12-2018

Better inpatient health quality at lower cost: Should I participate in the online healthcare community first?

Kai LUO

Qiu-Hong WANG
Singapore Management University, qiuhongwang@smu.edu.sg

Hock Hai TEO

Xi CHEN

Better Inpatient Health Quality at Lower Cost: Should I Participate in the Online Healthcare Community First?

Completed Research Paper

LUO Kai
National University of Singapore
15 Computing Drive, Singapore 117418
luokai@comp.nus.edu.sg

WANG Qiu-Hong
Singapore Management University
80 Stamford Road, Singapore 178902
qiuhongwang@smu.edu.sg

TEO Hock Hai
National University of Singapore
15 Computing Drive, Singapore 117418
teohh@comp.nus.edu.sg

CHEN Xi
Zhejiang University
Hangzhou, Zhejiang, China 310058
chen_xi@zju.edu.cn

Abstract

As policy makers across the globe look to health information technology (HIT) as a means of improving the efficiency of the healthcare systems, it has sparked significant interest in understanding how HIT might help achieve that. While researchers have examined and documented the efficiency-improving effect of various institution HITs (e.g., electronic clinic pathways and telemedicine), the impacts of consumer HITs such as online healthcare communities have been generally overlooked. Utilizing two unique datasets from both an online healthcare community and a general hospital, we study the impact of online healthcare community on offline inpatient care efficiency. Through rigorous analysis, we find that communications between physicians and patients on the online healthcare community leads to decreased inpatient cost and improved care quality. Our findings also underscore the importance of online physician-generated information in this impact. Implications for healthcare practitioners and policy makers are discussed.

Keywords: Healthcare, Cost, Quality, Online Healthcare Community

Introduction

Soaring health expenditures worldwide and the fast-growing demand for health service have put an increasing focus on cost containment and efficiency of health service delivery (Kittelsen et al. 2015). Undeniably, quality improvement and cost containment are the dual central concerns for any development of current healthcare systems (Hussey et al. 2013; Orszag 2016). They are also important mandates of the Affordable Care Act (ACA) in the US (Obama 2016; Orszag 2016). In this regard, the continuing development of healthcare information technologies (HIT) represents one of the major contemporary efforts in achieving such objectives (Agarwal et al. 2010).

To the extent that various HITs help improve healthcare efficiency, researchers have mainly focused on HITs implemented within the healthcare institutions, such as the automation of clinical processes, digitization of health records, and health information exchanges (Lee et al. 2016; Lion et al. 2016; Yaraghi et al. 2014). While these institutional HITs are doubtlessly important and necessary, there also exist consumer HITs (e.g., online healthcare communities) that are not developed or controlled by traditional healthcare institutions, thus raising the question of whether they can contribute to healthcare efficiency.

In this study, we investigate the impact of online healthcare communities on offline healthcare efficiency. In previous literature, online healthcare communities are generally defined as online communities that facilitate the creation, sharing, and dissemination of health information by patients or physicians (Goh et al. 2016). While consumer inferences about a product/service based on available information has been identified as a crucial factor in improving market efficiency (Chernev and Carpenter 2001), this important efficiency-improving effect has so far received limited scholarly attention in the healthcare context. Compared to offline context, the significantly more abundant supply of healthcare information on the online healthcare community (Agarwal et al. 2010), therefore, has the potential to influence both the healthcare quality and cost. Although previous studies have documented such offline efficiency-improving effect of other Internet-mediated health technologies such as telemedicine and teleconsultation (Albert et al. 2011; Hickson et al. 2015), they are still considered institutional technologies and their effectiveness is derived mainly from the direct substitution of non-acute offline healthcare service (e.g., general practitioner visit for minor illness) (Hickson et al. 2015; Watson et al. 2016). Considering that online healthcare communities are increasingly being used by patients as an important supplementary information source to offline acute healthcare visit (Agarwal et al. 2010), it is important to further explore questions related to its impact on non-substitutable offline care efficiency.

We pose the following question in this study: *How do patients' participation on the online healthcare communities influence the subsequent offline inpatient care quality and cost?* Recent studies have suggested that the locus of power in healthcare is shifting towards patients especially with the advent of online healthcare communities (Visser et al. 2016). Hence, patients' participation on the online healthcare communities plausibly provides an alternative way to improve offline healthcare efficiency through patient empowerment. This assertion, however, remains understudied. Anchoring this study in the theory of physician agency (McGuire 2000) and patient welfare (Arrow 1963), we suggest that patients' utilization of physician-generated information on the online healthcare communities can be instrumental in influencing physician behaviors and improving patient welfare in an offline inpatient visit.

To empirically answer our research question, we utilize two unique datasets: the inpatient data from one major general hospital in Shanghai, China, and the online data from a leading online healthcare community in China. The online healthcare community in China has been widely used by patients to consult nationwide physicians as well as making offline reservations. By linking the online healthcare community data and the offline inpatient data at the physician level, we first develop a valid proxy (the weekly number of physicians' online patients) for offline patients' participation on the online healthcare community. Then, we empirically test the impact of patients' online participation on the inpatient length of stay (quality of care measure) and the medical billings (inpatient cost measure). We find that higher patients' online participation will lead to shortened mean inpatient length of stay and lowered mean medication billing. However, such effects are significant only if the online patients are from the same geographical/administrative location where offline healthcare service is provided. To explore the mechanism underlying such impacts, we also conduct additional content analysis on the physician's online consultation texts. We find that the quality of online physician-generated information is positively related to offline inpatient quality of care. Our findings highlight the potentially important role of online healthcare communities as an effective information sharing and communication channel in alleviating the inherent information asymmetry problem and improving the efficiency of the overall healthcare systems.

This research contributes to the extant literature in several ways. First, to the best of our knowledge, we are among the first to empirically study the impact of online healthcare communities on the offline care quality and cost in the inpatient setting. Prior literature has identified the substitution effect of the online healthcare on non-acute offline healthcare service. Our analysis reveals that online healthcare could also positively influence offline care efficiency when it works as the supplement to the offline inpatient service. Second, our findings also help resolve a major concern raised in the health economics literature that the patients are generally incapable of fighting the severe information asymmetry problem. Our analysis reveals that patients may acquire the capability in influencing their inpatient care outcome through the participation of the online healthcare communities. Third, our study extends the theory of physician agency by further elucidating the value of online physician-generated information to the patients in combating the problem that physician often acts as an imperfect agent to them due to the existence of asymmetric information. Finally, by revealing several boundary conditions (e.g., geography, patient incentive, physician's ability to influence price) of the online to offline inpatient healthcare effect, our study adds to a more nuanced theoretical understanding of how online healthcare service could potentially influence the offline healthcare outcomes. The practical implications of this study will also be discussed.

Literature Review

The Economics of Healthcare Cost Containment and Quality Improvement

Although past empirical studies have identified multiple external factors that influence healthcare cost containment and quality improvement, conceptual works in the healthcare field have long suggested that lying at the heart of the healthcare cost/quality issues are the fundamental questions about physician motives and market power, or best known as the "physician agency" (McGuire 2000). Due to the lack of medical expertise, patients often depend on their physicians acting as their "agents" to decide the appropriate levels of care they should receive (McGuire 2000). The physician agency problem arises because of non-contractible physician effort and patients' general inability to assess the quality of healthcare services both *ex ante* and *ex post* (McGuire 2000). In essence, under such condition, a selfish physician have the incentive to induce demand for healthcare service for his/her own interest, leading to decreased healthcare efficiency and patient welfare (Phelps 1986). As such, past efforts that tried to deal with this agency problem have predominantly focused on institutional approaches that limit physician's market power or manage physician's incentives through, for example, imposing resource restriction (Taheri et al. 2000), and optimizing the contracting and payment model (Choné and Ma 2011). There are also growing literature that try to address this problem by adopting the patients' perspective. Among them, De Jaegher (2012) developed a game-theoretical model where he tied the quality of patients' private information about the medical treatment directly to the health outcome and patient welfare. Tunis and Pearson (2006) showed that when provided with independent and unbiased evidence-based treatment-related information, patients do not always seek the newest and most expensive mode of treatment. Of particular relevance to this research, we are interested in further understanding the value of online healthcare information to patients in face of physician agency.

