Received: 31 December 2016 | Revised: 23 May 2017 | Accepted: 30 May 2017
DOI: 10.1002/hec.3545

SUPPLEMENT PAPER

WILEY Health Economics

The effect of waiting times on demand and supply for elective surgery: Evidence from Italy

Andrea Riganti[1] | Luigi Siciliani[2] | Carlo V. Fiorio[1,3,4]

[1] Department of Economics, Management and Quantitative Methods, University of Milano, Milan, Italy
[2] Department of Economics and Related Studies, University of York, York, UK
[3] IRVAPP-FBK, Trento, Italy
[4] Dondena, Bocconi University, Milan, Italy

Correspondence
Andrea Riganti, Departments of Economics, Management and Quantitative Methods, University of Milano Via Conservatorio 7, 20122 Milano (MI), Italy.
Email: andrea.riganti@unimi.it

JEL Classification: I10

Abstract
Waiting times are a major policy concern in publicly funded health systems across OECD countries. Economists have argued that, in the presence of excess demand, waiting times act as nonmonetary prices to bring demand for and supply of health care in equilibrium. Using administrative data disaggregated by region and surgical procedure over 2010–2014 in Italy, we estimate demand and supply elasticities with respect to waiting times. We employ linear regression models with first differences and instrumental variables to deal with endogeneity of waiting times. We find that demand is inelastic to waiting times while supply is more elastic. Estimates of demand elasticity are between −0.15 to −0.24. Our results have implications on the effectiveness of policies aimed at increasing supply and their ability to reduce waiting times.

KEYWORDS
demand, elective surgery, supply, waiting times

1 | INTRODUCTION

Waiting times in health care sector are a major health policy concern across many OECD countries (Siciliani, Borowitz, & Moran, 2013). Waiting times for elective surgeries can last several months (Siciliani, Borowitz, & Moran, 2014) and generate dissatisfaction to patients and the general public. Patients' disutility from waiting includes postponed health benefits, potential worsening of health status while waiting, and uncertainty about receipt of treatment.

In many publicly funded systems, the combination of capacity constraints and limited or no user charges generates an excess demand. Patients are added to a waiting list and are asked to wait. Economists have argued that in the absence of price rationing, waiting times act as a form of nonprice rationing that brings together the demand for and the supply of health care (see seminal papers by Lindsay & Feigenbaum, 1984, and Martin & Smith, 1999). On the demand side, a longer wait will induce some patients to go private at a fee (or a reduced fee if they hold private health insurance) or to seek a less intensive drug treatment, therefore reducing the demand for public surgery. On the supply side, waiting times may induce hospitals to work harder and provide more treatments if doctors are altruistic (i.e., they feel bad about the patients waiting excessively) or if penalties are in place for hospitals exceeding maximum waiting time guarantees (see Martin & Smith, 1999, for a theoretical model, and Propper, Burgess, & Gossage, 2008, on penalties).

From a policy perspective, it is critical to establish the extent to which demand and supply respond to waiting time. For example, if demand is highly elastic, an exogenous increase in supply will have only minimal effect in reducing waiting times. In turn, this will make policymakers more reluctant to fund additional resources. Similarly, if supply is elastic, an exogenous increase in demand (e.g., due to ageing population or technology) will imply that waiting time will increase only to a small extent.

There is extensive empirical evidence on demand and supply elasticities from the U.K. Lindsay and Feigenbaum (1984) and Martin and Smith (1999) find that the elasticity of demand is generally low. The finding is also confirmed by more recent studies (Gravelle, Dusheiko, & Sutton, 2002; Gravelle, Smith, & Xavier, 2003 and Martin, Rice, Jacobs, & Smith, 2007). In most studies, demand elasticity is below −0.2. Estimates of supply elasticity are less stable and vary depending on methods, sample, and time period considered (see Siciliani & Iversen, 2012 for a more detailed discussion of the literature).

We know however very little about demand and supply elasticities from other OECD countries. These are likely to differ on the basis of institutional arrangements (gatekeeping system, use of user charges, and payment arrangements) and funding levels. Administrative data on waiting times have been collected within the English National Health System (NHS) since its inception but only in the last years in other countries (Siciliani et al., 2014).

We advance the literature by filling this gap in knowledge and study demand and supply elasticities within the Italian context. Using administrative data in 2010–2014, we employ linear regression models exploiting variability in waiting times by region, surgical procedure, and time. We first estimate pooled cross-section models using ordinary least squares (OLS). Second, we use an instrumental variables (IV) approach to deal with the endogeneity of waiting time due to simultaneity of demand and supply. Finally, we use a first-difference (FD) estimation procedure to address the remaining endogeneity of waiting time due to its possible correlation with time-invariant unobserved factors (e.g., regional factors). Differently from fixed-effect models, which require strict exogeneity (i.e., the error term is uncorrelated to past, present as well as future values of the control variables), FD models only require a weak-exogeneity assumption (i.e., there is no feedback from the idiosyncratic shock today to a covariate tomorrow). This is a considerably weaker assumption as it permits future values of the regressors to be correlated with the error, which is particularly important for the use of past values of regressors as controls or as instruments.

Our key finding is that demand is inelastic to waiting times and in the range of −0.15 and −0.24. This result is important for policy. It implies that an increase in publicly funded supply will reduce waiting times to a great extent because reductions in waiting are only offset by a small increase in demand. Conversely, governments under financial pressure who withdraw resources from the public system will experience large increases in waiting times.

As far as the authors are aware, this is the first study that uses administrative data to estimate demand and supply elasticities within the Italian context. We are only aware of another study that estimates demand elasticity for Italy (Fabbri & Monfardini, 2009). This study focuses on specialist consultations as opposed to elective surgeries. It makes use of a survey in 2000 rather than recent administrative data. The methodology and period covered is different. We are also not aware of studies estimating demand and supply elasticities from other OECD countries (in addition to the U.K.) except for one study from Australia, which finds that demand of public hospitals is elastic to waiting times and the elasticity is equal to −1.7 (Stavrunova & Yerokhin, 2011). This may be explained by the large private sector that generates a more extensive margin between public and private provision compared to England. It also confirms that demand estimates can vary significantly across countries.

The paper is structured as follows. In Section 2, we set out the theoretical framework for the estimation of demand for and supply of elective surgeries in the Italian NHS. In Section 3, we briefly describe the institutional background and sources of data. Sections 4 and 5 describe empirical implementation and provide descriptive statistics. Section 6 contains empirical results. Section 7 concludes and discusses policy implications.

2 | THEORETICAL FRAMEWORK

We adopt the theoretical framework outlined by Martin and Smith (1999, 2003). We assume that waiting times act as a nonmonetary price, which brings the demand for and the supply of elective surgery in equilibrium in a NHS. The demand for publicly funded elective surgery is described by the following function and we include (in parentheses) the expected direction of each of the effects:

$$D = f(waiting\ time(-), need(+), quality(+), private\ availability(-)). \tag{1}$$

Demand for publicly funded surgery is assumed to decrease in waiting times. Longer waiting times may induce some patients at the margin to look for treatment in the private sector by paying out of pocket (or if they hold private health insurance) and therefore to opt out of the public system (i.e., the NHS). In addition, longer waits may induce some patients to substitute surgery with a pharmaceutical treatment therefore reducing demand for publicly funded surgery.

