Università degli Studi del Molise
Dipartimento di Economia

Economics and Statistics Discussion Paper
No. 090/23

Enhancing Healthcare Cost Forecasting: A Machine Learning Model for Resource Allocation in Heterogeneous Regions

Nicola Caravaggio
Giuliano Resce

(This page intentionally left blank.)

Enhancing Healthcare Cost Forecasting: A Machine Learning Model for Resource Allocation in Heterogeneous Regions*

Nicola Caravaggio[a,†], Giuliano Resce[a,††]

[a]Department of Economics, University of Molise, Campobasso, Italy

[†]nicola.caravaggio@unimol.it, [††]giuliano.resce@unimol.it

Abstract

Accurate forecasting of healthcare costs is essential for making decisions, shaping policies, preparing finances, and managing resources effectively, but traditional econometric models fall short in addressing this policy challenge adequately. This paper introduces machine learning to predict healthcare expenditure in systems with heterogeneous regional needs. The Italian NHS is used as a case study, with administrative data spanning the years 1994 to 2019. The empirical analysis utilises four machine learning algorithms (Elastic-Net, Gradient Boosting, Random Forest, and Support Vector Regression) and a multivariate regression as a baseline. Gradient Boosting emerges as the superior algorithm in out-of-the-sample prediction performances; even when applied to 2019 data, the models trained up to 2018 demonstrate robust forecasting abilities. Important predictors of expenditure include temporal factors, average family size, regional area, GDP per capita, and life expectancy. The remarkable effectiveness of the model demonstrates that machine learning can be efficiently employed to distribute national healthcare funds to areas with heterogeneous needs.

Keywords: Machine Learning, National Health System, Healthcare expenditure.

JEL Classification: C54, H51, I10.

*This work was fully developed and funded by two primary research initiatives. One of these initiatives falls under Spoke 5 of the project Age-IT *Consequences and challenges of aging*. This project is one of 14 nationally funded Extended Partnerships under the National Recovery and Resilience Plan (NRRP), through Mission 4 *Education and Research*, component 2 *From research to business*. The second project, titled *Reconciling efficiency and equity in the Italian healthcare system after the pandemic crisis*, received funding from the 2022 National Interest Research Projects (PRIN) call, with the project code: 2022NKLHHT. While the authors bear sole responsibility for the content presented in this paper, gratitude is extended to Professors Cecilia Tomassini, Raffaele Lagravinese, and Vincenzo Carrieri, as well as the collaborative group of researchers engaged in the two funded projects. Their contributions have been immensely valuable.

1 Introduction

Predicting healthcare expenditure provides crucial information for decision-making, policy formulation, financial planning, and resource management. It supports the effective and efficient delivery of healthcare services, ensuring that individuals and societies can access affordable and sustainable healthcare. From a policy perspective, accurate predictions play a fundamental role, as they are used by several National Health Systems (NHS) to allocate funds at the territorial level to guarantee equity. Especially notable is the concept of horizontal equity, which stipulates that individuals with comparable health requirements should enjoy equitable entry to healthcare services, irrespective of their geographical location (Gravelle, Santos, & Siciliani, 2014; Rice & Smith, 2001). The pragmatic execution of this principle encounters various challenges, primarily stemming from the fact that decision-makers often lack the requisite information to gauge healthcare needs accurately. Hence, a majority of frameworks for distributing national resources rely on statistical models calibrated to healthcare service usage patterns. A case in point is the healthcare system in the United Kingdom, which has employed this methodology to analyse health consumption trends (Carr-Hill et al., 1994); other examples can be found in the work of Caballer-Tarazona, Guadalajara-Olmeda, and Vivas-Consuelo (2019), which predicted healthcare expenditure using a two-part model which relied on classic statistical approaches for the Spanish case.

From an empirical perspective, while a large share of the literature focused on investigating the core determinants of healthcare expenditure among European countries, US, and developing economies (Gerdtham & Jönsson, 2000; Martín, Puerto Lopez del Amo Gonzalez, & Dolores Cano Garcia, 2011; Newhouse, 1977), another course adopted in the literature focused on understanding the data generating process behind healthcare costs. Considering their unique distribution characteristics and the likely non-linear relationship between those costs and covariates, led to an indiscriminant modelling complexity (Basu & Manning, 2009; Jones, 2011). A large portion of the literature focused on the prediction of the conditional mean of the distribution of healthcare costs, generally through generalised linear models (Mullahy, 2009). Nonetheless, other approaches went beyond the mean in attempting to model the entire cost distribution, for example, by means of flexible parameters distributions (Jones, Lomas, & Rice, 2014; Manning, Basu, & Mullahy, 2005), finite mixture models (Deb & Burgess, 2003; Deb & Trivedi, 1997) or quantile regressions (Cook & Manning, 2009). In this regard, there is not a dominant approach, as demonstrated by Jones, Lomas, and Rice (2015), which compared the performance of 14 specifications by using the English National Health Service inpatient data.

However, it has been shown that the standard statistical and econometric tools, widely used to estimate needs and predict expenditure, are not appropriate for targeting policies (Kleinberg, Ludwig, Mullainathan, & Obermeyer, 2015). In this context, recent economic literature has debated that addressing policy challenges, like allocating public expenditure based on healthcare needs, does not necessarily require retrospective correlations or causal inference solutions. Instead, predictive inference would hold more significance, as argued by Kleinberg et al. (2015). Empirically, it has been emphasised that conventional econometric models are inadequate for resolving such predictive policy problems. These models are designed to yield unbiased coefficient estimates, rather than minimising prediction errors, as highlighted by (Einav & Levin, 2014; Kleinberg et al., 2015). On the other hand, the advancements in the realm of Machine Learning (ML) have demonstrated significant potential in tackling predictive problems (Varian, 2016). ML techniques are gaining momentum for solving problems connected to poverty targeting (Jean et al., 2016) and the effectiveness of public programmes and spending (Andini, Ciani, de Blasio, D'Ignazio, & Salvestrini, 2018).

This paper proposes a ML model to predict health care expenditure, which can be used to allocate national health funds to regions with heterogeneous needs. We use the Italian NHS as a case study since it allocates funds among citizens living in different regions as the regions have organisational responsibility for their healthcare systems (Lagravinese, Liberati, & Resce, 2019; Turati, 2013). Previous applications of ML models in similar contexts have been focused on the prediction of bankruptcy of local governments (Antulov-Fantulin, Lagravinese, & Resce, 2021), vaccine hesitancy (Carrieri, Lagravinese, & Resce, 2021), and local mortality and local inequality during the COVID-19 pandemic (Cerqua, Di Stefano, Letta, & Miccoli, 2021; Cerqua & Letta, 2022).

The analysis relies entirely on administrative data provided by Istat (2023a) through a specific tool tailored on health data named Health for All (HFA). Since the goal of our study is to predict Italian healthcare expenditure, our target variable is the total percapita healthcare expenditure at current prices. To identify crucial features in our investigation, we first conducted an extensive literature review of the determinants of healthcare expenditure (Section 2) starting from the pioneering works of Kleiman (1974) and Newhouse (1977). Moreover, the features' selection was driven by a two-fold choice: (i) the willingness to preserve longitudinal data, consequently several variables with relatively short time span, with several missing data or reported with specific yearly periods have been discarded; (ii) the choice to exclude variables which may depend directly or indirectly on healthcare expense itself such as hospital equipment, number of beds, and health personnel as well as mortality rates associated with specific diseases. The final database is composed of 23 features with a time coverage that spans the years 1994 to 2019.

Starting from the set of features at time $t-1$, we predicted healthcare expenditure in time t over the period 1994–2018 by implementing four different ML algorithms (Elastic-Net, Gradient Boosting, Random Forest, and Support Vector Regression) with an additional multivariate regression as the baseline scenario. We identified Gradient Boosting as the best predictive algorithm based on the lowest out-of-sample MAPE (3.07) and highest R^2 (0.97). Furthermore, we tested our model, trained until 2018, on 2019 data obtaining good forecasting performances (MAPE equal to 2.89 and R^2 to 0.82) stressing how, based on the selected features, our model is good in forecasting healthcare expenditure at the regional level. We replicate the analysis also by unpacking Italian health care expenditure into its public and out-of-pocket components, obtaining also good predictive performance.

Among the features with the highest predictive importance, we find time, the average number of family members, and regional area. Furthermore, some of the most important—and debated—determinants in literature resulted among the top features such as GDP per capita and life expectancy. Eventually, by relying on partial dependence plots (PDP) we graphically disentangle the relationship between each feature and the healthcare per capita expenditure without the necessity of a previous mathematical model in the functional relationship. For example, the linear time trend (year) has a positive relationship with expenditure which becomes flat after 2010 while that of the private component depicts an abrupt increase, instead. GDP per capita shows an increasing ladder pattern, hence characterised by some thresholds. Household size shows a negative association with healthcare expenditure, especially the private one, stressing how individuals spread fewer resources as family size increases but also how young family members may present life insurance for elderly members, potentially able to provide free caregiver assistance. Furthermore, large families may also be structurally younger therefore leading to lower medical costs per capita. Life expectancy at birth represents another crucial determinant of healthcare expenditure, positively associated with it, characterised by a strong increase after 80 years confirming the strong relevance of the elderly population in Italy in driving the cost of the NHS.

The remainder of the paper is structured as follows. In Section 2 we retraced the main determinants of healthcare expenditure as investigated by the literature with a focus also on specific Italian case-studies. Section 3 describes the data and the methodology adopted while Section 4 evaluates the predictive task by also identifying the core expenditure determinants for the Italian healthcare systems. Section 5 concludes the paper.

2 Determinants of healthcare expenditure in the literature

This section presents a literature review of the determinants of healthcare expenditure focusing on four main groups of factors: income related (Section 2.1); population aging (Section 2.2); technological progress (Section 2.3); and other determinants (Section 2.4). Additionally, specific Italian-based studies are reviewed in Section 2.5.

2.1 Income related

The first contributions in literature stressed out the strong positive relationship between per capita health expenditure and per capita gross domestic product (GDP) back in the seventies. A primer attempt could be found in the work of Kleiman (1974) while the most famous pioneer work is the one of Newhouse (1977) who investigated income elasticity to healthcare services for 13 OECD countries with data from 1971. Results, greater than one under different specifications, led to the identification of health expenditure as a luxury good giving rise to a vast line of literature which identifies GDP as core driver of it. According with White (2007), general economic growth, explains half of the real healthcare expenditure growth in both US and OECD countries between 1970 and 2002.

Within the so-called first generation group of studies (Gerdtham & Jönsson, 2000), which investigate international health expenditure, Newhouse's model has been enlarged through a public-choice approach by Leu (1986) who introduced other relevant exogenous variables such as population share under 15 and over 65, urbanization, and health system differences. Nonetheless, the analysis, based on 1974 data for 19 OECD countries, confirmed the prominent effect of income. Similarly, also the work of Gerdtham, Søgaard, Andersson, and Jönsson (1992) corroborated this primarily role played by income by analyzing 1987 data for 18 OECD countries. Therefore, both Leu (1986) and Gerdtham et al. (1992) confirmed the idea of healthcare expenditure as a luxury good.

The second-generation group of studies enhanced their analysis through pooled cross-sectional and time series approaches. Gerdtham (1992) analysed 22 OECD countries over the period 1972–1987 by juxtaposing inflation, fraction of public financing, and population to GDP as right-hand variables. However, among the different models performed, the income elasticity tended to be equal to or lower than one, hence confuting previous outcomes. A continuity in the outcome, instead, was obtained by Hitiris and Posnett (1992) which conducted their analysis for 20 OECD countries over the period 1960–1987 relying only on GDP per capita and share of population over 65 as explanatory variables. They both showed a positive impact of income on health expenditure with the former, again, revealing an elasticity greater than one. Nonetheless, Hansen and King (1996) show no evidence of a long-run relationship between healthcare expenditure, income, and other dependent variables for 20 OECD countries between 1960 and 1987, confuting the findings of Hitiris and Posnett (1992) which do not take into account variables' stationarity and cointegration. Similar critiques have been also advanced by Blomqvist and Carter (1997) and Okunade and Karakus (2001).

Alongside econometrics improvements, better data availability allowed for a widening in the number of tested determinants. For example, Gerdtham, Jönsson, MacFarlan, and Oxley (1998) analyzed 22 OECD countries over the period 1970–1991 by considering in their model, alongside GDP and population variables, also elements referred to labour market, population's addictions, and health system infrastructures. Among non institutional variables, tobacco consumption, and GDP were the most relevant, the latter with an income elasticity lower than unity.

Within the literature review conducted by Martín et al. (2011), which analyses 20 articles published between 1998 and 2007, GDP per capita still represents a widely used explanatory variable. Nonetheless, only four of them identified it as a primary determinant, two of them jointly with population aging. Moreover, elasticity generally resulted in being lower than unity. Therefore, by considering this shrink in the number of contributions which identify GDP as a primary determinant of healthcare expenditure and those whose elasticity is greater than one, healthcare may be better characterised as a necessity—thus not a luxury—good. Among the works which considered income as the main determinant of healthcare expenditure, the one of Barros (1998) analyses 24 OECD countries between 1960 and 1990 with different dependent variables with an elasticity for GDP per capita ranging between 0.62 and 0.92. Afterwards Roberts (2000), despite the econometric critiques directed to the work of Hitiris and Posnett (1992), came to similar conclusions in his work with an income elasticity of demand greater than one, between 1.21 and 1.25, depending on the estimated model. Roberts' model analyses 10 countries of the former European Community (EC) over the period 1980–1991. Also Clemente, Marcuello, Montañés, and Pueyo (2004) concludes for an income elasticity greater than one with their work, focused on 22 OECD countries for the period 1960–1977, which demonstrates a long-run relationship between total healthcare expenditure and GDP when structural breaks are considered in the estimation. Moreover, also Giannoni and Hitiris (2002), by examining the regional Italian case, identified GDP as a core determinant although with an income elasticity lower than one.

Still focusing on OECD economies, with a panel of 20 countries analszed from 1971 up to 2004, Baltagi and Moscone (2010) conclude for an income elasticity lower than one, as did Moscone and Tosetti (2010) as well as Murthy and Okunade (2016) for the specific US case. The former by studying 49 states between 1980 and 2004 while the latter by conducting a country-level analysis over the period 1960–2012. Moreover, Baltagi, Lagravinese, Moscone, and Tosetti (2017) conducted a global analysis of the role played by income in determining healthcare expenditure by considering 167 countries from 1995 up to 2012. They demonstrate how, at global level, healthcare represents a necessary good since its income elasticity is lower than one, ranging from 0.84 to 0.87. Results are even lower for high- and upper-middle-income economies. Nonetheless, for poorer countries income elasticities exceeded unity, hence ascribing it as a luxury good for lower-middle and low-income economies.

2.2 Population aging and closeness to death

Generally, the healthcare demand, distinct by age group, follows a J-shape distribution. There is first a local maximum relative to the early years of life, followed by a progressive decrease during childhood. Then the curve starts to rise again around fifty years up to the maximum level in correspondence of 75-80 years and so on (Altavilla, Mazza, & Monaco, 2016; Protonotari et al., 2007; Rebba, 2005).

Besides income, population aging has always been identified as a core driver of healthcare expenditure starting from the first works of Newhouse (1977) and Leu (1986). Gerdtham et al. (1992) state that age distribution impacts the demand for healthcare services with an older population characterized by a relatively higher propensity to spend for health services. Similarly, Murthy and Ukpolo (1994), by investigating the US healthcare sector for the period 1960–1987, showed how the population's age structure as well as the number of practicing physicians emerge as the two core determinants while GDP is ascribed as a normal good. Nonetheless, some macro-level studies conducted on OECD countries show no evidence of a linkage between aging and health expenditure (Barros, 1998; Getzen, 1992). Furthermore, while studying the determinants of healthcare expenditure growth between US and OECD countries, White (2007) showed how the role of aging was limited (0.3% in the US and 0.5% in the OECD) stressing how the increase in per capita spending interested each age group (Meara, White, & Cutler, 2004).

Within the literature review of Martín et al. (2011), six studies supported the aging hypothesis. Di Matteo and Di Matteo (1998) investigate the determinants of healthcare expenditure for some Canadian provinces from 1965 to 1991, finding that population aging represented the major driver while income, with an elasticity of demand of 0.77, was classified as a normal good. A similar conclusion emerges even from the work of Karatzas (2000), based on the US case with data from 1962 to 1989 where the percentage of population aged over 65 was shown to have a strong influence in explaining healthcare expenditure. By working on both Canadian (1975–2000) and US (1980–1998) cases, Di Matteo (2005a) shows how in both cases, the eldery population was mainly responsible for the increase in healthcare expenditure. Moreover, even the works of Roberts (2000) and Giannoni and Hitiris (2002) stress how population aging represent a core driver of healthcare expenditure jointly with income.

In analyzing historical data from 1960 to 2012, Murthy and Okunade (2016) demonstrate how GDP, aging population, and healthcare technology significantly drive expenditures for the US. The variable which embodies population aging showed the higher elasticity, equal to 1.74, while that of income elasticity was lower than unity (0.92).

Other studies aimed at investigating healthcare expenditure's determinants do not assume that only aged people may affect it but also include the proportion of young people in their models (Ang, 2010; Baltagi & Moscone, 2010; Crivelli, Filippini, & Mosca, 2006; Wang, 2009). Moreover, while there is evidence that the larger share of personal health expenditure is concentrated in the final years of life EC (2009), differences in the expenditure between young and older individuals are not due to differences in calendar age but rather in terms of time to death. Therefore, the demand for health services depends on health status and proximity to death and not to age, per se. In fact, with the work of Zweifel, Felder, and Meiers (1999) proximity to death also started to be included among possible determinant of healthcare expenditure leading the way for the so-called "red herring" hypothesis: the higher correlation between aging and healthcare expenditure is due to the fact that the latter rises steeply in the last months before death and since elderly persons are in their last years of life, this explain the higher propensity for older age groups to spend more on medical care. Zweifel et al. (1999) estimated a model for individuals under two healthcare insurance companies in Switzerland showing how the higher cost is concentrated in the four months prior to death, independently of the individual's age. These conclusions were also confirmed by Shang and Goldman (2008) using US data that enriched the analysis by also including predicted life expectancy.

Results similar to those of Zweifel et al. (1999) were obtained also by Felder, Meier, and Schmitt (2000) still focusing on Switzerland over the period 1986–1992. Data from Switzerland were also used by Werblow, Felder, and Zweifel (2007) to evaluate the role of proximity to death while Breyer and Felder (2006), instead, used it to calculate the demographic impact on German healthcare expenditure for 2050. Results obtained by Zweifel et al. (1999), despite being methodologically criticised by Seshamani and Gray (2004a), were confirmed in their study from the English county of Oxfordshire with data between the years 1970–1999, also replied based on different age cohorts by the same authors (Seshamani & Gray, 2004b). Wong, van Baal, Boshuizen, and Polder (2011) enlarged the analysis by considering the role of age and closeness to death for 94 different conditions in the Netherlands confirming the higher impact of the latter over healthcare expenditure. In the same direction the work of Blanco-Moreno, Urbanos-Garrido, and Thuissard-Vasallo (2013) could also be positioned. This work projected Spanish healtcare expenditure figures up to 2060 healthcare expenditure in Spain stressing intensity of use rather than aging as the main diver.