The Value of Online Healthcare Information

By facilitating the creation, sharing, and dissemination of health information among patients and physicians, online healthcare communities have the potential to resolve the physician agency problem caused by the severe information asymmetry between physicians and patients. The extant literature concerned with the Internet and healthcare have already examined the multi-faceted value of online healthcare information. For example, Bhandari et al. (2014) found that online health information help reduce the overall healthcare cost by decreasing patients' health seeking cost. Researchers have also demonstrated that communication on the online healthcare communities could help reduce the rural-urban health disparity (Goh et al. 2016) and improve patients' mental health outcomes (Yan and Tan 2014). At a more general level, several studies have attributed the ability of online healthcare information in influencing health related outcomes to the process of patient empowerment (Johnston et al. 2013), which reflects the ability of patients to proactively and positively influence their health outcomes and health behaviors. However, in most of these studies, researchers have mainly investigated the value of online healthcare information from the acquisition of medical knowledge and social support perspectives. As Johnston et al. (2013) suggested, online healthcare provides valuable informal information such as personalized health experiences, informal treatments, and personal success stories that serve to both spread medical knowledge as well as build confidence in health behaviors. Clearly amiss here is the element of online information about the physicians, which is more relevant to the physician agency problem. Although a handful of studies have examined impacts of online information about the physicians on the online healthcare communities, they have mainly focused on issues such as physician's motivation to participate (Guo et al. 2017), and physician online ratings (Gao et al. 2012). Our knowledge of how the online physician information empowers patients to mitigate the physician agency problems remains limited to-date.

Hypotheses Development

Based on the literature, this research proposes several hypotheses to explore 1) the impact of patients' participation on the online healthcare communities on offline inpatient care quality and cost, and 2) the value of physician-related information in this online to offline effect.

As the online healthcare communities are not formal medical institutions, legal physician-patient contracts are not explicitly established there. Institutional approaches such as hospital resource restriction or government regulations are not applicable there. Instead, the online healthcare communities serve as an information medium that facilitates patients' health seeking behaviors (Goh et al. 2016). Patients' utilization of the online information of physicians would likely exert influence over their subsequent healthcare visit. As such, we now describe the critical theoretical difference between direct offline healthcare visits and online then offline healthcare visits.

Although healthcare service as a typical credence good is well-acknowledged (Dulleck et al. 2011), healthcare economists have also stressed the important difference between the demand for physician services and the demand for a physician. In the offline healthcare context, McGuire (2000) suggested that, as opposed to demand for physician service, a patient has a demand for the services of a particular physician that fit his/her preferences. Hence, the physicians are imperfect substitutes in the eyes of patients. In a perfect market, efficiency would demand matching patients with the right physician (McGuire 2000). However, due to the limited and imperfect information about physicians in the offline healthcare market, patients would likely to develop allegiance to a known physician even if there are sufficient competitions in the market (Wong 1996). This inefficient match between physician and patient in the offline context closely mimics the *cheap-talk signaling* games with a single information sender (Farrell and Rabin 1996) where the optimal move for a consumer (patient) is to delegate his/her decision right to the sole expert (physician).

On the online healthcare communities, however, patients are provided with access to rich information of a large pool of reputable physicians. In our focal online healthcare community, patients are allowed to not only consult with different online physicians but also observe all the past online consultation information (text) between each physician and other online patients. We suggest that this online context resembles the *multi-sender communication* games (Battaglini 2002), which evolves from the *cheap-talk signaling* games but with multiple information senders. Theoretical works have shown that, with multiple experts imperfectly observing the true state of the information sender, the information receiver will be able to learn the true state and make the optimal decision for him/herself (Ambrus and Lu 2014).

The common consequences of the imperfect match of between physician and patient in the offline context, other than worsen health outcome, also include efficiency-decreasing activities such as over-treatment or over-prescription (De Jaegher 2012). Hence, from the patient's perspective, price/cost-conscious patients have the incentive to seek higher care quality at a lower cost. With the ample supply of information about multiple physicians on the online healthcare community, they could use these useful knowledges about different physicians to find a better match for their subsequent offline visits. On the other hand, from the physicians' perspective, better matched patient-physician pairs often mean that patients are more capable of constraining the physicians' behavior (De Jaegher and Jegers 2001). Therefore, the offline inpatient care efficiency is expected to improve as the physicians are facing more demand from online patients. As with previous studies, we measure offline inpatient care efficiency using the inpatient care quality and cost. Hence, we hypothesize as follows:

H1*: Patients' participation on the online healthcare community will lead to improved offline inpatient care quality.*

H2*: Patients' participation on the online healthcare community will lead to decreased offline inpatient care cost.*

As discussed above, a necessary condition of our H1 and H2 is that patients have the incentive to seek most efficient offline healthcare option. That is, if certain online patients are not sensitive to healthcare price/cost, then their participation on the online healthcare community would less likely to result in the efficiency-improving effect for their offline healthcare visit. Hence, differentiating patients' incentive helps us to further elucidate the mechanism behind the efficiency-improving effect of online healthcare community participation. Inferred by the healthcare tourism literature (Pocock and Phua 2011), we identify patients' geographical origin as a key differentiator for patients' incentive.

Healthcare tourism is generally defined as the organized travel outside one's local environment for the maintenance, enhancement or restoration of the individual's health (Pocock and Phua 2011). In our research context, since Shanghai City is easily one of the most developed and costliest cities in China, non-local patients who travel to Shanghai to seek healthcare service would likely to be less sensitive to price/cost comparing to local Shanghai residences. Those non-local patients may interpret "efficiency" differently than local patients. They highly value the healthcare outcomes but are generally insensitive/less sensitive to healthcare price. Thus, they will likely seek expensive care options even if it is inefficient to the healthcare system in general. Hence, we expect to see no efficiency-improving effect if the online patients are from outside Shanghai City. Therefore, we hypothesize as follows:

H3: *The offline quality improvement and cost reduction effect of patients' participation on the online healthcare community will only be observable among local patients.*

We also explore how physician-related information on the online healthcare community empowers patients to find the better-matched physician. McGuire (2000) has identified two general types of information that are distributed asymmetrically between patients and physicians: physician actions and physician characteristics. The problem about physician actions is that they cannot be verified by the patients to see if the physicians are diligent or not (Arrow 1986). We argue that this problem also exists in the online context, if not further aggravated, especially due to the lack of social presence in the online environment in general. Physician characteristics, on the other hand, represent unalterable characteristics of a physician such as empathy and medical acumen (McGuire 2000). As Gaynor (1994) noted that physician is an "experience good", an offline patient literally has to try a physician, and then make a valid inference about the characteristics of the physician. Gaynor (1994) further suggested that the learning of physician characteristics in the offline context is imperfect and slow, and the primary information source is often through personal contact or friends and relatives, which tends to distort or dilute the quality of such information. However, in an ideal context where public institutions such as the government provide patients with authentic information about physician characteristics prior to selecting a physician, empirical research has shown improved quality of care (Reid et al. 2010) and patient satisfaction (Kalda et al. 2003). On the online healthcare communities, there exist ample physician-generated information about how a physician diagnoses or treats a medical problem as well as how a physician interacts with a patient. Such information would allow patients to extract more useful inference about the physician characteristics compared to the offline information sources. However, the quality of the physician-generated information will influence patients' ability to extract such inference, as information quality is the key determinant of the quality of decisions and actions. According to the information quality taxonomy developed by Wang and Strong (1996), there are two main objective dimensions of information quality: information completeness and information amount. Hence, we hypothesize as follows:

H4a: *Higher information completeness of online physician-generated information will lead to higher offline inpatient quality improvement and cost reduction.*

H4b: *Higher information amount of online physician-generated information will lead to higher offline inpatient quality improvement and cost reduction.*

Research Methodology

Research Context and Data

The primary context of this research is a 3A-ranking (the highest ranking in China's nationwide hospital ranking system) general teaching hospital (referred to as Hospital A from now on) located in Shanghai City[1]. We obtained the inpatient discharge data of this hospital during a period from March 2003 to October 2013. To complement the offline hospital data, we further collected all the online consultation records of physicians from this hospital on a leading online healthcare community in China. This online community provides an online catalog of nationwide hospitals and physicians that is accessible to nationwide patients. Interested physicians could further open a personal website on this community to answer consultation questions from online patients. The text consultation on the community before 2015 is free-of-charge to all patients. More importantly, the online patients could post question on any physician's community and the physicians have their freedom to answer the posted question. Hence, the online community becomes an ideal place where online patients could sample and observe a physician, and online physicians could build and spread their reputations. Since the launch of this community, more than 18 million online questions

---

[1] Due to privacy agreement, we do not provide any other information that can help identify the name of the involved hospital.

have been answered by physicians[2]. The physicians of the Hospital A first entered the community in the first quarter of 2008. Hence the feasible time period for this research is between the time of first entry and October 2013 when we have both the online and offline data are available.

The inpatient discharge data from Hospital A provides detailed individual patient level information including admission/discharge date, patient age, patient gender, admission department, attending physician, as well as detailed medical billings (before insurance deduction). The patients' identity is strictly anonymized in our data. However, we have physician information such as name, gender, job title, and specialty. In total, we have around 1,000 unique physicians in our data with more than 400,000 inpatient discharges during our time period.

We developed a web crawler to collect the corresponding online consultation data from the website of this community. The collected data include the detailed text communications between each online patient and each physician. The timestamp of each conversation is also recorded. The demographic and identity information about the online patients is also completely anonymized on the online communities. However, we are able to collect the additional location data of the online patients by crawling the WAP version of this community. We find that 116 physicians from Hospital A have online consultation records during our time period with around 60,000 unique online patient-physician encounters[3].

To provide a clearer picture of the online consultation process as well as to inform what types of physician-generated consultation information exist in the online healthcare context, we now describe the online healthcare consultation process. According to Serrano and Karahanna (2016), the physician consultation process include three main stages: Information Gathering, Analysis and Diagnosis, and Explanation and Planning stages. Adapting to our context, online patient initiates the consultation by posting a medical question on the website of any physician. If the physician chose to answer the question (as almost all of them did), three general types of task-relevant information could be observed. The information gathering stage begins if the physician seeks additional medical-related information from the patients. Then, the physician will analyze the information exchanged and give his/her analysis of the problem in the diagnosis and analysis stage. Finally, in the explanation and planning stage, upon analysis of the problem, the physician would further explain recommend the course of actions (e.g., type of treatment, place of treatment, etc.). Due to the unsynchronized nature of text consultation on this community, however, not all three types of the online physician-generated information concur in every online patient-physician encounter. In addition to the three general types of physician-generated information above, there also exist other types of information such as physician's emotional support, greeting, courtesy message to the patients. Although they are not directly related to the focal healthcare consultation task, they are still important information a patient seeks when communicating with a physician (Ong et al. 1995). Hence, we conclude that there exist four general categories of online physician-generated information: 1) information gathering, 2) problem analysis, 3) treatment explanation, and 4) emotional support.

Empirical Strategy

To empirically test our proposed hypotheses, we first need to develop a valid proxy for the offline patients' participation on the online community. Due to the anonymity of both online and offline patients, we are unable to identify among the offline patients who has participated on the online community prior to offline admission. To overcome this empirical challenge, we link our online and offline data by physician and aggregate at the physician-weekly level. We define offline patients' participation on the online community as the number of patients who had online consultation first in the total offline patients admitted in a specific week to a specific physician. We then use the lagged (one week) number of online patients who consulted one physician as the proxy for the offline patients' participation on the online healthcare community.

To conceptually demonstrate the validity of the proxy, the key is to show that there exist overlaps between a physician's online patients pool and offline patients pool. Figure 1 summarizes the types of online and

---

[2] Indicated by the total number of consultations on http://zixun.haodf.com (in Chinese), accessed on April 14, 2018.

[3] We define one patient-physician encounter as the whole online consultation exchanges between a unique patient and a unique physician on this online healthcare community

offline patients for a physician if he/she joined the online healthcare communities. It shows that joining the online healthcare communities adds additional sources to a physician's offline patient pool. There exist common types (i.e., A(X) and B(X)) in both the online and offline pool of patients for a physician. For online patients who consulted other online physicians rather than physician X, their chance of converting into physician X's offline patients may also increase with the number of physician X's online patients. Hence, not only should the lagged number of online patients who consulted a physician positively correlate with the number of offline patients who were admitted by this physician, but also that it should further proportionally reflect the number of this physician's offline patients who participated on the online healthcare community prior to the offline admission. To empirically validate our proxy, we use two different approaches for better triangulation: 1) we estimated a fixed effect-model to test the effect of the lagged-week number of online patients on the weekly number of a physician's offline normal and surgical admissions; 2) we use relevant online anecdotal information as additional qualitative evidence for the existence of the online-to-offline correlation.

	Types of Physician X's Online Patients	**Types of Physician X's Offline Patients**
Common Types	• Patients who have already decided to visit physician X offline but chose to first consult her on the online healthcare community (***A(X)***) • Patients who decided to visit physician X offline after consulting her first on the online healthcare community (***B(X)***)	
Distinct Types	• Patients who consulted physician X on the online healthcare community but decided to subsequently visit other physicians offline (***C(X)***) • Patients who consulted physician X on the online healthcare community but did not plan any subsequent offline visit (***D(X)***)	• Patients who decided to visit physician X offline but after consulting other online physicians first on the online healthcare community (***C′ (X)***) • Patients who directly visited physician X offline without any online participation (***E(X)***)

Figure 1. Types of Online and Offline Patients

The second empirical challenge is the potential existence of a physician selection issue. It is highly possible that physicians who joined the online community differ systematically from those who did not join in terms of both observable and unobservable factors. Such systematic differences may be correlated with care quality and cost outcomes, leading to biased estimation results. To overcome this potential selection issue, we use propensity score matching (PSM) technique to create a matched sample of physicians who did not join the online community to those who joined. The matching variables include factors that could influence the care outcomes, such as physician department, medical title, gender, total inpatient admissions, share of admitted patients that belong to the 0-1, 2-17, 18-34, 35-64, and 65+ age groups, share of admitted female patients. We used the one-to-three with replacement matching specification which yielded us with 201 matched offline physicians who did not join the online community during 2008 to 2013. For simplicity, to demonstrate that the common support required for PSM is met (Lechner 2002), we plot the propensity score distributions for the online physicians and not-online physicians both before and after the matching in Figure 2. The almost identical distribution of the propensity scores after matching indicates a good matched sample. In our final matched sample, online and not-online physicians are evenly distributed across different departments. For the potential patient-side selection issue, we demonstrate in the analysis part that it should not bias our result.

