Does prescriptive appropriateness reduce health expenditure? Main effects and some unintended outcomes

Claudio Lucifora*[1], Antonio Russo†[2], and Daria Vigani‡[3]

[1]ATS Milano
[2]Università Cattolica di Milano and IZA
[3]Università Cattolica di Milano

February 28th, 2018

Abstract

Using register data from the Agency for Health Protection of Milan, we exploit the introduction of the "appropriateness" decree to evaluate the impact of prescriptive constraints on GPs' behavior with respect to expenditure and volumes of outpatient visits, exploring heterogeneity across family doctors and unintended consequences on patients' outcomes. Overall, we find a remarkable reduction in both outpatients' spending and volumes after the introduction of the decree, suggesting that the policy was effective in containing health costs. However, the behavioral response of GPs is heterogeneous and some unintended effects are unveiled: outpatients spending was reduced also on vulnerable patients and larger reductions in spending were associated with increasing resort to emergency care.

JEL classification: H51 I38 I18

Keywords: Appropriateness, Prescriptive constraints, Health expenditure, Impact evaluation

*Corresponding author: Department of Economics and Finance, Università Cattolica, Largo Gemelli, 1 (20123) Milan. e-mail: claudio.lucifora@unicatt.it

†ATS Città Metropolitana di Milano, Corso Italia, 19 (20122) Milan
e-mail: agrusso@ats-milano.it

‡Department of Economics and Finance, Università Cattolica, Largo Gemelli, 1 (20123) Milan.
e-mail: daria.vigani@unicatt.it

1 Introduction

For over twenty years health care spending has increased dramatically in most industrialized countries. From 1980 to 2009 the annual growth of per capita health expenditure in real terms has been around 3% on average across OECD countries (OECD, 2005-2017), with an upsurge in public expenditure in the early 2000s that outpaced GDP growth in all countries. Moreover, as the aging of populations has proven to be a long-lasting trend characterizing most countries, raising serious concerns about its financial sustainability, the quality and cost of health care provision have come under increasing scrutiny. In response to increasing health expenditures and overconsumption of medical services, several cost-containment policies have been implemented over the years in a number of countries, according with the three main dimensions that define health care systems: financing, service provision, and regulation.

Overall, cost-sharing and deductibles have been identified as one main tool to reduce moral hazard in medical consumption and excess spending (Pauly and Blavin, 2008; van Kleef et al., 2009). Empirical evidence on the effects on the demand for medical care generally suggests that cost-sharing at the point of consumption does lead to a reduction in health care utilization (Baicker and Goldman, 2011; Goodell and Swartz, 2010; Ziebarth, 2014), even though with significant heterogeneity across type of plans (Schreyögg and Grabka, 2010) and across patients with different price-sensitivity. On top of cost-sharing options, health systems that share the physician 'gatekeeper' feature are provided with additional means that can be used for cost-containment purposes. The gatekeeper simultaneously acts as a clinical expert that steers patients to proper services and as a rationing agent, thus resulting as a valuable resource in cutting relatively ineffective or unnecessary health care services. Pay-for-performance programs and financial incentive schemes on gatekeeping doctors have been designed worldwide, but empirical evidence on the effect of different plans on quality and costs of health care is rather controversial (Dusheiko et al., 2006; Fiorentini et al., 2013).

Regardless of the chosen strategy, a central concern for policy-makers in most publicly-funded health systems is to achieve budgetary goals while retaining the necessary and desirable demand of vulnerable groups. As a matter of fact, a prominent argument against all the above-mentioned cost-containment measures is that they might induce specific groups of fragile patients to avoid seeking necessary care, with potentially larger future costs.

In this context, policies aimed at both containing health care spending and improving the

efficiency of resource allocation are guided by the concept of *appropriateness*, being defined principally in terms of maximizing efficiency and minimizing costs. According to this concept, that became established in the '90s when the *new public management* organizational model started spreading, an appropriate service is the one that is delivered according to clinical indications with proven efficacy, with precise timing and within a technically correct and least expensive setting (Lavis and Anderson, 1996).

In Italy, appropriateness of health care is gaining increasing interest among both providers and policy-makers, as a preeminent issue in the lively health care debate is that a substantial proportion of provided care is deemed to be inappropriate. For this reason, on December 9, 2015 the Italian Ministry of Health presented the "appropriateness" decree, that established appropriateness requirements and prescriptive constraints on a selected number of health services. In particular, the decree restrained the ability of General Practitioners (GPs) to prescribe 203 diagnostic tests unless the patients meet some predetermined criteria, thus manipulating their discretionary power as gatekeepers for secondary care. Although exemption rules based on income and chronic disease status were defined to avoid a deterrent effect on vulnerable subjects, this action was preceded and followed by an active dispute, mainly centered on possible adverse effects on health outcomes, that culminated with the successive repealing of the decree in mid-July 2016.

In this paper we exploit this policy shift within a difference-in-difference framework to evaluate the effectiveness of prescriptive constraints in reducing health care utilization and spending. At the same time, we explore whether the introduction of prescriptive appropriateness requirements had some unintended consequences on the demand of medical services of vulnerable groups or on patients' resort to emergency care and health outcomes. Moreover, as the decree was designed to tighten the gatekeeper role of family doctors, by limiting their discretionary power over prescriptive behavior, we are further interested in analyzing whether GPs' response to the introduction of appropriateness requirements has been heterogeneous according with several individual characteristics.

To the best of our knowledge, our work is the first to provide causal evidence on the effects of the "appropriateness" decree on health care utilization and spending. In particular, drawing on administrative health archives of the largest Italian Local Health Authority, that match the *Outpatients Record* with the *Regional Health Roster*, the *Hospital Discharge Record* and the

Emergency Care Access Record through a unique anonymous personal identifier, we are able not only to assess the overall impact of the decree on spending and volumes of outpatient visits induced by general practitioners, but also to analyze heterogeneity in GPs' prescriptive behavior as well as the relationship between the latter and patients' resort to emergency care or hospital admissions.

The rest of the paper is structured as follows. In Section 2 we provide a brief overview of the Italian health care system and a description of the institutional setting. Section 3 describes the data and the empirical strategy used, and provides some descriptive evidence. Section 4 describes the results and Section 5 concludes.

2 Institutional Background

2.1 The Italian health care system in brief

The Italian health care system is a regionally based National Health Service (*Sistema sanitario nazionale* - SSN): the central government is responsible for general legislation and financing, while regional governments are in charge of the organization, management and provision of health care. The SSN provides universal coverage to all citizens and residents largely free of charge and it is funded through national and regional taxes, supplemented by co-payments for pharmaceuticals and outpatient care. As a matter of fact, while there are no user charges for family doctors' consultations and hospitalizations, patients pay a co-payment for procedures, specialist visits and pharmaceuticals up to a ceiling. A 25 euro co-payment also exists for the "unwarranted" use of emergency care, defined as non-critical or non-urgent conditions.

