How do hospital-specialty characteristics influence health system responsiveness? An empirical evaluation of in-patient care in the Italian region of Emilia-Romagna

(Article begins on next page)

19 April 2024

How do hospital-specialty characteristics influence health system responsiveness? An empirical evaluation of in-patient care in the Italian Region of Emilia-Romagna

April 2017

Abstract

Studies of health system responsiveness mostly focus on the demand-side by investigating the association between socio-demographic characteristics of patients and their reported level of responsiveness. However, little is known about the influence of supply-side factors. This paper addresses that research gap by analysing the role of hospital-specialty characteristics in explaining variations in patients' evaluation of responsiveness from a sample of about 38,700 in-patients treated in public hospitals within the Italian Region of Emilia-Romagna. The analysis is carried out by adopting a two-step procedure. First, we use patients' self-reported data to derive five measures of responsiveness at the hospital-specialty level. By estimating a generalised ordered probit model, we are able to correct for variations in individual reporting behaviour due to the health status of patients and their experience of being in pain. Secondly, we run cross-sectional regressions to investigate the association between patients' responsiveness and potential supply-side drivers, including waiting times, staff workload, the level of spending on non-clinical facilities, the level of spending on staff education and training, and the proportion of staff expenditure between nursing and administrative staff. Results suggest that responsiveness is to some extent influenced by the supply-side drivers considered.

1. Introduction

Health system responsiveness has been recognized as one of the intrinsic goals of health systems and as a valid tool to assess their performance (World Health Organisation, 2000). The importance of this tool for the evaluation of health systems' performance has been further reinforced in recent years by both the European Ministerial Conference on Health Systems (World Health Organization, 2008) and the National Institute for Health and Care Excellence (NICE, 2012).[1]

Responsiveness can be defined as "the way in which individuals are treated and the environment in which they are treated, encompassing the notion of an individual's experience of contact with the health system" (Valentine et al., 2003a). The concept embraces eight dimensions of quality of care, which concern aspects of the interaction of patients with the health system and of patient satisfaction, including *autonomy*, *choice*, *clarity of communication*, *confidentiality of personal information*, *dignity*, *prompt attention*, *quality of basic amenities* and *access to family and community support*.[2]

Most of previous studies on health system responsiveness have focused on the demand-side by investigating the association between socio-demographic characteristics of patients and their reported level of responsiveness. Some of these studies have adopted an international comparison perspective (e.g. Valentine et al., 2008; Sirven et al., 2012; Rice et al., 2012), while others have adopted a national one (Adesanya et al., 2012; Ebrahimipour et al., 2013; Luo et al., 2013; Röttger et al., 2014; Fiorentini et al., 2015). However, little is yet known about the influence of supply-side factors on responsiveness. A few studies have

---

[1] Guidelines of NICE - a non-departmental public body within the UK Department of Health - have recently indicated the users' perspective as an instrument through which to evaluate the UK health system (NICE, 2012).

[2] Table I provides a definition of each of these domains. See Valentine et al. (2003a) for a more detailed description about the concept of responsiveness.

investigated the supply-side determinants of responsiveness with countries taken as units of observation (World Health Organisation, 2000; Anderson and Hussey, 2001; Blendon et al., 2001; Valentine et al., 2003b; Robone et al., 2011). By contrast, in this paper we conduct a supply-side analysis at a more disaggregated level by analysing the influence of hospital-specialty characteristics in explaining variations in patients' evaluation of responsiveness. The use of data at disaggregated level has been shown of paramount importance to increase the reliability of analyses about hospital systems, because it allows, for example, to take into account differences in accounting principles and in the definitions in the outcomes of interest, which may exist across different observational units (e.g. health-care providers) (Gravelle and Backhouse, 1987; Rice et al., 2010). Earlier work has shown that hospital characteristics play an important role in explaining differences in patients' evaluation of health-care (e.g. Young et al., 2000; Sjetne et al., 2007; Robone et al., 2014; Fiorentini et al., 2015). However, differently from previous studies on responsiveness, which treated hospitals as "black-boxes", the present paper investigates which supply factors affect responsiveness not only *across* hospitals, but also *within* hospitals.

In order to address the above research question, we use a unique dataset that combines information about responsiveness evaluations drawn from a sample of about 38,700 in-patients treated in public hospitals in the Italian Region of Emilia-Romagna with data on hospital-specialty characteristics from the hospital's annual financial and administrative reports. In this sample, patients' data on responsiveness are self-reported and measured on an ordinal categorical scale. A common issue in survey research is that individuals are likely to interpret the meaning of the available response categories in a way that systematically differs across populations or population subgroups (King et al., 2004). If this happens, a given level of responsiveness is unlikely to be rated equally by all respondents, thereby compromising data comparability. To address this issue, known as "reporting

heterogeneity" (Rice et al., 2012), we perform a two-step analysis following an established methodology (e.g. Fally et al., 2010; Huber and Stanig, 2011; Street et al., 2012). First, we purge differences in individual reporting behaviour due to patients' health status and their experience of physical pain by estimating a generalized ordered probit (GOP) model (Terza, 1985). Secondly, we run cross-sectional regression models in order to investigate the association between patients' responsiveness and potential supply-side drivers.

Our results **are consistent with a theoretical model of hospital behaviour in which the Chief Executive Officer maximises an objective function where revenues depend on her effort to improve quality along the dimensions directly perceived by patients (responsiveness). More precisely, our findings** indicate a strong negative association between the workload of nurses and responsiveness for the domains *communication* and *confidentiality*. They also reveal that overall patients' responsiveness is positively affected by the proportion of total staff expenditure invested in nursing staff. For the domain *prompt attention*, we also find a positive and significant effect on responsiveness of the proportion of expenditure invested in administrative staff, and a negative and significant effect of waiting times. Last, the level of spending on non-clinical facilities and on staff education and training activities have a strong positive effect on responsiveness.

This study innovates over previous literature on health system responsiveness by shedding some light into the "black-box" of the determinants of responsiveness at hospital-specialty level. Moreover, as far as we know, this is the first study to provide evidence in favour of the use of self-reported measures of responsiveness as credible indicators of how patients are treated by the health system. One issue raised by previous studies on responsiveness is the scarcity of *objective* measures (e.g. waiting times or space available in the hospital rooms) for several aspects of responsiveness (Rice et al., 2011). This has made it difficult to compare *self-reported* measures with more *objective* ones, as has been

done, for instance, with regard to health status (see, for example, Crossley and Kennedy, 2002; Baker et al., 2004; Jurges, 2007). Our results point to a strong correlation between the two sets of measures. This suggests that patients' self-reports may be good predictors of more objective measures of health system responsiveness, which are more costly to collect.

2. Background

2.1 Institutional setting and theoretical background

Our analysis focuses on a sample of elective inpatients treated in public hospitals located within the Italian Region of Emilia-Romagna. The Italian health-care system is a regionally based National Health Service (NHS) that provides comprehensive care financed by general taxation, and where patients are free to choose any publicly funded hospital, even outside their area of residence. The major public hospitals (Aziende Ospedaliere – AOs) are granted an independent legal status similar to that of Foundation Trusts in England, and are financed via a DRG-based Prospective Payment System (PPS). The remaining public hospitals are managed directly by the Local Health Authorities (LHAs), public organisations responsible for purchasing hospital and community services from a fixed budget to their local population. For treating patients who reside within the Region but outside the hospital catchment area, LHA hospitals are refunded by the LHA to which the patient is enrolled through the use of regional DRG tariffs. There are also private (accredited) hospitals that supply less than a fifth of the total hospital care (average of 18%), and whose output-mix is

designed to be complementary to that of public hospitals.[3] Hospital reimbursements for patients who reside outside the Region are paid on the basis of national DRG tariffs.