Demand for public treatment will be higher in areas with higher need, for example, areas with an older and sicker population, and in areas where the quality of healthcare is higher making hospital services more attractive to patients (though quality is potentially endogenous if low demand reduces quality due to learning-by-doing effects). Similarly, private hospital availability is assumed to reduce demand for public treatment: smaller access costs to the private sector will induce some patients to switch from the public to the private sector (Martin & Smith, 2003; Martin et al., 2007).

The supply of (publicly funded) elective surgery is assumed to be determined by waiting time and local resources:

$$S = g(\text{waiting time}(+), \text{capacity}(+)). \tag{2}$$

We assume that long waits induce the provider to increase the supply, for given level of inputs, for both altruistic and non-altruistic motivations. Doctors may be willing to work harder when waiting times are longer because they care about the patients. Waiting times are regularly used as performance indicators or targets for public providers (and for private providers treating publicly funded patients). When waiting times are longer, hospitals with a higher proportion of patients waiting longer than expected may be under tighter scrutiny from the regulator (Lindsay & Feigenbaum, 1984; Propper et al., 2008; Siciliani & Iversen, 2012). Longer waits may also reduce idle capacity due to random patient arrivals and therefore increase efficiency and the number of patients treated, though this effect is likely to be modest when waiting times are generally long (Iversen, 1997; Siciliani, Stanciole, & Jacobs, 2009). Finally, the supply of care in a region is a function of its inputs, such as the number of available beds in publicly funded hospitals and their personnel, which determine the overall capacity.

3 | INSTITUTIONAL BACKGROUND AND DATA

The Italian healthcare system is publicly funded with hospitals reimbursed by Diagnosis Related Group according to volumes performed. The system is decentralised: Italy is divided in 19 regions and two autonomous provinces (Trento and Bolzano). The Italian National Health System was founded in 1978, provides full coverage to every citizen, and is funded through national and regional taxation. In 2001, the Constitutional reform gave regions the freedom to choose the type of healthcare model, generating great variability in institutional arrangements across regions.

Every region can decide its own organisational and regulatory scheme for public and private sector, how to allocate resources, define prevention and budgetary policies, strategic plans (e.g., building new hospitals), and elective admission rules. To avoid excessive territorial disparities, the Italian Ministry of Health sets the Essential Levels of Assistance, which are minimum healthcare requirements that regions have to provide, whose compliance is annually verified by the national government. Heterogeneity in regional policies has emerged in relation to waiting times (Fattore, Mariotti, & Rebba, 2013) driven by differences in co-payment schemes, unified booking centres, and promotion of private health insurance, providing a fragmented framework with regional disparities.

There are similarities but also differences between the Italian and the English National Health Service. In England, to which most of the empirical literature refers to, hospitals are also paid by a Diagnosis Related Group-type payment system (known as Healthcare Resources Groups) and patients have choice of hospital. Patients are also heavily insured with no co-payments for surgery or specialist visits and some co-payments for drugs. In both countries, there are exemptions for persons with disabilities or chronic conditions, pregnant women, elderly, and children (Paris, Devaux, & Wei, 2010). There are more pronounced differences between England and other health systems in the U.K. (e.g., Scotland and Wales) but arrangements vary less across different regions within England. In this respect, Italy has much more pronounced differences across regions in organisational arrangements and regulatory schemes.

The proportion of private health expenditure is similar for both countries. The share of public health expenditure out of total health expenditure (at the beginning of our period of observation) was about 76.5% for Italy and 81.7% for the U.K. Although the public–private mix is similar on the funding side, this is not the case on the hospital provision one. 96% of acute care beds in the U.K. is provided by public hospitals, but this is only 81.5% in Italy where 16.7% of the total number of acute beds is provided by not-for-profit private hospitals (Paris et al., 2010).

In this paper, we use information on waiting times provided by the Italian Ministry of Health's Statistical Office.[1] Waiting times are available for 19 regions and the two autonomous provinces for several procedures during the period 2010–2014. Waiting times are calculated for elective publicly funded patients who receive treatment in a public or private hospital. They are published annually in the Hospital Discharges Report (HDR) by the Ministry of Health.

Waiting times are defined as the number of days elapsed between the time the patient has been added to a hospital waiting list for elective surgery and the day the patient is admitted to the hospital to receive the treatment. From the same source and for

[1]Data are publicly available on the Italian Ministry of Health website (www.salute.gov.it) under Section "Temi e professioni" (Figures), subsections "Assistenza, ospedale e territorio" (Assistance, hospitals and local areas), "Ricoveri Ospedalieri" (Hospital discharges) and it is possible to select and download annual reports and data:

www.salute.gov.it/portale/temi/p2_6.jsp?lingua=italiano&id=1237&area=ricoveriOspedalieri&menu=vuoto

We use the original data and no data cleaning was performed apart for the exclusions mentioned in this section.

each year, region, and procedure, we collect data on hospital utilisation, that is, the total number of elective and emergency discharges. Hospital utilisation rates are computed for each procedure as the ratio of the total number of discharges to the regional population in a given year. Hospital utilisation also refers to publicly funded patients regardless of the type of provider (public or private) in which they receive treatment. The annual report for hospital discharges refers to patients treated in public hospitals and from the same source of data we calculate within each combination of region and year the overall share of patients treated in public hospitals paying with their own resources. The share is small and on average only about 2%, which is in line with findings as in Vittadini, Moscone, and Tosetti (2012).

We use data on waiting time for elective surgical (as opposed to medical) treatments because only these are available from administrative sources and are used as hospital targets. The 10 procedures included in the HDR are as follows: prostatectomy, breast cancer, colon cancer, uterus cancer and lung cancer surgeries, coronary bypass, percutaneous transluminal coronary angioplasty, carotid endarterectomy, hip replacement, and tonsillectomy.[2] We exclude tonsillectomy because regions show heterogeneous clinical attitudes and protocols which in turn reduce comparability across regions (Materia et al., 2005; see also national guidelines provided by the Italian Institute of Health for this clinical area[3]). We also compute the proportion of emergency discharges as the ratio between the number of emergency discharges and the total number of discharges by procedure, region, and year. In summary, the HDR data used in this analysis vary along three dimensions: surgical procedure, region, and year.

Control variables are obtained from demographic indicators available from the Italian National Institute of Statistics, which vary only by year and region, not by procedure. They include number of residents, age distribution in the regional population, and age and sex adjusted mortality rates. From the age distribution of residents, we calculate the proportion of population over 60 years old. We use risk-adjusted mortality rates and proportion of the elderly as need indicators.