Each individual is assigned to a general practitioner (pediatrician for children under 14) who provides family medicine services free of charge and acts as a gatekeeper to higher levels of care and pharmaceuticals. Then, for the latter services, the level of cost-sharing ranges from total exemption (i.e. free access) to a coverage of part of the costs. Total exemptions are applied to people aged 65 and over, children below 6, unemployed individuals or those with a gross family income below a given threshold and patients with severe disabilities. Moreover, exemptions also apply to individuals with chronic or rare diseases, HIV-positive and to pregnant women as far as the needed treatments are related with their condition. Finally, all individuals with out-of-pocket payments above a set amount (currently 129 euros) in a given year are eligible for a tax credit equal to roughly one-fifth of their spending.

2.2 The D.M. *Lorenzin* 9^{th} December, 2015

On December 9, 2015 the Italian Ministry of Health issued the "appropriateness" decree (so-called D.M. *Lorenzin*, from the name of the Minister of Health), that entered into force on January 21, 2016. The decree introduced appropriateness requirements and prescriptive constraints for 203 health services and diagnostic tests, with the aim of putting a filter to the increasing overconsumption of medical services driven by aging and health awareness of the population, as well as limiting the so-called 'defensive medicine' (i.e. over-prescription of exams and medications as a safeguard from malpractice litigation). The latter phenomenon is in fact plaguing many health care systems, with an estimated cost about 10 billion euros a year for Italy (around 0,75% of GDP and 10% of total NHS expenditure[1]).

In practice the decree introduced a constraint on the ability of physicians and general practitioners to prescribe a selected number of diagnostic tests - such as laboratory analyses, diagnostic imaging and genetic exams[2] -, unless predetermined conditions are met. For example, Magnetic Resonance Imaging (MRI) ought to be restricted to cancer patients or those suffering traumas, while total cholesterol tests should only be prescribed to individuals over 40 or with certified chronic illnesses. In other words, following the guidelines of the D.M., any doctor could reject the prescription of an exam for lack of "appropriateness": either because the patient is below the critical age, or because too little time has elapsed since the last check or because, at doctor's discretion, it is not deemed to be appropriate. Conversely, when the appropriateness criteria are satisfied, the doctor can prescribe the test, which is then provided by the NHS (free of charge or with the "ticket" co-payment). Finally, penalties were planned for doctors who repeatedly refuse to comply with the decree, by prescribing inappropriate medical services to their patients. This way, the gate-keeping role of general practitioners is tightened and an additional filter to the growing medicalization is introduced.

The approval of the decree followed a long dispute with associations representing physicians and general practitioners. The arguments for opposing the implementation by the medical associations were that it would compromise the relationship between patients and doctors, demean the role of health professionals and, more importantly, expose patients' health at risk for not receiving adequate treatments and therapies. Also, the new rules were expected to introduce ambiguities in the allocation of responsibilities and discretion, as far as the limitations on deliv-

[1] Source: Defensive Medicine Report from the Ministry of Health and Panella et al. (2017).

[2] The 203 health services belong to 7 specialties: laboratory analyses, diagnostic imaging, dentistry and maxillofacial surgery, physical therapy and rehabilitation, dermatology, nuclear medicine and genetic tests.

erability are concerned. Doctors also opposed the decree saying that it would transfer national health responsibilities to the private sector and increase patients' costs for treatments. It was questioned whether adherence to appropriateness criteria in doctors' prescriptions would not increase resort to emergencies and patients' hospitalization rates.

After a long dispute, in mid-July 2016, the decree was repealed by a new decree concerning standard/minimum health care provision (*D.P.C.M. LEA - Livelli Essenziali di Assistenza*), which reduced the number of health services subjected to appropriateness criteria to 40 (mainly radiotherapy, some genetic test and a few more tests) and softened the restrictions for doctors.

3 Data and Methods

3.1 Data and descriptive statistics

We use register data from the Health Information System of the Agency for Health Protection of the Province of Milan (and Lodi), a collection of databases that store information on the universe of health care services for the whole population of the largest metropolitan area of northern Italy (3.2 million inhabitants).

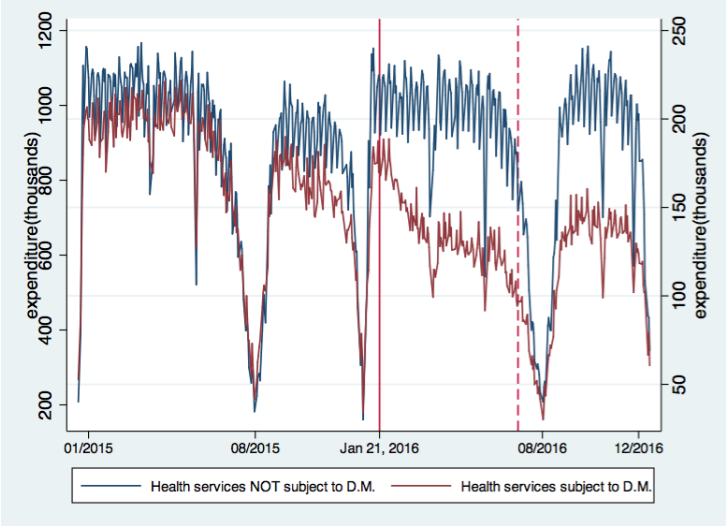
We first draw information on volumes and expenditure for outpatient visits prescribed by General Practitioners (GP) over the period 01/01/2015–31/12/2016 from the *Outpatients Record* (46 million prescriptions for a sample of 2,255,326 individuals). Through a unique anonymous personal identifier we then match this record with the *Regional Health Roster*, the *Hospital Discharge Record* and the *Emergency care Access Record*, to gather additional information on patients' health and demographic characteristics, hospital admissions, access to emergency care and GPs' characteristics. From the individual-level archive we extract information on patients' age, presence of any diagnosed chronic disease[3] and cost-sharing exemptions (based on income, chronic disease and severe disability), while GPs' characteristics include age, years of practice, number of patients, union membership and participation to the Chronic Related Group (CReG) project[4].

[3]We focus on Cardiopathies, Chronic Obstructive Pulmonary Diseases (COPD), Diabetes and Cancer that are mostly diffused.

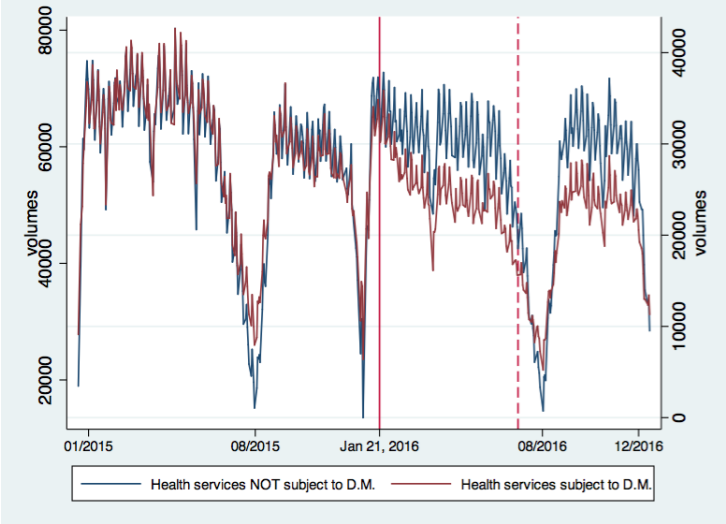
[4]In 2011 Lombardy's regional government launched an experimental model of organization - CReG - aimed at promoting continuity of care for chronic patients after hospital discharge, as chronic diseases represent a growing burden for regional health. The core innovation with the CReG model is the creation of a new subject - provider - whose endowed with predefined economic resources intended for the provision of established treatment plans to chronic patients outside hospitals. This new organizational model is currently being tested in several Local Health Units and involves patients affected by hypertension, diabetes, heart failure, and chronic obstructive pulmonary disease. The CReG pilot involved more than 60,000 patients and 450 GPs in the role of providers.

