UNIVERSITY OF CATANIA

DEPARTMENT OF ECONOMICS AND BUSINESS

PhD program in Economics and Management (XXVIII cycle)

Decentralization and health performance in Italy: theoretical and empirical issues

LIVIO FERRANTE

_____

DOCTORAL THESIS

_____

Supervisor
Prof. Marina Cavalieri

PhD Coordinator
Prof. Isidoro Mazza

TABLES OF CONTENTS

INTRODUCTION

In the last decades, few comparative political economy debates have been as intense and exciting as those concerning whether decentralization enhances public service delivery. Advocates of fiscal decentralization have argued that greater local authority in decision-making improves the efficiency of public service delivery, because government outputs can be provided in small units and tailored directly to local preferences (Besley and Coate, 2003; Oates, 1972). In addition, decentralization creates competition for capital and labour, leading to improved governance outcomes (Tiebout, 1956). A number of issues pertaining the design and the implementation of decentralization have also been raised by the first generation literature, which mainly concern the existence of inter-jurisdictional spillovers, vertical fiscal relations, limited devolution and overlapping responsibilities. Along with these, other potential concerns have been emphasized by second generation theories of fiscal federalism (e.g. Lockwood, 2002; Besley and Coate, 2003; Weingast, 2009) and focus on political economy and corruption problems.

The above theoretical arguments in favour of decentralization have been highly influential for policy decisions, leading many countries worldwide to decentralize administrative, fiscal, and political functions to subnational levels of government. Over the time, these country experiences have provided the basis for empirical investigation of the effects of decentralization. Nevertheless, empirical evidence remains limited and often points to inconsistent, inconclusive and, ultimately, irrelevant findings. Litvack et al. (1998) summarize the literature this way: "It is not much of an exaggeration to say that one can prove, or disprove, almost any proposition about decentralization by throwing together some set of cases or data" (p.30). Similarly, by reviewing previous

works on the topic, Treisman's (2007) concludes: "To date, there are almost no solidly established, general empirical findings about the consequences of decentralization" (p.250). "Almost nothing that is robust or general has emerged" (p.268).

The lack of conclusive evidence along with the dissatisfaction of many countries with the results of their decentralization reforms (Mansuri and Rao, 2012) have recently led to a re-thinking of the promising paradigm of decentralization and to a new trend toward re-centralization (Dickovick, 2011; Malesky et al., 2014).

Based on the above picture, the primary objective of this dissertation is to deepen the current understanding of the relationship between decentralization and public policy results. In doing this, the work builds further on the existing theoretical framework of the relationship between decentralization and service delivery while trying to overcome some of the well-known limits of the previous empirical studies. Firstly, existing analyses mainly explore the effects of decentralization by looking at the issues of efficiency, economic growth and, to a less extent, equity. Very few researches focus the attention on individual outcomes, that should be the ultimate result of any public policy aiming at promoting wellbeing among citizens. Secondly, the majority of such empirical contributions concerns less developed or developing countries, thus almost neglecting decentralization experiences in developed contexts. The latter fact along with the marked preference for cross-country comparisons prevent drawing general conclusions. This is because the effects of decentralization are highly context-dependent and shaped by numerous factors that are difficult to be systematically isolated and assessed when heterogeneous countries are considered. It has been also argued that the ambiguousness of the existing empirical findings on the relationship between decentralization and the efficiency of public service provision reflects the lack of clarity on the concept of decentralization and on its related "determs" (deconcentration, delegation,

denationalisation, destatisation, and, devolution) (Tomaney et al., 2011). A further limit of the current literature concerns the measurement of a phenomenon that is itself complex, multidimensional and intertwined, and, accordingly, difficult to be synthesized by a single comprehensive indicator. In general, indicators refer to just one of the three core dimensions of decentralization, that is fiscal, administrative and political, thus providing a partial picture of the full phenomenon. Finally, spatial and temporal patterns have seldom been explicitly considered by scholars, when analysing the effects of decentralization processes.

In this regard, the present dissertation has many empirical advantages over the above studies. First of all, a within-country approach is used, which makes the findings less vulnerable to the heterogeneity problems of cross-county studies and, hence, more robust. Specifically, both the empirical analyses carried out for this work concern a developed country, such as Italy, that has been involved in an intense, complex and unfinished process of decentralization during the last 20 years. This process has deeply reshaped the pre-existing intergovernmental relations, devolving new legislative powers and sectorial competencies to sub-national governments, especially regional ones.

To assess the effects of the Italian decentralization process, objective healthcare outcomes (i.e. infant mortality rates and life expectancy at birth) are chosen as measures of public performance. This choice is motivated by different reasons. Firstly, the health sector has been at the heart of the decentralization reforms in Italy, with responsibilities for the funding and delivery of healthcare services devolved to regional and local authorities. Secondly, healthcare outcomes in terms of either mortality or life expectancy are suggested to be not only a quantity but also a quality indicator of citizens' well-being (Sen, 1998; Becker et al., 2005), thus superior to income measures.

With regard to the measures of decentralization, a multiple approach is used, consistently with the multidimensional nature of the study process. Therefore, quantitative continuous indicators of fiscal decentralization are considered along with discrete variables that capture the political changes following the introduction of specific reforms. The choice of quantitative indicators has been justified by the fact that a great variation exists in the way Italian regions have used their tax autonomy. In this regard, Italian regions represent a unique "natural laboratory" to test the effects of decentralization on public performance.

Finally, in this dissertation decentralization has been explored along multiple spatial dimensions and with a multi-level perspective. Indeed, as better explained next, the phenomenon and its effects have been analysed both at a macro-level, looking at the autonomous regional health systems, and at a micro-level, looking at the administratively decentralized hospital structures.

Three chapters constitute the main structure of this contribution. The first chapter investigates the relationship between fiscal decentralization and regional health outcomes, as measured by infant mortality rates, in Italy. The paper employs a panel of all Italian regions over a period of 17 years (from 1996 to 2012), applying a linear Fixed-Effect model. Two different quantitative measures of fiscal decentralization are used, which capture the degree of regional decision-making autonomy in the allocation of tax revenues and the extent of regional transfer dependency from the central government (i.e. vertical fiscal imbalance). Methodologically, to account for the temporal dynamics of the decentralization impact, the robustness of the findings is checked, among others, with respect to the use of an Error Correction Model, which allows to disentangle short and long run effects. The analysis also deals with the issue of heterogeneous distributional geographical responses by modelling the asymmetric

impact of decentralization on infant mortality rates according to the level of regional wealth.

The second chapter addresses the issue of convergence. Here the main research questions are whether health outcomes in Italy converge/diverge over time and, more importantly, whether decentralization has played a somewhat role in the convergence/divergence process. Using a pooled dataset with the same time span as the previous one, the conventional measures of σ- and β- (both absolute and conditional) convergence are estimated for two different regional health outcomes (*i.e.* infant mortality rate and life expectancy at birth). Again, two measures of decentralization are employed in order to catch both the degree of fiscal regional decision-making autonomy (*i.e.* the same indicator as in chapter 1) and the political decentralization dynamics (*i.e.* a dummy variable taking the value of 1 after the introduction of the 2001 constitutional reform). From a methodological point of view, the real novelty of the analysis is to take spatial dependence and neighbourhood effects among the regions into consideration. Modelling the impact of decentralization through an interaction term, the speed of convergence is found to be significantly affected by the level of decentralization.

The third and last chapter deals with the issue of the effects of decentralization from a different but related viewpoint. Compared to the previous chapters, it examines descriptively the administrative aspects of decentralization by a lower (micro) level perspective, looking at the managerial autonomy of local healthcare structures. The emphasis is here on the role of intrinsic and extrinsic motivations in enhancing accountability and improving the performance of healthcare system, in general, and the quality of hospital care, in particular. Though the focus is not specifically on the Italian system, the analysis is particularly relevant for this country, where regional governments, in charge of the responsibilities for the financing and the delivery of

healthcare, act through a network of Local Health Authorities – i.e. public entities with their own budgets and management, which directly run small public hospitals -, public hospital trusts with full managerial autonomy and accredited for-profit private providers. The understanding research hypothesis here is that the way in which the financial incentive schemes for providers are designed and structured is likely to affect their effectiveness in pursuing the expected results (e.g. improved efficiency and quality of healthcare service delivery). However, the same incentive is expected to work differently according to the provider's degree of decision-making autonomy and its utility function.

References

Becker, G.S., Philipson, T. J., Soares, R.R. (2005). The quantity and quality of life and the evolution of world inequality. *The American Economic Review*, *95*(1), 277-291.

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

Dickovick, J.T. (2011). *Decentralization and recentralization in the developing world: Comparative studies from Africa and Latin America*. Penn State Press.

Litvack, J.I., Ahmad, J., Bird, R.M. (1998). *Rethinking decentralization in developing countries*. World Bank Publications.

Lockwood, B. (2002). Distributive politics and the costs of centralization. *The Review of Economic Studies*, *69*(2), 313-337.

Malesky, E.J., Nguyen, C.V., Tran, A. (2014). The impact of recentralization on public services: A difference-in-differences analysis of the abolition of elected councils in Vietnam. *American Political Science Review*, *108*(01), 144-168.

Mansuri, G., Rao, V. (2012). *Localizing development: does participation work?*. World Bank Publications.

Oates, W. (1972). *Fiscal federalism*. Harcourt Brace Jovanovich. New York, NY.

Sen, A. (1998). Mortality as an indicator of economic success and failure. *The Economic Journal*, *108*(446), 1-25.

Tiebout, C.M. (1956). A pure theory of local expenditures. *The journal of political economy*, 416-424.

Tomaney, J., Pike, A., Rodríguez-Pose, A., Tselios, V., Torrisi, G. (2011). *Decentralisation outcomes: a review of evidence and analysis of international data.* Final Report to Department for Communities and Local Government. London: Department of Communities and Local Government.

Treisman, D. (2007). What have we learned about the causes of corruption from ten years of cross-national empirical research?. *Annual Review of Political Science*, *10,* 211-244.

Weingast, B.R. (2009). Second generation fiscal federalism: The implications of fiscal incentives. *Journal of Urban Economics*, *65*(3), 279-293.

CHAPTER 1[*]

Does fiscal decentralization improve health outcomes? Evidence from infant mortality in Italy

ABSTRACT

Despite financial and decision-making responsibilities having been increasingly devolved to lower levels of government worldwide, the potential impact of these reforms remains largely controversial. This paper investigates the hypothesis that a shift towards a higher degree of fiscal autonomy of sub-national governments could improve health outcomes, as measured by infant mortality rates. Italy is used as a case study since responsibilities for healthcare have been decentralized to regions, though the central government still retains a key role in ensuring all citizens uniform access to health services throughout the country. A linear fixed-effects regression model with robust standard errors is employed for a panel of 20 regions over the period 1996-2012 (340 observations in the full sample). Decentralization is proxied by two different indicators, capturing the degree of decision-making autonomy in the allocation of tax revenues and the extent to which regions rely on fiscal transfers from the central

government. The results show that a higher proportion of tax revenues raised and/or controlled locally as well as a lower transfer dependency from the central government are consistently associated with lower infant mortality rates, *ceteris paribus*. The marginal benefit from fiscal decentralization, however, is not constant but depends on the level of regional wealth, favouring poorest regions. In terms of policy implications, this study outlines how the effectiveness of decentralization in improving health outcomes is contingent on the characteristics of the context in which the process takes place.

Acknowledgements

1. Introduction

Over the past few decades, decentralization has been implemented by an increasing number of countries, becoming a key element of the public-sector reform (Manor, 1999). In this broad process, healthcare services have occupied a central position (Saltman et al., 2007; Costa-Font and Greer, 2013). Devolution of healthcare responsibilities to lower levels of government exists, under different forms, in Scandinavian countries, traditionally National Health Service (NHS) systems (e.g. Spain, Italy and the UK), federal states (e.g. Switzerland, Canada and Australia) and developing nations as diverse as China, India, the Philippines and Tanzania.

Although reasons for decentralization remain ultimately country-specific, behind this worldwide trend is the general conviction that the transfer of powers and responsibilities to lower tiers of government allows a better match between citizens' preferences and public policies (Oates, 1972), rooted in the implicit assumption of welfare improving mobility (Tiebout, 1956). A decentralized structure of government is claimed to improve service provision efficiency essentially by reducing information asymmetries, by enhancing accountability of locally elected policy makers, by promoting community participation, by fostering competition among jurisdictions and by encouraging innovation in government policies and diffusion of best practices (Weingast, 2009).

Besides these benefits, a number of reasons have also been suggested to explain why decentralization may not enhance or even hinder the efficiency of public services provision (Prud'homme, 1995). These focus on the failure to exploit economies of scale in decentralized provision, the risk of local elites' capture of the decision-making process (Bardhan and Mookherjee, 2000) and the lack of organization and administrative capacity by local governments (Smith, 1985).

As for decentralization in healthcare, the characteristics of health goods and services further complicate the already ambiguous normative predictions (Costa-Font, 2012). Indeed, healthcare is regarded as an 'experience good' where the consumer faces an adverse selection problem and quality dimension is filtered by intermediate agents (e.g. physicians). These features may limit the ability of patients to evaluate objectively the health system performance, which is the prerequisite for holding local politicians accountable for their policy choices. Furthermore, the costs of decentralization are generally traced back to the presence of inter-jurisdictional spillovers, public good characteristics and diseconomies of scale (Bardhan and Mookherjee, 1998; Besley and Coate, 2003), all of which are very common in the healthcare sector (Alves et al., 2013). For example, health prevention initiatives promoted by one jurisdiction are likely to benefit neighbours; hence, a strong incentive exists for this to free-ride and to invest sub-optimally. In terms of cost savings, the relevant advantages arising from the collective purchasing of many healthcare resources (e.g. drugs, equipment, medical devices), the aggregate production and provision of health services, and the joint administration of healthcare structures (e.g. hospitals) may strongly prompt for a centralized solution. Even the assumption of perfect mobility of citizens, which is at the heart of the Oates' argument, may not reflect reality of the healthcare sector, especially with regard to chronically ill patients and elderly ones with relevant health needs (Jimenez-Rubio, 2011a,b).

Notwithstanding, the main argument for healthcare decentralization remains the potential efficiency gains achieved by mitigating information asymmetries and by better tailoring programs to heterogeneous local needs, preferences and providers' features (Levaggi and Smith, 2005). The mechanism through which all this is likely to happen relies largely on inter-jurisdictional competition. In a decentralized self-financing

setting, sub-national governments (SNGs) compete with each other to provide high-quality healthcare services at low user charges or financed through lower taxes. Under the assumption that individuals are well informed and able to 'vote with their feet', better performing jurisdictions will attract mobile citizens and, hence, tax base (Levaggi and Zanola, 2007). Furthermore, the opportunity for citizens to benchmark the policy choices made by their local representatives with the neighbour counterparts' actions helps to enhance the political accountability of local health systems ('yardstick competition'; Shleifer, 1985): local politicians providing low-quality health services are expected not to be re-elected.

Both theoretical and empirical literature offers conflicting evidence regarding the consequences of inter-jurisdictional interactions, especially on health expenditure (Costa-Font et al, 2015). Competition among jurisdictions may result in either a downward ('race to the bottom') or an upward ('race to the top') bias in public spending for healthcare. The latter situation is more likely to occur in presence of soft-budget constraints and expectations of future bailing out of regional health deficits (Bordignon and Turati, 2009).

When fiscal responsibilities for healthcare are decentralized, equity concerns could be a major issue. Vertical fiscal imbalances may arise due to a mismatch between revenue raising powers and health expenditure responsibilities at a sub-national level. Common solution to the problem requires the use of piggybacked and shared taxes as well as grants from the central government (CG). However, these may weaken accountability by SNGs, leading to soft budget constraints and common pool problems. Inter-jurisdictional healthcare spillovers provide a further economic rationale for vertical grants, to the extent that the CG is able at identifying them and targeting grants optimally.

More importantly, horizontal imbalances can emerge primarily as a result of exogenous regional differences in healthcare needs and fiscal capacity. Specifically, given the positive relationship between income and health status (Marmot, 2002), poorer regions are expected to experience higher healthcare needs. However, when taxes are locally collected, the same "tax effort" generates different levels of revenues between rich and poor regions due to the effect of differing tax bases. Hence, horizontal fiscal equalisation schemes based on solidarity principles and risk sharing agreements are the usual answer.

The present study aims at contributing to the above debate by analysing empirically the effect of decentralization on health outcomes, as measured by infant mortality rates (IMRs). A panel of 20 Italian regions over a 17-year period (1996-2012) is considered. Italy is taken as a case study since responsibilities for healthcare have been progressively decentralized to regions, though the CG still retains a key role in ensuring all citizens uniform access to health services throughout the country. In this respect, Italian regions represent a unique "natural laboratory" to test the theoretical predictions concerning the effects of decentralization policies, avoiding problems of cross-country heterogeneity and comparability.

The study draws on the previous literature analysing the effects of decentralization on various health outcomes (for a review: Channa and Faguet, 2016). In particular, numerous papers have explored the issue of decentralization in Italy, focusing on its effects in general (Tediosi et al., 2009; Ferrè et al., 2012) or in terms of efficiency of healthcare policies (Bordignon and Turati, 2009; Porcelli, 2014), and on its distributional consequences (Toth, 2014; Di Novi et al., 2015).

In the light of this literature, the contribution of our paper is twofold. First, provided that a quasi-experimental setting is not available for Italy (i.e. no control group exists), we decide not to model the decentralization intervention through discrete variables (e.g. reform dummies) as in other works (Porcelli, 2014). Contrarily, continuous variables are preferred, which allow to account for the evolution of the degree of fiscal decentralization over time. In doing this, we try to overcome the well-known shortcomings of the usual indicators, mainly accounting for the fact that tax autonomy does not always correlate with expenditure autonomy. Specifically, we rely on two different fiscal indicators: the ratio of tax revenues controlled by the regional government (RG) to total tax revenues and the ratio of transfers from the State to regions to total regional expenditures. While the former is intended to capture the degree of regional decision-making power in allocating tax revenues, the latter measures the degree of regional vertical fiscal imbalance (VFI), that is, the extent to which RGs rely on fiscal transfers from the central level.

Second, we also investigate whether the impact of decentralization on health outcomes varies according to the level of local wealth. From this point of view, our work relates to the strand of research analysing the distributional effects of decentralization across heterogeneous jurisdictions (Galiani et al., 2008; Caldeira et al., 2014; Soto et al., 2012). Opposite conjectures are theoretically possible. On the one hand, benefits from decentralization in terms of improved services may leave poorer regions behind as citizens can lack the ability to voice and support their preferences. On the other hand, deprived regions may be more incentivized to use their fewer resources in accordance to local needs. The issue is particularly relevant for Italy, where marked income disparities exist among regions.

Our analysis suggests that, on average, decentralization is positively associated with a reduction in IMRs, especially in the long run. However, its impact is heterogeneous between poor and non poor regions with the latter being favoured.

The remainder of the paper is organized as follows. The next section reviews the relevant literature on the relationship between fiscal decentralization and health outcomes. The Italian process of healthcare decentralization is briefly described in Section 3. Section 4 presents our data and empirical strategy, while Section 5 reports and discusses the results. Robustness tests for our findings are provided in Section 6. Finally, some conclusions are drawn.

2. Previous literature

A growing number of studies have investigated the relationship between decentralization and health outcomes. Here, we limit the discussion to those that have explicitly modelled the degree of decentralization using fiscal data. Existing evidence can be grouped into two broad categories: single-country and cross-country case studies. Table 1 summarizes the methodological approaches and the main results of this literature.

Single-country analyses have considered different institutional contexts such as Argentina (Habibi et al., 2003), India (Asfaw et al., 2007), Spain (Cantarero and Pascual, 2008), China (Uchimura and Jütting, 2009), Nigeria (Akpan, 2011), Canada (Jiménez Rubio, 2011a), Colombia (Soto et al. 2012). With the only exceptions of Robalino, Picazo and Voetberg (2001) and Jiménez-Rubio (2011b), all other cross-

country analyses have focused on low and middle income economies (Ebel and Yilmaz, 2001; Khaleghian, 2004).

In all the reported studies, health outcomes are measured by objective indicators, without explicitly controlling for quality. In nine out of eleven papers IMR is the dependent variable. Besides it, Cantarero and Pascual (2008) also use life expectancy (LE) to test the effects of decentralization in 15 Spanish regions over the period 1992-2003. Two works examine the impact of decentralization on immunization services. Using data for six developing countries from 1970 to 1999, Ebel and Yilmaz (2001) analyze decentralization for its effects on immunization against diphtheria, pertussis and tetanus (DPT) and measles for children under-12 months of age. Khaleghian (2004) investigates a similar relationship between decentralization and immunization coverage rates for the third vaccine DTP and measles in 1-year-old children for 140 countries with low and medium per capita incomes during the years 1980 to 1997. Indeed, due to its public good characteristics and externalities, immunization is an example of health services on which decentralization is expected to have a negative effect. This is likely to happen since the presence of shared benefits leads local authorities to free-ride on the provision of immunization programs. However, both papers are unable to reach definite conclusions.

Regarding the control variables, cross-country studies are generally constrained by limited availability of comparable data to using a reduced-form relationship between decentralization and health outcomes. In this respect, the assessment of the studied relationship at a single-country level allows to overcome this problem and to control better for unobserved heterogeneity between countries (e.g. institutional and cultural differences, differences in the quality of data, etc.), yielding firmer results.

The key point of all of these analyses remains, however, the choice of the measure of fiscal decentralization, which is likely to critically affect the study findings. Though the proper decentralization fiscal index is a highly debated issue in the literature (OECD/KIPF, 2013), a first difference is usually made between indicators on the revenue-side and on the expenditure-side. Two conventional measures are mostly employed: the ratio of local government revenues to total government revenues and the ratio of local government expenditures to total government expenditures. The first measure indicates the extent to which local governments are involved in mobilizing public resources through their system of taxes and user charges. Nonetheless, it has the limit of ignoring a possible greater responsibility of local governments for the delivery of goods and services financed through external sources. This kind of public activities is better accounted for when the expenditure-side indicator is used. However, the latter suffers from the fact that local governments acting just as spending agents of the CG are not always fully fiscal autonomous (i.e. tax sharing exists).

As shown in Table 1, only three papers have controlled for the shortcomings of the conventional indicators of decentralization. Specifically, a study by Habibi et al. (2003) on the relationship between decentralization and human capital development in Argentina during the period 1970-94 considers the proportion of revenue raised locally as well as the proportion of controlled revenue over the total. More recently, Jimenez-Rubio (2011b) uses both a conventional and a new non-conventional measure of revenue decentralization to investigate the effects of a higher decision-making autonomy on IMRs in 20 developed OECD countries over the period 1970-2001.

Table 1

Previous studies on the impact of fiscal decentralization on health outcomes (chronological order).