Finally, to identify the effect of patients' participation on the online healthcare community, we first control the physician individual-level fixed effect, week fixed effect, and department-level fixed effect. This allows us to eliminate the influence of unobserved but potentially endogenous individual, time and institutional effect that influence the offline inpatient care quality and cost. After controlling for the fixed effects, the natural variations in the lagged number of online patients of different physicians in different weeks help us isolate the true effect of patients' participation on the online healthcare community. More specifically, there are two sources of exogenous variations: 1) since all online consultation cases are initiated by the patients rather than the physicians, for the same physician, the variation in the number of online patients across different weeks[4] should be exogenous to the inpatient care outcomes; and 2) for the same week, the variation in the number of online patients across different physicians[5] after matching, should also be exogenous to the inpatient care outcomes.

Kernel Density Plot

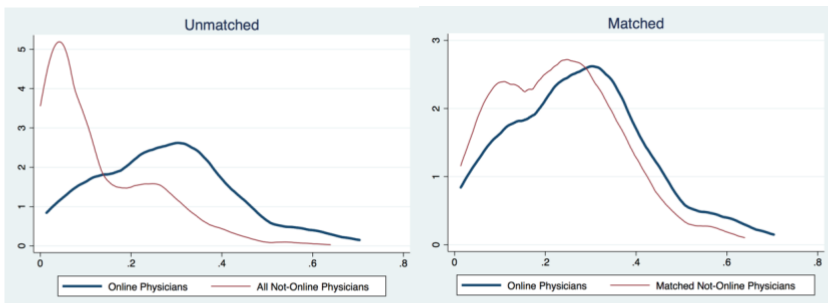

Figure 2. Distribution of Propensity Score Before and After Matching

Variable Operationalizations

To measure the offline inpatient care quality, we use the *Inpatient_Stay*$_{it}$, which measures the average length of stay of patients admitted to physician i in week t. Length of stay is a commonly used measure for inpatient quality of care in many medical studies. To measure the offline inpatient care cost, for simplicity, we use only two set of cost measures: *Medication_Cost*$_{it}$ and *Non_Medication_Cost*$_{it}$. *Medication_Cost*$_{it}$ measures the average total medication bill before insurance deduction of patients admitted to physician i in week t. *Non_Medication_Cost*$_{it}$ measures the average total medical bill before insurance deduction excluding medication of patients admitted to physician i in week t. We measure only the medication and non-medication cost due to two important reasons. First, as the health economics literatures have suggested, one of the most common consequences of physician agency problem is over-prescription. Hence, measuring medication cost fits well with the theory in this research. Second, kickbacks (commissions) to physicians from drug sales are rampant in China's health system (Today 2017). Physicians have strong incentive to prescribe more drugs. Hence, medication (prescription) cost is also very relevant to our context.

From the online data, we create three proxy measures for the offline patients' participation on the online healthcare community: *Online_Local*$_{i,t-1}$, *Online_Near*$_{i,t-1}$, and *Online_Others*$_{i,t-1}$, which measure the total number of online patients from Shanghai City, near Shanghai provinces, and all other provinces in China who started online consultation with physician i in week t-1 respectively. The near Shanghai provinces include three most adjacent provinces (i.e. Jiangsu, Zhejiang, Anhui) of Shanghai. Figure 3 shows the geo-location of Shanghai and its adjacent provinces on the map of China as well as the distribution of online patients from these three different regions.

We use two variables to measure the information quality of online physician-generated information: *information completeness* and *information amount*. These two measures belong to the information quality taxonomy developed by Wang and Strong (1996). Although there are other dimensions (e.g., accuracy, objectivity, and interpretability) of information quality in the taxonomy, measuring them require the subjective evaluations from the patients which are not available or feasible in our context. Moreover, due to the information asymmetry problem of physician actions we have reviewed earlier, it is also very hard to ensure the validity of those subjective measures. As such, we argue that, especially at the aggregated level,

---

[4] Before the physician have joined the online community, the weekly number of online patients are all zero. After the physician has joined the online community, the weekly number of online patients fluctuate from week to week and, in some weeks, the number could also be zero.

[5] Not every physician joined the online community. And not every online physician joined the online community at the same week.

the objectively-measured completeness and amount of the online physician-generated information are the most relevant and appropriate information quality dimensions in our research context.

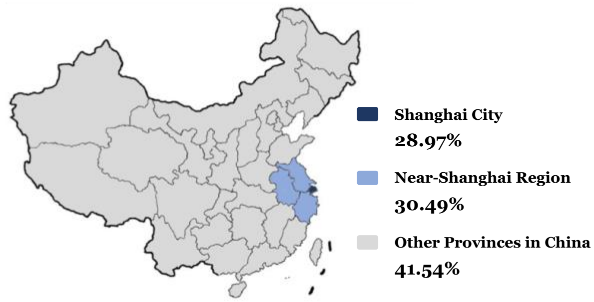

Figure 3. Distribution of Online Patients by Geo-Location

Information completeness is generally defined as the breadth and scope of the information (Wang and Strong 1996). Adapting to our context, we operationalize online physician-generated information completeness in one online patient-physician encounter using an index variable (0 - 4). It measures the coverage (breadth and scope) of information across the four general categories that we have identified earlier. Information amount is simply defined as the amount of the information (Wang and Strong 1996). Due to the dyadic nature of our research context, we operationalize online physician-generated information amount in one online patient-physician encounter as the ratio between the length of physician-generated information and length of patient-generated information in one patient-physician encounter. It gives the relative amount of physician-generated information. As with the proxy variable for patients' online participation, we measure the lagged one-week information completeness and amount of online physician-generated text information. It is reasonable to assume that online patients would more likely to view some of the most recent consultation records available to them as the basis for inference of the physician characteristics. Thus, $Information\_Completeness_{i,t-1}$ measures the average information completeness in all online patient-physician encounters happened to physician *i* in week *t-1*[6]. And $Information\_Amount_{i,t-1}$ measures the average amount of information in all online patient-physician encounters happened to physician *i* in week *t-1*. Alternative lag specifications of the measures are included in the analysis section.

Measuring the information completeness requires us to give precise labels that indicate whether a particular physician consultation text record belong to one of the four categories of physician consultation information or not. Due to the sheer volume of the online text data, it is not feasible for us to manually label all the online consultation text. Hence, we utilize the recent advancement of the deep learning natural language processing (NLP) technology to do the classification task automatically for us. To achieve this, we first develop a comprehensive labeling protocol of the online consultation text based on the classification framework we described eariler. Then we train two independent human coders to manually label a randomly selected sample of around 27,000 online consultation texts to form our training data. We conducted three rounds of manual labelling to ensure the quality of the training labels. Finally, we trained a state-of-the-art Convolutional Neural Networks (CNN) classification model (Kim 2014) on the training data. With both the classification precision and recall rate exceed 90%, we deployed the trained CNN model on the full online text and generated the corresponding labels for each text records, which we used to construct the information completeness variable.

Finally, we use the general demographic information of patients admitted to physician *i* in week *t* to control for the patients' effect on the outcome variables. They include the share of patients belong to the 0-1, 2-17, 18-34, 35-64, and 65 above age groups as well as the gender group. To account for the time-varying physician-side effects (e.g., the supply of service), we also further control for the number of his/her weekly offline admissions/surgeries. Table 1 summarizes descriptive statistics of main variables used in this study.

[6] In case of an encounter lasted more than a week, we only use the text records that happened prior to the focal week.

Table 1. Summary Statistics

Variables	All weeks (93,515 observations)				Weeks with non-zero admissions (42,753 observations)			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
$Offline\_Admission_{it}$	2.249	4.357	0	87	4.919	5.329	1	87
$Offline\_Surgery_{it}$	1.351	3.346	0	67	2.955	4.443	0	67
$Online\_Local_{i,t-1}$	0.191	1.139	0	44	0.255	1.290	0	25
$Online\_Near_{i,t-1}$	0.206	1.303	0	41	0.248	1.304	0	29
$Online\_Others_{i,t-1}$	0.278	2.276	0	109	0.260	1.650	0	54
$Inpatient\_Stay_{it}$	-	-	-	-	9.532	9.718	0	462.5
$Medication\_Cost_{it}$ *(in RMB)*	-	-	-	-	4294.476	7331.308	0	403767.3
$Non\_Medication\_Cost_{it}$ *(in RMB)*	-	-	-	-	11289.15	13543.33	0	380057.8
$Information\_Completeness_{i,t-1}$	-	-	-	-	0.089	0.369	0	4
$Information\_Amount_{i,t-1}$	-	-	-	-	0.025	0.256	0	22.165

Note: the full panel includes observations of 317 physicians (after matching) across 295 weeks.