Overall, our baseline sample consists of 2,723 doctors for a total of 2,255,326 individuals, 37% living in Milan, 56% in municipalities belonging to its province and 7% in Lodi. Average age is around 49 and more than half of the patients meet at least one criterion for cost-sharing exemptions (48% are income-exempt, 38% have access to a disease-related exemption and 8% have severe disability). Patients with chronic diseases account for 40% of the sample, most of which suffer from cardiopathies (27%).

Fig. 1 Time trends in volumes and expenditure for outpatient visits



(a) Expenditure(thousands)



(b) Volumes

Figure 1 depicts the trends in daily expenditure (panel *a*) and volumes (panel *b*) of outpatient visits prescribed by GPs over the period under consideration, for the 203 procedures subjected to appropriateness criteria (*treatment* group - red line) and for all the remaining medical services

unaffected by the introduction of the decree (*control* group - blu line)[5]. The vertical red solid and dashed lines represent respectively the introduction of the decree (January 21^{st},2016) and the successive repealing of the latter in mid-July.

Overall, health care utilization and related expenses show some cyclical patterns, with downturns during summer and around Christmas holidays. However, both expenditure and, to a lesser extent, volumes of prescriptions for health services subjected to appropriateness criteria show a remarkable drop after January 21^{st} 2016 until mid-July (when the repealing of the decree was announced), with considerably lower levels with respect to the same months of the previous year. Such contraction among treated health services is particularly pronounced for Diagnostic Imaging and Laboratory analyses, while spending and volumes for diagnostic tests belonging to the other 5 specialties subjected to appropriateness criteria (residual category) do not significantly change (see Figure A1 in the Appendix). Trends in the second semester are rather stable around pre-summer levels, albeit much lower as compared with the same semester in 2015. On the other hand, we do not observe relevant differences in the trends of the control group before and after the introduction of the decree, with a blue line for both expenditure and volumes that is relatively flat, net of cyclical patterns.

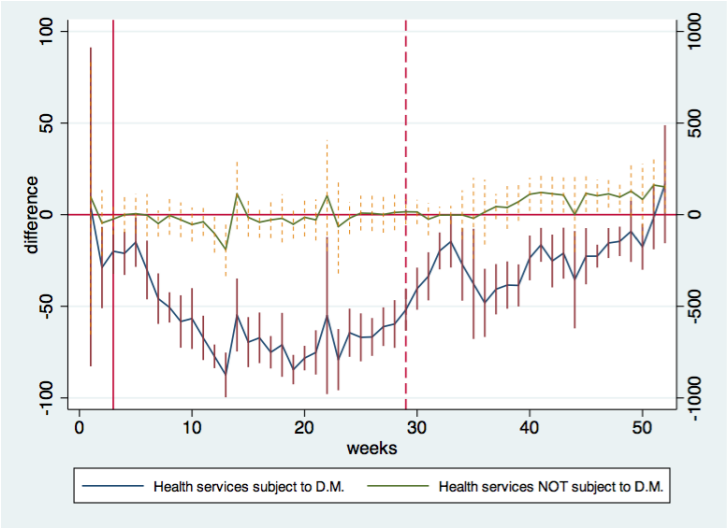
As a further investigation, in Figure 2 we test the difference in the levels of expenditure (panel *a*) and volumes (panel *b*) of treated and control health services between each week in 2016 and the same week in 2015. The relatively flat line around zero observed in Panel *a*, representing the mean difference in outpatients' expenditure between 2016 and 2015 for the control group, confirms a stable trend for the latter, with no significant changes after the introduction of the decree. Conversely, the blue line suggests a remarkable gap in weekly figures for health care spending in the treatment group, that progressively lessens when comparing the second semester of the two years.

Weekly figures for mean differences in total volumes of prescribed health services (panel *b*) present analogous patterns. Starting from mid-February, the level of prescriptions for diagnostic tests affected by appropriateness constraints is substantially lower in each week of 2016 compared with the same week of the previous year, while no statistically significant difference is found for health services belonging to the control group. However, this preliminary analysis suggests that a larger response to prescriptive appropriateness comes from GPs' spending rather than volumes

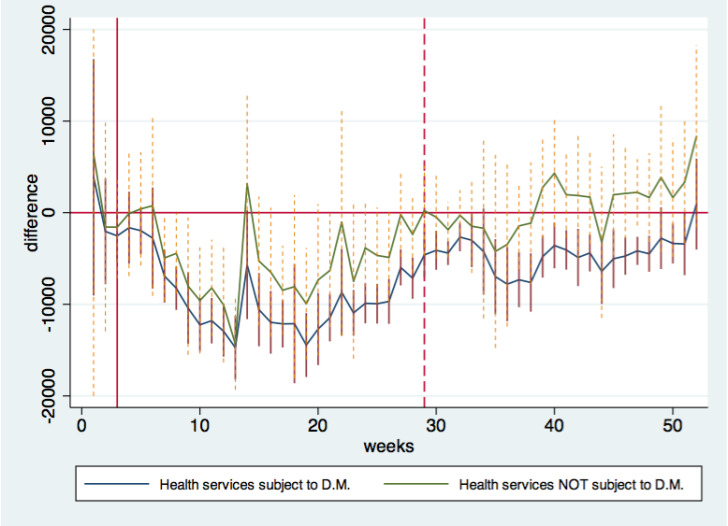
[5]Outpatient visits provided on Saturdays, Sundays and holidays are excluded from the following descriptive analysis.The *y-axis* on the left is for the control group, the right-side axis for the treatment group.

of prescriptions.

Fig. 2 Week-by-week difference between 2016 and 2015's expenditure and volumes of treatment and control groups of health services



(a) Expenditure(thousands)



(b) Volumes

3.2 Empirical strategy

In our empirical exercise we start from the overall evaluation of the impact of the D.M 9^{th} December, 2015 on expenditure and volumes of outpatient visits prescribed by general practitioners. To this end, we pool together the information on outpatient visits from January 2015

to December 2016 and estimate a difference-in-difference model of the form:

$$logY_{ijt} = \alpha + \beta T_i + \gamma Post_t + \delta \underbrace{(T_i * Post_t)}_{DD_{it}} + \phi_j + \epsilon_{ijt} \qquad (1)$$

where $logY_{ijt}$ is our outcome of interest and represents either total expenditure or volumes in logs for each health service i prescribed by GP j in period $t = 1, ..., 23$[6] . T_i is a binary variable defining treated services, that are those whose prescriptive criteria have been modified by the decree, and it captures possible differences between treatment and control groups prior to the policy change. $Post_t$ is a dummy for the time period after the introduction of the D.M.[7], capturing aggregate factors that would cause changes in outpatient visits even in the absence of the decree. δ is our coefficient of interest (the Diff-in-Diff coefficient), measuring the effect of the introduction of the D.M. as $\delta = [Y_{T,1} - Y_{T,0}] - [Y_{C,1} - Y_{C,0}]$. ϕ_j are GP fixed-effects that account for time-invariant unobserved characteristics of the doctor that might confound the effect of the decree and ϵ_{ijt} is an error term[8].