Study	Country/ies	Methodology	Variables	Decentralization Index	Results/conclusions
			Single-country studies		
Habibi *et al.* (2003)	Argentina (23 provinces; period: 1970-1994)	OLS with robust SEs, GLS with heteroskedastic and autocorrelated errors and fixed effects	*Dependent variable/s*: students enrolled in secondary school per 1,000 primary students and IMR; *Independent variables:* Two decentralization indicators: the ratio of royalties to controlled revenue, the ratio of conditional transfers from the centre to controlled revenue; per capita provincial GDP, public employees per 1,000 population, total per capita provincial expenditure	Proportion of revenue raised locally and proportion of controlled revenue over the total	The two decentralization indicators are associated positively with educational output and negatively with infant mortality. Comparing decentralization patterns across low-income and high-income provinces, disparities in regional IMRs decline significantly over the period after decentralization reforms are undertaken
Asfaw *et al.* (2007)	India (14 major states; Period: 1990-1997)	Fixed and random effects	*Dependent variable/s*: rural IMR *Independent variables*: fiscal decentralization index, index of political decentralization constructed from a number of variables using factor analysis (total voters' turnout, women's participation in polls, number of polling stations per electors in each state), per capita state GDP, women literacy	Fiscal decentralization index obtained by factor analysis on the basis of three variables: the share of local (rural) expenditure on the total state expenditure, the total local expenditure per rural population, the share of local own revenue from the total local expenditures	Fiscal decentralization reduces IMRs but its effectiveness increases with the level of political decentralization
Cantarero and Pascual (2008)	Spain (15 regions; period: 1992-2003)	Fixed and random effects	*Dependent variable/s*: IMR and LE *Independent variables*: fiscal decentralization indicator, per capita GDP, acute care beds per 1,000 population, general practitioners density per 1,000 population	Ratio of sub-national healthcare expenditure to total health expenditure for all the levels of government	Fiscal decentralization is negatively related to IMRs and positively related to LE
Uchimura and Jütting (2009)	China (26 provinces; period: 1995-2001)	Fixed effects with White corrected SEs	*Dependent variable/s*: IMR *Independent variables:* two fiscal decentralization indicators, per capita provincial GDP, rural/urban ratio, provincial birth and illiteracy rates, provincial government size	Ratio of counties' expenditure aggregated at the provincial level to counties' own revenue aggregated at the provincial level (as a measure of vertical balance) and ratio of county's aggregate expenditure to total provincial expenditure	More decentralized provinces perform better with respect to health outcomes if two conditions are met: 1) a functioning transfer system is established between the province and county levels; 2) county governments' own fiscal capacity is strengthened

Akpan (2011)	Nigeria (36 states and federal capital territory; period: 2002-2009)	Random effects	*Dependent variable/s:* IMR *Independent variables:* fiscal decentralization index, adult literacy rate, state population growth, state own revenue	Ratio of state revenue to federal revenue	Higher degrees of fiscal decentralization are consistently associated with lower mortality rates
Jiménez-Rubio (2011a)	Canada (10 provinces: period: 1979-1995)	Fixed effects with White corrected SEs	*Dependent variable/s*: IMR *Independent variables:* healthcare decentralization index, provincial per capita GDP, per capita healthcare block grants from the federal government, per capita federal expenditure in healthcare, per capita municipal healthcare expenditure, per capita private healthcare expenditure, educational level, female prevalence of daily smoking, low birth weight	Ratio of provincial healthcare expenditure to total healthcare expenditure	Decentralization has a substantial positive influence on the effectiveness of public policy in improving population's health (as measured by infant mortality)
Soto, Farfan and Lorant (2012)	Colombia (1080 municipalities; period: 1998–2007)	Fixed effects with year dummies and White corrected SEs	*Dependent variable/s:* IMR *Independent variables:* fiscal decentralization index, transfers from CG as a proportion of total health expenditure, sum of years of municipality with certification status, percentage of households with unsatisfied basic needs, level of urbanization	Locally controlled health expenditure as the proportion of total health expenditure	Decentralization decreases IMRs. However, the improved health outcome effects depend on localities' socio-economic conditions: they are greater in non-poor municipalities than in poor ones

Cross-country studies

Ebel and Yilmaz (2001)	Six developing countries (Argentina, Brazil, Colombia, Philippines, South Africa and Venezuela); period: 1970-1999	Intervention analysis based on a GLS estimation model with fixed effects	*Dependent variable/s:* immunization rate for DPT (diphtheria, pertussis and tetanus) and measles of children under 12 months. *Independent variables:* one-year lagged immunization rate, intervention variable	Decentralization is measured by an intervention variable that takes on the value of zero prior to and unity after intervention	Decentralization intervention by SNGs is statistically and positively associated only with immunization rates for measles
Robalino, Picazo, and Voetberg (2001)	Between 45 and 70 low and high income countries; period: 1970-1995	Fixed effects	*Dependent variable/s:* IMR *Independent variables:* decentralization index, per capita GDP, corruption, ethno-linguistic fractionalization (i.e. probability that two individuals picked at random in a given country belong to different ethnic groups)	Ratio between total expenditures of SNGs and total expenditures of the CG	Decentralization is associated with lower IMRs. The marginal benefits from fiscal decentralization is noticeably higher for poor countries
Khaleghian (2004)	140 low and middle income countries; period: 1980-97	Standard OLS multiple regression with year dummies and White corrected SEs	*Dependent variable/s:* immunization coverage rates for DPT3 (the third of a three-vaccine series against diphtheria, pertussis and tetanus) and measles of children at one year of age *Independent variables*: two decentralization indexes, per capita GDP, population size, population density, illiteracy rate, democracy score, participation in the UNICEF's Vaccine Independence Initiative, institutional quality index, ethnic tension score	Binary variable defined as the presence of taxing, spending, or regulatory authority on the part of sub-national authorities; a combination of two variables: the share of sub-national expenditures on total government expenditures and the share of health spending on total sub-national expenditures	The relationship between decentralization and immunization coverage is positive in low-income countries while negative in middle-income ones. Pathways of effects are also different in each country
Jiménez-Rubio (2011b)	20 developed OECD Countries; period: 1970-2001	Error Correction Model with both time and year dummies and Newey-West corrected SEs	*Dependent variable/s:* IMR *Independent variables:* decentralization index, percentage of total healthcare expenditure on GDP, per capita GDP, alcohol and tobacco consumption, educational level	A new measure of fiscal decentralization: the ratio of sub-national own tax revenue over general government total revenue (taxes in the numerator include only those where the SNGs can change the tax rate, the tax base or both). A conventional decentralization index: the ratio of sub-national tax revenue to general government total revenue	Fiscal decentralization has a considerable and positive long-term effect on reducing infant mortality only if a substantial degree of autonomy in the sources of revenue is devolved to local governments (i.e. the new measure of fiscal decentralization is used)

IMR = infant mortality rate; LE = life expectancy; GDP = gross domestic product; SE = standard error; OLS = ordinary least squares; GLS = generalized least squares.

Following the tax classification by Stegarescu (2005), the non-conventional measure is computed as the share of local government taxes (including only those where the local government controls the tax rate, the tax base or both) over the general government taxes. By comparing the results obtained with the two measures, the author concludes that decentralization has a considerable and positive long-term effect on reducing infant mortality (IM) only if it entails a substantial degree of autonomy in the sources of revenue as measured by the "new" indicator.

On the expenditure side, by examining the effects of decentralization on IMRs in Columbia during a 10-year period, Soto et al. (2012) employ the share of locally controlled health expenditures over total health expenditures, thus accounting for that part of health spending that is not financed internally but is managed by the local government. Apart from Akpan (2011) that opts for a standard revenue-side measure of decentralization, all the other studies rely on conventional expenditure-side indicators. Among these, healthcare related measures are highly preferred to overall ones. Generally, the use of overall indicators (not health-related) is considered to be inappropriate as countries differ in the types of expenditure that are decentralized. Hence, an identification problem of the relationship between health decentralization and outcomes may arise. However, such a problem is more relevant for cross-country analyses than for single-country ones.

Despite the existing methodological differences, the revised literature generally agrees on the beneficial impact of decentralization on health outcomes. However, decentralization *per sé* does not seem to be an enough powerful mechanism to enhance population health. A series of conditions should also be met, which include, among others, the quality of the local institutional context and other local socio-economic characteristics. Above all, the level of local development plays an important role in

explaining country differences in pathways of decentralization effects. Using a panel data of low and high income countries during the period 1970-1995, Robalino et al. (2001) find that the curve of the benefits associated with fiscal decentralization have a U shape with respect to GDP per capita, implying that low and high income countries are more likely to take advantage from fiscal decentralization reforms than middle income ones. In the already cited paper by Khalegian (2004) the benefit curve of fiscal decentralization is found to be L shaped for immunization: after a per capita GDP of 1,400 (1995 USD) the negative relationship stabilizes.

3. Institutional background

The Italian NHS (*Sistema Sanitario Nazionale)* was established in 1978 to guarantee uniform and comprehensive care to all citizens throughout the country. The system was initially funded through general taxation by the CG but a set of reforms has been progressively undertaken to assign responsibilities for the financing and the delivery of healthcare to regions (France et al., 2005). Together with the fact that the process is not yet completed, the analysis of the inter-governmental relationships for healthcare is further complicated by the existence of a complex network of political and institutional rules (Piacenza and Turati, 2014). Moreover, the Italian trend towards healthcare decentralization does not emerge as a linear one. Rather, it resembles a somewhat contradictory stop-and-go process where increasing powers to regions have gone together with a significant role retained by the CG in regulating the system, in governing health expenditure as well as in maintaining inter-regional solidarity (Tediosi et al., 2009).

The country is divided into 20 regions: 15 ordinary statute regions (OSRs) and 5 special statute regions (SSRs), one of which is further divided into 2 autonomous Provinces. Regions differ markedly in terms of socio-demographic, economic, structural and institutional characteristics with a clear-cut North-South dualism (Toth, 2014). Since their establishment, SSRs have enjoyed a higher degree of fiscal (they could retain revenues from main national taxes) and legislative autonomy. However, before the regionalization process, this wide autonomy did not extend to the healthcare system, whose financing and delivering was managed directly by the CG. Therefore, both OSRs and SSRs have experienced the 90's decentralization reforms, though with some differences. Along with the constitutional mandate to ensure uniform health services throughout the country that induces the CG to equalise per capita resources across regions, the above considerations help to explain why per-capita public health expenditures are very similar over time in both groups of regions (Figure 1).
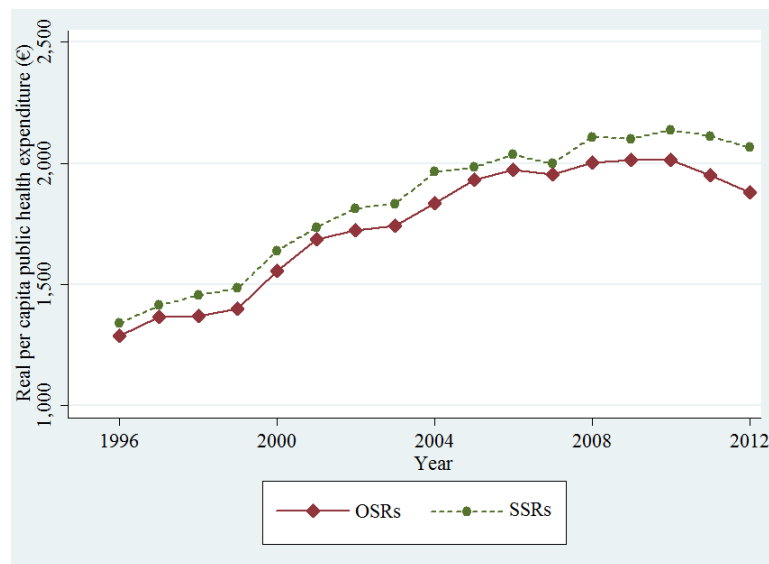
Fig. 1. Evolution of real per capita public health expenditure in Italy. Source: Health for All.

The regionalization of healthcare started in the beginning of the 90's when a wave of reforms was introduced with the threefold aim of enhancing efficiency within the health system, creating an internal market for health services and increasing the autonomy of regions in planning, organizing and financing healthcare in their own territory. Regions were, thus, entitled to decide on different organizational aspects of their health systems, including the number and size of the local healthcare authorities, the level of integration between local authorities and autonomous hospital trusts, the involvement of private providers (Jommi et al., 2001).

In 1998 a process of fiscal decentralization took place when previous inter-governmental grants earmarked for the health sector were replaced by two regional taxes: 1) a value added tax on productive activities (*Imposta Regionale sulle Attività Produttive,* IRAP), earmarked for 9/10 to finance health expenditure; 2) a surcharge on the national personal income tax (*Addizionale IRPEF*). The rationale for the reform was to reduce the extent of VFI so as to make regions more accountable. These could now choose how much to rely on either regional taxes or patient co-payments for financing healthcare for their population, and whether or not to modify (within a defined range) regional tax rates compared with national standards.

Since the tax bases of both IRAP and IRPEF is positively related to the per capita GDP which varies greatly among Italian regions, a distributional issue arises, strongly calling for an equalizing transfer scheme. With the legislative decree n. 56/2000 an inter-regional incomplete (i.e. 90% solidarity coefficient) equalization fund was defined, financed by a revenue sharing on the Value Added Tax (VAT) and on a petrol tax. The fund was intended to redistribute financing to regions on the basis of geographic and population size, healthcare needs and fiscal capacity. Although the new allocation mechanism was expected to come into force gradually, it was never applied as CG and

regions failed to agree on the funding arrangements. Therefore, the previous equalization method continues to be applied, which implies a yearly negotiation between CG and RGs to define the total amount of public resources assigned to the NHS. These are, then, allocated among regions according to a formula based on regional expenditure needs indicators (i.e. population size weighted by age, gender and epidemiological indicators) (Ferrario and Zanardi, 2011).

The devolution of political and fiscal powers to RGs has been further strengthened by the Constitutional amendment of 2001 (Constitutional law n. 3/2001), which introduced the definition of essential levels of care (*Livelli Essenziali di Assistenza,* LEAs) that must be guaranteed over the entire national territory. Responsibilities for setting and ensuring the general objectives and fundamental principles of the system are maintained on the central level. In particular, the CG is in charge of using equalizing transfers to top up regional own resources to fully cover the expenditure standards. Regions are responsible for ensuring the delivery of LEAs but are free to administer, organize and finance it in accordance with their population needs. They can also offer additional health services over the LEAs, provided that they finance them with their own resources.

Table 2 shows the structure of the NHS funding and its evolution over the period 2001-2012. With respect to this national picture, marked variability exists in the funding composition for healthcare across regions. In northern regions the share of revenues from both and the surcharge on IRPEF accounts for about three times that of southern regions (Turati, 2013). On the opposite, the latter rely more heavily on equalizing transfers (VAT and excise on petrol) from the CG.

Table 2.

NHS funding structure and deficits (millions of Euros and percentages). Source: Data from Tediosi et al. (2009) and Ministry of Health.

	Years		
	2001	2006	2012
Financing	71,878 (100.00%)	95,129 (100.00%)	112,641 (100.00%)
IRAP and surcharge on IRPEF	30,295 (42.15%)	37,280 (39.19%)	39,902 (35.42%)
VAT and excise on petrol	27,288 (37.96%)	41,932 (44.08%)	52,969 (47.02%)
Other transfers from central state and private sector	12,029 (16.74%)	13,179 (13.85%)	15,018 (13.33%)
Own sources of local healthcare units	2,266 (3.15%)	2,739 (2.88%)	3,077 (2.73%)
Other			1,675 (1.49%)
Deficits	-4,121 (5.73%)	-4,519 (4.75%)	-1,043 (0.93%)

Note: the heading "other" includes "capitalized costs" and "adjustments and funding use", introduced since 2012.

From its inception, the fiscal decentralization process was not straightforward. The power of regions to raise the surcharges on IRPEF and IRAP was suspended in 2003 and 2004 with the aim to contain the global fiscal pressure. Moreover, during the first years of fiscal decentralization (2001-2005), the CG partially bailed out the previous healthcare deficits of the regions. Therefore, expectations of future bailing outs weakened the incentives for regions to be financially accountable, thus resulting in increasing deficits (Bordignon and Turati, 2009). In order to prevent these, legislation was approved that introduced a new turnaround strategy for regions with deficits (Ferrè et al., 2012). The Financial Stability Law 2004 (L. 311/2004) and the Health Pact signed with regions in 2006 (further enforced by the Financial Stability Law n. 296/2006) regulated the main aspects of this strategy. The allocation of special national funds to the regions with high deficits (i.e. higher than 7% of the funding) was subjected to the sign of specific agreements (the so-called "repayment plans") to restore the financial stability of their health systems. Typical measures of a repayment plan included the restructuring of the public hospitals network, the total block on staff

turnover and the automatic increase of the regional tax rates to the maximum allowed level. Disciplinary consequences were provided for non-compliance, which enabled increased CG's interference in regional autonomy (e.g. the formal replacement of the President of the region by an *ad acta* commissioner). Overall, ten regions were exposed to repayment plans, eight of which still continue to be (Table 3).

From the above description it follows that the Italian decentralization process for healthcare has proceeded by alternating phases of acceleration and deceleration. This fact, together with the impossibility to use a quasi-experimental setting due to the lack of a control group, induces us to employ continuous variables as proxies for regional decentralization, which are better suited to quantitatively capture the changes in the degree of the phenomenon over time.

Table 3.
Healthcare repayment plans. Source: Ministry of Health.

Regional repayment plans		
	Starting date	**Status**
Lazio	February 28, 2007	Ongoing
Abruzzo	March 6, 2007	Ongoing
Liguria	March 6, 2007	Exit after 2007-2009
Campania	March 13, 2007	Ongoing
Molise	March 27, 2007	Ongoing
Sicilia	July 31, 2007	Ongoing
Sardegna	July 31, 2007	Exit after 2007-2009
Calabria	December 17, 2009	Ongoing
Piemonte	July 29, 2010	Ongoing
Puglia	November 29, 2010	Ongoing

4. Data and methods

We employ a balanced panel of 20 Italian regions over a 17-year period (1996-2012). To account for the different status of the Italian regions, we consider two samples, with and without the five SSRs (i.e. Friuli V.G., Sardegna, Sicilia, Trentino A.A. and Valle D'Aosta). Therefore, the number of available observations varies between 255 and 340, depending on the sample. If not differently specified, data are taken from the WHO Health For All database. *IMR*, the dependent variable, is used as a measure of health outcome. This indicator is generally assumed to be more reliable than other alternative indicators of population health status since it not only reflects both child's and pregnant women's health but it is also more sensitive to policy reforms, such as health decentralization (Jimenez-Rubio, 2011b). Particularly, three different aspects make IMR superior to LE (Porcelli, 2014): the relationship between the event and the characteristics of the regional health system where it occurred is more straightforward (less affected by spillover effects); the short-run changes of the healthcare system are better captured; biases due to statistical manipulation are less.

Figure 2 shows the falling trend in Italian IMRs over the study period, separately for OSRs and SSRs. Overall, the indicator decreases from 57.1 per 10,000 births in 1996 to less than 29.3 in 2012 (-48.7%). A decline is displayed in Figure 3 with regard to each of the 20 regions. However, considerable differences exist across regions: in 2012, IMRs ranged from a minimum of 16.7 per 10,000 births in the Region Marche to a maximum of 45.2 in the Region Calabria, with a clear gradient moving from the North (29.6) to the South (39.0) of Italy.
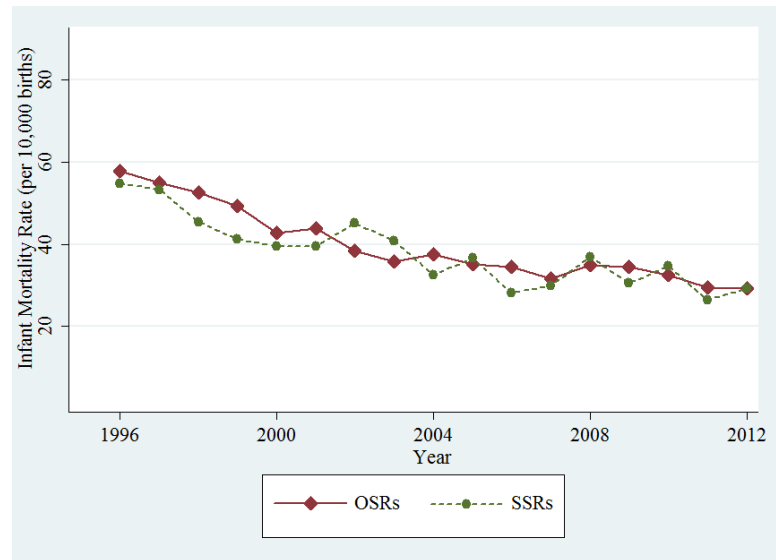
Fig. 2. Evolution of infant mortality rates in Italy.

**Fig. 3.** Evolution of infant mortality rates by region.

To measure the degree of regional decentralization, we employ two different indicators based on fiscal data from the Territorial Public Accounts (*Conti Pubblici Territoriali*, TPA). These are produced by the Italian Ministry of Economy and provide

the allocation of revenues and expenditure flows collected/paid by each level of government (central, regional, local) among the 20 Italian regions. Revenue flows are regionalized according to where the resources are collected while expenditure flows are allocated to the region where the means of production for public services or investments are located. Following Grisorio and Prota (2015a,b), the first indicator (Fiscal Decentralization, *FDEC*) considers the ratio of tax revenues raised and/or controlled by the RG to the total tax revenues collected in the region. Specifically, the numerator of the ratio is represented by the sum of the regional own-source revenues ("entrate proprie") and the transfers from other governments ("tributi devoluti da altre amministrazioni"). The denominator is the same sum but with regard to central, regional and local governments. Hence, by construction, the *FDEC* indicator is intended to capture the degree of regional decision-making autonomy in the allocation of tax revenues. Other things being equal, an increase in *FDEC* is expected to have a positive impact on health outcomes, thus reducing IM. The preference for a measure of decentralization from the revenue-side is motivated by the fact that in Italy the financing side is mostly affected by the ongoing decentralization process. Moreover, the use of an overall measure of decentralization that is not health-related is consistent with the consideration that the revenue of the main regional tax (i.e. IRAP) is earmarked to finance healthcare expenditure.

The second indicator (Vertical Fiscal Imbalance, *VFI*) measures the degree to which regions are dependent from CG revenues to support their expenditure levels. It is computed as the share of government transfers to a region over the total regional expenditures. From a political economy perspective, a high degree of reliance on transfers from the CG is likely to have adverse effects on the size and the efficiency of local choices, as decision-makers are less accountable to local voters and more prone to

rent-seeking behaviours. On the contrary, a greater reliance on own taxes is expected to lessen "common pool" problems such as fiscal illusion and "flypaper effects" (Gramlich, 1977; Fisher, 1982). However, in this "common pool versus own resources" issue, the structure and the composition of local revenues are found to play an important role (Liberati and Sacchi, 2013).

Figure 4 shows an increased trend of our variable FDEC during the period 1996-2012 for both OSRs and SSRs, though more pronounced for the former (from 3% in 1996 to 21% in 2012) than for the latter (from 30% in 1996 to 38% in 2012). The growing tendency becomes more evident after 2001, when the regional VAT sharing was introduced. This is especially true for OSRs, which mostly experienced the reform changes. Figure 5 displays a high level of variability in the dynamics of regional trends.
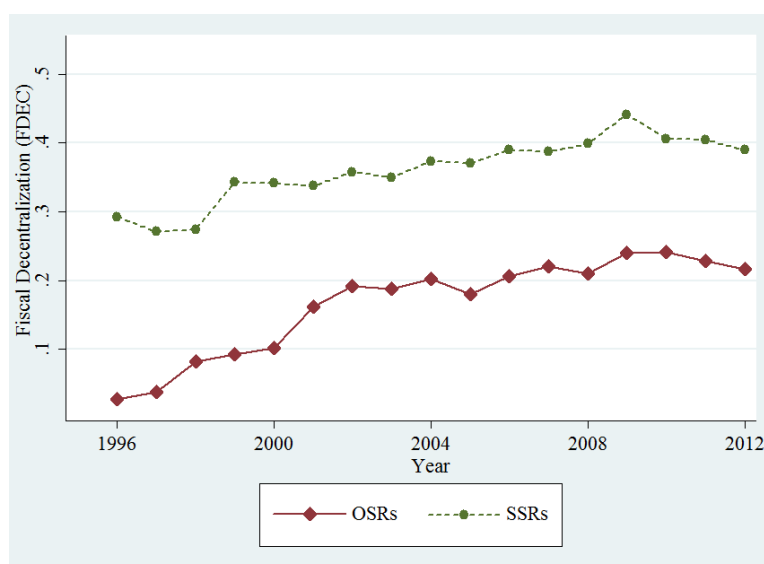
Fig. 4. Evolution of fiscal decentralization (*FDEC*) in
Italy.
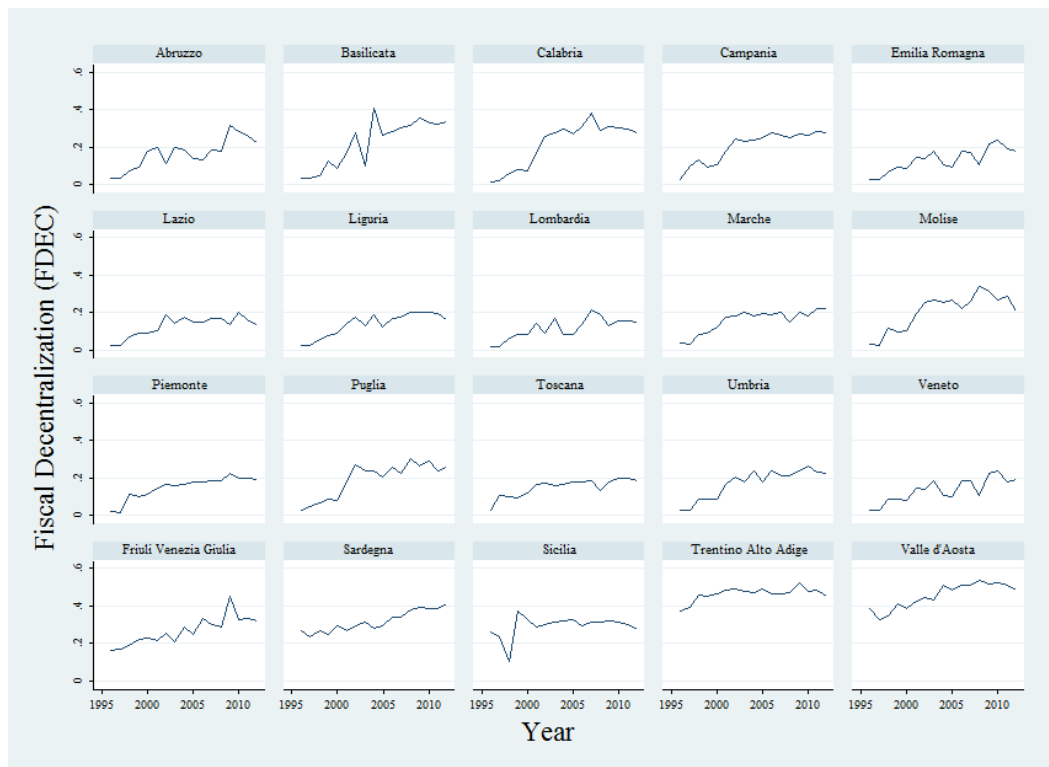
Fig. 5. Evolution of fiscal decentralization (*FDEC*) by region.