In such a setting, where regulated prices are used to allocate public funding and where patients have free choice of public sector treatment that is largely free of charge at the point of delivery, hospitals might have an incentive to raise quality. This may allow to retain local patients, thereby avoiding losses due to cost duplication, and to attract patients residing outside the hospital catchment area as a way of raising revenues (Lippi Bruni and Mammi, 2015). In the following, we present a theoretical framework based on the Bloom et al (2015) model to illustrate the behaviour and the underlying incentive structure of the hospital decision-makers. According to the model, the hospital Chief Executive Officer (CEO) chooses the level of effort e* that maximises her utility function:

$$U = p \cdot q(z(e), S) - c(q(z(e), S) - F \tag{1}$$

where p is a fixed price paid to the hospital for the provision of health-care; $q(z(e), S)$ is the demand of health-care, which is a function of quality of care, $z(e)$, and a vector of exogenous factors, S; and $c(q(z(e), S)$ and F are the variable and fixed costs, respectively. By assumption, quality increases with hospital effort: $z_e = (\partial z / \partial e) > 0$; and variable costs increase with both demand and hospital effort: $c_q = (\partial c / \partial q) > 0$ and $c_q = (\partial c / \partial e) > 0$. The first-order condition for utility maximisation is:

$$[(p - c_q)/c_e] \cdot \eta_e^q = (e/q) \tag{2}$$

---

[3] Private (accredited) hospitals subscribe contracts with LHAs. Under such contracts, the activity of private hospitals is tightly regulated by rigid budgetary constraints, so that they have no financial incentive to over-supply and, therefore, do not represent a significant competitive threat for public hospitals.

Equation (2) suggests that effort intensity, (e/q), increases with the price-cost margin, $(p - c_q)$, and decreases with the marginal cost of effort, c_e. Moreover, effort intensity increases with the elasticity of demand to the level of effort exerted by the CEO, η_e^q, *ceteris paribus*.

One of the assumptions underlying the result of equation (2) is that hospital quality is supposed to be an increasing function of managerial effort. However, the asymmetry of information between providers and patients may introduce a wedge between actual and perceived quality, and it is a matter for empirical investigation whether, and to which extent, increasing effort will translate into improvements in perceived quality. In our empirical application, we test this assumption by estimating whether quality of care, measured in terms of patients' responsiveness, does respond to changes in a vector of supply-side factors that we use to proxy managerial effort. In the next sub-section, we highlight the role of such supply-side factors, by reviewing a selected set of previous studies on health system responsiveness and on patient satisfaction.

2.2 Health system responsiveness and its potential determinants

We refer to two strands of literature to identify the supply-side factors **that might reflect managerial effort and** that might affect responsiveness.[4] The first consists of those studies that perform international comparisons of health system responsiveness and attempt to identify which country characteristics affect responsiveness, while the second consists of studies about "patient satisfaction", a concept that partly overlaps with "responsiveness".[5]

---

[4] Some frameworks have been recently proposed by Valentine et al. (2009) and Robone et al. (2011), but they refer to factors affecting responsiveness at the country level, while our analysis is performed at the hospital-specialty level.

[5] The notion of patient satisfaction does not coincide with the concept of responsiveness because it "may not capture what actually happens when people come in contact with the health system, and the responses are

In the first strand of literature, most studies report a positive association between health-care spending per capita and responsiveness (World Health Organization, 2000; Anderson and Hussey, 2001; Robone et al., 2011). Low levels of *prompt attention*, for instance, could be due to shortage of resources to remunerate the health staff or managerial failures (Valentine et al., 2003a). Similarly, low levels of *quality of facilities* could refer to shortages in hospital funding (World Health Organisation, 2000). Therefore, in analogy with the macro-level, we consider health spending at the hospital-specialty level, focusing particularly on the role of non-clinical facilities, as they are conceptually the more closely related to the concept of responsiveness.

It has been argued, however, that responsiveness does not depend only on health spending, and that a bivariate association may be a spurious one (World Health Organization, 2000). While high costs are required to guarantee some elements of responsiveness (e.g., *quality of facilities*), this is not necessarily the case for other elements (e.g., *dignity* and *communication*) - these may simply require improvements in staff training and awareness (World Health Organization, 2000; Blendon et al., 2001). For this reason, we investigate the effect of hospital staff training on patients' responsiveness. More generally, institutional and organizational factors play a role at hospital-specialty level in shaping how health systems meet patients' expectations (Blendon et al., 2001; Azfar and Gurgur, 2008), and can be important in mediating the relationship between health-care spending and responsiveness.[6] To identify such factors, we refer to our second strand of literature (i.e. studies of "patient satisfaction").

---

strongly influenced by prior expectations of what will or should happen" (Valentine et al., 2003a). The actual experience of people, differently, should be shown by the responsiveness measurement (Coulter and Cleary, 2001).

[6] A positive and significant relationship between health-care spending and responsiveness has been found for specific groups of countries and specific types of treatments only (Valentine et al. 2003b, 2009).

In the literature on patient satisfaction, staff workload has been found as one of the factors influencing patient satisfaction negatively (Ansmann et al., 2013). This could be particularly relevant for patient satisfaction with respect to those domains of responsiveness that might imply a strong interaction between patients and staff. As an example, if staff are subject to an excessive workload, their capacity to treat patients respectfully may be impaired (Valentine et al., 2003a). Therefore, we consider as possible supply-side drivers of responsiveness both the medical and nursing staff workloads. In particular, we anticipate nursing staff to have the greatest impact on patients' responsiveness, since they have the closest contact with patients. The importance of the nursing role is suggested by Leiter et al. (1998), Vahey et al. (2004), and Kutney-Lee at al. (2009), who provide evidence of the key influence of nursing staff, rather than physicians, for patient satisfaction. For this reason, we also investigate the effects of staff composition by considering the expenditure on nursing staff as a share of total staff expenditure. Moreover, since the share of expenditure on administrative staff may affect at least some domains - most notably *confidentiality* and *prompt attention* - we also include this variable in our model.

Further, we examine the effect of waiting times on patients' responsiveness. We expect this variable to be strongly correlated to the domain *prompt attention*. Moreover, this factor could also influence patients' opinion with other domains, since patients frustrated by long waiting times could be more critical when judging such domains.

Finally, many studies have investigated the influence on patient satisfaction of hospital size (e.g. Young et al., 2000; Sietne et al., 2007; Bacon and Mark, 2009; Hekkert et al., 2009; Murante et al., 2014) and of case-mix (e.g. Bacon and Mark, 2009; Ansmann et al., 2013). Although these variables are not strictly linked to responsiveness, we include them as controls for the hospital organizational models.

3. Data

3.1. Patient-level data

The data are drawn from several sources. First, to derive our measures of responsiveness, we use survey data collected by the Agency for Health Care and Social Services of Emilia-Romagna (*Agenzia Sanitaria e Sociale Regionale dell'Emilia-Romagna*, ASSR) in years 2008-2012. The survey aimed at investigating in-patient satisfaction with the health-care services provided by public hospitals located in the Region of Emilia-Romagna.[7] The questionnaire was developed by the ASSR, and respondents were selected through sampling procedures based on a probabilistic approach to help ensure that the sample is statistically representative of the group of in-patients treated in each hospital ward. The survey was carried out using a self-administered, pen and paper questionnaire. The help of a caregiver was permitted for patients with severe clinical conditions. More details on the survey design and the sampling procedure are reported in Appendix I.

We pool the survey data across all years, ending up with a final sample of 38,696 respondents. The data are clustered into 3 levels: patients within specialties within hospitals. Different hospitals participated to the survey in different years between 2008 and 2012. The largest collection of data took place in 2010 (44%), with 2% in 2008, 14% in 2009, 22% in 2011 and 18% in 2012. In each survey year, individuals were questioned about their satisfaction with a number of aspects of the health system (21 items), which are part of the standard domains of responsiveness. Table II illustrates the correspondence between the items of the ASSR questionnaire and the standard domains of

---

[7] Because the survey did not cover day-hospital cases, only ordinary in-patients are included in our sample.

responsiveness as defined by the World Health Organization (WHO) and elaborated by Valentine et al. (2003a). The latter domains comprise *communication* (e.g. clarity of information provided by hospital personnel), *confidentiality* (e.g. whether sensitive information were privately communicated to patients), *dignity* (politeness and respectfulness of hospital staff), *prompt attention* (waiting times) and *quality of facilities*.[8] Within each responsiveness domain, patients were asked to rate their experience according to six categories: "completely dissatisfied", "very dissatisfied", "dissatisfied", "satisfied", "very satisfied" and "completely satisfied". We aggregate "completely dissatisfied" and "very dissatisfied", because of the very low proportion of responses in these categories.

The dataset also contains information on patients' characteristics, which we exploit to build the variables used as regressors in the first-stage analysis. Table III provides a brief description and summary statistics for these variables.