We further exploit the repealing of the decree that occurred in July to estimate a difference-in-difference model with an additional treatment, that allows us to investigate whether the withdrawal had any impact on outpatient visits. Specifically, we estimate a model of the form:

$$logY_{ijt} = \alpha + \beta T_i + \gamma_1 PostJAN_t + \gamma_2 PostJUL_t + \delta_1 \underbrace{(T_i * PostJAN_t)}_{DD_{it}\text{introduction D.M.}} + \delta_2 \underbrace{(T_i * PostJUL_t)}_{DD_{it}\text{repealing D.M.}} + \phi_j + \epsilon_{ijt} \qquad (2)$$

where δ_1 measures the effect of the enforcement of the decree as in equation (1), while δ_2 refers to the repealing of the D.M..

Finally, we explore the heterogeneous effects of the decree on health services related with

[6] In the empirical analysis we exclude January 2016 as the decree entered into force on the 21^{st}.

[7] As we exclude Jan 2016 from the sample, the $Post_t$ dummy takes on value 0 from Jan 2015 to Dec 2015, 1 from Feb 2016 onward.

[8] As an additional exercise, we also set up an event-study specification (Autor, 2003) at the prescription level to verify whether the parallel trends assumption of difference-in-difference models is satisfied. To this end, we set the last period before policy intervention - i.e. December 2015 - as a baseline and estimate a model that allows for leads and lags of the treatment. In other words, we include interaction terms between time dummies and the treatment indicator for each pre and post-treatment month over the period of interest:

$$logY_{it} = \phi_i + \sum_{k \neq -m}^{q} \gamma_{t+k} Time_{t+k} + \sum_{k \neq -m}^{q} \delta_{i,t+k} (T_i * Time_{t+k}) + \epsilon_{it}$$

where ϕ_i are fixed-effects for health services and all other elements are defined as in equation (1). If the assumption holds, no statistically significant coefficient should be associated with any of the pre-treatment interaction terms. Results from this estimation procedure, plotted in Figure A2 of the Appendix, show that both outcomes for treatment and control group in the pre-treatment months follow a common trend, while after the introduction of the decree we observe a decrease in expenditure and volumes of outpatient visits.

vulnerable patients by estimating equation (1) separately on diagnostic tests prescribed to patients belonging to different age groups, with and without cost-sharing exemptions, and with selected chronic diseases.

4 Results

4.1 Baseline results

In order to carry out our impact evaluation exercise we build our dataset as to have the total amount of prescriptions and expenditure for each medical service, prescribed by each GP in each month and year of the period under consideration (January 2016 is excluded). We then estimate equation (1) on the pooled sample of health services and separately by major specialties (Diagnostic Imaging, Laboratory Analyses and a residual category[9]).

Tab. 1 Estimated impact of the decree on expenditure and volumes of outpatient visits

	Pooled Sample	***Category***		
		DI	*Lab*	*Others*
Expenditure (log)				
treat	-0.395**	1.189***	-0.200	-0.00327
	(0.179)	(0.232)	(0.167)	(0.255)
post	0.0161**	-0.00409	-0.00942	0.0244**
	(0.00750)	(0.0108)	(0.00873)	(0.00993)
treat*post	-0.238***	-0.111***	-0.214***	-0.0718**
	(0.0253)	(0.0350)	(0.0261)	(0.0349)
constant	3.464***	4.317***	2.974***	3.730***
	(0.0638)	(0.116)	(0.0940)	(0.204)
R^2	0.0660	0.174	0.0769	0.0278
Volume (log)				
treat	0.750***	0.128	0.683***	0.157
	(0.180)	(0.146)	(0.211)	(0.355)
post	-0.00916	0.0245***	-0.0225**	0.0176**
	(0.00612)	(0.00775)	(0.0106)	(0.00685)
treat*post	-0.117***	-0.0956*	-0.118***	-0.109***
	(0.0231)	(0.0515)	(0.0277)	(0.0332)
constant	0.831***	0.617***	0.999***	1.010***
	(0.0618)	(0.0683)	(0.107)	(0.0872)
R^2	0.0795	0.0434	0.0807	0.0665
Physician fixed-effects	✓	✓	✓	✓
N	7,787,261			

Significance: * p<.1, ** p<.05, *** p<.01. Robust standard errors in parentheses, clustered at the prescription level.

[9]The residual category encompasses dentistry and maxillo-facial surgery, physical therapy and rehabilitation, dermatology, nuclear medicine and genetic tests, that are the other 5 specialties whose diagnostic tests were subjected to the decree. We choose to group the above specialties into a residual category as the level of spending for this group only represents 8% of total expenditure (DI and Laboratory analyses represent 30% each) and the volume of prescriptions accounts for 5% (80% of total health services prescribed are lab. results).

Results in Table 1 suggest an overall negative impact of the decree on both expenditure and volumes of prescribed outpatient visits. The introduction of prescriptive constraints reduced the level of health care services by almost 12%, with a 24% associated contraction in spending. The largest reduction in prescriptions is associated with laboratory analyses, that also implied a 21% saving in health care expenses, while for diagnostic imaging and the residual category the estimated impact is milder.

As an additional exercise to test the effectiveness of prescriptive constraints in shaping GPs' behavior, we exploit the repealing of the decree to estimate a difference-in-difference model that includes both treatments (see Table 2). If prescriptive appropriateness policies are effective in reducing unnecessary spending and over-prescription behaviors, one would expect a strong negative reaction of GPs to the introduction of the decree in January 2016 and a milder upturn in outpatient visits after the repealing of the latter, that loosens constraints and softens sanctions on non-compliant GPs. Estimation results for equation (2) show a 21% (10%) reduction in outpatients' spending (volumes) for health services subjected to appropriateness criteria following January 2016, and a weaker effect after the decree was withdrawn in mid-July.

Tab. 2 Estimated impact of the decree and its' repealing (July 2016)

	Expenditure (log)	***Volume (log)***
treat	-0.395**	0.750***
	(0.179)	(0.180)
post (Jan 2016)	0.0360***	0.00218
	(0.00716)	(0.00607)
post (Jul 2016)	-0.0461***	-0.0263***
	(0.00462)	(0.00444)
treat*post (Jan 2016)	-0.211***	-0.0961***
	(0.0259)	(0.0216)
treat*post (Jul 2016)	-0.0605***	-0.0467***
	(0.0177)	(0.0179)
constant	3.464***	0.831***
	(0.0638)	(0.0618)
R^2	0.0663	0.0796
Physician fixed-effects	✓	✓
N	7,787,261	7,787,261

Significance: * p<.1, ** p<.05, *** p<.01. Robust standard errors in parentheses, clustered at the prescription level.

The empirical analysis presented so far thus indicates that the D.M. *Lorenzin* did achieve its primary goal of health care cost containment. However, when we explore the heterogeneous effect of the decree on outpatients' spending and volumes across different groups of patients, we uncover some unintended effects on vulnerable patients.