Eventually, in these studies, when controlling for proximity to death, there emerges a reduction—or lack of significance—for population aging variables. Nonetheless, even in this domain some different outcomes stand out. For the specific Italian case, for example, Atella and Conti (2014) examined primary care expenditures for a sample of 750,000 adult individuals between 2006 and 2009 concluding that age is a much better predictor of healthcare expenditure rather than proximity do death. Similar conclusions—

in opposition to the red herring hypothesis—were also found by De Meijer, Koopmanschap, d'Uva, and Van Doorslaer (2011) for the Netherlands and by Karlsson and Klohn (2014) for Sweden.

Another confirmation of the role played by closeness to death we may find the work of Hyun, Kang, and Lee (2016) which analyses the role of population aging in Korea showing how, when closeness to death is included in their model, the expenditure in healthcare decreases as a function of age while it increases as a function of time to death during individuals' terminal years.

2.3 Technological progress

Newhouse (1992) first stressed the relevance of technological progress as a determinant of medical expenditure increase. According to the author, after identifying the three major factors known to affect the demand for medical services in US from 1960 to 1987 (aging, the spread of insurance, and the growth of income), with their contributions to the increase in healthcare expenditure growth, the role played by technological change could be identified in a residual manner. In a similar residual way also White (2007) identified the technological change as a possible driver—alongside other factors, including expansions in insurance coverage and changes in healthcare financing and delivery systems—of the higher growth of healthcare spending in US compared to other OECD countries, especially after the mid-eighties, once the two major components of population aging and general economic growth were isolated.

In order to assess changes in medical care technology, several proxies have been used in literature. For example, Baker and Wheeler (1998) and Weil (1995) used surgical procedures and the number of specific medical equipments while Dreger and Reimers (2005) relied on life expectancy and infant mortality, instead. Among time-series model, authors have included a time index (Gerdtham et al., 1992), time-specific intercepts (Di Matteo, 2005b) or time-trend (Roberts, 2000).

In the work of Okunade and Murthy (2002), reference was made to the US case over the period 1969–1999, in which technological change represented the main driver of healthcare expenditure, expressed in terms of either total research and development (R&D) expenditure or healthcare R&D expenditure. Its impact remained significant and positive even in the analysis of Murthy and Okunade (2016) although its elasticity was lower compared to that of GDP and aging. Moreover, still focusing on the US case, the work of Koenig et al. (2003) relied on different variables which referred to specific services and equipment to assess the role of medical technological on physician services expenditure. It transpired that the most important variable appeared as the percent of surgeries performed on an outpatient basis.

Eventually, for the Australian case, You and Okunade (2017) specifically analysed the technological role played in determining healthcare national expense between 1971 and 2011 through several proxies: economy-wide research and development expenditures, hospital research expenditures, mortality rate, and two technology indexes based on medical devices. Results suggest that healthcare is a technical necessity—as well as a normal good—showing an elasticity between 0.3 and 0.35.

2.4 Other determinants

Alongside income, aging, and technological change starting from the pioneer works in this literature, other variables have aslo been included in the model as possible determinants of national, regional or individual healthcare expenditure.

The literature provides a plethora of other variables included in the analysis as possible determinants for healthcare expenditure. For example Mosca (2007) identified the decentralisation of healthcare systems as a core determinant in its cross-country analysis of 20 OECD countries between 1999 and 2000. Results from the work demonstrate how healthcare services are a normal good and how an increase of 10% in the percentage of over-80s lead to an increase in health expenditure equal to 1.4%, thus supporting the aging hypothesis. Also López-Casasnovas and Saez (2007) investigated the role of decentralisation on 110 regions from eight OECD countries, including Italy, in 1997. Their results show an income elasticity far lower than one and a positive role of population over 65 in rising healthcare expenditure.

Even the intrinsic characteristics of each NHS may present another important discriminant determinant (Wagstaff, 2009; Wagstaff & Moreno-Serra, 2009). In fact, differences in health systems also have repercussions on general individual health status as demonstrated by Fonseca, Langot, Michaud, and Sopraseuth (2023) through a general equilibrium heterogeneous agent model which compares European countries with the US. For the latter, prices are estimated to be 33% higher compared to the former and they explained more than 60% in healthcare expenditure and half of individual health status.

Several authors included variables related to the labour market participation in their works. For example, Crivelli et al. (2006) and Mosca (2007) use the unemployment rate, Dormont, Grignon, and Huber (2006) use social and occupational groups for each individual. Still at individual level, Zweifel et al.

(1999), Felder et al. (2000), and Seshamani and Gray (2004b) found the inverse of the Mills ration; thus, the propensity for each individual to participate in the labour market. Unemployment has a negative impact at the OECD levels but not for Swiss cantons. Conversely, the Mills ratio, investigated for older population, has a positive impact on the expenditure.

Moreover, the work of Koenig et al. (2003), which aimed at identifying the factors that contributed to the increase in the US healthcare expenditure between 1990 and 2000, relied on a set of 41 dependent variables grouped into nine categories: demographic and general economic conditions; health status; provider payment; healthcare insurance; supply of physicians and specialists; market structure of providers; running costs; healthcare regulation; treatment patterns and technology. Mostly all implemented variables were found to have a statistically significant impact, either positive or negative, on healthcare expenditure. The percentage of people aged 65 years and more emerged as the variable with the highest positive impact followed by per capita income (with elasticity lower than 1), but also the percentage of beds. On the contrary, the percentage of black and Hispanic population were the variables with the highest negative impact on healthcare expenditure.

Nonetheless, several other interesting determinants can also be found in the literature, for example, the number of cities with more than 1,000 inhabitants and per capita administrative expenses (Karatzas, 2000) or democracy indices Crivelli et al. (2006).

2.5 Specific Italian-based studies

Within the Italian framework, Giannoni and Hitiris (2002) investigated the determinants of healthcare expenditure at regional level over the period 1980–1995 stressing how decentralisation had not increased expenditure. Real GDP represented the main determinant, although with an elasticity equal to 0.33 and thus in line with the idea of healthcare as a normal good for the Italian case. Population aging emerged as the second most important variable with an elasticity equal to 0.16. The authors also included in their model some variables aimed at capturing the health offer such as the number of beds per hospital and the number of healthcare and non-healthcare personnel per hospital. The former variable was found to be negatively associated with per capita healthcare expenditure which is an indicator of economies of scale at regional level; the latter, instead, showed a positive sign supporting the fact that the intensive use of labour in the healthcare sectors increases expenditure.

Lopreite and Mauro (2017) analysed the impact of demographic changes on healthcare expenditure in Italy over the period 1990-2013. However, due to the relative short time-span, they preferred to rely on a Bayesan rather than a frequentist approach. Their response-pulse function shows how per capita healthcare expenditure is more influenced by aging rather than GDP per capita and life expectancy in the short-run. This result is also corroborated over a long-run forecast horizon (from 1 to 10 years). Moreover, changes in health spending have a minimal impact on the other three determinants: aging index (the ratio between people aged 65 and over and the youth population between 0 and 14 years), life expectancy, and GDP per capita. The variance-decomposition analysis further corroborates those findings showing how GPD per capita, life expectancy, and especially the aging index are important determinants of healthcare expenditure in Italy.

A reverse perspective is analysed by Golinelli et al. (2017) where health expenditures is studied in terms of their effects on all-cause mortality within the regional Italian healthcare systems. Over the period 1999–2003, the authors show how services directly provided by the Italian SSN has a positive impact in reducing short-term mortality while other services supplied via private healthcare providers show no significant association with it.

A more disaggregated analysis, based on individual data collected routinely by General Practitioners between 2006 and 2009 has been conducted by Atella and Conti (2014). They aimed at testing whether time to death was a better predictor for primary care costs—which represent around 33% of total healthcare costs and are highly related to the end-of-life health costs—, instead of aging. Results show how primary care cost are mainly driven by age in opposition to Raitano et al. (2007). In fact, while in the former survivor costs increase more than decedent costs as far as people age, for the latter, decedent costs decrease by age more than survivor costs increase by age.

Eventually, still using disaggregated data, Sarti, Terraneo, and Bordogna (2017) studied the effects of the 2008 financial crisis on health behaviours of Italian households using a family survey covering the period between 1997 and 2013. Results showed that, after the crisis, the propensity to spend for poorer families is decreased compared to expenditure on families which were not classified as poor.

3 Data and methodology

3.1 Data

The analysis relies entirely on administrative data provided by the Italian Statistic Institute Istat (2023a) through a specific tool tailored on health data named Health for All (HFA, latest version December 2022). This source represents a database of indicators on the Italian healthcare system and on population health structured in such a way that it can be queried by the HFA software provided by the World Health Organization (WHO) adapted to national needs. Currently, the database contains 4,000 indicators disaggregated at the regional level and this is the stratification level at which we conducted our analysis.

The literature review proposed in Section 2 represented for us an essential instrument to identify the core variables implemented as possible determinants of regional or national healthcare systems' expenditure in order to make rational choices in selecting the features to include in our predictive model.

Since the goal of our study is to predict Italian healthcare expenditure, our target variable is the total per capita healthcare expenditure at current prices. We investigated the variable also by dividing it into its two main components: public and private (households) per capita expenditure. The public component represents the largest share of health expenditure although it has decreased over time, from 83.2% in 1990 to 74.8% in 2019[1]. Moreover, the share distribution is heterogeneous among regions with those of the North with a larger share of family-private expenditure as opposed to Southern regions.[2]

The features' choice was driven by a two-fold criterion: (i) the willingness to preserve longitudinal data, rather than several variables with relative short time span, with several missing data or reported with specific yearly periods having been discarded; (ii) the choice to exclude variables which may depend directly or indirectly on healthcare expense itself such as hospital equipment, number of beds, and health personnel as well as mortality rates associated with specific diseases. All selected features with their description and relative descriptive statistics are reported in Table 1. The final database is composed of 23 features with a time coverage which spans from 1994 to 2019.

[1] There has been a new increase in 2020 (77.1%) during the first year of the pandemic crisis.

[2] Southern Italy, also known as *Meridione* or *Mezzogiorno* comprises the administrative regions that correspond to Abruzzo, Apulia, Basilicata, Calabria, Campania, Molise, Sicily, and Sardinia.

Table 1: *Selected features for prediction (1995–2019)*

Code	Description	Mean	Median	Sd	Min	Max
target_var_tot	Total healthcare expenditure	2029.0	2164.5	551.9	876.0	3228.0
target_var_pub	Public healthcare expenditure	1552.3	1715.0	419.7	694.0	2309.0
target_var_fam	Private healthcare expenditure	476.7	461.5	165.2	182.0	1054.0
area_5	Categorical variable divided into North-West, North-East, Center, and South					
area	Region surface	15103.2	14446.1	7250.8	3260.9	25832.5
age_ind	Aging index	154.6	156.2	40.0	59.8	259.8
dens	Population density	178.8	157.8	107.6	35.7	444.0
fam_com_avg	Average number of family members	2.5	2.5	0.2	1.9	3.1
lav_tas_att_15ov	Activity rate for individuals aged 15 and over	47.7	47.9	5.4	33.1	57.9
lav_tas_dis_15ov	Unemployment rate for individuals aged 15 and over	10.5	8.8	5.9	2.6	28.0
lav_tas_occ_15ov	Employment rate for individuals aged 15 and over	43.4	45.1	6.7	30.4	55.6
lf_fum_big_p	Heavy smokers aged 15 and over (%)	7.3	7.0	3.1	0.6	16.3
life_exp	Life expectancy (average between male and female)	80.9	81.4	1.8	76.3	84.2
ls_fum_65ov_p	Smokers aged 65 and over (%)	9.7	9.6	2.1	4.2	16.8
ls_fum_p	Smokers aged 15 and over (%)	22.0	21.8	2.8	15.5	30.7
ls_obe_65ov_p	Obese people aged 65 and over (%)	13.7	13.5	3.0	6.3	22.8
ls_obe_p	Obese people aged 18 and over (%)	9.8	9.8	1.8	4.8	15.0
ls_sov_65ov_p	Overweight people aged 65 and over (%)	43.7	43.9	3.9	30.3	55.0
ls_sov_p	Overweight people aged 18 and over (%)	34.9	34.8	3.0	28.0	41.8
pil_cap	Per capita gross domestic product at current prices	23856.9	23273.6	7480.7	9071.9	43967.2
pop_res	Total resident population	2937798.2	1882683.5	2358371.0	116522.0	10053264.0
pop_res_0.14_p	Resident population between 0 and 14 years (%)	13.7	13.3	1.9	10.1	20.2
pop_res_65ov_p	Resident population aged 65 and over (%)	20.6	20.8	3.2	12.1	28.7
pop_res_f_p	Female residents (%)	51.5	51.4	0.4	50.6	52.8
pop_res_m_p	Male residents (%)	48.5	48.6	0.4	47.2	49.4
str_res_p	Foreign residents (%)	4.6	3.3	3.5	0.3	12.1

All variables have been retrieved from Istat (2023a).

Subsequently, we conducted a second analysis which enlarges the number of selected features by relying not only on HFA data but also on other sources by including variables related to the population's education level, migratory balance, and institutional quality (Istat, 2023b; Nifo & Vecchione, 2021). However, this extension in features' number occurs at the cost of a reduction in the time coverage resulting in a new data set which spans from 2005 to 2019 for a total of 42 features. All variables from this second data set with their description and relative descriptive statistics are reported on Table 2.

Table 2: *Selected features for prediction (2005–2019)*

Code	Description	Mean	Median	Sd	Min	Max
target_var_tot	Total healthcare expenditure	2379.7	2381.0	283.3	1681.0	3228.0
target_var_pub	Public healthcare expenditure	1831.4	1828.0	168.1	1356.0	2309.0
target_var_fam	Private healthcare expenditure	548.3	530.0	153.1	272.0	1054.0
area_5	Categorical variable divided into North-West, North-East, Center, and South					
area	Region surface	15103.2	14446.1	7261.5	3260.9	25832.5
age_ind	Aging index	164.5	166.4	33.6	82.8	259.8
dens	Population density	181.4	162.0	109.7	37.5	444.0
fam_com_avg	Average number of family members	2.4	2.4	0.2	1.9	3.0
inc_fam_cap	Disposable income of consumer households per inhabitant	17498.6	18516.8	3463.5	11363.1	24059.6
iqi_cor	Corruption (IQI)	0.793	0.840	0.183	0.214	0.990
iqi_gov	Government effectiveness (IQI)	0.368	0.391	0.163	0.000	0.690
iqi_reg	Regulatory quality (IQI)	0.541	0.586	0.197	0.087	0.966
iqi_rol	Rule of law (IQI)	0.580	0.619	0.229	0.063	1.000
iqi_voi	Voice and accountability (IQI)	0.587	0.621	0.203	0.118	0.979
lav_tas_att_15ov	Activity rate for individuals aged 15 and over	48.8	50.5	5.3	37.9	57.9
lav_tas_dis_15ov	Unemployment rate for individuals aged 15 and over	9.9	8.9	5.1	2.8	23.4
lav_tas_occ_15ov	Employment rate for individuals aged 15 and over	44.1	46.3	6.8	30.4	55.6
lf_fum_big_p	Heavy smokers aged 15 and over (%)	5.7	5.6	2.2	0.6	11.5
life_exp	Life expectancy (average between male and female)	82.1	82.1	0.9	79.3	84.2
ls_fum_65ov_p	Smokers aged 65 and over (%)	9.2	9.1	1.9	4.2	15.4
ls_fum_p	Smokers aged 15 and over (%)	20.8	20.7	2.3	15.5	27.4
ls_no_spo_p	People who do not practice sport or physical activity (%)	38.9	38.0	11.1	12.8	60.2
ls_obe_65ov_p	Obese people aged 65 and over (%)	14.6	14.5	2.8	8.4	22.8
ls_obe_p	Obese people aged 18 and over (%)	10.5	10.5	1.5	6.6	15.0
ls_sov_65ov_p	Overweight people aged 65 and over (%)	44.5	44.6	4.0	30.3	55.0
ls_sov_p	Overweight people aged 18 and over (%)	35.6	35.2	2.8	29.4	41.8
pil_cap	Per capita gross domestic product at current prices	26443.1	26559.3	7066.1	15233.1	43967.2
pop_res	Total resident population	2987354.6	1882683.5	2418947.5	122246.0	10053264.0
pop_res_0_14_p	Resident population between 0 and 14 years (%)	13.5	13.3	1.3	10.9	17.9
pop_res_65ov_p	Resident population aged 65 and over (%)	21.8	21.9	2.7	14.8	28.7
pop_res_f_p	Female residents (%)	51.5	51.4	0.4	50.7	52.8
pop_res_m_p	Male residents (%)	48.5	48.6	0.4	47.2	49.3
pov_inc_pov_fam	Incidence of poverty at household level	11.8	9.0	7.5	2.3	35.3
sal_mig_est	External migration balance (per thousand inhabitants)	3.1	2.4	2.8	-3.0	13.6
sal_mig_int	Internal migration balance (per thousand inhabitants)	-0.2	0.3	2.4	-7.4	4.8
str_res_p	Foreign residents (%)	6.2	6.2	3.3	0.9	12.1
stu_pop_lau_p	University graduate population (% tot)	11.6	11.4	2.5	6.4	20.5
stu_pop_lic_ele_p	Population with at least a primary school certificate (%)	22.4	22.2	4.9	11.7	33.3
stu_pop_lic_med_inf_p	Population with at least a middle school certificate (%)	31.6	31.2	3.1	25.2	40.1
stu_pop_lic_med_sup_p	Population with at least a high school certificate (%)	34.4	34.5	3.5	25.7	41.6
tas_due_mal_cro	Rate of people with at least two chronic diseases	585.8	584.9	73.2	375.1	762.2
tas_fec	Fertility rate (total)	1316.4	1318.0	123.0	979.0	1637.0
tas_mor	Mortality rate	103.4	102.5	13.2	78.2	141.8
tas_nat	Birth rate	8.3	8.4	1.2	5.4	11.3
tas_nuz	Marriage rate	3.5	3.5	0.6	2.6	5.5

The variables iqi_cor, iqi_gov, iqi_reg, iqi_rol, and iqi_voi have been retrieved from the Institutional Quality Index dataset (Nifo & Vecchione, 2021). The variables sal_mig_est, sal_mig_int, and inc_fam_cap have been retrieved from Istat (2023b). All remaining variables have been retrieved from retrieved from Istat (2023a).