With regard to the *VFI* variable, Figure 6 shows a consistent decrease over the study period, more marked for OSRs because of their limited statute autonomy. For OSRs, the value of *VFI* decreases rapidly since the 1998 tax reform, ranging from over 0.9 in 1996 to less than 0.14 in 2012 (-84%), when it approaches the value of SSRs. Once again, important differences exist across regions (Figure 7).
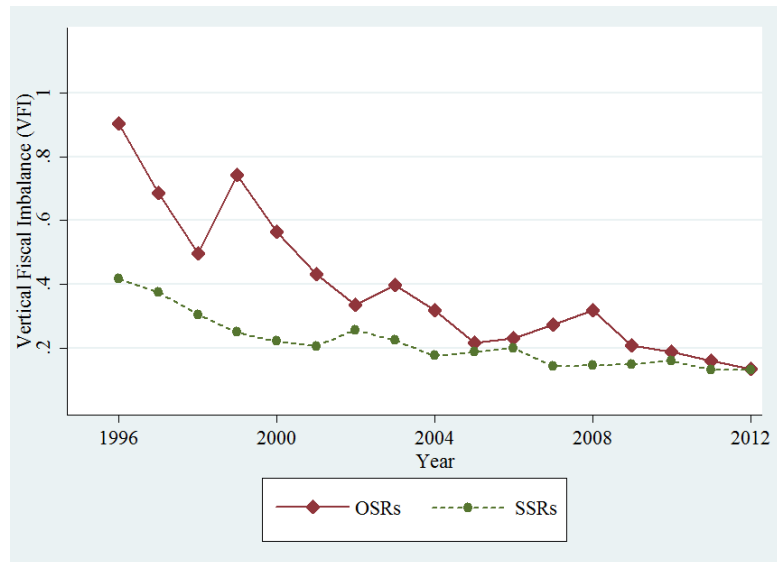
Fig. 6. Evolution of Vertical Fiscal Imbalance (*VFI*) in Italy.
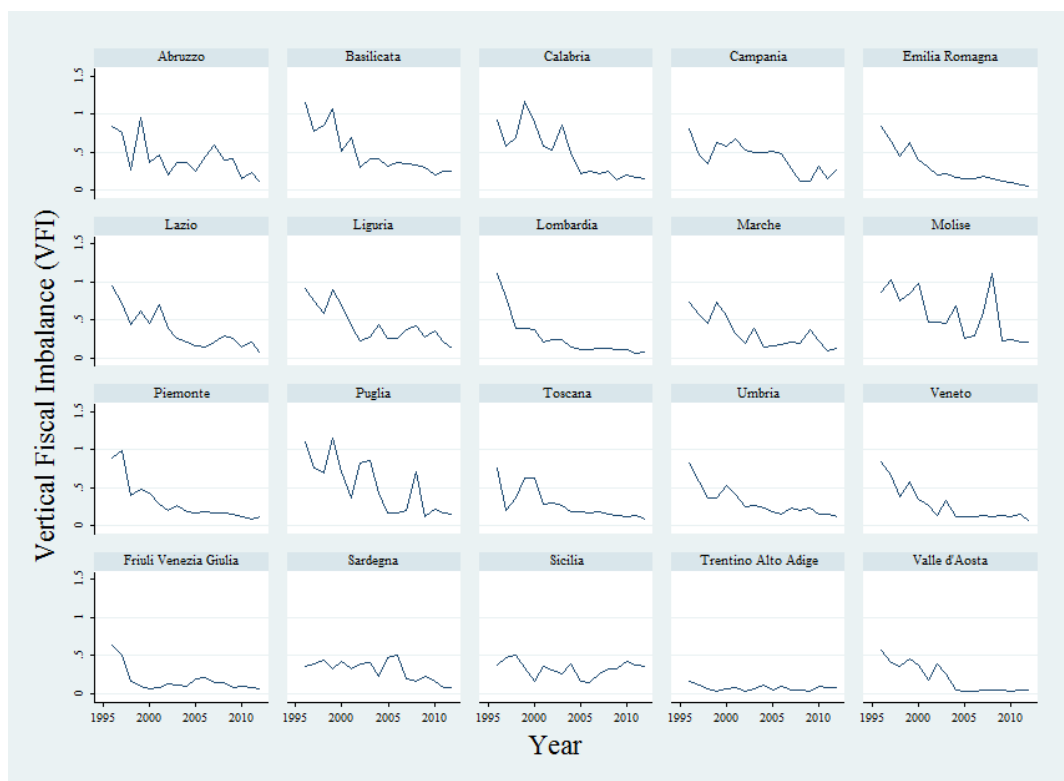
Fig. 7. Evolution of Vertical Fiscal Imbalance (*VFI*) by region.

To empirically assess the impact of decentralization on health outcomes in Italy, the following general specification is applied:

$$IMR_{it} = \alpha + \beta DEC_{it} + \delta Z_{it} + \epsilon_{it}$$ Equation (1)

where *IMR* denotes the infant mortality rate measured as the number of deaths of children aged under one year per 10,000 live births, *DEC* is our fiscal decentralization indicator (either *FDEC* or *VFI*), *Z* represents a vector of control (environmental) variables, *ε* is the disturbance term, *i* indicates region (*i* = 1, …,15 or 20, depending on the sample) and *t* year (*t* = 1996,…, 2012).

In selecting the control variables, we have adopted a quite parsimonious approach, including only medical and non-medical regional characteristics (Jiménez-Rubio, 2011a), such as the income level measured by the gross domestic product per capita (at 2012 prices, *GDP*), the share of total expenditures devoted to healthcare (*HEALTH_EXP*) as a proxy of the level of medical care inputs, the level of female education (*EDUC*) and the consumption of tobacco (*SMOKE*) as life-style indicators. In particular, *GDP* allows controlling for differences in both living conditions and the size of tax bases across regions; *HEALTH_EXP* is expected to have a negative impact on our dependent variable if an increased percentage of resources employed in the health sector is associated with improvements in the quality and/or the levels of healthcare services, *ceteris paribus*. The variable is lagged one year to deal with the potential endogeneity problem arising from reverse causality: IM is supposed to be affected by the share of healthcare expenditure but the RGs could also decide on the allocation of their budget to healthcare based on the current or expected level of child deaths. To account for restrictions in regional decision-making autonomy imposed by the repayment plans, we use a dummy variable (*PLAN*), which takes the value 1 in year *t* if a region is exposed to

a repayment strategy and 0 otherwise. Descriptive statistics for all the variables included in the analysis are provided in Table 4.

Table 4
Descriptive statistics for the selected variables.

Variable	Source		15 Ordinary statute regions				All 20 regions			
			Mean	SD	Min	Max	Mean	SD.	Min	Max
IMR (Infant mortality rate expressed per 10,000 births)	HFA	Overall	39.68	13.04	2.06	83.26	39.24	13.73	2.06	90
		Between		7.36	30.80	53.85		8.21	26.22	56.54
		Within		10.92	-5.30	75.90		11.15	-5.74	75.46
FDEC (Ratio of tax revenues of the regional government to total tax revenues)	TPA	Overall	0.17	0.08	0.01	0.41	0.21	0.12	0.01	0.54
		Between		0.03	0.12	0.22		0.10	0.12	0.46
		Within		0.08	-0.03	0.35		0.07	0.01	0.40
VFI (Transfers from central government to regions as a share of total regional expenditures)	TPA	Overall	0.39	0.27	0.06	1.17	0.34	0.26	0.03	1.17
		Between		0.10	0.27	0.57		0.12	0.07	0.57
		Within		0.57	-0.01	1.22		0.23	-0.05	1.18
GDP (GDP per capita - 2012 PPP)	HFA	Overall	26,074	6,247	14,685	37,013	26,582	6,712	14,685	38,582
		Between		6,316	17,079	35,327		6,756	17,079	36,385
		Within		1,280	22,627	28,575		1,251	23,135	29,222
HEALTH_EXP (Regional health expenditure as a share of regional total expenditure)	ISTAT	Overall	0.75	0.08	0.52	0.90	0.67	0.17	0.13	0.90
		Between		0.07	0.59	0.83		0.17	0.21	0.83
		Within		0.04	0.64	0.86		0.04	0.50	0.82
EDUC (Percentage of female population aged 6 and over with at least an upper secondary school degree)	HFA	Overall	35.90	5.83	23.35	50.30	35.75	5.79	23.35	50.30
		Between		3.23	28.76	42.52		3.14	28.76	42.52
		Within		4.91	26.71	50.62		4.91	26.56	50.47
SMOKE (Percentage of population who are daily smokers)	HFA	Overall	23.10	2.29	17.18	30.68	22.83	2.34	15.64	30.68
		Between		1.76	20.43	26.80		1.74	20.43	26.80
		Within		1.52	19.30	27.33		1.61	17.71	30.21
PLAN (Exposition to a repayment plan)	HFA	Overall	0.14	0.34	0	1	0.13	0.34	0	1
		Between		0.15	0	0.35		0.15	0	0.35
		Within		0.31	-0.22	1.02		0.3	-0.22	1.01
FDECxGDP	HFA	Overall	4,192	1,911	257	7,893	5,799	4,181	257	20,492
		Between		336	3,584	4,581		3,824	3,584	16,656
		Within		1,883	4.32	7,930		1,883	333	11,158
VFIxGDP	HFA	Overall	9,503	6,229	1,870	36,962	8,472	6,032	1,011	36,962
		Between		1,423	7,418	12,337		2,336	2,613	12,337
		Within		6,074	253	36,757		5,584	-779	35,726

HFA = Health For All-Italy; TPA = Territorial Public Accounts (*Conti Pubblici Territoriali*).

We consider the following two models:

$$IMR_{it} = f(FDEC_{it}, GDP_{it}, HEALTH\_EXP_{it-1}, EDUC_{it}, SMOKE_{it}, PLAN_{it}, TIME_{it})$$

Model (1)

$$IMR_{it} = f(VFI_{it}, GDP_{it}, HEALTH\_EXP_{it-1}, EDUC_{it}, SMOKE_{it}, PLAN_{it}, TIME_{it})$$

Model (2)

In both models, we include a linear time trend (*TIME*) to control for technological advances. The above models are estimated for each of our two samples (i.e. 15 OSRs and all the 20 Italian regions). Moreover, since the effect of decentralization on *IMR* may differ according to the level of regional *GDP*, we also augmented each of the two models with the interaction term between the decentralization indicator (either *FDEC* or *VFI)* and *GDP*.

Before model estimation, we conduct necessary tests. First, we perform the Im-Pesaran-Shin test (Im, Pesaran and Shin, 2003) for unit roots in panel datasets, which rejects the null hypothesis of non-stationarity for our key variables in levels. Secondly, we run a Breush-Pagan Lagrange Multiplier test, which confirms significant differences across regions (*p*-value=0.000 for different models, specifications and samples). Consequently, a simple pooled OLS estimation is not an appropriate solution and fixed effects (FEs) and random effects (REs) models are standard alternatives. In our analysis, though the Hausman test is not always significant, there is a strong theoretical preference for FEs: as we are dealing with regions, it is very likely that there are unobserved time-invariant region-specific effects that can be controlled for using FEs models. Furthermore, the FEs approach is usually appropriate when the data exhausts the population, which is our case. However, the results with REs (available on request)

are broadly similar to those with FEs. Moreover, the Wald test confirms that coefficients of regional dummies are jointly significant (*p*-value=0.000).

As in our dataset the cross-sectional dimension of the panel is larger than the time-series one, heteroskedasticity could be a possible problem in estimates. In particular, *IMR* might exhibit a different variability according to the degree of regional decentralization, eventually implying heteroskedastic residuals. Therefore, we apply the modified Wald statistic for groupwise heteroskedasticity in FEs models, which strongly rejects the null hypothesis of homoskedasticity for both models, all samples and specifications (*p*-value=0.000). Consequently, for all estimates, we provide robust standard errors (SEs). As serial correlation of the error term biases the SEs and causes the results to be less efficient, we also run the Wooldridge test for panel-data models. The test results (*p*-value>0.1) strongly fail to reject the null hypothesis of no first order autocorrelation.

5. Results and discussion

Tables 5 and 6 report our regression results, respectively, for model (1) and model (2), using FEs with robust SEs. Overall, these are quite robust and consistent with our prior expectations and rather stable across the different models, samples and specifications. A good fit to the data is always shown, with an adjusted *R*-squared statistic of about 70%. Moreover, the *F*-tests indicate that the coefficients are always jointly significant (*p*-values=0.000). With regard to model (1), a statistically significant and negative coefficient for *FDEC* is always found, thus suggesting that, other things being equal, regions with a higher degree of decisional and accountable autonomy in the management of tax revenues tend to have a lower *IMR*. The coefficients of the variable

GDP are all negative and highly significant, meaning that higher standards of living as proxied by the level of per capita income are expected to reduce *IMR*.

In the specification with the interaction term, the positive and significant signs of *FDECxGDP* coefficients show the presence of a moderating effect of *GDP* on the relationship between regional decision-making autonomy in revenue allocation and *IMR*. Finally, time trends are negative and highly statistically significant.

Table 5
Estimation results (fixed effects): model (1).[a]

Regressors	Infant mortality							
	Excluding SSRs				All regions			
	Coef.	t	Coef.	t	Coef.	t	Coef.	t
FDEC	-23.62	-2.55**	-60.48	-2.66**	-19.26	-1.94*	-61.76	-3.40***
GDP	-0.0014	-3.16***	-0.0018	-3.55***	-0.0017	-3.79***	-0.0021	-4.12***
FDECxGDP			0.0019	2.02*			0.0022	2.75**
HEALTH_EXP	2.89	0.13	4.27	0.20	-23.80	-1.16	-19.53	-1.01
EDUC	0.014	0.04	-0.073	-0.21	0.29	0.63	0.18	0.44
SMOKE	-0.31	-1.06	-0.30	-1.03	-0.70	-0.28	-0.041	-0.16
PLAN	2.79	1.51	3.22	1.74	1.16	0.68	1.76	1.02
TIME	-1.43	-3.66***	-1.46	-4.22***	-1.48	-3.33***	-1.50	-3.69***
_cons	98.41	2.77**	108.38	3.23***	107.79	4.24***	118.02	4.96***
Adj R^2	0.71		0.71		0.68		0.68	
F test (p-value)	69.35 (0.000)		64.87 (0.000)		56.93 (0.000)		53.89 (0.000)	
N	255		255		340		340	
Regions	15		15		20		20	

[a] T statistics computed with robust SEs.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Looking at the estimation results for model (2), the *VFI* indicator shows a positive significant association with *IMR* when the interaction terms are considered, while the

coefficients of *GDP* are always highly significant and with the expected sign. As for *VFIxGDP*, the coefficients are negative and significant, suggesting that the positive association between the extent of transfer dependency and the level of regional *IMR* is moderated by *GDP*. Therefore, a rise in *VFI* is likely to increase *IMR* more in poorer regions than in richer ones. When SSRs are not considered in the sample and the interaction term is included in the model, the coefficient of the dummy *PLAN* is significant at a 5% level, indicating that a reduction in the regional decision-making autonomy following the submission of a repayment plan is associated with a higher *IMR*, other things being equal. Time trends are highly significant and with the expected signs.

Table 6
Estimation results (fixed effects): model (2).[a]

	Infant mortality							
	Excluding SSRs				All regions			
Regressors								
	Coef.	t	Coef.	t	Coef.	t	Coef.	t
VFI	4.33	1.72	25.97	3.33***	4.95	1.45	22.44	2.72**
GDP	-0.0016	-4.48***	-0.0014	-4.12***	-0.0018	-4.52***	-0.0016	-4.47***
VFIxGDP			-0.0009	-3.59***			-0.0008	-2.36**
HEALTH_EXP	5.53	0.24	8.83	0.41	-20.78	-1.04	-18.33	-0.94
EDUC	-0.026	-0.06	-0.21	-0.64	0.22	0.47	0.12	0.27
SMOKE	-0.39	-1.33	-0.28	-0.95	-0.14	-0.53	-0.024	-0.09
PLAN	2.84	1.73	3.80	2.14**	1.13	0.71	1.81	1.07
TIME	-1.54	-2.90**	-1.44	-3.56***	-1.48	-3.07***	-1.43	-3.37***
_cons	100.57	2.86**	95.60	3.28***	106.62	4.26***	101.30	4.61***
Adj R^2	0.70		0.72		0.68		0.68	
F test (p-value)	63.27 (0.000)		65.97 (0.000)		55.09 (0.000)		51.34 (0.000)	
N	255		255		340		340	
Regions	15		15		20		20	

[a] T statistics computed with robust SEs.

*Significant at 10%; **significant at 5%; ***significant at 1%.

To better disentangle the impact of regional wealth on our study relationships, in Figures 8 and 9 we employ the previous regression results with the interaction terms (Tables 5 and 6) to graphically illustrate the distributional effects of fiscal decentralization. Specifically, the predicted values for *IMR* are plotted against each of the two indicators of decentralization (*FDEC* and *VFI*), according to three different levels of *GDP* (low = mean-½ standard deviation; mean; and high = mean + ½ standard deviation). In doing this, all the other control variables are kept constant at their mean values.
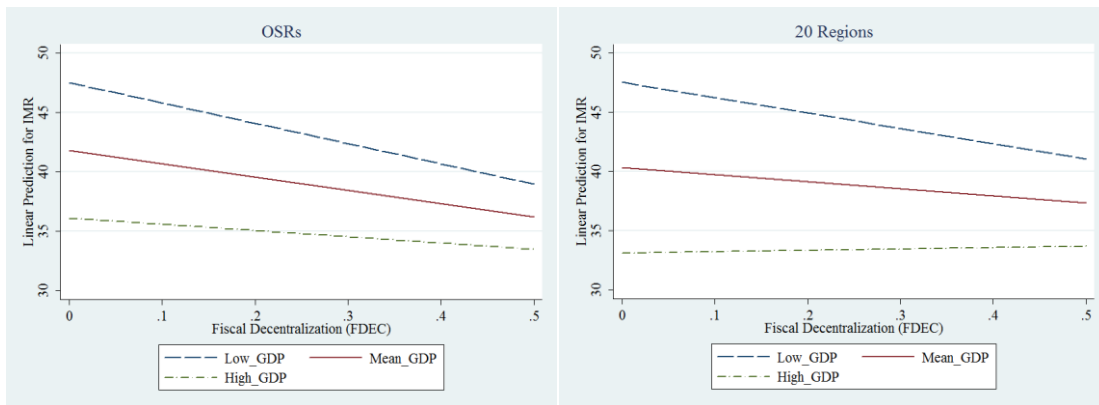
Fig. 8. Predictive margins for IMR - variable FDEC.
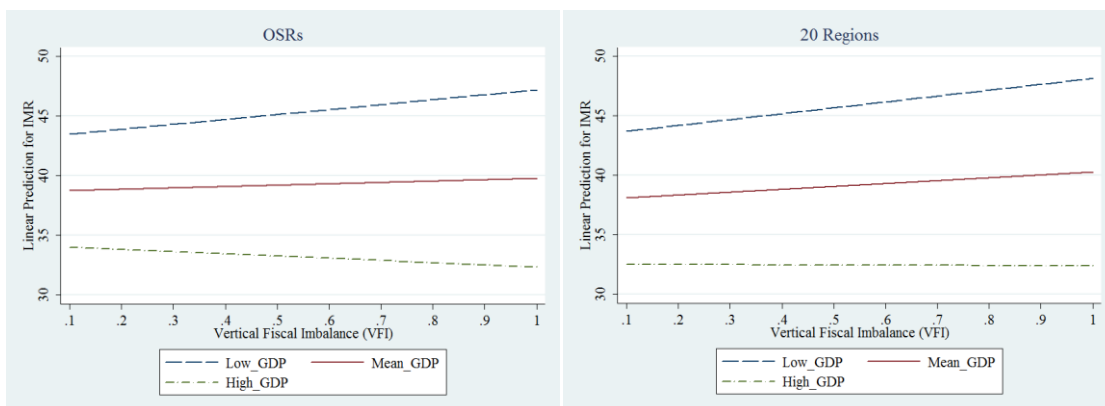
Fig. 9. Predictive margins for IMR - variable VFI.

For both samples in Figure 8, as *GDP* increases, the positive marginal effect of *FDEC* in reducing *IMR* decreases progressively to the extent of being not significantly different from zero for high levels of income. Regarding the variable *VFI* (Figure 9), the marginal benefit of a reduction of transfer dependency (that is, of a higher degree of tax autonomy) is higher for poorer regions while becomes not significantly different from zero at high GDP levels. In other words, regions with lower fiscal capacity seem to be penalized more than richer ones (in terms of rise of IMRs) by an increase in the degree of transfer dependency, other things being equal.

To sum up, findings from both the figures are consistent with each other and suggest that a higher degree of fiscal decentralization (or, conversely, a lower transfer dependency) is associated with a reduction of *IMR* that is significantly higher for less wealthy regions (Robalino et al., 2011). One possible explanation for this is that when IMRs are high, as in poor regions, these can be more easily lowered with well targeted healthcare interventions at local levels, all else being equal. On the opposite, it is more difficult to obtain the same beneficial effect from local healthcare interventions in rich regions where *IMR* exhibits low values, close to the best frontier. In this respect, a relative "late-comers advantage" in terms of marginal benefit from decentralization exists for less developed areas. An alternative explanation for the moderating effects of *GDP* is more outcome-specific and relies on the possibility that less developed regions may give to IM a greater priority among their health objectives, given the higher rate of child deaths.

The results confirm evidence from other studies that responses to decentralization are likely to vary when differences across jurisdictions are accounted for. With regard to the direction of these heterogeneous distributional responses, the majority of empirical researches have focused on developing countries where weaker institutions, higher

levels of corruption, lower levels of community participation and less ability to raise financial resources may prevent fiscal decentralization from exerting its beneficial effects. This literature agrees that decentralization mostly benefits the wealthier regions, leaving the poorer ones behind (Galiani et al, 2008; Caldeira et al., 2014; Soto et al., 2012). On the contrary, our analysis considers a developed country where a given degree of economic development has been already reached in all regions, though inter-regional differences still exist. In this regard, our results are more similar to those of Robalino et al. (2001), who found that the effect of fiscal decentralisation on IMRs was higher in poor-income countries than in middle-income countries. However, compared to this, our results should be regarded as more robust, being based on within-country analysis that makes them less vulnerable to heterogeneity problems.

6. Robustness tests

In this section, we examine the reliability and robustness of our previous results. First of all, we test for problems of spurious relationships. As previously outlined, we have applied the Im-Pesaran-Shin (2003) test to ensure unit roots in our heterogeneous panel data. However, this test, which assumes cross-sectional independence among panel units, is recognized to suffer from limited statistical power and is likely to yield biased results when applied to panel data with cross-sectional dependency. Therefore, for both our samples, we re-estimate all the models and specifications using a generalized one-equation error correction model (ECM) with regional fixed effects, which allows exploring the short- and long-run dynamics between *IMR* and our relevant explanatory variables. Equation (1) can be easily translated into our ECM form as:

$$\Delta IMR_{it} = \alpha + \gamma IMR_{it-1} + \beta_1 \Delta DEC_{it} + \beta_2 DEC_{it-1} + \delta_1 \Delta Z_{it} + \delta_2 Z_{it-1} + \epsilon_{it}$$

Equation (2)

where Δ is the first differences operator, β_1 provides the short-run response of our decentralization indicator (either *FDEC* or *VFI*), while its long-run response is computed as $\varphi_1 = -\frac{\beta_2}{\gamma}$.

Tables 7 and 8 report estimates from the ECMs, which widely confirm the robustness of our previous relationships in the long-run. It is worth noting that in both the tables we consider a linear time trend. The inclusion of a linear time trend within an ECM that regresses the first difference of the dependent variable can be considered analogous (from an observational point of view) to the inclusion of a quadratic time trend in our baseline regression model. In this respect, the above results can be also interpreted as a robustness check for the use of a different functional form for our time trend parameters. For completeness, we have also estimated ECMs without a time trend. The results, which are available on request by the authors, again confirm our previous predictions. Furthermore, with regard to our baseline regression models, the Hausman test always shows that the choice between a linear and a quadratic time trend specification does not imply systematic differences in the estimated coefficients (*p*-value>0.1).

