The Effect of Waiting Times on Demand and Supply for Elective Surgery: Evidence from Italy

Andrea Riganti\textsuperscript{1} Luigi Siciliani\textsuperscript{2} Carlo Fiorio\textsuperscript{1}

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 non-monetary prices to bring demand for and supply of health care in equilibrium. Using an administrative dataset from Italy in 2010-2014, we estimate demand and supply elasticities with respect to waiting times. Employing linear regression models with instrumental variables and fixed effects to deal with endogeneity of waiting times, we find that both demand and supply are inelastic to waiting times (estimated elasticity respectively equal to -0.29 for demand and 0.52 for supply). Our results have implications on the effectiveness of policies aimed at increasing supply and their ability to reduce waiting times.

Keywords: Waiting times; Elective surgery; Demand; Supply.

\textit{JEL}: I10

\textsuperscript{1}Departments of Economics, Management and Quantitative Methods, University of Milano, IT.

\textsuperscript{2}Departments of Economics and Related Studies, University of York.
1. Introduction

Waiting times in health care sector are a major health policy concern across many OECD countries (Siciliani, Borowitz and Moran, 2013). Waiting times for elective surgeries can last several months (Siciliani, Moran and Borowitz, 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 non-price rationing which brings together the demand for and the supply of treatments (see seminal papers by Lindsay and Feigenbaum, 1984; Martin and Smith, 1999). On the demand side, a higher 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 of 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 and Smith, 1999, for a theoretical model, and Propper et al, 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 only have 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 United Kingdom. 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 and Sutton, 2002; Gravelle, Smith and Xavier, 2003, and Martin, Jacobs, Rice and 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 and 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 based on differ institutional arrangements (gatekeeping system, use of user charges, payment arrangements) and funding levels. Administrative data on waiting times have been collected within the English NHS since its inception, but only in the last years in other countries (Siciliani, Moran and Borowitz, 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. To deal with endogeneity of waiting due to simultaneity of demand and supply we use an instrumental variables approach and to address the remaining endogeneity of waiting due to its possible correlation with unobserved variables, we introduce a large set of fixed-effects controlling for group-constant variables. Our key finding is that both demand and supply are inelastic to waiting times (estimated elasticity in the saturated model is \(-0.3\) for demand and \(0.5\) for supply).

As far as the authors are aware, this is the first study which uses administrative data to estimate demand and supply elasticities within the Italian context. We are only aware of another study which estimates demand elasticity for Italy (Fabbri and Monfardini, 2009). This study focuses on specialist consultations as opposed to elective surgeries. It makes use of survey in 2000 rather than recent administrative data, and the methodology and period covered is therefore different. We are also not aware of studies estimating demand and supply elasticities from other OECD countries (in addition to the UK) except for one study from Australia, which finds that demand is elastic and equal to \(-1.7\) (Stavrunova and Yerokhin, 2011).

The paper is structured as follows. In Section 2 we set out the theoretical model for estimation of demand and supply of elective surgeries in 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 discusses the findings and concludes.
2. Theoretical Model

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

\[ D = f(waiting\ time(-), private\ availability(-), need(+), treatment\ quality(+), management\ quality(+)) \]  

(1)

Demand is assumed to decrease in waiting times. Longer waiting times may induce some patients at the margin to opt out of the NHS and look for treatment in the private sector. Similarly, private availability is presumed to reduce demand: some patients may prefer to pay out-of-pocket and receive immediate care rather than waiting to be treated for free by the public sector. Alternatively, patients may also opt for less invasive pharmaceutical treatments. Demand will be higher in areas with higher need, e.g. areas with an older and sicker population, and in areas where the quality of healthcare is higher making hospital services more attractive for residents as well as for people residing in other regions (Martin and Smith, 2003, Martin et al., 2007). Finally, worse management quality, which for instance might cause patients to be allocated to the wrong department, will reduce demand.

