Social Interaction in Patients' Hospital Choice: Evidences from Italy *

F. Moscone[†]
Brunel University

E. Tosetti[‡]
University of Cambridge

G. Vittadini[§]
University of Milan, Bicocca

April, 2009

Abstract

In this paper we study the influence of social interaction on patients' hospital choice and its relationship with quality delivered by hospitals, using Italian data. We explore the impact on individual choices of a set of variables such as travel distance, individual- and hospital-specific characteristics, as well as a variable capturing the effect of the neighbourhood. The richness of our data allows us to disentangle contextual effects from the influence of information sharing on patients' hospital choices. We then use this framework to assess how such interaction is related to clinical hospital quality. Results show that network effect plays an important role in hospital choices, although it is less relevant for larger hospitals. Another empirical finding is the existence of a negative relationship between the degree of interaction among individuals and the quality delivered by hospitals. The absence of a source of information on the quality of hospitals accessible to all individuals, such as guidelines or star ratings, exacerbates the importance of information gathered locally in hospital choices, which may result in a lower degree of competition among hospitals and lower quality.

Keywords: health care, social interaction, quality.
JEL classification: I18, H4.

*We thank Paolo Berta, Arne Risa Hole, and Luca Merlino for helpful discussion.
†e-mail: francesco.moscone@brunel.ac.uk
‡e-mail: et268@cam.ac.uk.
§e-mail: giorgio.vittadini@unimib.it.

1 Introduction

In this paper we empirically study the role that social interaction has on the demand for health care in Italy. We investigate whether the choice of an hospital by patients with cardiac illness is influenced by information shared with their peers. Our hypothesis is that individuals, before deciding in which hospital to be treated, may seek advice by speaking with friends, relatives or trusted persons experiencing similar health problems in what has been termed by Freidson (1960) as *lay referral network*.

Empirical literature has supported the important role of social influences in explaining individual choices regarding a large variety of economic, social, and health behaviours (see Brock and Durlauf (2001) and Birke (2009) for a survey). For example, there is evidence that interaction among economic agents has an impact on unemployment (Conley and Topa (2002)), criminality (Glaeser, Sacerdote, and Scheinkman (1996)), the demand for addictive goods (Jones (1994)), or the adoption of technological standards (Skinner and Staiger (2005)). The study of how social interaction affects health services utilization has been first investigated by Freidson (1960). The author argued that a patient, before seeking professional advice, usually consults an informal network made of, for example, family and friends. A number of works, also by the means of interviews and surveys, have attempted at identifying such network effect[1] on the choice of an health specialist, and its influence on individual's health status (e.g., see Schoenberg et al. (2003), and Chaix et al. (2008), Cornford and Cornford (1999)). For example, Schoenberg et al. (2003), and Chaix et al. (2008) have provided evidence of a relationship between lay referral patterns and medical care seeking for patients with myocardial infarction, emphasising an increasing effect of the neighbourhood on patients' survival probability. Aizer and Currie (2004) showed that the use of public prenatal and delivery services in California was correlated within groups defined by race, ethnicity and zip-code, though such correlation was not found to be linked to information sharing. A recent study by Deri (2005) on Canadian data, has detected strong interdependence in the decisions of neighbouring people to visit a GP or a dentist, due in particular to norm, and transmission of information.

In the literature that studies the determinants of hospital choices, the role of social interaction has not been explored yet. However, network effects in patients' choices are likely to be strong, especially in health care systems where there is no comparative information on the quality of hospitals available to all citizens, like in the Italian case. If social influence in hospital choices is found to exist, one important research question is whether using information from the network increases the likelihood of choosing an high quality hospital. Thus, in this paper we will also investigate how the sensitiveness

[1] We use the terms social interaction, network effects and peer influences as synonimous, to indicate what has been called by Manski (1993) endogenous effect.

of patient's choice to local information is associated with hospital quality indicators based on health outcomes. Interacting and sharing information with neighbours does not necessary help to choose a high quality hospital. For example, the reference group may give importance to attributes, such as appearance, comfort, and convenience of hospitals (the so-called amenities, see Goldman and Romley (2008) and Romano and Mutter (2004)), which are not necessarily related with clinical quality; in the absence of comparative information on the quality of all hospitals, local GPs may involuntary direct patients to poor quality institutions.

Studying the effect of interaction among individuals on hospital choice, and its impact on the quality delivered by hospitals has important policy implications. If such interaction is found to exist, and is negatively related with quality then policy makers (e.g., the local authority) should put efforts on implementing mechanisms of information diffusion, for example, by making available to citizens guidelines and comparative information on hospitals quality. In the case of a positive relation between network effects and hospital quality, although on average information sharing leads to better quality, policy makers are called to intervene to reduce geographical inequality in the access to information. As shown (at aggregate level) by a recent strand of literature in public economics (see, for example, Revelli (2006)), interaction may be reduced over time by introducing, for example, publicly released star rating indicators on the performance of hospitals.[2]

We use data on 144 Italian hospitals from the Lombardy region and on all patients being admitted to these hospitals for a cardiac illness, in the years from 2004 to 2007. Focusing the analysis on this region, rather than the entire nation, allows us to restrict the attention to a competitive health system, as established by the 1997 regional health reform (see Section 2 for a description of this reform and how it has introduced competition among hospitals). Another advantage of limiting the study to Lombardy hospitals is that we reduce the heterogeneity that arises from different rules underlying the health systems of other Italian regions.

We consider all patients whose source of admission was elective[3] or the emergency room. We note that flows of patients admitted via the emergency room are governed by rules that are different from those driving flows of elective patients. An individual requiring emergency often cannot choose her hospital, since her admission is mainly determined by external factors such as the availability of beds and the ambulance service. Tay (2003), using data on patients admitted for AMI to US hospitals in 1994, detected that one-half of heart attack patients arrive at the hospital via ambulance. Even

[2] In particular, Revelli (2006) has provided evidences that interaction among municipalities in social care resources distribution reduces over time after the publicly released star rating indicators for the performance of local authorities.