Table 7
Estimation results (Error Correction Model - fixed effects): model (1).[a]

Regressors	Infant mortality							
	Excluding SSRs				All regions			
	Coef.	t	Coef.	t	Coef.	t	Coef.	t
IMR_{t-1}	-1.05	-11.23***	-1.06	-12.35***	-0.97	-10.70***	-0.99	-11.63***
$FDEC_{t-1}$	-42.77	-4.64***	-90.65	-3.23***	-30.00	-2.49**	-80.62	-3.29***
$\Delta FDEC$	-12.65	-1.82*	-3.18	-0.38	-13.52	-1.62	-5.53	-0.59
GDP_{t-1}	-0.0014	-3.01***	-0.002	-3.01***	-0.0014	-3.41***	-0.0021	-3.50***
ΔGDP	-0.0011	-1.34	-0.001	-1.23	-0.0012	-1.87*	-0.0011	-1.77*
$FDEC_{t-1} \times GDP_{t-1}$			0.0027	2.28**			0.0028	2.79**
$HEALTH\_EXP_{t-2}$	1.11	0.05	4.69	0.21	-36.87	-1.39	-27.78	-1.17
$\Delta HEALTH\_EXP_{t-1}$	-3.24	-0.15	1.77	0.09	-23.75	-1.23	-14.88	-0.89
$EDUC_{t-1}$	0.099	0.18	-0.045	-0.09	0.62	1.08	0.47	0.83
$\Delta EDUC$	-0.50	-0.93	-0.63	-1.22	-0.69	-1.42	-0.75	-1.58
$SMOKE_{t-1}$	-0.86	-2.57**	-0.79	-2.38**	-0.18	-0.41	-0.062	-0.12
$\Delta SMOKE$	0.0065	0.03	0.10	0.41	0.10	0.45	0.18	0.78
$PLAN_{t-1}$	3.85	1.42	4.63	1.74	2.58	1.21	3.64	1.71
$\Delta PLAN$	0.87	0.73	1.57	1.21	-0.63	-0.47	0.01	0.07
TIME	-1.36	-2.57**	-1.46	-3.31***	-1.55	-3.24***	-1.68	-3.88***
_cons	112.44	2.50**	127.32	2.83**	102.09	2.77**	113.61	3.01***
Adj R^2	0.49		0.50		0.46		0.47	
F test (p-value)	11.68 (0.000)		11.24 (0.000)		13.18 (0.000)		12.89 (0.000)	
N	240		240		320		320	
Regions	15		15		20		20	

[a] T statistics computed with robust SEs.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 8

Estimation results (Error Correction Model - fixed effects): model (2).[a]

Regressors	Infant mortality							
	Excluding SSRs				All regions			
	Coef.	t	Coef.	t	Coef.	t	Coef.	t
IMR_{t-1}	-1.02	-10.49***	-1.08	-10.94***	-0.96	-11.03***	-0.99	-11.26***
VFI_{t-1}	9.05	1.70	36.01	2.78**	7.47	1.46	29.45	2.35**
ΔVFI	4.07	1.18	3.40	1.06	4.70	1.36	4.39	1.29
GDP_{t-1}	-0.0017	-3.52***	-0.0015	-3.62***	-0.0015	-3.84***	-0.0014	-3.65***
ΔGDP	-0.0015	-2.01*	-0.0009	-1.51	-0.0014	-2.12**	-0.0011	-2.03*
VFI_{t-1} x GDP_{t-1}			-0.0012	-2.93**			-0.001	-2.39**
$HEALTH\_EXP_{t-2}$	4.54	0.20	11.20	0.48	-31.86	-1.33	-27.82	-1.19
$\Delta HEALTH\_EXP_{t-1}$	-0.60	-0.03	3.95	0.20	-18.36	-1.02	-15.34	-0.94
$EDUC_{t-1}$	0.045	0.07	-0.26	-0.45	0.50	0.84	0.37	0.63
$\Delta EDUC$	-0.42	-0.78	-0.57	-1.02	-0.64	-1.26	-0.69	-1.34
$SMOKE_{t-1}$	-1.07	-2.92**	-0.81	-2.52**	-0.27	-0.56	-0.033	-0.07
$\Delta SMOKE$	-0.15	-0.62	0.068	0.28	0.019	0.08	0.16	0.71
$PLAN_{t-1}$	3.73	1.61	5.21	2.02*	2.54	1.32	3.53	1.64
$\Delta PLAN$	1.55	1.33	2.00	1.48	-0.25	-0.18	0.24	0.17
TIME	-1.54	-2.25**	-1.40	-2.49**	-1.59	-2.98***	-1.52	-3.21***
_cons	113.96	2.35**	110.55	2.87**	100.11	2.71**	92.93	2.89***
Adj R^2	0.47		0.51		0.45		0.47	
F test (p-value)	11.51 (0.000)		12.05 (0.000)		13.37 (0.000)		12.53 (0.000)	
N	240		240		320		320	
Regions	15		15		20		20	

[a] T statistics computed with robust SEs.

*Significant at 10%; **significant at 5%; ***significant at 1%.

A second concern relates to the linearity assumption assumed in our estimation. In presence of bounded dependent variables, as in our case, linear regression models may produce predicted values that lie outside of the bounded interval, are likely to be subject to floor and ceiling effects and often display non constant responses to changes as they approach to bounds (Papke and Wooldridge, 1996). We, thus, re-estimate, for both samples, all models and specifications using a Tobit regression with regional dummies. As shown in Tables 9 and 10, previous results are overall confirmed.

Table 9
Estimation result (tobit model with regional dummies): model (1).[a]

	Infant mortality							
Regressors	Excluding SSRs				All regions			
	Coef.	t	Coef.	t	Coef.	t	Coef.	t
FDEC	-23.62	-2.13**	-60.47	-2.67***	-19.26	-1.90*	-61.76	-2.51**
GDP	-0.0014	-3.88***	-0.0018	-4.45***	-0.0016	-4.75***	-0.0021	-5.27***
FDECxGDP			0.0019	2.23**			0.0021	1.92*
HEALTH_EXP	2.84	0.17	4.27	0.25	-23.80	-1.74*	-19.53	-1.45
EDUC	0.014	0.04	-0.073	-0.21	0.29	0.92	0.18	0.56
SMOKE	-0.31	-0.94	-0.30	-0.90	-0.70	-0.20	-0.041	-0.12
PLAN	2.79	1.94*	3.22	2.23**	1.16	0.89	1.76	1.32
TIME	-1.43	-3.50***	-1.46	-3.63***	-1.48	-4.09***	-1.50	-4.20***
_cons	100.44	3.68***	109.99	3.83***	109.30	4.85***	119.65	4.96***
F test (p-value)	49.93 (0.000)		53.46 (0.000)		42.13 (0.000)		51.86 (0.000)	
N	255		255		340		340	
Regions	15		15		20		20	

[a] T statistics computed with robust SEs.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 10
Estimation result (tobit model with regional dummies): model (2).[a]

Regressors	Excluding SSRs				All regions			
	Coef.	t	Coef.	t	Coef.	t	Coef.	t
VFI	4.33	1.21	25.97	2.84***	4.95	1.59	22.44	2.48**
GDP	-0.0017	-4.14***	-0.0014	-3.67***	-0.0017	-4.95***	-0.0016	-4.46***
VFIxGDP			-0.0009	-3.08***			-0.0008	-2.30**
HEALTH_EXP	5.53	0.32	8.83	0.53	-20.78	-1.52	-18.33	-1.35
EDUC	-0.025	-0.07	-0.21	-0.58	0.22	0.68	0.12	0.35
SMOKE	-0.39	-1.16	-0.28	-0.82	-0.14	-0.39	-0.024	-0.07
PLAN	2.84	1.99**	3.80	2.68***	1.13	0.87	1.81	1.38
TIME	-1.54	-3.83***	-1.44	-3.62***	-1.48	-4.11***	-1.43	-3.98***
_cons	101.88	3.57***	97.31	3.53***	108.28	5.69***	103.04	5.46***
F test (p-value)	45.63 (0.000)		54.02 (0.000)		38.35 (0.000)		42.07 (0.000)	
N	255		255		340		340	
Regions	15		15		20		20	

[a] T statistics computed with robust SEs.

*Significant at 10%; **significant at 5%; ***significant at 1%.

Finally we check the robustness of our findings with regard to the choice of different outcome variables. Results obtained using both neonatal mortality and LE at age 65 (available on request) are in line with those presented in this study.

7. Conclusions

This study investigates the overall and distributional effects of fiscal decentralization on health outcomes at the regional level in Italy. The results suggest that fiscal decentralization, measured as either a higher decision-making autonomy in the allocation of tax revenues or a lower dependence from central transfers, has an unambiguous positive overall effect on reducing IMRs. This positive effect is in line with other empirical evidence on developed countries experiencing a similar shift towards a higher degree of fiscal autonomy at the sub-national level (in Spain: Cantarero and Pascual, 2008; in Canada: Jimenez-Rubio, 2011a).

Beyond this pattern, fiscal decentralization also yields some distributional consequences: it affects IMRs differently according to the regional wealth, having a more positive effect in the poorest regions. These results contribute to the previous literature predicting heterogeneous response to decentralization, in presence of differences in the characteristics of local governments and populations. In particular, compared to the other existing contributions, our study uses a within-country approach for investigating the distributional effects of decentralization in a developed country, thus yielding more robust results.

In terms of policy implications, our findings outline the importance to consider fiscal decentralization not as a goal in itself but as an instrument to make local authorities more accountable in the management of public resources so as to ultimately improve health outcomes. However, the effectiveness of such an instrument should always be regarded together with the characteristics of the context in which the decentralization process takes place.

Nevertheless, no research is without its limitations and attention should be paid in interpreting our results. First of all, causal inferences must be drawn with caution. This

is a standard problem in statistical interpretation but it is frequently overlooked in practice. Attributions of cause and effect should remain speculative, although strong statistical relationships are observed. Other study limitations relate to the choice of both the measures of fiscal decentralization and the health outcome variable, leaving space for further research. In this study, we have investigated only one aspect of the multiple and complex processes involved in decentralization: the fiscal perspective. However, decentralization is much more than this; it is a very complex process, implying not only fiscal but, among others, political, administrative, and managerial issues. Given the multidimensional nature of the decentralization phenomenon, there is the need to rely on multiple and more sophisticated indicators able to capture different aspects of the problem. Furthermore, since decentralization is supposed to affect differently public sector functions and sub-functions, more function-specific measures of decentralization could help to better disentangle these differences. To this purpose, though IMR is actually considered to be the most reliable indicator of population health, other more specific and less general measures of objective health outcomes could help to account for the differential effects of decentralization. Finally, as decentralization is likely to result in better policy preferences match, subjective measures of outcomes could fit better. Therefore, their use should be considered in future analyses.

References

Akpan, E.O. (2011). Fiscal decentralization and social outcomes in Nigeria. *European Journal of Business and Management*, 3 (4), 167-183.

Alves, J., Peralta, S., Perelman, J. (2013). Efficiency and equity consequences of decentralization in health: an economic perspective. *Revista Portuguesa de Saúde Pública*, 31 (1), 74-83.

Asfaw, A., Frohberg, K., James, K.S., and al. (2007). Fiscal decentralization and health outcomes: empirical evidence from rural India. *The Journal of Developing Areas,* 41 (1), 17-35.

Bardhan, P., Mookherjee, D. (1998). Expenditure decentralization and the delivery of public services in developing countries. *Working Paper C98-104. Center for International and Development Economics Research*. Berkeley, CA.

Bardhan, P., Mookherjee, D. (2000). Capture and governance at local and national levels. *American Economic Review*, 90 (2), 135-139.

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

Bordignon, M., Turati, G. (2009). Bailing out expectations and public health expenditure. *Journal of Health Economics,* 28 (2), 305-21.

Caldeira, E., Foucault, M., Rota-Graziosi, G. (2014). Does decentralization facilitate access to poverty-related services? Evidence from Benin. In *African Successes: Government and Institutions, Volume 1*. University of Chicago Press.

Cantarero, D., Pascual, M. (2008). Analysing the impact of fiscal decentralization on health outcomes: empirical evidence from Spain. *Applied Economics Letters*, 15 (2), 109-11.

Channa, A., Faguet, J.P. (2016). Decentralization of health and education in developing countries: a quality-adjusted review of the empirical literature. *The World Bank Research Observer*, lkw001.

Costa-Font, J. (2012). Fiscal federalism and European health system decentralization: a perspective. *LEQS working paper 55*, London School of Economics, London.

Costa-Font, J., Greer S., (Eds.) (2013). *Federalism and decentralization in European health and social care*, Palgrave MacMillan, London.

Costa-Font, J., De-Albuquerque, F., Doucouliagos, H. (2015). Does inter-jurisdictional competition engender a "race to the bottom"? A meta-regression analysis. *Econonimcs & Politics,* 27 (3), 488-508.

Di Novi, C., Piacenza, M., Robone, S., Turati, G. (2015). How does fiscal decentralization affect within-regional disparities in well-being? Evidence from health inequalities in Italy. *HEDG working papers 15/23, Health, Econometrics and Data Group*, University of York, York.

Ebel, R.D., Yilmaz S. (2001). Fiscal decentralization: is it happening? How do we know? *paper prepared for the Conference on Public Finance in Developing Countries, Georgia State University, Atlanta*.

Ferrario, C., Zanardi, A. (2011). Fiscal decentralization in the Italian NHS: What happens to interregional redistribution?. *Health Policy*, 100 (1), 71-80.

Ferrè, F., Cuccurullo, C., Lega F. (2012). The challenge and the future of health care turnaround plans: evidence from the Italian experience. *Health Policy,* 106 (1), 3-9.

Fisher, R.C. (1982). Income and grant effects on local expenditure: The flypaper effect and other difficulties. *Journal of urban Economics*, 12 (3), 324-45.

France, G., Taroni, F., Donatini A. (2005). The Italian health care system. *Health Economics,* 14(Suppl. 1), S187-202.

Galiani, S., Gertler, P., Schargrodsky E. (2008). School decentralization: helping the good get better, but leaving the poor behind. *Journal of Public Economics,* 92 (10-11), 2106-2120.

Gramlich, E.M. (1977). Intergovernmental grants: A review of the empirical literature, in W.E. Oates (Ed.).*The political economy of fiscal federalism*. Health-Lexington. Lexington, MA, 219-239.

Grisorio, M.J., Prota, F. (2015a). The short and the long run relationship between fiscal decentralization and public expenditure composition in Italy. *Economics Letters,* 130 (C), 113-116.

Grisorio, M.J., Prota, F. (2015b). The impact of fiscal decentralization on the composition of public expenditure: panel data evidence from Italy. *Regional Studies,* 49 (12), 1941-1956.

Habibi, N., Huang, C., Miranda, D., et al. (2003). Decentralisation and human development in Argentina. *Journal of Human Development,* 4 (1), 73-101.

Im, K.S., Pesaran M.H., Shin Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics,* 115 (1), 53-74.

Jimenez Rubio D. (2011a). The impact of decentralization of health services on health outcomes: evidence from Canada. *Applied Economics,* 43 (26): 3907-3917.

Jiménez-Rubio, D. (2011b). The impact of fiscal decentralization on infant mortality rates: evidence from OECD countries. *Social Science & Medicine,* 73 (9), 1401-1407.

Jommi, C., Cantù E., Anessi-Pessina E. (2001). New funding arrangements in the Italian National Health Service. *The international journal of health planning and management,* 16 (4), 347-368.

Khaleghian, P. (2004). Decentralization and public services: the case of immunization. *Social Science & Medicine,* 59 (1), 163-183.

Levaggi, R., Smith, P.C. (2005). Decentralization in health care: lessons from public economics, in P.C. Smith, L. Ginelly, M. Sculpher (Eds). *Health Policy and Economics. Open University Press. Maidenhead*, United Kingdom, 223-247.

Levaggi, R., Zanola, R. (2007). Patients' migration across regions: the case of Italy. *Applied Economics,* 36 (16), 1751-1757.

Liberati, P., Sacchi A. (2013). Tax decentralization and local government size. *Public Choice*, 157 (1-2), 183-205.

Manor, J. (1999). *The political economy of democratic decentralisation. Directions in Development Series*. World Bank. Washington, DC.

Marmot, M. (2002). The influence of income on health: views of an epidemiologist. *Health Affairs* 21 (2), 31-46.

Oates, W. (1972). *Fiscal federalism*. Harcourt Brace Jovanovich. New York, NY.

OECD/KIPF (2013). *Measuring fiscal decentralisation: concepts and policies*. OECD Fiscal Federalism Studies. Organisation for Economic Co-operation and Development, Paris.

Papke, L.E, Wooldridge, J. M.. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics,* 11 (6), 619-632.

Piacenza, M., Turati, G. (2014). Does fiscal discipline towards sub-national governments affect citizens' well-being? Evidence on health. *Health Economics,* 23 (2): 199-224.

Porcelli, F. (2014). Electoral accountability and local government efficiency: quasi-experimental evidence from the Italian health care sector reforms. *Economics of Governance* 15 (3), 221-251.

Prud'homme, R. (1995). The Dangers of Decentralization. *The world bank research observer,* 10 (2), 201-220.

Robalino, D.A., Picazo, O.F., Voetberg, A. (2001). *Does fiscal decentralization improve health outcomes? Evidence from a cross-country analysis*. Policy Research Working Paper 2565. World Bank. Washington, DC.

Saltman, R.B., Bankauskaite, V., Vrangbaeck, K. (2007). *Decentralization in health care: strategies and outcomes*. European Observatory on Health Systems and Policies Series. Open University Press. Maidenhead.

Shleifer, A. (1985). A theory of yardstick competition. *The RAND Journal of Economics,* 16 (3), 319-327.

Smith, B.C. (1985). *Decentralization: the territorial dimension of the State*. George Allen and Unwin, London.

Soto, V.E., Farfan M.I., Lorant V. (2012). Fiscal decentralisation and infant mortality rate: the Colombian case. *Social Science & Medicine,* 74 (9), 1426-34.

Stegarescu, D. (2005). Public sector decentralization: measurement concepts and recent international trends. *Fiscal studies,* 26 (3), 301-333.

Tediosi, F., Gabriele, S., Longo, F. (2009). Governing decentralization in healthcare under tough budget constraint. *Health Policy,* 90 (2): 303-312.

Tiebout, C.M. (1956). A pure theory of local expenditures. *The journal of political economy*, 64 (5), 416-424.

Toth, F. (2014). How health care regionalization in Italy is widening the North-South gap. *Health Economics Policy and Law,* 9 (3): 231-249.

Turati, G. (2013). The Italian Servizio Sanitario Nazionale: A renewing tale of lost promises. In J. Costa-Font, S.L. Greer (Eds). *Federalism and decentralization in European health and social care*. Palgrave Macmillan, United Kingdom, 47-66.

Uchimura, H., Jutting, J. (2009). Fiscal decentralization, Chinese style: good for health outcomes? *World Development,* 37 (12), 1924-1936.

Weingast, B.R. (2009). Second generation fiscal federalism: the implications of fiscal incentives. *Journal of Urban Economics,* 65 (3), 279-293.

CHAPTER 2[2]

Do health outcomes across the Italian regions converge? The role of decentralization and neighbours' effects

ABSTRACT

This paper examines convergence hypothesis for regional health outcomes (in terms of infant mortality rate and life expectancy at birth) in Italy over the period 1996-2012, with the aim to disentangle the role played by decentralization and neighbours' effects. In terms of sigma convergence, no evidence is found of reduction in disparities for the two health outcomes. On the opposite, beta-convergence affects the time dynamics of health outcomes across regions. Independently from the indicator actually used to measure decentralization (either a continuous or a dichotomous variable), shifting responsibilities at regional level is proved to fasten the rate of convergence in Italy. However, beyond certain levels of performance a trade-off between higher levels of decentralization and further improvements arises. Along with decentralization, neighbours' effects are also found to be relevant in explaining the regional convergence process. Different cutting-edge spatial convergence metrics confirm the pushing up effect of decentralization on the speed of convergence, especially for infant mortality.

Keywords: convergence, decentralization, spatial effects, infant mortality, life expectancy, Italy

---

[2]Co-authored by Prof. Marina Cavalieri

1. Background

Despite the long-lasting theoretical and political debate on the potential benefits and pitfalls arising from a decentralized structure of government, existing empirical evidence is still lacking or inconclusive, thus calling for deeper research efforts (Shah et al., 2004, Treisman, 2007). The uncertainty surrounding the 'true' effects of decentralization has recently led in many countries to a process of re-thinking of the promising paradigm of decentralization and to a new trend toward re-centralization (Dickovick, 2011; Malesky et al., 2014).

A scant stream of literature has considered the relationship between decentralization and health (see Cavalieri and Ferrante, 2016, for a review of empirical studies). Attention has been primarily paid to the issue of efficiency, analyzing the effects of decentralization on service provision, health expenditure and, to a less extent, health outcomes. The latter is however a particularly interesting topic since health outcomes, in terms of mortality and life expectancy, are suggested to be valid measures for quality of life (Maynou et al., 2015) and, hence, for citizens' well-being (Sen, 1998; Becker et al., 2005), even superior to income. Notwithstanding, empirical research has failed to reach firm and unequivocal conclusions, mainly due to the difficulties of analyzing a phenomenon (i.e. decentralization) that is, by its nature, highly context-dependent and shaped by numerous factors that are difficult to be systematically isolated and assessed in a heterogeneous cross-country setting. In this regard, by reviewing the relevant literature, Channa and Faguet (2016) conclude that decentralization appears to enhance preference matching and technical efficiency in health only when empirical evidence is organized by substantive themes and restricted to higher-quality studies.

The consequences of decentralization on equity, in general, and on equity in health, in particular, have been widely neglected in the literature. A small number of papers has applied the standard convergence analytical framework of the economic growth literature (Baumol, 1986; Barro et al., 1991; Barro and Sala-i-Martin, 1992; Quah, 1993) to health outcomes without, however, explicitly considering the role played by decentralization on the converging/diverging process. Employing indicators of life expectancy, few of these studies find sigma-convergence (Becker et al., 2005; Wilson, 2011), while others do not find sigma-convergence in older ages (Glei et al., 2010) or across countries (Edwards, 2011), although evidence of beta-convergence seems always to emerge from them. When measures of mortality are considered, neither evidence of sigma-convergence in industrialized countries (Edwards and Tuljapurkar, 2005), nor beta-convergence across countries appear (Clark, 2011; Edwards, 2011). However, as different sub-samples of more homogeneous countries are considered, the previous results seem to be overturned (d'Albis et al., 2012). More recently, Maynou et al. (2015) extend the previous literature by investigating the speed of (beta) convergence of (cause-specific) mortality and life expectancy at birth in EU regions between 1995 and 2009. Exploiting variations within a country and controlling for spatial correlations across regions, they find no evidence of sigma-convergence but of beta- convergence on average, though with marked differences in the catching-up process across both time and regions.

A step forward in the analysis of convergence in health has been done by Montero-Granados et al. (2007), who investigate the influence of decentralization on the health convergence process among the Spanish Autonomous Communities and provinces according to two different health outcome indicators (i.e. infant mortality rate and life expectancy at birth). Using a quasi-experimental design, the authors conclude

that the process of decentralization either does not affect convergence or leads to divergence in health outcomes. Indeed, in the case of infant mortality the so-called change of role scenario seems to have occurred, where certain provinces with initially poor indicators have improved, overtaking those that were originally in a better position. The final result, however, is of greater dispersion than initially.

Drawing on the above literature, this paper employs a balanced panel of 20 Italian regions over a 17-year period (1996-2012) to address two related research questions: whether health outcomes of Italian regions, as measured in terms of infant mortality rates and life expectancy at birth, are converging or diverging and to what extent the degree of regional decentralization and the neighbours' effects influence the converging/diverging process. The contribution of the paper is manifold. First of all, a more robust within-country analysis is carried out, which allows to avoid problems of cross-country heterogeneity and comparability of findings. Specifically, the case-study of Italy is considered where a decentralization process of reform has been implemented over the time with the goal of increasing the autonomy of regional governments in planning and financing healthcare services, albeit within a national regulatory framework (for further details, see Cavalieri and Ferrante, 2016). In this respect, Italy represents a unique "natural laboratory" to test the effects of decentralization on the time dynamics of health outcomes across regions. As a further contribution to the previous empirical analyses on convergence in health, this paper explicitly accounts for the geographical components of the regional convergence phenomenon. More in depth, the paper employs different spatial econometric techniques to control for the fact that each region is not a geographically independent decision-making unit but is likely of being affected by its neighbours' behaviours (Rey and Montouri, 1999). Third, with regard to the measures of decentralization, a multiple approach is used, consistently

with the multi-faceted nature of the phenomenon, which involves both administrative, fiscal and political dynamics (Falleti, 2005). Therefore, in the next analysis, a quantitative continuous indicator of regional decision-making autonomy is considered along with a dummy variable that captures the political changes following the introduction of the Constitutional law n.3/2001[3].

The remainder of the paper is organized as follows. The next section describes the dataset and the methodology employed in the paper. In the third section, the results of the empirical analysis are reported and discussed. Then, different robustness tests are conducted. In the final section, concluding remarks are offered.

