources
C: people affected by cancer
NC: people do not affected by cancer
N: North area; S: South area

6 Conclusions

In this paper, we seek to provide evidence on the effect of exemption on the utilization of diagnostic cares. We decide to consider diagnostic care as they are health services for both preventative care and controls for overt diseases. Thus, their use into the analysis allow to have a dimension of potential basic needs unmet. In line with the literature (Arrow, 1963) on the effect of health insurance on healthcare spending that shows overconsumption of healthcare services in insured people ("moral hazard"), we assume people with co-payment exemption increase the utilization of diagnostic care. First, we adopt an IV strategy to address potential co-payment exemption endogeneity. We use the Regional Global Competitiveness Index (RGCI) as an instrument and proxy of administrative and bureaucracy slowness. However, limits on the validity of the instrument have led us to adopt an alternative empirical strategy to capture the effect of co-payment exemption. Thus, we apply an inference on intersection bound as developed by Chernozhukov,et al. (2013, 2015). that allows to estimate lower and upper bounds for different intervals confidence, where we can see the effect of co-payment exemption on the average number of diagnostic care, conditional at the average standardized level of Regional Global Competitiveness Index. We consider two measures of the number of diagnostic cares, distinguish between the utilization in the total population and in the female population. As several diagnostic cares are specific only for women, we thought to consider these two measures separately. Our preliminary results show a general increase of the number of diagnostic cares utilization due to the co-payment exemption. Moreover, in order to investigate potential opportunistic behaviours, we select people affected and do not affected by cancer, with low and high self-perceived economic resources and living in northern and southern regions. Findings show that the intervals of confidence are wider in both geographical areas in the group of people with low economic resources, given the presence of cancer. These results may reveal potential opportunistic behaviours. Moreover, the size of the confidence intervals is broader in southern regions, showing a heterogeneous effect of cop-payment exemption due to geographical differences.

This work highlights a relevant starting point for future analysis. At the moment, the analysis is only preliminary due to several lacks of data information. We use survey data with data collected during 2011-2012, were the information on the number of diagnostic cares is related only to the four weeks before the interview. Thus, we can limit our results only on partial time information. Moreover, we are interested in the effect of co-payment exemption distinguished by income status. Data do not provide any income information, so we use a proxy of income, considering the self-perceived household economic resources. The lack of information on real income can limit the interpretation of possible opportunistic effects. Also, we cannot distinguish between the type of co-payment reason (such as age, invalidity status, chronics disease, etc.). The available information is only between total and partial co-payment exemption. We aggregate this information because, even keeping separately, we could not distinguish the reason for exemption. In conclusion, our work presents several drawbacks, but it could be a good starting point for further analysis.

Moreover, the estimation of the average number of diagnostic cares by inference of intersection bounds permits to understand the dimension of the phenomenon when the exact average treatment effect could not reach. Furthermore, this preliminary analysis may be relevant for future healthcare policy evaluation on efficiency improvements. As the high healthcare expenditure is the most relevant concern in Italy, as well as in other western countries, it should be necessary to understand which are the real inefficiencies in the health system and the cause of their high costs without reducing the access and utilization to healthcare services. We reserve in future analysis to enrich our study on the effect of co-payment by considering several healthcare services and also collecting more useful data information.

References

Abadie A., J. Angrist and Imbens G. (2002). "Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings". In: *Econometrica* 70.1, pp. 91–117. DOI: 10.1111/1468-0262.00270.

Angrist J., G. Imbens and Rubin D. (1996). "Identification of Causal Effects Using Instrumental Variables". In: *Journal of the American Statistical Association* 91.434, pp. 444–455. DOI: 10.2307/2291629.

Angrist, J. and Krueger A. (1991). "Does Compulsory School Attendance Affect Schooling and Earnings?" In: *The Quarterly Journal of Economics* 106.4, pp. 979–1014.

Angrist, J. and Imbens G. (1995). "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity". In: *Journal of the American Statistical Association* 90.430, pp. 431–442. DOI: 10.1080/01621459.1995.10476535.