Analysis and Results

Proxying Offline Patients' Online Participation

We first demonstrate that the number of the physician's past one-week online patients could serve as a valid proxy for offline patients' online participation. We expect that the number of the physician's online patients in week *t-1* is positively associated with the number of offline admissions in week *t*. The fixed-effect regression model is specified as follows:

$$Log(Y_{it}) = \alpha_i + \gamma_t + \beta_1 Online\_Local_{i,t-1} + \beta_2 Online\_Near_{i,t-1} + \beta_3 Online\_Others_{i,t-1} + \eta' Z_i + \varepsilon_{it}, \quad (1)$$

Where Y_{it} is the number of offline admissions or surgeries to physician *i* in week *t*. And α_i captures the physician fixed effect while γ_t captures the week fixed effect. We also control for higher level physician's medical title and department fixed effect using vector Z_i. The number of online patient variables are all in lagged forms. The estimation uses the full panel of 93,515 observations. Weeks with zero offline admission are included to ensure there is no bias in the estimation results. Table 2 shows the result of this regression. Models (1)-(4) report the result when using number of offline admissions as dependent variable and Models (5)-(8) report the result when using number of offline surgeries as dependent variable. In all our results in this paper, we report the robust standard errors clustered at the physician level in the parentheses.

Table 2. Proxying Offline Patients' Online Participation

Variables	DV: *Log(Offline_Admission)*				DV: *Log(Offline_Surgery)*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Online_Local	0.018** (0.006)	0.018** (0.006)	0.008 (0.005)	0.005 (0.004)	0.020** (0.007)	0.020** (0.007)	0.009 (0.005)	0.006 (0.004)
Online_Near	0.026* (0.011)	0.026* (0.011)	0.018* (0.008)	0.016* (0.006)	0.024* (0.011)	0.024* (0.011)	0.017* (0.008)	0.013* (0.007)
Online_Others	-0.003 (0.004)	-0.003 (0.004)	-0.003 (0.003)	-0.003 (0.002)	-0.003 (0.004)	-0.003 (0.004)	-0.003 (0.003)	-0.003 (0.002)
Lag Specification	Past 1 week	Past 1 week	Past 2 weeks	Past 3 weeks	Past 1 week	Past 1 week	Past 2 weeks	Past 3 weeks
Subsample	Unmatched	Matched	Matched	Matched	Unmatched	Matched	Matched	Matched
Adjusted R^2	0.605	0.599	0.600	0.600	0.606	0.618	0.618	0.618

Note: *p<0.05; **p<0.01; ***p<0.001.

We focus on interpreting the results presented in models (2)-(4) here. Sorting by the geographical distance, it suggests that one additional local online patient who consulted the focal physician in week *t-1* will lead to, on average, 1.8% increase in the total number of offline admissions to this physician in week *t*. And one additional online patient from near Shanghai region who consulted the focal physician in week t-1 will lead to, on average, 2.6% increase the total number of offline admissions to this physician in week *t*. However, as distance grows further, this online to offline correlation becomes insignificant. This is reasonable. As the travelling cost increases with geographical distance, it will lead to higher inconveniences cost that prevent the online to offline conversion. Along the temporal distance, the positive and significant online to offline correlation of local online patient is transient. The effect becomes insignificant if the number of online local patients include beyond past one week. However, the positive and significant online-offline correlation of online patient from near Shanghai region is much more persistent. The effect remains significant even when lagged three weeks of observations are included. We argue that this is also reasonable due to that local online patients enjoy more flexibility about when to make the offline visit, whereas non-local patients would be more likely to schedule ahead to offset the geographical and inter-provincial administrative inconvenience. Such intuitive results suggest that the number of past week online patients by regions is indeed a valid proxy for the offline patients' online participation in a particular week. Using the number of offline surgeries as alternative dependent variable also gives us the similar result, which further lends support to our claim on the validity of our proxy variables.

For triangulation purpose, we further demonstrate with qualitative evidences that the online-to-offline correlation do exists in real world. Below is an excerpt from the online certified personal blog of a pediatric physician published in 2014[7]:

> *"In 2013 … we had in total 4,011 inpatient admission, which represents a 320% increase from 2009 … Besides, most of our patients were from non-local regions … and more than 50% of patients are coming from various online channels."*

It confirms that not only do the online-to-offline patient conversion exist, but also that the online channels could account for a significant portion of the offline patient pool. Combining this qualitative evidence with our earlier quantitative evidence, we are confident in the validity of our proxy variable approach.

Impact on Offline Inpatient Care Quality and Cost

We estimate the effect of our proxy variables of the offline patients' online participation on the offline inpatient care quality and cost by employing the fixed-effect estimator. Equation (2) outlines our model specification:

$$Log(Y_{it}) = \alpha_i + \gamma_t + \beta_1 Online\_Local_{i,t-1} + \beta_2 Online\_Near_{i,t-1} + \beta_3 Online\_Others_{i,t-1} + \delta' E_{it} + \eta' Z_i + \varepsilon_{it}, \quad (2)$$

where our outcome of interest Y_{it} is either the weekly average inpatient length of stay $Inpatient\_Stay_{it}$, average medication cost $Medication\_Cost_{it}$, or average non-medication cost $Non\_Medication\_Cost_{it}$. Similarly, α_i captures the physician fixed effect while γ_t captures the week fixed effect. And we also control for time-varying offline patient effect such as aggregated weekly patients' age and gender compositions, average length of stay (not in controls if the dependent variable is *Inpatient_Stay*), share of surgeries, and number of admissions using vector E_{it}. Vector Z_i is used to control for time-invariant higher-level physician fixed effects such as medical title and department. This estimation (and in all subsequent analysis) uses only weeks with non-zero offline admissions. We also only use one-week lagged number of online patients to proxy offline patients' online participation in the subsequent analysis. Table 3 reports the estimation results.

Assessing the impact on inpatient care quality, from model (2), we can find that the coefficient for *Online_Local* is negative and significant ($p<0.001$). It suggests that one additional local online patient who consulted the focal physician in week *t-1* will lead to, on average, 1.7% decrease in the average length of stay for offline patients admitted to this physician in week *t*. However, there is no significant effect of the number of near Shanghai online patients on the offline inpatient length of stay. Moreover, assessing the impact on inpatient care cost, from model (4), we can also find that the coefficient for *Online_Local* is negative and significant, suggesting that one additional local online patient consulted the focal physician in week *t-1* will lead to, on average, 1.3% decrease in the average medication cost for offline patients admitted to this physician in week *t*. The effect of the number of near Shanghai online patients is also insignificant. However, similar results do not apply to non-medication cost. On the contrary, while we do not find a significant effect of the number of local online patients on the offline inpatient non-medication cost, we find that the number

[7] http://blog.sina.com.cn/s/blog_6200895d0101gzyl.html (in Simplified Chinese)

of online patients from near Shanghai region has a positive and significant effect on offline inpatient non-medication cost. And the coefficient suggests that one additional near Shanghai region online patient who consulted the focal physician in week *t-1* will lead to, on average, 1.5% increase in the average non-medication cost for offline patients admitted to this physician in week *t*.

