Hospital Competition and Quality: Evidence from the Entry of the High-Speed Train in South Korea*

Hyesung Yoo† Maria Ana Vitorino‡ Song Yao§

April 2022

Abstract

This paper leverages the entry of a high-speed train (HST) system in South Korea as a natural experiment to establish the causal effect of competition among hospitals on health care quality and consumer welfare. We implement a difference-in-differences research design that exploits the differential effect of the HST entry on hospitals based on their distance to train stations. Our results suggest that increased competition intensity caused by increased hospital substitutability leads to better quality of clinical care. To evaluate the overall impact of the entry of the HST on patients welfare, we estimate a structural model of hospital choice, allowing for a flexible formation of patients' consideration sets. We find that patients living near an HST station experience an improvement in welfare from improvements in hospital quality in addition to reduction in travel time. Patients living further away from HST stations also experience an improvement in welfare – even though they do not gain from the reduced travel time, these patients benefit from the improvement in the quality of the hospitals that are located close to HST stations. We also find that the HST can have a beneficial impact on patients' health by facilitating patients' sorting to better hospitals, even while holding constant the quality of clinical care.

*We thank Dinara Akchurina, Mark Bergen, Andrew Ching, Thomas Holmes, Nitin Mehta, Amil Petrin, Mengze Shi, Joel Waldfogel and Linli Xu for helpful comments. We also thank the discussants Stephan Seiler and Ying Xie for their feedback as well as seminar participants at the University of Toronto, University of Minnesota, Sorbonne University, Cornell University, University of Chicago, Washington University in St. Louis, CEPR Industrial Organization Seminar Series, European Quantitative Marketing Seminar, INFORMS Marketing Science Conference 2018, and UT Dallas Bass FORMS Conference 2021. We are also grateful to Gautam Gowrisankaran and Kelli Marquardt for their help with the ggt program. The opinions expressed in this article are the authors' and do not reflect the view of the South Korean National Health Insurance Services and of the South Korean Health Insurance Review Services.

†University of Toronto: hyesung.yoo@utoronto.ca

‡INSEAD: maria-ana.vitorino@insead.edu

§Washington University in St. Louis: songyao@wustl.edu

1 Introduction

How does competition affect clinical quality in the health care industry? This is an important question because many countries, including United Kingdom, Netherlands, Belgium, Israel, and Australia, which historically have been providing healthcare through centralized non-market means have recently to adopted or are considering market oriented reforms, despite weak evidence on its effects on patient outcomes. Due to the size and impact of the health care industry on welfare, correctly assessing the impact of competition on health outcomes is of crucial importance for policymakers.

Assessing the impact of competition on health care is complex due to the fact that competition in health care markets is based on geography. Hospitals compete in geographical markets because patients have a strong preference, among other things, for hospitals that are located closed to their home. Since some geographies are intrinsically more competitive than others, it is difficult to separate the effect of competition from geographical factors using only cross-sectional analysis. Therefore, many researchers have used changes in cross-sectional variation in levels of market structure over time to identify the impact of competition. However, the challenge again lies in the fact that the market structure is endogenous: quality of incumbent hospitals and potential entrants in a given geographical region may affect their strategic entry and exit decisions.

To address this issue, more recent papers have been exploiting changes in market structure induced by health-related policies, which are seen as exogenous shocks that spur competition (Gaynor et al. 2013). Yet when policies themselves are *health*-related, the analyses can be complicated by the fact that they may affect the behavior of the agents involved in ways unanticipated by researchers. For example, the U.K. government mandated in 2006 that patients be offered a choice of five hospitals when referred to a hospital by their physician. However, there is evidence that not all primary case physicians thought that patients were able to or wanted to make choices (Gaynor et al. 2013). If such behavior are not accounted for in the analysis, conclusions may be biased.

In this article we leverage the entry of the high-speed train (henceforth HST) system in South Korea to examine the effects of competition on the quality of health care. In April 2004, Korea Train eXpress (KTX) started operating in South Korea, connecting many large and small cities via high-speed rail system. The An important aspect of the South Korean healthcare industry is that patients have the full freedom to go to any hospital of their choice with some financial incentives, and the fee for each medical procedure is fixed by the South Korean National Health Insurance

(NHI) Corporation. The introduction of the HST represents an exogenous shock to the healthcare market in that it greatly reduced patients' travel time, and enabled patients to consider hospitals that were previously unreachable due to long travel distances, thereby increasing substitutability between hospitals. According to media reports, the proportion of rural patients choosing the top four largest hospitals in Seoul increased from 41.2% in 2002 to 48.5% in 2007 as a result of the HST.[1] In addition, a survey by Kim et al. (2008) of HST passengers arriving in Seoul by train reports that 36% of passengers had at some point used the HST to seek hospital treatment in Seoul.