3.2 The prediction task

The prediction task is formulated as follows, for each region i at the year t, based on the set of lagged features $Features_{i,t-1}$, find the function $f(.)$ (machine learning model) that predicts the regional health expenditure percapita $Exp_{i,t}$:

$$\left\{Features_{i,t-1}\right\} \xrightarrow{f(.)} Exp_{i,t}. \tag{1}$$

Following the predominant approach in the literature using ML models for health decision-making (Brnabic & Hess, 2021), we randomly divide the database as 80 percent for training and 20 percent for the out-of-sample testing set (test). As the healthcare data are only available at the regional level, and with fewer observations, different splits of the data may result in different results, we have performed 100 different random splittings and we have averaged model performances across the repetitions. The hyper-parameter optimisation is only done on the training set using a repeated (10 times) five-fold cross-validation. To carry out our analysis, we use four different ML predicting algorithms plus a more "classical" multivariate regression model:

- Elastic Net (EN): a regression statistical method that performs features selection and regularisation with a mix of L1 (LASSO-type) and L2 (ridge-type) penalization to reduce over-fitting and increase prediction accuracy and interpretability (Tibshirani, 1996; Zou & Hastie, 2005);
- Random Forest (RF): a family of randomised tree-based classifier decision trees which uses different random subsets of the features at each split in the tree (Breiman, 2001);
- Gradient Boosting Machines (GBM): an ensemble method which works in an iterative way where at each stage, a new learner tries to correct the pseudo-residual of its predecessors (Friedman, 2001);
- Support Vector Regression (SVR): an algorithm which aims to find a hyperplane in a high-dimensional feature space that best fits the data points while minimising the error between the predicted and actual target values (Drucker, Burges, Kaufman, Smola, & Vapnik, 1996);
- Multivariate Regression (MR): an Ordinary Least Square (OLS) regression (Wooldridge, 2015).

The analysis has been conducted firstly by considering the primary data set (wider time coverage, less variables) by excluding the year 2019 from the sample whose data has been used instead to further validate the predictability of our model. We conducted this analysis for the three different target variables: total, public, and private healthcare expenditures. Afterward, by preserving the same analysis structure we analyzed also the second data set (narrow time coverage, more variables). All performance measures are calculated on the test set (20% of the data, the same for each algorithm). Since we predict a continuous outcome, the performances of the models are primarily compared by the Mean Absolute Percentage Error (MAPE) which is independent of the unit measure of the outcome and, for this reason, is the most widely used goodness-of-fit measure (Moreno, Pol, Abad, & Blasco, 2013).

4 Results

In this section, we present the results of the model predicting regional health care expenditure. The focus will be on two main aspects: the predictability of our dependent variables (Section 4.1) and the features' importance of independent variables used for predictions (Section 4.2).

4.1 Model's performance

Once trained using the 1994–2018 data, we computed several out-of-sample performance statistics, based on tested data, for each function (ML algorithm) in order to evaluate their predictive power. Performance output are reported on Table 3 referring to total per capita healthcare expenditure as target variable. Because we initially computed 100 different splits of the initial database, we reported, for each outcome and each model, the mean value, as well as other information related to their variability. The GBM algorithm with a MAPE value equal to 3.07 proved to be the better one. The value of MAPE is below 10, which according to Lewis (1982) is the threshold for highly accurate forecasting. The GBM model is also characterized by the highest R^2 score with a value equal to 0.97 while the RF ranks as the second best model. When considering public healthcare expenditure as the target variable, GBM still emerged as the best performing model with a MAPE value equal to 3.08 and a R^2 to 0.97 (Table A1 in Appendix). Conversely, when considering private healthcare expenditure, the preferred model was RF with a MAPE equal to 5.13 and R^2 to 0.93 (Table A2 in Appendix)[3].

[3] More detailed performances' output can be found in the Appendix in Tables A3, A4, and A5 for total, public, and private healthcare expenditure, respectively.

Table 3: *Models' performances, total healthcare expenditure (1994–2018)*

	Mean	Median	Min	Max	sd
Gradient Boosting					
RMSE	78.7966	78.0233	54.7692	103.0542	9.1457
MAE	59.5427	59.7027	44.8861	74.3375	5.9083
R-squared	0.9787	0.9793	0.9637	0.9896	0.0049
MAPE	3.0714	3.0496	2.4227	3.8341	0.2859
Elastic Net					
RMSE	124.4908	124.989	98.4416	154.2244	11.3467
MAE	92.5567	92.5716	71.8705	112.0285	8.1383
R-squared	0.9471	0.9481	0.9254	0.9643	0.009
MAPE	4.7981	4.7758	4.018	5.7668	0.4097
Random Forest					
RMSE	87.0089	87.0555	66.5716	118.5396	11.0057
MAE	63.9104	64.1321	50.0988	79.1371	6.2705
R-squared	0.974	0.9744	0.9572	0.9849	0.0063
MAPE	3.3382	3.3654	2.67	3.9923	0.2948
Regression					
RMSE	125.759	125.8533	100.8659	154.57	11.2939
MAE	93.7925	94.6845	76.4449	113.6491	8.4725
R-squared	0.9461	0.9472	0.9251	0.9645	0.009
MAPE	4.858	4.8059	4.0485	6.0326	0.4503
Support Vector Regression					
RMSE	96.0728	95.4064	74.6998	119.3491	8.9812
MAE	72.5331	71.9648	54.4561	91.2275	6.6547
R-squared	0.9685	0.9688	0.9545	0.9798	0.0057
MAPE	3.8068	3.7936	2.9572	4.8582	0.3583

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

In the second step of our analysis, we used trained models on data over the period 1994–2018 to test their predictive performance over 2019 data, hence to predict information over a specific year the machine has never encountered. Results from this second evaluation are reported in Table 4 for total healthcare expenditure. Even in this case, GBM was identified as best algorithm based on the MAPE value (2.89). Again as the value of MAPE is below 10 we can consider the out-of-sample results highly accurate forecasting (Lewis, 1982). Nonetheless, considering the relative small test sample, the R^2 slightly dropped down to 0.82; nonetheless, still showing how the proportion of variance in the target variable explained by the model is high. When considering public expenditure the best performing model with 2019 data was RF with a MAPE equal to 3.43. However, all algorithms tend to show a performance downgrade in terms of R^2 which resulted equal to 0.43 for RF (Table A6 in Appendix). When testing the model for private expenditure over 2019 data the best model remained the RF with a MAPE equal to 4.87 and an R^2 equal to 0.88, substantially in line with what was observed for total healthcare expenditure per capita (Table A7 in Appendix).

Table 4: *Models' performances, total healthcare expenditure (1994–2018: 2019)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	91.4334	91.6125	70.7628	117.6461	9.471
MAE	76.227	76.5399	54.1729	95.0928	8.593
R-squared	0.8274	0.8286	0.7173	0.8977	0.0358
MAPE	2.8908	2.8976	2.059	3.5892	0.3205
Elastic Net					
RMSE	150.0621	150.3403	139.8028	163.1436	4.9014
MAE	123.0379	122.6312	111.4755	136.2586	4.5741
R-squared	0.5395	0.5383	0.4563	0.6008	0.03
MAPE	4.7015	4.6768	4.2282	5.2433	0.1879
Random Forest					
RMSE	95.4757	94.8162	85.0461	119.829	6.202
MAE	78.7251	77.8901	69.8621	96.1539	5.2826
R-squared	0.813	0.8164	0.7067	0.8523	0.0255
MAPE	2.9499	2.9242	2.6163	3.536	0.1909
Regression					
RMSE	156.3306	156.4624	144.0385	170.9412	5.7005
MAE	127.6733	126.9828	113.3113	144.9846	5.9958
R-squared	0.5001	0.5	0.4031	0.5762	0.0365
MAPE	4.8743	4.8545	4.3141	5.5648	0.2376
Support Vector Regression					
RMSE	129.3177	127.6828	109.8442	169.0696	11.3342
MAE	106.6276	104.8977	89.4801	138.7281	9.9033
R-squared	0.6558	0.667	0.4161	0.7535	0.0623
MAPE	4.0071	3.9572	3.4236	5.1258	0.3518

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

After evaluating the performance over the primary data set, we repeated the same analysis also for the second data set (2005–2018) for all three target variables. For total healthcare expenditure, RF emerged as best model (MAPE equal to 2.62 and R^2 to 0.89) while performing on 2019 the Elastic Net algorithm outperformed the others (MAPE equal to 2.69 and R^2 to 0.83) (Tables 5 and 6). For public healthcare expenditure, instead, GBM was the preferred model (MAPE equal to 2.58 and R^2 to 0.82) while with 2019 data we preferred Elastic Net (MAPE equal to 3.31 and R^2 0.34). Notice that even in this case there emerges a lower proportion of variance of the target variable explained by the model (Tables A8 and A10 in Appendix). Eventually, when considering out-of-pocket healthcare expenditure GBM resulted as the preferred model both on the test sample (MAPE equal to 4.38 and R^2 to 0.96) and on 2019 data (MAPE equal to 4.42 and R^2 to 0.82) (Tables A9 and A11 in Appendix). Also in these cases, the values of MAPE are all below 10 which indicates highly accurate forecasting (Lewis, 1982).

Table 5: *Models' performances, total healthcare expenditure (2005–2018)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	86.5748	83.165	64.4941	128.1902	13.0464
MAE	64.1009	62.7564	48.2888	85.9879	7.9333
R-squared	0.8945	0.901	0.7939	0.9421	0.0292
MAPE	2.6688	2.614	1.9506	3.7082	0.3278
Elastic Net					
RMSE	108.8337	109.0883	81.7951	151.2694	13.2546
MAE	83.7404	83.1287	63.8595	120.1623	9.9299
R-squared	0.8336	0.8374	0.7221	0.9083	0.0376
MAPE	3.5102	3.4803	2.675	5.1272	0.4328
Random Forest					
RMSE	87.6928	83.5645	64.1278	129.8281	13.9751
MAE	63.0297	60.8982	48.2409	95.3599	8.5674
R-squared	0.8923	0.8966	0.8002	0.941	0.0282
MAPE	2.6289	2.5554	2.0492	4.0733	0.3563
Regression					
RMSE	106.7749	106.4242	76.1639	155.0853	14.2482
MAE	82.7612	82.6521	60.304	114.8571	10.3254
R-squared	0.8394	0.8388	0.7248	0.925	0.0395
MAPE	3.4813	3.479	2.521	4.9528	0.4565
Support Vector Regression					
RMSE	95.8131	93.4934	67.5015	147.2978	15.9512
MAE	66.9473	65.4701	52.0674	105.3669	9.2146
R-squared	0.8723	0.8762	0.7856	0.9208	0.0303
MAPE	2.7656	2.7219	2.135	4.5305	0.3841

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

Table 6: *Models' performances, total healthcare expenditure (2005–2018: 2019)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	96.1441	95.8584	72.0043	117.1209	9.0097
MAE	78.6433	78.5507	56.9251	95.2911	8.0837
R-squared	0.8095	0.8123	0.7198	0.8941	0.035
MAPE	2.9916	3.0016	2.1679	3.6446	0.3045
Elastic Net					
RMSE	90.238	88.9419	76.2422	107.1552	6.5178
MAE	72.0241	72.3014	59.4292	85.6122	5.256
R-squared	0.8328	0.8384	0.7655	0.8813	0.0244
MAPE	2.6994	2.6998	2.2379	3.1992	0.1933
Random Forest					
RMSE	102.6907	101.733	92.9286	126.2474	5.2903
MAE	83.2135	82.1311	71.4257	101.2787	5.3167
R-squared	0.784	0.7886	0.6744	0.8236	0.0228
MAPE	3.106	3.075	2.6699	3.7828	0.1956
Regression					
RMSE	112.3121	111.7608	88.1477	136.8525	10.8379
MAE	82.1757	81.7713	62.35	101.3859	7.4798
R-squared	0.74	0.7449	0.6174	0.8413	0.0502
MAPE	3.0151	2.9951	2.2964	3.7162	0.274
Support Vector Regression					
RMSE	118.9434	119.066	104.0825	137.6207	7.0614
MAE	96.2955	96.1351	82.2521	108.4876	6.1233
R-squared	0.71	0.7104	0.6131	0.7787	0.0346
MAPE	3.5892	3.5944	3.0626	4.0136	0.2208

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

Generally, our models are good at predicting the total national healthcare expenditure where GBM and RF alternates as best algorithms generally with a minimal margin of difference. Even dividing total expenditure in its two macro components, public and private, still we get optimal predictive models. Furthermore, when testing our models on 2019 data (not used for training) we still get good performances, especially for total and private expenditure. Moreover, also when including more features in the analysis with a narrow time coverage, models remained predictive despite having lower R^2 values.

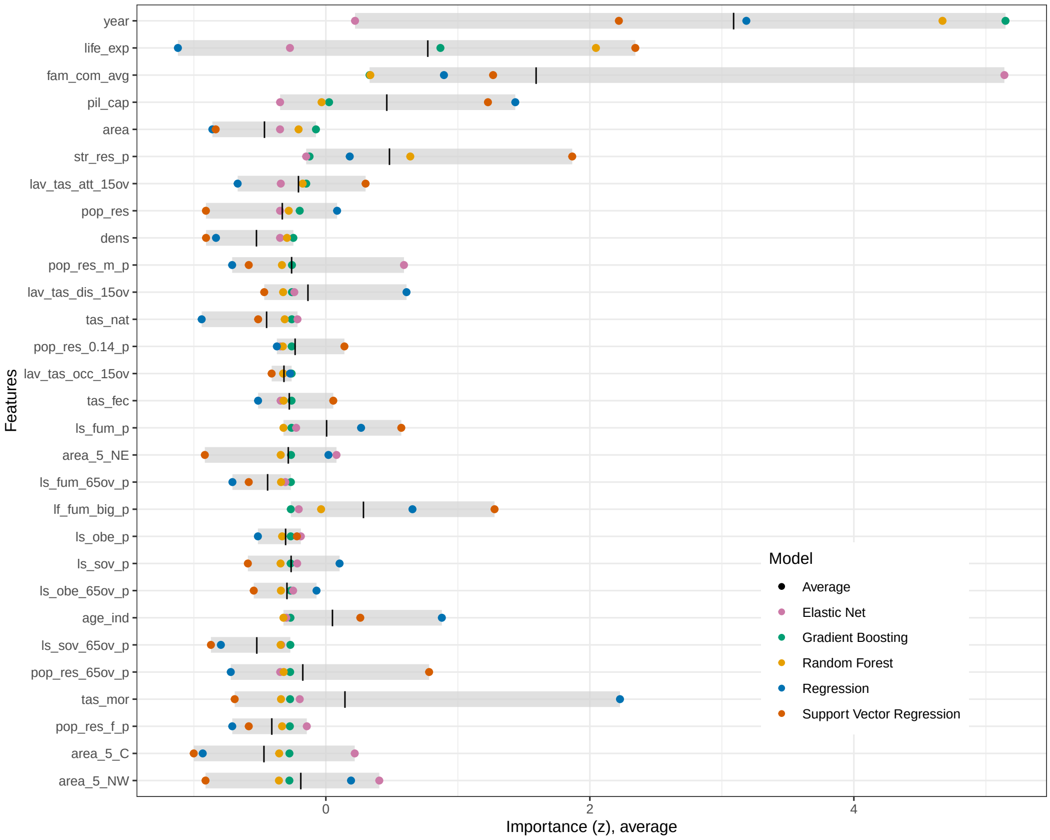
4.2 The determinants of health expenditure

Thereafter, once the good predictability of our model has been verified, we classified our features based on their importance (meaning how much the model's performance decreases when a specific feature is randomly permuted during the prediction process). The graphical representation in Figure 1 helps to better understand the core determinants in predicting total per capita healthcare expenditure in Italy at the regional level. Features' importance is ordered based on the best algorithm, in this case, GBM—selected in the Section 4.1—, and the average importance value is selected for each model and each feature[4]. Furthermore, to make them comparable we standardised the values as well as reporting a general average importance value for each feature.

The variable of `year` takes first place in the importance ranking followed by life expectancy (`life_exp`) and the average number of members in a family (`fam_com_avg`). In fourth place we find one of the most debated determinants of healthcare expenditure, GDP per capita (`pil_cap`). Moreover, among the first ten core features we find not only variables related to population, both in absolute (`pop_res`) and density terms (`dens`), but also the share of foreign residents (`str_res_p`), that of male population (`str_res_m_p`), and the labour activity rate (`lav_tas_att_15ov`).

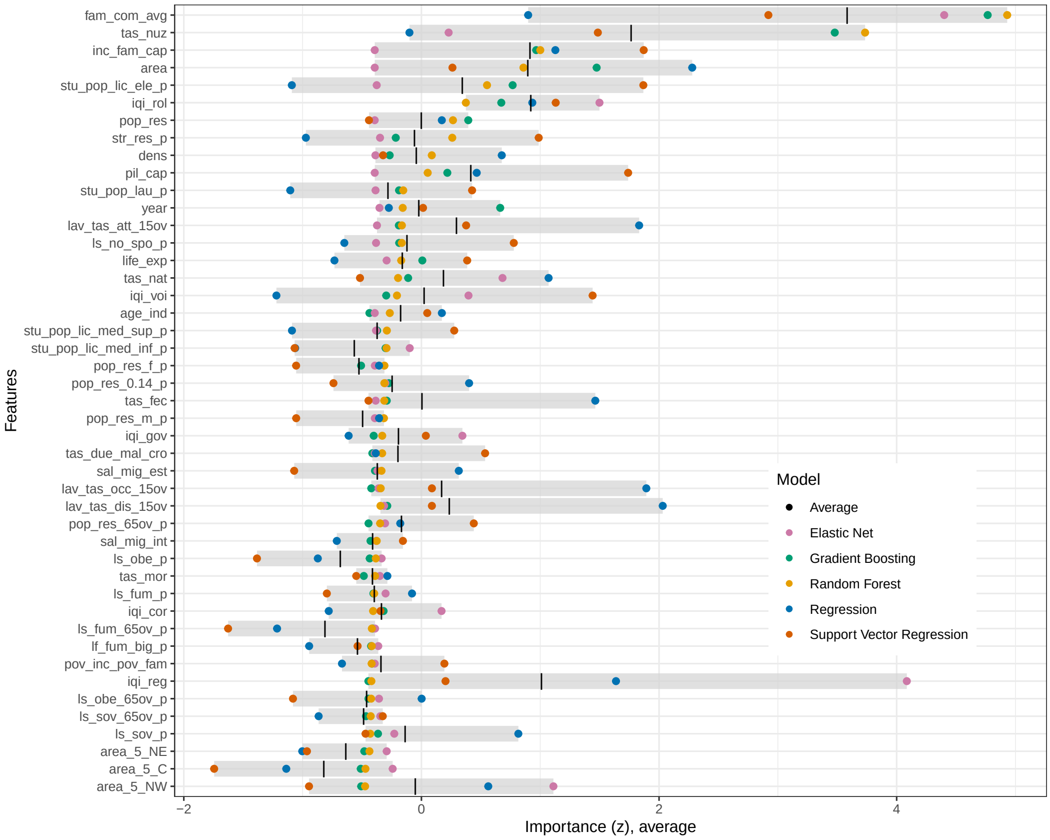
[4] The feature importance is calculated over the entire data frame (train and test) by selecting the best hyper parameters obtained from the trained models.