As a matter of fact, the cost-containment aim of the decree was essentially targeted on

unnecessary spending and over-prescription behaviors, so that vulnerable patients - i.e. the elderly, patients with cost-sharing exemptions or suffering from chronic diseases - in actual need of care should not be affected by prescriptive constraints, as they usually meet appropriateness criteria. Nevertheless, results from our heterogeneity exercise show a persistent negative effect across all groups of individuals (see Tables A1, A2 and A3). The reduction in spending and volumes of prescriptions associated with the introduction of the D.M. is in fact statistically different from zero also among patients older than 65 and with cost-sharing exemptions, albeit smaller in magnitude with respect to younger patients and full-payers. As for chronic illnesses, the impact of prescriptive constraints on outpatients' expenditure and volumes is largest among subjects suffering from cardiopathies and cancer patients, slightly lower for COPD and diabetic patients.

4.2 GPs' behavior

As the decree introduced constraints on the ability of physicians and general practitioners to prescribe health services, thus limiting their gate-keeping role, we are further interested in exploring heterogeneity across GPs' behaviors. In other words, we ask whether individual characteristics of family doctors are associated with different prescriptive behaviors in response to the D.M. In this context, we focus on a restricted sample of 2,545 GPs that are observed in the Outpatients Registry every month over the period under consideration (January 2016 is still excluded) - i.e. induced at least one outpatient visit a month[10] - and we concentrate on outpatients' spending, as the estimated impact of the decree is much larger in terms of expenditure than in volumes of prescriptions[11].

The average GP in the sample is 58 years-old, has 1,343 patients and 23 years of working experience (see Table A4). More than half of the sample of doctors is unionized, while only 7% participated in the CReG project.

We then evaluate the differential effect of the decree on outpatients' spending according with age of doctors, years of practice, union membership and active participation in the CReG project[12]. To this end, we add to our baseline specification the above vector of GPs' characteris-

[10]Note that not being present in the Outpatients Registry in a given month for a GP means no prescriptions at all, that is a very unlikely event considering that the average number of patients is above 1,200. As a matter of fact, the 178 doctors excluded are notably different from those in the core sample, with a considerably smaller number of patients (300 on average as compared to 1,340) and lower seniority (more than 50% of excluded doctors have less than 5 years of practice vs. a median value of 24 among sampled GPs).

[11]Results on volumes of outpatient visits are available in the Appendix.

[12]We dichotomize age and years of practice so as to have two binary variables for older (above age 63) and

tics, their interaction with both $Treat_i$ and $Post_t$ dummies, as well as a triple interaction among $Treat_i$, $Post_t$ and GPs' characteristics. The coefficient on this last term measures heterogeneity in the effect of treatment across doctors. Results are reported in Table 3.

Tab. 3 Estimated impact of the decree on outpatient visits' expenditure - heterogeneity by GP characteristics

	Years of Practice	*Age*	*Creg ass.*	*Union member*
Expenditure (log)				
treat	-0.404**	-0.387**	-0.398**	-0.343*
	(0.178)	(0.180)	(0.180)	(0.186)
post	0.0206***	0.0188**	0.0158**	0.0211***
	(0.00743)	(0.00754)	(0.00747)	(0.00769)
GP characteristics	-0.0152***	-0.00364		
	(0.00569)	(0.00384)		
treat*post	-0.246***	-0.250***	-0.241***	-0.233***
	(0.0259)	(0.0267)	(0.0260)	(0.0259)
treat*GP	0.0700***	-0.00277	0.124***	-0.0874***
	(0.0199)	(0.0120)	(0.0255)	(0.0218)
post*GP	-0.0171***	-0.0144***	-0.000848	-0.0102***
	(0.00218)	(0.00255)	(0.00340)	(0.00220)
treat*post*GP	0.0137**	0.0493***	0.00419	-0.0168***
	(0.00667)	(0.00905)	(0.00976)	(0.00603)
constant	3.475***	3.472***	3.471***	3.471***
	(0.0638)	(0.0640)	(0.0640)	(0.0640)
R^2	0.0596	0.0595	0.0596	0.0597
Physician fixed-effects	✓	✓	✓	✓
N	7,732,045	7,732,045	7,732,045	7,732,045

Significance: * p<.1, ** p<.05, *** p<.01. Robust standard errors in parentheses, clustered at the prescription level.

Age and seniority are associated with a lower compliance with prescriptive constraints: the reduction in spending on outpatient visits is 5% (1%) lower for older (experienced) family doctors with respect to their younger (less experienced) counterpart, suggesting that long-standing GPs might be more reluctant to change their prescriptive behavior in response to the introduction of appropriateness requirements. On the contrary, unionized doctors show a stronger reaction to the introduction of prescriptive appropriateness requirements (-25% spending) with respect to non-unionized GPs (-23.3%), and we tentatively interpret this finding in the light of a greater awareness about penalties and sanctions associated with non-compliance. Finally, no behavioral differences in terms of spending are found regarding participation in the CReG's project.

4.3 Emergency Room services and Hospital admissions after the D.M.

The empirical analysis conducted so far revealed that 1) overall the D.M. *Lorenzin* was effective in reducing health care spending on outpatient visits and, to a lesser extent, the volume

experienced GPs (more than 31 years of practice), that are those in the last quartile of the distribution. Union membership and CReG participation are also dichotomous indicators and are time invariant.

of prescriptions, 2) the decree also had an effect on the patterns of outpatient visits for vulnerable patients, possibly exposing them to higher health risks, 3) the impact of the decree was heterogeneous across doctors. As a final step, in this last section we investigate whether the significant contraction in outpatients' spending that followed the introduction of the decree also had some unintended consequences on patients outcomes in terms of increasing resort to emergencies and hospitalizations. In particular, we focus on those GPs who responded to the decree by lowering expenditure for outpatient visits (i.e. *compliers*) and explore whether there exists a correlation between stronger reactions in terms of prescriptive behavior and percentage changes in Emergency Room services or hospital admissions between 2015 and 2016. To this end, we first estimate our baseline difference-in-difference model separately by GP, so as to retrieve a measure of each GP's reaction to the D.M. (the coefficient of the treat/post interaction - i.e. δ in equation 1). We then focus on compliant GPs - those with negative δs - and use the absolute value of the δs to predict the 2016-2015 percentage change in ER services and hospital admissions at the GP level. In other words, we are interested in studying whether 1) a larger contraction in outpatients' spending among compliant GPs is associated with a higher probability of an increase in ER services (substitution effect) or hospital admissions (health issues), 2) GPs with larger δs also show greater variation in ER services or hospital admissions, 3) the variation in outcomes differs for vulnerable individuals.

Table 4 presents the results of three different estimation exercises. In column 1 (column 4) we regress a binary indicator for increasing ER services (hospital admissions) on GPs' δs, thus estimating whether compliant GPs showing larger reductions in expenditure are more likely to show also an increase in ER services (hospital admissions) between 2015 and 2016. In column 2 (column 5) the δs are used to predict the magnitude of the percentage increase in ER services (hospital admissions). In this setting, a positive coefficient suggests that GPs that reacted more to the D.M. (i.e. larger contraction in outpatients' spending) also experienced a larger percentage increase in ER services or hospital admissions between 2015 and 2016. Finally, after cutting the distribution of the δs into four quartiles, in column 3 (column 6) we estimate whether the percentage increase in ER services (Hospital admissions) differs across quartiles.