As measure of local resources in the supply equation, we measure the capacity of private and public providers within each region. These are measured as (a) the total number of acute care beds in *public* and *private* hospitals (standardised by the number of residents) and (b) the ratio between beds in private hospitals and total number of beds within each region. Private hospitals treat both publicly and privately funded patients, and our data do not allow to make a distinction whether the treatment is paid by the NHS or privately. Because regulatory policies vary across regions in relation to reimbursements to private providers, it is not possible to identify the number—or the proportion—of publicly funded patients who are treated by the private sector. Therefore, private hospitals contribute to the capacity available to publicly funded patients. Variable (b) measures the public–private mix in provision in each region.

On the demand equation, we use control variables from the annual National Survey on Householders' Lifestyles and compute the proportion of regional population smoking more than 11 cigarettes per day on the total number of smokers as a proxy of unhealthy behaviour. We use the C-section rate as a proxy of poor appropriateness of care, which is a form of quality. This is computed as the total number of C-section deliveries to the total number of births within each region, which is provided in the HDR by the Italian Ministry of Health. C-section rates have been used by international organisations (OECD, 2015; WHO, 2015) as markers of appropriateness of care in health system performance. High C-section rates (on total births) are positively associated with complications and maternal and infant morbidity. According to the OECD, Italy has a surprisingly high C-section rate compared to other OECD countries although there are marked differences across regions, which we exploit in our analysis.

To measure the availability of private supply to privately funded patients, which could potentially reduce demand for public services, we measure the number of acute care beds in *private* hospitals (standardised by the number of residents).[4] Ideally, we would have liked to measure the number of private hospital beds available to privately funded patients, therefore excluding beds available to publicly funded patients. Information on private beds is however available only at hospital level, and is not split between publicly- and privately funded patients.

We do not have information on the fraction of patients who die on the waiting list. However, the proportion of patients who are likely to die while on the waiting list is negligible for most of the elective procedures (e.g., hip replacement). Even for most of the more serious conditions (e.g., cancer), elective patients experiencing a worsening of their health status are treated quickly or as emergencies.

[2]Other six procedures have been added in 2011 but there are consistency issues across regions, which prevented their use here.

[3]http://www.snlg-iss.it/pubblico_tonsillectomia_adenoidectomia

[4]Beds refer to the number of beds available in each Region on the 1st of January of each year. We only considered beds for elective patients, thus excluding beds for day cases and day surgeries. Source: Ministero della Salute - Dipartimento della programmazione e dell'ordinamento del Servizio sanitario nazionale - Direzione generale del sistema informativo e statistico sanitario;

4 | ECONOMETRIC SPECIFICATION

We use linear models to estimate the impact of waiting times on the demand for and supply of surgical treatments. We estimate separate models for demand and supply. We assume that the system is in equilibrium and that demand y^d_{irt} in Equation 1 equates supply y^s_{irt} in Equation 2, so that $y^d_{irt} = y^s_{irt} = y_{irt}$.

The empirical specification of the *demand* equation is

$$y_{irt} = \alpha + w^{'}_{irt}\beta + x^{'}_{irt}\gamma + z^{'}_{rt}\delta + h_t + h_i + \varepsilon_{irt}, \tag{3}$$

where subscript i indicates the type of elective surgery (e.g., hip replacement and surgeries for breast cancer, with, $i = 1, \ldots, I$), r the region (with $r = 1, \ldots, R$), t the year (with $t = 2010, \ldots, 2014$). Utilisation rate (y_{irt}) and waiting time (w_{irt}) are log transformed so that the key coefficient of interest (β) can be interpreted as the elasticity of demand with respect to waiting time. Utilisation rates are the total number of discharges for a given surgical procedure in a region and year standardised by population (the total number of residents, in thousands) of the region in the same year.

The vector x_{irt} includes control variables that vary over time, procedures, and regions, such as the proportion of emergency discharges. The vector z_{rt} includes variables that vary only over time and region and, in the demand equation, it includes the proportion of residents over 60, smoking prevalence, age and sex adjusted mortality rates at time $t-1$, the number of private beds per capita, and the C-section rate as indicator of poor quality.

The empirical model also includes time dummies h_t to capture common time trends and surgical procedure dummies h_i to control for differences in waiting times by procedure that among other factors reflect different degree of urgencies (e.g., cancer patients waiting less than hip replacement patients). We therefore exploit variations of waiting times *across* regions pooled across several years, controlling for the type of procedure, to identify the effect of waiting on demand. ε_{irt} is the error term.

The empirical specification of the *supply* of elective surgery (y_{irt}) is analogous to Equation 3 but uses a different set of controls (x_{irt}, z_{rt}). x_{irt} includes the proportion of emergency discharges. z_{rt} includes the per capita number of acute beds in public and private hospitals and the proportion of beds in private hospitals on the total amount of available beds within each region.

Because average waiting times in Italy are relatively short (about one month), we model demand for elective care as contemporaneously responding to waiting time, given the yearly frequency of the data used in the empirical analysis. We model the relation between supply and waiting time also as simultaneous because providers can quickly react to waiting time that they observe with no time lag and are also aware that waiting time are annually assessed by the Ministry of Health.

The OLS estimation of Equation 3, which again is estimated separately for the demand and supply equation, might produce a biased and inconsistent estimate of the coefficient of interest β. As mentioned above, longer waiting times may reduce demand for public treatment (because some patients opt for swifter private treatment) and also increase the supply of public treatments (due to targets or altruistic motives): therefore, waiting times are endogenous and have a simultaneous effect on both demand for and supply of treatment. Following previous literature (e.g., Martin & Smith, 1999, 2003), we instrument waiting time in the supply equation with a selection of exogenous demand shifters and we instrument waiting time in the demand equation with a selection of exogenous supply shifters. As the latter proved to be weak instruments, we also used the lag of waiting time as an instrument in demand models.[5]

To eliminate the remaining time-invariant unobservable factors, for example, at regional level, that might simultaneously affect the dependent variable as well as the controls in the regression, we also estimated FD models. This specification will, for example, control for any time-invariant regional factor (e.g., proportion of individuals holding private insurance, which is unlikely to vary quickly over time). We prefer the FD models over the (region) fixed-effect models, because the latter require the strong exogeneity assumption,[6] which is violated when we use the lag of waiting time as an instrument. Although the FD models are less efficient of fixed-effects ones, they only require weak exogeneity (i.e., that there is no feedback from the idiosyncratic shocks today to a covariate or an instrument tomorrow). The first difference version of the previous model is

$$\Delta y_{ir\ t,t-1} = \beta_1 \cdot \Delta w_{ir\ t,t-1} + \Delta x^{'}_{ir\ t,t-1}\gamma + \Delta z^{'}_{r\ t,t-1}\delta + h_t - h_{t-1} + \Delta\varepsilon_{ir\ t,t-1}, \tag{4}$$

We estimate an analogous model for the supply (i.e., $\Delta y^s_{ir\ t,t-1}$).

[5]This instrument is used also in Martin and Smith (2003). Admittedly, this is not an ideal instrument because persistency of waiting time over the years can cause the error term in the base equation to remain correlated, to some extent, with the instrument.