To capture patients' health status we include two dummies. The first is for individuals reporting a poor or moderate health status (*poor/moderate health* = 1, 0 otherwise), with those being in good or excellent health being the reference category. The other dummy is for individuals stating they have experienced pain during their hospital stay (*pain* = 1, 0 otherwise). Moreover, the dummies *emergency* (= 1 if Accident & Emergency hospitalization, 0 otherwise) and *previous hospitalisation* (= 1 if admitted to the same ward before, 0 otherwise) are also included to account for the kind of treatments received by

---

[8] In the current study, we adopt a version of the domain *quality of facilities*, which is as close as possible to the one proposed by the WHO Report (2000). Indeed, following the operalization made by the WHO, only items regarding *cleanliness* and *space* have been used to form this domain. As a robustness check, we have also replicated our empirical analysis by using a more comprehensive version of *quality of facilities*, which also includes items regarding patient satisfaction with meals and heating, comfort of the bedding, level of quietness and maintenance of the facilities and of the rooms in the ward. Results are qualitatively similar to those shown in the paper and are available upon request.

patients. Further, as shown in Table III, we also control for other patients' characteristics, including gender, marital status, education, employment status and location of residence.[9]

Finally, the dataset contains information about both the hospital and - *within each hospital* - the specialty where patients are treated. As previous work has noted, general hospitals provide health-care services across a range of specialties, which can be viewed as having separate production functions (e.g., Linna and Häkkinen, 2006; Laudicella et al., 2010).[10] Given this heterogeneity across specialty units, earlier studies claim that comparing hospital-specialties is more appropriate than comparing hospitals (e.g. Harper et al, 2001). Therefore, in the first-stage regressions we include hospital-specialty fixed effects, which allow us to control for unobservable hospital-specialty-specific characteristics that may affect responsiveness.[11]

3.2. Hospital-specialty data

The final sample used at the second step includes 165 hospital-specialty units, referring to 30 different specialties and 28 public hospitals within Emilia-Romagna. **On average, the number of specialties per hospital is 6, ranging between 1 and 20.** [12] To generate the explanatory variables used in the second step analysis, we employ data drawn from the administrative hospital discharge dataset (*Schede di Dimissione Ospedaliera*, SDO), complemented by data on available beds, labour inputs and hospital costs provided

---

[9] Income and age are not included in the model for either data availability or collinearity reasons. In our sample the majority of respondents are retired, and thus age and the employment status variables are highly correlated. Data on patients' income is not collected by the ASSR to protect patients' confidentiality.

[10] Specialties are groups of the hospital wards that in each hospital share several distinct features, such as the diseases and the type of patients treated, the activities performed and the type of labour and capital endowment. Examples of specialties included in our sample are *orthopaedics*, *ophthalmology*, *obstetrics and gynaecology* and *general medicine*, among the others.

[11] The base category is the specialty *general medicine* in the hospital "Guglielmo da Saliceto" (Piacenza). We chose this as the base category because it is the largest hospital-specialty in our sample.

[12] The specialties included in the final sample are listed in Appendix II.

by the Regional Department of Health. Table V offers some descriptive statistics. Since the survey data refer to a sample of patients discharged in years 2008-2012, the data we use to generate the explanatory variables are averaged over the period 2008-2011.[13]

As for our key explanatory variables, we use measures of specific supply-side factors related to the items included in each responsiveness domain. First, we calculate the median of inpatient waiting time for patients discharged from each hospital-specialty unit. On average, the median waiting time is approximately 3 weeks.[14] We expect this variable to have a negative impact on the dependent variable capturing the waiting times domain of responsiveness (i.e. *prompt attention*).

Additionally, we examine whether patients' responsiveness is affected by the workload of doctors and nurses. Our workload measures are given by the ratios between the total number of inpatient cases (measured in 1000 cases) and the total number of full-time equivalent doctors and nurses employed in each hospital-specialty unit. In the sample, there are on average about 137 in-patients per doctor and 136 in-patients per nurse.

The last set of key explanatory variables is derived from data on costs recorded at Trust-specialty level rather than at hospital-specialty level.[15] We include the expenditure per patient (in 1000 Euros) on non-clinical goods and services (food, laundry, beddings and

---

[13] More precisely, data on beds, labour inputs and hospital discharges are extracted for years 2008-2011, and then averaged across those years. The average values of these variables remain fairly constant through that time period. For brevity, we do not report the sample statistics separately for each year, but they are available upon request. Due to lack of data availability, we could not extract the same information for year 2012. However, since the extracted data remain constant over the period 2008-2011, and there were neither institutional nor organisational changes for the hospitals included in our sample in 2012, we do not expect our results to be affected by this exclusion. Finally, cost data were available for 2010 only.

[14] We use the median as opposed to the mean waits, in order to account for the fact that the distribution of waiting times is skewed to the left, with a large proportion of patients who had to wait for a short time before they obtained their required treatments, and a long right-hand tail of individuals for whom we observe longer waiting times.

[15] The majority of hospitals included in our sample are under the direct control of the ~~Local Health Authorities (Aziende Sanitarie Locali – ASLs)~~ LHAs~~, which are public health enterprises responsible of the health care needs of their catchment population~~. The only exception represented is a teaching hospital that enjoys the status of independent hospital Trust, with considerable financial and decision-making autonomy (responsibility for the budget, management and technical functioning).

cleaning materials), and the expenditure (in 1000 Euros) on staff education and training activities over the **sum of the DRG weight units from each** hospital-specialty unit.[16] Due to the high degree of correlation (0.91) between the expenditure on non-clinical goods and services and on staff education and training activities, to avoid multicollinearity problems we run separate regressions including each of these measures in turn. On average, the expenditure on non-clinical goods and services is about 107 Euros per patient, while the expenditure on staff education and training activities is about 19 Euros per DRG weight unit. Concerning the proportion of total staff expenditure on nurses and on administrative staff, they are approximately 43% and 5%, respectively.

Finally, we include **hospital** Trust dummies to allow for unobserved Trust-specific factors, such as management style and human resource policies, that may affect responsiveness, and use measures of hospital-specialty size and patient severity as additional control variables. Previous studies examining the determinants of patient satisfaction find evidence that larger hospital units are associated with lower levels of satisfaction (e.g. Young et al., 2000; Meterko et al., 2004; Sjetne et al., 2007; Hekkert et al., 2009). In this study, we use the number of beds recorded at hospital-specialty level as a proxy for the hospital-specialty size.[17] The average number of beds for hospital-specialty unit is about 25. Since part of the variation in patients' responsiveness may be due to the heterogeneity across hospital-specialties with respect to case-mix, we further control for

---

[16] The choice of the denominator for the latter variable is due to the fact that spending in education and training - which is not compulsory and does not depend on the sheer number of patients - can be more naturally framed as a portfolio choice. In this respect, the value of the delivered DRGs represents the financial budget to be allocated to alternative uses, including, among the others, staff education and training activities.
[17] The number of inpatient cases is an alternative variable frequently used as a measure of size. However, since this variable is highly correlated with the number of beds (the degree of correlation being equal to 0.81), we could not include it in the regression model.

patient case-mix, as proxied by the average DRG weight assigned to patients treated in each hospital-specialty unit. In our sample, on average, this is equal to 1.224.[18]

4. Methods

The aim of this study is to examine the impact of hospital-specialty characteristics on patients' responsiveness, conditional on patients' reporting behaviour. For this purpose, we perform a two-stage analysis using STATA 12.

4.1 First-step analysis

In the first stage, we **pool all survey years in a single cross-section and** estimate a generalized ordered probit (GOP) model of the following form (Terza, 1985): **[19]**

$$y_{id} = j \quad if \quad \mu_{j-1} < y^*_{id} < \mu_j, \quad j = 1,..., m \tag{3}$$

where the dependent variable y_{id} is an ordered variable that rank patients' evaluation of responsiveness according to five different values from 1 ("completely dissatisfied" or "very dissatisfied") to 5 ("completely satisfied") for each respondent i and for each of the five responsiveness domain d listed in Table II and described in section 3.1; the latent variable y^*_{id} is modelled as a linear function of a vector of exogenous variables (x) plus a random error term ε:

---

[18] DRG weights reflect the average amount of resources necessary to treat patients with a given diagnosis. Higher DRG weights indicate a higher complexity of patient clinical conditions.
[19] In each year, a different set of hospitals collected the data, and we pool observations across all years.

$$y*_{id} = x_{id}\beta_d + \varepsilon_i \tag{4}$$

$$\varepsilon_{id} \sim N(0,1)$$

and the unknown threshold values $\mu_{j-1} = -\infty, \mu_j \leq \mu_{j+1}, \mu_m = \infty$ are assumed to be functions of some explanatory variables (z_{id}):

$$\mu_{id}^j = z_{id}\gamma^j, \quad j=1,...,4 \tag{5}$$

Since four cut-points (μ) divide the ordered categorical outcomes, we need to estimate four cut-point equations in the GOP model. The nonlinearity of the model makes it possible to provide only a qualitative interpretation of the estimated parameters β_d. Higher values of β_d indicate higher levels of patients' responsiveness.