[3] We define as *elective* all booked or planned admissions, where patients have been given a date or approximate date at the time the decision to admit was made.

in the case where these individuals are able to make a decision, it is very unlikely that they have the opportunity to engage into social interaction before deciding the hospital where to be treated. Therefore, it is reasonable to assume for these patients that they do not use information from the network to make a choice. On the contrary, elective patients have the time to gather information, consult other people with similar health problems before selecting their hospitals. For these individuals it is plausible that the neighbourhood have some influence on their choices. In this paper we exploit differences between urgent and elective patients in terms of use of network information to identify and measure the effect of social interaction on hospital choices. Under the assumption that urgent patients do not exploit the network to make a decision, we use information on these patients to identify contextual (and correlated) factors in hospital choices (see Manski (1993)), and interpret the remaining correlation within neighbourhood as the effect of pure social interaction.

Another point to observe is that, since urgent patients need immediate care, typically at the closest hospital with an emergency department, their choice set will be very limited and confined around the place where they live. For this reason, the literature studying admissions of urgent patients typically considers as potentially relevant market for each patient the set of hospitals within a short, predetermined distance (for example, 10 miles) from the patient (Tay (2003), Romano and Mutter (2004), and Volpp, William, Waldfogel, Silber, Schwartz, and Pauly (2003)). On the contrary, for elective patients the set of choice will be wider and less constrained by geographical factors than in the case of emergency care. For these people, all hospitals compete with each other, and their potentially relevant geographical market is the entire region. Accordingly, in our model of hospital choice, as it includes both types of patients, we will allow the choice set to vary across patients to account for local and global choice behaviour.

The remainder of the paper is organized as follows. Section 2 briefly describes the Lombardy health system and the reform that has introduced competition among hospitals. Section 3 describes the data set. Section 4 discusses the role of social interaction in the choice of the health provider. Accordingly, it introduces a model for patients' choice that includes a measure of social interaction. Section 5 estimates the relationship of local interaction and a set of quality indicators. Section 6 comments on the empirical findings. Section 7 concludes.

2 The health care pro-competition reform in Lombardy

The last decade has witnessed a deep institutional change in the Italian National Health Service (NHS), which has gradually transferred the responsibilities of financing and managing health care services from the central system to the regions. This has lead to a marked heterogeneity in the supply of health care services across Italian regions.

The Lombardy region has been the first to implement, through the 1997 regional health reform, an innovative health care model that promotes competition among agents and increases patients' choice, with the ultimate aim to improve the quality of health care services and reduce costs. The reform has introduced a net distinction between the role of Local Health Authorities (ASL) and that of hospitals within the health care system. While the ASL are responsible for programming, financing and controlling the quality and quantity of NHS activities in their target area, hospitals provide health care services purchased by the ASL. Such distinction between the purchaser (the ASL) and the provider (the hospital) has lead the former to develop tools for monitoring the quality of providers, and the latter to search for quality and technical efficiency. The health reform has also introduced competition between public and private hospitals, by allowing the latter to provide free health care. To enter such competition, health providers are required to satisfy minimum technology and organisational standards set by the region[4]. While patients are assigned to the ASL on the basis of their place of residence, they have choice of receiving free heath care in any (accredited) hospital of the region. In Section 6 we provide some statistics on migration flows of patient between areas.

Since 1995, the Lombardy region has implemented the financing mechanism known as Prospective Payment System (PPS). This is a financing system where a predetermined, fixed reimbursement is paid to the hospital for each patient, on the basis of his/her Diagnosis Related Group (DRG), established using clinical information reported in the Hospital Discharge Chart (HDC). The reimboursement for a particular DRG does not vary if the length of stay falls within a threshold. The tariff and the threshold rule paid for each DRG is set at a regional level[5], and covers all health care services relative to hospital admissions, as well as outpatients activity[6]. The tariff scheme is updated at irregular intervals, and may change even more than once within a year.

A great proportion of resources for financing the Lombardy health care system is tax based, although private resources, like insurances and out-of-pocket, are also significant, accounting for circa 27 per cent of total spending. These funds are employed to cover inpatients and outpatients activities, and are also linked with the variability of hospital functions, as well as hospital debts that need to be compensated.

The changes introduced by the 1997 reform and discussed above have determined a significant transformation in the supply and demand for health care. First, the number of private health care

[4] Private hospitals satisfying such standards are indicated as accredited.

[5] Since it is set at a regional level, it does not vary across hospitals.

[6] Notice that also other Italian regions are in the process of adopting the PPS. However, to date, the Lombardy is the only region that has de facto implemented it.

facilities has increased, boosting the total number of providers from 181 to 193 in the years from 1995 to 2006. Public hospitals have reduced their number of beds for ordinary admission, while increasing those for day hospital, and rehabilitation. Another effect of this reform is that hospital attraction of patients from other Italian regions has significantly increased. In the years from 1995 to 2003, the number of patients from other regions admitted to Lombardy hospitals has increased by 34 per cent. For further details on the Lombardy reform we refer to Amigoni, Marchetti, Merlino, and Zangrandi (1998) and Zangrandi (1998).

3 Data sources and sample construction

We gathered administrative data on all patients admitted to any hospitals in Lombardy, in the years from 2004 to 2007, whose principal diagnosis is an ischemic heart disease.[7] According to the International Classification of Diseases, 9h revision, Clinical Modification (ICD-9-CM)[8] published by the World Health Organization, these can be subdivided into five categories: acute myocardial infarction (AMI), other acute and subacute forms of ischemic heart disease, old myocardial infarction, angina pectoris, and other forms of chronic ischemic heart disease. We removed from the data set any patient whose source of admission was other than the emergency room or elective (planned or ordinary admission). As shown in Table 2, 60 to 80 per cent of AMI and other acute forms of ischemic hearth disease admissions (i.e., for ICD-9-CM=410, 411) are an emergency. In contrast, only 15-24 per cent of admissions for diseases belonging to the remaining ICD-9-CM categories are an emergency, the rest being ordinary or planned. In this case patients have the time to gather information, perhaps consulting other people, and plan their choice.