[6]The within transformation of error term ($\varepsilon_{it} - \overline{\varepsilon}_i$) and of the log of waiting times ($w_{it} - \overline{w}_i$) are correlated through their means.

To control for endogeneity caused in the demand model by the presence of waiting time, we instrument $\Delta w_{ir\ t,t-1}$ with $\Delta w_{ir\ t-2,t-3}$. This is a valid instrument because

$$Cov\left(\Delta\varepsilon_{ir\ t,t-1}, \Delta\varepsilon_{ir\ t-2,t-3}\right) = Cov((\varepsilon_{irt}-\varepsilon_{ir\ t-1}), (\varepsilon_{ir\ t-2}-\varepsilon_{ir\ t-3})) = 0.$$

To check the validity of the instruments, we use the *F* statistic on the excluded instruments, both for robust and cluster standard errors, under the null of weak instruments. Following Stock, Wright, and Yogo (2002), we conclude that instruments are valid if the *F* statistic is larger than 10.

5 | DESCRIPTIVE STATISTICS

We use data from 19 Italian regions and two autonomous provinces for 5 years and nine surgical procedures. Table 1 reports some descriptive statistics for variables entering the supply and demand equations, respectively. Due to the presence of lagged variables in estimation procedure, we report descriptive statistics for 4 years of observation (732 units—left-hand side of Table 1) and for the last 2 years (366 units—right-hand side of Table 1), which are respectively the maximum and minimum sample size used, depending on whether the lagged values of waiting time is used as instrument in the FD estimation.

On average, the per capita utilisation is of about 0.6 procedures per thousand residents, of which about 22% is emergency discharges. Waiting time is about 31 days across all procedures. The number of total beds for acute care per thousand residents is about 2.84, whereas the availability of private beds per thousand residents is 0.48. The proportion of population over 60 years old is about 28%, and the proportion of smokers who smoke more than 11 cigarettes per day is close to 40%. C-section rates are on average around 35%, and adjusted mortality rate per thousand residents at time $t-1$ is about 1%. The summary statistics are similar across the two samples used in the analysis.

Figure 1 shows the average waiting time and utilisation rate for different procedures, at the beginning and at the end of the study period. It suggests that there is larger variability across treatments than over time, with hip replacement procedures having the longest wait. Lung and uterus cancer surgeries, percutaneous transluminal coronary angioplasty exhibit the shortest wait. Figure 2 shows that waiting times and utilisation rates exhibit high variability across regions both at the beginning and at the end of the period.

6 | EMPIRICAL RESULTS

In Table 2, we report the results for the demand equation. We first estimated the model in Equation 3 using OLS over the period 2011–14. The model includes year- and procedure-fixed effects and is reported in the first column. In line with previous literature (e.g., Martin & Smith, 2003), we first considered a cross-sectional specification, exploiting variability across procedures, regions, and years. Given the likely endogeneity of waiting times, we estimate the same model with IV, using as instrument the lagged value of waiting times. The model is reported in the second column. Because of the inclusion of the 1-year lagged value

TABLE 1 Descriptive statistics

	2011–2014			2013–2014		
	Obs.	***M***	***SD***	**Obs.**	***M***	***SD***
Utilisation rate (per 1,000 residents)	732	0.60	0.65	366	0.61	0.67
Waiting time (days)	732	31.33	0.55	366	30.87	0.53
Demand shifters						
Emergency discharges (%)	732	21.64	22.44	366	22.86	22.95
C-section rates (%)	732	35.09	9.45	366	34.77	9.24
Heavy smokers (%)	732	39.59	4.59	366	38.46	4.62
Population over 60 years old (%)	732	27.75	2.66	366	28.00	2.60
Private beds (per 1,000 residents)	732	0.48	0.28	366	0.46	0.27
Mortality rate (per 1,000 residents, at $t-1$)	732	10.29	1.26	366	10.21	1.23
Supply shifters						
Emergency discharges (%)	732	21.64	22.44	366	22.86	22.95
Beds (public and private, per 1,000 residents)	732	2.84	0.29	366	2.73	0.28
Private beds (%)	732	0.17	0.10	366	0.17	0.10

Note. Our calculations using ISTAT and HDR data.

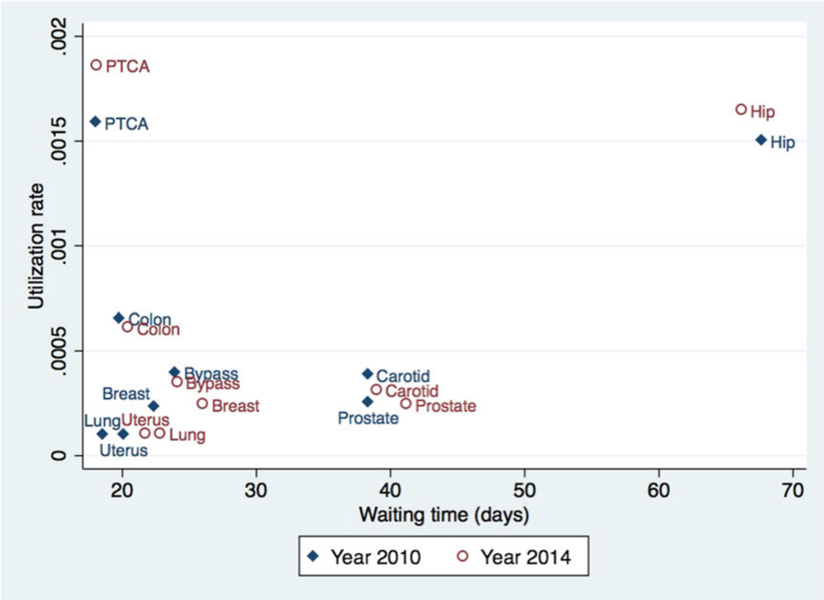

FIGURE 1 Utilisation rate and waiting time across procedures in 2010 and 2014. Source: Our calculations over HDR data. Notes: procedures considered are breast cancer (breast), prostatectomy (prostate), colon cancer (colon), uterus cancer (uterus) and lung cancer (lung) surgeries, coronary bypass (bypass), percutaneous transluminal coronary angioplasty (PTCA), carotid endarterectomy (carotid), hip replacement (hip) and tonsillectomy (ton) [Colour figure can be viewed at wileyonlinelibrary.com]