The importance of the tradeoff between quality and travel time that patients face has been highlighted in Tay (2003). This tradeoff between quality and travel time is what gives hospitals market power, especially when patients have the full freedom to choose any hospital. The entry of the HST alleviates the tradeoff between quality and travel time, leading to increased competition between hospitals. Standard models of hospital non-price competition predict that, conditional on price being set above the marginal cost, competition becomes intensified with more hospitals and this leads to higher quality, regardless of their ownership status (public, not-for-profit, for profit). Low levels of NHI fees have been a subject of recurrent complaint by providers in South Korea, and it is a well known fact that due to low price margins, attracting many patients is of vital importance for hospitals in South Korea. [2] In fact, when "Super Rapid Train (SRT)", a different high-speed rain system run by a private company, entered in 2016, several major hospitals started operating shuttle busses to- and from SRT stations to their hospitals.[3] Although in 2004 hospitals didn't respond with shuttle busses, the entry of the HST facilitates access to patients' preferred hospitals, implying that hospitals that were previously competing for patients locally are now competing with those located further away. In fact, there is an anecdotal evidence suggesting that hospitals responded to the increased competition caused by the HST: according to one medical personnel, several hospitals in non-Seoul regions started adopting expensive equipment and strengthening their services in response to the entry of HST. [4]

The objectives of this paper are twofold: First, we examine the impact of competition on hospital quality. Second, we decompose changes in patients' welfare and health outcomes into those caused

[1] Source: http://news20.busan.com/controller/newsController.jsp?newsId=20110804000124 (in Korean), accessed on July 10, 2018.

[2] Source: https://www.rapportian.com/news/articleView.html?idxno=15659 (in Korean), accessed on April 18, 2022.

[3] Source: https://www.docdocdoc.co.kr/news/articleView.html?idxno=1043608 (in Korean), accessed on April 18, 2022.

[4] Source: https://www.donga.com/news/It/article/all/20050323/8172270/1 , accessed on April 18, 2022.

by patient sorting to better hospitals, and those caused by improvements in hospital quality. To achieve our first objective, we rely on the fact that the HST stations do not extend to all regions, making the intensity of the competition induced by the HST to vary depending on the hospital's proximity to the nearest train station. We exploit the exogenous variation in hospitals' proximity to the HST station to identify the impact of competition on hospital clinical quality, as measured by 30-day mortality outcomes following patients' admissions for a surgical procedure.[5]

We argue that the proximity of a HST station and therefore the impact of competition to any given hospital is exogenous for several reasons: First, the HST didn't enter into all the major cities. For example, the HST did not enter into metropolitan areas such as Ulsan or Incheon, and other cities such as Pohang, Jeonju, Chuncheon, Cheonju, etc. Second, while it is true that the HST entered several major, it also entered into many rural regions that are aligned in between the stations connecting the cities. Finally, even within a given city, there is substantial variation in hospital's proximity to a HST station. For example, there is evidence that a few hospitals in Seoul that are located particularly close to the HST station experienced a disproportionately huge influx of rural patients post HST compared to other hospitals in Seoul that are located further away from the HST station.

To achieve our second objective, we develop and estimate a structural model of hospital choice wherein patients are time-constrained and use the model estimates to perform counterfactuals when the HST is removed. Specifically, in our counterfactuals, we decompose the effects of the HST on patient welfare and health outcomes along two dimensions: First, we quantify the changes in welfare and health outcomes caused by patients sorting themselves to better hospitals via HST, keeping the quality of care constant. Second, we quantify the changes in welfare and health outcomes caused by improvement in hospital quality, keeping travel time constant.

In our analysis, we consider all surgeries that were conducted during the period of study and for which mortality rate can be used as a measure of hospital quality.[6]However, using raw mortality rates as a measure of clinical quality is problematic due to patient selection issue: patients' hospital choice is non-random. Therefore, to minimize the contamination of hospital quality with patient

[5]Unlike some papers that only consider in-hospital deaths within 30 days of admission, we consider all deaths within 30 days of admission, regardless of where the death occurs. This is because regulated prices give hospitals an incentive to discharge patients prematurely Kosecoff et al. (1990). As pointed out by Gaynor et al. (2013), focusing only on in-hospitals deaths disregards this damaging response.

[6]Although ideally we would have wanted to look at patients suffering from one specific illness, or who underwent one specific type of surgery, this prevents us from doing any meaningful analysis because it leaves us with too few observations per hospital due to the fact that our data is a 2 percent random sample of the entire population.

selection, we employ the Bayesian inference method developed by Geweke et al. (2003) to obtain selection-adjusted measures of clinical quality.