Figure 1: *Feature importance, total healthcare expenditure (1994–2018)*



With the same approach we calculated features' importance also for the second data frame (2005–2018) reported on Figure 2. We can clearly observe how the feature `year` loses its previous importance within this different framework placing only tenth. The most important feature now is represented by the average size of families (`fam_com_avg`), followed by the marriage rate (`tas_nuz`), and the household per capita income (`inc_fam_cap`). Subsequently, we find the amount of resident population (`pop_res`) and one institutional variable, that associated to rule of law (`iqi_rol`). Among the ten most important features we find two related to educational attainment, first the share of population with at least a primary school certificate (`stu_pop_lic_ele_p`) and then the share of university graduated population (`stu_pop_lau_p`). Moreover, among these most important determinants is still GDP per capita (`pil_cap`) which represents a core feature.

Figure 2: *Feature importance, total healthcare expenditure (2005–2018)*



For the private and public component of our target variable, features' importance could be found in the Appendix in Figures A1 and A2 for public and in Figures A3 and A4 for private healthcare expenditure. Differences with respect to the total expenditure are minimal, especially for the private side where, for example, the most important feature is that of is that of the average number of family components.

After identifying the most important features in predicting healthcare expenditure, it could be useful to investigate the functional impact of each feature on our target variable with the possibility to identify also possible non-linearity. By relying on a partial dependence plot (PDP) we could graphically disentangle this relationship without the necessity of a previous mathematical model in the functional relationship. PDPs relative to the first ten features are reported in Figure 3. Regarding public and private healthcare expenditure, PDPs plots could be found in the Appendix showmn in Figures A5 and A6. Since we initially performed 100 different splittings into train and test samples, in each graph we reported an equal number of PDP (grey lines), all centered on zero and averaged (red line).

The time (`year`) component represents the leading determinant for total healthcare expenditure in Italy and this relationship is strictly positive meaning that over time the expenditure has grown, expeditiously in the early years and slower around 2005 becoming flat after 2010. It is interesting to compare the PDP for this feature also divided by its private and public share. Furthermore, within this latter component, we also performed an additional disaggregation by considering only the spending on health services provided directly, which represents the largest share of public expenditure. For this additional target variable, we repeated the same methodology previously described, and its relative PDP plot is computed by considering the GBM algorithm. We can clearly observe how while the positive relationship between year and total expenditure stopped to grow starting from 2010, instead, that of private expenditure showed an abrupt vertical growth which continued, with a lower grade, in the next years. Therefore, the two components after that year had a diametrically opposite intensity impact on healthcare expenditure. Furthermore, by observing the public component, we can clearly observe how the public expenditure for direct services clearly showed a negative downturn. The importance of the time (`year`) component and the trends in PDPs can be partially explained by the technological progress (Gerdtham et al., 1992; Newhouse, 1992; Roberts, 2000; White, 2007) and partially by the huge institutional reforms involving the Italian NHS during the last decades (Lagravinese et al., 2019). Indeed, after one of the most important federal reforms (Legislative Decree 56/2000), each region became responsible for the organization of the health system, following the guidelines defined by the central government. However, the separation of financing responsibilities from expenditure responsibilities in the provision of uniform levels of service has provided a non-negligible incentive to the uncontrolled growth of Italian health expenditure. This has historically contributed to creating bailing out expectations in regional behaviour (Liberati, 2003), in a context of often inadequate regional health governance and accountability (Carinci, Caracci, Di Stanislao, & Moirano, 2012; Lagravinese & Paradiso, 2014). For these reasons, since 2007, seven regions—Lazio, Abruzzo, Liguria, Campania, Molise, Sicily, and Sardinia—have been placed under strict repayment plans. Calabria, Piedmont, and Puglia were also included in these plans in 2009 and 2010. These plans take the form of comprehensive programmes designed for industrial restructuring, aiming to rein in spending factors that have gone beyond the regions' control. Starting in these years, there was a notable deceleration in the growth of current expenditures, particularly evident in the southern regions. Therefore, the reduction in the growth of public spending, as evident in Figure 4c, can be partly attributed to the limitations imposed by strict repayment plans and financial constraints stemming from the economic downturn, as noted by Lagravinese and Resce (2020). However, as depicted in Figure 4b, the decline in public spending coincided with a significant rise in private expenditure. This indicates that the budget reductions led to an escalation in private expenses for healthcare, potentially heightening the risk of unequal access to medical services, as highlighted, among others, by De Matteis, Ishizaka, and Resce (2019) and Cirulli and Marini (2023).

The Italian NHS, once considered one the of the most efficient healthcare system worldwide (WHO, 2000), nowadays shows a GDP share of expenditure lower when compared to other European and OECD countries. In 2022 the total public expenditure was about 131 billion euros, 6.8% of GDP, while the average OECD share was equal to 7.1%. This amount corresponds to a per capita expenditure slightly over € 3,091, which is more than € 800 lower compared to the average OECD expenditure (€ 3,920) OECD (2023). This performance places Italy at the 20th place among OECD countries in terms both of total share as well as per capita expenditure (13th and 16th in Europe, respectively). This data depicts a decreasing trend in public healthcare expenditure in Italy which started with the debt sovereign crisis after 2010 and for the upcoming future the trend will continue to follow the same decreasing pathway since the expected expenditure for the upcoming years would be decreasing to 6.2% of GDP in 2024 (Documento di Economia e Finanza 2023).

Regarding life expectancy (`life_exp`), as expected, it depicts a positive relationship with total healthcare expenditure which becomes stiffer once exceeding 80 years becoming flatter after 82. When comparing this PDP with the public and private components the relationship remained the same (Figure 5b. By looking at the average family size (`fam_com_avg`) we do observe a negative relationship—with the exception of a small initial positive effect derived from the public expenditure—where larger families tend to reduce healthcare expenditure (Figure 6b. This result is in line with the fact that families spread less resources as family size increases and the negative trend is indeed more linear for the out-of-pocket component (Marmot & Wilkinson, 2005). In fact, the study of Dormont et al. (2006) shows a negative impact of the variable household size on health expenditure. Furthermore, the presence of one or more children in a family can also represent a kind of life insurance for older members by providing potentially free caregiver assistance.

When considering GDP per capita (`pil_cap`) we do observe a clearly positive role of it on determining healthcare expenditure in line with the general literature Kleiman (1974) and specific Italian studies (Giannoni & Hitiris, 2002; Lopreite & Mauro, 2017). Although positive, the relationship is not clearly linear but rather a stepwise trend which clearly derives from its private component (Figure 7). In fact, while public expenditure's PDP follows a more linear trend, that of private expenditure does not, characterised, instead, by a steep growth after about € 22,000 of income per capita representing a possible threshold effect. Related to this feature we can also describe the path showed by the labour activity rate (`lav_tas_att_15ov`) which show a clear similar path, even when investigating separately between the private and the public components as showed in Figure 8. Moreover, some studies (Felder et al., 2000; Seshamani & Gray, 2004b; Zweifel et al., 1999) used the inverse of the Mills ratio as individuals' propensity to participate in the labour market and coefficients were positive, hence in line with what emerges from our PDPs.

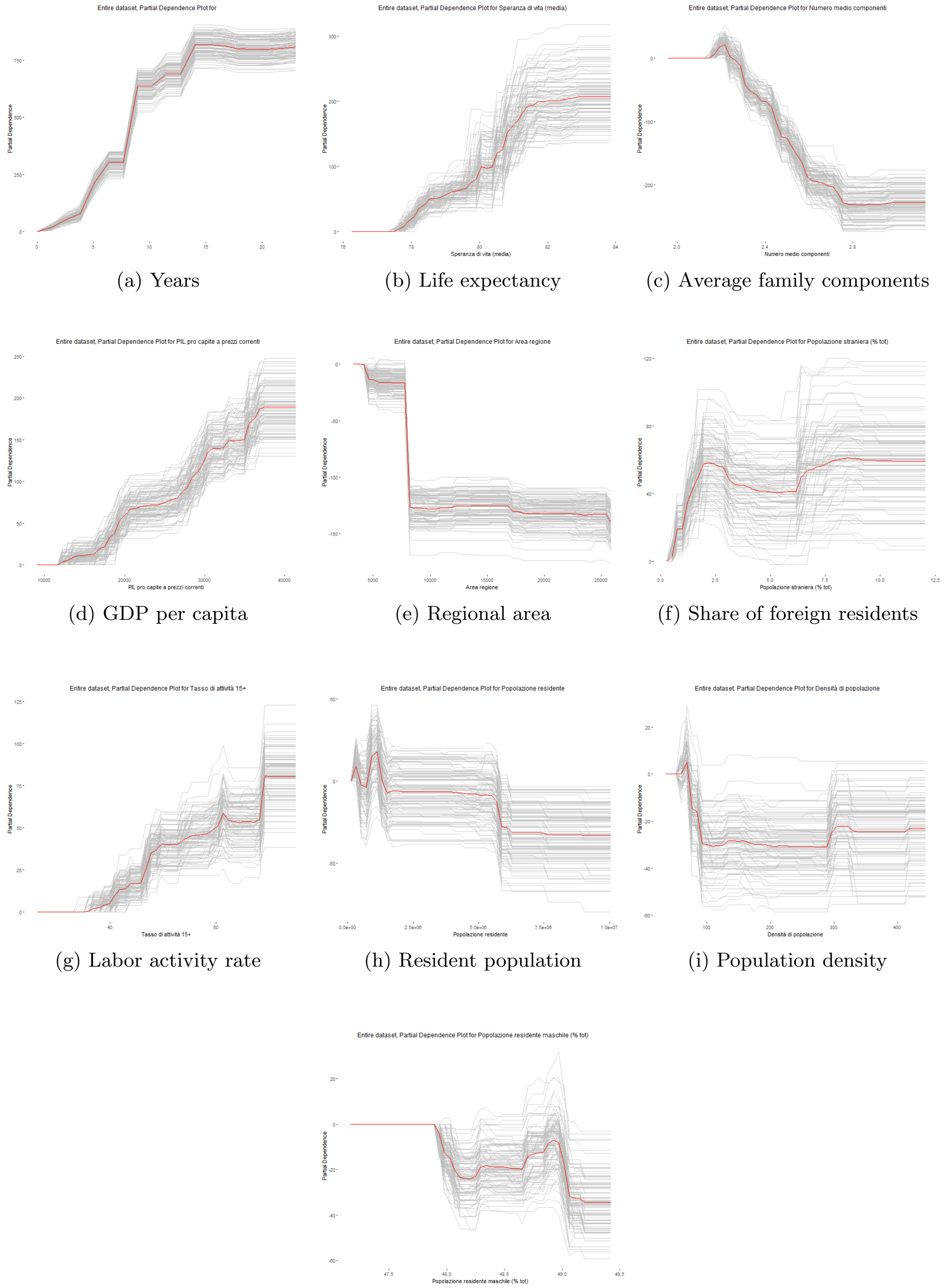
Foreign residents (`str_res_p`), among the core determinants of total healthcare expenditure, depicts a positive relationship, although nonlinear as shown in Figure 9. It steeply grows for lower share of foreign residents reaching its ceiling although characterised by a slight contraction between 2.5% and 6%. Nonetheless, when comparing public and out-of-pocket PDPs we observe how, for the latter, the relationship is far more linear and positive stressing how foreign residents tend to spend mainly on private healthcare. These findings exhibit some variance from the findings presented in the study by Bettin and

Sacchi (2020), where it was discovered that a rise in the immigrant population as a proportion of the overall residents results in a reduction of public health spending. These disparities could be attributed in part to the different identification strategies (they were interested in causal inference while our results show just associations between feature and the outcome), and in part to the distinct specification models employed. Bettin and Sacchi (2020) employed conventional regression methods, which lack the ability to account for potential nonlinear and threshold effects. In contrast, the GBM, a tree-based technique utilised in this study, can effectively capture such effects.

The regional area size (`area`) depicts another negative relationship but highly nonlinear with an initial vertical fall due to the fact that some small regions such as Valle d'Aosta or Molise despite a small area surface, do have a relatively costly sanitary system. This effect could be depicted also by looking at the variables of the resident population and population density since these same small regions have few residents and therefore low population density. These results show the effects of economies of scale which have been widely documented in healthcare (Hurley, 2000). In particular, it has been shown that increasing the dimension of health systems can yield operational efficiencies reducing, among others, administrative costs (Woolhandler & Himmelstein, 1991).

Although not identified as a core determinant of healthcare expenditure, it is interesting to investigate the PDP for the aging index (`age_ind`) as depicted in Figure 10. For total expenditure, we find a substantial null effect or slightly negative up to 150 (when for each 100 young individuals there corresponds 150 elderly ones). Here the relationship becomes rapidly positive remaining stable up to the value of 200 on the index, where, substantially for each young individual there are two elders. Here the PDP shows a new jump which strengthens the already positive role played by this variable in boosting healthcare expenditure. This staple pattern, characterized by two main thresholds emerges also for the public components. Conversely, by looking at the out-of-pocket expenditure, the relationship is highly positive and far more linear. Generally, the greater growth of the elderly population compared to the young follows what already showed with life expectancy (`life_exp`) ad it boost health expenditure as shown for the Italian case by Lopreite and Mauro (2017).

Figure 3: *Partial Dependence Plots for selected features, total healthcare expenditure (1994–2018)*



(a) Years

(b) Life expectancy

(c) Average family components

(d) GDP per capita

(e) Regional area

(f) Share of foreign residents

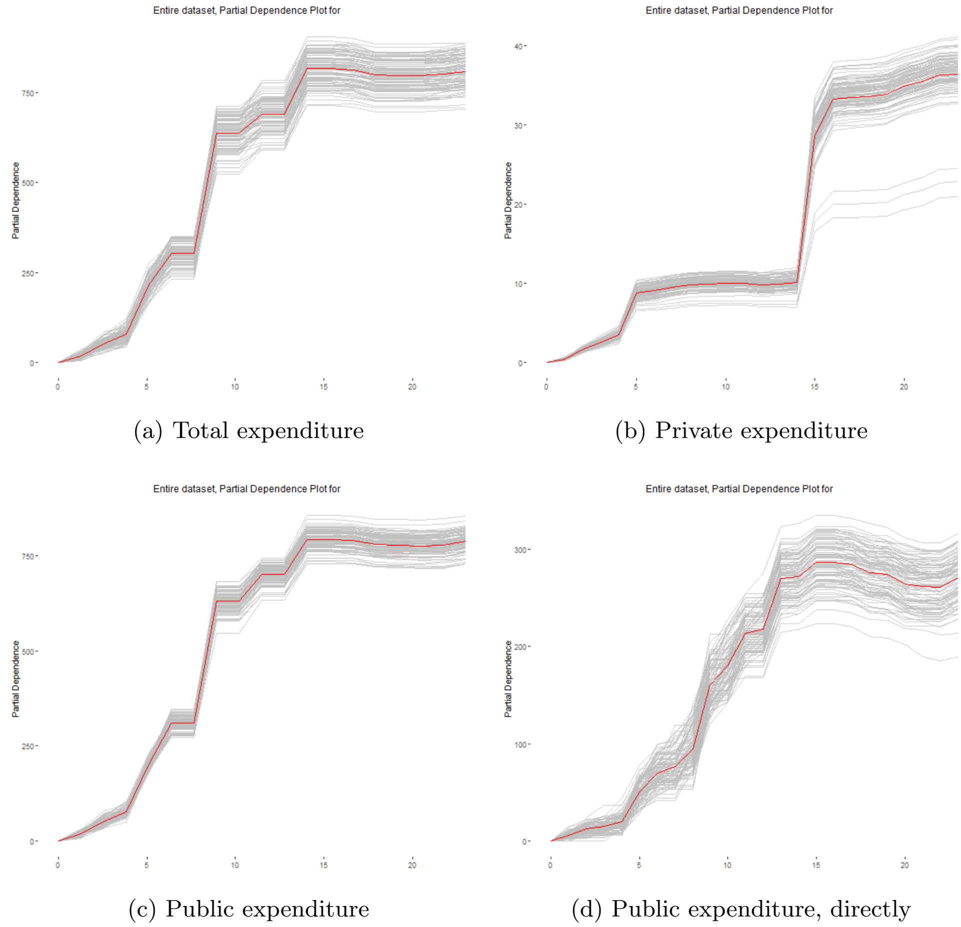
(g) Labor activity rate

(h) Resident population

(i) Population density

(j) Share of male population

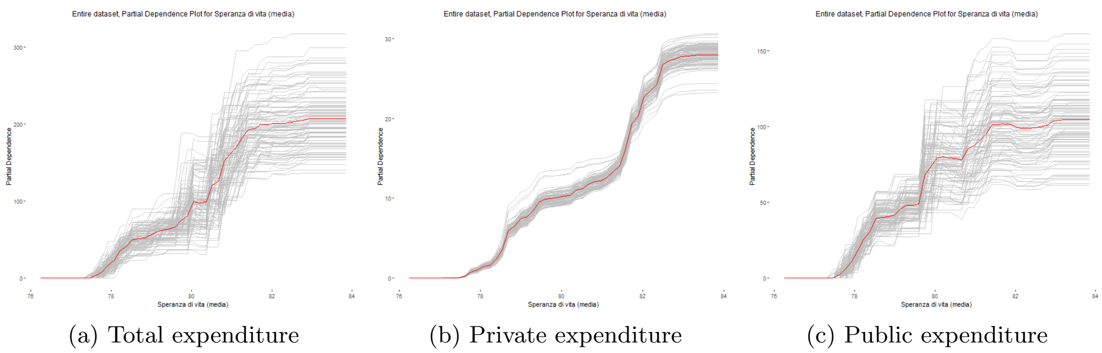
Figure 4: *Partial Dependence Plots for the feature* `year`*, comparison between different target variables (1994–2018)*