The supply of elective care is assumed to be determined by waiting time and local resources:

\[ S = g(waiting\ time(+), local\ resources(+)) \]  

(2)

We hypothesise that long waits may induce an increase in supply. Doctors may be willing to work harder based on altruistic motives (Lindsay and Feigenbaum, 1984; Propper et al., 2008). When waiting times are longer hospitals with a higher proportion of patients waiting longer than expected may be under tighter scrutiny from the regulator (Propper et al., 2008; Siciliani and Iversen, 2012). Longer waits may also reduce idle capacity due to random patient arrivals, though this effect is likely to be modest when waiting times are generally long (Iversen, 1997; Siciliani, Stanciole and Jacobs, 2009). The supply of care in a region is a function of publicly-funded hospitals capacity, such as the number of available beds, doctors and nurses (i.e. local NHS endowments in capital and labour).
3. Institutional background and sources of data

The Italian health-care system is a public reimbursement DRG based type, since local authorities and hospitals receive money from central government according to volumes performed. The system is regional-based: Italy is divided in 19 regions and two autonomous provinces (Trento and Bolzano). The Italian National Health Systems (NHS) was settled in 1978 and grants all Italian citizens full coverage funded through national and regional taxation. In 2001 the Constitutional reform gave regions and autonomous provinces autonomy in the choice of the healthcare model producing large variability.

Every region can decide its own regulatory scheme for the public and the private sector, set reimbursements levels to hospital, allocate resources and define budgetary and prevention policies, strategic plans (e.g. building new hospitals) and elective admission rules. To avoid excessive territorial disparities, the Italian Ministry of Health set the Essential Levels of Assistance, which are minimum healthcare requirements that each region has to provide, whose compliance is annually verified by the national government. Large heterogeneity on waiting times regional policies has emerged (Fattore et al., 2013) and has been affected by regional differences in co-payment schemes, in unified booking centres and in the promotion of private health insurances, providing a fragmented framework with lots of regional disparities.

In this paper we use information on waiting times provided by the Italian Ministry of Health’s Statistical Office. Waiting times are available for each of the 19 regions and the 2 autonomous provinces (Trento and Bolzano) and for several treatments during the five-year period 2010-2014. They are published annually in the Hospital Discharges Report (HDR) by the Ministry of Health. Waiting times are calculated as the average number of days that elapsed between the time the patient has been added to a hospital waiting list for elective surgery and the day in which the patient is admitted to the hospital to receive the treatment. From the same source we collect data on hospital utilization rates for each region and in each year and procedure. These are computed as the ratio of the total number of discharges in each procedure and the regional population in a given year.

Here we only use data on elective surgical treatments since these only are available from administrative sources and are targets for providers. The ten procedures included in the HDR are: breast cancer, prostatectomy, colon cancer, uterus cancer and lung cancer surgeries, coronary bypass, percutaneous transluminal coronary angioplasty, carotid endarterectomy,
hip replacement and tonsillectomy. We exclude tonsillectomy since regions show heterogeneous clinical attitudes and protocols reducing comparability across regions (Matera et al., 2005. See also the national guidelines provided by the Italian Institute of Health for this clinical area). These nine surgical procedures represent on average about 7% of total discharges for elective patients in any given year of observation. All HDR data vary along three dimensions: surgical procedure, region and year.

Control variables are mostly obtained from demographic indicators available from ISTAT (the Italian National Institute of Statistics), which however vary only by year and region. They include number of residents and age distribution, standardized mortality rates in the regional population and unemployment rate. From the age distribution of residents we calculate the proportion of population over 60 and over 80 years old. A measure of local resources is obtained by the number of hospital beds in wards for elective surgeries in public and private and accredited hospitals, providing treatment funded by the NHS. We standardize the number of beds by the region’s population, which we refer to as hospital public bed rate. We compute expenditure for medical staff per capita (in Euros). This is computed from wages of medical staff working in hospitals. We compute the rate of beds in private hospitals as the standardized public bed rate to measure private healthcare availability, which can complement public healthcare provision. We aim to use this variable as proxy of private availability instead of private cost of surgeries. Lastly we calculate the share of total private beds within each region and use it as control for supply model.

We also use the annual National Survey on Householders’ Lifestyles, computing the proportion of regional population who smokes more than 11 cigarettes per day as a measure of unhealthy behaviour in the regional population over time.