2. Data and methodology

A balanced panel of 20 Italian regions over the period 1996 to 2012 is employed. Health outcome is proxied by two different indicators: Infant Mortality Rate (*IMR*) and Life Expectancy at Birth (*LEB*). These are commonly considered two of the most exhaustive indicators of health in a society, though the former is more sensitive to policy changes (e.g. decentralization reforms) than the latter, which is widely influenced by events beyond the direct control of healthcare systems (Jiménez-Rubio, 2011). Specifically, *IMR* refers to the number of deaths of children aged less than one year per 10,000 live births while LEB is computed as the arithmetic mean between male and

---

[3] For a comprehensive description of the decentralization reforms undertaken in Italy since the '90s, see the previous chapter of this dissertation.

female values. Regional data on both indicators are taken from the WHO Health for All-Italy (HFA) database[4].

Figures 1 and 2 show how the above two outcome indicators are distributed across regions in the first (i.e. 1996) and last (i.e. 2012) year of the study period: the darker the colour in the map, the higher either the IMR or the LEB values. The maps clearly indicate a high cross-regional variation, with clear-cut differences between the best performing Northern and Central regions and the worst performing regions of the South of Italy. Moreover, a pure inspection of the spatial patterns of the two outcome indicators reveals insights of neighbourhood effects, thus calling for an explicit analysis of spatial dependence among territorial units.
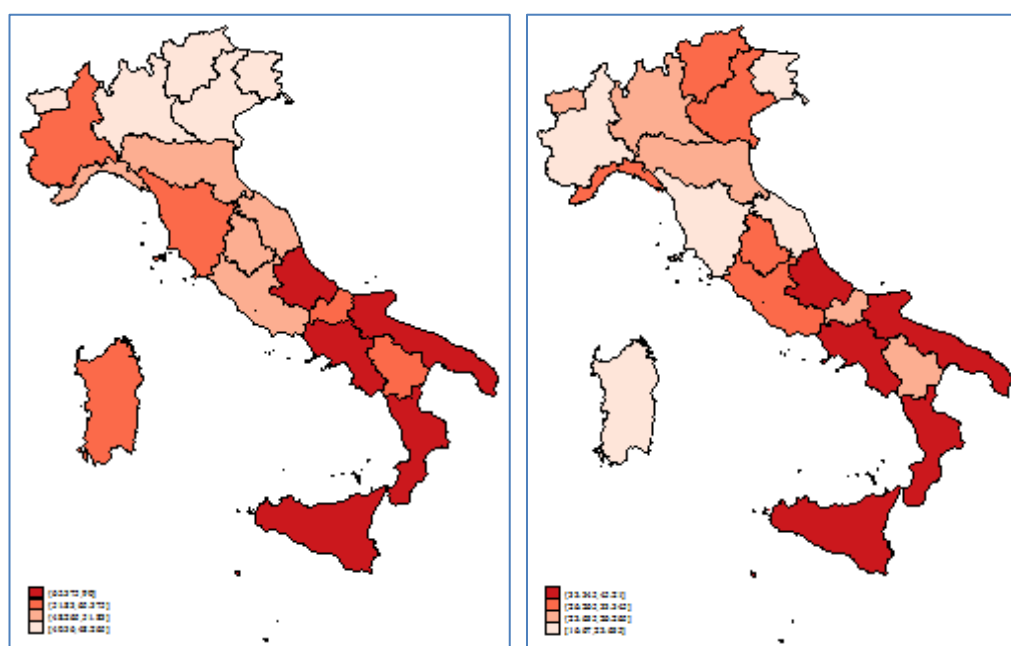
Fig. 1. Spatial distribution of *IMR* (in quartile) – year 1996 (on the left) and year 2012 (on the right)
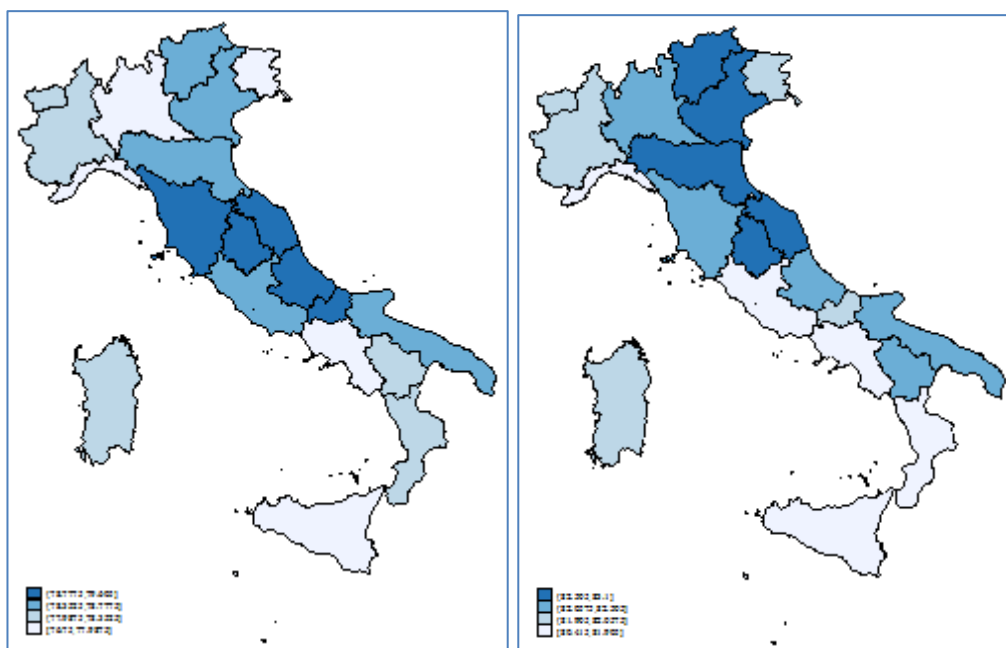
Fig. 2. Spatial distribution of *LEB* (in quartile) - year 1996 (on the left) and year 2012 (on the right)

To investigate the convergence/divergence of the above health outcomes across Italian regions, in this paper two concepts from the economic growth literature are applied (Barro et al., 1991; Barro, and Sala-I-Martin, 1992): *sigma* (σ-) and *beta* (β-) convergence. The former concept of convergence (i.e. σ-convergence) occurs when the dispersion of a given variable across regions, as measured by different dispersion indicators (i.e. variance, standard deviation or coefficient of variation), falls over time. The latter concept of convergence (i.e. β-convergence) measures the so-called "catching-up effect", where poorer economies tend to grow at faster rates than wealthier ones (Sala-i-Martin, 1996). With regard to this paper, β-convergence occurs when worse performing regions experience greater improvements in health outcomes than better performing ones. This can be assessed by testing the relationship between the growth rate of the health outcome indicator (either *IMR* or *LEB*) and its starting level for each period. In contrast with the standard literature, in this paper a pooled model is preferred

to a cross-section one. The main advantage of a pooled approach, with respect to a cross-section model, relies on the increase in the total number of observations that are equal to N (i.e. number of regions) × T (i.e. years). Therefore, to estimate (unconditional) β-convergence, the following Ordinary Least Squares (OLS) regression is run:

$$ln\left(\frac{y_{i(t+1)}}{y_{i,t}}\right) = \alpha + \beta \; ln(y_{i,t}) + \varepsilon_{i,t} \tag{1}$$

where y represents the health outcome variable (either *IMR* or *LEB*), i indicates the region (i=1, 2, …, 20), t is the year of observation (t=1996, 1997, …, 2012), β is the strength of the convergence process, and the error terms ($\varepsilon_{i,t}$) are assumed to be identical, independent and normally distributed. A statistically significant and negative sign of β implies the presence of absolute β-convergence.

From Eq. (1), two indicators can be derived that are often used to characterise the β-convergence process: the *speed of convergence* and the *half-life* (Arbia, 2006). The speed of convergence (b) indicates how fast economies converge towards the steady-state and is computed according to the following formula:

$$b = - \frac{ln \; (1+ \beta)}{T} \tag{2}$$

where T is the number of periods for which data on growth rates are available. Therefore, in the pooled approach applied in this paper $T = 1$.

The half life is defined as the time required to eliminate half of the initial gap from the steady-state and is calculated as follows:

$$Half\text{-}life = \frac{ln \; (2)}{b} \tag{3}$$

where *b* indicates the speed of convergence.

In the economic growth literature a distinction is made between unconditional (absolute) and conditional β-convergence. The unconditional β-convergence relies on the assumption that all regions converge to the same steady-state and that there is homogeneity among their structural characteristics. However, this is not always the case if cross-regional heterogeneity exists and factors other than differences in the study variable may condition the convergence process. The need to isolate and to control for these variables leads to develop the conditional concept of β-convergence, by modifying Eq. (1) as follows:

$$ln\left(\frac{y_{i(t+1)}}{y_{i,t}}\right) = \alpha + \beta\ ln(y_{i,t}) + \gamma\ ln(Z_{i,t}) + \varepsilon_{i,t} \qquad (4)$$

where $Z_{i,t}$ denotes a matrix of explanatory variables (of convergence) and γ is the associated vector of (unknown) parameters. Again, the conditional β-convergence hypothesis can be accepted if the estimated value for β is significantly negative. A higher determination coefficient value (R^2) than the unconditional approach, gives a measure of goodness of fit.

The concepts of σ-convergence and β-convergence are closely related. Formally, β-convergence is a necessary but not a sufficient condition for σ-convergence to take place (Barro and Sala-i-Martin, 1992; Sala-i-Martin, 1996; Young et al., 2008). Moreover, if conditional convergence or convergence clubs exist, the variance approach is biased by wrong inferences (Plümper and Schneider, 2009). Therefore, the two measures of convergence should be regarded as complementary and not exclusive. Indeed, Quah (1993) has argued that σ-convergence is of greater interest since it

provides straightforward information on whether the distribution of the study variable across regions is becoming more equitable.

In this paper, both σ-convergence and β-convergence are considered. Although models of absolute β-convergence are generally plausible when the object of study is within-country convergence (i.e. regions share common steady-states due to their similar characteristics. See, Chocholatá and Furková, 2016), the marked heterogeneity existing across Italian regions requires opting for a conditional approach. Furthermore, the latter approach is also the only one consistent with the specific purpose of this study, that is to disentangle the impact of decentralization on the convergence process.

In the choice of the explanatory variables to be inserted in the conditional β-convergence, a quite parsimonious approach is applied. Thus, to control for the effect of decentralization (*DEC*) on convergence two different measures are employed, alternatively. Following Grisorio and Prota (2015a and 2015b), the first measure is a continuous fiscal decentralization indicator (*FDEC*) that captures the degree of regional decision-making autonomy in the allocation of tax revenues. This is based on fiscal data from the Italian Territorial Public Accounts (*Conti Pubblici Territoriali*) and is computed as the ratio of tax revenues raised and/or controlled by the regional government to the total tax revenue collected in the region (Cavalieri and Ferrante, 2016). The second indicator of decentralization (*REF*) is a dummy variable equal to 1 after the year 2001, when the Italian Constitutional law n.3/2001 was issued, aiming to capture the political and administrative changes introduced by the reform. From a strictly fiscal point of view, the *REF* variable also allows to control for the year of the introduction of the new mechanism of funding based on a regional sharing of the national Value Added Tax. As additional control, real per-capita gross domestic product

(*GDP*) is also included to account for cross-regional differences in both the size of tax bases and the living conditions. Therefore, Eq. (4) is reformulated as follows:

$$ln\left(\frac{y_{i(t+1)}}{y_{i,t}}\right) = \alpha + \beta \, ln(y_{i,t}) + \gamma \, ln(DEC_{i,t}) + \delta \, ln(GDP_{i,t}) + \varepsilon_{i,t} \qquad (5)$$

where *DEC* is the decentralization indicator (either *FDEC* or *REF*) and all other terms are as previously defined. However, Eq. (5) doesn't explain anything about whether the rate of adjustment of the convergence process is conditioned by the decentralization process undertaken in Italy (as measured by either *FDEC* or *REF*). Therefore, following the approach originally proposed by Plümper and Schneider (2009) and then applied by Schmitt and Starke (2011), a third model including an interaction term (*INT*) between the decentralization indicator (*DEC*) and the starting level of health outcome *y* is also estimated:

$$ln\left(\frac{y_{i(t+1)}}{y_{i,t}}\right) = \alpha + \beta \, ln(y_{i,t}) + \gamma \, ln(DEC_{i,t}) + \delta \, ln(GDP_{i,t}) + \theta(INT_{i,t}) + \varepsilon_{i,t} \qquad (6)$$

In the above equation, the causal relationship between rate of convergence and decentralization is read as follows: 1) if $\beta = 0$ and $\theta \neq 0$ the speed of adjustment depends entirely on the level of decentralization; 2) if $\beta, \theta \neq 0$ the speed of adjustment depends partly on the level of decentralization; 3) if $\beta \neq 0$ and $\theta = 0$ convergence independence exists. Descriptive statistics of the study variables for the pooled sample (1996-2012) are reported in Table 1.

Table 1
Descriptive statistics for the pooled sample

Variable	Source	Obs.	Mean	SD	Min	Max	Description
IMR	HFA	340	39.24	13.73	2.06	90	Infant Mortality Rate per 10,000 birth
LEB	HFA	340	80.47	1.38	76.75	83.11	Life Expectancy at Birth for male and female
FDEC	TPA	340	0.21	0.12	0.01	0.54	Degree of regional decision-making autonomy
REF	-	340	0.65	0.48	0	1	Dummy equal to 1 after Const. Law n.3/2001
GDP	HFA	340	26,582	6,712	14,685	38,582	GDP per capita in 2012 PPP

The above models are based on the assumptions that each region is an independent entity and no spatial interaction exists among observations. However, during the last years the traditional convergence modelling has been modified to account for spatial effects. It has been proved that, if ignored or not properly modelled, spatial effects may give rise to serious econometric problems such a misspecification and biased or inefficient estimates, depending on the form of spatial dependence.

A popular indicator used to detect spatial dependence among the geographic units is the global Moran's I statistic (Moran, 1950), defined as:

$$I_t = \left(\frac{n}{s_0}\right) \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} x_{i,t} x_{j,t}}{\sum_{i=1}^{n} \sum_{j=1}^{n} x_{i,t} x_{j,t}} \tag{7}$$

where w_{ij} is an element of a spatial matrix W; $x_{i,t}$ is the natural log of the outcome variable (either *IMR* or *LEB*) in region i in year t, n is the number of regions (i.e. 20), and s_0 is a scaling factor equal to the sum of all the elements of W. In this paper the spatial interaction between regions is modelled by a contiguity spatial weight matrix (W): a square, nonstochastic and symmetric matrix, whose elements (w_{ij}) measure the intensity of the spatial connection between regions i and j and take on a finite and nonnegative value. Specifically, in building the contiguity matrix W the

queen's criteria in the game of cheese is followed (Chocholatá and Furková, 2016), giving the value 1 to those contiguous regions that share any part of a common border (zero otherwise).

As for the form of spatial dependence, Anselin and Rey (1991) distinguish between substantive spatial dependence and nuisance dependence. The latter refers to spatial autocorrelation that pertains to the error term and can be caused by measurement problems such as a boundary mismatch between the spatial pattern of the analysed phenomenon and the units of data observation. The substantive form of dependence characterises economic phenomena that incorporate spatial interactions. Following the approach by Rey and Montouri (1999), this paper focuses on the substantive forms of spatial autocorrelation in convergence analysis, specifically spatial cross-regressive models. In contrast to the alternative spatial lag model where the spatial lag of the health outcome growth rates is incorporated into the original non-spatial specification of β-convergence, in a spatial cross-regressive model the spatial lag of starting health outcomes is added. Therefore, by choosing the latter spatial specification, this paper implicitly assumes that regional behaviours are influenced by the levels of health outcomes of their neighbours, rather than by their growth rates. This is actually a quite reasonable hypothesis in a context, such as the Italian one, where citizens are free to decide where to receive healthcare and regions are in competition against each other to provide care to them. Under these circumstances, it is expected that the neighbours' health outcome values are seen as a benchmark to reach.

In order to improve the predictive power of previous models with regard to their capacity of disentangling the potential role of decentralization on the converge process of regional health outcomes, in this paper the non-spatial Eq. (5) and (6) are re-estimated as follows:

$$ln\left(\frac{y_{i(t+1)}}{y_{i,t}}\right) = \alpha + \beta\ ln(y_{i,t}) + \gamma\ ln(DEC_{i,t}) + \delta\ ln(GDP_{i,t}) + \mu\ W^*ln(y_t) + \varepsilon_{i,t}$$

(8)

$$ln\left(\frac{y_{i(t+1)}}{y_{i,t}}\right) = \alpha + \beta\ ln(y_{i,t}) + \gamma\ ln(DEC_{i,t}) + \delta\ ln(GDP_{i,t}) + \theta(INT_{i,t}) +$$

$$\mu\ W^*ln(y_t) + \varepsilon_{i,t} \qquad\qquad (9)$$

where the term $W^*ln(y_t)$ is the spatial lag of starting health outcomes (either *IMR* or *LEB*) and W^* is our previous contiguity matrix that is now weighted by the number of regions that share a common border. One advantage of a cross regressive model is that OLS continues to be an appropriate estimation procedure, since both the starting health outcome variable and its spatial lag are exogenous.

3. Results and discussion

In this Section the paper's research question - i.e. understanding the role played by decentralization as well as by neighbours' effects on the convergence of regional health outcomes - is addressed step by step. First, attention is directed towards σ-convergence and the potential for observational interactions across space is considered through an exploratory data analysis. Then, the effects of decentralization are disentangled, using a *β*-convergence approach.

3.1 Sigma convergence and spatial correlation

To examine σ-convergence the standard deviations for the natural log of each of the two health outcome variables are computed for the 20 Italian regions for the period from 1996 to 2012. Figures 3 and 4 depict the results of σ-convergence for *IMR* and

LEB, respectively. When the entire period is considered, the Figures provide no clear indication of either long-term convergence or long-term divergence. Specifically, with regard to (the log of) *IMR*, the level of dispersion for the last year (0.26) is slightly higher than that for the first year (0.22), though the year-by-year analysis show quite fluctuating values with the highest peak of 0.66 in the year 2007. On the contrary, the value of the standard deviation of (the log of) *LEB* has slightly declined from 0.008 in 1996 to 0.007 in 2012, showing a less fluctuating trend. No particular patterns of standard deviations can be observed as a consequence of the implementation of the 2001 Constitutional reform. To detect whether the differences in standard deviations are significant, the Levene's *F* test (Levene, 1960) is carried out. For both health outcomes, the test fails to reject the nil hypothesis of equality of variance between the starting (1996) and final (2012) year (H$_o$: $\sigma^2_{1996}=\sigma^2_{2012}$; test values equal to 0.674 for *IMR* and to 0.251 for *LEB*; *p*-values > 0.10 in both cases)[5].

The other series presented in Figures 3 and 4 are the Moran's I statistics that display the paths of spatial correlation by year and are computed following Eq. (7). For both health outcomes, the Moran's coefficients are positive and statistically significant for almost all years[6], thus ruling out the hypothesis of regions as independent observations and supporting that of distributions being clustered in nature.

Looking at the relationship between the σ-convergence and the Moran's I statistic, a strong negative correlation exists over the study period for *IMR* (r = - 0.673): higher (lower) levels of dispersion in regional *IMR* values are associated with lower (higher) spatial correlation. On the contrary, in the case of *LEB*, evidence of a co-movement

---

[5] The Levene's *F* test implicitly assumes that data are normally distributed. Indeed, according to the Kolmogorov-Smirnov test for normality, this is proved to be the case for the majority of the health outcomes data employed in this study (over 90%).
[6] The coefficients are significant at 1% for almost all years, with the exception of 2006, 2007 and 2012 for the variable *IMR*.

between dispersion in regional data and spatial dependence emerges (r = + 0.259), with the Moran's I statistics showing an overall increasing trend during the study period. Following Rey and Montouri (1999), a higher spatial dependence can be the consequence of two different effects: cluster becoming more similar in their health outcome values and/or the onset of newly formed clusters. However, the global Moran's I statistics employed in the Figures 3 and 4 do not allow disentangling these two effects.
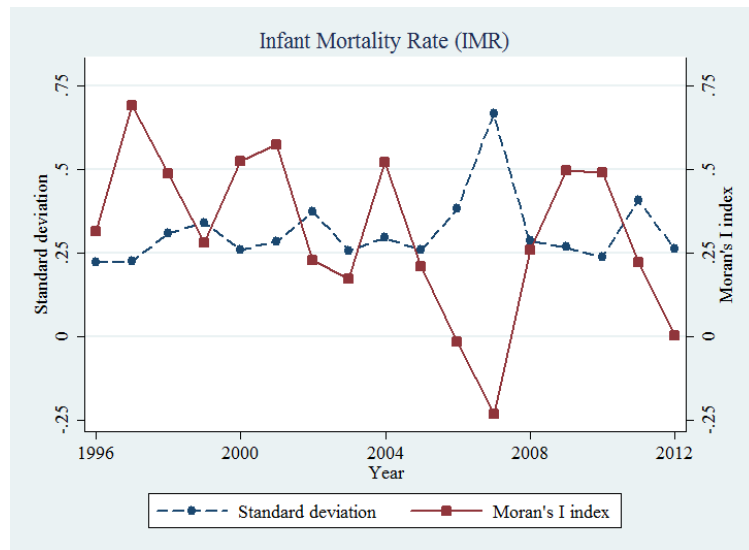

Fig. 3. Standard deviation and spatial correlation of (log of) *IMR*, 1996-2012
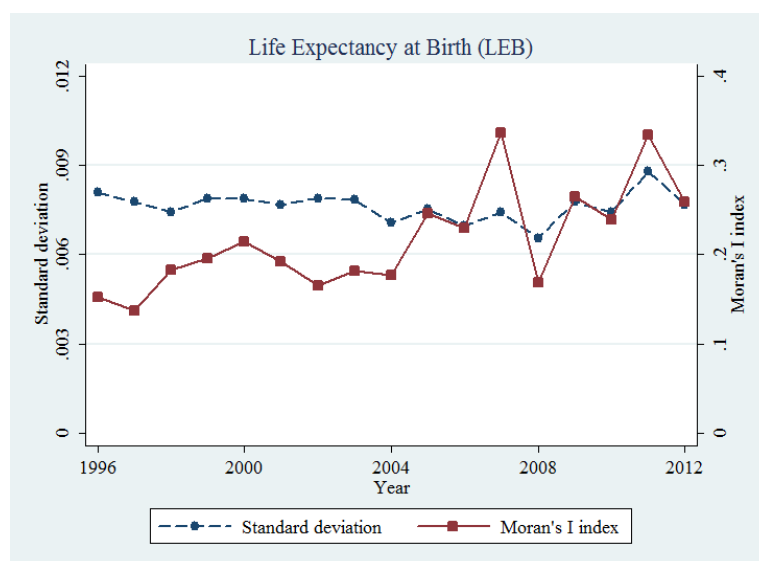
Fig. 4. Standard deviation and spatial correlation of (log of) *LEB*, 1996-2012

A common visual tool for exploratory spatial data analysis is provided by the Moran's scatterplot (Anselin, 1993), where the (standardized) values for each unit on the x-axis are plotted against the respective spatial lag value on the y-axis. Figures 5 and 6 display the Moran's scatterplots for the pooled data (1996-2012) of *IMR* and *LEB*, respectively. In both cases, the spatial lag of the (log of the) regional health outcome variable is the average of the (log of the) health outcome values of the neighbouring regions. All values are then standardized year by year to make data comparable across time.