(a) Total expenditure (b) Private expenditure

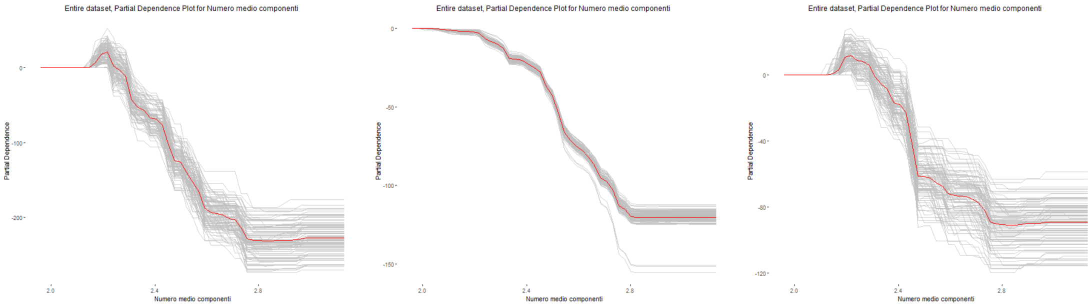
(c) Public expenditure (d) Public expenditure, directly

Figure 5: *Partial Dependence Plots for the feature* `life_exp`*, comparison between different target variables (1994–2018)*



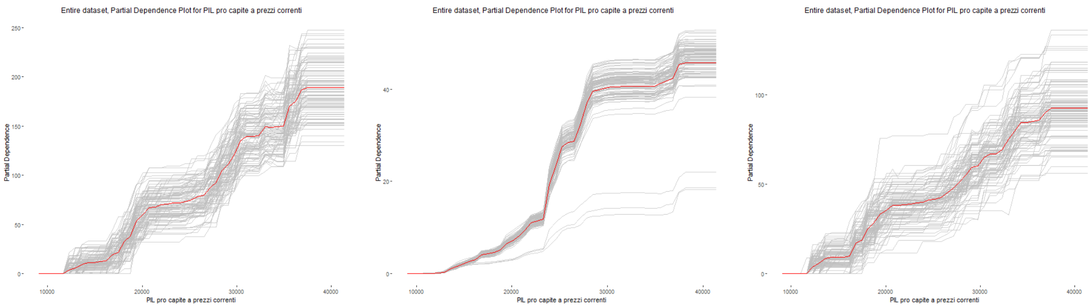
(a) Total expenditure (b) Private expenditure (c) Public expenditure

Figure 6: *Partial Dependence Plots for the feature* `fam_com_avg`*, comparison between different target variables (1994–2018)*



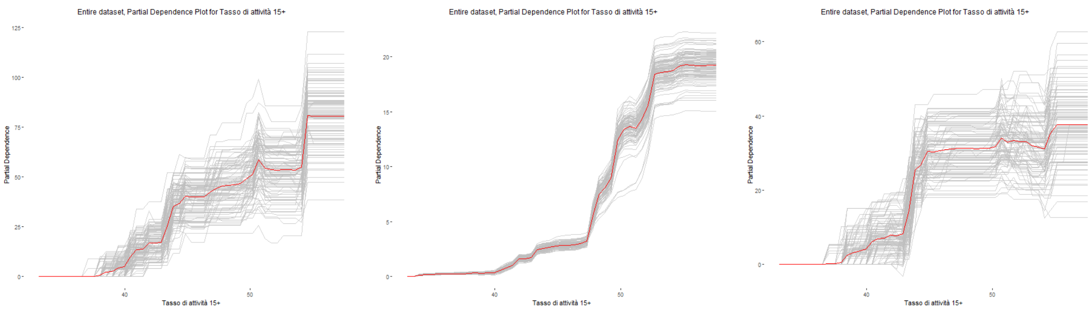
(a) Total expenditure (b) Private expenditure (c) Public expenditure

Figure 7: *Partial Dependence Plots for the feature* `pil_cap`*, comparison between different target variables (1994–2018)*



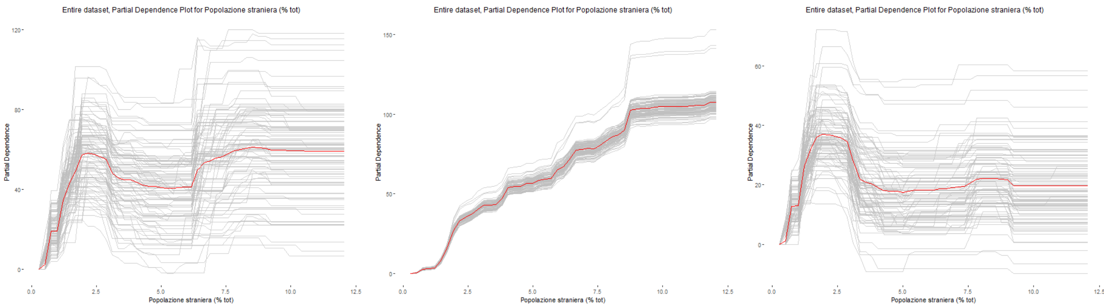
(a) Total expenditure (b) Private expenditure (c) Public expenditure

Figure 8: *Partial Dependence Plots for the feature* `lav_tas_att_15ov`*, comparison between different target variables (1994–2018)*



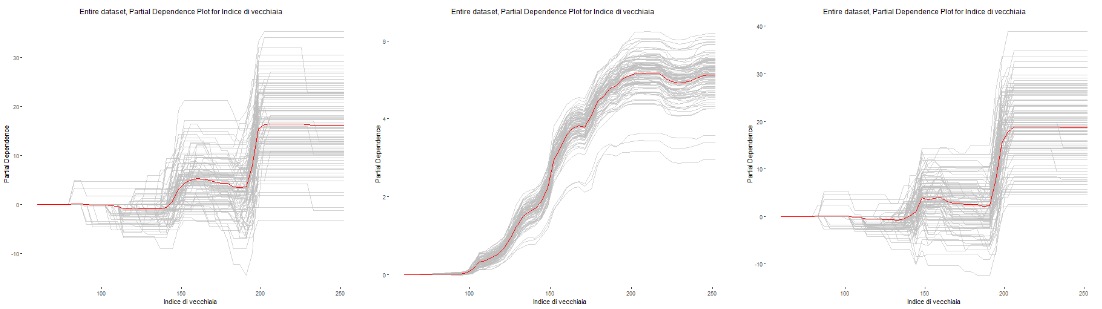
(a) Total expenditure (b) Private expenditure (c) Public expenditure

Figure 9: *Partial Dependence Plots for the feature* `str_res_p`*, comparison between different target variables (1994–2018)*



(a) Total expenditure (b) Private expenditure (c) Public expenditure

Figure 10: *Partial Dependence Plots for the feature* `age_ind`*, comparison between different target variables (1994–2018)*



(a) Total expenditure (b) Private expenditure (c) Public expenditure

5 Conclusions

Forecasting healthcare costs offers vital insights for making decisions, shaping policies, financial strategising, and resource allocation. This aids in the streamlined and resourceful provision of healthcare, fostering accessible and lasting medical services for both individuals and communities. Precise predictions play a pivotal role from a policy standpoint, as they guide numerous National Health Systems (NHS) in distributing funds locally to ensure fairness.

The literature showed how empirically ML tools are more suitable for addressing prediction problems rather than standard econometric models, hence showing a potentially greater representation, a potentially better tool for the political decision-maker. Therefore, the allocation of healthcare resources, in a fragmented and heterogeneous context characterized by an ever-increasing elderly population, represents a fundamental issue within the national context suitable for being addressed with these new evaluation tools.

This study develops a ML model for forecasting healthcare expenses whose outcomes can guide the allocation of national health funds towards Italian regions with heterogeneous needs. We constructed our model by relying on data provided by Istat over the period 1994—2018 by identifying Gradient Boosting as the best predictive algorithm. Moreover, we tested our model using 2019 data obtaining good forecasting results (MAPE equal to 2.89 and R^2 to 0.82). This underscores the effectiveness of our model in predicting healthcare expenditures across different Italian regions, highlighting its proficiency based on the chosen features. Furthermore, we also disentangled the total healthcare expenditure between public and private components.

Among the principal determinant features, significant in predicting total health regional expenditure, we identified time, the average number of family members, and the size of the regional area. Moreover, several determinants widely debated in the literature, such as GDP per capita and life expectancy, also emerged as top contributors. To explore these relationships, we employed partial dependence plots (PDP), a graphical tool able to unravel the relationship between each feature and per capita healthcare expenditure, our target outcome. For instance, the time feature displays a positive correlation with healthcare expenditure that levels off after 2010, while the private component depicts a marked increase. GDP per capita exhibits an incremental step pattern with distinct thresholds. Household size exhibits a negative link with healthcare expenditure, particularly in the private sector, highlighting the trend of distributing fewer resources as family size grows. Additionally, it underscores how younger family members might serve as life insurance for older people, potentially offering cost-free caregiver support. Notably, life expectancy emerges as another pivotal determinant, positively influencing healthcare expenditure. Its impact is especially pronounced after 80 years of age, underscoring the profound influence of Italy's elderly population on driving NHS costs.

Despite the potential of ML algorithms, since the proposed predictive tool is a data-driven approach, it would benefit from an extension of the data. Unfortunately, since the NHS is distributed at the regional level, this represents a significant limit in terms of sample size. Nonetheless, we tried to turn around this issue in several ways: (i) by performing 100 different splittings of our samples, then training our ML models that many times; (ii) by selecting two different samples, the primary one with a larger longitudinal component while the other with had a shorter time coverage offset by more features. Also, with this latter solution we were able to obtain a good predictive model. Another possibility would be to disaggregate data at the local health authority (*i.e.*, *Azienda Sanitaria Locale* – ASL). However, while it would enhance data granularity, it would represent a burden (if not a cost) in terms of data availability and comparability.

References

Altavilla, A. M., Mazza, A., & Monaco, L. (2016). Effetti dell'invecchiamento della popolazione sulla spesa del sistema sanitari. *Rivista Italiana di Economia, Demografia e Statistica*, *LXVIII*, 151–158.

Andini, M., Ciani, E., de Blasio, G., D'Ignazio, A., & Salvestrini, V. (2018). Targeting with machine learning: An application to a tax rebate program in italy. *Journal of Economic Behavior & Organization*, *156*, 86–102.

Ang, J. B. (2010). The determinants of health care expenditure in Australia. *Applied Economics Letters*, *17*(7), 639–644.

Antulov-Fantulin, N., Lagravinese, R., & Resce, G. (2021). Predicting bankruptcy of local government: A machine learning approach. *Journal of Economic Behavior & Organization*, *183*, 681–699.

Atella, V., & Conti, V. (2014). The effect of age and time to death on primary care costs: The Italian experience. *Social Science & Medicine*, *114*, 10–17.

Baker, L. C., & Wheeler, S. K. (1998). Managed care and technology diffusion: The case of mri: Could managed care inhibit the spread of beneficial technologies? one case study—magnetic resonance imaging—provides some answers. *Health Affairs*, *17*(5), 195–207.

Baltagi, B. H., Lagravinese, R., Moscone, F., & Tosetti, E. (2017). Health care expenditure and income: A global perspective. *Health Economics*, *26*(7), 863–874.

Baltagi, B. H., & Moscone, F. (2010). Health care expenditure and income in the OECD reconsidered: Evidence from panel data. *Economic Modelling*, *27*(4), 804–811.

Barros, P. P. (1998). The black box of health care expenditure growth determinants. *Health Economics*, *7*(6), 533–544.

Basu, A., & Manning, W. G. (2009). Issues for the next generation of health care cost analyses. *Medical Care*, *47*(7), S109–S114.

Bettin, G., & Sacchi, A. (2020). Health spending in italy: the impact of immigrants. *European Journal of Political Economy*, *65*, 101932.

Blanco-Moreno, Á., Urbanos-Garrido, R. M., & Thuissard-Vasallo, I. J. (2013). Public healthcare expenditure in Spain: measuring the impact of driving factors. *Health Policy*, *111*(1), 34–42.

Blomqvist, Å. G., & Carter, R. A. (1997). Is health care really a luxury? *Journal of Health Economics*, *16*(2), 207–229.

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32.

Breyer, F., & Felder, S. (2006). Life expectancy and health care expenditures: a new calculation for Germany using the costs of dying. *Health Policy*, *75*(2), 178–186.

Brnabic, A., & Hess, L. M. (2021). Systematic literature review of machine learning methods used in the analysis of real-world data for patient-provider decision making. *BMC medical informatics and decision making*, *21*(1), 1–19.

Caballer-Tarazona, V., Guadalajara-Olmeda, N., & Vivas-Consuelo, D. (2019). Predicting healthcare expenditure by multimorbidity groups. *Health Policy*, *123*(4), 427–434.

Carinci, F., Caracci, G., Di Stanislao, F., & Moirano, F. (2012). Performance measurement in response to the tallinn charter: experiences from the decentralized italian framework. *Health Policy*, *108*(1), 60–66.

Carr-Hill, R. A., Sheldon, T. A., Smith, P., Martin, S., Peacock, S., & Hardman, G. (1994). Allocating resources to health authorities: development of method for small area analysis of use of inpatient services. *British Medical Journal*, *309*(6961), 1046.

Carrieri, V., Lagravinese, R., & Resce, G. (2021). Predicting vaccine hesitancy from area-level indicators: A machine learning approach. *Health Economics*, *30*(12), 3248–3256.

Cerqua, A., Di Stefano, R., Letta, M., & Miccoli, S. (2021). Local mortality estimates during the covid-19 pandemic in italy. *Journal of Population Economics*, *34*(4), 1189–1217.

Cerqua, A., & Letta, M. (2022). Local inequalities of the covid-19 crisis. *Regional science and urban economics*, *92*, 103752.

Cirulli, V., & Marini, G. (2023). Are austerity measures really distressing? evidence from italy. *Economics & Human Biology*, *49*, 101217.

Clemente, J., Marcuello, C., Montañés, A., & Pueyo, F. (2004). On the international stability of health care expenditure functions: are government and private functions similar? *Journal of Health Economics*, *23*(3), 589–613.

Cook, B. L., & Manning, W. G. (2009). Measuring racial/ethnic disparities across the distribution of health care expenditures. *Health Services Research*, *44*(5p1), 1603–1621.

Crivelli, L., Filippini, M., & Mosca, I. (2006). Federalism and regional health care expenditures: an empirical analysis for the Swiss cantons. *Health Economics*, *15*(5), 535–541.

Deb, P., & Burgess, J. F. (2003). *A quasi-experimental comparison of econometric models for health care expenditures.* New York, NY.

Deb, P., & Trivedi, P. K. (1997). Demand for medical care by the elderly: a finite mixture approach. *Journal of applied Econometrics*, *12*(3), 313–336.

De Matteis, D., Ishizaka, A., & Resce, G. (2019). The 'postcode lottery'of the italian public health bill analysed with the hierarchy stochastic multiobjective acceptability analysis. *Socio-Economic Planning Sciences*, *68*, 100603.

De Meijer, C., Koopmanschap, M., d'Uva, T. B., & Van Doorslaer, E. (2011). Determinants of long-term care spending: age, time to death or disability? *Journal of Health Economics*, *30*(2), 425–438.

Di Matteo, L. (2005a). The macro determinants of health expenditure in the United States and Canada: assessing the impact of income, age distribution and time. *Health Policy*, *71*(1), 23–42.

Di Matteo, L. (2005b). What Drives Provincial Health Expenditure. *Canadian Tax Journal*, *52*(4), 1102–1120.

Di Matteo, L., & Di Matteo, R. (1998). Evidence on the determinants of Canadian provincial government health expenditures: 1965–1991. *Journal of Health Economics*, *17*(2), 211–228.

Dormont, B., Grignon, M., & Huber, H. (2006). Health expenditure growth: reassessing the threat of ageing. *Health Economics*, *15*(9), 947–963.

Dreger, C., & Reimers, H.-E. (2005). *Health care expenditures in OECD countries: a panel unit root and cointegration analysis.* Bonn, DE. (IZA Discussion Papers, No. 1469)

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. *Advances in neural information processing systems*, *9*, 155–161.

EC. (2009). *The 2009 ageing report: economic and budgetary projections for the EU-27 Member States (2008-2060).* Luxembourg. (Joint Report prepared by the European Commission (DG ECFIN) and the Economic Policy Committee (AWG))

Einav, L., & Levin, J. (2014). Economics in the age of big data. *Science*, *346*(6210), 1243089.

Felder, S., Meier, M., & Schmitt, H. (2000). Health care expenditure in the last months of life. *Journal of Health Economics*, *19*(5), 679–695.

Fonseca, R., Langot, F., Michaud, P.-C., & Sopraseuth, T. (2023). *Understanding cross-country differences in health status and expenditures.* Cambridge, MA. (NBER Working Papers)

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189–1232.

Gerdtham, U.-G. (1992). Pooling international health care expenditure data. *Health Economics*, *1*(4), 217–231.

Gerdtham, U.-G., & Jönsson, B. (2000). International comparisons of health expenditure: theory, data and econometric analysis. In A. J. Culyer & J. P. Newhouse (Eds.), *Handbook of health economics* (Vol. 1, pp. 11–53). Amsterdam, NL: Elsevier.

Gerdtham, U.-G., Jönsson, B., MacFarlan, M., & Oxley, H. (1998). The determinants of health expenditure in the OECD countries: a pooled data analysis. In P. Zweifel (Ed.), *Health, the medical profession, and regulation* (pp. 113–134). New York, NY-US: Springer.

Gerdtham, U.-G., Søgaard, J., Andersson, F., & Jönsson, B. (1992). An econometric analysis of health care expenditure: A cross-section study of the OECD countries. *Journal of Health Economics*, *11*(1), 63–84.

Getzen, T. E. (1992). Population aging and the growth of health expenditures. *Journal of Gerontology*, *47*(3), S98–S104.

Giannoni, M., & Hitiris, T. (2002). The regional impact of health care expenditure: the case of Italy. *Applied Economics*, *34*(14), 1829–1836.

Golinelli, D., Toscano, F., Bucci, A., Lenzi, J., Fantini, M. P., Nante, N., & Messina, G. (2017). Health Expenditure and All-Cause Mortality in the 'Galaxy' of Italian Regional Healthcare Systems: A 15-Year Panel Data Analysis. *Applied Health Economics and Health Policy*, *15*, 773–783.

Gravelle, H., Santos, R., & Siciliani, L. (2014). Does a hospital's quality depend on the quality of other hospitals? a spatial econometrics approach. *Regional science and urban economics*, *49*, 203–216.

Hansen, P., & King, A. (1996). The determinants of health care expenditure: a cointegration approach. *Journal of Health Economics*, *15*(1), 127–137.

Hitiris, T., & Posnett, J. (1992). The determinants and effects of health expenditure in developed countries. *Journal of Health Economics*, *11*(2), 173–181.

Hurley, J. (2000). An overview of the normative economics of the health sector. *Handbook of health economics*, *1*, 55–118.