Data on patients have been extracted from the Hospital Discharge Chart (HDC) available for each patient. These include socio-demographic characteristics such as age, gender, and place of residence (the municipality); clinical information like principal diagnosis, severity of the illness, length of stay, the type of admission (planned or via the emergency room) the ward of admission, type of discharge (e.g. death); financial information such as the Diagnosis Related Group (DRG), and the HDC reimbursement. We also gathered information on zip-code of residence patients and their mortality from the General Register Office. The characteristics of the hospital include its capacity expressed in number of beds, the number of doctors employed, its ownership (e.g. private or public), teaching status,

[7] These data have been kindly provided by the Region of Lombardy, in conformity to all privacy regulations.
[8] For further information, see http://icd9cm.chrisendres.com/icd9cm/.

whether it is mono-specialist and if it has an catheterization laboratory[9], and the ASL to which the hospital belongs. We refer to Table 1 for a description of the variables used in our analysis.

We cleaned the data by eliminating records with missing entries on either the hospital or the patient identifier. Also, we only kept records for private hospitals that are accredited by the region (see Section 2 on this). After this selection process, our data set contains around 230,600 patients admitted in 144 hospitals.

4 Local interaction in patients' hospital choice

In health care systems with fixed prices and where patients have free choice of hospitals, the quality delivered by hospitals is an important determinant of individuals' choices. However, quality differences among hospitals may be difficult for people to observe. These constraints have raised concerns among policy makers on whether hospital markets are competitive, and encouraged initiatives to diffuse information about "true" hospital quality. Institutions such as the National Committee for Quality Assurance in the US and the Care Quality Commission in the UK diffuse reports with comparative information, or star-rating indicators, on the quality of hospitals in terms of rates of post-operative mortality, hospital acquired infections, and readmission rates. We observe that the influence of quality reports on individual choice of hospital is still controversial. A number of studies at individual level have determined a low influence of these reports on the selection of the health provider (see, for example, Schneider and Epstein (1996)). Other studies have found a positive relationship between the published hospital quality and the market share of an hospital, showing that the demand reacts especially when the published actual quality deviates significantly from expected quality (see, for example, Pope (2009) and Romano and Zhou (2004)).

Since true hospital quality is difficult to observe, and choosing a low quality hospital could be costly, individuals try to get as much information as possible when making a choice. Therefore, it may be sensible to use information about the decisions of others with the same pathology, who have had a comparable decision to make. Friends, relatives or trusted persons who have experienced a similar health problem may act as filters for the quality of hospitals, thus shaping preferences of individuals. Individuals may also seek assurance as to whether their thinking is reasonable, by looking if people with similar features have come to the same conclusion. These processes may be more relevant in health care systems where there are no measures publicly available on the performance of hospitals, like in the Italian case. We observe that in systems that provide star-rating indicators, information

[9] A catheterization laboratory is an examination room with diagnostic imaging equipment used to support catheterization procedures.

gathered locally can reinforce or be in contrast with that provided at central level by the star ratings.

It is plausible that an individual admitted at a particular point in time may not observe choices of people admitted in the same period but can easily gather information from the choices of patients admitted in the past. Therefore, in this paper we assume as neighbours for a patient admitted at time t all individuals sharing the same pathology, admitted (and discharged) to any hospital in the region in the 11 months before time t, alive after the hospitalization and living in the same zip-code.

As suggested by Manski (1993) and Brock and Durlauf (2001), correlation between the behaviour of a patient and that of her neighbourhood could reflect not only social influences, but also the effect of other factors. In particular, such interdependence may arise because of contextual effects, if individual action varies with observed attributes that define her group membership, or correlated effects, if individuals in the same group tend to behave similarly because they have similar characteristics or they face similar opportunities and constraints. For example, the decision to refer patients for hospital admission by the local general practitioner may induce contextual effects in hospital choices. An example of correlated effects is the behaviour of hospitals towards certain categories of patients. Indeed, some hospitals may encourage or discourage groups of individuals from presenting on the basis of whether it is profitable or not to treat them (the so-called cream-skimming effect, see Berta et al. (2009) on this). It is reasonable to assume that most contextual and correlated effects in Italy are the same regardless the type of admission of patients, i.e. whether it is elective or via the emergency room.

Our strategy to disentangle between social interaction and contextual or correlated effects is based on estimating the correlation between the behaviour of a patient and that of her neighbourhood for patients both on emergency and non-emergency care. Under the assumption that patients in emergency care cannot engage in communication with other people when choosing their hospital, the additive neighbourhood effect of routine patients is likely to reflect pure social interaction.

In the following section we introduce an econometric model for individuals' hospital choices, and define a measure of social interaction.

4.1 Modelling individual patient choices

Consider an individual i with cardiac illness from a population of N agents choosing from a set of H_i hospitals, i.e. from $\{1, 2, ..., H_i\}$, at time t. We assume that each individual is drawn randomly from a set of neighbourhoods, and that within each neighbourhood, all individuals interact with each other. Thus, membership to various neighbourhoods is not endogenously determined. Suppose that the observable choice of individual i of being admitted to hospital h at time t , $y_{ih,t}$, is related to the expected utility of i choosing h, $y^*_{ih,t}$, according to $y_{ih,t} = 1\left[y^*_{ih,t} > 0\right]$, where $1[.]$ is an indicator function.

The literature on social interaction decomposes $y^*_{ih,t}$ into three components, the private utility, the social utility, and a random utility term (Brock and Durlauf (2003)). In this paper we hypothesise that private utility of the ith patient from choosing hospital h at time t depends on her characteristics, $\mathbf{x}_{it}$, the characteristics of hospital h, $\mathbf{z}_{ht}$, and the distance of i from h relative to the distance from her nearest hospital, $d_{ih,t}$. We further assume that social utility depends on $\bar{y}_{ih,t-1}$, the percentage of people with identical disease category and living in the same zip-code who made the same choice in the 11 months prior to the admission of the ith patient and that are alive after hospitalization.[10] Accordingly, we model patient i's expected utility at time t from choosing h out of a total of H_i hospitals, $y^*_{ih,t}$, as

$$y^*_{ih,t} = \alpha + \theta_h \bar{y}_{ih,t-1} + \delta_h \left(\bar{y}_{ih,t-1} \cdot elective_{it}\right) + \boldsymbol{\beta}'_h \mathbf{x}_{it} + \boldsymbol{\gamma}' \mathbf{z}_{ht} + \eta d_{ih,t} + \varepsilon_{ih,t}, \tag{1}$$

where α, θ_h, δ_h, $\boldsymbol{\beta}_h$, $\boldsymbol{\gamma}$, and η are parameters to be estimated. Under the further hypothesis that each individual makes the choice that maximizes her total utility, and the double exponential assumption for the random utility term, the multinomial logit structure can be derived for the conditional probability that i chooses h (see Brock and Durlauf (2003)).