Table 3. Impact on Offline Inpatient Care Quality and Cost

Variables	Log(*Inpatient_Stay*)		Log(*Medication_Cost*)		Log(*Non_Medication_Cost*)	
	(1)	(2)	(3)	(4)	(5)	(6)
Online_Local	-0.017*** (0.002)	-0.017*** (0.002)	-0.014** (0.005)	-0.013** (0.005)	0.001 (0.003)	0.001 (0.003)
Online_Near	-0.005 (0.004)	-0.005 (0.004)	0.006 (0.008)	0.006 (0.008)	0.015*** (0.005)	0.015*** (0.004)
Online_Others	-0.003 (0.003)	-0.003 (0.003)	0.004 (0.005)	0.004 (0.005)	0.004 (0.003)	0.003 (0.003)
Share_Female	-0.037*** (0.007)	-0.026* (0.011)	-0.120*** (0.017)	-0.040 (0.026)	-0.005 (0.011)	-0.003 (0.018)
Share_Age0_1	Omitted	Omitted	Omitted	Omitted	Omitted	Omitted
Share_Age2_17	-0.045 (0.027)	-0.083 (0.1130	0.446*** (0.088)	0.331 (0.216)	-0.246*** (0.041)	-0.101 (0.157)
Share_Age18_34	-0.057 (0.030)	-0.145 (0.127)	0.785*** (0.108)	0.553** (0.218)	-0.129* (0.055)	0.038 (0.168)
Share_Age35_64	0.030 (0.030)	-0.046 (0.126)	1.198*** (0.110)	0.940*** (0.212)	0.110* (0.056)	0.274 (0.169)
Share_Age65	0.249*** (0.032)	0.178 (0.126)	1.297*** (0.114)	1.326*** (0.213)	0.332*** (0.056)	0.491** (0.167)
Share_Surgical	0.291*** (0.004)	0.328*** (0.029)	0.680*** (0.009)	0.734*** (0.111)	1.382*** (0.036)	1.505*** (0.063)
Offline_Admission	0.003* (0.001)	0.001 (0.001)	0.030*** (0.005)	0.029*** (0.004)	0.023*** (0.004)	0.023*** (0.005)
Inpatient_Stay	-	-	0.027*** (0.002)	0.034*** (0.004)	0.020*** (0.001)	0.023*** (0.003)
Lag Specification	Past 1 week	Past 1 week	Past 1 week	Past 1 week	Past 1 week	Past 1 week
Subsample	Unmatched	Matched	Unmatched	Matched	Unmatched	Matched
Adjusted R^2	0.329	0.213	0.483	0.435	0.625	0.624
Note: *p<0.05; **p<0.01; ***p<0.001.						

In terms of the magnitude of the effect, since the number of local online patients is just a proxy measure of the offline local patients' online participation and not every offline patient would participate in the online healthcare community prior to offline visit, we suggest that our result underestimates the true effect of the offline local patients' online participation on the offline length of stay and medication cost outcome. The magnitude of the true effects should be much larger. In models (1), (3), and (5), we report the estimation results using unmatched sample to serve as a benchmark for comparison with the matched sample. The comparable coefficients between the matched and unmatched samples suggest that physician selection issue might not be a serious concern in this study.

Although we conducted the PSM on physicians and controlled for various relevant physician and time fixed effects, our estimation results can still be biased if patients who chose to participate in the online healthcare community are systematically different from those who chose not to participate. To demonstrate that this potential patient-side selection issue does not affect our estimation results, we further estimated the effect of a physician's number of online patients in week *t-1* on the age and gender composition of offline patients admitted to this physician in week *t*. Table 4 shows the estimation results. The insignificant coefficients dismiss the concern that the potential patient-side selection issue may contaminate our results. Another potential source of bias may come from the influence of outliners in our dependent variables. We define outliners as data points that are three standard deviations away from the sample mean. Dropping the outliners does not change neither the significance nor the direction of the estimation results but slightly shifts the magnitude of some coefficients. Hence, we also dismiss the outliner concerns.

Table 4. Impact on Offline Inpatients' Age and Gender Compositions

Variables	Share (%) of patients admitted in week *t* to physician *i* by age and gender groups					
	Age 0-1	Age 2-17	Age 18-34	Age 35-64	Age 65+	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Online_Local	0.000 (0.000)	-0.001 (0.001)	0.000 (0.002)	0.000 (0.002)	-0.014 (0.012)	0.003 (0.002)
Online_Near	-0.000 (0.000)	0.000 (0.001)	0.000 (0.001)	-0.001 (0.001)	-0.039** (0.0014)	-0.002 (0.003)
Online_Others	-0.000 (0.000)	0.001 (0.001)	0.000 (0.001)	-0.000 (0.002)	-0.013 (0.007)	0.000 (0.003)
Lag Specification	Past 1 week	Past 1 week	Past 1 week	Past 1 week	Past 1 week	Past 1 week
Subsample	Matched	Matched	Matched	Matched	Matched	Matched
Adjusted R^2	0.024	0.431	0.288	0.134	0.363	0.260
Note: *p<0.05; **p<0.01; ***p<0.001.						

The Influence of Online Physician-Generated Information Quality

To estimate the influence of online physician information quality on the offline care outcomes, we further incorporate the online physician information quality variables. The new model is specified in Equation (3):

$$Log(Y_{it}) = \alpha_i + \gamma_t + \beta_1 Online\_Local_{i,t-1} + \beta_2 Online\_Near_{i,t-1} + \beta_3 Online\_Others_{i,t-1} + \beta_4 Information\_Completeness_{i,t-1} + \beta_5 Information\_Amount_{i,t-1} + \delta' E_{it} + \eta' Z_i + \varepsilon_{it}, \quad (3)$$

where $Information\_Completeness_{i,t-1}$ and $Information\_Amount_{i,t-1}$ capture the overall online physician information quality to the online patients who consulted patient *i* in week *t-1*. The interpretation of all other terms in Equation (3) is identical to Equation (2). In previous section, we have shown that the offline patients' online participation positively influences the offline inpatient care quality, but the cost reduction effect is limited only to the medication cost. Hence, in this section, we will only further explore the influence of online physician information quality on the length of stay and medication. Table 5 reports the estimation results (estimation results of the control variables are not reported here due to paper length limitation).

Table 5. The Influence of Online Physician-Generated Information Quality

Variables	Log(*Inpatient_Stay*)			Log(*Medication_Cost*)		
	(1)	(2)	(3)	(4)	(5)	(6)
Online_Local	-0.014*** (0.002)	-0.013*** (0.002)	-0.012*** (0.002)	-0.017*** (0.004)	-0.017*** (0.005)	-0.018*** (0.003)
Online_Near	-0.005 (0.004)	-0.004 (0.004)	0.004 (0.004)	0.006 (0.008)	0.007 (0.009)	0.007 (0.009)
Online_Others	-0.002 (0.003)	-0.001 (0.003)	0.001 (0.003)	0.005 (0.004)	0.005 (0.004)	0.005 (0.004)
Information_ Completeness	-0.021 (0.014)	-0.031* (0.014)	-0.033* (0.014)	-0.002 (0.023)	0.002 (0.024)	0.009 (0.024)
Information_ Amount	-0.023*** (0.006)	-0.024*** (0.007)	-0.024*** (0.007)	0.003 (0.013)	-0.013 (0.034)	-0.015 (0.034)
Online Patient Lag Specification	Past 1 week	Past 1 week	Past 1 week	Past 1 week	Past 1 week	Past 1 week
Online Information Lag Specification	Past 1 week	Past 25 consultations	Past 50 consultations	Past 1 week	Past 25 consultations	Past 50 consultations
Subsample	Matched	Matched	Matched	Matched	Matched	Matched
Adjusted R^2	0.214	0.214	0.214	0.437	0.437	0.437
Note: *p<0.05; **p<0.01; ***p<0.001.						