---

1 Other six procedures have been added in 2011 without consistency across regions and have been ignored.
2 http://www.snlg-iss.it/pubblico_tonsillectomia_adenoidectomia
3 As beds we considered the number of beds available in each Region every Jan 1. We only considered beds for elective patients, thus excluding Day Hospitals and Day Surgery beds. Source: Ministero della Salute - Dipartimento della programmazione e dell’ordinamento del Servizio sanitario nazionale - Direzione generale del sistema informativo e statistico sanitario;
4 The classification of structures is the following: Azienda Ospedaliera (Hospital agency), Ospedale a gestione diretta (Hospital directly managed by Local Health Authority), Azienda Ospedaliera universitaria integrata con il Servizio Sanitario Nazionale (Hospital managed by National Health System), Azienda Ospedaliera integrata con l'Università (University Hospital), Policlinico universitario privato (Private university polyclinic), Istituto di Ricerca e Cura a Carattere Scientifico o IRCCS (Scientific Institute for Cure and Research) either Public or Private, Casa di cura privata accreditata (Accredited Private Nursing Home), Istituto qualificato presidio della U.S.L. and Ente di ricerca (Research Centre).
The Italian Ministry of Health publishes in HDR a series of healthcare quality indicators. We use the proportion of surgical discharges in non-surgical wards and we interpret it as a proxy of hospitals’ mismanagement and percentage of C-sections as a proxy of quality.

Finally, we exclude from the sample the smallest region, Valle D’Aosta, because it presents a large number of missing observations, leaving us a cohort of 20 Health Authorities.

4. Econometric Specification

We use linear models to estimate the impact of waiting times on the demand and supply and we assume that the system is in equilibrium and that demand equates supply. The supply is modelled similarly and we omit its presentation here, discussing below the differences in the control variables used. The basic empirical specification of equation (1) is:

\[
\log(y_{i,rt}^d) = \beta_0 + \beta_1 \log(wt_{i,rt}) + x_{i,rt} \cdot \beta_2 + z_{rt} \cdot \beta_3 + h_t + h_i + \epsilon_{irt},
\]

where utilization rate \(y_{i,rt}^d\) and waiting time \(wt_{i,rt}\) are log-transformed to interpret the coefficient of interest \(\beta_1\) as the elasticity of demand with respect to waiting time. Subscript \(i\) indicates the type of elective surgery (e.g. hip replacement, surgeries for breast cancer etc., with, \(i = 1, \ldots, I\)), \(r\) the region (with \(r = 1, \ldots, R\) ), \(t\) the year (with \(t = 2010, \ldots, 2014\) ) and the superscript \(d\) denotes demand. Utilization rates are measured by dividing the total number of discharges for a given surgical procedure in a region and year by the total number of residents (in million) of the region in the same year.

The matrix \(x_{i,rt}\) includes control variables that vary over time, procedures and regions, such as the number of emergency admissions. The matrix \(z_{rt}\) vary only over time and region and, in the demand equation, it includes the proportion of residents over 60, the percentage of unemployed, smoking prevalence, mortality rates at time \(t - 1\), the (log) number of private beds per capita, the proportion of emergency admissions, the share of C-sections on total number of birth and the mismanagement rate. The basic empirical model also includes a time dummy \(h_t\), to capture common time trends, and an procedure dummy \(h_i\) to control for procedure specific characteristics. The term \(\epsilon_{irt}\) is the error term.

The empirical specification of the supply of elective surgery \(y_{i,rt}^s\) is analogous to equation (3) but for the set of controls \((x_{i,rt}, z_{rt})\), which includes the share of emergency admissions,
the (log) total beds per capita for publicly-funded patients in wards where the surgeries are performed, the proportion of beds of private providers (serving publicly-funded patients) on the total amount of available beds in each region, and the (log) per capita euros spent by local health authorities or hospitals for medical staff as a proxy of the number of doctors and surgeons working in every region.

The ordinary least square estimation of equation (3) is however likely to produce a biased and inconsistent estimate of the coefficient of interest, $\beta_1$, because of the endogeneity of log-waiting times due to the simultaneity of demand and supply. Following previous literature (e.g. Martin and Smith, 1999, 2003), we instrument waiting time in the demand equation with a selection of exogenous supply shifters and waiting time in the supply equation with a selection of exogenous demand shifters. To check instrument validity we use the F statistics on the excluded instruments under the null of weak instruments and, following Stock, Wright and Yogo (2002), we conclude that instruments are valid if the F-statistic is larger than 10. Following Martin and Smith (2003) we also add the lag of waiting time as an instrument in demand models.