GOP regressions have already been estimated in a number of empirical studies to control for reporting heterogeneity (see, for instance, Pudney and Shields, 2000; Etilè and Milcent, 2006; Jurges, 2007).[20] The GOP model is a flexible extension of the standard ordered probit (OPROBIT) estimator. Whilst the latter model assumes that individuals adopt homogeneous reporting behaviour by mapping the latent variable to the available response categories in the same way, through a set of constant cut-points (Rice et al., 2012), the former model eliminates the need for such restrictive assumption, allowing for the presence of differences in reporting style.[21]

---

[20] See also Jones and Schurer (2011) for more references to GOP models.
[21] If the assumption of homogeneous reporting behaviour does not hold, then the effects on responsiveness estimated using the OPROBIT will be biased, reflecting both the "true" responsiveness effect and the reporting heterogeneity effect.

In the original dataset, 21 different items of responsiveness were rated by each patient. Although it would be possible to estimate 21 different regression models, this would lead to a burdensome interpretation of the results. Thus, we stratify our sample into 5 sets, one for each domain of responsiveness. Given that for each responsiveness item the level of satisfaction is recorded, we reshape the dataset from a wide to a long form within each set. A variable called "Satisfaction with responsiveness" was created for each of the 5 responsiveness domains; it represents the patient satisfaction with the domain of responsiveness considered, regardless of the specific item of responsiveness under evaluation. For each responsiveness domain, the percentage of individuals choosing each response category is illustrated in Table IV. The domain which appears to be best rated is *dignity*, while *prompt attention* appears as the worst rated.

In order to separate the "true" responsiveness effect from the "reporting heterogeneity" effect, the GOP model requires that the regressors included in the cut-points equation (5) and in the main responsiveness equation (4), i.e. z_{id} and x_{id} respectively, are distinct vectors (Pudney and Shields, 2000). We assume that the variables *poor/moderate health* and *pain* do not influence "true" responsiveness. Therefore, we include these variables in the cut-points equation, while excluding them from the main responsiveness specification.[22] The exclusion from the main responsiveness equation of *poor/moderate health* and *pain* is based on the observation that it does not appear plausible that the Health Service in Emilia-Romagna is systematically less responsive to individuals who suffer from poor/moderate health or who are in pain. First, from an institutional point of view, health is recognised by the Italian Constitution as a fundamental right of citizens. To guarantee this

---

[22] **This assumption implies that individuals who suffer more pain or experience bad health have higher expectations from the health service. Their higher expectations produce a bias in their reporting of responsiveness, which is taken into account by the GOP model.** Differently, with regard to x_i variables, we are not able to assume that they solely affect "true" responsiveness, and we interpret them as having combined effects ("true" responsiveness effect plus "reporting heterogeneity" effect).

right, in 1978 the Italian National Health Service (NHS) was established as a regional administrative system based on the principles that health-care should be financed through general taxation and should be allocated on the basis of health needs. ~~The working of e~~Each Regional Health Service is inspired by concerns about the equity of health-care access and financial contribution (France et al. 2005, Ferré et al. 2014). Second, from an empirical point of view, Glorioso and Subramanian (2014) provided evidence that in the Italian NHS the objective of horizontal equity is met for access to hospitalization, with no discrimination due to patients' health status. On the basis of such evidence, since the concepts of "responsiveness" and "access to care" share several features (e.g. waiting times) (Valentine et al. 2003a), the Italian NHS is unlikely to be considered inequitable with regard to responsiveness too.[23]

4.2 Second-step analysis

In the second-step analysis, we use an estimated dependent variable (EDV) model (Lewis and Linzer, 2005) to analyse the role of hospital-specialty characteristics in explaining variations in patients' responsiveness. By each responsiveness domain, the estimated coefficients for the hospital-specialty dummies derived from the first stage are regressed against a set of supply-side factors. An example of this modelling strategy can be found in Robone et al. (2011), who examine how health system responsiveness can be affected by country characteristics at aggregate level. The estimated coefficients reflect the relative levels of responsiveness across hospital-specialties.[24] By each responsiveness

---

[23] The same identification assumption was adopted in Fiorentini et al. (2015).

[24] For example, if hospital-specialty A shows a greater (positive) coefficient than hospital-specialty B, we can interpret this difference as hospital-specialty A being more responsive to patients' expectations than hospital-specialty B, *ceteris paribus*.

domain, the model is estimated on a pool of 165 hospital-specialties. Huber–White robust

standard errors (White 1980) are applied since heteroskedasticity from sampling variation

in the estimated levels of hospital-specialty-specific responsiveness might be induced when

using an EDV model (Lewis and Linzer, 2005).[25]

5. Results

Table VI contains the results from the GOP model in the first step. Across the 5

responsiveness domains, the coefficients of *poor/moderate health* and *pain* are positive and

mostly statistically significant in the cut-point 1 and cut-point 2 equations, while they are

negative and mostly statistically significant in the cut-point 4 equation. This evidence

suggests that patients in *poor/moderate health* and in *pain* tend to use more the extreme

categories of responsiveness ("completely or very dissatisfied" or "completely satisfied")

compared to patients who are not in such conditions. [26]

---

[25] **In alternative to our two-stage model, we might have explored the use of a multilevel single-stage model. Multilevel single-stage models have been employed in previous studies of health system responsiveness (e.g. Valentine et al., 2015). In a multilevel single-stage model, the unexplained error term is decomposed into different random components attributable to each level of the hierarchy, and assumed to have zero mean and constant variance (Rice and Jones, 1997). However, the estimation of a single-stage model requires the assumption that the random components across hierarchical levels are uncorrelated. This assumption may be restrictive in our application, since unobserved patient severity may be correlated with provider units.**

[26] We have also estimated an OPROBIT model by regressing patients' rating of responsiveness on the sets of variables illustrated in section 3.1 (results are available on request). In the OPROBIT model, the estimated coefficients on *poor/moderate health* and *pain* are negative and strongly statistically significant for all responsiveness domains. However, for the reasons explained in section 4.1, it is implausible that the Health Service in Emilia-Romagna is systematically less responsive to patients who are in *poor/moderate health* and in *pain*. We interpret these results as evidence of the negative effect of such conditions on patients' reporting behaviour, supporting the use of the GOP model instead of the OPROBIT one, and the validity of our exclusion restrictions. Other evidence supports this choice. For both *poor/moderate health* and *pain*, results from a Wald test across the four cut-points allow to reject the null hypothesis that the estimated coefficients are jointly equal to zero, thus indicating that the cut-points are functions of these variables (results are available on request). Moreover, for each responsiveness domain, the values of the Akaike and Bayesian information criteria show that the best fit to the data is obtained using the GOP model rather than the OPROBIT one (results are also available on request).

Tables VII and VIII present, for each responsiveness domain, the results obtained from the second-step regressions. The difference between Tables VII and VIII is that, while the former shows the effect of spending on non-clinical facilities, the effect of spending on staff education and training activities is estimated in Table VIII. The results of the Ramsey regression specification-error test (RESET), reported in the last row of the two tables, provide no evidence of functional form misspecification.

The coefficient of the waiting time variable is negative and highly significant for the domain *prompt attention*, suggesting a strong and negative association between actual waiting times and the patients' reported level of responsiveness with respect to waiting times. Whilst we find a strong negative association between nursing staff workload and patients' responsiveness for the *communication* and *confidentiality* domains, the effect of medical staff workload is never significant. Overall, the proportion of total staff expenditure in nursing staff has a positive and significant impact, while the proportion of total staff expenditure in administrative staff has a positive and significant effect only for the domain *prompt attention*. Not surprisingly, these findings highlight the crucial role of nursing staff on the patients' experience of contact with the health system (Leiter et al., 1998, Vahey et al., 2004, and Kutney-Lee et al., 2009). In addition, our result for the impact of the proportion of total staff expenditure on administrative staff for the *waiting times* domain supports the hypothesis advanced by Valentine et al. (2003a) suggesting that increasing the proportion of administrative staff on total staff expenditure improves the hospitals' ability to manage the waiting lists. Concerning the expenditure on non-clinical goods and services, we find a positive and significant association for all domains. This finding is in line with previous evidence provided at country-level (World Health Organization, 2000; Anderson and Hussey, 2001; Robone et al., 2011). A positive and significant effect on all domains of responsiveness is also found for the expenditure on staff education and training activities.