The estimated coefficient of interest in column 1 indicates that GPs that reduced outpatients' spending more are also more likely to record an increasing resort to emergencies among their patients, while the opposite holds for hospital admissions (column 4), suggesting that there might be some substitution going on between outpatient and emergency visits but no health

Tab. 4 GP's response to the decree and patients' admission to Hospital and ER

	2016-2015 % Variation (Δ)					
	ER services			***Hospital admissions***		
Expenditure	*Prob. of* $\Delta > 0$	Δ	Δ	*Prob. of* $\Delta > 0$	Δ	Δ
$\|\delta\|$ DD	0.145*	0.0363**		-0.207**	0.0235	
	(0.0864)	(0.0180)		(0.0864)	(0.0273)	
2^{nd} quartile of $\|\delta\|$'s			0.00441			-0.00356
			(0.00779)			(0.0108)
3^{rd} quartile of $\|\delta\|$'s			0.00347			0.0193*
			(0.00766)			(0.0110)
4^{th} quartile of $\|\delta\|$'s			0.0178**			0.00501
			(0.00767)			(0.0111)

Significance: * p<.1, ** p<.05, *** p<.01.

worsening in the short-run. However, very large contractions in spending are associated with greater increments in both ER and hospital admissions (columns 3 and 6). Supporting evidence of substitution effect also comes from the fact that the increase in ER accesses associated with outpatients' spending cuts is essentially driven by non-urgent care (white and red codes), while we find no correlation between outpatients and urgent emergency care.

Finally, we explore whether the relation between GPs' behavior and changes in ER or hospital admissions differs across groups of patients. In particular, we estimate whether compliant GPs with larger δs are more likely to report an increase in ER and hospital admissions related with vulnerable patients - i.e. older than 65 or with a cost-sharing exemption. In the upper panel of Table 5 we report estimates of the probability that compliant GPs with larger δs experienced an increase in ER services among four different groups of patients: those with or without a payment exemption, older than 65 and up to 65 years-old. The lower panel replicates the same exercise on hospital admissions.

Tab. 5 Vulnerable patients

	Patient's type			
Expenditure	*w/Exemption*	*wo/Exemption*	*over 65yo*	*up to 65yo*
Prob. of $\Delta > 0$ **- ER services**				
$\|\delta\|$ DD	0.211**	-0.0307	0.0947	0.231***
	(0.0863)	(0.0886)	(0.0905)	(0.0858)
Prob. of $\Delta > 0$ **- Hospital admissions**				
$\|\delta\|$ DD	-0.0688	-0.000300	-0.105	-0.0884
	(0.0864)	(0.0905)	(0.0921)	(0.0860)

Significance: * p<.1, ** p<.05, *** p<.01.

Estimation results show no correlation between outpatient and emergency care among payers and the elderly, while tightening outpatients' spending is associated with an increasing resort

to ER for patients with cost-sharing exemptions. No heterogeneous effects are found on hospital admissions, suggesting no health damaging effects of outpatients' cost containment also considering vulnerable patients in the short run.

5 Conclusions

Increasing health care spending and the aging of populations characterizing most industrialized countries over the last decades are one of the most prominent issues and a matter of great concern for policy-makers around the world. Faced with the current economic pressure of health care cost containment, the main target of health care systems is to achieve budgetary goals while ensuring the provision of "appropriate, scientifically sound, socially acceptable and universally accessible methods and technologies" (Organization, 1978) to avoid possible adverse effects on health outcomes. In this context, the concept of *appropriateness* in health care has been guiding several cost-containment policies aimed at tackling overconsumption of medical services and increasing unnecessary spending.

In Italy, a lively debate over the appropriateness of health care delivered within the National Health System led to the establishment of a set of appropriateness requirements for the prescription of a selected number of health services and diagnostic tests by general practitioners. The so-called "appropriateness" decree, issued by the Ministry of Health in December 2015, introduced constraints on the ability of physicians and family doctors to prescribe 203 health services, unless predetermined conditions are met by the patient.

In this paper we exploit this policy shift within a difference-in-difference framework to assess the impact of prescriptive appropriateness requirements on health care utilization and spending, as well as some unintended outcomes. In particular, drawing on administrative health records of the Agency for Health Protection of the Province of Milan, we explore 1) the overall effect of the decree on outpatients' spending and volumes, 2) whether the introduction of appropriateness requirements had some unintended consequences on the demand of medical services of vulnerable groups, 3) whether GPs' response to the introduction of prescriptive constraints has been heterogeneous according with several individual characteristics and 4) whether changes in outpatients' spending are associated with patients' resort to emergency care or increasing hospitalization.

We find that the introduction of the decree was followed by a 12% reduction in the level of

health care services prescribed by family doctors, with a 24% associated contraction in spending, suggesting that the policy did achieve its primary goal of cost containment. However, when we explore the heterogeneous effect of prescriptive constraints on outpatients' spending and volumes across different groups of patients, we uncover some unintended effects on vulnerable patients (i.e. the elderly, patients with cost-sharing exemptions or suffering from chronic diseases). We find a statistically significant reduction in both spending and volumes also among patients older than 65 and with cost-sharing exemptions, albeit smaller in magnitude with respect to younger patients and full-payers.

Moreover, some degree of heterogeneity is also found in GPs' behavior, where age and seniority are associated with a lower compliance with prescriptive constraints. Older and experienced family doctors show a 1 to 5% milder reduction in spending as compared to their younger and less experienced counterpart, suggesting that a long-stending medical career might be associated with some reluctance in changing prescriptive behavior.

Finally, when we analyze the relationship between outpatients' spending and patients' outcomes we find that compliant GPs associated with larger reductions in outpatients' spending are also more likely to record an increasing resort to emergency care among their patients, while the opposite holds for hospital admissions. This last finding suggests that the decree might have triggered some substitution effects between outpatient and emergency visits, without directly worsening patients' health at least in the short-run.

References

Autor, D. H. (2003), 'Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing', *Journal of Labor Economics* **21**(1), 1–42.

Baicker, K. and Goldman, D. (2011), 'Patient Cost-Sharing and Healthcare Spending Growth', *Journal of Economic Perspectives* **25**(2), 47–68.

Dusheiko, M., Gravelle, H., Jacobs, R. and Smith, P. (2006), 'The effect of financial incentives on gatekeeping doctors: evidence from a natural experiment', *Journal of Health Economics* **25**(3), 449–478.

Fiorentini, G., Bruni, M. L. and Ugolini, C. (2013), 'GPs and hospital expenditures. Should we keep expenditure containment programs alive?', *Social Science & Medicine* **82**, 10–20.

Goodell, S. and Swartz, K. (2010), 'Cost-sharing: effects on spending and outcomes', *Policy Brief* **20**, 42–45.

Lavis, J. N. and Anderson, G. M. (1996), 'Appropriateness in health care delivery: definitions, measurement and policy implications', *CMAJ: Canadian Medical Association Journal* **154**(3), 321.

OECD (2005-2017), 'Health at a Glance: OECD Indicators', *OECD Publishing* .

Organization, W. H. (1978), Declaration of Alma Ata, *in* 'International Conference on Primary Health Care Alma-Ata, 6-12 September, 1978'.