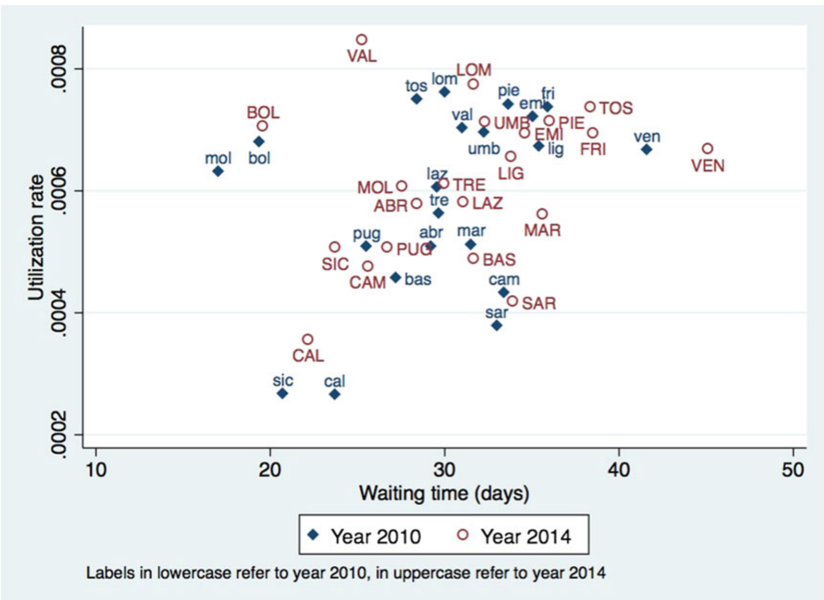

FIGURE 2 Utilisation rate and waiting time across regions in 2010 and 2014. Source: Our calculations over HDR data. Notes: Regions considered are Abruzzo (Abr), Basilicata (Bas), Autonomous Province of Bolzano (Bol), Campania (Cam), Calabria (Cal), Emilia-Romagna (Emi), Friuli Venezia Giulia (Fri), Lazio (Laz), Lombardia (Lom), Liguria (Lig), Marche (Mar), Molise (Mol), Piemonte (Pie), Puglia (Pug), Sardegna (Sar), Sicilia (Sic), Toscana (Tos), Autonomous Province of Trento (Tre), Umbria (Umb), Veneto (Ven) [Colour figure can be viewed at wileyonlinelibrary.com]

of waiting time among controls of the demand model, we lose the first year of observations. Hence, we omit the first year from all estimation samples to maintain data consistency.

Estimation results for the cross-sectional analysis show a negative and significant at 10% level coefficient for waiting time. The OLS estimation suggests an elasticity of demand to waiting time of −0.1, which increases (in absolute value) to −0.15 once waiting time is instrumented. In the third and fourth column of Table 2, we estimate the same models but restrict the sample to the last 2 years, to compare the results with those obtained in the FD specification (see model 4 above). In this restricted, hence less informative sample, the elasticity of waiting time coefficient loses statistical significance but the magnitude of the coefficient remains similar. In the last two columns of Table 2, which uses FD estimation, we test whether the results for the demand equation are robust once controlling for time-invariant characteristics in the error term, which might be an additional source of endogeneity. The results show that the elasticity is equal to −0.04 using OLS and increases (in absolute value) to −0.24 with an IV estimation, which uses the 2-years-lagged difference in waiting times as an instrument, albeit statistical evidence is weak.

In relation to the control variables, we find that lower hospital quality (i.e., higher C-section rates) reduces demand. A higher proportion of emergency discharges and of heavy smokers (unhealthy lifestyle), both reduce the demand. Because the FD

TABLE 2 Demand estimates

Variables	Pooled cross section 2011–14 OLS	Pooled cross section 2011–14 IV	Pooled cross section 2013–14 OLS	Pooled cross section 2013–14 IV	First difference 2013–14 OLS	First difference 2013–14 IV
Waiting time (log)	−0.102* (0.060)	−0.149* (0.086)	−0.045 (0.067)	−0.128 (0.105)	−0.04* (0.023)	−0.238* (0.142)
Emergency discharges (%)	−0.011*** (0.003)	−0.012*** (0.003)	−0.016*** (0.005)	−0.016*** (0.005)	−0.002** (0.001)	−0.004** (0.001)
C-section rates (%)	−0.018*** (0.004)	−0.018*** (0.004)	−0.015*** (0.004)	−0.016*** (0.004)	−0.006 (0.006)	−0.016 (0.051)
Proportion of over 60 years old (%)	0.018 (0.034)	0.021 (0.035)	−0.004 (0.039)	0.003 (0.041)	0.004 (0.045)	−0.012 (0.057)
Proportion of heavy smokers (%)	−0.018*** (0.005)	−0.018*** (0.005)	−0.014*** (0.005)	−0.015*** (0.005)	0.000 (0.002)	0.002 (0.002)
Private beds (per 1,000 residents)	0.047 (0.033)	0.050 (0.032)	0.041 (0.036)	0.047 (0.035)	0.054 (0.056)	0.05 (0.061)
Mortality rate (per 1.000)	0.029 (0.064)	0.024 (0.064)	0.057 (0.069)	0.049 (0.069)	−0.01 (0.019)	−0.028 (0.024)
Constant	6.078*** (0.509)	6.175*** (0.513)	6.002*** (0.576)	6.147*** (0.561)	−0.009 (0.013)	−0.002 (0.0159)
Observations	732	732	366	366	366	366
R^2	0.902		0.9		0.025	
First-stage F stat		196.7		116.3		11.15

Note. In cross-sectional specifications, we also include year and procedure dummy variables, whose estimated coefficients are omitted here. The instrument used in pooled cross-section models is the 1-year-lagged waiting time; in the first difference specification, it includes 2-years-lagged values of first differences. Standard errors are robust and clustered. IV = instrumental variables; OLS = ordinary least squares.

The dependent variable is utilisation rate and standard errors are given in parentheses. Significance levels:

*$p < 0.1$.

**$p < 0.05$.

***$p < 0.01$.

models control for time-invariant characteristics, the statistical significance of some control variables reduces. This is not the case for the emergency admission rate, which remains highly statistically significant.

The F statistics for the instrument used in the first-stage regression show that the instrument is valid, even when FD models are used. In Table A1, we report first-stage regressions, which show positive associations of waiting time over years, and of waiting time with population over 60, whereas the associations between waiting time and share of emergency discharges and C-section rate tend to be negative. In the FD model (Table A1, third column), we find a negative coefficient of lagged waiting times, which is negative by construction as the autocorrelation coefficient for the change of waiting time is positive.[7]

Table 3 contains results for the supply equation. Again, it presents first the OLS and then the IV estimation for each model. In the first two columns, we present the pooled cross-sectional specification over the period 2011–14. Estimates show an elastic supply of elective surgery to waiting time only when waiting time is instrumented. Here, we use as instrument the proportion of population over 60 years old, which is a key driver of demand and exogenous to supply. The following two columns show that results are qualitatively similar when only observations in 2013–14 are used. In all pooled cross-sectional models, the number of beds in public and private hospitals, measuring local endowments, has a positive and significant coefficient, suggesting that the higher capacity increases supply. A higher rate of emergency discharges reduces the supply for elective interventions, because emergency discharges require more stand-by capacity. The results are robust to the use of alternative instruments, including C-section rates. All instruments are valid in the cross-sectional models, and first-stage results are presented in Table A2.

In the last four columns of Table 3, we report estimation results for FD estimates, which controls for time-invariant characteristics. The waiting time coefficients are still positive but smaller and not statistically significant. Given the aggregate nature of our data, the lack of statistical significance might be due to the loss of information caused by the first difference transformation. Moreover, all instruments used are weak and vary only by region and year.