This result also provides evidence for what has been previously suggested in the literature (World Health Organization, 2000; Blendon et al., 2001).

The remaining rows in Tables VII and VIII report the estimates for our control variables. We find evidence of a negative and significant association between the number of available beds and responsiveness across all domains. This is consistent with the results of previous studies, which suggests a negative association between hospital size and patient satisfaction (e.g., Young et al., 2000; Meterko et al., 2004; Sjetne et al., 2007; Hekkert et al., 2009), and may reflect possible diseconomies of scale due to, for example, problems of coordination and cooperation in larger hospital-specialty units. Overall, patient severity appears also to be a relevant determinant of responsiveness. The positive influence of patient case-mix on responsiveness implies, on average, a higher ability of staff to meet patients' expectations in those hospital specialties where there are more severe cases. Possible explanations are that, in those hospital-specialty units where patient complexity is higher, doctors and nurses are characterised by a higher degree of motivation, and/or make greater efforts to reduce the risk of malpractice claims.

6. Discussion and conclusions

This paper investigates the role of hospital-specialty characteristics in explaining variations in the health system' responsiveness rated by a sample of about 38,700 in-patients treated in public hospitals located in the Italian Region of Emilia-Romagna. **The hospital-specialty characteristics that we analysed are used as proxies of managerial effort, and allow us to test the theoretical assumption underlying previous studies (Bloom et al, 2015) according to which hospital quality, which we measured in terms of patients' evaluation of responsiveness, is an increasing function of managerial effort. Overall,**

our findings suggest that the supply-side factors considered play a relevant role for responsiveness. More precisely, we find a strong negative association between nursing staff workload and responsiveness for the *communication* and *confidentiality* domains. In addition, our analysis reveals a positive and significant association between responsiveness and the proportion of total staff expenditure invested in nursing staff. Moreover, **our results point to a strong positive effect on responsiveness of the expenditure on non-clinical facilities and on staff education and training activities.** Last, **for the domain *prompt attention*, we find a positive and significant effect on responsiveness of the proportion of total staff expenditure invested in administrative staff, and a negative and significant effect of actual waiting times.**

Our study has some limitations. Although the sample used at the first step of our analysis includes a large number of patient-level observations, the one employed at the second step comprises a relatively small number of hospital-specialty units. **This data limitation has prevented us from exploring potential differences in the patients' evaluation of responsiveness across hospital specialties.** Future studies might benefit from using a larger sample for the second-step estimation. This may also allow future research to examine possible differences in the impact of supply-side factors on responsiveness *across* specialties within hospitals. We might expect such differences to arise between, for example, surgical and medical specialties, given the different experience of patients with medical and nursing staff across specialties. Further, future studies may also take advantage from the use of additional information on the type of investments on non-clinical goods, and on education and training activities, to investigate the transmission channels linking the supply-side factors to responsiveness.

Since vignettes have not been included in the questionnaires administered by the ASSR, we could not apply the hierarchical ordered probit (HOPIT) model, which has

instead been used by several studies investigating the issue of reporting heterogeneity for responsiveness (e.g., Valentine et al., 2003b; Sirven et al., 2012; Rice et al., 2012).[27] However, although the GOP model does not allow the disentanglement of the "true" effects on responsiveness from the "reporting style" effect, it enables us to account for some forms of biases in reporting style, thanks to the identification restrictions we make.[28] Whilst the identification assumption imposed by the GOP model is not required by the HOPIT model, the latter requires other assumptions, known as *vignette equivalence* and *response consistency*, whose tenability has been questioned and is currently a subject of extensive debate (see, for example, Bago d'Uva et al, 2011; Peracchi and Rossetti, 2013; Robone, 2016). More generally, since the inclusion of vignettes within a survey increases both the financial costs of implementing that survey and the risk of inducing fatigue effects in the respondents (King and Wand 2007, Rice at al. 2012, Peracchi and Rossetti 2013), few datasets include vignettes. Therefore, using the GOP model can result helpful in purging self-reports from some reporting heterogeneity when vignettes are not available.

Despite its limitations, our paper is the first attempt, as far as we know, to shed some light into the "black box" of the determinants of responsiveness at hospital-specialty level. By investigating the supply-side effects on responsiveness, we are also able to provide some evidence in favour of the use of self-reported measures of responsiveness as credible

---

[27] "Vignettes represent hypothetical descriptions of fixed levels of a latent construct, such as responsiveness […] Since the vignettes are fixed and predetermined, any systematic variation across individuals in the rating of the vignettes can be attributed to differences in reporting behavior" (Rice et al. 2012).

[28] **Both the HOPIT and the GOP models are specified through the use of two equations, where the first equation represents the "true" effect of some covariates on the variable of interest in the analysis (~~we call it~~ main equation), while the second equation represents the effect of some covariates on the reporting behavior of the respondents (~~we call it~~ reporting bias equation). Whilst when estimating the HOPIT model, thanks to the additional information provided by vignettes, the same set of covariates can be included both in the main equation and in the reporting bias equation, this is not possible when estimating the GOP model. Because of the lack of additional information provided by vignettes, identifying the GOP model requires that the variables ~~which are~~ included in the reporting bias equation are used exclusively in that equation. A second set of variables has to be used in the main equation. The impossibility of including a variable both in the main equation and in the reporting bias equation is the main limitation of the GOP model.**

indicators of how patients are treated by the health system. In particular, for the domains *prompt attention* and *quality of facilities*, we are able to compare subjective and objective measures of responsiveness, as has been done, for instance, with regard to health status (Crossely and Kennedy, 2002; Baker et al., 2004; Jurges, 2007). The two types of measures appear to be strongly correlated, suggesting that patients' self-reported measures can be considered as valid predictors of more objective measures of responsiveness. Therefore, NHS managers and policy-makers can regard self-reported measures of responsiveness as reliable information sources that can be used as governance tools to improve the performance of decentralised units.

However, our analysis also shows that taking action based on the results of a responsiveness survey alone, without controlling for the supply-side variables, may lead to serious misallocations of resources. Indeed, if relatively poor results (in terms of patients' evaluation of health service) in given hospital units are reported, more intensive monitoring or the enforcement of incentives/sanctions should be contemplated only after taking into consideration possible shortages in critical resources, such as nursing staff, non-clinical goods and services or investment in staff training.

More generally, in the last decades there has been a growing need to adopt patients' self-reported measures of quality to balance the charge that the supply of health services is increasingly shaped in response to providers' interests. This calls for further methodological refinements in both the design and the interpretation of such measures, starting from the need to take into account the role of the supply-side factors.

References

Adesanya T, Gbolahan O, Ghannam O, Miraldo M., Patel B., Verma R, Wong H. 2012. Exploring the responsiveness of public and private hospitals in Lagos, Nigeria. *Journal of Public Health Research*, volume 1 (e2).

Anderson G, Hussey P. 2001. Comparing Health System Performance in OECD Counties., *Health Affairs* **20** (3): 219-232.

Ansmann L, Kowalski C, Ernstmann N, Ommena O, Jung J, Visser A, Pfaff H, 2013. Do breast cancer patients receive less support from physicians in German hospitals with high physician workload? A multilevel analysis. *Patient Education and Counseling* **93**: 327–334.

Azfar O, Gurgur T. 2008. Does Corruption Affect Health and Education Outcomes in the Philippines?. *Economics of Governance* **9**: 197–244.

Bacon CT, Mark B. 2009. Organizational Effects on Patient Satisfaction in Hospital Medical-Surgical Units. *Journal of Nursing Administration* **39**(5): 220–227.

Bago d'Uva T, Lindeboom M, ODonnell O.,Van Doorslaer E. 2011. Slipping anchor?: testing the vignettes approach to identification and correction of reporting heterogeneity. *Journal of Human Resources* **46**: 872–903.

Baker M, Stabile M, Deri C. 2004. What do self-reported, objective, measures of health measure?. *Journal of Human Resources* **39**: 1067-1093.

Blendon RJ, Kim M, Benson JM. 2001.The Public Versus The World Health Organization on Health System Performance. *Health Affairs* **20** (3): 10–20.

Bloom N, Propper C, Seiler S, Van Reenen J. 2015. The impact of competition on management quality: Evidence from public hospitals. *The Review of Economic Studies* **82** (2): 457-489.

Coulter A, Cleary P D. 2001. Patients' Experiences With Hospital Care In Five Countries. *Health Affairs* **20** (3): 24-52.