Model (1)-(3) report the estimation results when using inpatient length of stay as the dependent variable while model (4)-(6) use inpatient medication cost as the dependent variable. Model (1) and (4) use the standard one-week lag specification for the online physician-generated information quality. We also additionally include two alternative online information lag specifications (i.e., past 25 or 50 consultations)[8]. We find negative and significant coefficients for the effect of online information quality on offline length of stay, but the effect of information completeness only become significant with the two alternative lag specifications. Since we do not know exactly how many past consultations one patients would go through to extract inference about the physician characteristics, our results in general suggest that the online information quality to an online patient would positively influence the subsequent offline care quality. However, we find no significant effect of online information quality on the offline medication cost.

Discussion and Conclusion

Key Findings

This paper reveals three important findings regarding the impact of online healthcare communities on the offline inpatient care quality and cost. First, we find that the offline patients' participation on the online healthcare community positively impact the offline care efficiency as we observe decreased inpatient length of stay and medication cost. However, we find no effect on the offline inpatient non-medication cost. We suggest that it is because non-medication costs including medical examination, testing, surgery, as well as administrative fees, compared to medication cost, are less influenced by an individual physician but the hospital system in general. Since we have argued that offline patients' online participation influence offline care and cost through enabling patients to find a better matched individual physician for them, the effect of online participation will also be likely to have limited effect on the non-medication cost.

Second, we also find that the geographic origin of the online patients constrains the positive effect of patients' online participation on the offline inpatient cost reduction and quality improvement. Not only is there no effect on both the inpatient length of stay and medication cost, but also that the non-medication cost will increase if the online patients are of non-local origins. This finding confirms that non-local patients may have different motivation/interpretation for offline care efficiency. They are less sensitive to price and the travelling cost and inconvenience may even induce them to stay longer.

Finally, we find that the quality of online physician-generated information negative influences the offline inpatient length of stay. But it does not have any effect on the offline inpatient medication cost. One possible interpretation of such results may be that online patients mainly use the online physician-generated information to make inference about the skills/quality aspect of physician characteristic as the main criteria for selecting which physician to consult offline. Hence, the online information quality will have a direct impact on offline inpatient care quality. However, the channel through which offline inpatients' online participation influences the subsequent offline cost reduction may not be simply the selection of physician based on their online information quality but rather an empowered patient also has more negotiation power in the cost aspect of the subsequent offline healthcare visit. This offline negotiation process, however, cannot be captured by the quality of online physician-generated information.

Limitations

There are several limitations in our findings. First, the care quality measurement, if available, should further include patients' self-reported care satisfaction. Second, individual level analysis may potentially provide more insight if future study could obtain relevant dataset that can identify individual patients. Finally, it will also be interesting to investigate the influence of patients' online participation on the outcomes in repeated offline visit settings. We are currently not able to study this long-term effect with our dataset.

Theoretical Contributions

Notwithstanding the limitations, our study also makes several contributions to the extant literature. First, this research makes the first effort to investigate the influence of online healthcare in the offline inpatient care settings. The past literatures of online healthcare have predominantly focused on investigating its impact on offline healthcare in the telemedicine settings, where the observed efficiency-improvement effects are primarily explained using the substitution effect. Some researchers have also argued that the

[8] On this online community, the maximum number of past consultations shown per page is 25.

overall health benefits in this setting may be marginal and limited (Watson et al. 2016). By examining the impact of online healthcare community on the offline inpatient care outcomes, this study adds to a more comprehensive theoretical understanding of the impact of online healthcare.

Second, our findings also contribute to the health economics literature. Due to the serious information asymmetry problem, one major concern in the health economics literature is that, without sufficient institutional protections, patients are generally incapable of effectively influencing their healthcare outcomes. Hence the literature that focused on healthcare efficiency and patient welfare have mainly adopted the institutional or physician perspectives. Our findings, however, help resolve this concern by demonstrating that patients' participation on the online healthcare community would lead to improved care quality and decreased care costs in their subsequent offline healthcare visit. The patients are able to achieve such outcomes solely by participating in the online community without government or hospital interference. Our study hence highlights the increasing theoretical importance of examining economic role/power of patient in the physician-patient interaction in this information age.

Third, our findings also extend the theory of physician agency (McGuire 2000) by further investigating the value of online physician-generated information to patients. Although this theory has identified the information about physician characteristics as one of the asymmetrically distributed information between the physicians and the patients, it does not offer useful discussion on how this asymmetric information problem could be resolved. Our results show that the quality of online physician-generated information to online patients positively influences the subsequent offline inpatient care quality, indicating possible online patients' selections of better-match physicians through extracting inference of physician characteristics from online physician-generated information. As such, our study adds to a more nuanced understanding of the value of online healthcare information to the patients in face of physician agency problem.

Finally, our findings further reveal several boundary conditions of the effect of patients' online participation on offline care quality and cost. They help further elucidate the theoretical mechanisms underlying such influence. We find that the online to offline healthcare efficiency-improvement effect only exist among local patients. We also find that the cost reduction effect is only effective in reducing the medication cost not the non-medication cost. Together they imply that there exist critical antecedents from both the patient (e.g., patient incentives) and physician (e.g., physician's ability to influence the price) side to this online to offline healthcare efficiency improvement effect. Acknowledging them will be important in studying the impact of online healthcare services on offline outcomes.

Implication to Practice

There are several important managerial implications in this study. To healthcare policymakers, our study suggests that supporting the development of online healthcare communities could provide a potentially effective way to improve the efficiency of healthcare systems. However, in doing so, the government also need to pay close attention to efficiency-decreasing activity such as healthcare tourism that is further enabled by such platform. Decreasing the healthcare resource disparity across different geo-regions should still be a focus in national health policy. To online healthcare platform owners, our study implies that there exists tremendous social/economic value in providing open healthcare and physician information to the general public. Hence, the platform should increase rather than constrain the supply of such information.

In conclusion, this research is among the first to study the impact of online healthcare communities on offline inpatient care quality and cost. Our findings suggest that offline patients' participation on the online healthcare communities leads to improved offline inpatient care efficiency. We also underscore the unique value of online physician-generated information to patients in face of the physician agency. Policymakers and online healthcare platform owners can create specific strategies based on our findings to further promote the economic as well as social value of the platform through providing online healthcare services.

Acknowledgements

This research was supported by the Singapore Ministry of Education (MOE) Academic Research Fund (AcRF) Tier 1 grant. And the authors would like to thank the Track Chairs, Associate Editor and anonymous reviewers for their valuable comments and suggestions.

References

Agarwal, R., Gao, G., DesRoches, C., and Jha, A. K. 2010. "Research Commentary—the Digital Transformation of Healthcare: Current Status and the Road Ahead," *Information Systems Research* (21:4), pp. 796-809.

Albert, S. M., Shevchik, G. J., Paone, S., and Martich, G. D. 2011. "Internet-Based Medical Visit and Diagnosis for Common Medical Problems: Experience of First User Cohort," *Telemedicine and e-Health* (17:4), pp. 304-308.

Ambrus, A., and Lu, S. E. 2014. "Almost Fully Revealing Cheap Talk with Imperfectly Informed Senders," *Games and Economic Behavior* (88), pp. 174-189.

Arrow, K. J. 1963. "Uncertainty and the Welfare Economics of Medical-Care," *American Economic Review* (53:5), pp. 941-973.

Arrow, K. J. 1986. "Chapter 23 Agency and the Market," in *Handbook of Mathematical Economics*. Elsevier, pp. 1183-1195.

Battaglini, M. 2002. "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica* (70:4), pp. 1379-1401.

Bhandari, N., Shi, Y. F., and Jung, K. 2014. "Seeking Health Information Online: Does Limited Healthcare Access Matter?," *Journal of the American Medical Informatics Association* (21:6), pp. 1113-1117.