The Moran's scatterplot allows identifying four different quadrants, each of which presents a different type of spatial association between a region and its neighbours. In the below Figures, the concentration of the observations in the first (*high-high* association) and third (*low-low* association) quadrant suggests a positive spatial dependence for both health outcome variables: the slope of the regression line (i.e. the global Moran's I), is positive (equal to 0.18 and 0.21, respectively) and statistically significant at the 1% level for both *IMR* and *LEB*.
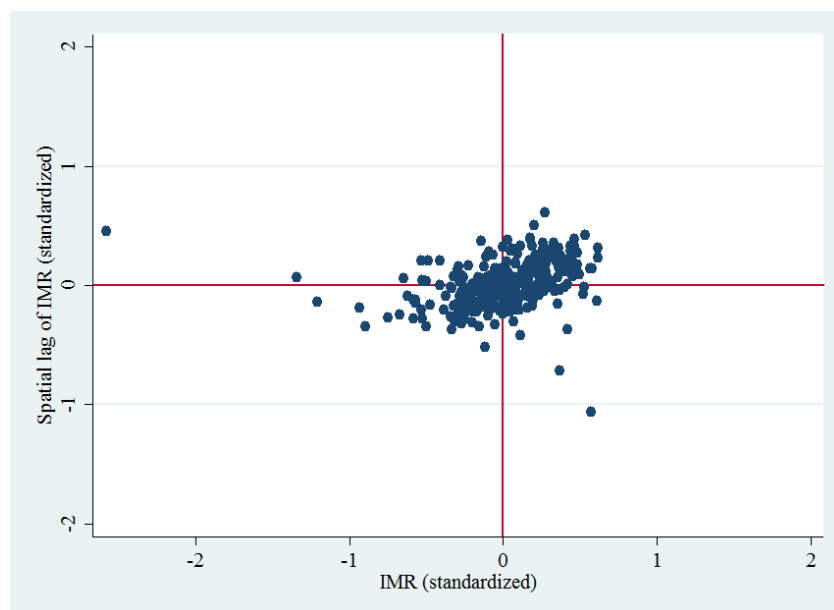
Fig. 5. Moran scatterplot of (log of) *IMR*, 1996-2012
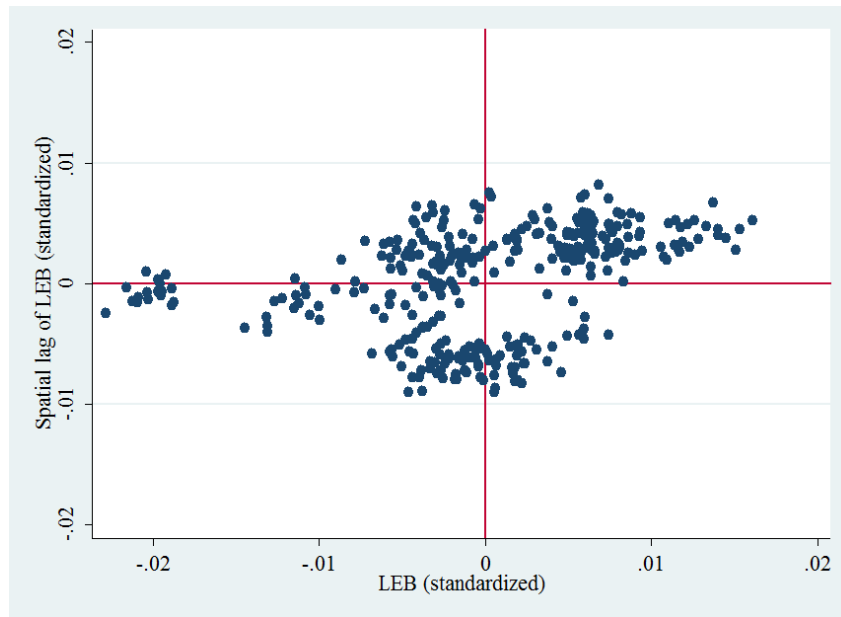
Fig. 6. Moran scatterplot of (log of) *LEB*, 1996-2012

To better investigate the dynamics of the spatial association between a region and its neighbours over the period 1996-2012, Table 2 reports, for both *IMR* and *LEB*, the number of years the local Moran's value falls in each of the four quadrants of the above Moran's scatterplots. Several points are worth noting. First of all, for both health outcomes, around 70% of the local Moran's statistics fall in either the first or the third quadrant of the scatterplot. As for the remaining statistics, the prevailing form of clustering is that of "doughnut", with indicators revealing negative associations more concentrated in the fourth quadrant for *IMR* (i.e. a high *IMR* region with low *IMR* neighbours) and the second quadrant for *LEB* (i.e. a low *LEB* region with high *LEB* neighbours). Out of 20 regions, 16 regions for *IMR* and 15 regions for LEB have most of their local Moran's values located in either the first or third quadrant. Some geographic clusters persist throughout the 17 years. Specifically, with regard to *IMR*, two clusters appear. The first cluster includes the regions in the north and central part of Italy (with the only exceptions being Lazio and Marche), each of which mostly appears

in the third quadrant. The second cluster comprises the southern regions (with the only exceptions being Molise and Abruzzo), each of which falls in the first quadrant the majority of years. In line with Rey and Montouri (1999), it could be suggested that the negative correlation between the global Moran's I and the *IMR* dispersion is due to a strengthening of the regional clusters during periods of income convergence, rather than to the appearance of newly formed clusters. The clustering picture emerging for LEB is, however, very mixed, making more difficult to reach similar conclusions.

Table 2
Summary of local measures of spatial association, 1996-2012

	Infant Mortality Rate (*IMR*)				Life Expectancy at Birth (*LEB*)			
Region	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**
Abruzzo	6	1	1	9	14	3	0	0
Basilicata	14	0	0	3	0	0	12	5
Calabria	16	0	0	1	0	0	16	1
Campania	14	0	0	3	0	3	14	0
Emilia-Romagna	0	1	10	6	17	0	0	0
Friuli-Venezia Giulia	0	1	14	2	2	15	0	0
Lazio	9	2	0	6	1	13	3	0
Liguria	1	0	10	6	1	16	0	0
Lombardia	1	0	16	0	9	8	0	0
Marche	1	6	5	5	17	0	0	0
Molise	8	9	0	0	0	0	3	14
Piemonte	1	3	10	3	0	7	10	0
Puglia	15	0	0	2	0	0	3	14
Sardegna	10	7	0	0	0	0	15	2
Sicilia	17	0	0	0	0	1	16	0
Toscana	0	3	12	2	17	0	0	0
Trentino-Alto Adige	1	0	11	5	16	0	0	1
Umbria	1	3	10	3	17	0	0	0
Valle d'Aosta	1	3	7	6	0	0	17	0
Veneto	0	1	14	2	16	0	0	1

3.2. Beta convergence and decentralization influence

This section introduces the issue of *β*-convergence. It begins by considering unconditional (absolute) *β*-convergence models and then proceeds by investigating the role of decentralization in the convergence process, either in absence or presence of spatial dependence.

Table 3 displays the OLS results for the (unconditional) *β*-convergence pooled model of Eq. (1). For both *IMR* and *LEB*, the estimated *β* coefficients are negative and highly statistically significant (at 1% level), meaning that regions with higher starting values of health outcomes tend to grow less than regions with lower starting values. All these will lead to a natural reduction in differences over time. Table 3 also provides the estimated speed at which regions converge to their steady-state. This is more than ten times higher for *IMR* (66.94%) than for *LEB* (5.67%), leading regions to halve the gap from the predicted equilibrium over around 1 and 12 years, respectively.

Table 3
Unconditional *β*-convergence (pooled sample)

Regressors	Eq. (1)				
	IMR	*Growth rate*		*LEB*	
	Coef.	T		Coef.	t
Y	-0.488	-3.90***		-0.0551	-12.52***
_cons	1.717	3.84***		0.245	12.64***
R^2	0.256			0.089	
Speed	66.94%			5.67%	
Half-life	1.04			12.23	
N	340			340	

Source: our elaboration on HFA data.
Notes: ***, **, and * denote significance at the 1%, 5% and 10% level.
All variables expressed in natural log form. *t*-statistics computed with
robust standard errors clustered at regional level.

Under a conditional β-convergence hypothesis it is assumed that region-specific factors exist leading regions to converge to different steady-states. Among these factors, the different degree of regional decentralization is here considered. Table 4 reports the results from estimation of Eq. (5), (6), (8) and (9) for the variable *IMR*. In the Table, models are labelled with the letter *a* or *b* according to whether the variable *FDEC* or *REF* is used to measure decentralization at regional level. As for Eq. (5*a*) and (5*b*), evidence of a stronger convergence than in the unconditional approach emerges since the coefficients of β are both statistically significant and equal to -0.718 and -0.774, respectively. Looking at the coefficients of the decentralization variables (i.e. *FDEC* and *REF*), they are negative and statistically significant, meaning that higher decentralization levels push for a reduction in regional *IMR* growth rates over time. A similar thing can be said for the variable *GDP*. In Eq. (6*a*) and (6*b*) the interaction term (*INT*) between decentralization and the starting level of *IMR* is added to the basic specifications so as to disentangle the impact of decentralization on the health convergence process. The sign of the variable *INT* is always negative and significant (p-value < 0.01 in model 6*a* and p-value < 0.05 in model 6*b*), suggesting that a higher degree of decentralization is likely to enhance convergence across regions. The strength of the decentralization effect on the convergence process is better displayed by the speed of convergence and the half-life. In Eq. (6*a*) and (6*b*) both these measures are reported according to three different levels of *FDEC* (respectively: high decentralization = mean + std. dev.; mean decentralization = mean value; low decentralization = mean - std. dev) and two different states of *REF* (respectively: high decentralization if *REF* = 1; low decentralization if *REF* = 0). Overall, speed of convergence increases as decentralization becomes higher, thus lowering the half-life of convergence. It is worth

noting that the extremely high values of the speed of convergence signal a process of overshooting, whereby regions exceed the steady-state level.

In the Eq. (8) and (9) of Table 4 the previous specifications are augmented with the spatial lag of starting IMR (W_{IMR}) to account for the effects of substantive spatial spillovers across regions on the convergence process. Results from the OLS cross-regressive models suggest a strong influence of neighbouring starting levels on regional IM growth rates: the coefficient of the spatial lag is positive and statistically significant at 5% level in all specifications. Concerning the effects of the other variables, all the previous conclusions remain valid. The importance of including spatial effects in the estimation is further proved by the goodness of fit of the spatial cross-regressive models (with adjusted R^2 around 0.41) that are superior to those of the corresponding aspatial models. Furthermore, taking spatial dependence into account results in even faster annual rates of convergence (181.08% and 191.73% in model 9a and 9b against 156.45% and 183.89% in model 6a and 6b for high decentralization levels).

Table 5 reports similar estimates for the variable LEB. Once again, the conditional approach increases the goodness of fit (adjusted R^2 from 10% to 11%), even though it just explains a small portion of the overall variance. Indeed, this result is consistent with the fact that the growth rate of LEB is likely to be influenced by many factors, such as socio-economic, demographic, technological, genetic, environmental ones, whose analysis is beyond the scope of this paper. Looking at the conditional variables DEC and GDP, they exert a positive and overall statistically significant impact on the growth rate of LEB. The interaction terms in Eq. (6) and (9) have the same statistically negative sign as the β coefficients, demonstrating a strong ability of decentralization to fasten the health convergence process: high levels of decentralization levels almost double the speed of convergence compared to low levels (Eq. 6a and 9a), thus halving the half-life.

Table 4

Conditional β-convergence for *IMR* (pooled sample)

	Growth rate - Infant mortality Rate (*IMR*)															
Regressors	Eq. (5a)		Eq. (5b)		Eq. (6a)		Eq. (6b)		Eq. (8a)		Eq. (8b)		Eq. (9a)		Eq. (9b)	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	T	Coef.	t
IMR	-0.718	-5.93***	-0.774	-5.93***	-0.053	-0.28	-0.521	-4.47***	-0.793	-7.16***	-0.804	-6.76***	-0.311	-2.49**	-0.595	-5.48***
FDEC	-0.113	-2.84**			0.682	2.88***			-0.0648	-1.77*			0.495	2.86***		
REF			-0.227	-3.12***			0.977	2.26**			-0.133	-1.76*			0.824	1.94*
INT					-0.205	-3.22***	-0.320	-2.67**					-0.146	-3.13***	-0.258	-2.19**
GDP	-0.471	-4.87***	-0.479	-4.91***	-0.410	-4.71***	-0.429	-4.77***	-0.324	-3.43***	-0.354	-3.36***	-0.297	-3.45***	-0.335	-3.36***
W*$_{IMR}$									0.359	2.98***	0.288	2.95***	0.317	3.16***	0.241	2.76**
_cons	4.399	5.80***	4.450	5.86***	1.624	1.75*	3.316	4.92***	2.761	3.88***	3.057	3.55***	0.977	1.10	2.368	2.98***
R^2	0.371		0.391		0.389		0.407		0.407		0.409		0.416		0.419	
Speed (b)	126.59%		148.72%		156.45% / 101.53% / 66.27%		183.89% / 73.61%		157.50%		162.96%		181.08% / 129.98% / 96.33%		191.73% / 90.39%	
Half-life	0.55		0.47		0.43 / 0.68 / 1.05		0.37 / 0.94		0.44		0.43		0.38 / 0.53 / 0.72		0.36 / 0.77	

Source: our elaboration on HFA and TPA data.

Notes: ***, **, and * denote significance at the 1%, 5% and 10% level. All variables expressed in natural log form. *t*-statistics computed with robust standard errors clustered at regional level.

Table 5

Conditional β-convergence for *LEB* (pooled sample)

Growth rate - Life Expectancy at Birth (*LEB*)

| Regressors	Eq. (5a)		Eq. (5b)		Eq. (6a)		Eq. (6b)		Eq. (8a)		Eq. (8b)		Eq. (9a)		Eq. (9b)	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
LEB	-0.0705	-9.54***	-0.0691	-7.56***	0.0223	0.68	-0.0168	-1.47	-0.0568	-3.65***	-0.057	-4.11***	0.0351	1.05	-0.0089	-0.93
FDEC	0.00045	2.39**			0.138	2.75**			0.00049	2.43**			0.137	2.82**		
REF			0.0004	1.32			0.344	4.25***			0.00063	1.90*			0.338	3.90***
INT					-0.0315	-2.74**	-0.0785	-4.26***					-0.0313	-2.81**	-0.0771	-3.89***
GDP	0.0013	2.53**	0.0012	3.26***	0.0011	2.05*	0.0011	2.95***	0.0013	2.62**	0.0013	3.56***	0.0011	2.22**	0.0012	3.03***
W*$_{LEB}$									-0.0175	-0.97	-0.020	-1.14	-0.0167	-0.94	-0.0148	-0.76
_cons	0.307	9.73***	0.302	7.50***	-0.098	-0.69	0.0733	1.49	0.323	8.85***	0.338	7.51***	-0.0812	-0.58	0.104	1.25
R^2	0.105		0.099		0.112		0.115		0.107		0.100		0.119		0.116	
Speed (*b*)	7.31%		7.16%		9.54% / 6.99% / 4.50 %		10.02% / 1.69%		5.85%		5.87%		8.07% / 5.57% / 3.13%		8.99% / 0.89%	
Half-life	9.48		9.68		7.26 / 9.91 / 15.40		6.92 / 40.91		11.85		11.81		8.58 / 12.44 / 22.16		7.71 / 77.53	

Source: our elaboration on HFA and TPA data.

Notes: ***, **, and * denote significance at the 1%, 5% and 10% level. All variables expressed in natural log form. *t*-statistics computed with robust standard errors clustered at regional level.

In contrast with the results for *IMR*, the spatial lag variable for *LEB* does not add much to the model in terms of goodness of fit since in either Eq. (8) or (9) the coefficient W_{LEB} is not significantly different from zero and the adjusted R^2 increases slightly.Jointly taking into account the results for both health outcomes, it can be observed that the effects of decentralization on the convergence process do not exhibit a uniform dynamics. Indeed, in either Eq. (6) and (9) the sign of the decentralization variable is the opposite of that of the interaction, thus exhibiting a moderating effect of the former. Indeed, a convergence process toward a common trend can be favoured by higher levels of decentralization but this does not necessarily ensure that the best performance in terms of growth rate is achieved. To better understand this effect, in Figures 7 and 8 the results from Eq. (9*a*) and (9*b*) are employed to depict the relationship between growth rates and starting levels of the health outcome variable according to the previously defined levels of decentralization (i.e. high, mean and low for *FDEC*; high and low for *REF*). More specifically, the predicted values of *IMR* (Figure 7) and *LEB* (Figure 8*)* growth rates are plotted, keeping constant at their mean values the other control variables. In Figure 7 (left and right hand sides), decentralization is shown to clearly improve the convergence process (the slope of the lines becomes higher as the level of decentralization increases). However, as *IMR* reach low values (i.e. to the left of the intersection of the lines), a moderating effect prevails, leading regions with a lower level of decentralization to overtake the other regions in terms of performance (i.e. less *IMR* growth rates). A similar effect appears also for the other health outcome variable (Figure 8, both sides) as regions overcome a given high level of *LEB* (i.e. to the right of the intersection of the lines). Therefore, for high levels of LEB a trade-off between decentralization and growth rates emerges: to secure higher growth rates a certain degree of decentralization has to be given up. To conclude,

decentralization appears to enhance the converge process by flattening the performance of the leading regions rather than by favouring their increase.
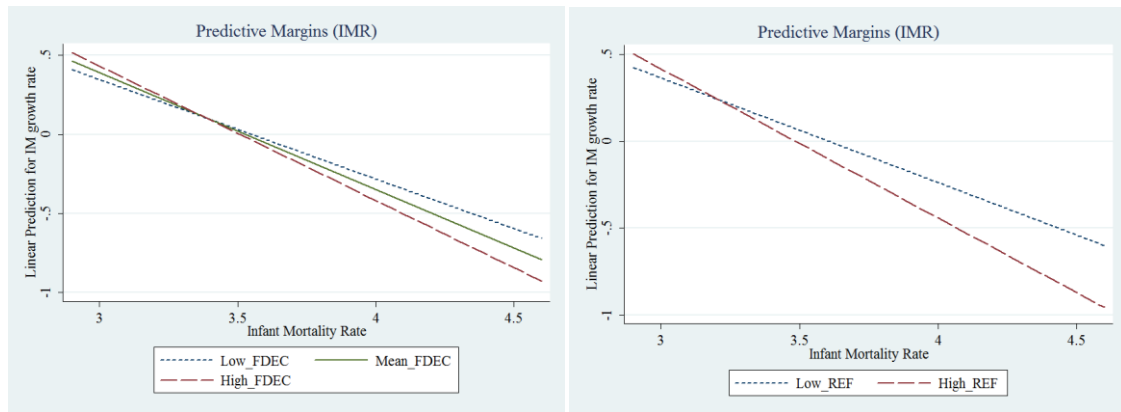
Fig. 7. Predictive margins for (log of) *IMR* – variables *FDEC* (left side) and *REF* (right side)
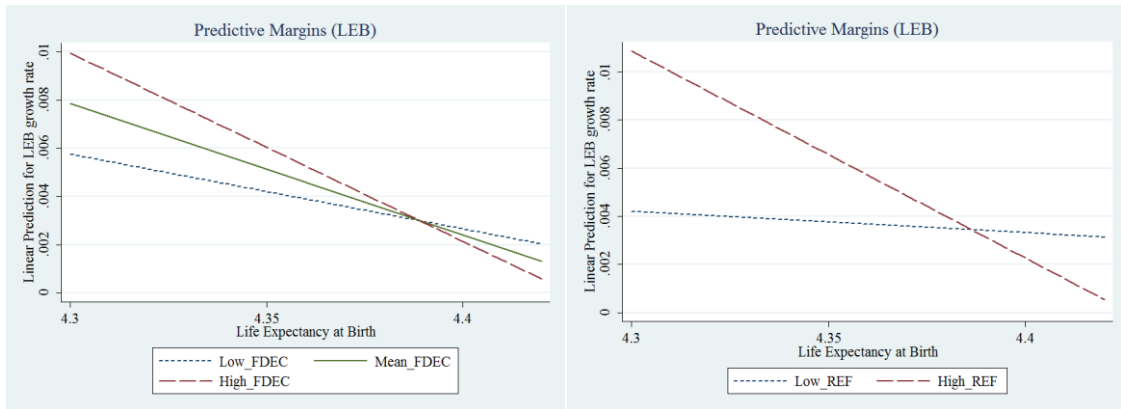
Fig. 8. Predictive margins for (log of) *LEB* – variables *FDEC* (left side) and *REF* (right side)

4. Robustness checks

In this section a battery of additional robustness tests is provided to assess the reliability of the previous results. Firstly, the potential pitfalls in estimating convergence by means of a pooled approach (as opposed to the more traditional cross-regional one) are considered. In the growth literature authors such Bianchi and Menegatti (2007) have argued that pooled convergence models are not able to distinguish between two different simultaneous phenomena: the possible tendency for poor countries to grow faster than rich ones and the possible tendency for each country to grow at a decreasing rate over time. The risk is that of reaching misleading conclusions on the actual convergence, even in presence of a negative sign of the β coefficient. To overcome this problem, it has been suggested to augment the original pooled β-convergence models with time-specific fixed effects, which in the case of this paper can be done just for those specifications including the continuous variable of decentralization (i.e. *FDEC*).

Secondly, the robustness of the findings with respect to alternative specifications of spatial dependence is also verified. Particularly, both spatial error models and spatial lag models are estimated based on maximum likelihood (ML)[*]. From a methodological point of view, a spatial error model assumes spatially auto-correlated error terms. Formally, it requires to rewrite Eq. (6) as:

$$ln\left(\frac{y_{i(t+1)}}{y_{i,t}}\right) = \alpha + \beta \ln(y_{i,t}) + \gamma \ln(DEC_{i,t}) + \delta \ln(GDP_{i,t}) + \theta(INT_{i,t}) + \varepsilon_{i,t} \quad (10)$$

where $\varepsilon_{i,t} = \lambda W^* \varepsilon_{i,t} + u_{i,t}$, while all other terms are as previously defined.

A way of dealing with substantive dependence alternative to cross-regressive ones is through the use of a spatial lag model, which implies to modify Eq. (9) as follows:

---

[*] For a more detailed explanation of the differences among the alternative specifications of spatial dependence see Ray and Montouri, 1999.

$$ln\left(\frac{y_{i(t+1)}}{y_{i,t}}\right) = \alpha + \beta \ln(y_{i,t}) + \gamma \ln(DEC_{i,t}) + \delta \ln(GDP_{i,t}) + \theta(INT_{i,t}) +$$

$$\mu\, W^* ln\left(\frac{y_{i(t+1)}}{y_{i,t}}\right) + \varepsilon_{i,t} \qquad (11)$$

where the term $W^* ln\left(\frac{y_{i(t+1)}}{y_{i,t}}\right)$ is now the spatial lag of the growth rates of the

health outcomes (either *IMR* or *LEB*).

Results of all robustness tests for the models with interaction and spatial effects
are presented in Tables 6 and 7 for *IMR* and *LEB*, respectively. As for mortality, the
previous findings remain overall valid when time-specific fixed effects as well as
different specifications of spatial dependence are taken into account. In particular, the
speed of convergence under the hypothesis of a high level of decentralization is almost
twice that under the hypothesis of a low level of decentralization. Looking at the results
for *LEB*, when time fixed-effects are included in the models, both the coefficients of
FDEC and the interaction terms are always not statistically significant, thus denying any
role of decentralization in shaping the convergence process dynamics.

Table 6
Robustness checks for *IMR*

Regressors	Cross-regressive (OLS)		Spatial Error (ML)		Spatial Lag (ML)	
	Coef.	t	Coef.	t	Coef.	t
IMR	-0.369	-1.76*	-0.311	-1.38	-0.299	-1.30
FDEC	0.496	2.24**	0.575	2.27**	0.570	2.24**
INT	-0.140	-2.22**	-0.158	-2.39**	-0.159	-2.37**
GDP	-0.378	-3.44***	-0.473	-6.87***	0.467	-6.83***
W$_{IMR}$	0.196	1.78*				
_cons	1.94	1.40	2.766	2.80***	2.726	2.71***
Time Dummies	YES		YES		YES	
R^2	0.461		-44.34 (log l.) 0.462 (variance ratio)		-42.65 (log l.) 0.461 (variance ratio)	
Speed (b)	206.26% / 146.23% / 108.98%		211.75% / 143.40% / 103.18%		205,00% / 139.63% / 100.44%	
Half-life	0.34 / 0.47 / 0.64		0.33 / 0.48 / 0.67		0.34 / 0.50 / 0.69	

Source: our elaboration on HFA and TPA data.
Notes: ***, **, and * denote significance at the 1%, 5% and 10% level. All variables expressed in
natural log form. *t*-statistics computed with robust standard errors clustered at regional level.

Table 7
Robustness checks for *LEB*

Regressors	Cross-regressive (OLS)		Spatial Error (ML)		Spatial Lag (ML)	
	Growth rate *LEB*					
	Coef.	t	Coef.	t	Coef.	t
LEB	-0.0551	-2.60**	-0.051	-1.50	-0.0597	-1.77*
FDEC	0.0017	0.07	0.01	0.21	-0.0048	-0.10
INT	-0.0004	-0.07	-0.0023	-0.22	0.0011	0.10
GDP	0.0014	2.65**	0.0013	2.86***	0.0012	2.61***
W_{LEB}	-0.0157	-0.57				
_cons	0.307	2.42**	0.22	1.50	0.259	1.77*
Time Dummies	YES		YES		YES	
R^2	0.622		1639.44 (log l.) 0.622 (variance ratio)		1339.59 (log l.) 0.623(variance ratio)	
Speed (*b*)	5.82% / 5.79% / 5.76%		6.11% / 5.93% / 5.75%		5.74% / 5.82% / 5.91%	
Half-life	11.91 / 11.97 / 12.04		11.34 / 11.69 / 12.06		12.09 / 11.90 / 11.73	

Source: our elaboration on HFA and TPA data.
Notes: ***, **, and * denote significance at the 1%, 5% and 10% level. All variables expressed in natural log form. *t*-statistics computed with robust standard errors clustered at regional level.