Hyun, K.-R., Kang, S., & Lee, S. (2016). Population aging and healthcare expenditure in Korea. *Health Economics*, *25*(10), 1239–1251.
Istat. (2023a). *Health for All.* Rome, Italy. (Italian Statistic Institute. Retrieved from: https://www.istat.it/it/archivio/14562 [Accessed April 20, 2023])
Istat. (2023b). *Istat Statistics.* Rome, Italy. (Italian Statistic Institute. Retrieved from: http://dati.istat.it/?lang=en [Accessed April 20, 2023])
Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790–794.
Jones, A. M. (2011). Models for health care: Oxford handbook of economic forecasting. In M. Clements & D. F. Hendry (Eds.), *Oxford Handbook of Economic Forecasting* (pp. 625–654). Oxford, UK: Oxford University Press.
Jones, A. M., Lomas, J., & Rice, N. (2014). Applying beta-type size distributions to healthcare cost regressions. *Journal of Applied Econometrics*, *29*(4), 649–670.
Jones, A. M., Lomas, J., & Rice, N. (2015). Healthcare cost regressions: going beyond the mean to estimate the full distribution. *Health Economics*, *24*(9), 1192–1212.
Karatzas, G. (2000). On the determination of the US aggregate health care expenditure. *Applied Economics*, *32*(9), 1085–1099.
Karlsson, M., & Klohn, F. (2014). Testing the red herring hypothesis on an aggregated level: ageing, time-to-death and care costs for older people in Sweden. *The European Journal of Health Economics*, *15*, 533–551.
Kleiman, E. (1974). The determinants of national outlay on health. In M. Perlman (Ed.), *The Economics of Health and Medical Care: Proceedings of a Conference Held by the International Economic Association at Tokyo* (pp. 66–88). London/Basingstoke, UK.
Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction policy problems. *American Economic Review*, *105*(5), 491–495.
Koenig, L., Siegel, J. M., Dobson, A., Hearle, K., Ho, S., & Rudowitz, R. (2003). Drivers of healthcare expenditures associated with physician services. *The American Journal of Managed Care*, *9*(S1), 34–42.
Lagravinese, R., Liberati, P., & Resce, G. (2019). Exploring health outcomes by stochastic multicriteria acceptability analysis: An application to Italian regions. *European Journal of Operational Research*, *274*(3), 1168–1179.
Lagravinese, R., & Paradiso, M. (2014). Corruption and health expenditure in italian regions. *Economics Bulletin*, *34*(3), 1498–1511.
Lagravinese, R., & Resce, G. (2020). Fiscal federalism and national health standards in Italy: implications for redistribution. In G. Coco & C. De Vincenti (Eds.), *Una Questione Nazionale: Il Mezzogiorno da ≪problema≫ a ≪opportunità≫*. Bologna, IT: il Mulino.
Leu, R. (1986). The public-private mix and international health care costs. *The public and private health services*, 41–63.
Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting.* London, UK: Butterworths.
Liberati, P. (2003). Fiscal federalism and national health standards in Italy: implications for redistribution. In D. Franco & A. Zanardi (Eds.), *I sistemi di welfare tra decentramento regionale e integrazione europea* (pp. 241–273). Milan, IT: Franco Angeli.
López-Casasnovas, G., & Saez, M. (2007). A multilevel analysis on the determinants of regional health care expenditure: a note. *The European Journal of Health Economics*, *8*, 59–65.
Lopreite, M., & Mauro, M. (2017). The effects of population ageing on health care expenditure: a Bayesian VAR analysis using data from Italy. *Health Policy*, *121*(6), 663–674.
Manning, W. G., Basu, A., & Mullahy, J. (2005). Generalized modeling approaches to risk adjustment of skewed outcomes data. *Journal of Health Economics*, *24*(3), 465–488.
Marmot, M., & Wilkinson, R. (2005). *Social determinants of health.* Oxford, UK: Oxford University Press.
Martín, J. J. M., Puerto Lopez del Amo Gonzalez, M., & Dolores Cano Garcia, M. (2011). Review of the literature on the determinants of healthcare expenditure. *Applied Economics*, *43*(1), 19–46.
Meara, E., White, C., & Cutler, D. M. (2004). Trends in medical spending by age, 1963–2000. *Health Affairs*, *23*(4), 176–183.
Moreno, J. J. M., Pol, A. P., Abad, A. S., & Blasco, B. C. (2013). Using the r-mape index as a resistant measure of forecast accuracy. *Psicothema*, *25*(4), 500–506.
Mosca, I. (2007). Decentralization as a determinant of health care expenditure: empirical analysis for

OECD countries. *Applied Economics Letters*, *14*(7), 511–515.

Moscone, F., & Tosetti, E. (2010). Health expenditure and income in the United States. *Health Economics*, *19*(12), 1385–1403.

Mullahy, J. (2009). Econometric modeling of health care costs and expenditures: a survey of analytical issues and related policy considerations. *Medical Care*, *47*(7), S104–S108.

Murthy, V. N., & Okunade, A. A. (2016). Determinants of US health expenditure: Evidence from autoregressive distributed lag (ARDL) approach to cointegration. *Economic Modelling*, *59*, 67–73.

Murthy, V. N., & Ukpolo, V. (1994). Aggregate health care expenditure in the united states: evidence from cointegration tests. *Applied Economics*, *26*(8), 797–802.

Newhouse, J. P. (1977). Medical-care expenditure: a cross-national survey. *The journal of human resources*, *12*(1), 115–125.

Newhouse, J. P. (1992). Medical care costs: how much welfare loss? *Journal of Economic perspectives*, *6*(3), 3–21.

Nifo, A., & Vecchione, G. (2021). *Institutional Quality Index.* (IQI Dataset. Retrieved from: https://sites.google.com/site/institutionalqualityindex/dataset/iqi-dataset?authuser=0 [Accessed April 20, 2023])

OECD. (2023). *Health resources - Health spending.* Paris, FR. (Organisation for Economic Co-operation and Development. Retrieved from: https://data.oecd.org/healthres/health-spending.htm [Accessed September 15, 2023])

Okunade, A. A., & Karakus, M. C. (2001). Unit root and cointegration tests: timeseries versus panel estimates for international health expenditure models. *Applied Economics*, *33*(9), 1131–1137.

Okunade, A. A., & Murthy, V. N. (2002). Technology as a 'major driver' of health care costs: a cointegration analysis of the newhouse conjecture. *Journal of Health Economics*, *21*(1), 147–159.

Protonotari, A., Buriani, A., Mezzetti, F., Quargnolo, E., Castaldini, I., Catalano, N., ... Bragonzi, G. (2007). Analisi economica sull'assorbimento di risorse da parte della popolazione anziana dell'Ausl di Bologna. *Politiche sanitarie*, *8*(2), 89–94.

Raitano, M., Gabriele, C., Cislaghi, F., Costantini, F., Innocenti, F., Lepore, V., ... Valerio, F. (2007). *A survey of the impact of death-related costs on health-care expenditure and demographics and health costs in italy.* Brussels, BE. (ENEPRI Policy Brief, No. 5)

Rebba, V. (2005). Invecchiamento della popolazione e spesa per il long-term care. *Politiche sanitarie*, *6*(3), 109–115.

Rice, N., & Smith, P. C. (2001). Capitation and risk adjustment in health care financing: an international progress report. *The Milbank Quarterly*, *79*(1), 81–113.

Roberts, J. (2000). Spurious regression problems in the determinants of health care expenditure: a comment on Hitiris (1997). *Applied Economics Letters*, *7*(5), 279–283.

Sarti, S., Terraneo, M., & Bordogna, M. T. (2017). Poverty and private health expenditures in italian households during the recent crisis. *Health Policy*, *121*(3), 307–314.

Seshamani, M., & Gray, A. (2004a). Ageing and health-care expenditure: the red herring argument revisited. *Health Economics*, *13*(4), 303–314.

Seshamani, M., & Gray, A. M. (2004b). A longitudinal study of the effects of age and time to death on hospital costs. *Journal of Health Economics*, *23*(2), 217–235.

Shang, B., & Goldman, D. (2008). Does age or life expectancy better predict health care expenditures? *Health Economics*, *17*(4), 487–501.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Turati, G. (2013). The Italian Servizio Sanitario Nazionale: a renewing tale of lost promises. In C.-F. Joan & L. G. Scott (Eds.), *Federalism and Decentralization in European Health and Social Care* (pp. 47–66). New York, NY-US: Palgrave Macmillan.

Varian, H. R. (2016). Causal inference in economics and marketing. *Proceedings of the National Academy of Sciences*, *113*(27), 7310–7315.

Wagstaff, A. (2009). *Social health insurance vs. tax-financed health systems-evidence from the OECD.* Washington D.C., US. (World Bank policy research Working Papers, No. 4821)

Wagstaff, A., & Moreno-Serra, R. (2009). Europe and central Asia's great post-communist social health insurance experiment: Aggregate impacts on health sector outcomes. *Journal of health economics*, *28*(2), 322–340.

Wang, Z. (2009). The determinants of health expenditures: evidence from US state-level data. *Applied Economics*, *41*(4), 429–435.

Weil, T. P. (1995). Comparisons of medical technology in Canadian, German, and US hospitals. *Journal*

of *Healthcare Management*, *40*(4), 524–533.
Werblow, A., Felder, S., & Zweifel, P. (2007). Population ageing and health care expenditure: a school of 'red herrings'? *Health Economics*, *16*(10), 1109–1126.
White, C. (2007). Health care spending growth: how different is the United States from the rest of the OECD? *Health Affairs*, *26*(1), 154–161.
WHO. (2000). *The world health report 2000. Health systems: improving performance. World Health Organization.* Geneva, CH.
Wong, A., van Baal, P. H., Boshuizen, H. C., & Polder, J. J. (2011). Exploring the influence of proximity to death on disease-specific hospital expenditures: A carpaccio of red herrings. *Health Economics*, *20*(4), 379–400.
Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach.* Mason, OH-US: South-Western Cengage learning.
Woolhandler, S., & Himmelstein, D. U. (1991). The deteriorating administrative efficiency of the us health care system. *New England Journal of Medicine*, *324*(18), 1253–1258.
You, X., & Okunade, A. A. (2017). Income and technology as drivers of Australian healthcare expenditures. *Health Economics*, *26*(7), 853–862.
Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *67*(2), 301–320.
Zweifel, P., Felder, S., & Meiers, M. (1999). Ageing of population and health care expenditure: a red herring? *Health Economics*, *8*(6), 485–496.

Appendix

Table A1: *Models' performances, public healthcare expenditure (1994–2018)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	58.6057	57.3748	48.2828	75.6431	6.0012
MAE	45.1327	44.356	37.5431	56.0848	4.0923
R-squared	0.9796	0.9805	0.9656	0.9867	0.0042
MAPE	3.0834	3.0431	2.5038	3.7648	0.2706
Elastic Net					
RMSE	102.1749	102.1495	80.99	125.7477	9.4094
MAE	80.0897	80.066	66.7746	99.98	6.7484
R-squared	0.9382	0.9391	0.9046	0.9605	0.0115
MAPE	5.4409	5.4206	4.5096	6.8787	0.4282
Random Forest					
RMSE	64.604	64.7502	51.4161	80.7652	6.6465
MAE	48.5669	49.1518	39.1583	60.3321	4.5321
R-squared	0.9753	0.9753	0.961	0.9849	0.0049
MAPE	3.3468	3.3919	2.626	4.2553	0.3141
Regression					
RMSE	102.7652	103.1196	81.2362	127.9384	9.5852
MAE	80.3164	80.0339	67.1147	99.8419	6.6513
R-squared	0.9374	0.9389	0.9012	0.9578	0.0118
MAPE	5.4378	5.4048	4.5115	6.8774	0.4316
Support Vector Regression					
RMSE	80.6154	80.1822	64.0426	109.0997	8.3402
MAE	61.9114	61.3858	48.4322	78.138	5.7176
R-squared	0.9614	0.9618	0.9282	0.9768	0.0081
MAPE	4.2511	4.1918	3.5139	5.4198	0.397

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

Table A2: *Models' performances, private healthcare expenditure (1994–2018)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	36.9831	36.4195	25.0963	54.0609	5.9887
MAE	25.7657	25.9219	18.7412	34.4748	2.8429
R-squared	0.9381	0.9408	0.8754	0.9747	0.0178
MAPE	5.3446	5.3688	3.9454	6.8857	0.538
Elastic Net					
RMSE	64.1264	63.5309	49.3741	78.4542	6.4832
MAE	48.2641	47.7847	38.3056	57.2414	3.6932
R-squared	0.8169	0.8176	0.7402	0.8735	0.0246
MAPE	10.8494	10.889	8.6499	12.6688	0.7913
Random Forest					
RMSE	36.4362	34.7011	25.836	59.2833	7.3935
MAE	24.6236	23.913	19.5839	32.1741	2.8124
R-squared	0.9399	0.9454	0.8617	0.9697	0.0213
MAPE	5.1636	5.2027	3.9366	6.2581	0.4901
Regression					
RMSE	64.8505	63.2436	49.748	79.8961	6.5512
MAE	49.0771	48.7772	38.9823	58.4861	3.897
R-squared	0.8125	0.8136	0.7341	0.8618	0.0271
MAPE	11.0723	11.1235	8.921	12.841	0.8238
Support Vector Regression					
RMSE	42.9922	42.2908	29.0999	59.5742	7.3009
MAE	27.7728	27.4256	21.8087	34.4935	2.9168
R-squared	0.9172	0.9214	0.8575	0.9543	0.0215
MAPE	5.7538	5.686	4.7794	6.8601	0.5223

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

Table A3: *Models' performances, total healthcare expenditure (1994–2018)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	78.7966	78.0233	54.7692	103.0542	9.1457
MAE	59.5427	59.7027	44.8861	74.3375	5.9083
R-squared	0.9787	0.9793	0.9637	0.9896	0.0049
Adj. R-squared	0.9645	0.9655	0.9395	0.9827	0.0081
MAPE	3.0714	3.0496	2.4227	3.8341	0.2859
SMAPE	3.0642	3.061	2.4191	3.7883	0.2846
MSLE	0.0016	0.0016	0.001	0.0025	0.0003
EVS	0.979	0.9799	0.9639	0.9896	0.0047
MedAE	47.9918	48.1244	34.8008	59.2323	5.5757
MSPE	0.0015	0.0015	0.0007	0.0025	0.0003
Quantile Loss	-15.3878	-15.0292	-25.7446	-7.2091	2.9107
Elastic Net					
RMSE	124.4908	124.989	98.4416	154.2244	11.3467
MAE	92.5567	92.5716	71.8705	112.0285	8.1383
R-squared	0.9471	0.9481	0.9254	0.9643	0.009
Adj. R-squared	0.9119	0.9135	0.8756	0.9405	0.0149
MAPE	4.7981	4.7758	4.018	5.7668	0.4097
SMAPE	4.8032	4.7783	3.9543	5.777	0.4071
MSLE	0.0041	0.0041	0.0029	0.0057	0.0006
EVS	0.948	0.9487	0.9255	0.9661	0.0089
MedAE	71.1891	70.6181	51.3324	102.2561	9.4982

Table A3 *continued from previous page*

	Mean	**Median**	**Min**	**Max**	**sd**
MSPE	0.0038	0.0037	0.0025	0.0059	0.0007
Quantile Loss	-23.4129	-23.2787	-39.9571	-14.5625	4.2571
Random Forest					
RMSE	87.0089	87.0555	66.5716	118.5396	11.0057
MAE	63.9104	64.1321	50.0988	79.1371	6.2705
R-squared	0.974	0.9744	0.9572	0.9849	0.0063
Adj. R-squared	0.9566	0.9573	0.9286	0.9749	0.0106
MAPE	3.3382	3.3654	2.67	3.9923	0.2948
SMAPE	3.3299	3.3478	2.6656	3.9504	0.2923
MSLE	0.002	0.0019	0.0012	0.0028	0.0003
EVS	0.9744	0.9748	0.9573	0.985	0.0061
MedAE	48.8505	48.5128	33.1229	65.5895	6.0094
MSPE	0.0019	0.0018	0.0011	0.0034	0.0005
Quantile Loss	-15.8092	-15.6283	-27.2887	-9.1122	3.0878
Regression					
RMSE	125.759	125.8533	100.8659	154.57	11.2939
MAE	93.7925	94.6845	76.4449	113.6491	8.4725
R-squared	0.9461	0.9472	0.9251	0.9645	0.009
Adj. R-squared	0.9101	0.9119	0.8751	0.9409	0.015
MAPE	4.858	4.8059	4.0485	6.0326	0.4503
SMAPE	4.8644	4.8131	4.0733	6.0103	0.4446
MSLE	0.0042	0.0041	0.003	0.0058	0.0007
EVS	0.947	0.9482	0.927	0.9647	0.0089
MedAE	72.0085	71.9345	50.9393	100.4555	10.244
MSPE	0.0039	0.0038	0.0025	0.0059	0.0007
Quantile Loss	-23.7187	-23.3856	-40.789	-13.833	4.4985
Support Vector Regression					
RMSE	96.0728	95.4064	74.6998	119.3491	8.9812
MAE	72.5331	71.9648	54.4561	91.2275	6.6547
R-squared	0.9685	0.9688	0.9545	0.9798	0.0057
Adj. R-squared	0.9474	0.948	0.9242	0.9664	0.0095
MAPE	3.8068	3.7936	2.9572	4.8582	0.3583
SMAPE	3.7797	3.7724	2.9162	4.7193	0.3439
MSLE	0.0025	0.0025	0.0016	0.0053	0.0005
EVS	0.9691	0.9695	0.9565	0.98	0.0056
MedAE	56.0487	56.1183	35.9954	75.3596	8.4558
MSPE	0.0023	0.0022	0.0014	0.0035	0.0004
Quantile Loss	-17.9648	-17.3945	-29.3528	-10.8647	3.8617

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error; SMAPE stands for Symmetric Mean Absolute Percentage Error; MSLE stands for Mean Squared Logarithmic Error; EVS stands for Explained Variance Score; MedAE stands for Median Absolute Error; MSPE stands for Mean Squared Percentage Error.