In the above model, $elective_{it}$ is an indicator variable taking value of 1 if individual ith's source of admission is elective and zero otherwise. The vector of individual-specific characteristics, $\mathbf{x}_{it}$, contains demographic, health, and geographic attributes, such as gender, age, whether patient ith's source of admission is elective, the disease category (ICD09), and a dummy variable indicating the province where the patient lives. We also include in the model interactions of $\bar{y}_{ih,t-1}$ with the distance patient-to-hospital and with a dummy variable indicating whether the patient is old[11], to see if the network is more influential with certain categories of population. The vector of hospital-specific characteristics, $\mathbf{z}_{ht}$, contains the size, the ASL of the hospital, the number of doctors per number of beds, ownership and teaching status, dummy variables indicating whether the hospital is mono-specialist, and if it has a catheterization laboratory. We also add in the model the interaction of $\bar{y}_{ih,t-1}$ with a dummy variable indicating whether the hospital is large, to check if network effects are weaker or stronger towards larger hospitals. We have incorporated dummies for the ASL in order to capture unobserved heterogeneity in health policies at ASL level, and province dummies to account for contextual effects, including recommendations by the local general practitioner. The influence on patients choice of the

[10] See Brock and Durlauf (2001) for a discussion of the dependence of social utility on past society behaviour.

[11] In this paper we define old patients as patients older than 75 years of age. See Table 1 for definition of variables.

specialist or the local GPs may also be captured by the hospital-specific characteristics (see Luft and et al. (1990) on this).

As explained at the beginning of this section, the coefficient δ_h attached to the variable $(\bar{y}_{ih,t-1} \cdot elective_{it})$ measures the correlation between individual behaviour and the behaviour of her neighbours as effect of pure social interaction. We allow this parameter to vary across hospitals. We remark that estimation of one coefficient for each hospital is possible only if there exists some geographical variability in patient flows within hospitals that enables identification of parameters. This is not a problem for our analysis, as shown in our exploratory data analysis. It is also worth nothing that, since the above model contains neighbours' lagged decisions, it does not entail any restriction on the size of the interaction effects, δ_h and θ_h. The interpretation of δ_h plays a central role in our study. $\hat{\delta}_h$, obtained from estimating equation (1), measures the average effect of neighbourhood choices on the probability that a patient chooses the hth hospital. A positive and significant $\hat{\delta}_h$ means that a subset of the population, sharing information on the quality of the hth hospital, increases the conditional probability of choosing it for each member of this subset. A negative and significant $\hat{\delta}_h$ implies that, *ceteris paribus*, a patient on average will make a choice different from that of her neighbours, in relation to the hth hospital. Namely, individuals choosing the hth hospital are surrounded by people that, on average, have not been admitted to that hospital in the past. The key mechanism underlying a significant $\hat{\delta}_h$, either positive or negative, is the existence of clusters of information on the quality of the hth hospital. Such information shapes the preferences of individuals, ultimately influencing their decisions. An insignificant $\hat{\delta}_h$ means that patients do not use information from the network to choose that hospital, and hence their choice is only driven by personal- and hospital-level characteristics.

Before concluding, we remark that in equation (1) we allow the choice set to vary across patients. As pointed in Section 1, the potentially relevant geographical market differs between patients admitted via the emergency room and all other patients. Therefore, for patients in emergency care we restrict the choice set to the hospitals within a distance of 15 kilometers from where they live. On the contrary, for patients not under emergency we extend the choice set to include all hospitals in the region. This allows us to distinguish in our choice model a more localised market where hospitals compete for urgent patients, against a global market where all hospitals compete with each others for non-urgent patients.

Whether there exists a relationship between δ_h and quality has important policy implications. In the next section we will introduce a measure of quality and discuss its link with social interaction.

5 Local interaction and the quality of health care

We estimate the relationship between social interaction and the quality of hospitals at individual level, using a regression framework. As quality indicators, we consider the outcome variables readmission and mortality within 30 days from the discharge, which are commonly used indicator in the literature. We refer to Romano and Mutter (2004) for a review of quality indicators adopted in the literature.

If the correlation between $\hat{\delta}_h$ (in absolute value) and quality[12] is either insignificant or negative and significant[13] then, on average, sharing information within the neighbourhood does not help to select an hospital with better quality. In this case, social interaction does not help to choose a high quality hospital. One possible reason behind such mismatch is, for instance, that the group identifies hospital quality not only with clinical quality, but also with amenities such as convenience, good food, attentive staff, and pleasant surroundings (see Goldman and Romley (2008) on this). In the case of positive and significant correlation between $\left|\hat{\delta}_h\right|$ and the level of hospital quality[14], local interaction is related to higher quality. In this case, the reference group is able to identify and suggest high quality hospitals.

We consider the following regression model for the latent continuous variable, $r^*_{ih,t}$, underlying our quality indicators[15]

$$r^*_{ih,t} = \boldsymbol{\beta}'\mathbf{x}_{it} + \boldsymbol{\gamma}'\mathbf{z}_{h,t-1} + \lambda\hat{\delta}_h + \varphi pr_{it} + \epsilon_{ih,t}, \tag{2}$$

where $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, λ and φ are parameters to be estimated, $\mathbf{x}_{it}$ indicates the individual-specific characteristics - age, gender, disease category and a dummy variable indicating whether the patient's source of admission was elective; $\mathbf{z}_{ht}$ is the vector of hospital attributes, namely number of beds, number of doctors per n. of beds, ASL dummies, ownership and teaching status, and whether the hospital is mono-specialist. The variables n. of beds and n. the doctors per n. beds have been lagged at time $t-1$ to avoid potential endogeneity problems. Other hospital-specific characteristics, such as whether the hospital is private, mono specialist or teaching-oriented can be considered as fixed attributes, established well before the time period considered in this analysis. The variable pr_{it} is the regulated (hence exogenous) price attached to the HDC of the ith individual. This variable is included in the regression to control for the effect of different reimboursments that are disease-specific, and for variations in the DRG reimboursements that occur within a year. The coefficient, λ, attached to $\hat{\delta}_h$

[12] In this paper we assume quality is inversely related to our quality indicators.