Crossley TF, Kennedy S. 2002. The reliability of self-assessed health status. *Journal of Health Economics* **21**: 643–658.

Ebrahimipour H, Vafaei Najjar A, Khani Jahani A, Pourtaleb A, Javadi M, Rezazadeh A, Vejdani M, Shirdel A. 2013. Health System Responsiveness: A Case Study of General

Hospitals in Iran. *International Journal of Health Policy and Management* **1**(1): 85-90.

Etilé F, Milcent C. 2006. Income-related reporting heterogeneity in self-assessed health: evidence from France. *Health Economics* **15**: 965–981.

Fally T, Paillacar R, Terra C. 2010. Economic geography and wages in Brazil: Evidence from micro-data. *Journal of Development Economics* **91**: 155–168.

Ferré F, de Belvis AG, Valerio L, Longhi S, Lazzari A, Fattore G, Ricciardi W, Maresso A. 2014. Italy: Health System Review. *Health Systems in Transition* **16**(4):1–168.

Fiorentini G, Ragazzi G, Robone S. 2015. Are bad health and pain making us grumpy? An empirical evaluation of reporting heterogeneity in rating health system responsiveness. *Social Science & Medicine* **144**: 48-58.

France G, Taroni F, Donatini A. 2005. The Italian Health-care System. *Health Economics* **14** (S1): S187–202.

Glorioso V, Subramanian SV. 2014. Equity in Access to Health Care Services in Italy. *Health Services Research* **49** (3): 950-970.

Gravelle H, Backhouse ME. 1987. International Cross-Sectional Analysis of the Determination of Mortality. *Social Science and Medicine* **25**(5): 427‑41.

Hekkert KD, Cihangir S, Kleefstra SM, Van den Berg B, Kool RB. 2009. Patient satisfaction revisited: A multilevel approach. *Social Science & Medicine* **69**: 68-75.

Huber JD, Stanig P. 2011. Church-state separation and redistribution. *Journal of Public Economics* **95**: 828–836.

Harper J, Hauck K, Street A. 2001. Analysis of costs and efficiency in general surgery specialties in the United Kingdom. *European Journal of Health Economics* **2**: 150-157.

Jones AM, Rice N, Bago d'Uva T, Balia S. 2007. Applied Health Econometrics, Routledge eds.

Jones A.M., Schurer S. 2011. How does heterogeneity shape the socioeconomic gradient in health satisfaction? *Jounral of Applied Econometrics* **26**: 549-579.

Jurges H. 2007. True Health Vs Response Styles: Exploring Cross-country Differences In Self-Reported Health. *Health Economics* **16**: 163–178.

King G, Murray CJL, Salomon J, Tandon A. 2004. 'Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review* **98** (1): 184–91.

King G, Wand J. 2007. Comparing incomparable survey responses: New tools for anchoring vignettes. *Political Analysis* **15**: 46–66.

Kutney-Lee A, McHugh MD, Sloane DM, Cimiotti JP, Flynn L, Neff DF, Aikem LH. 2009. Nursing: A key to patient satisfaction. *Health Affairs* **28**: w669-w677.

Laudicella M, Olsen KR, Street A. 2010. Examining cost variation across hospital departments – a two-stage multi-level approach using patient-level data *Social Science & Medicine* **71**: 1872-1881.

Leiter MP, Harvie P, Frizzell C. 1998. The correspondence of patient satisfaction and nurse burnout. *Social Science & Medicine* **47**: 1611-1617.

Lewis JB, Linzer DA. 2005. Estimating Regression Models in Which the Dependent Variable Is Based on Estimates. Political Analysis 13: 345–64.

Linna M, Häkkinen U. 2006. Reimbursing for the costs of teaching and research in Finnish hospitals: A stochastic Frontier Analysis. *International Journal of Health Care Finance and Economics* **6**: 83-97.

Lippi Bruni M, Mammi I. 2015. Spatial effects in hospital expenditures: a district level analysis. *Working paper DSE*, n. 1027.

Luo Q, Wang Q, Lu Z, Liu J. 2013. Evaluation of Responsiveness of Community Health Services in Urban China: A Quantitative Study in Wuhan City. *PLoS ONE* **8**(5), e62923. doi:10.1371/journal.pone.0062923.

Meterko M, Mohr DC, Young G. 2004. Teamwork Culture and Patient Satisfaction in Hospitals *Medical Care* **42** (5): 492-498.

Murante AM, Seghieri C, Brown A, Nuti S. 2014. How do hospitalization esperience and institutional characteristics influence inpatient satisfaction? A multilevel approach. *The International Journal of Health Planning and Management* **29**: e247-e260.

Murray C, Frenk J. 2000. A Framework for Assessing the Performance of Health Systems. *Bulletin of the World Health Organization* **78**: 717–31.

National Institute for Health and Clinical Excellence (NICE). 2012. *Patient experience in adult NHS services: improving the experience of care for people using adult NHS services*, National Clinical Guideline Centre, London.

Peracchi F, Rossetti C. 2013. The heterogeneous thresholds ordered response model: identification and inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **176**: 703–722.

Pudney S, Shields M. 2000. Gender, Race, Pay and Promotion in the British Nursing Profession: Estimation of a Generalized Ordered Probit Model. *Journal of Applied Econometrics* **15** (4): 367-399.

Rice N, Jones A. 1997. Multilevel models and health economics. *Health Economics* **6**: 561-75.

Rice N, Robone S, Smith PC. 2010. International Comparison of Public Sector Performance: The Use of Anchoring Vignettes to Adjust Self-Reported Data. *Evaluation* **16** (1): 81-101

Rice N, Robone S, Smith PC. 2011. Analysis of the Validity of the Vignette Approach to Correct for Heterogeneity in Reporting Health System Responsiveness. *The European Journal of Health Economics* **12** (2): 141-162.

Rice N, Robone S, Smith PC. 2012. Vignettes and health systems responsiveness in cross-country comparative analyses. *The Journal of The Royal Statistical Society*, *Series A* **175** (2): 337-369.

Robone S, Rice N, Smith PC. 2011. Health systems' responsiveness and its characteristics: a cross-country comparative analysis. *Health Services Research* **46**(6): 2079-2100.

Robone S, Fiorentini G, Nicoli MA, Rodella S. 2014. La responsiveness dei sistemi sanitari: un'analisi empirica sull'assistenza ospedaliera nel Servizio Sanitario Regionale dell'Emilia-Romagna *Politiche sanitarie* **15**: 1-14.

Robone S. 2016. Can we use vignettes to address response-scale heterogeneity in the EQ-5D? Not if but how. *Health Economics* (Commentary) *forthcoming*.

Röttger J, Blümel M, Fuchs S, Busse R. 2014. Assessing the responsiveness of chronic disease care - Is the World Health Organization's concept of health system responsiveness applicable?. *Social Science & Medicine* **113**: 87-94.

Sirven N, Santos-Eggimann B, Spagnoli J. 2012. Comparability of Health Care Responsiveness in Europe. Using anchoring vignettes from SHARE. *Social Indicators Research* **105**(2): 255-271.

Sjetne IS, Veenstra M, Stavem K. 2007. The Effect of Hospital Size and Teaching Status on Patient Experiences with Hospital Care: A Multilevel Analysis. Medical Care 45 (3): 252-258.

Street A, Kobel C, Renaud T, Thuilliez J. on behalf of the eurodrg group. 2012. How well do diagnosis-related groups explain variations in costs or length of stay among

patients and across hospitals? methods for analysing routine patient data. *Health Economics* **21**(Suppl. 2): 6–18.

Vahey DC, Aiken LH, Sloane DM, Clarke SP, Vergas D. 2004. Nurse Burnout and patient satisfaction. *Medical Care* **42**: II57-II66.

Valentine NB, De Silva A, Kawabata K, Darby C, Murray CJL, Evans D. 2003a. Health System Responsiveness: Concepts, Domains and Operationalization. In *Health Systems Performance Assessment: Debates, Methods and Empiricism*, edited by CJL Murray and DB Evans, pp. 573–96. Geneva: World Health Organisation.

Valentine NB, Ortiz JP, Tandon A, Kawabata K, Evans DB, Murray CJL. 2003b. Patient Experiences with Health Services: Population Surveys from 16 OECD Counties. In *Health Systems Performance Assessment: Debates, Methods and Empiricism*, edited by CJL Murray and DB Evans, pp. 643–52. Geneva: World Health Organisation.