5. Conclusions

This paper contributes to the small amount of empirical literature on the impact of decentralization as well as neighbours' effects on the regional convergence process of health outcomes. By analysing data on infant mortality rates and life expectancy at birth in Italy during the period 1996-2012, no reduction is found in dispersion levels but evidence of unconditional *β*-convergence among regions. Independently from the way the degree of regional decentralization is actually measured - whether by means of a continuous fiscal indicators or by means of a dichotomous variable capturing legislative changes -, decentralization enhances catching-up effects for both health outcomes, pushing up the speed of convergence. However, the effect of decentralization on health

outcomes' convergence does not seem to follow a standard uniform path. Once an acceptable level of health performance is achieved by a region, the beneficial effects of a higher degree of decentralization, though always present, should be traded-off against further increases in performance (in terms of growth rates).

Accounting for the spatial effects among regions in estimations does not change the overall emerging picture but results in less misspecified models. Finally, after controlling for potential biases arising from the use of pooled data, decentralization continues to improve the convergence process of *IMR*. However, similar statistically significant conclusions cannot be reached regarding the positive effects of decentralization on the convergence process of *LEB*, probably because of the limited time span used in this paper. Indeed, a 17-year could be a not enough long period to fully capture the convergence dynamics of the *LEB* variable since this outcome indicator is, by its nature, quite insensitive to short-run shocks (e.g. policy changes) but is likely to be affected by many factors in the long-run.

References

Anselin, L., Rey, S. (1991). Properties of tests for spatial dependence in linear regression models. *Geographical analysis*, *23*(2), 112-131.

Anselin, L. (1993). *The Moran scatterplot as an ESDA tool to assess local instability in spatial association*. Morgantown, WV: Regional Research Institute, West Virginia University.

Arbia, G. (2006). *Spatial econometrics: statistical foundations and applications to regional convergence*. Springer Science & Business Media.

Barro, R.J., Sala-i-Martin, X., Blanchard, O.J., Hall, R. E. (1991). Convergence across states and regions. *Brookings papers on economic activity*, 107-182.

Barro, R.J., Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy*, 100(2), 223-251.

Baumol, W.J. (1986). Productivity growth, convergence, and welfare: what the long-run data show. *The American Economic Review*, 1072-1085.

Becker, G.S., Philipson, T. J., Soares, R.R. (2005). The quantity and quality of life and the evolution of world inequality. *The American Economic Review*, *95*(1), 277-291.

Bianchi, C., Menegatti, M. (2007). On the potential pitfalls in estimating β-convergence by means of pooled and panel data. *Applied Economics Letters*, *14*(13), 963-967.

Cavalieri, M., Ferrante, L. (2016). Does fiscal decentralization improve health outcomes? Evidence from infant mortality in Italy. *Social Science & Medicine*, *164*, 74-88.

Channa, A., Faguet, J.P. (2016). Decentralization of health and education in developing countries: a quality-adjusted review of the empirical literature. *The World Bank Research Observer*, lkw001.

Chocholatá, M., & Furková, A. (2016). Does the location and the institutional background matter in convergence modelling of the EU regions?. *Central European Journal of Operations Research*, 1-19.

Clark, R. (2011). World health inequality: Convergence, divergence, and development. *Social Science & Medicine*, 72(4), 617-624.

d'Albis, H., Esso, L.J., Pifarré, H. (2012). Mortality convergence across high-income countries: An econometric approach. *Documents de Travail du Centre d'Economie de la Sorbonne*, Université Paris 1, 2012.76.

Dickovick, J.T. (2011). *Decentralization and recentralization in the developing world: Comparative studies from Africa and Latin America*. Penn State Press.

Edwards, R.D., Tuljapurkar, S. (2005). Inequality in life spans and a new perspective on mortality convergence across industrialized countries. *Population and Development Review*, *31*(4), 645-674.

Edwards, R.D. (2011). Changes in world inequality in length of life:1970-2000. *Population and Development Review*, 37(3), 499-528.

Falleti, T.G. (2005). A sequential theory of decentralization: Latin American cases in comparative perspective. *American Political Science Review*, *99*(03), 327-346.

Glei, D.A., Meslé, F., Vallin, J. (2010). Diverging trends in life expectancy at age 50: A look at causes of death. In Eileen, M; Crimins, S.H.P. and Cohen, B. (eds.).

International Differences in Mortality at Older Ages: Dimensions and Sources. pp. 103-151. The National Academies Press: Washington, DC.

Grisorio, M.J., Prota, F. (2015a). The short and the long run relationship between fiscal decentralization and public expenditure composition in Italy. *Economics Letters,* 130 (C), 113-116.

Grisorio, M.J., Prota, F. (2015b). The impact of fiscal decentralization on the composition of public expenditure: panel data evidence from Italy. *Regional Studies,* 49 (12), 1941-1956.

Jiménez-Rubio, D. (2011). The impact of fiscal decentralization on infant mortality rates: Evidence from OECD countries. *Social Science & Medicine, 73*(9), 1401-1407.

Levene, H. (1960). Robust tests for equality of variances1. *Contributions to probability and statistics: Essays in honor of Harold Hotelling*, *2*, 278-292.

Malesky, E.J., Nguyen, C.V., Tran, A. (2014). The impact of recentralization on public services: A difference-in-differences analysis of the abolition of elected councils in Vietnam. *American Political Science Review*, *108*(01), 144-168.

Maynou, L., Saez, M., Bacaria, J., Lopez-Casasnovas, G. (2015). Health inequalities in the European Union: an empirical analysis of the dynamics of regional differences. *The European Journal of Health Economics*, *16*(5), 543-559.

Montero-Granados, R., de Dios Jiménez, J., Martín, J. (2007). Decentralisation and convergence in health among the provinces of Spain (1980–2001). *Social Science & Medicine*, 64(6), 1253-1264.

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*(1/2), 17-23.

Plümper, T., Schneider, C. J. (2009). The analysis of policy convergence, or: how to chase a black cat in a dark room. *Journal of European Public Policy*, *16*(7), 990-1011.

Quah, D. (1993). Galton's Fallacy and the Convergence Hypothesis. *Scandinavian Journal of Economics,* 95, 427–443.

Rey, S.J., Montouri, B.D. (1999). US regional income convergence: a spatial econometric perspective. *Regional studies*, *33*(2), 143-156.

Sala-i-Martin, X. X. (1996). The classical approach to convergence analysis. *The economic journal*, 1019-1036.

Schmitt, C., Starke, P. (2011). Explaining convergence of OECD welfare states: a conditional approach. *Journal of European Social Policy*, *21*(2), 120-135.

Sen, A. (1998). Mortality as an indicator of economic success and failure. *The Economic Journal*, *108*(446), 1-25.

Shah, A., Thompson, T., Zou, H.F. (2004). The impact of decentralization on service delivery, corruption, fiscal management and growth in developing and emerging market economies: A synthesis of empirical evidence. *CESifo DICE Report*, 1/2004: 10-14

Treisman, D. (2007). What have we learned about the causes of corruption from ten years of cross-national empirical research?. *Annual Review of Political Science*, *10*, 211-244.

Wilson, C. (2011). On the scale of global demographic convergence 1950-2000. *Population and Development Review*, 27(1):155-171.

Young, A.T., Higgins, M.J., Levy, D. (2008). Sigma convergence versus beta convergence: Evidence from US county-level data. *Journal of Money, Credit and Banking*, *40*(5), 1083-1093.

CHAPTER 3[*]

Quality-enhancing incentive initiatives for hospital care: policy implications and management requirements

ABSTRACT

During the last decades, in many countries retrospective payment systems for hospital care have been replaced by prospective payment methods, mainly DRG-based ones. The latter, though encouraging a more efficient use of hospital resources, present several undesired effects, some of which are likely to jeopardize the quality of care. The aim of the paper is to examine the main issues in designing and implementing an effective scheme of financial incentives for enhancing quality of hospital care. The main international experiences regarding the adoption of hospital performance-based incentives as well as the existing empirical evidence on their capacity to effectively boost quality of care are also critically analysed. Finally, some managerial remarks are considered in order to provide indications for the concrete implementation process and to contribute to the definition of a quality-enhancing incentive scheme well suited to the peculiarities of each organization system.

Keywords: Payment systems, quality of hospital care, incentives, performance, P4P, PQ4

---

[*] Co-authored by Prof. Marina Cavalieri and Prof. Pierluigi Catalfo.

1. Introduction

Starting from the '80s, Diagnosis Related Groups (DRGs) have gradually became the prospective case-based hospital payment scheme in many countries. It was expected that the implementation of the new system would have primarily resulted in a more efficient use of resources by hospital providers (micro-efficiency) compared to the previous retrospective payment systems (RPSs). Over the time, however, several types of undesired effects of prospective payment systems (PPSs) have been described in the literature. Among others, potential perverse responses by hospitals include: upcoding, unbundling, cost-shifting and cream skimming (Cots et al., 2011). These strategic behaviours are likely to produce negative consequences in terms of both macro-efficiency and equity but have also the potential to adversely affect quality of hospital care (e.g. bloody discharges). Therefore, concerns have been raised that DRG-based payments might not contain enough explicit incentives to improve the quality of hospital care. Nonetheless, empirical evidence on the impact of DRG-based payment systems on quality of hospital care is quite inconsistent both in direction and magnitude and no clear-cut conclusions can be drawn from a systematic review of it (Or and Häkkinen, 2011).

In many countries, new strategy are now being implemented to incentivize quality improvements for inpatient care. The approaches follow two main directions. On the one hand, marginal changes to the design of existing DRG methods (e.g. more sophisticated risk adjustment and trimming methods, etc.) are being introduced to make less ambiguous their effects in terms of quality enhancement. On the other hand, DRG-based payments are being supplemented with specific incentives for achieving specific quality goals. The latter strategy includes the so called 'pay for performance' (P4P) or 'pay for quality' (P4Q) initiatives, which make use of explicit financial incentives to reach targets

on predefined performance measures. These initiatives, albeit different in nature, seem to respond to the common approach of moving the priority of policy makers away from the logic of achieving better 'value for money' in the hospital sector to that of paying for healthcare outcome (and not output) and explicitly rewarding quality of care.

The aim of this paper is to examine some key issues in designing and implementing a P4Q incentive model in hospital care and to report the most important international initiatives. In the final part of the work some managerial aspects are also underlined to better support the implementation process of quality incentive systems, taking into account that incentives should play the role of an alignment system for healthcare organization management.

2. Issues in designing effective quality improvement incentives for hospital care

The enthusiasm behind the quality enhancing experiences for hospital care is driven by the promising idea that providers respond to incentives. Therefore, if quality improvement incentives are created in a way that the objectives of both 'principal' and 'agent' are fully aligned, better care should occur (Dranove and White, 1987; Blomqvist, 1991; Blomqvist and Léger 2005). However, the relationship between incentives and performance has been proved not to be straightforward and many different factors are likely to affect it. Moreover, some aspects of quality are hard to assess (e.g. diligence of care) and incentives based only on verifiable aspects may decrease the effectiveness of the non-verifiable ones, thus reducing patients' benefit. Hospitals are, indeed, multi-tasking providers and a mix of incentives should be enhanced to motivate their effort across various tasks (Eggleston, 2005). Furthermore, whenever quality dimensions are

substitutes, the levels of altruism for which providers care about the welfare of the patients may also impact on the strength of the optimal incentive scheme that should be implemented (Kaarboe e Siciliani, 2011).

As suggested by Christianson and Conrad (2011), the desired outcomes are not likely to be achieved by financial incentives alone; rule-based strategies mainly rooted on regulation (e.g. accreditation rules) and monitoring are required as well. The design, adoption, monitoring and fine-tuning of financial and non-financial incentives require, as an essential prerequisite, the development of a good-quality health information and reporting system. Without them, not only what might be theoretically the "best" approach is *de-facto* constrained by what it is actually possible to implement but also payment schemes are more likely to be "gamed" and, thus, to lose credibility in the eyes of the providers they are targeted on. The implementation of a valid information and reporting system is, however, a complex, costly and time-consuming operation, which could potentially conflict with the need by many countries to implement quality improvement strategies as soon as possible.

Literature has highlighted various theoretical and practical issues in designing the right incentives to encourage quality improvement and P4P (Rosenthal and Dudley, 2007; Cronwell et al., 2011; Charlesworth et al., 2012; Eijkenaar, 2013; Van de Voorde et al. 2013). Though the topic is very complex, in the following the most relevant ones will be briefly summarized.

Quality indicators

The choice of quality indicators and the way in which quality achievements should be measured are strictly dependent on the agreed definition of quality of hospital care, which is, however, out of the scope of this paper. Here, it is just worth reminding that quality of hospital care is a multidimensional concept embodied in structures, processes and outcomes (Donabedian, 1988). Although the former two concepts are imperfect surrogates for the latter, whose improvements should be the ultimate goal of any quality enhancing scheme, they are used frequently because of the difficulty in directly measuring outcomes. Moreover, compared to outcomes, they are clearly under the control of providers and, thus, do not require to be risk adjusted.

Economic theory suggests that the choice of the objective quality measure should be contingent to the level of knowledge of patients' health and its production function as well as the way this information is shared among patients, providers and payers (Nicholson et al., 2008). As a general rule, the more information is imperfect, the more providers' incentives should be directly based on outcome measures. To secure a high engagement in the incentive scheme, a key factor is also that the indicators chosen are clinically meaningful for both clinicians and patients and represent important aspects of hospital care (Jha, 2013).

Evidence on the effects of different measures in enhancing quality of care is scarce but a systematic review by Van Herck et al. (2010) found larger effects when measures with more room for improvement (i.e. process ones) are used. In this respect, combining process and intermediate outcome measures could represent a solution for eluding the pitfalls of the former alone that can encourage gaming, while avoiding the disadvantage of basing incentives solely on outcomes that may be difficult to achieve and somewhat beyond the control of the provider. The advantages of using many different measures

should, however, be traded off against the increased costs and complexity of the programme, which are likely to reduce providers' adherence to it (Eijkenaar, 2013).

Performance benchmarks

Once quality indicators have been chosen, benchmarks can be define as *absolute* (achieving or not a pre-determined target), *relative* (being among the top x%) or *improved* (positive changes in performance). Compared to relative benchmarks, absolute ones are generally considered more transparent and straightforward since the incentive gained does not depend on other providers' achievements. They, however, are not adequate when the objective of the scheme is to improve the overall performance of the hospital system. Indeed, under this approach, while providers already at or above the targets at the baseline will be rewarded without being incentivized to improve their performance, providers with improvements either beyond or not reaching targets will not receive any payment as compensation for their 'investment' in quality. Differentiating performance targets across groups depending on groups' baseline performance could help alleviate the problem.

The most used relative benchmark approach ("tournament approach") involves hospital providers to outperform others in order to be eligible for the performance payment. It, thus, stimulates continuous and higher levels of improvement by single providers because no one knows in advance how high performance is required in order to get the payment. However, it is expected not to foster collaboration and dissemination of best practices among providers but has the further advantage that the funder can calculate *ex-ante* the total amount of incentive payments, keeping better costs under control.

Under the 'improved performance' approach, each hospital provider who improves compared to the past performance is rewarded. This method is, however, considered to be inequitable since all performers are rewarded in the same way, irrespective of the fact that improving from a low rather than a high starting point is easier. A possible solution to the problem could be rewarding providers in proportion to their improvement achievements or to phase out rewards based on improvement after a given period of time (Cromwell et al., 2011).

Recent reviews of the empirical literature have found more positive effects when absolute rather than relative targets are used, though the relationship is not always straightforward (Van Herck et al., 2010; Eijkenaar et al., 2013). The choice of the benchmark is ultimately a critical decision that each quality improvement scheme should tailor to its specific goals (Werner and Dudley, 2009). Current quality incentive schemes seem not to consider the various approach as mutually exclusive but complementary. Once again, however, the potential advantages of a combined scheme in terms of both quality incentive effects and ability to respond to multiple objectives should be weighed against the disadvantage in terms of improved complexity, which is likely to adversely affect effectiveness (Conrad and Christianson, 2004).

Voluntary versus mandatory quality improvement schemes

The major advantage of a voluntary programme is that it is easier to be implemented since not all providers need be willing to participate. However, they are expected to create 'adverse selection' problems by attracting good performers while leaving out the poor ones. As for mandatory programmes, their main advantage is fairness and the ability to promote quality across the overall system. The actual

realization of the latter objective is, however, highly influenced by the structure of the programme's payment, which can sometimes make voluntary and mandatory programmes comparable in terms of responses to incentive schemes.

Rewards versus penalties: the 'non-payment for non-performance' approach

Quality financial incentives can be structured as either positive (e.g. rewards, bonuses, etc.) or negative (e.g. penalties, withholds, etc.). Though economic theory predicts that behavioural responses by risk adverse individuals are expected to be larger if individuals perceive the incentive as a potential loss as opposed to a potential gain (Kahneman and Tversky, 1979), negative incentives are generally considered as unfair since providers' income (absolute or relative) position deteriorates. Therefore, by inducing negative reactions in terms of participation and compliance rates and by favouring opportunistic behaviours, the net effect of a 'stick' approach on quality of hospital care could be far from the one predicted by the theory. On the opposite, the adoption of a 'carrot' approach, though found to be more effective (Van Herck et al., 2010), could be questionable from an ethical standpoint since paying hospital providers for their performances comes down to reward what in fact should be standard of hospital care (Dudley and Rosenthal, 2006). The various options have also different consequences in terms of payer's funding commitment, since positive incentives requires 'new money' rather than just allocating existing funds (penalties) or generating savings (withholds). The choice of the type of incentives ultimately depends on the nature of the quality improvement scheme (negative incentives could be adopted only if the programme is mandatory) and the ability to effectively monitor providers' behaviour. An effective compromise may prove to be using penalties to set a minimum baseline for performance expectations and rewarding those providers above the floor. As overall

performance improves, the floor could be moved upward over time. Alternatively, providers could be rewarded for their performance, without paying for the extra costs of conditions that are considered as avoidable (Rosenthal, 2007). Two versions of this approach, called 'non-payment for non-performance' (NP4NP), exist. In a most narrow implementation ('never events' approach), payments are not provided for those complications that are clearly preventable and should never occur (e.g. wrong site surgery or retained foreign object post-operation). A wider implementation comprises all hospital-acquired conditions (HACs; e.g. infections, falls and trauma following a select surgery) or readmissions. Such initiatives have been argued to be unfair under the assumption that some complications are not fully preventable with current medical knowledge. A possible solution to the problem could be to identify a risk-adjusted rate of complications for each hospital, and then not to fund activities in hospitals with a rate of complications higher than the expected one (Fuller et al., 2011).

Generally speaking, the NP4NP approach is expected to improve patient safety and quality of care by increasing compliance to evidence-based guidelines and promoting innovations as well as organizational and cultural changes. Over time, all these should result in a reduction of costs due to medical errors. However, some scholars (Wachter et al., 2008; Mookherjee et al., 2010) have underlined that, under this approach, only the included complications will receive extra attention, perhaps at the cost of other equally important medical areas not covered by the policy (the so-called 'teaching to the test' problem described by Holmstrom and Milgrom, 1991). Moreover, providers may also be induced to adopt defensive measures (e.g. large use of antibiotics to prevent infections), to "cream skim" patients according to their probability of developing HACs and to behave strategically (e.g. change documentation to minimize the negative impact of the scheme).

The level of payment adjustment

There are three possible options for adjusting hospital payments for the quality of care provided, respectively at (Or and Häkkinen, 2011): 1) hospital level; 2) DRG level; 3) individual patient level. Under the first option, hospital-level quality indicators are used to reward hospitals for improvements in the quality of care provided. This system does not require patient-level data but it is only effective whenever quality is independent on the volume of activity. When patient level data on outcomes and/or treatments is available, it is possible to adjust DRG-based payments for the quality of all patients treated within a given DRG. This involves moving away from payments by reference to average costs to payment based on best practice. The third option requires to provide adjustments at the individual patient level. It, however, implies the availability of reliable indicators of patient outcomes, which are not always easily identifiable. Under the latter scheme, quality can also be integrated into the DRG-based payment for the individual patient by extending the treatment episode for which a DRG based payment is granted (e.g. to include outpatient, readmissions, etc.).

The size of the incentive payment

The size of the incentive payment is a central consideration in the design of quality improvement programmes. As correctly argued by Jha (2013), incentives that target organizations should be more rationally designed than those that appeal to individuals since too little incentives risk "to fail to motivate organizations to invest in care redesign and quality improvement and will have little benefit for patient care".

Economic theory suggests that reward should be commensurate with the incremental net costs of undertaking the desired action, including the lost revenue

(opportunity cost) that the provider could generate in other activities (i.e., the 'participation constraint' in an agency theory framework). Though this prediction seems to further support the use of rewards as opposed to penalties, it does not tell anything about the 'optimal' size of the payment. Moreover, in presence of a diminishing marginal utility of income, the relationship between incentive size and performance is expected to be positive with diminishing marginal increases in performance above a certain payment level (the well-known 'target-income hypothesis", described by Rizzo and Blumenthal, 1996; Rizzo and Zeckhauser, 2003). Additionally, large payments are likely to 'crowd out' providers' intrinsic motivations, determining undesired behaviours by providers (Frey, 1997; Deci et al., 1999; Damberg et al., 2007).

It has sometimes been suggested that the size of the incentive payments should rather reflect the benefit of the targeted activity. However, this 'shared savings' approach does not ensure that the reward will be sufficient to cover the full cost of quality improvement, thus discouraging providers to exert the required effort (Eijkenaar, 2013). Accounting for the cost of meeting the performance target may help to increase the effectiveness of the incentive, though it should be noticed that the stream of costs that the adherence to a given project will produce during its life it is not constant, thus requiring to vary payments over the time accordingly.

Very little empirical research exists on the dose-response relationships of quality improvement schemes, especially in the hospital sector. Randomized controlled trials of P4P are rare and small in scale, thus explaining the inconsistency of the findings (Rosenthal et al., 2005; Christianson et al., 2007; Felt-Lisk et al., 2007; Mullen et al., 2010). In a systematic review of the literature, Van Herck et al. (2010) find no clear-cut relationship between incentive size and the reported P4P results.

Payment frequency

Theoretically, a high payment frequency is expected to contribute to incentive strength since future gains are discounted at a positive rate. This is also preferred due to the existence of a diminishing marginal utility of income, which makes less effective a large lump-sum payment than a series of smaller, more frequent, payments. Finally, a high payment frequency is expected to increase incentive salience (Eijkenaar, 2013). Contrarily, the need to collect data on performance and validate them may require to delay payments and to reduce the number of them. Furthermore, when payments are expected to reward performance on outcomes that occur in the long run, a high payment frequency is likely to be a feasible solution.

Empirical evidence on this point is once again quite inconclusive (Chung et al., 2009). Although variation in evaluated performance measures makes difficult to compare results, Emmert et al. (2012) found that programmes with little delay between care delivery and payment were all relatively successful.

3. International quality-enhancing initiatives for hospital care

The most interesting initiatives concerning the adoption of quality enhancing incentives for hospital care are those developed in the United Kingdom and the United States. Minor experiences of P4P programmes exists in other European countries, especially Nordic ones.

In England, attempts to address quality concerns under a DRG-based PPS are included in the so-called Payment by Results (PbRs) scheme, which comprises four main

elements: Commissioning for Quality and Innovation (CQUIN), Best Practice Tariffs (BPTs), the Advancing Quality (AQ) and the 'never events' initiatives.

Introduced in 2009, the CQUIN scheme operates, at hospital level, through a list of quality improvement (in terms of safety; effectiveness and patient experience) and innovation goals (DoH, 2008a and 2008b), which are periodically agreed between hospitals and local commissioners to ensure flexibility to local priorities and to generate local enthusiasm. A small amount of providers' income (reaching the 2.5% of the actual contract value in 2015/16) is conditional on achieving these goals. Accounting for the risk of selection bias in estimation, McDonald et al. (2013) found no evidence of an impact of the content of local schemes on performance improvement, except for hip fracture. In another evaluation, Kristensen et al. (2013) found the scheme to be largely unsuccessful in generating local enthusiasm and addressing the requirements set by the Department of Health (DoH), concluding that "..a somewhat firmer national framework would be preferable to a fully locally designed framework".