Table A4: *Models' performances, public healthcare expenditure (1994–2018)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	58.6057	57.3748	48.2828	75.6431	6.0012
MAE	45.1327	44.356	37.5431	56.0848	4.0923
R-squared	0.9796	0.9805	0.9656	0.9867	0.0042
Adj. R-squared	0.966	0.9674	0.9427	0.9779	0.007
MAPE	3.0834	3.0431	2.5038	3.7648	0.2706
SMAPE	3.0813	3.0386	2.4999	3.778	0.2697
MSLE	0.0016	0.0015	0.001	0.0025	0.0003
EVS	0.9799	0.9809	0.9656	0.9871	0.0041
MedAE	37.0321	36.7108	25.1026	47.5002	4.6436

Table A4 *continued from previous page*

	Mean	Median	Min	Max	sd
MSPE	0.0014	0.0014	0.001	0.0023	0.0003
Quantile Loss	-11.1141	-11.1413	-17.0707	-6.61	1.9654
Elastic Net					
RMSE	102.1749	102.1495	80.99	125.7477	9.4094
MAE	80.0897	80.066	66.7746	99.98	6.7484
R-squared	0.9382	0.9391	0.9046	0.9605	0.0115
Adj. R-squared	0.8969	0.8985	0.841	0.9341	0.0191
MAPE	5.4409	5.4206	4.5096	6.8787	0.4282
SMAPE	5.4408	5.4347	4.479	6.9341	0.447
MSLE	0.0048	0.0047	0.0031	0.0069	0.0008
EVS	0.9393	0.9401	0.9054	0.9618	0.0109
MedAE	67.201	67.4847	52.8599	95.2216	7.2671
MSPE	0.0044	0.0043	0.0027	0.0065	0.0008
Quantile Loss	-19.8041	-20.18	-27.9275	-11.9006	3.2659
Random Forest					
RMSE	64.604	64.7502	51.4161	80.7652	6.6465
MAE	48.5669	49.1518	39.1583	60.3321	4.5321
R-squared	0.9753	0.9753	0.961	0.9849	0.0049
Adj. R-squared	0.9588	0.9588	0.9349	0.9748	0.0081
MAPE	3.3468	3.3919	2.626	4.2553	0.3141
SMAPE	3.342	3.3875	2.6201	4.2422	0.3122
MSLE	0.002	0.002	0.0012	0.0032	0.0004
EVS	0.9756	0.9756	0.961	0.9856	0.0048
MedAE	36.9185	37.4718	24.8354	52.0894	4.9449
MSPE	0.0018	0.0017	0.0011	0.0027	0.0004
Quantile Loss	-11.6188	-11.6321	-17.4506	-7.4537	1.9596
Regression					
RMSE	102.7652	103.1196	81.2362	127.9384	9.5852
MAE	80.3164	80.0339	67.1147	99.8419	6.6513
R-squared	0.9374	0.9389	0.9012	0.9578	0.0118
Adj. R-squared	0.8957	0.8982	0.8354	0.9297	0.0196
MAPE	5.4378	5.4048	4.5115	6.8774	0.4316
SMAPE	5.4412	5.4002	4.4791	6.915	0.4545
MSLE	0.0049	0.0047	0.0033	0.0072	0.0009
EVS	0.9386	0.9399	0.9042	0.9585	0.0112
MedAE	66.4918	65.5517	52.6529	92.2232	7.196
MSPE	0.0044	0.0044	0.0027	0.0067	0.0008
Quantile Loss	-19.8378	-20.1063	-27.7663	-11.8603	3.3259
Support Vector Regression					
RMSE	80.6154	80.1822	64.0426	109.0997	8.3402
MAE	61.9114	61.3858	48.4322	78.138	5.7176
R-squared	0.9614	0.9618	0.9282	0.9768	0.0081
Adj. R-squared	0.9357	0.9363	0.8803	0.9613	0.0134
MAPE	4.2511	4.1918	3.5139	5.4198	0.397
SMAPE	4.2317	4.1659	3.505	5.4054	0.3883
MSLE	0.0031	0.003	0.002	0.0061	0.0007
EVS	0.962	0.9623	0.935	0.977	0.0077
MedAE	49.0306	49.4416	37.3775	62.8556	5.8741
MSPE	0.0027	0.0027	0.0017	0.0049	0.0006
Quantile Loss	-14.6012	-14.4663	-22.5352	-9.3288	2.4021

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error; SMAPE stands for Symmetric Mean Absolute Percentage Error; MSLE stands for Mean Squared Logarithmic Error; EVS stands for Explained Variance Score; MedAE stands for Median Absolute Error; MSPE stands for Mean Squared Percentage Error.

Table A5: *Models' performances, private healthcare expenditure (1994–2018)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	36.9831	36.4195	25.0963	54.0609	5.9887
MAE	25.7657	25.9219	18.7412	34.4748	2.8429
R-squared	0.9381	0.9408	0.8754	0.9747	0.0178
Adj. R-squared	0.9077	0.9116	0.814	0.9622	0.0266
MAPE	5.3446	5.3688	3.9454	6.8857	0.538
SMAPE	5.3055	5.3164	3.9805	6.8281	0.5281
MSLE	0.0047	0.0046	0.0026	0.0079	0.001
EVS	0.9389	0.9416	0.8772	0.9752	0.0177
MedAE	18.9154	18.9195	13.3468	27.3533	2.6208
MSPE	0.0063	0.0059	0.0027	0.0127	0.002
Quantile Loss	-6.5053	-6.637	-8.9997	-3.9226	1.1514
Elastic Net					
RMSE	64.1264	63.5309	49.3741	78.4542	6.4832
MAE	48.2641	47.7847	38.3056	57.2414	3.6932
R-squared	0.8169	0.8176	0.7402	0.8735	0.0246
Adj. R-squared	0.7268	0.7279	0.6123	0.8113	0.0367
MAPE	10.8494	10.889	8.6499	12.6688	0.7913
SMAPE	10.7441	10.79	8.6609	12.2541	0.7251
MSLE	0.0192	0.0193	0.0131	0.0254	0.0027
EVS	0.8191	0.8195	0.7402	0.8745	0.0243
MedAE	40.3981	39.9648	31.7265	52.894	4.3057
MSPE	0.0185	0.018	0.011	0.027	0.0035
Quantile Loss	-11.9981	-12.1285	-16.9599	-7.4662	1.9263
Random Forest					
RMSE	36.4362	34.7011	25.836	59.2833	7.3935
MAE	24.6236	23.913	19.5839	32.1741	2.8124
R-squared	0.9399	0.9454	0.8617	0.9697	0.0213
Adj. R-squared	0.9103	0.9185	0.7937	0.9548	0.0317
MAPE	5.1636	5.2027	3.9366	6.2581	0.4901
SMAPE	5.1195	5.1478	3.9237	6.2692	0.4804
MSLE	0.0047	0.0045	0.0026	0.0085	0.0011
EVS	0.9406	0.946	0.862	0.9698	0.0209
MedAE	17.9406	17.7012	13.3726	23.4697	2.2829
MSPE	0.0061	0.0055	0.003	0.0155	0.0026
Quantile Loss	-5.9503	-5.8834	-8.6259	-3.7104	1.0421
Regression					
RMSE	64.8505	63.2436	49.748	79.8961	6.5512
MAE	49.0771	48.7772	38.9823	58.4861	3.897
R-squared	0.8125	0.8136	0.7341	0.8618	0.0271
Adj. R-squared	0.7202	0.7219	0.6032	0.7938	0.0404
MAPE	11.0723	11.1235	8.921	12.841	0.8238
SMAPE	10.9437	10.9432	8.9412	12.5766	0.7529
MSLE	0.0199	0.0201	0.0135	0.0268	0.0029
EVS	0.8146	0.8149	0.7344	0.8625	0.027
MedAE	40.7022	40.5147	30.4621	51.0565	4.512
MSPE	0.0189	0.0182	0.0115	0.0283	0.0036
Quantile Loss	-12.183	-12.3363	-17.0005	-7.7682	1.9686
Support Vector Regression					
RMSE	42.9922	42.2908	29.0999	59.5742	7.3009
MAE	27.7728	27.4256	21.8087	34.4935	2.9168
R-squared	0.9172	0.9214	0.8575	0.9543	0.0215
Adj. R-squared	0.8765	0.8827	0.7873	0.9318	0.0321
MAPE	5.7538	5.686	4.7794	6.8601	0.5223

Table A5 *continued from previous page*

	Mean	**Median**	**Min**	**Max**	**sd**
SMAPE	5.711	5.6095	4.7356	6.9073	0.5083
MSLE	0.0061	0.006	0.0037	0.0092	0.0013
EVS	0.9183	0.9224	0.8628	0.9548	0.021
MedAE	18.4908	18.2716	12.6169	26.289	2.0264
MSPE	0.0085	0.008	0.0039	0.0155	0.0028
Quantile Loss	-6.5136	-6.4028	-10.0495	-4.003	1.1132

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error; SMAPE stands for Symmetric Mean Absolute Percentage Error; MSLE stands for Mean Squared Logarithmic Error; EVS stands for Explained Variance Score; MedAE stands for Median Absolute Error; MSPE stands for Mean Squared Percentage Error.

Table A6: *Models' performances, public healthcare expenditure (1994–2018: 2019)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	87.4315	88.3363	74.565	100.8289	5.6038
MAE	69.8031	70.3689	57.5245	82.6375	5.3673
R-squared	0.4436	0.4343	0.263	0.5969	0.0708
MAPE	3.5099	3.5419	2.9189	4.1624	0.2698
Elastic Net					
RMSE	123.7064	123.9915	114.4721	134.2849	3.6014
MAE	96.6833	96.7835	84.6546	110.3777	4.3397
R-squared	-0.1103	-0.1145	-0.3073	0.05	0.0646
MAPE	4.8426	4.8485	4.2297	5.5432	0.2239
Random Forest					
RMSE	88.2685	87.5609	82.8846	103.565	3.1588
MAE	68.8333	68.2877	61.5934	80.1812	3.3469
R-squared	0.4345	0.4442	0.2224	0.502	0.0415
MAPE	3.4365	3.4137	3.0708	3.987	0.1671
Regression					
RMSE	126.0889	125.6404	117.8714	140.8783	4.3671
MAE	100.3634	99.9615	88.1518	117.3565	5.3101
R-squared	-0.1539	-0.1444	-0.4388	-0.0072	0.0809
MAPE	5.0337	5.0166	4.4044	5.9016	0.2738
Support Vector Regression					
RMSE	112.5763	112.4006	95.9891	134.1641	7.4531
MAE	91.7427	91.4201	73.7906	107.9866	6.2983
R-squared	0.0773	0.0841	-0.3049	0.332	0.1232
MAPE	4.5721	4.5605	3.6867	5.37	0.3102

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

Table A7: *Models' performances, private healthcare expenditure (1994–2018: 2019)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	44.2132	44.1414	31.9526	62.326	5.6744
MAE	34.613	34.6863	26.8819	43.9761	3.9796
R-squared	0.9186	0.9202	0.8409	0.9582	0.0212
MAPE	5.1621	5.1918	4.0102	6.4666	0.5963
Elastic Net					
RMSE	76.3463	75.8481	73.2833	82.0406	2.0137
MAE	60.0352	59.6345	56.0788	65.6361	2.0499
R-squared	0.7611	0.7643	0.7243	0.78	0.0127
MAPE	9.1563	9.1251	8.1095	10.2582	0.4107
Random Forest					
RMSE	50.3613	49.0449	40.0275	73.6319	6.1005
MAE	34.854	34.65	28.7374	45.938	3.0707
R-squared	0.8946	0.9015	0.7779	0.9344	0.0268
MAPE	4.8038	4.7772	3.9992	5.9772	0.3705
Regression					
RMSE	75.37	74.6188	69.8358	86.675	3.3435
MAE	61.5848	61.2222	56.2628	69.4527	2.9884
R-squared	0.7668	0.7719	0.6923	0.8002	0.0211
MAPE	9.6447	9.6223	8.526	10.7916	0.5083
Support Vector Regression					
RMSE	59.3753	58.2366	44.675	82.8933	6.5901
MAE	37.0949	36.8342	30.2999	47.4426	2.8232
R-squared	0.8538	0.8611	0.7185	0.9182	0.0334
MAPE	4.9742	4.9337	3.9998	6.0254	0.3264

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

Table A8: *Models' performances, public healthcare expenditure (2005–2018)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	63.5926	62.4432	45.4191	100.9656	10.0809
MAE	47.2559	47.1605	32.6739	65.6084	6.4148
R-squared	0.828	0.8365	0.608	0.9081	0.051
MAPE	2.5816	2.5698	1.7679	3.6144	0.351
Elastic Net					
RMSE	85.2777	84.16	63.3376	113.3749	9.7085
MAE	65.7846	64.9616	50.1899	85.5217	7.5279
R-squared	0.6914	0.6951	0.4622	0.8261	0.069
MAPE	3.5606	3.5244	2.7432	4.7156	0.3965
Random Forest					
RMSE	63.4852	61.9355	46.3064	93.9296	9.61
MAE	47.5113	47.0585	32.8142	65.9996	6.4157
R-squared	0.8297	0.8353	0.6698	0.9012	0.0434
MAPE	2.5935	2.5912	1.8426	3.6785	0.3536
Regression					
RMSE	88.3067	86.8577	68.2786	115.3653	9.8298
MAE	68.5343	67.9359	54.286	86.3209	7.5363
R-squared	0.6688	0.6786	0.4288	0.8087	0.075
MAPE	3.7224	3.6908	2.9491	4.6788	0.3947
Support Vector Regression					
RMSE	75.9666	75.6814	56.0179	98.1622	8.9977
MAE	56.2229	56.4391	41.1511	72.0256	6.2391
R-squared	0.7573	0.7587	0.638	0.8555	0.0442
MAPE	3.0459	3.0682	2.2768	3.9125	0.3327

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

Table A9: *Models' performances, private healthcare expenditure (2005–2018)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	51.2041	51.0099	40.455	62.9837	4.4672
MAE	39.1638	39.1821	31.3318	48.2049	3.2162
R-squared	0.9638	0.9644	0.9474	0.9767	0.0061
MAPE	4.3833	4.3834	3.2898	5.6291	0.3965
Elastic Net					
RMSE	74.4725	74.4341	61.2209	93.1745	4.993
MAE	57.5752	57.2082	46.925	72.188	4.1556
R-squared	0.9236	0.9247	0.8795	0.9453	0.0097
MAPE	6.4365	6.4052	4.9261	7.8592	0.477
Random Forest					
RMSE	54.0674	53.9444	43.0519	66.0616	4.9532
MAE	41.387	41.1869	33.8161	52.1289	3.3651
R-squared	0.9597	0.9602	0.9411	0.9711	0.0066
MAPE	4.671	4.6143	3.7943	5.9685	0.4043
Regression					
RMSE	74.4827	74.9286	60.6306	91.143	5.0986
MAE	57.4855	57.3861	46.5148	70.9605	4.1922
R-squared	0.9236	0.9237	0.8847	0.9456	0.0099
MAPE	6.428	6.4066	4.9116	7.9167	0.4891
Support Vector Regression					
RMSE	51.6146	51.7423	38.9	65.4823	4.472
MAE	39.7791	39.9712	29.6835	48.6715	3.3646
R-squared	0.9632	0.9634	0.9466	0.9784	0.006
MAPE	4.457	4.4391	3.3686	6.1628	0.4151

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

Table A10: *Models' performances, public healthcare expenditure (2005–2018: 2019)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	88.553	88.2326	75.6892	104.3385	4.8061
MAE	67.0553	67.0674	56.4612	80.3154	4.5934
R-squared	0.4299	0.4356	0.2108	0.5847	0.0621
MAPE	3.3495	3.3547	2.8249	4.0249	0.2355
Elastic Net					
RMSE	93.1126	91.7604	81.1054	114.1178	6.2522
MAE	64.2898	63.7883	54.016	83.7516	6.0098
R-squared	0.3687	0.3896	0.0559	0.5231	0.0868
MAPE	3.1875	3.1626	2.6723	4.1594	0.3019
Random Forest					
RMSE	87.0645	86.8649	80.3363	95.095	2.8818
MAE	63.5136	62.709	55.2204	73.2173	3.8448
R-squared	0.4499	0.453	0.3444	0.5321	0.0366
MAPE	3.1487	3.1102	2.7438	3.6367	0.1913
Regression					
RMSE	110.9925	109.3662	94.4336	134.7538	8.4827
MAE	78.4187	76.9273	60.2193	109.8017	9.4349
R-squared	0.1018	0.1329	-0.3164	0.3535	0.1386
MAPE	3.8903	3.8092	2.9537	5.4703	0.4801
Support Vector Regression					
RMSE	95.7054	95.7177	85.3336	107.2569	4.183
MAE	75.2698	74.4886	64.7141	85.6718	4.4582
R-squared	0.3347	0.3358	0.166	0.4721	0.0582
MAPE	3.7431	3.7084	3.245	4.2544	0.2201

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

Table A11: *Models' performances, private healthcare expenditure (2005–2018: 2019)*

	Mean	**Median**	**Min**	**Max**	**sd**
Gradient Boosting					
RMSE	63.3012	62.911	47.6255	89.1998	6.6672
MAE	52.7627	51.5399	39.1321	75.0273	6.0914
R-squared	0.8248	0.8288	0.6559	0.9019	0.0375
MAPE	4.4272	4.3167	3.3152	6.2785	0.5078
Elastic Net					
RMSE	82.1062	81.8562	75.4778	89.4701	2.7163
MAE	61.7421	61.6988	56.1471	70.6133	2.7252
R-squared	0.7081	0.7102	0.6538	0.7536	0.0194
MAPE	4.9904	4.98	4.4985	5.7643	0.226
Random Forest					
RMSE	69.0496	69.1818	59.3547	79.6605	4.1452
MAE	58.1751	57.924	49.6023	69.538	3.7431
R-squared	0.7931	0.793	0.7256	0.8476	0.0249
MAPE	4.6884	4.6632	4.0153	5.6015	0.3078
Regression					
RMSE	86.7672	86.2073	79.9779	96.788	3.5903
MAE	69.0408	68.3509	61.8792	78.8769	3.5257
R-squared	0.6739	0.6786	0.5949	0.7234	0.0273
MAPE	5.5721	5.5259	4.9094	6.3439	0.2849
Support Vector Regression					
RMSE	85.8566	85.473	74.8113	102.3791	6.0997
MAE	71.8695	71.5487	59.5186	87.1739	5.9214
R-squared	0.6796	0.6841	0.5467	0.758	0.0461
MAPE	5.7964	5.7673	4.8401	7.0633	0.4722

Notes: RMSE stands for Root Square Mean Error; MAE stands for Mean Absolute Error; MAPE stands for Mean Absolute Percentage Error.