Panella, M., Rinaldi, C., Leigheb, F., Knesse, S., Donnarumma, C., Kul, S., Vanhaecht, K. and Di Stanislao, F. (2017), 'Prevalence and costs of defensive medicine: a national survey of Italian physicians', *Journal of Health Services Research & Policy* .

Pauly, M. V. and Blavin, F. E. (2008), 'Moral hazard in insurance, value-based cost sharing, and the benefits of blissful ignorance', *Journal of Health Economics* **27**(6), 1407–1417.

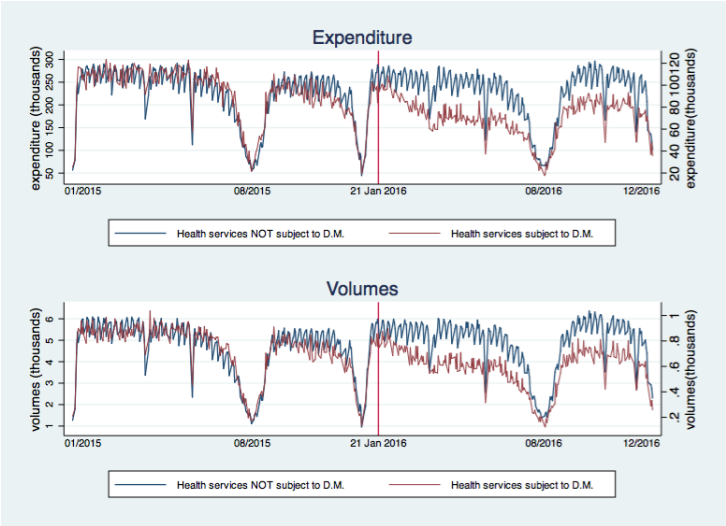
Schreyögg, J. and Grabka, M. M. (2010), 'Copayments for ambulatory care in Germany: a natural experiment using a difference-in-difference approach', *The European Journal of Health Economics* **11**(3), 331–341.

van Kleef, R. C., van de Ven, W. P. and van Vliet, R. C. (2009), 'Shifted deductibles for high risks: More effective in reducing moral hazard than traditional deductibles', *Journal of Health Economics* **28**(1), 198–209.

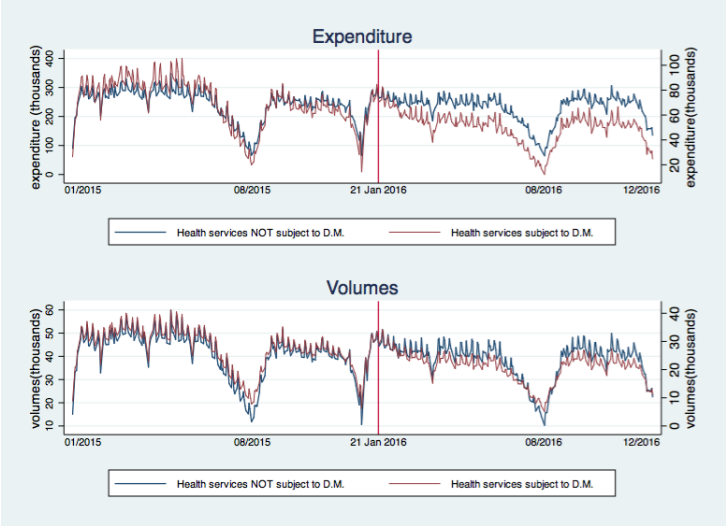
Ziebarth, N. R. (2014), 'Assessing the effectiveness of health care cost containment measures: evidence from the market for rehabilitation care', *International Journal of Health Care Finance and Economics* **14**(1), 41–67.

6 Appendix

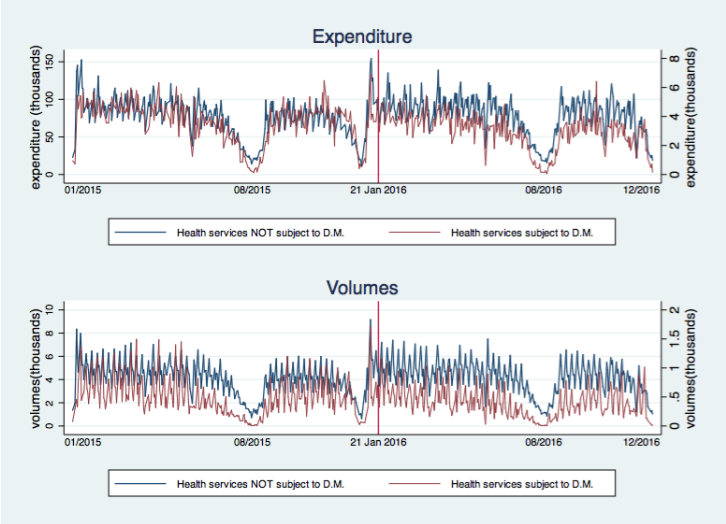
Fig. A1 Time trends of outpatient visits' expenditure and volumes across categories of medical services



(c) Diagnostic Imaging

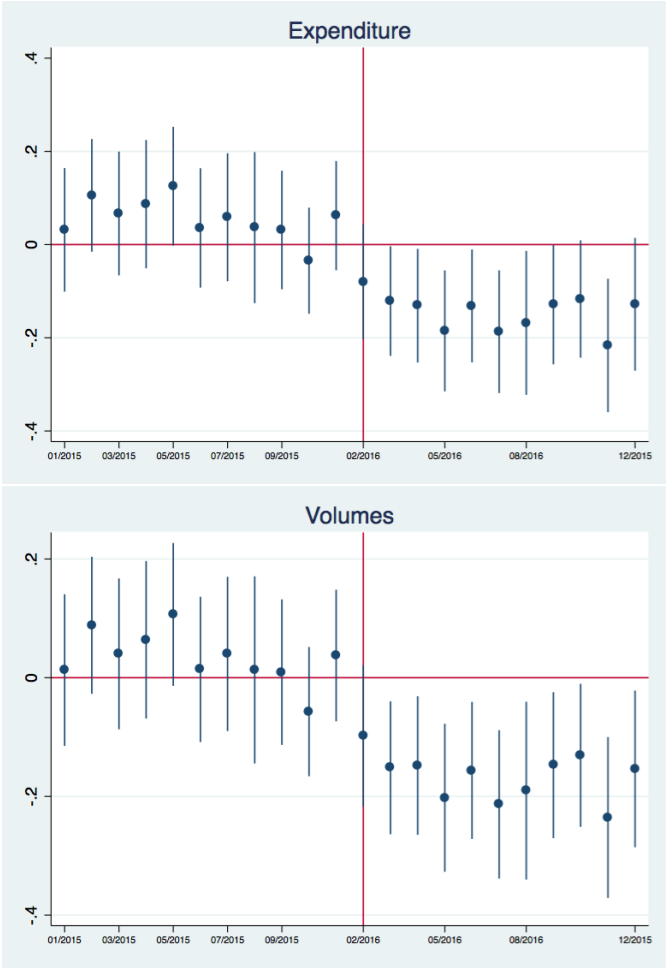


(d) Laboratory Analyses



(e) Other health services subjected to D.M.