[13] Namely, if the correlation between $\left|\hat{\delta}_j\right|$ and our quality indicator - readmission within 30 days - is insignificant, or positive and significant.

[14] Namely, if the correlation between $\left|\hat{\delta}_j\right|$ and $q_{ij,t}$ is negative and significant.

[15] Also in this case we assume a logistic specification for the conditional probability.

indicates the sensitiveness of quality indicators to changes in social interaction.

6 Results

6.1 Exploratory data analysis

Table 3 shows some descriptive statistics that can be recovered from patients' HDC. As expected in the case of heart diseases, the number of males in the data set is high, accounting for around 75 per cent of the sample. The average age of patients is 68 years, and roughly 30 per cent of the sample has more than 75 years. The length of stay of patients reduces over time, passing from 9.14 to 8.56 number of days on average. The bottom panel of Table 3 summarizes the variables used to capture hospital quality in equation (2), namely readmission and mortality within 30 days from the date of discharge. The readmission variable has been constructed by including all patients that have been readmitted at least once in the period considered, also via the emergency room. Readmission within a fixed length of time as quality indicator has been employed in various studies on hospital quality, such as Kessler and McClellan (2000), and Ho and Hamilton (2000). The relatively high likelihood of 30-day readmission (around 9 per cent of the sample, see Table 3) suggests that this is an appropriate measure of hospital quality. As for 30-day mortality, the other quality indicator used in our analysis, Table 3 shows that this outcome concerns circa 5 per cent of our sample. Such small figure can be explained by the lower risk of dying of elective patients in our sample.

Table 4 summarizes the characteristics of hospitals in the data sets. We observe an increasing pattern in the average number of total beds, passing from around 260 to 276, indicating that hospitals tend to expand in size over time. Such trend is largely explained by the rise in number of ordinary beds. The number of doctors per number of beds ranges between 0.53-0.55.

Table 5 reports some descriptive statistics on migration flows of patients, as well as join-count measures of spatial correlation. The upper panel shows that around one third of the sample moves to an hospital based in a province different from that where they live. We observe that about 9 per cent of the sample comes from outside the region, making the average distance patient-to-hospital around 55 kilometers. However, when restricting the sample only to those living in Lombardy, the average distance patient-to-hospital drops to 12 kilometers. The central panel reports the average number of people living in the same neighbourhood that choose the same hospital. It is interesting to note that a large fraction of people living in the same zip-code and with similar disease (i.e., the same disease category) make similar choices, and that these figures tend to remain constant over time.

The lower panel of the table shows join-count statistics of spatial correlation. We adopt the following statistic

$$n_{H,t} = \frac{1}{2} \sum\nolimits_{h=1}^{H} \left(\sum_{i=1}^{N} \sum_{j=1, i \neq j}^{N} s_{ij} c_{ih,t} c_{jh,t} \right),$$

where $c_{ih,t} = 1$ if at time t individual i chooses hospital h, and zero otherwise, and $s_{ij} = 1$ when i and j belong to the same zip-code, and zero otherwise. Under the null hypothesis of independence in hospital choices of neighbouring individuals, this statistic has approximately mean zero. We refer to Epperson (2003) for a detailed discussion on the theoretical moments and the distribution of this statistic. If the null hypothesis of absence of spatial correlation is rejected and the statistic is significantly larger than its expected value, it indicates positive spatial autocorrelation, meaning that patients with similar hospital choices are more spatially clustered than could be caused by chance. The estimated $n_{H,t}$ statistic is positive and significant in all years, although it shows a slight decrease over time.

6.2 Modelling local interaction in patients' hospital choice

We estimated model (1) by maximum likelihood for each year, from 2005 to 2007. In the estimation of this model we only focused on patients living in the Lombardy region, to avoid potential heterogeneity in patients flows from other provinces of Italy.[16] We also dropped hospitals with less than 50 observations (i.e., patients) within a year, since estimation of one set of regression coefficients for each hospital requires enough observations for each hospital. As a robustness check, we also incorporated unobserved heterogeneity in (1) by including individual fixed effects, interacting individual-specific characteristics with hospital dummies, and then estimating the model via a conditional likelihood approach (see Greene (2008)). We further tried a specification with homogeneous coefficients and hospital dummies to account for hospital-specific characteristics that may affect individual choice and that are unobserved. Results for the parameters of the key variable $\bar{y}_{ih,t-1}$ are very similar across different specifications.

For descriptive purpouses, in Table 6 we report the output for the estimation of equation (1) for the years 2005 to 2007, imposing that all regression coefficients are homogeneous across hospitals. The coefficient of $\bar{y}_{ih,t-1}$ is positive and significant. This coefficient measures the correlation between the behaviour of a patient and that of her neighbourhood, due to contextual or correlated factors. For example, such correlation may arise if hospitals encourage (or discourage) certain categories of patients from being admitted, because their disease (i.e., their DRG), or the treatment they need is lucrative (or not) (Berta et al. (2009)). The coefficient of the variable $(\bar{y}_{ih,t-1} \cdot elective_{it})$ captures the effect

[16] When excluding patients living in provinces outside the Lombardy region, we dropped around 8 per cent of the sample. See Table 5 on this.

on hospital choices of the correlation within groups that is attributable to information sharing, since relative only to elective patients. This is positive and significant in all years, indicating that, after controlling for contextual and correlated effects, individuals try to access information on the quality of hospitals by observing the choices of their neighbours with similar pathology admitted in the past.

Gender and age of patients do not seem to play a role in the choice of the hospital. As expected, distance patient-to-hospital (relative to the distance from the nearest hospital) has a strong, negative impact on choices, implying that longer distance to undertake decreases the probability of choosing that hospital. The interaction term between $\bar{y}_{ih,t-1}$ and distance patient-to-hospital is also negative and significant, suggesting that network effects are less important when the choice concerns hospitals that are distant from the place of residence. This may be implied by the existence of a trade off between imitating neighbours' decisions and bearing the cost of moving to a distant hospital.