Valentine NB, Darby C, Bonsel GJ. 2008. Which aspect of non-clinical quality of care are most important? Results from WHO's general population surveys of "health system responsiveness" in 41 countries. *Social Science and Medicine* **66**: 1939-1950.

Valentine NB, Prasad A, Rice N, Robone S, Chatterji S. 2009. Health Systems Responsiveness – A Measure of the Acceptability of Health Care Processes and Systems. In *Performance Measurement for Health System Improvement: Experiences, Challenges and Prospects*, edited by PC Smith, E Mossialos, and S Leatherman, pp. 138–86. London: WHO European Regional Office.

Valentine N, Verdes-Tennant E., Bonsel G. 2015. Health systems' responsiveness and reporting behaviour: Multilevel analysis of the influence of individual-level factors in 64 countries. *Social Science & Medicine* **138**: 152-160.

Terza JV. 1985. Ordinal probit: a generalization. *Communication in Statistics*. **14**(1): 1-11.

Young GT, Meterko M, Desai KR. 2000. Patient satisfaction with hospital care: Effects of demographic and institutional characteristics. *Medical Care* **38**(3): 325-334.

White H.1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for the Heteroskedasticity. *Econometrica,* **48**: 817–38.

World Health Organization. 2000. *The World Health Report 2000. Health Systems: Improving Performance*. Geneva: World Health Organization.

World Health Organization. 2008. *The Tallin Charter: Health Systems for Health and Wealth*. WHO European Ministerial Conference on Health Systems, Tallin, Estonia, 25-27 June, 2008.

Table I. Domains of responsiveness

Autonomy: respect for patients' views of what is appropriate, and allowing patients to make informed choices.
Choice: an individual's right or opportunity to choose a health-care institution and health provider, and to request a second opinion and access specialist services when required.
Clarity of communication: the offering of a clear explanation to patients and family regarding the nature of the illness, together with details of treatment and of any available options.
Confidentiality of personal information: privacy in the environment in which consultations are conducted, and the concept of the privileged communication and confidentiality of medical records.
Dignity: the opportunity for patients to receive care in a respectful, caring, non-discriminatory setting.
Prompt attention: the opportunity to receive care rapidly in emergencies, or readily with short waiting times in the case of non-emergencies.
Quality of basic amenities: the physical environment and services often referred to as 'hotel facilities', including clean surroundings, regular maintenance, adequate furniture, sufficient ventilation and adequate space in waiting rooms.
Access to family and community support: the extent to which patients have access to their family and friends when receiving care, and the maintenance of regular activities (e.g. the opportunity to carry out religious and cultural practices).

Notes: Source: Rice et al. (2012). The eight domains of responsiveness are defined by the World Health Organization (see Valentine *et al.* (2003a) for a full exposition of these domains).

Table II. Correspondence between the standard domains of responsiveness, as defined by Valentine et al. (2003a), and the items present in the questionnaire administered by the Agency for Health Care and Social Services (ASSR) of Emilia-Romagna

Responsiveness domains	Items included in the questionnaire of the ASSR
	How would you evaluate the:
Communication	- information received before being hospitalized?
	- information received when hospitalized?
	- information received from your doctor about your state of health?
	- willingness of nurses to clarify things for you?
	- information received from your doctor about possible risks of your treatment?
	- information received about the treatment required?
	- information received about the checks up you require following your discharge from
	- information received about the symptoms to be monitored?
	- information received about the staff to contact in case of need?
	- information received about how to cope with physical pain?
Confidentiality	- respect of your privacy when dealing with sensitive information?
	- respect of your privacy when receiving confidential treatment?
Dignity	- kindness and respectfulness of nurses?
	- kindness and respectfulness of doctors?
	- ability of nurses to make you feel comfortable?
Prompt attention	- waiting time between the booking of a hospital admission and the admission itself?
	- waiting time between your arrival at the ward and the admission itself?
Quality of facilities	- cleanliness of the rooms and corridors?
	- cleanliness of bathrooms?
	- cleanliness of bedding?
	- availability of space in the rooms?

Notes: Source: Fiorentini et al. (2015).

Table III. Descriptive statistics of respondents (No. of observations = 38,696)

Variable name	Variable description	Mean	SD
woman	= 1 if female, 0 otherwise	0.509	0.500
single	= 1 if single, 0 otherwise	0.434	0.496
high education	= 1 if secondary education or higher, 0 otherwise	0.630	0.483
work	= 1 if currently working, 0 otherwise	0.301	0.459
outside the province	= 1 if residing outside the hospital's catchment area, 0 otherwise	0.124	0.329
outside the region or country	= 1 if residing outside the region where the hospital is located, 0 otherwise	0.096	0.294
poor/moderate health	= 1 if reporting poor/moderate health = 1, 0 otherwise	0.342	0.475
pain	= 1 if experienced pain during the hospital stay, 0 otherwise	0.511	0.500
emergency	= 1 if Accident & Emergency hospitalization, 0 otherwise	0.496	0.500
previous hospitalisation	= 1 if already admitted to the same ward before, 0 otherwise	0.330	0.470

Table IV. Frequency of individuals choosing each response category, for each responsiveness domain

Response category	Communication		Confidentiality		Dignity		Prompt attention		Quality of facilities	
	N	%	N	%	N	%	N	%	N	%
completely/very unsatisfied	2,198	1.39	420	1.21	503	0.94	737	2.54	1,058	1.48
unsatisfied	5,774	3.65	1188	3.42	983	1.83	1542	5.32	2,848	4.00
satisfied	66,879	42.29	15,267	43.99	14,525	27.04	11,742	40.53	28,156	39.50
very satisfied	43,807	27.7	9,076	26.15	18,831	35.06	6,877	23.74	21,048	29.53
completely satisfied	39,496	24.97	8,758	25.23	18,872	35.13	8,074	27.87	18,170	25.49
Total	158,154	100	34,709	100	53,714	100	28,972	100	71,280	100

Table V. Descriptive statistics

Variable name	Variable description	Obs	Mean	SD
Waiting times	Median waiting times (weeks)	165	3.269	3.890
Medical staff workload	Ratio between the number of patients discharged (in 1000 cases) and total medical staff	165	0.137	0.063
Nursing staff workload	Ratio between the number of patients discharged (in 1000 cases) and total nursing staff	165	0.136	0.174
Level of spending on non-clinical goods and services	Per patient hospital expenditure on non-clinical goods and services (Euros 000), including food, laundry, materials for wardrobe and for cleanliness	165	0.107	0.106
Level of spending on staff education and training activities	Per DRG weight hospital expenditure on staff education and training activities (Euros 000)	165	0.019	0.021
Proportion of spending on nursing staff	Proportion of expenditure invested on nursing staff over hospital expenditure on total staff	165	0.432	0.095
Proportion of spending on administrative staff	Proportion of expenditure invested on administrative staff over hospital expenditure on total staff	165	0.052	0.010
Beds	Number of available ordinary beds	165	24.738	21.836
DRG case mix	Average DRG weight assigned to patients discharged from each hospital-specialty unit	165	1.224	0.500