[7]This follows from computing the correlation coefficient of $(\Delta w_{t,t-1}, \Delta w_{t-2,t-3})$, which has opposite sign with respect to the correlation coefficient of (w_t, w_{t-1}). Intuitively, assuming for simplicity, a simple autoregressive model for our first-stage regression, the sign of the autocorrelation coefficient, $\rho(\Delta w_{t,t-1}, \Delta w_{t-2,t-3})$ is given by the sign of its numerator, that is, the covariance $C(\Delta w_{t,t-1}, \Delta w_{t-2,t-3})$. The latter can be written as $C(\Delta w_{t,t-1}, \Delta w_{t-2,t-3}) = C(w_t, w_{t-2}) - C(w_t, w_{t-3}) - C(w_{t-1} . w_{t-2}) + C(w_{t-1} . w_{t-3})$, and under weak stationarity (i.e. $C(w_t, w_{t-h}) = q$ for all integer values of t,h), $C(\Delta w_{t,t-1}, \Delta w_{t-2,t-3}) = -q(q-1)^2$. This implies that for a positive autocorrelation coefficient $q > 0$, from which follows that the sign of $\rho(\Delta w_{t,t-1}, \Delta w_{t-2,t-3})$ is negative.

TABLE 3 Supply estimates

	Pooled cross section 2011–14		Pooled cross section 2013–14		First difference 2011–14		First difference 2013–14	
Variables	**OLS**	**IV**	**OLS**	**IV**	**OLS**	**IV**	**OLS**	**IV**
Waiting time (log)	0.010	1.069***	0.079	0.718***	0.025	0.627*	−0.036	0.176
	(0.057)	(0.354)	(0.057)	(0.237)	(0.045)	(0.360)	(0.022)	(0.206)
Emergency discharges (%)	−0.010***	−0.001	−0.018***	−0.016**	0.001	0.003	−0.002**	−0.000
	(0.003)	(0.006)	(0.006)	(0.007)	(0.002)	(0.003)	(0.001)	(0.002)
Beds (public and private, per 1,000 residents)	1.567***	1.313***	1.408***	1.124***	0.018	0.382	−0.085	0.009
	(0.314)	(0.416)	(0.352)	(0.386)	(0.220)	(0.373)	(0.193)	(0.233)
Private beds (%)	−0.399*	−0.203	−0.116	0.092	0.100	−0.379	0.367	0.591
	(0.237)	(0.362)	(0.213)	(0.287)	(0.696)	(1.130)	(0.461)	(0.556)
Constant	−7.041***	−8.416**	−5.902**	−5.732*	0.006	0.002	−0.013	−0.011
	(2.537)	(3.323)	(2.811)	(2.981)	(0.025)	(0.013)	(0.010)	(0.011)
Observations	732	732	366	366	732	732	366	366
R^2	0.887		0.896		0.001		0.021	
First-stage F stat		14.04		18.90		4.439		2.622

Note. In cross-section specifications we also include year and procedure dummy variables, whose estimated coefficients are omitted here. The instrument set in pooled cross-section models includes proportion of population over 60 years old; in the first-difference specifications it also includes C-section rates. Robust and clustered standard errors computed. IV = instrumental variables; OLS = ordinary least squares.

The dependent variable is utilisation rate and standard errors are given in parentheses. Significance levels:

***$p < 0.01$.

**$p < 0.05$.

*$p < 0.1$.

Other factors should ideally be included in the demand equation, such as (average) distance to the hospital or co-payments, which are likely to deter some patients and reduce demand. A variable capturing the average distance would require detailed access to patient level data and geographical coordinates between patients' residence and hospital address. We conjecture the bias caused by the omission of these variables is likely to be negligible in our FD model, which controls for time-invariant factors including regional effects. The average distance to hospital is unlikely to have changed significantly over time.

We check the robustness of our findings on the demand side excluding the variable that is related to smoking prevalence. Coefficients and significance levels are similar to the one presented in base model at the cost of losing estimate precision in the waiting time coefficient for the pooled IV cross section (see Table A3).

As a robustness check of the supply equation estimation, we also measure the per capita cost of medical staff (wages) in public hospitals as an additional input in the production function of publicly funded treatments. The results are very similar in terms of magnitude and significance of coefficients (see Table A4), though the power of the instruments is marginally reduced due to the high correlation between wages and beds (equal to 0.9).

The small sample size does not allow us to perform sensitivity analysis by intervention type. Nonetheless, we test the robustness of our results by excluding from the sample procedures with more than 40% of emergency discharges (i.e., percutaneous transluminal coronary angioplasty whose fraction of emergency discharges is 67%), and we found similar pointwise estimates, though with some precision loss (results omitted but available from the authors).

7 | CONCLUSIONS AND POLICY IMPLICATIONS

We have used administrative data on waiting times and volume of elective surgeries across different procedures and regions in Italy over the period 2010–2014. Our key finding suggests that the demand for elective surgery is inelastic to waiting times, and the elasticity is in the range of −0.15 and −0.24. This is in line with the literature in England (Martin & Smith, 1999, 2003; Martin et al., 2007), which provides a comparable demand elasticity of −0.2. The statistical significance of our results is however weaker, and this is likely to be due to our use of aggregated data at regional level as opposed to electoral ward and a shorter time series used in the English context (Martin & Smith, 2003; Martin et al., 2007).

The similar elasticity between Italy and the U.K. could be the result of the similarities between health systems (both with a National Health Service, similar public–private funding mix, and financial arrangements for hospitals). But there are also differences (such as the differences in public–private mix in hospital provision) and a priori elasticities could have been different. The importance of different institutional arrangements across health systems is indeed confirmed by the different demand elasticity for Australia, which has been estimated at −1.7 (Stavrunova & Yerokhin, 2011). This is because, in Australia, the public–private mix is very different on the funding side, with more than half of the population holding private health insurance and therefore more susceptible to switch from the public to the private sector when waiting times are longer.

Although our study suggests a weak effect of waiting times on demand, the results have important policy implications in relation to the effectiveness of policy initiatives that encourage an expansion in supply to reduce waiting times (through more funding, an extension of working hours, revision of contracts, contracting out to existing private providers, etc.). Some policymakers have argued that such supply-side policy initiatives can be ineffective because an increase in supply can be offset by large increases in demand (Siciliani & Hurst, 2005). Whether there are merits to this argument depends critically on the demand elasticity to waiting times. Our results show that within the Italian institutional context, the demand is inelastic. Therefore, policies aimed at increasing supply would be effective in reducing waiting times.

At times of great financial pressure following the economic crisis, governments have introduced or are introducing measures to keep health expenditure under control, which is likely to imply a reduction or a slower growth in supply. Driven by the ageing population and technology, the gap between demand and supply may increase and, based on our findings, so will waiting times and waiting lists. Governments therefore need to consider policy interventions that act on the demand, for example, by reducing unnecessary referrals through better coordination between general practitioners and specialists (Mariotti, Siciliani, Rebba, et al., 2014) or improving the prioritisation of the list to minimise the impact of delays (Siciliani et al., 2013).