Among the hospital-specific characteristics, the coefficients of the variables n. of beds, n. doctors per n. of beds and whether the hospital is mono-specialist are positive and significant in all years considered, while the ownership status does not have a clear effect on the dependent variable. The interaction term between $\bar{y}_{ih,t-1}$ and the dummy variable indicating whether the hospital is large (i.e., with at least 300 patients) negatively impact on hospital choices. This indicates that network effects for large hospitals are less important than those of small and medium size hospitals. This may be explained by the fact that information on the quality of large hospitals is often more accessible to a larger fraction of the population, for example through national and regional media.

Table 7 reports results for the estimation of (2) for the years 2005 to 2007. Notice that, unlike the regression reported in Table 6, $\hat{\delta}_h$ has been obtained by estimating (1) where parameters are allowed to vary across hospitals.[17]

We first observe that, when using 30-day readmission as quality indicator, the coefficient attached to the variable $\hat{\delta}_h$ is positive and significant in all years. This result indicates that, *ceteris paribus*, a higher sensitiveness of patient's decision to local information decreases the probability of choosing a high quality hospital. One explanation for this result is that network effects, implied by asymmetric information, are a signal of low competition in the market, which in turn decreases quality (Kessler and McClellan (2000)). Another explanation for the negative relationship between network effects and quality of hospitals is that the reference group may give weight to hospital attributes, such as convenience or single room accommodation, which are not related with clinical quality, as measured by our health outcomes indicators. When adopting 30-day mortality as quality indicator, the coefficient attached to $\hat{\delta}_h$ is statistically insignificant. Thus, on average, using network information does not

[17] We obtained that most estimated coefficients $\hat{\delta}_j$, for $j = 1, ..., H$ are positive over time, with mean 1.89 and standard deviation 2.84. Only three hospitals showed a negative significant coefficient.

lead to select a higher (or lower) quality hospital, than if this was chosen without using network information.

As for the remaining regressors, the (exogenous) price variable shows a negative impact on quality indicators in all regressions, thus indicating that higher reimboursments lead to higher quality. The coefficient attached to n. of beds (lagged at time $t-1$) has a positive and significant effect on the 30-day readmission variable in all years. This may be explained by the fact that the severity of illness in larger hospitals is higher, thus inducing a higher readmission rates. However, this variable does not seem to play a role in explaining 30-day mortality. The n. of doctors per n. of beds variable (lagged at time $t-1$) shows a positive and significant effect on readmission, at the beginning and at the end of the sample period.

7 Concluding remarks

In this paper we have explored the effect of social interaction on individuals' hospital choice for patients with cardiac illness. To our knowledge, this is the first attempt at identifying, testing and modelling social interaction in patients' decisions on the hospital where to be treated. Our findings support the existence of strong correlation between an individual's hospital choice and those of her neighbours. The strategy that we have proposed in this paper allows us to conclude that part of this correlation is due to social interaction among patients. Therefore, individuals rely also on information gathered from neighbours when choosing their health provider. Our empirical findings also show that this network effect is less important when it concerns large hospitals. We have then investigated how such network effect is related to the quality of hospitals. When adopting mortality as a proxy of hospital quality, the use of neighbourhood information does not seem to have any impact on the likelihood of choosing a high quality of hospitals. However, it is interesting to observe that when we adopt 30-day readmission as quality indicator, our results show that the higher the strength of interaction among individuals for a particular hospital, the lower is likely to be its quality.

One important implication for these results is that policy makers should put efforts in implementing central and local mechanisms of information diffusion such as guidelines or star rating indicators, that reduce geographical inequalities in the access to information. To be effective, the strategy of increasing the supply in the region of Lombardy to boost competition, as attempted with the last regional reform, should be accompanied by central and local mechanisms of information diffusion that diminish territorial inequality in the access to information.

References

Aizer, A. and J. Currie (2004). Networks or neighborhoods? correlations in the use of publicly-funded maternity care in california. *Journal of Public Economics 88*, 2573–2585.

Amigoni, M., M. Marchetti, L. Merlino, and A. Zangrandi (1998). Offerta di servizi sanitari, regole di finanziamento e sistema competitivo in regione lombardia. *Mecosan 27*.

Berta, P., G. Callea, G. Martini, and G. Vittadini (2009). The effects of upcoding, cream skimming and readmissions on the efficiency of italian hospitals: a population-based investigation. *Economic Modelling*, Forthcoming.

Birke, D. (2009). The economics of networks: A survey of the empirical literature. *Journal of Economic Surveys Forthcoming*.

Brock, W. and S. Durlauf (2001). Interactions-based models. In H. J. and L. E. (Eds.), *Handbook of Econometrics, Volume V*. North Holland, Amsterdam.

Brock, W. A. and S. N. Durlauf (2003). Multinomial choice with social interactions. NBER Technical Working Paper No. 288.

Chaix, B., M. Lindström, M. Rosvall, and J. Merlo (2008). Neighbourhood social interactions and risk of acute myocardial infarction. *Journal of Epidemiology and Community Health 62*, 62–68.

Conley, T. G. and G. Topa (2002). Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Econometrics 17*, 303–327.

Cornford, C. S. and H. M. Cornford (1999). I'm only here because of my family.' a study of lay referral networks. *British Journal of General Practice 49*, 617–620.

Deri, C. (2005). Social networks and health service utilization. *Journal of Health Economics 24*, 1076Ű1107.

Epperson, B. K. (2003). Covariances among join-count spatial autocorrelation measures. *Theoretical Population Biology 64*, 81–87.

Freidson, E. (1960). Client control and medical practice. *American Journal of Sociology 35*, 374.

Glaeser, E. L., B. I. Sacerdote, and J. A. Scheinkman (1996). Crime and social interactions. *Quarterly Journal of Economics 111*, 507–548.

Goldman, D. and J. A. Romley (2008). Hospitals as hotels: The role of patient amenities in hospital demand. NBER Working Paper No. 14619.

Greene, W. (2008). *Econometric Analysis* (Sixth ed.). Prentice Hall.

Ho, V. and B. H. Hamilton (2000). Hospital mergers and acquisitions: Does market consolidation harm patients? *Journal of Health Economics 19*, 767–791.