Arrow, K. (May 1963). "Agency and the welfare economics of medical care". In: *American Economic Review* 53.

Atella, V. and Kopinska J.A. (2014). "The impact of cost-sharing schemes on drug compliance in Italy: evidence based on quantile regression." In: *Journal of the European Economic Association* 59, pp. 329–339.

Barbadoro P., D'Alleva A. Galmozzi S. Zocco G. Di Stanislao F. Prospero E. et al. (2018). "Variations in diagnostic testing utilization in Italy: Secondary analysis of a national survey". In: *PLoSONE* 13.e0196673. DOI: https://doi.org/10.1371/journal.pone.0196673.

Chernozhukov V., Kim W. S. Lee and Rosen A. (2015). "Implementing intersection bounds in Stata". In: *Stata Journal* 15.1, 21–44(24). URL: http://www.stata-journal.com/article.html?article=st0369.

Chernozhukov, V., S. Lee, and A. Rosen (2013). "Intersection Bounds: Estimation and Inference". In: *Econometrica* 81.2, pp. 667–737.

Fireman, S.J. and Swain B. (1996). "Effect of a Copayment on use of the Emergency Department in a Health Maintenance Organization". In: *New England Journal of Medicine* 334, pp. 635–641.

G. Costa, T. Spadea and Petrelli A. (2012). “Cost sharing, salute ed equità nella salute, Il ruolo del copayment nella sanità italiana”. In: *Seminario di studio e di discussione promosso da Agenas e Aies, Roma.*

Imbens, G. and Angrist J. (1994). “Identification and Estimation of Local Average Treatment Effects”. In: *Econometrica* 62, 467–475.

ISTAT (2013). “Multiscopo ISTAT – Health Conditions and Use of Health Services (2011-2012”. In.

Italian, Ministry of Health (2016). “Quaderni del Ministero della Salute”. In.

J., Angrist and Pischke J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist’s Companion.* 1st ed. Princeton University Press.

Manning G., Newhouse P. Duan N. Keeler B. and A. Leibowitz (1987). “Health insurance and the demand for medical care: evidence from a randomized experiment”. In: *American Economic Review*, pp. 251–277.

Mourifié, I. and Wan Y. (2017). “Testing Local Average Treatment Effect Assumptions”. In: *The Review of Economics and Statistics* 99.2, pp. 305–313.

Neri, S. (Sept. 2019). “Economic Crisis, Decentralisation and Health Inequalities: The Case of Italy”. In: DOI: 10.5772/intechopen.89006.

Reichard A., Gulley S. P., E. K. Rasch, and Leighton C. (2015). “Diagnosis isn’t enough: Understanding the connections between high health care utilization, chronic conditions and disabilities among U.S. working age adults”. In: *Disability and Health Journal* 8.4, pp. 535 –546. ISSN: 1936-6574. DOI: https://doi.org/10.1016/j.dhjo.2015.04.006.

Sargan, J. (1958). “The Estimation of Economic Relationships using Instrumental Variables”. In: *Econometrica* 26, pp. 393–415. ISSN: 0167-6296. DOI: doi:10.2307/1907619.

Tamblyn, R. et al. (2001). “Adverse Events Associated with Prescription Drug Cost-Sharing Among Poor and Elderly Persons”. In: *Journal of the American Medical Association* 285, pp. 421–429.

Toth, F. (2014). “How health care regionalisation in Italy is widening the North–South gap”. In: *Health Economics, Policy and Law* 9.3, 231–249. DOI: 10.1017/S1744133114000012.

V. Atella, V. Conti and Kopinska J. (2015). "Efficacia dei meccanismi di compartecipazione sulla spesa sanitaria: le differenze regionali e gli effetti in termini di salute e spesa sanitaria". In: *I Quaderni della Fondazione Farmafactoring* 2.

WHO (2015). *Roadmap for action, 2014-2019.*