Health systems differ to a great extent across the OECD countries on funding, provision, and organisational arrangements. As data on waiting times become increasingly available, future work could replicate our analysis in other health systems to inform the policy debate on supply-side initiatives aimed at reducing waiting times.

REFERENCES

Fabbri, D., & Monfardini, C. (2009). Rationing the public provision of healthcare in the presence of private supplements: Evidence from the Italian NHS. *Journal of Health Economics*, *28*(2), 290–304.

Fattore, G., Mariotti, G., & Rebba, V. (2013). “Italy” in Siciliani, L., Borowitz, M., Moran, V., 2013. Waiting time policies in the health sector: What works? In *OECD Health Policy Studies*OECD Publishing.

Gravelle, H., Dusheiko, M., & Sutton, M. (2002). The demand for elective surgery in a public system: Time and money prices in the UK National Health Service. *Journal of Health Economics*, *21*(3), 423–449.

Gravelle, H., Smith, P., & Xavier, A. (2003). Performance signals in the public sector: The case of health care. *Oxford Economic Papers*, *55*(1), 81–103.

Iversen, T. (1997). The effect of a private sector on the waiting time in a national health service. *Journal of Health Economics*, *16*(4), 381–396.

Lindsay, C. M., & Feigenbaum, B. (1984). Rationing by waiting lists. *The American Economic Review*, *74*(3), 404–417.

Mariotti, G., Siciliani, L., Rebba, V., et al. (2014). Waiting time prioritisation for specialist Services in Italy: The homogeneous waiting time groups approach. *Health Policy*, *117*(1), 54–63.

Martin, S., Rice, N., Jacobs, R., & Smith, P. C. (2007). The market for elective surgery: Joint estimation of supply and demand. *Journal of Health Economics*, *26*(2), 263–285.

Martin, S., & Smith, P. C. (1999). Rationing by waiting lists: An empirical investigation. *Journal of Public Economics*, *71*, 141–164.

Martin, S., & Smith, P. C. (2003). Using panel methods to model waiting times for National Health Service surgery. *Journal of the Royal Statistical Society*, *166*(Part 2), 1–19.

Materia, E., Baglio, G., Bellussi, L., Marchisio, P., Perletti, L., Pallestrini, E., & Calia, F. (2005). The clinical and organisational appropriateness of tonsillectomy and adenoidectomy—An Italian perspective. *International Journal of Paediatric Otorhinolaryngology*, *69*(4), 497–500.

OECD. (2015). *Health at a Glance 2015: OECD indicators*. Paris: OECD Publishing.

Paris, V., Devaux, M., Wei, L. (2010). Health systems institutional characteristics: A Survey of 29 Oecd Countries. OECD Health Working Papers, No.50, OECD Publishing, Paris.

Propper, C., Burgess, S., & Gossage, D. (2008). Competition and quality: Evidence from the NHS internal market 1991–9. *The Economic Journal*, *118*(525), 138–170.

Siciliani, L., Borowitz, M., & Moran, V. (2013). *Waiting time policies in the health sector: What works?* OECD Health Policy Studies: *OECD Publishing*, Paris.

Siciliani, L., Borowitz, M., & Moran, V. (2014). Measuring and comparing health care waiting times in OECD countries. *Health Policy*, *118*(3), 292–303.

Siciliani, L., & Hurst, J. (2005). Tackling excessive waiting times for elective surgery: A comparison of policies in 12 OECD countries. *Health Policy*, *72*(2), 201–215.

Siciliani, L., & Iversen, T. (2012). Waiting times and waiting lists. In *The Elgar companion to health economics* (Second ed.) chapter 24.

Siciliani, L., Stanciole, A., & Jacobs, R. (2009). Do waiting times reduce hospital costs? *Journal of Health Economics*, *28*(4), 771–780.

Stavrunova, O., & Yerokhin, O. (2011). An equilibrium model of waiting times for elective surgery in NSW public hospitals. *The Economic Record*, *87*(278), 384–398.

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics, American Statistical Association*, *20*(4), 518–529.

Vittadini, G., Moscone, F., & Tosetti, E. (2012). Social interaction in patients' hospital choice: Evidence from Italy. *Journal of the Royal Statistical Society*, *175*(2), 453–472.

World Health Organisation (2015). WHO statement on caesarean section rates, WHO/RHR/15.02.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Riganti A, Siciliani L, Fiorio CV. The effect of waiting times on demand and supply for elective surgery: Evidence from Italy. *Health Economics*. 2017;26(S2):92–105. https://doi.org/10.1002/hec.3545

APPENDIX A

Additional results

TABLE A1 First-stage estimates; demand equation

Variables	Pooled cross section 2011–14	Pooled cross section 2013–14	First difference 2013–14
Waiting time (log) at t-1	0.653*** (0.047)	0.610*** (0.057)	
Waiting time (log) at t-2			−0.181*** (0.054)
Emergency discharges (%)	−0.005*** (0.002)	0.002 (0.003)	−0.007*** (0.002)
C-section rate (%)	0.001 (0.002)	−0.005** (0.002)	−0.025* (0.014)
Proportion of heavy smokers (%)	−0.000 (0.002)	0.004 (0.003)	0.008** (0.004)
Proportion of over 60 years old (%)	0.040*** (0.013)	0.052*** (0.017)	−0.075 (0.103)
Private beds (per 1,000 residents)	0.012 (0.015)	0.039** (0.018)	−0.042 (0.129)
Mortality rate (per 1,000 residents, at $t-1$)	−0.047* (0.024)	−0.063** (0.027)	−0.076* (0.044)
Constant	0.447* (0.250)	0.228 (0.329)	0.036 (0.030)
Observations	732	366	366
R^2	0.798	0.824	0.109

In all specifications we also include year and procedure dummy variables, whose estimated coefficients are omitted here. Robust and clustered standard errors computed. The dependent variable is waiting time and standard errors are given in parentheses. Significance levels:

***$p < 0.01$.

**$p < 0.05$.

*$p < 0.1$.

TABLE A2 First-stage estimates; supply equation.

	Pooled cross section	Pooled cross section	First difference	First difference
	2011–14	2013–14	2011–14	2013–14
Proportion of over 60 years old (%)	0.040*** (0.011)	0.050*** (0.011)	0.050 (0.059)	−0.045 (0.106)
C-section rate (%)			−0.028*** (0.011)	−0.033** (0.015)
Emergency discharges (%)	−0.010*** (0.003)	−0.003 (0.003)	−0.005 (0.003)	−0.008*** (0.002)
Beds (public and private, per 1,000 residents)	−0.040 (0.243)	0.148 (0.222)	−0.586 (0.503)	−0.243 (0.463)
Private beds (%)	0.171 (0.254)	0.138 (0.258)	1.471 (1.038)	0.030 (1.245)
Constant	2.399 (1.881)	0.626 (1.674)	−0.024 (0.035)	0.002 (0.035)
Observations	732	366	732	366
R^2	0.590	0.637	0.037	0.062

The dependent variable is waiting time and standard errors are given in parentheses. Significance levels:

***$p < 0.01$.

**$p < 0.05$.

*$p < 0.1$.