Jones, A. (1994). Health, addition, social interaction and the decision to quit smoking. *Journal of Health Economics 13*, 93–110.

Kessler, D. and M. McClellan (2000). Is hospital competition socially wasteful? *Quarterly Journal of Economics 115*, 577–615.

Luft, H. and et al. (1990). Does quality influence choice of hospital? *JAMA 263*, 2899–2906.

Manski, C. (1993). Identification of endogenous social effects: the reaction problem. *Review of Economic Studies 60*, 531–542.

Pope, D. G. (2009). Reacting to rankings: Evidence from ŞamericaŠs best hospitalsŤ. *Journal of Health Economics*, Forthcoming.

Revelli, F. (2006). Performance rating and yardistick competition in social service provision. *Journal of Public Economics 90*, 459–475.

Romano, P. S. and R. Mutter (2004). The evolving science of quality measurement for hospitals: Implications for studies of competition and consolidation. *International Journal of Health Care Finance and Economics 4*, 131–157.

Romano, P. S. and H. Zhou (2004). Do well-publicized risk-adjusted outcomes reports affect hospital volume? *Medical Care 42*, 367Ű377.

Schneider, E. C. and A. M. Epstein (1996). Influence of cardiac-surgery performance reports on referral practices and access to care. a survey of cardiovascular specialists. *New England Journal of Medicine 335*, 251–256.

Schoenberg, N. E., C. H. Amey, and et al. (2003). Lay referral patterns involved in cardiac treatment decision making among middle-aged and older adults. *The Gerontologist 43*, 493–502.

Skinner, J. S. and D. Staiger (2005). Technology adoption from hybrid corn to beta blockers. National Bureau of Economic Research, Working Paper 11251.

Tay, A. (2003). Assessing competition in hospital care markets: The importance of accounting for quality differentiation. *The RAND Journal of Economics 34*, 786–814.

Volpp, K. G., S. V. William, J. Waldfogel, J. H. Silber, J. S. Schwartz, and M. V. Pauly (2003). Market reform in new jersey and the effect on mortality from acute myocardial infarction. *Health Services Research 38*, 515Ű533.

Zangrandi, A. (1998). I sistemi di contratti in sanità per la regolazione della produzione: quali reali aspettative? *Mecosan 26*, 45–55.

Table 1: Definition of variables

Variable	Description
Patient characteristics:	
Distance_{ih}	Distance of patient i to hospital h / distance of i to her nearest hospital
Old_i	1 if patient i is over 75 years of age
Males_i	1 if patient i is male
ICD-9-CM=410_i	1 if patient i suffers from AMI
ICD-9-CM=411_i	1 if patient i suffers from other acute and subacute forms of ischemic heart disease
ICD-9-CM=412_i	1 if patient i suffers from old myocardial infarction
ICD-9-CM=413_i	1 if patient i suffers from angina pectoris
ICD-9-CM=414_i	1 if patient i suffers from other forms of chronic ischemic heart disease
Elective_i	1 if patient i's admission is booked or planned
Price_i	Total expenditure for patient i
Hospital characteristics:	
N. beds_h	Total number of beds (ordinary + day hospital) in hospital h
$\text{N. doct./n. of beds}_h$	Number of doctors in hospital h / n. of beds in hospital h
Teaching_h	1 if hospital h is teaching (i.e., it provides clinical education and training to doctors/nurses etc.)
Mono-specialist_h	1 if hospital h is mono-specialist
Private_h	1 if hospital h is private
Technology_h	1 if hospital h has a catheterization laboratory
Large_h	1 if hospital h has more than 299 beds

Table 2: Number of observations and percentage of emergency cases by ICD9CM category

	Total	ICD-9-CM=410		ICD-9-CM=411		ICD-9-CM=412		ICD-9-CM=413		ICD-9-CM=414	
	N.	N.	% emerg.	N.	% emerg.	N.	% emerg.	N.	% emerg.	N.	% emerg.
2004	57,351	19,439	79.97	10,603	59.48	1,365	15.90	12,501	23.21	12,501	18.60
2005	58,291	21,067	80.91	10,864	58.71	1,095	17.99	11,273	23.16	13,992	19.63
2006	56,730	13,992	82.14	10,472	61.01	931	17.99	10,853	24.54	13,484	18.93
2007	58,230	21,232	83.60	10,346	63.05	852	21.24	10,688	23.41	15,112	18.79

Notes: See Table 1 for definition of variables.

Table 3: Descriptive statistics of patients

	2004	2005	2006	2007
% Males	68.98	69.07	69.86	69.42
Age (av. years)	67.83	68.28	68.43	68.43
% 65-74 years	29.51	29.44	29.28	29.15
% 75+ years	28.62	30.20	31.12	31.93
Length of stay (av. n. of days)	9.14	9.01	8.78	8.56
Expenditure per patient (av., in Euro)*	5,117.4	5,208.5	5,288.5	5,400.4
30-day readmission	9.54	9.00	8.71	8.13
30-day mortality	4.72	5.08	5.01	5.05

(*): The aggregate has been deflated using the consumer price index (2005=100)

Table 4: Lombardy hospital characteristics

	2004	2005	2006	2007
N. hospitals	128	132	129	127
Catheterization lab.	51	60	63	69
Teaching (n.)	10	10	10	10
Mono-specialist (n.)	10	10	8	6
Public (n.)	85	86	83	80
Patients (average n.)	448.7	441.5	439.7	458.5
Medium+ (%)	49.22	47.73	48.06	51.97
Large++ (%)	29.69	30.30	30.23	31.50
Beds (average n.)	259.8	262.53	267.4	275.7
ordinary	233.0	235.3	239.3	245.5
day-hospital	26.83	27.22	28.06	30.23
Doctors per n. beds	0.53	0.54	0.55	0.55

(+): Medium hospitals are those with a number of beds between 100 and 299.
(++): Large hospitals are those with more than 299 beds.