However, endogeneity of log of waiting time might still arise if it remains correlated with the error term because of some unobserved control variables. Hence, we augment equation (3) by first including the interaction between year- and procedure-fixed effects ($h_{ti}$), and finally also the interaction between procedure- and region-fixed effects ($h_{ir}$). By saturating the model with all these fixed-effects we address a residual source of endogeneity given by the correlation of waiting time ($wt_{irt}$) with omitted variables that do not vary within the year-procedure and within the procedure-region groups:

$$\log(y_{irt}^d) = \beta_0 + \beta_1 \log(wt_{irt}) + x_{irt}^t \cdot \beta_2 + z_{rt}^r \cdot \beta_3 + h_{ti} + h_{ir} + \epsilon_{irt},$$

(4)

This is a particularly demanding specification for the presence of a large number of fixed-effects among controls. The estimation procedure is even more troublesome as we deal with endogeneity of log-waiting time with instrumental variables, which tend to be highly collinear for the administrative nature of the data and their limited variability within regions.
5. Descriptive Statistics

We consider twenty Italian regions for five years of study and nine different surgical procedures. Our sample includes a maximum of 900 observations.

Table 1 reports some descriptive statistics for variables entering the supply and demand equations, respectively. Different number of observations reflects the fact that in the demand equation we used also the one-year lagged waiting time, reducing the estimating sample. On average the per capita utilization is of about 340 procedures (obtained by taking the exponential of 5.84) per million residents, of which about 22% is emergency admissions. Waiting time is about 27 days (exponential of 3.28).

The number of public beds for elective care per million population is about 1260 (exponential of 7.14) and if we compare it with the 288 private beds (Table 1, exponential of 5.66), their ratio shows that for each bed in private sector for elective patients there are about 4.37 beds in the public sector. The standardized cost per million residents for wages of doctors and medical staff is 215.72 (exponential of 5.374) thousands euro.

Table 1 shows that the proportion of elderly is about 28% and fraction of smokers smoking more than 11 cigarettes per day is higher, close to 40%. Unemployment rate is equal to 11.39 and it shows high variability during the period considered. In our data, mismanagement rate, measured as the share of misallocated operations, is quite high (32.38%) and C-section rate is 35.16% and with large variability.

Figure 1 shows the average (over regions) variability in log-waiting time for elective surgeries and log-utilisation rate at the beginning and at the end of the period considered. It suggests that there is limited variability over time whereas the highest variability is among treatments, with hip replacement procedures (Hip) requiring the longest wait and lung cancer surgeries (Lung) requiring the shortest. Similarly, Figure 2 shows the average (over procedures) variability in regional log-waiting time and log-utilisation rate showing more variation over time and large variability by regions.
Table 1: Descriptive statistics for the variables used in demand and supply equations.

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utilization per capita (in log)</td>
<td>716</td>
<td>5.844</td>
<td>1.080</td>
</tr>
<tr>
<td>Waiting time (in log)</td>
<td>716</td>
<td>3.293</td>
<td>0.557</td>
</tr>
</tbody>
</table>

**Supply shifters**
- Emergency admissions: 716, 21.686, 22.015
- Public beds per capita: 716, 7.138, 0.095
- Euros for staff wages: 716, 5.374, 0.760
- Share private market: 716, 16.002, 10.362

**Demand shifters**
- Proportion of over 60 years old: 716, 27.764, 2.685
- Lagged Mortality rate 1 year: 716, 10.249, 1.307
- Emergency admissions: 716, 21.686, 22.015
- Private beds per capita: 716, 5.664, 1.459
- Proportion of smokers: 716, 39.655, 4.621
- Unemployment rate: 716, 11.396, 5.149
- Mismanagement rate: 716, 32.380, 5.671
- C-section: 716, 35.163, 9.539

Source: Our calculations using ISTAT and HDR data, 2010-2014.
Figure 1: Utilisation rates (in logs) and waiting time (in logs) across procedures, at the start and the end of the period.