In all specifications, we also include year and procedure dummy variables, whose estimated coefficients are omitted here. The instrument set in pooled cross-section models includes proportion of over 60 years old in the population; in first-difference specifications, it also includes low hospital quality measured by C-section rate. Robust and clustered standard errors computed.

TABLE A3 Robustness check: Demand estimates without fraction on heavy smokers among demand controls.

	Pooled cross section Estimation period: 2011–14		Pooled cross section Estimation period: 2011–14		First difference Estimation period: 2013–14	
Variables	**OLS**	**IV**	**OLS**	**IV**	**OLS**	**IV**
Waiting time (log)	−0.077 (0.059)	−0.101 (0.085)	−0.028 (0.067)	−0.07 (0.103)	−0.040* (0.023)	−0.243* (0.146)
Emergency discharges (%)	−0.012*** (0.003)	−0.012*** (0.003)	−0.017*** (0.005)	−0.017*** (0.005)	−0.002** (0.001)	−0.004** (0.001)
C-section rates (%)	−0.015*** (0.004)	−0.015*** (0.004)	−0.013*** (0.005)	−0.014*** (0.005)	−0.006 (0.006)	−0.012 (0.008)
Proportion of over 60 years old (%)	0.051 (0.040)	0.054 (0.040)	0.021 (0.043)	0.026 (0.044)	0.005 (0.044)	−0.008 (0.049)
Private beds (per 1,000 residents)	0.047 (0.034)	0.049 (0.033)	0.050 (0.035)	0.053 (0.035)	0.054 (0.056)	0.051 (0.061)
Mortality rate (per 1,000 Residents, at $t-1$)	−0.024 (0.072)	−0.027 (0.072)	0.022 (0.075)	0.017 (0.075)	−0.010 (0.019)	−0.028 (0.025)
Constant	4.776*** (0.516)	4.807*** (0.510)	4.926*** (0.587)	4.980*** (0.571)	−0.010 (0.012)	−0.006 (0.014)
Observations	732	732	366	366	366	366
R^2	0.898		0.897		0.024	
First-stage F stat		196.6		109.9		10.62

In cross-sectional specifications, we also include year and procedure dummy variables, whose estimated coefficients are omitted here. The instrument used in pooled cross-section models is the 1-year-lagged waiting time; in the first difference specification, it includes 2-years-lagged values of first differences. Standard errors are robust and clustered. IV = instrumental variables; OLS = ordinary least squares.

The dependent variable is utilisation rate and standard errors are given in parentheses. Significance levels:

*$p < 0.1$.

**$p < 0.05$.

***$p < 0.01$.

TABLE A4 Robustness check: Supply estimates including wages per capita.

	Pooled cross section Estimation period: 2011–14		Pooled cross section Estimation period: 2013–14		First difference Estimation period: 2011–14		First difference Estimation period: 2013–14	
Variables	**OLS**	**IV**	**OLS**	**IV**	**OLS**	**IV**	**OLS**	**IV**
Waiting time (log)	−0.016 (0.056)	1.126** (0.443)	0.063 (0.057)	0.739*** (0.269)	0.022 (0.045)	0.596* (0.356)	−0.038* (0.022)	0.160 (0.193)
Emergency discharges (%)	−0.011*** (0.003)	−0.000 (0.007)	−0.018*** (0.006)	−0.016** (0.007)	0.001 (0.002)	0.003 (0.003)	−0.002**(0.001)	−0.001 (0.002)
Beds (public and private, per 1,000 residents)	1.778*** (0.333)	1.238** (0.489)	1.503*** (0.370)	1.092** (0.430)	−0.061 (0.235)	0.324 (0.398)	−0.107 (0.194)	−0.005 (0.235)
Private beds (%)	−0.465** (0.233)	−0.174 (0.394)	−0.165 (0.207)	0.110 (0.300)	0.416 (0.711)	−0.195 (1.181)	0.558 (0.491)	0.636 (0.543)
Wages (log)	−0.064*** (0.022)	0.019 (0.048)	−0.037 (0.026)	0.010 (0.035)	0.706** (0.328)	0.361 (0.517)	0.356 (0.316)	0.115 (0.418)
Constant	−8.290*** (2.617)	−8.105** (3.491)	−6.400** (2.887)	−5.599* (3.115)	0.005 (0.007)	0.002 (0.013)	−0.012 (0.010)	−0.011 (0.011)
Observations	732	732	366	366	732	732	366	366
R^2	0.890		0.897		0.005	−0.492	0.024	−0.189
First-stage F stat		9.310		14.31		4.106		2.904

In cross-section specifications we also include year and procedure dummy variables, whose estimated coefficients are omitted here. The instrument set in pooled cross-section models includes proportion of population over 60 years old; in the first-difference specifications it also includes C-section rates. Standard errors are robust and clustered. IV = instrumental variables; OLS = ordinary least squares.

The dependent variable is utilisation rate and standard errors are given in parentheses. Significance levels:

*$p < 0.1$.

**$p < 0.05$.

***$p < 0.01$.

APPENDIX B

DATA SOURCES

We download from Italian Ministry of Health website Annual report on Hospital Discharges for years 2010–2014. Data are free and the Ministry of Health does not require any special permission for access. Data are available in xls format.

For year 2014 the relevant file is:

"C_17_pubblicazioni_2396_ulterioriallegati_ulterioreallegato_0_alleg" and data on waiting times are in tables 3.10–3.10(5). We use the total number of discharges ("Totale dimisisoni") and the relative average of waited time (in days, "Attesa media in giorni").

For year 2013 the relevant file is:

"C_17_tavole_18_allegati_iitemAllegati_0_fileAllegati_itemFile_1_file" and data on waiting times are in tables 3.10–3.10(5). We use the total number of discharges ("Totale dimisisoni") and the relative average of waited time (in days, "Attesa media in giorni").

For year 2012 the relevant file is:

"C_17_tavole_16_allegati_iitemAllegati_0_fileAllegati_itemFile_1_file" and data on waiting times are in tables 3.10–3.10(5). We use the total number of discharges ("Totale dimisisoni") and the relative average of waited time (in days, "Attesa media in giorni").

For year 2011 the relevant file is:

"C_17_tavole_1_allegati_iitemAllegati_0_fileAllegati_itemFile_11_file" and data on waiting times are in tables 3.10–3.10(5). We use the total number of discharges ("Totale dimisisoni") and the relative average of waited time (in days, "Attesa media in giorni").

For year 2010 the relevant file is:

"C_17_pubblicazioni_1690_ulterioriallegati_ulterioreallegato_0_alleg" and data on waiting times are in tables 3.10–3.10(5). We use the total number of discharges ("Totale dimisisoni") and the relative average of waited time (in days, "Attesa media in giorni").