Table 5: Migration flows of patients and their concentration across territory

	2004	2005	2006	2007
Migration characteristics				
% Movers$^{(+)}$	28.12	27.97	28.37	30.07
% Movers from outside Lombardy	9.03	8.67	8.73	8.83
Av. distance patient-to-hospital (Km)	56.99	54.83	54.12	53.70
Av. distance patient-to-hospital (Lombardy only)	12.37	12.26	12.37	12.61
Median distance patient-to-hospital (Km)	6.86	6.79	6.85	7.16
N. pat. in the neighbourhood with same choice	50.23	46.06	44.22	41.54
% pat. in the neighbourhood with same choice	38.40	37.59	37.96	37.04
$n_{H,t}$	2.16*	2.03*	2.18*	2.13*

$^{(+)}$: By mover we mean a patient admitted in an hospital that is located in a province different from the province of residence of the patient. (*): significant at the 5% level.

Table 6: The determinants of patient's hospital choice (years 2005-2007)

	2005		2006		2007	
	coeff	Std.err.	coeff	Std.err.	coeff	Std.err.
Elective_i	-2.539*	0.032	-2.329*	0.033	-2.298*	0.033
$\bar{y}_{ih,t-1}$	6.994*	0.095	7.313*	0.096	6.499*	0.092
$\bar{y}_{ih,t-1}$*Elective_i	2.380*	0.081	2.070*	0.086	2.002*	0.084
Age_i	0.000	0.001	-0.001	0.001	-0.001	0.001
Male_i	-0.022	0.017	-0.013	0.017	-0.010	0.017
Distance_{ih}	-0.173*	0.006	-0.148*	0.006	-0.146*	0.005
ICD-9-CM=411_i	0.039	0.024	0.045	0.024	0.104*	0.024
ICD-9-CM=412_i	0.066	0.066	0.131*	0.069	0.111	0.069
ICD-9-CM=413_i	0.112*	0.025	0.118*	0.026	0.114*	0.026
ICD-9-CM=414_i	0.105*	0.025	0.108*	0.026	0.120*	0.026
$\bar{y}_{ih,t-1}$*age_i	0.088	0.080	0.113	0.082	0.134	0.079
$\bar{y}_{ih,t-1}$*distance_{ih}	-0.023*	0.004	-0.028*	0.004	-0.010*	0.004
Catheterization lab._h	-0.172*	0.026	-0.142*	0.030	-0.117*	0.034
N. beds_h/100	0.105*	0.002	0.102*	0.003	0.090*	0.003
$\bar{y}_{ih,t-1}$*Large hospitals_h^+	-0.866*	0.088	-0.666*	0.087	-0.171*	0.082
N. doctors per n. of beds_h	0.968*	0.045	0.735*	0.051	1.039*	0.059
Teaching_h	0.088*	0.021	-0.069*	0.023	-0.090*	0.023
Private_h	0.042*	0.017	-0.011	0.017	0.014	0.016
Mono-specialist_h	1.086*	0.040	1.495*	0.050	1.571*	0.048
Intercept	-4.256*	0.247	-4.358*	0.249	-4.591*	0.256
N	39,820		39,820		39,914	
Log-likelihood	-92,413.7		-87,496.483		-91,874.56	

Notes: ASL and province dummies have been included. ($^+$): hospitals with more than 299 beds. (*): significant at the 5% level. See Table 1 for definition of variables.

Table 7: The impact of social influences on quality indicators (years 2005-2007)

Dep. var.	30-day readmission						30-day mortality					
	Year 2005		Year 2006		Year 2007		Year 2005		Year 2006		Year 2007	
	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.
Elective$_i$	-0.061*	0.056	-0.096*	0.059	-0.120*	0.062	-0.666*	0.081	-0.743*	0.086	-0.802*	0.088
$\hat{\delta}_h$	0.032*	0.016	0.030*	0.011	0.081*	0.006	0.004	0.020	0.016	0.017	-0.006	0.009
Age$_i$	-0.007*	0.002	-0.006*	0.002	-0.004*	0.002	0.089*	0.003	0.082*	0.003	0.093*	0.003
Males$_i$	0.254*	0.045	0.287*	0.046	0.228*	0.047	0.049	0.051	0.089	0.051	-0.002	0.050
ICD-9-CM=411$_i$	-0.311*	0.054	-0.303*	0.055	-0.302*	0.056	-1.477*	0.084	-1.653*	0.091	-1.472*	0.085
ICD-9-CM=412$_i$	-1.425*	0.224	-1.431*	0.247	-1.200*	0.231	-1.931*	0.413	-0.973*	0.289	-1.435*	0.361
ICD-9-CM=413$_i$	-0.606*	0.067	-0.585*	0.069	-0.659*	0.074	-2.531*	0.177	-2.280*	0.160	-2.481*	0.185
ICD-9-CM=414$_i$	-1.236*	0.080	-1.032*	0.079	-0.932*	0.080	-1.268	0.090	-1.277	0.097	-1.040	0.091
N. Beds$_{h,t-1}$/100	0.058*	0.007	0.060*	0.008	0.083*	0.007	-0.028*	0.010	-0.005	0.010	-0.020*	0.010
N. doct./n. of beds$_{h,t-1}$	0.231	0.166	-0.062	0.185	0.685*	0.167	-0.288	0.191	-0.493*	0.220	0.135	0.204
Price$_i$/1,000	-0.087*	0.008	-0.100*	0.007	-0.094*	0.008	0.043*	0.005	0.043*	0.005	0.037*	0.005
Teaching$_h$	-0.278*	0.053	-0.440*	0.058	-0.358*	0.055	0.107	0.078	-0.012	0.078	0.088	0.079
Mono-specialist$_h$	-0.123	0.151	-0.417*	0.189	0.137	0.178	-0.723*	0.194	-0.746*	0.225	-0.561*	0.205
Private$_h$	-0.431*	0.070	-0.298*	0.059	-0.475*	0.056	-0.075	0.088	-0.175*	0.081	-0.129	0.073
Intercept	-1.610*	0.177	-1.514*	0.177	-2.400*	0.180	-8.654*	0.270	-8.189*	0.277	-9.221*	0.286
N. obs.	45,976		46,234		45,658		45,658		45,976		46,234	
Log-likelihood	-13,225.2		-12,818.8		-12,244.3		-6,806.1		-6,851.6		-6,801.85	

Notes: ASL dummies have been included. (*): significant at the 5% level. See Table 1 for definition of variables.