Source: Our calculations over HDR data. Notes: Local health authorities 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).
**Figure 2:** Utilisation rates (in logs) and waiting time (in logs), across regions at the start and the end of the period.

Source: Our calculations over HDR data. Notes: Local health authorities 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).
6. Empirical results

In Table 2 we report results of the demand side of the model. We first estimated the model in equation (3) using ordinary least squares (OLS), including year- and procedure-fixed effects, and we report results in the first column. In the second column we report the estimation of the same model obtained instrumenting (log-) waiting times with some supply shifter variables and the one-year lagged value of waiting time. In the third column we augment the set of controls by including the interactions \( h_{ti} \) and in the third we also add the \( h_{ir} \) interaction. Because of the inclusion of the one-year lagged value 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 preserve data consistency.

The sign of the waiting time coefficient is statistically significant and negative, as expected. The OLS estimation suggests an elasticity of demand to waiting time of -0.12, which increases (in absolute value) to -0.16 as log-waiting time is instrumented. As candidate instruments for waiting time tend to be affected by multicollinearity, we selected those that allow for the larger first stage F-statistics to test for weak instruments, namely the one-year lagged value of waiting time and the log of beds in public hospitals. \(^5\) The elasticity of waiting times increases (in absolute value) even more if all interaction dummies are introduced and model (4) is estimated. The estimation reported in column (4) suggests that a 10% increase of log waiting times lowers demand by about 3%.

As for the other control variables used, they enter the regressions with the expected sign, although they are not always statistically significant. In particular, the higher is private care availability, measured by the number of beds provided by the private sector (in log), the lower is NHS healthcare utilization; the larger the share of needy people, measured by the share of elderly population, the higher is demand; the higher is the mismanagement, measured by the rate of misallocated surgery procedure, the lower is demand.

In all IV estimation the first-stage F-statistics is higher than 10 suggesting that the first-stage regression exists.

\(^5\) Although different sets of instruments are less satisfactory in terms of first-stage F-statistics, results are robust to different groups of instruments used.
Table 2: Demand estimates

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>IV</td>
<td>IV</td>
<td>IV</td>
</tr>
<tr>
<td>Waiting time</td>
<td>-0.115***</td>
<td>-0.163***</td>
<td>-0.164***</td>
<td>-0.292***</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.049)</td>
<td>(0.049)</td>
<td>(0.086)</td>
</tr>
<tr>
<td>Proportion of over 60 y.o.</td>
<td>0.039**</td>
<td>0.043**</td>
<td>0.043**</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.018)</td>
<td>(0.018)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Mortality rate</td>
<td>-0.057</td>
<td>-0.063*</td>
<td>-0.064*</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.035)</td>
<td>(0.034)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>Emergency admissions</td>
<td>-0.008***</td>
<td>-0.008***</td>
<td>-0.008***</td>
<td>-0.007***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Private beds</td>
<td>-0.007</td>
<td>-0.005</td>
<td>-0.005</td>
<td>-0.034***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.009)</td>
<td>(0.009)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Fraction of smokers</td>
<td>-0.007**</td>
<td>-0.007**</td>
<td>-0.007**</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-0.036***</td>
<td>-0.036***</td>
<td>-0.036***</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Misallocation rate</td>
<td>-0.019***</td>
<td>-0.020***</td>
<td>-0.020***</td>
<td>-0.019***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>C-section (%)</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Constant</td>
<td>8.923***</td>
<td>9.182***</td>
<td>9.236***</td>
<td>9.439***</td>
</tr>
<tr>
<td></td>
<td>(0.360)</td>
<td>(0.363)</td>
<td>(0.364)</td>
<td>(0.392)</td>
</tr>
</tbody>
</table>

Observations: 716
Year Effects: Yes
Procedure Effects: Yes
Year x Procedure Effects: No
Procedure x Region: No
R-squared: 0.920
First stage F-stat: 337

Notes: The dependent variable is utilisation rate and standard errors are in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1; the instrument set used to estimate waiting times comprehends demand shifters, (log of) public beds, (log of) euros spent in wages for doctors and nurses and 1-year-lagged values of waiting times. These variables are used to predict the variable waiting time variable which is in the models above.