Fig. A2 Estimated impact of the decree on expenditure and volumes of outpatient visits for months before, during, and after the introduction



Tab. A1 Estimated impact of the decree on expenditure and volumes of outpatient visits - by age classes

	Age classes		
	0-44	*45-64*	*65+*
Expenditure (log)			
treat	-0.826***	-0.699***	-0.753***
	(0.155)	(0.185)	(0.162)
post	0.00226	0.0166*	0.0200**
	(0.00752)	(0.00965)	(0.00852)
treat*post	-0.266***	-0.246***	-0.200***
	(0.0310)	(0.0290)	(0.0263)
constant	2.911***	3.147***	3.401***
	(0.0651)	(0.0693)	(0.0709)
R^2	0.137	0.114	0.0814
Volume (log)			
treat	0.394***	0.560***	0.650***
	(0.128)	(0.160)	(0.165)
post	-0.0288***	-0.0104*	-0.00569
	(0.00881)	(0.00547)	(0.00511)
treat*post	-0.171***	-0.119***	-0.0823***
	(0.0305)	(0.0295)	(0.0244)
constant	0.546***	0.605***	0.765***
	(0.0618)	(0.0694)	(0.0732)
R^2	0.0498	0.0745	0.0616
Physician fixed-effects	✓	✓	✓

Significance: * p<.1, ** p<.05, *** p<.01. Robust standard errors in parentheses, clustered at the prescription level.

Tab. A2 Estimated impact of the decree on expenditure and volumes of outpatient visits - by exemption status

	Expenditure(log)		***Volume(log)***	
	Ex	*Non-ex*	*Ex*	*Non-ex*
treat	-0.574***	-0.610***	0.686***	0.561***
	(0.168)	(0.182)	(0.171)	(0.157)
post	0.0183**	0.0205**	-0.00671	-0.0211**
	(0.00753)	(0.00883)	(0.00532)	(0.00864)
treat*post	-0.220***	-0.272***	-0.0968***	-0.145***
	(0.0256)	(0.0295)	(0.0229)	(0.0319)
constant	3.412***	3.013***	0.774***	0.587***
	(0.0644)	(0.0667)	(0.0650)	(0.0685)
R^2	0.0916	0.0972	0.0808	0.0727
Physician fixed-effects	✓	✓	✓	✓

Significance: * p<.1, ** p<.05, *** p<.01. Robust standard errors in parentheses, clustered at the prescription level.

Tab. A3 Estimated impact of the decree on expenditure and volumes of outpatient visits - by chronic disease

	Chronic diseases			
	Cardiopathies	*COPD*	*Diabetes*	*Cancer*
Expenditure (log)				
treat	-0.744***	-1.215***	-1.174***	-1.014***
	(0.162)	(0.145)	(0.133)	(0.158)
post	0.0225***	0.0370***	0.0249***	0.0281***
	(0.00819)	(0.00748)	(0.00729)	(0.00775)
treat*post	-0.194***	-0.126***	-0.139***	-0.155***
	(0.0256)	(0.0168)	(0.0184)	(0.0194)
constant	3.334***	2.584***	2.837***	2.927***
	(0.0718)	(0.112)	(0.0967)	(0.0888)
R^2	0.128	0.250	0.252	0.193
Volume (log)				
treat	0.668***	0.172***	0.280***	0.314***
	(0.169)	(0.0662)	(0.107)	(0.102)
post	-0.00408	-0.00607**	-0.00955**	-0.00855**
	(0.00513)	(0.00289)	(0.00425)	(0.00394)
treat*post	-0.0723***	-0.0430***	-0.0521***	-0.0652***
	(0.0259)	(0.00977)	(0.0152)	(0.0121)
constant	0.721***	0.325***	0.504***	0.465***
	(0.0746)	(0.0541)	(0.0769)	(0.0672)
R^2	0.0910	0.0549	0.0610	0.0641
Physician fixed-effects	✓	✓	✓	✓

Significance: * p<.1, ** p<.05, *** p<.01. Robust standard errors in parentheses, clustered at the prescription level.

Tab. A4 GPs' descriptive statistics

	Statistics					
Variable	*Mean*	*p1*	*p25*	*p50*	*p75*	*p99*
N. patients	1,340	251	1,111	1,511	1,579	1,754
Age	58	37	54	59	63	68
Years of practice	23	1	18	24	31	38
Union member	.53					
CReG member	.072					
Total expenditure - 2015	101,894	5,739	55,974	110,695	142,863	214,319
Total expenditure - 2016	99,009	6,193	58,623	106,869	137,684	207,508
ER services - 2015	370	57	308	377	440	616
ER services - 2016	375	54	314	383	442	640
Hosp. admissions - 2015	149	27	119	157	182	238
Hosp. admissions - 2016	149	23	117	158	183	238

Note: Total spending, ER and Hospital admissions do not include 01/2015, to make figures for 2015 comparable with 2016.

Tab. A5 Estimated impact of the decree on volumes of outpatient visits' - heterogeneity by GP characteristics

	Years of Practice	*Age*	*Creg ass.*	*Union member*
Volume (log)				
treat	0.746***	0.757***	0.748***	0.791***
	(0.179)	(0.182)	(0.181)	(0.190)
post	-0.00709	-0.00746	-0.00898	-0.00518
	(0.00615)	(0.00618)	(0.00614)	(0.00657)
GP characteristics	-0.00662	-0.00322		
	(0.00492)	(0.00283)		
treat*post	-0.125***	-0.128***	-0.121***	-0.116***
	(0.0237)	(0.0246)	(0.0239)	(0.0235)
treat*GP	0.0447**	-0.00476	0.104***	-0.0661***
	(0.0186)	(0.0102)	(0.0250)	(0.0187)
post*GP	-0.00780***	-0.00823***	-0.00360	-0.00776***
	(0.00139)	(0.00195)	(0.00225)	(0.00173)
treat*post*GP	0.0156***	0.0419***	0.0194**	-0.00889*
	(0.00524)	(0.00849)	(0.00882)	(0.00460)
constant	0.836***	0.835***	0.834***	0.834***
	(0.0615)	(0.0619)	(0.0619)	(0.0619)
R^2	0.0782	0.0781	0.0782	0.0783
Physician fixed-effects	✓	✓	✓	✓
N	7,732,045	7,732,045	7,732,045	7,732,045

Significance: * p<.1, ** p<.05, *** p<.01. Robust standard errors in parentheses, clustered at the prescription level.

Tab. A6 GP's response to the decree (volumes of prescriptions) and patients' admission to Hospital and ER

	2016-2015 % Variation (Δ)					
	ER services			***Hospital admissions***		
Expenditure	*Prob. of* $\Delta > 0$	Δ	Δ	*Prob. of* $\Delta > 0$	Δ	Δ
$\|\delta\|$ DD	-0.0415	0.0292		0.313**	-0.0336	
	(0.156)	(0.0339)		(0.156)	(0.0450)	
2^{nd} quartile of $\|\delta\|$'s			0.0105			-0.00210
			(0.00940)			(0.0134)
3^{rd} quartile of $\|\delta\|$'s			0.00454			0.0102
			(0.00947)			(0.0135)
4^{th} quartile of $\|\delta\|$'s			0.0167*			-0.0111
			(0.00964)			(0.0131)

Significance: * p<.1, ** p<.05, *** p<.01.