Figure A1: *Feature importance, public healthcare expenditure (1994–2018)*

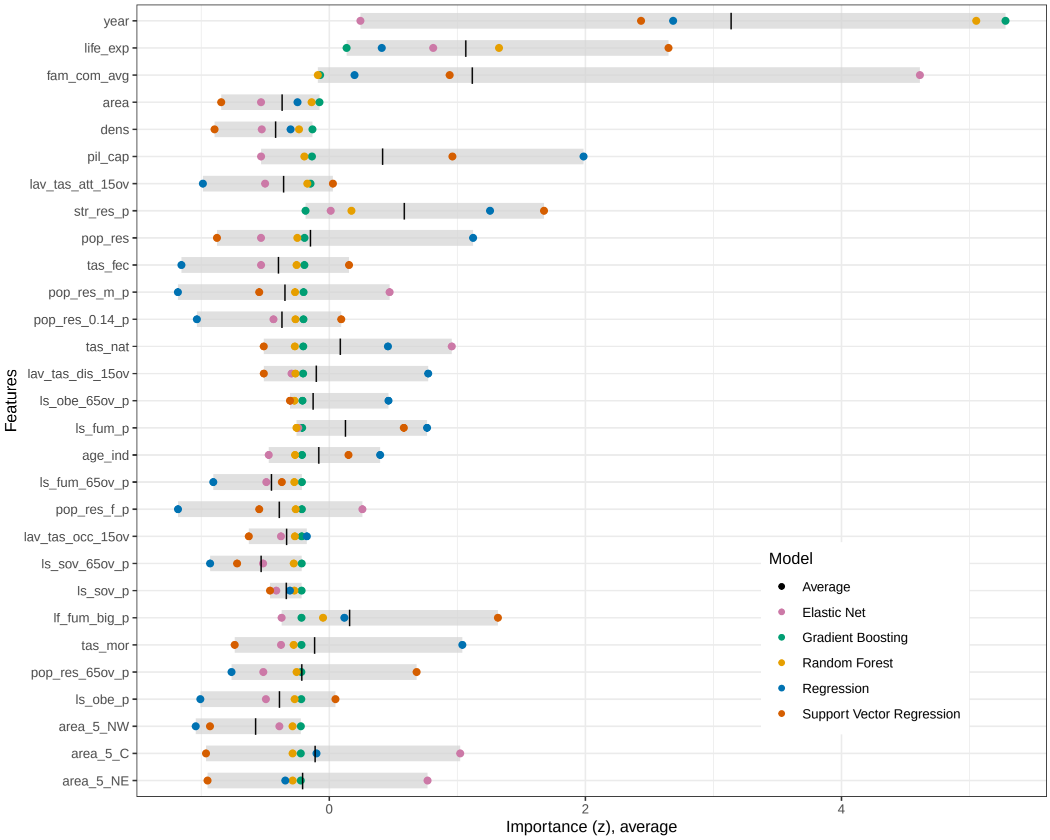


Figure A2: *Feature importance, public healthcare expenditure (2005–2018)*

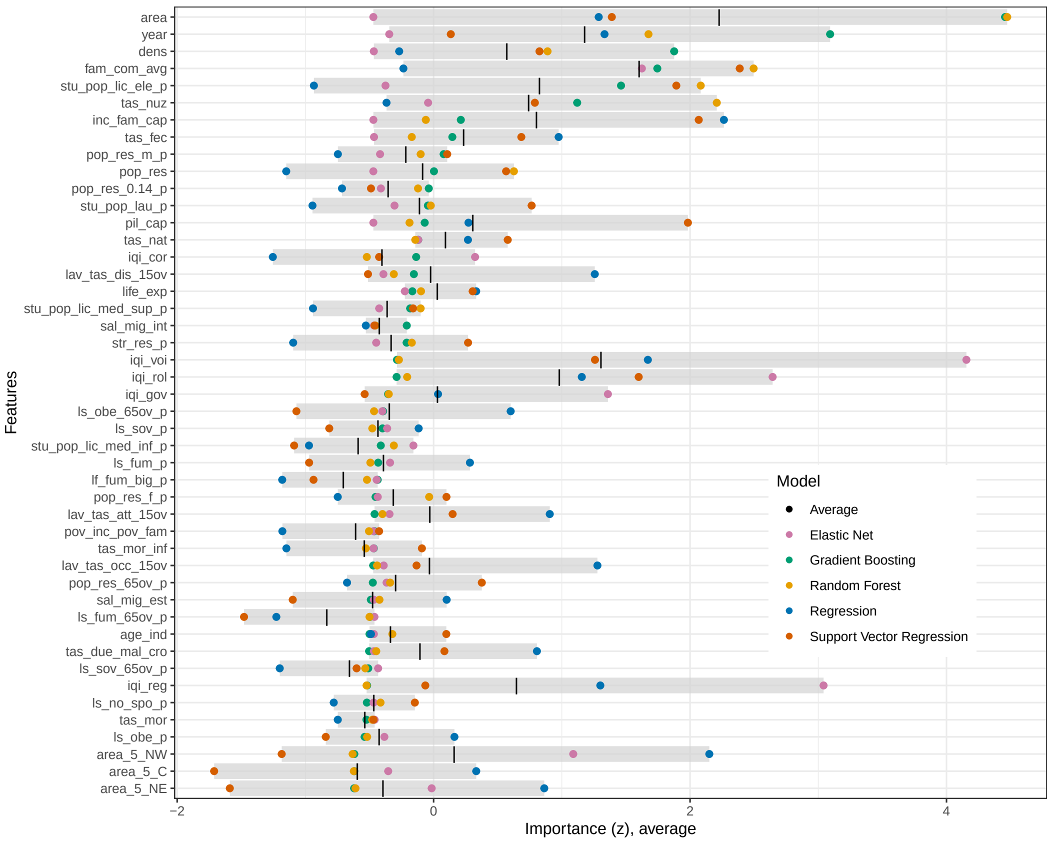


Figure A3: *Feature importance, private healthcare expenditure (1994–2018)*

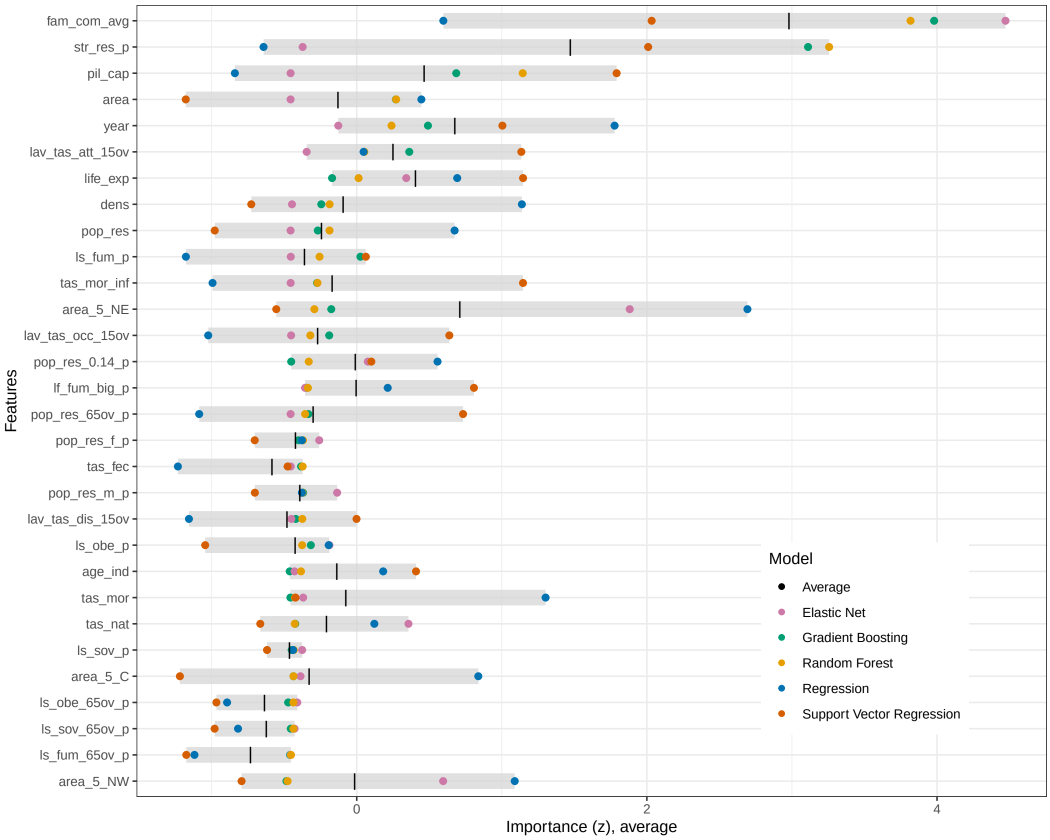


Figure A4: *Feature importance, private healthcare expenditure (2005–2018)*

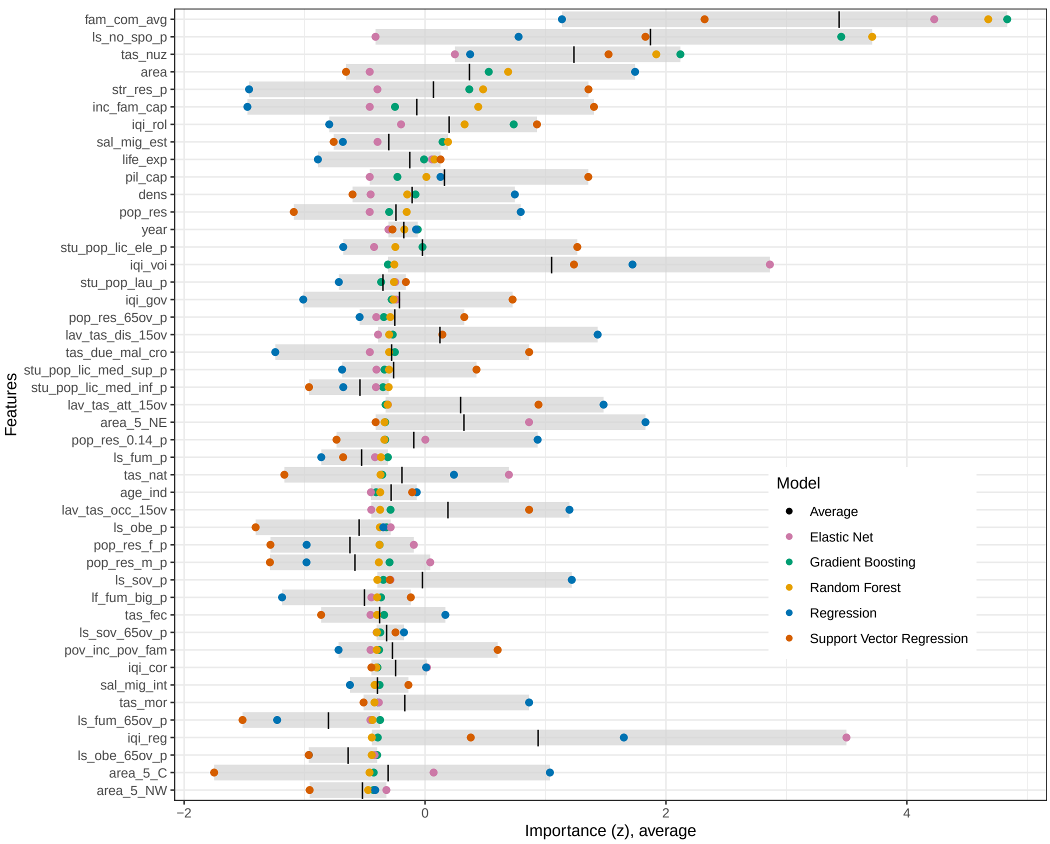
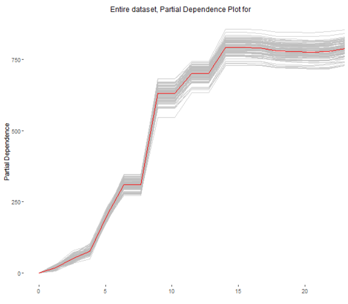
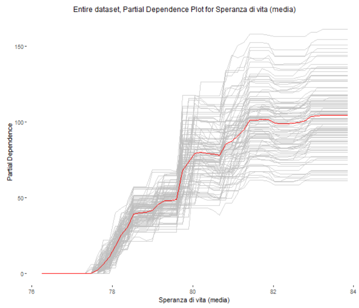


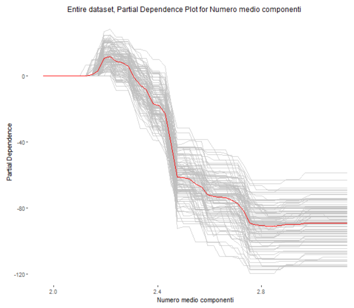
Figure A5: *Partial Dependence Plots for selected features, public healthcare expenditure (1994–2018)*



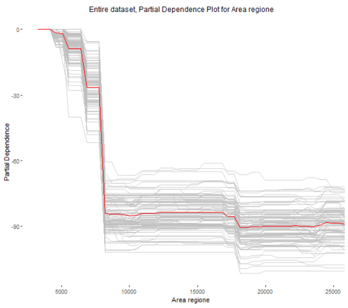
(a) Years



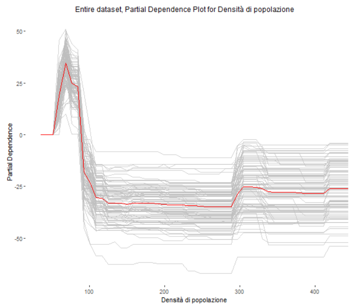
(b) Life expectancy



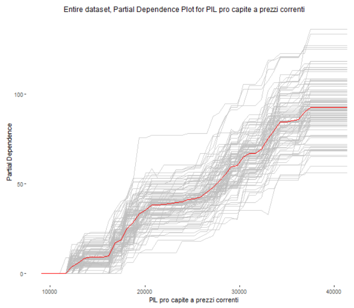
(c) Average family components



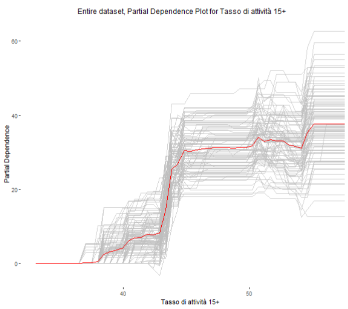
(d) Regional area



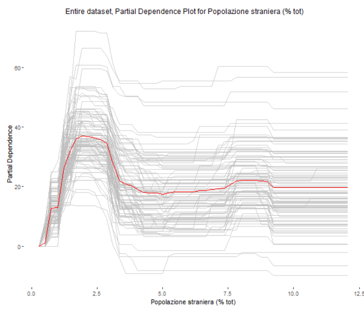
(e) Population density



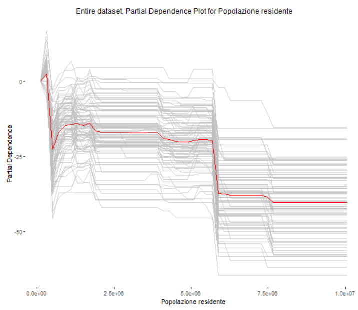
(f) GDP per capita



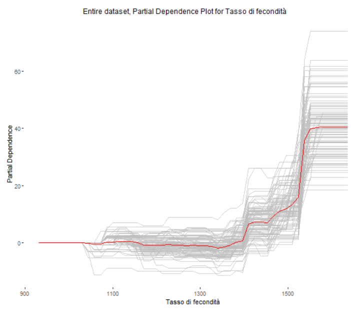
(g) Labor activity rate



(h) Share of foreign residents

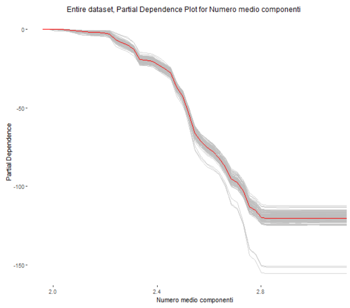


(i) Resident population

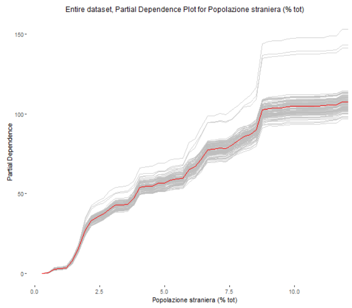


(j) Fertility rate (total)

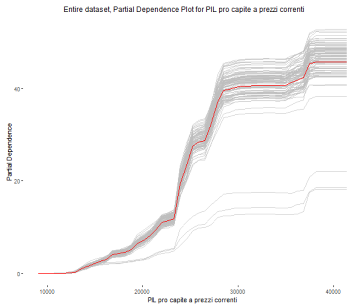
Figure A6: *Partial Dependence Plots for selected features, private healthcare expenditure (1994–2018)*



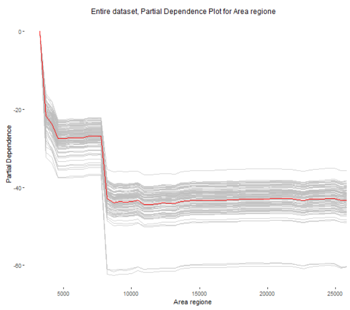
(a) Average family components



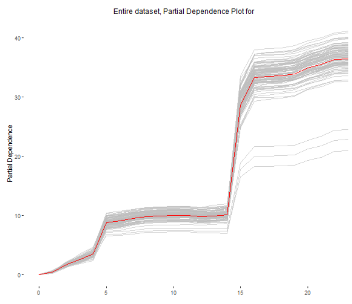
(b) Share of foreign residents



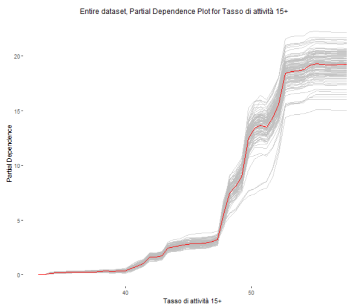
(c) GDP per capita



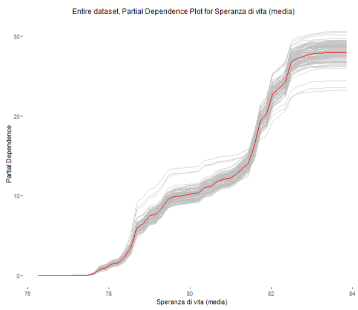
(d) Regional area



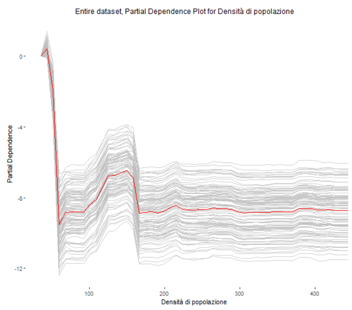
(e) Years



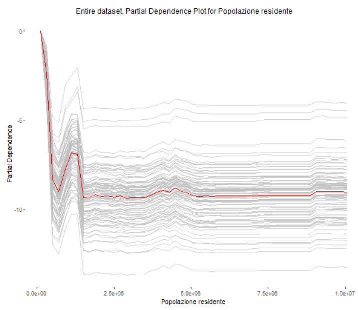
(f) Labor activity rate



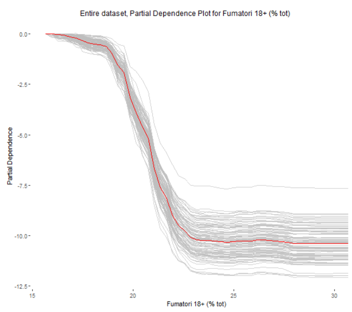
(g) Life expectancy



(h) Population density



(i) Resident population



(j) Share of smokers