Chernev, A., and Carpenter, G. S. 2001. "The Role of Market Efficiency Intuitions in Consumer Choice: A Case of Compensatory Inferences," *Journal of Marketing Research* (38:3), pp. 349-361.

Choné, P., and Ma, C.-t. A. 2011. "Optimal Health Care Contract under Physician Agency," *Annals of Economics and Statistics* (101/102), pp. 229-256.

De Jaegher, K. 2012. "The Value of Private Patient Information in the Physician-Patient Relationship: A Game-Theoretic Account," *Computational and Mathematical Methods in Medicine* (2012), pp. 1-16.

De Jaegher, K., and Jegers, M. 2001. "The Physician–Patient Relationship as a Game of Strategic Information Transmission," *Health Economics* (10:7), pp. 651-668.

Dulleck, U., Kerschbamer, R., and Sutter, M. 2011. "The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition," *American Economic Review* (101:2), pp. 526-555.

Farrell, J., and Rabin, M. 1996. "Cheap Talk," *Journal of Economic perspectives* (10:3), pp. 103-118.

Gao, G. G., McCullough, J. S., Agarwal, R., and Jha, A. K. 2012. "A Changing Landscape of Physician Quality Reporting: Analysis of Patients' Online Ratings of Their Physicians over a 5-Year Period," *Journal of medical Internet research* (14:1), p. e38.

Gaynor, M. 1994. "Issues in the Industrial Organization of the Market for Physician Services," *Journal of Economics & Management Strategy* (3:1), pp. 211-255.

Goh, J. M., Gao, G. D., and Agarwal, R. 2016. "The Creation of Social Value: Can an Online Health Community Reduce Rural-Urban Health Disparities?," *Mis Quarterly* (40:1), pp. 247-263.

Guo, S., Guo, X., Fang, Y., and Vogel, D. 2017. "How Doctors Gain Social and Economic Returns in Online Health-Care Communities: A Professional Capital Perspective," *Journal of Management Information Systems* (34:2), pp. 487-519.

Hickson, R., Talbert, J., Thornbury, W. C., Perin, N. R., and Goodin, A. J. 2015. "Online Medical Care: The Current State of "Evisits" in Acute Primary Care Delivery," *Telemedicine and e-Health* (21:2), pp. 90-96.

Hussey, P. S., Wertheimer, S., and Mehrotra, A. 2013. "The Association between Health Care Quality and Cost: A Systematic Review," *Annals of internal medicine* (158:1), pp. 27-34.

Johnston, A. C., Worrell, J. L., Di Gangi, P. M., and Wasko, M. 2013. "Online Health Communities: An Assessment of the Influence of Participation on Patient Empowerment Outcomes," *Information Technology & People* (26:2), pp. 213-235.

Kalda, R., Põlluste, K., Lember, M., Nordiska, m., and Nordic School of Public Health, N. H. V. 2003. "Patient Satisfaction with Care Is Associated with Personal Choice of Physician," *Health policy* (64:1), pp. 55-62.

Kim, Y. 2014. "Convolutional Neural Networks for Sentence Classification," *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1746–1751.

Kittelsen, S. A., Anthun, K. S., Goude, F., Huitfeldt, I., Häkkinen, U., Kruse, M., Medin, E., Rehnberg, C., and Rättö, H. 2015. "Costs and Quality at the Hospital Level in the Nordic Countries," *Health economics* (24:S2), pp. 140-163.

Lechner, M. 2002. "Some Practical Issues in the Evaluation of Heterogeneous Labour Market Programmes by Matching Methods," *Journal of the Royal Statistical Society. Series A (Statistics in Society)* (165:1), pp. 59-82.

Lee, V. S., Kawamoto, K., Hess, R., Park, C., Young, J., Hunter, C., Johnson, S., Gulbransen, S., Pelt, C. E., and Horton, D. J. 2016. "Implementation of a Value-Driven Outcomes Program to Identify High Variability in Clinical Costs and Outcomes and Association with Reduced Cost and Improved Quality," *Jama* (316:10), pp. 1061-1072.

Lion, K. C., Wright, D. R., Spencer, S., Zhou, C., Del Beccaro, M., and Mangione-Smith, R. 2016. "Standardized Clinical Pathways for Hospitalized Children and Outcomes," *Pediatrics* (137:4), pp. peds. 2015-1202.

McGuire, T. G. 2000. "Physician Agency," in *Handbook of Health Economics*. Elsevier, pp. 461-536.

Obama, B. 2016. "United States Health Care Reform: Progress to Date and Next Steps," *Jama* (316:5), pp. 525-532.

Ong, L. M. L., Dehaes, J., Hoos, A. M., and Lammes, F. B. 1995. "Doctor-Patient Communication - a Review of the Literature," *Social Science & Medicine* (40:7), pp. 903-918.

Orszag, P. R. 2016. "Us Health Care Reform: Cost Containment and Improvement in Quality," *Jama* (316:5), pp. 493-495.

Phelps, C. E. 1986. "Induced Demand — Can We Ever Know Its Extent?," *Journal of Health Economics* (5:4), pp. 355-365.

Pocock, N. S., and Phua, K. H. 2011. "Medical Tourism and Policy Implications for Health Systems: A Conceptual Framework from a Comparative Study of Thailand, Singapore and Malaysia," *Globalization and Health* (7:1), pp. 1-12.

Reid, R. O., Friedberg, M. W., Adams, J. L., McGlynn, E. A., and Mehrotra, A. 2010. "Associations between Physician Characteristics and Quality of Care," *Archives of Internal Medicine* (170:16), pp. 1442-1449.

Serrano, C. I., and Karahanna, E. 2016. "The Compensatory Interaction between User Capabiities and Technology Capabilities in Influencing Task Performance: An Empirical Assessment in Telemedicine Consultations," *MIS Quarterly* (40:3), pp. 597-621.

Taheri, P. A., Butz, D., Griffes, L. C., Morlock, D. R., and Greenfield, L. J. 2000. "Physician Impact on the Total Cost of Care," *Annals of surgery* (231:3), p. 432.

Today. 2017. "Kickbacks from Drug Firms Plague China's Health System." Retrieved April 19, 2018, from https://www.todayonline.com/world/kickbacks-drug-firms-plague-chinas-health-system

Tunis, S. R., and Pearson, S. D. 2006. "Coverage Options for Promising Technologies: Medicare's 'Coverage with Evidence Development'," *Health Affairs* (25:5), pp. 1218-1230.

Visser, L. M., Bleijenbergh, I. L., Benschop, Y. W., Van Riel, A. C., and Bloem, B. R. 2016. "Do Online Communities Change Power Processes in Healthcare? Using Case Studies to Examine the Use of Online Health Communities by Patients with Parkinson's Disease," *BMJ open* (6:11), p. e012110.

Wang, R. Y., and Strong, D. M. 1996. "Beyond Accuracy: What Data Quality Means to Data Consumers," *Journal of management information systems* (12:4), pp. 5-33.

Watson, J., Salisbury, C., Atherton, H., Campbell, J., McKinstry, B., and Ziebland, S. 2016. "Proliferation of Private Online Healthcare Companies." British Medical Journal Publishing Group, p. i1076.

Wong, H. S. 1996. "Market Structure and the Role of Consumer Information in the Physician Services Industry: An Empirical Test," *Journal of Health Economics* (15:2), pp. 139-160.

Yan, L., and Tan, Y. 2014. "Feeling Blue? Go Online: An Empirical Study of Social Support among Patients," *Information Systems Research* (25:4), pp. 690-709.

Yaraghi, N., Du, A. Y., Sharman, R., Gopal, R. D., and Ramesh, R. 2014. "Health Information Exchange as a Multisided Platform: Adoption, Usage, and Practice Involvement in Service Co-Production," *Information Systems Research* (26:1), pp. 1-18.