Table VI. First-step regression results

	Communication	Confidentiality	Dignity	Prompt attention	Quality of facilities
	b/se	b/se	b/se	b/se	b/se
xb					
woman	0.020***	0.027*	0.016	-0.000	-0.016
	(0.006)	(0.013)	(0.010)	(0.014)	(0.009)
high education	0.099***	0.058***	0.113***	0.106***	0.049***
	(0.007)	(0.014)	(0.012)	(0.015)	(0.01)
work	-0.004	-0.014	0.064***	0.064***	-0.039***
	(0.007)	(0.015)	(0.012)	(0.016)	(0.011)
single	-0.047***	0.001	-0.026*	-0.016	0.002
	(0.006)	(0.013)	(0.010)	(0.014)	(0.009)
outside the province	0.034***	0.052**	0.040*	0.018	0.046***
	(0.009)	(0.020)	(0.016)	(0.021)	(0.014)
outside the region or the country	0.174***	0.172***	0.164***	0.142***	0.171***
	(0.010)	(0.022)	(0.018)	(0.024)	(0.016)
emergency	-0.074***	-0.025	-0.035**	-0.002	0.020*
	(0.007)	(0.014)	(0.011)	(0.015)	(0.010)
previous hospitalization	-0.019**	-0.045***	-0.050***	-0.040**	-0.062***
	(0.006)	(0.013)	(0.011)	(0.014)	(0.009)
cut-point 1					
poor/moderate health	0.360***	0.255***	0.226***	0.250***	0.221***
	(0.017)	(0.039)	(0.034)	(0.032)	(0.025)
pain	0.156***	0.108**	0.089**	0.077*	0.029
	(0.017)	(0.038)	(0.032)	(0.033)	(0.025)
constant	-2.443***	-2.388***	-2.452***	-1.991***	-2.289***
	(0.022)	(0.046)	(0.038)	(0.043)	(0.030)
cut-point 2					
poor/moderate health	0.200***	0.220***	0.471***	0.057	0.140***
	(0.026)	(0.058)	(0.063)	(0.049)	(0.036)
pain	0.014	0.006	0.218***	0.052	0.164***
	(0.026)	(0.055)	(0.056)	(0.049)	(0.037)
constant	-0.654***	-0.637***	-1.207***	-0.633***	-0.663***
	(0.028)	(0.055)	(0.065)	(0.044)	(0.035)
cut-point 3					
poor/moderate health	0.017*	0.042**	0.008	0.031	0.066***
	(0.007)	(0.015)	(0.016)	(0.016)	(0.011)
pain	-0.058***	-0.044**	-0.066***	-0.042**	-0.036***
	(0.007)	(0.015)	(0.016)	(0.016)	(0.011)
constant	0.546***	0.564***	0.437***	0.381***	0.451***
	(0.006)	(0.013)	(0.015)	(0.013)	(0.009)
cut-point 4					
poor/moderate health	-0.114***	-0.123***	-0.095***	-0.179***	-0.056***
	(0.009)	(0.021)	(0.013)	(0.024)	(0.013)
pain	-0.044***	-0.034	-0.059***	-0.046*	-0.022
	(0.008)	(0.019)	(0.012)	(0.022)	(0.012)
constant	-0.182***	-0.243***	0.040***	-0.340***	-0.144***
	(0.006)	(0.014)	(0.010)	(0.016)	(0.009)
N	158,154	34,709	53,714	28,972	71,280

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors clustered at the hospital level are in parentheses.

Table VII. Regression results of responsiveness on supply-side factors including level of spending on non-clinical goods and services

	Communication b/se	Confidentiality b/se	Dignity b/se	Prompt attention b/se	Quality of facilities b/se
Waiting times	-0.0103	-0.00671	-0.00951	-0.0214***	-0.00237
	(0.00841)	(0.00821)	(0.00904)	(0.00645)	(0.00826)
Medical staff workload	0.0518	0.307	0.408	0.221	0.0923
	(0.504)	(0.430)	(0.361)	(0.445)	(0.450)
Nursing staff workload	-0.319**	-0.175***	-0.121	-0.0626	-0.155
	(0.129)	(0.0611)	(0.111)	(0.164)	(0.115)
Level of spending on non-clinical goods and services	0.943***	1.097***	0.825***	0.664***	1.094***
	(0.216)	(0.275)	(0.233)	(0.199)	(0.268)
Proportion of spending on nursing staff	0.405	0.588**	0.769**	0.532*	0.613*
	(0.279)	(0.236)	(0.287)	(0.262)	(0.350)
Proportion of spending on administrative staff	0.760	-0.367	-1.708	6.084***	-2.846
	(3.410)	(3.072)	(3.822)	(1.896)	(2.859)
Beds	-0.00306***	-0.00316***	-0.00388***	-0.00379***	-0.00357***
	(0.000549)	(0.000675)	(0.000731)	(0.000762)	(0.00107)
DRG case mix	0.0715*	0.0237	0.112***	0.0367	0.146***
	(0.0386)	(0.0372)	(0.0332)	(0.0269)	(0.0458)
Constant	-0.0111	-0.0265	-0.182	-0.197	-0.214
	(0.185)	(0.175)	(0.136)	(0.150)	(0.308)
N	165	165	165	165	165
R^2	0.309	0.278	0.308	0.313	0.229
RESET	p-value 0.519	p-value 0.867	p-value 0.933	p-value 0.269	p-value 0.680

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models also include hospital Trust dummies. Robust standard errors clustered at the hospital level are in parentheses.

Table VIII. Regression results of responsiveness on supply-side factors including level of spending on staff education and training activities

	Communication b/se	Confidentiality b/se	Dignity b/se	Prompt attention b/se	Quality of facilities b/se
Waiting times	-0.00889	-0.00512	-0.00830	-0.0199***	-0.000730
	(0.00866)	(0.00852)	(0.00919)	(0.00625)	(0.00864)
Medical staff workload	-0.0343	0.202	0.330	0.179	-0.00968
	(0.517)	(0.431)	(0.370)	(0.429)	(0.459)
Nursing staff workload	-0.354**	-0.206***	-0.146	-0.127	-0.192*
	(0.138)	(0.0586)	(0.102)	(0.153)	(0.0945)
Level of spending on staff education and training activities	4.790***	5.369***	4.060***	4.193***	5.470***
	(1.268)	(1.788)	(1.378)	(1.313)	(1.752)
Proportion of spending on nursing staff	0.446	0.637**	0.806**	0.552**	0.661*
	(0.287)	(0.245)	(0.294)	(0.263)	(0.363)
Proportion of spending on administrative staff	1.251	0.159	-1.308	6.612***	-2.296
	(3.370)	(3.033)	(3.836)	(1.796)	(2.923)
Beds	-0.00278***	-0.00283***	-0.00363***	-0.00358***	-0.00324***
	(0.000516)	(0.000716)	(0.000677)	(0.000705)	(0.00102)
DRG case mix	0.111**	0.0679	0.145***	0.0690***	0.191***
	(0.0422)	(0.0400)	(0.0396)	(0.0233)	(0.0536)
Constant	-0.0971	-0.121	-0.253	-0.278**	-0.312
	(0.184)	(0.187)	(0.151)	(0.134)	(0.320)
N	165	165	165	165	165
R^2	0.299	0.261	0.300	0.322	0.220
RESET	p-value 0.563	p-value 0.967	p-value 0.775	p-value 0.167	p-value 0.595

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models also include hospital Trust dummies. Robust standard errors clustered at the hospital level are in parentheses.

Appendix I.

The sample used at the first-step of our analysis contains patient-level data collected by the Agency for Health Care and Social Services of Emilia-Romagna (ASSR), as we briefly describe in section 3. Below we provide the details about the survey design and the sampling procedure.

The questionnaire was developed by the ASSR following the requirement of the Italian legal system, and a pilot study was designed to collect data for years 2008-2009. During the two-year pilot study period, patients were interviewed and asked to rank a number of aspects of their hospital experience, which were described in the following order: reception, relationship with the staff (in terms of communication, dignity and kindness), privacy, pain treatment, basic amenities (i.e. rooms, toilets, meals), discharge. Only five out of eight domains of responsiveness were retained in the final questionnaire, which were those rated as the most important by survey respondents.

The sampling design implied two steps. First, a four-months' time period was randomly selected in the year chosen to run the survey. Second, a stratified sampling method was adopted, with allocation proportional to the last year of discharge. The sample size calculation was based on the use of the following formula:

$$n = N * z^2 * p * (1-p) \, / \, [(N-1) * e^2 + z^2 * p * (1-p)]$$

where n is the sample size, N is the number of discharges within each single hospital in the last 12 months from data collection, z is the normal variable value (level of confidence = 95%), e is the sampling error (equal to 0.05), and p is the proportion of assumed satisfied patients in the total population (equal to 0.5).

The questionnaire form was received by patients the day before they were discharged. Once this was completed, patients were asked to put it into a box in the lobby of their ward of hospitalization. In every ward where the questionnaire was handed out, an internal supervisor of the survey, usually the head nurse, was given the aim to provide general information to patients and caregivers, to promote their participation to the survey and to collect the filled questionnaires. Internal supervisors trained ward staff or volunteers, who were in charge of informing patients and caregivers about the purpose of the survey, and of assuring them about the anonymity and confidentiality of their responses.

Appendix II. List of specialties included in the final sample used in the second step analysis

Specialty
Cardiology
Child neuropsychiatry
Dermatology
Ear, nose and throat
Endocrinology
Gastroenterology
General medicine
General paediatrics
General surgery
Geriatric medicine
Haematology
Heart surgery
Infectious diseases
Maxillofacial surgery
Nephrology
Neurology
Neurosurgery
Nuclear medicine
Obstetrics and gynaecology
Oncology
Ophthalmology
Orthopaedics
Paediatric surgery
Physiopathology
Plastic surgery
Pneumology
Renal transplantation
Thoracic surgery
Urology
Vascular surgery