Since 2010, the National Health Service (NHS) has incrementally introduced BPTs, at individual patient level, for a selection of high impact (i.e. high volumes, significant variation in practice or significant impact on outcomes) Healthcare Resource Groups (HRGs), the English version of DRGs. Under this scheme, payment is made dependent on whether or not the treatment is provided in the most appropriate setting or whether the best practice clinical treatment is offered. Specifically, BPTs imply two payment components: a base tariff which is paid to all activity irrespective of whether the characteristics of best practice are met, and the 'conditional' component, which is paid only if the treatment meets several characteristics of evidence-based best practice. The price of the conditional component is calculated on the basis of the additional costs to deliver best practice while the base tariff is set below the national average cost. This

tariff structure is intended to provide hospitals that are below average performers a further financial incentive to change their practice style. An evaluation of the impact of the first year of the BTP introduction on the quality of hospital care and patient outcomes has been made by McDonald et al. (2012), on commission of the DoH. Using a difference-in-differences (DiD) approach, they found that hospitals responded quickly to the increase in price for day-case cholecystectomy (+7 percentage points) but no beneficial impact of the stroke BPT on the selected process and outcome indicators. In contrast, the hip fracture BPT had substantial effects, being associated with: +4 percentage point in receipt of surgery within 48 hours of admission, -0.7 percentage point in mortality, and +2.1 percentage point in the proportion of patients discharged home within 56 days. Such differences in impact among treatment areas are attributable to the different structures of the BPTs, as the tariff for hip fracture was only paid if all criteria were met whereas providers were rewarded separately for each indicator in the stroke BPT.

Introduced in 2008 in all NHS hospitals in the Northwest region of England, the AQ initiative uses a relative benchmark approach where performance is measured by 54 quality indicators based on patient-level data and covering the following clinical areas: acute myocardial infarction, coronary artery bypass graft, heart failure, hip and knee replacement surgery, pneumonia, dementia, psychosis and stroke. Once assured by independent auditors, indicators are publicly reported and used to reward the top and second quartiles of hospitals with 4% and 2% bonuses, respectively. Bonuses should then be reinvested by hospitals in quality programs. The AQ initiative has been evaluated by Sutton et al. (2012), who used a triple DiD analysis to compare changes over time in mortality of patients treated for acute myocardial infarction, heart failure and pneumonia with mortality in two control groups: patients admitted for the same three conditions to

the rest of England and patients admitted for the non-incentivized conditions in England as a whole. They found that the introduction of P4P was associated -1.3 percentage points in the combined mortality for the three conditions included in the AQ program but also underlined that it was not yet clear how and why this program was associated with reduced mortality when other similar US programs (e.g. the Premier Quality Incentive Demonstration) were found not to be. They concluded that, though the more positive results found in AQ program could be attributed to the universal participation of providers, the higher generosity of bonus payments and the collaborative nature of the scheme, the context in which the incentives are introduced played a crucial role. Recently, Meacock et al. (2014a) have shown that the AQ program was cost-effective during its first 18 months, by generating 5,227 Quality-Adjusted Life Years (equals to £105m) and £4.4m of savings in reduced length of stays.

England first adopted a 'never events' policy in 2009, where HRG payments are not made for those "serious, largely preventable patient safety incidents that should not occur if the available preventative measures have been implemented by healthcare providers" (NPSA, 2010). The list of never events defined by the National Patient Safety Agency currently comprises 14 incident types such as wrong site surgery and implant, retained foreign object post-procedure, wrong route administration of medication. In England, payments are also denied for those emergency readmissions within 30 days of a discharge that are related to the original hospital admission. Finally, a 70% reduced payment for emergency admissions above a specified volume exists. A formal evaluation of the 'never events' initiative has not been published yet but a report by the DoH (2012) shown an increase in the number of reported events in the period 2009-2013, though a direct comparison of the data is made difficult by the fact that the never events' list has been enlarged over time.

As for the US, the number of P4P initiatives in the health care sector has grown rapidly over time (Rosenthal et al., 2004), sponsored by private health organizations, state and federal government agencies. Concerning hospital care, the Medicare's Premier Hospital Quality Incentive Demonstration (HQID) was launched in 2003 to test if providing financial incentives to hospitals that demonstrate high quality performance in a number of areas of acute inpatient care would improve patient outcomes and reduce overall costs. The HQID operated for three years under a budget neutrality criterion (Cronwell et al., 2011). In its first phase, it rewarded only hospitals in the top two deciles of quality performance (with 2% and 1% of the DRG value, respectively), providing penalties for those hospitals in the bottom two deciles. The programme was redesigned in 2006 (second phase) to also reward hospitals that achieved significant improvements. Afterwards, it was replaced by the new Value-based Purchasing (VBP) programme, which now reduces hospitals' base DRG rates by a 1% to fund value-based incentive payments according to the overall performance of the hospital measured by a set of quality indicators (*i.e.* 13 clinical process indicators, 8 measures of patient satisfaction, 5 outcome measures and 1 measure of efficiency.). Empirical evidence on the effects of the HQID programme failed to demonstrate a significant effect on mortality within either its first three (Glickman et al., 2007; Ryan, 2009) or six years (Jha et al., 2012) of life. Few studies shown an improvement in quality of care, as measured by process indicators, for those hospitals participating to the first phase of the programme as compared to the others (Lindenauer et al., 2007) and for the shift from the first to the second phase (Werner et al., 2011) but no significant effect of the incentive changes introduced in 2006 for the lowest-performing hospitals (Ryan et al., 2012).

With regard to the NP4NP experiences, since 2008 Medicare no longer pays hospitals for additional costs associated with 11 categories of HACs, selected by the

Center for Medicare and Medicaid Services under the following four criteria: 1) high cost or high volume or both, 2) harm the patient, 3) result in the assignment of a case to a DRG that has a higher payment when present as a secondary diagnosis, and 4) could reasonably have been prevented through the application of evidence-based guidelines. In 2012, it was decided to extend the policy of non-payment for preventable complications also to Medicaid. Evaluations of the HAC programme have mainly focused on its ability to prevent hospital infections after surgery, showing inconsistent findings (Hoff and Soerensen, 2011; Lee et al., 2011 and 2012). As for readmissions, the Medicare Readmission Program does not use a 'no pay' approach but actually penalizes (up to a 3% reduction in payment) those hospitals which had above-average readmission rates within 30 days from June 2010 through July 2013 for acute myocardial infarction, heart failure, pneumonia, chronic obstructive pulmonary disease, hip and knee replacement (Averill et al., 2009; Medicare, 2013).

Among the other few European experiences, Sweden has not developed a national P4P program but some county councils have decided to add to the existing hospital activity-based funding a performance-based compensation, which covers up to 4% of total hospital payment. Generally, these county initiatives withhold payment if an hospital does not meet certain targets in terms of wait times, patient safety or clinical indicators (Anell et al, 2012). In 2012, the Danish Government formed a Committee with the aim of analysing the effectiveness of activity-based reimbursement schemes in enhancing quality of hospital care (Ministeriet for Sundhed og Forebyggelse, 2013). The resulting report, though underlining the many drawbacks of the current hospital paying approach, was very cautious about the adoption of a national P4P scheme but encouraged regions to develop pilot experiences (Meacock et al., 2014b). Currently, some Danish regions also reduce hospitals' reimbursement if a hospital's intensity of treatment is

higher than 1.5%. Since treatment intensity is measured as the sum of the DRG value of the production divided by the yearly total number of patients treated by the hospital, the policy implicitly disincentives readmissions (Kristensen et al. 2015). Since 2014, Norway has launched a three-year P4P initiative distributing 500m NOK to the four Regions, on the basis of regional indicators of process, outcome and patient experience of care (Meacock et al., 2014b). Finally, in Germany performance-based payments are mainly used for outpatient care. However, regional P4P schemes for improving quality of hospital care exist throughout the country, forming part of selected contracts between sickness funds and hospitals. According to them, hospitals receive higher payments for births and patients with cardiac surgery and hip implants if they scores above the national average on a set of quality indicators. With regard to readmissions, German hospitals do not receive a second DRG payment if a patient is readmitted for the same condition within 30 days after discharge.

4. Some final managerial remarks

The peculiarities of the previously reported quality-enhancing indicators for hospital care seem to prevent a generic automatic approach in their adoption. From a managerial perspective, the specific circumstances and characteristics of the organization in which the performance indicators are used should be carefully considered before selecting them. In particular, a full understanding of the organization's cultural approach is crucial to methodologically implement an incentive system. In this respect, it could be said that no incentive system is appropriate for every context or organization (Van Dooren et al., 2010). The effectiveness of a given incentive indicator depends, among others, on the cultural managerial approach, the internal quality of the human capital's

commitment and the institutional and socio-economic environment. In particular, a key factor is the ability to manage internal resources, especially human ones, so as to improve employees' effort and to drive it towards greater efficiency and effectiveness.

Moreover, the organizational structure of the health care system and the complexity of the unitary management of some of its characterizing variables, such as volumes and quality features of expenditure, the equilibrium dynamics and the persistence of hospital units, the homologation of quality and efficacy levels of health services as well as of hospitals' actions, require to interpret the health system dynamics according to an inter-organizational approach (Benson, 1975; Osborn, e Hagedoorn, 1997; Ilyoo e Hong, 2002). Within this framework, incentive systems such as P4P and P4Q can be considered as useful tools to operatively realize the strategic alignment of the overall system (Dekker, 2004). This methodological prescription, though not exclusive, could represent an alternative or complementary option for the management of the inter-organizational systems, different from the logics suggested by the models of managerial decision-making based on transaction costs (Zajac e Olsen 1993; Speklé, 2001; Jha S., 2013). In this respect, the adoption of incentive methods is likely to produce a twofold effect: to consider in an unitary way all the organization levels (Klein et al., 2000) and to act at a level of single system entity, that is single hospital structure, so as to produce results that are coherent with the overall strategic direction planned for the health system as a whole. From this point of view, it becomes essential to find the way to align the main strategic objectives, that is the system of objectives that primarily constrains each of the system's entities to a unitary model.

Furthermore, overcoming the system of transaction costs and assuming incentive schemes as a model of organizational strategic alignment under a cognitive perspective is likely to favour both greater flexibility within the system and greater consideration of the

peculiarities of each hospital unit but also an higher level of internal and external accountability (Shapira, 2000; Solomon, 2007). Therefore, each item of the incentive system offers a unique occasion to impact on a specific strategic variable, with different potentials and various possibilities to determine effective managerial directions.

References

Anell, A., Glenngård A.H., and Merkur S. (2012). Sweden: Health system review. *Health Systems in Transition*, 14(5): 1-144.

Averill, R.F., McCullough, E.C., Hughes, J.S. et al. (2009). Redesigning the Medicare Inpatient PPS to reduce Payment to Hospitals with High Readmission Rates. *Health Care Financing Review*, 30(4): 1-15.

Benson, J.K. (1975). The Interorganizational Network as a Political Economy. *Administrative Science Quarterly,* 20: 229-249.

Blomqvist, A. (1991). The doctor as double agent: Information asymmetry, health insurance, and medical care. *Journal of Health Economics*, 10: 411-432.

Blomqvist, A., Léger, P.T. (2005). Information asymmetry, insurance, and the decision to hospitalize. *Journal of Health Economics*, 24(4): 775:793.

Charlesworth, A., Davies, A., Dixon, J. (2012). *Reforming payment for health care in Europe to achieve better value.* Research Report, Nuffield Trust.

Christianson, J., Leatherman, S., Sutherland, K. (2007). *Financial Incentives for Healthcare Providers and Quality Improvement: A review of the evidence*, The Health Foundation.

Christianson, J.B., Conrad, D. (2011). Provider Payment and Incentives. In Glied, S., Smith, P. (Eds.), *The Oxford Handbook of Health Economics*, New York: Oxford University Press Inc., 26: 624-648.

Chung, S., Palaniappan, L., Wong, E., et al. (2009). Does the frequency of pay-for-performance payment matter? Experience from a randomized trial. *Health Service and Research*, 45(2): 553–564.

Conrad, D.A., Christianson, J.B. (2004). Penetrating the 'black box': financial incentives for enhancing the quality of physician services. *Medical Care Research Review*, 61(3): 37S-68S.

Cots, F., Chiarello, P., Salvador, X., Castells, X., Quentin, W. (2011). DRG-based hospital payment: Intended and unintended consequences. In R. Busse, A. Geissler, W. Quentin, M. Wiley (ed)., *Diagnosis-Related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals*, Open University Press: Maidenhead.

Cromwell, J., Trisolini, M.G., Pope, G.C., et al. (2011). *Pay for Performance in Health Care: Methods and Approaches.* RTI Press, available at: http://www.rti.org/pubs/bk-0002-1103-mitchell.pdf.

Damberg, C. L., Sorbero, M. E., Mehrotra, A., Teleki, S., Lovejoy, S., Bradley, L. (2007). An environmental scan of pay for performance in the hospital setting: final report. *Washington, DC: Office of the Assistant Secretary for Planning and Evaluation (ASPE)*.

Deci, E.L., Koestner, R., Ryan, R.M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bullet*, 125(6): 627–668.

Dekker, H.C. (2004). Control of inter-organizational relationships: evidence on appropriation concerns and coordination requirements. *Accounting, Organizations and Society*, (29):27–49.

Department of Health (2008a). *Using the Commissioning for Quality and Innovation (CQUIN) Payment Framework,* London.

Department of Health (2008b). *High quality care for all: NHS next stage review final report*, London.

Department of Health (2012). *The Never Events Policy Framework: an update to the never events policy*, London.

Donabedian, A. (1988). The quality of care: How can it be assessed?. *The Journal of the American Medical Association*, 260(12): 1743-1748.

Dranove, D., White, W.D. (1987). Agency and the organization of health care delivery. *Inquiry*, 24(4): 405-415.

Dudley, R.A., Rosenthal, M.B. (2006). *Pay for Performance: A Decision Guide for Purchasers*. Report prepared for the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services, Rockville, MD, Available at: http://www.ahrq.gov/legacy/qual/p4pguide.pdf.

Eggleston, K. (2005). Multitasking and mixed systems for provider payment. *Journal of Health Economics*, 24(1):211–223.

Eijkenaar, F. (2013). Key issues in the design of pay for performance programmes. *European Journal of Health Economics,* 14(1): 117-31.

Emmert, M., Eijkenaar, F., Kemter, H. et al. (2012). Economic evaluation of pay-for-performance in health care: a systematic review. *The European Journal of Health Economics*, 13(6):755–67.

Felt-Lisk, S., Gimm, G., Peterson, S. (2007). Making Pay for Performance Work in Medicaid. *Health Affairs Web Exclusive*, 26(4): w516-17.

Frey, B.S. (1997). On the relationship between intrinsic and extrinsic work motivation. *International Journal of Industrial Organization*, 15(4), 427–439.

Fuller, R.L., McCullough, E.C., Averill, R.F. (2011). A new approach to reducing payments made to hospitals with high complication rates. *Inquiry*, 48(1): 68-83.

Glickman, S.W., Ou, F.S., DeLong, E.R., Roe, M.T., Lytle, B.L., Mulgund, J., Rumsfeld, J.S., Gibler, W.B., Ohman, E.M., Schulman, K.A., Peterson, E.D. (2007). Pay for performance, quality of care, and outcomes in acute myocardial infarction. *Journal of American Medical Association*, 297: 2373–2380.

Hoff, T.J., Soerensen, C. (2011). No payment for preventable complications: reviewing the early literature for content, guidance, and impressions. *Quality Management in Health Care*, 20(1): 62-75.

Holmstrom, B., Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization*, 7(1): 24-52.

Ilyoo, B., Hong, A. (2002). New framework for interorganizational systems based on the linkage of participants' roles. *Information & Management*, 39, 261–270.

Jha, A.K., Joynt, K.E., Orav, E.J., Epstein, A.M., (2012). The Long-Term Effect of Premier Pay for Performance on Patient Outcomes. *New England Journal of Medicine*, 366: 1606–1615.

Jha, A.K. (2013). Time to get serious about pay-for-performance. *Journal of American Medical Association*; 309(26): 123-135.

Jha, S. (2013). Learning orientation and performance orientation: scale development and its relationship with performance. *Global Business Review*, 14(1): 43-54.

Kaarboe, O., Siciliani, L. (2011). Multi-tasking, quality and pay for performance. *Health Economics*, 20(2): 225-238.

Kahneman, D., Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2): 263-91.

Klein, K.J., Palmer, S., Conn, A.B. (2000). Inter-organizational relationships: A multilevel perspective. In S. W. J. Kozlowski & K. J. Klein (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*, pp. 267–307, San Francisco, CA: Jossey-Bass.

Kristensen, S.R., McDonald, R., Sutton, M., (2013). Should pay-for-performance schemes be locally designed? Evidence from the commissioning for quality and innovation (CQUIN) framework. *Journal Health Services Research and Policy*, 18, 38–49.

Kristensen, S.R., Bech, M., Quentin, W. (2015). A roadmap for comparing readmission policies with application to Denmark, England, Germany and the United States. *Health Policy*, 119(3): 264–273.

Lee, G.M., Hartmann, C.W., Graham, D. et al. (2011). Perceived impact of the Medicare policy to adjust payment for health care-associated infections. *American Journal of Infection Control*, 40(4): 314-319.

Lee, G.M., Kleinman, K., Soumerai, S.B. et al. (2012). Effect of Nonpayment for Preventable Infections in U.S. Hospitals. *New England Journal of Medicine*, 367(15): 1428-1437.

Lindenauer, P.K., Remus, D., Roman, S., et al. (2007). Public reporting and pay for performance in hospital quality improvement. *New England Journal of Medicine*, 356(5): 486–96.

McDonald, R., Zaidi, S., Todd, S., et al. (2012). *A qualitative and quantitative evaluation of the introduction of Best Practice Tariffs*. Report commissioned by the Department of Health.

McDonald, R., Kristensen, S.R., Zaidi, S., Sutton, M., Todd, S., Konteh, F., Hussein, K., Brown, S., (2013). *Evaluation of the Commissioning for Quality and Innovation Framework*. Final Report.

Meacock, R., Kristensen, S.R., Sutton, M. (2014a). The cost-effectiveness of using financial incentives to improve provider quality: a framework and application. *Health Economics*, 23(1):1-13.

Meacock, R., Kristensen, S.R., Sutton, M. (2014b). Paying for improvements in quality: recent experience in the NHS in England. *Nordic Journal of Health Economics*, 2(1): 239-255.

Medicare (2013). *Hospital Readmissions Reduction Program*. Available at: http://www.medicare. gov/hospitalcompare/readmissionprogram.aspx

Ministeriet for Sundhed og Forebyggelse, (2013). *Bedre Incitamenter i Sundhedsvæsenet*. Copenhagen.

Mookherjee, S., Vidyarthi, A.R., Ranji, S.R., et al. (2010). Potential Unintended Consequences Due to Medicare's 'No Pay for Error Rule'? A Randomized Controlled Trial of an Educational Intervention with Internal Medicine Residents. *Journal of General Internal Medicine*, 25(10); 1097-1101.

Mullen, K.J., Frank, R.G., Rosenthal, M.B. (2010). Pay-for-performance and the quality of healthcare providers. *Rand Journal of Economics*, 41(1): 64-91.

National Patient Safety Agency (2010). *Never Events Annual Report 2009/10*. NPSA, available at: http://www.nrls.npsa.nhs.uk/resources/collections/never-events/?entryid45=83319.

Nicholson, S., Pauly, M.V., Ya, A. et al. (2008). Getting Real Performance Out of Pay-for-Performance. *The Milbank Quarterly*, 86(3): 435-457.

Or, Z., Häkkinen, U. (2011). DRGs and Quality: For Better or Worse?. In R. Busse, A. Geissler, W. Quentin and M. Wiley (ed.), *Diagnosis Related Groups in Europe. Moving Towards Transparency, Efficiency and Quality in Hospitals*, Maidenhead: Open University Press.

Osborn, R.N., Hagedoorn, J. (1997). The institutionalization and evolutionary dynamics of interorganizational alliances and networks. *Academy of Management Journal*, 40.2: 261-278.

Rizzo, J.A., Blumenthal, J.A. (1996). Is the target income hypothesis an economic heresy?. *Medical Care Research and. Review*, 53(3): 243–266.

Rizzo, J.A., Zeckhauser, R.J. (2003). Reference incomes, loss aversion, and physician behaviour. *Review Economics and Statistics*, 85(4): 909–922.

Rosenthal, M.B. (2007). Nonpayment for performance? Medicare's new reimbursement rule. *New England Journal of Medicine*, 357(16): 1573–1575.

Rosenthal, M.B, Fernandopulle, R., Song, H.R., Landon, B. (2004). Paying for quality: providers' incentives for quality improvement. *Health Affairs*, 23: 127–141.

Rosenthal, M.B., Frank, R.G., Zhonge, L., et al. (2005). Early Experience with Pay for Performance: From Concept to Practice. *Journal of the American Medical Association*, 294(14): 1788-179.

Rosenthal, M.B., Dudley, R.A. (2007). Pay-for-performance: will the latest payment trend improve care?. *Journal of American Medical Association*, 297(7): 740–744.

Ryan, A.M. (2009). Effects of the Premier Hospital Quality Incentive Demonstration on Medicare patient mortality and cost. *Health Services Research*, 44: 821–842.

Ryan, A.M., Blustein, J., Casalino, L.P. (2012). Medicare's flagship test of pay-for-performance did not spur more rapid quality improvement among low-performing hospitals. *Health Affairs*, 31(4): 797-805.

Salomon, J. (2007). *Corporate governance and accountability*. John Wiley and Sons, England.

Shapira, Z. (2000). Governance in organizations: a cognitive perspective. *Journal of Management and Governance*, 4: 53-67.

Speklé, R.F. (2001). Explaining management control structure variety: a transaction cost economics perspective. *Accounting, Organizations and Society* 26(4): 419-441.

Sutton, M., Nikolova, S., Boaden, R., Lester, H., Roland, M. (2012). Reduced mortality with hospital pay for performance in England. *New England Journal of Medicine*, 367(19): 1821-1828.

Van De Voorde, C., Gerkens, S., Van Den Heede, K., Swartenbroekx, N. (2013). *A comparative analysis of hospital care payments in five countries*. KCE Report, Belgian Health Care Knowlegde Centre.

Van Dooren, W., Bouckaert, G., Halligan, J. (2010). *Performance Management in the Public Sector.* London: Routledge.

Van Herck, P., Smedt, D., Annemans, L., et al. (2010). Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC Health Services Research*,10: 247.

Wachter, R.M., Foster, N.E., Dudley, R.A.(2008). Medicare's Decision to Withhold Payment for Hospital Errors: The Devil is in the Details. *The Joint Commission Journal on Quality and Patient Safety*, 34(2): 116-123.

Werner, R.M., Dudley, R.A. (2009). Making the 'pay' matter in pay-for-performance: implications for payment strategies. *Health Affairs* (Millwood). 28(5): 1498–1508.

Werner, R.M., Kolstad, J.T., Stuart, E.A., Polsky, D. (2011). The effect of pay-for-performance in hospitals: lessons for quality improvement. *Health Affairs*, 30(4): 690–8.

Zajac, E.J., Olsen C.P. (1993). From transaction cost to transactional value analysis: Implications for the study of interorganizational strategies. *Journal of Management Studies* 30(1): 131-145.

CONCLUSIONS

The present dissertation has tried to shed new light on the complex relationship between decentralization and public policy performance. Specifically, the issue has been investigated along a double research line: 1) at a macro level, by considering the effects of decentralization on regional health outcomes in Italy; 2) at a micro-level, by looking at how different degrees of provider's autonomy impact on the effectiveness of quality-enhancing incentive schemes for hospital care.

Some interesting insights derive from the previous chapters, which can help inform policy recommendations and provide further room for future researches. Firstly, an higher fiscal autonomy, *ceteris paribus*, is significantly associated with better regional health outcomes. However, the positive marginal effects of decentralization vary with the level of wealth, favouring poorer regions. In terms of policy implications, these findings highlight the importance of considering distributional issue when implementing a decentralization reform, along with the need for an adequate equalization system able to reduce the starting differences in term of resources among regions.

Secondly, the empirical analysis carried out in chapter 2 shows how decentralization favours convergence process, pushing up the speed of convergence. The study also employs a spatial econometric perspective with the aim to take into account the geographical dynamics among Italian regions. Nevertheless, fostering convergence doesn't imply straightforward the achievement of the highest level of performance. Once a certain level of healthcare outcome has been reached, more centralized policies seem to fit better for further improvements.

Finally, the last descriptive analysis highlights the importance, in a context of administrative decentralization, of designing appropriate incentives to improve providers' performance and, consequently, citizens' health.

A part from the specific conclusions drawn from each chapter, some more general comments can be made. The effects of decentralization are complex and intertwined. By itself, no single model of decentralization is able to guarantee the best results in terms of service delivery. This is because the impact of decentralization is highly context-specific and likely to depend on a variety of factors. In this uncertainty, only one thing seems to be plain: there is little to expect in terms of homogeneity. Notwithstanding, this work gives a significant contribution to the existing literature, defining decentralization as a useful tool able to affect the level of performance, and highlighting some aspects that should be carefully considered by policy-makers in designing an effective decentralization reform. Moreover, local levels play a crucial role in addressing this issue and incentives became critical motivators and performance drivers for gaining positive responses. Despite this, the high influence of socio-economic, cultural and legislative context, makes not easy to generalize the results to other countries. Decentralization is a more complex process than this research has tried to synthetically analyse, thus suggesting the need of further research in order to jointly consider the multi-dimensional aspects of this phenomenon and its effects on various measures of citizens well-being according to different governmental settings.