Table 3 contains results for supply and present the same structure as Table 2. The elasticity of supply of elective surgery to log-waiting time is not significant only if endogeneity is not taken into account. As an IV estimation is considered, instrumenting the log-waiting time
with the proportion of over 80 in the population (which presents a higher variability than the proportion of over 60) and with the misallocation rate, as proxies of the determinants of demand, the elasticity turns statistically significant. The magnitude in the most saturated model is 0.5 suggesting that for an increase of waiting time by 10%, supply increases by 5%. The first-stage F-statistic suggests that instruments are valid.

### Table 3: Supply estimates

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>IV</td>
<td>IV</td>
<td>IV</td>
</tr>
<tr>
<td>Waiting time</td>
<td>0.047</td>
<td>1.572***</td>
<td>1.559***</td>
<td>0.527***</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.239)</td>
<td>(0.234)</td>
<td>(0.148)</td>
</tr>
<tr>
<td>Emergency admission</td>
<td>-0.009***</td>
<td>0.004</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Public beds</td>
<td>0.631***</td>
<td>0.929***</td>
<td>0.934***</td>
<td>-0.057</td>
</tr>
<tr>
<td></td>
<td>(0.167)</td>
<td>(0.294)</td>
<td>(0.291)</td>
<td>(0.167)</td>
</tr>
<tr>
<td>Wages</td>
<td>-0.048**</td>
<td>0.005</td>
<td>0.004</td>
<td>-0.067***</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.035)</td>
<td>(0.034)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Private market share</td>
<td>-0.010***</td>
<td>-0.006**</td>
<td>-0.006**</td>
<td>0.006***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Constant</td>
<td>3.790***</td>
<td>-3.958*</td>
<td>-3.674</td>
<td>6.499***</td>
</tr>
<tr>
<td></td>
<td>(1.191)</td>
<td>(2.361)</td>
<td>(2.318)</td>
<td>(1.091)</td>
</tr>
</tbody>
</table>

Observations 716  716  716  716
Year Effects Yes  Yes  Yes  Yes
Procedure Effects Yes  Yes  Yes  Yes
Year x Procedure Effects No  No  Yes  Yes
Procedure x Region No  No  No  Yes
R-squared 0.880
First stage F-stat --  31.93  31.65  11.18

The dependent variable is utilisation rate and standard errors are in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1; the instrument set used to estimate waiting times comprehends proportion of over 80 and misallocation rate. These variables are used to predict the variable waiting time variable which is in the models above.

We conducted several sensitivity checks to assess the robustness of our empirical findings. First, we estimated our equations removing the Autonomous Provinces of Trento and Bolzano. Second, we also removed the Molise region. These regions are the smallest, with less supply of services and their residents tend to often move to nearby regions hence potentially affecting our results. Our findings are however robust to these sample size changes.
On supply side of the model we tried two different specifications, both of them excluding the log of per capita public beds in selected wards. In the first check we insert we total number of public beds (in log and standardised by population) regardless to wards and in the second one we also included beds in day hospital surgical wards. Estimates remain robust and instruments valid.

7. Discussion and Conclusion

We use administrative data provided by Italian Ministry of Health on waiting time and characteristics of Italian hospitals for different surgeries for elective patients, complemented with other administrative data provided by ISTAT, over the period 2010-2014 at the regional level.

We find negative and statistically significant coefficients of waiting time on aggregate demand suggesting that demand is inelastic to waiting. This implies that longer waiting time reduces demand for elective care although at a rate lower than one. We also find positive elasticity in supply, although smaller than one, suggesting that larger supply of public providers might be off-set by larger demand.

As expected the availability of public beds have a positive effect on supply of care and policy makers should be careful in deciding whether to reduce or not their number. Hospitals should improve their organisational skills since our findings suggest that high levels of mismanagement of patients are associated with a reduction in demand.

Our findings on demand are in line with previous UK literature as outlined by Martin and Smith (1999, 2003) and Martin et al. (2007), suggesting inelastic demand at around -0.2. This work employs a methodology which can be replicated as soon as newer and richer data for Italy become available as well as for application to other OECD countries.
Bibliography