We find that increased competition improves the clinical quality of hospitals, i.e., hospitals facing greater competition due to their close proximity to train stations experience a greater improvement in quality compared to hospitals located further away from the train stations. Our counterfactuals from the structural model shows that patients living close to train stations experience an improvement in welfare due to reduction in travel costs as well as enhanced clinical quality. Patients living far from HST stations do not benefit in terms of travel costs, but they also experience an increase in welfare because many of them choose to go to hospitals whose clinical quality were positively affected by the HST. We further use the model estimates to measure the impact of the entry of the HST on patient's health outcomes (i.e., surgery survival) by comparing the number of deaths in the post-HST period to those in a counterfactual scenario in which the HST is removed. From this analysis we find that a substantial number of lives can be saved annually with the HST, not only as a result of improved quality of hospitals, but also as a result of patients' sorting (due to lower travel costs) to better hospitals.

Our research contributes to the literature on hospital competition and quality in the health care industry. The empirical evidence on this topic is mixed. One of the initial studies on competition in health care markets and health outcomes is by Kessler and McClellan (2000), who examine the impact of market concentration on hospital quality in the US Medicare program. They find that higher market concentration leads to significantly higher mortality rates for heart attack patients. On the other hand, some papers find opposite results. Using similar methods to Kessler and McClellan (2000), Gowrisankaran and Town (2003) find that mortality rates are higher for Medicare heart attack and pneumonia patients that are treated in less concentrated markets. This is in contrast to the classical theoretical literature which predicts that increased competition under fixed prices results in improved quality. Gowrisankaran and Town (2003) suggest as a possible explanation for their results that a sufficiently low profit margin on Medicare patients coupled with increased competition can cause hospitals to focus on more profitable HMO patients at the expense of Medicare patients. While these papers use the predicted market share based on exogenous characteristics of the hospitals and patients to solve the endogeneity in market shares, they do not deal with a more serious problem that the number of hospitals itself may be endogenous due to entry and exits.

Other papers study changes in competition brought by health-related reforms. Propper et al. (2004) leverage on the 1991 Health Reform in the UK National Health Service, and find that the relationship between competition and AMI mortality rates is negative. Propper et al. (2008) further investigate this policy change and find that increased competition reduces waiting times, suggesting that hospitals facing more competition cut-on services that affect mortality rates (which are unobserved, in their setting, by consumers) in order to focus on other activities which are better observed by health-care buyers. Cooper et al. (2011) and Gaynor et al. (2013) exploit the 2006 English pro-competitive policy shift to study the impact of competition on quality using a difference-in-differences research design. Both papers find that increased competition improves the quality of clinical care. Leveraging on the same reform, Gaynor et al. (2016) find coronary artery bypass graft patients to become more responsive to clinical quality post-reform and hospitals to be responsive to changes in demand through quality improvements. Moscelli et al. (2021), on the other hand, find mixed results from the same reform.

Our paper adds to the prior literature by studying the effects of competition following an exogenous shock which is not related to hospital market structure or any other aspect of healthcare, thus providing a unique and novel natural-experiment to identify the impact of competition. Another paper that uses an identification strategy unrelated to aspects of the healthcare market is one by Bloom et al. (2015) in which the authors exploit the variation in hospital closures driven by the political process in the U.K. to study the impact of competition on hospital performance. While Bloom et al. (2015) use cross-sectional data for a single year, our data and setting allows us to not only leverage the cross-sectional variation in the degree of HST entry across regions, but also allows us to carry out a pre-post analysis. In addition, we explicitly model patients' choice sets to take into account changes in travel time induced by the HST. This allows us to decompose the effect of the HST along various dimensions, such as patient sorting and changes in quality of care.

The rest of this paper is structured as follows. In the next section we describe the relevant aspects of the health care industry and the entry of the high-speed train. Section 3 describes our data and section 4 describes our differences-in-differences estimation strategy and issues concerning measures of hospital quality. Section 5 describes our data and present descriptive statistics. In section 6 we present differences-in-differences regression results. Section 7 outlines the structural model of hospital choice, and section 8 presents the structural model estimates. In section 9 we measure patients' welfare changes and changes in health outcomes through a series of counterfactual exercises. Section 10 concludes.

2 Industry Details

2.1 Health Care Industry

The National Health Insurance (NHI) program in South Korea is a compulsory solo-payer public insurance system which covers the entire resident population. The social insurance system of South Korea was established in 1977, and initially covered only 8.79% of the population, but expanded to approximately 97% of the population by 1989. It operated as a multi-insurance fund system with more than 370 insurers until July 2000, when the funds were integrated to form a single-payer system. It is managed by a single insurer, the National Health Insurance Corporation (NHIC), and is supervised by the Ministry of Health, Welfare and Family Affairs (MIHWFA). The Health Insurance Review and Assessment Service (HIRA), also supervised by MIHWFA, reviews the cost and healthcare benefits and evaluates the appropriateness of health care services provided by hospitals. The system is funded by compulsory contributions from the entire resident population and government subsidies. The amount paid as NHIC contributions by an individual depends on his income and wealth; the elderly and disabled pay less.

The healthcare delivery system in South Korea is classified into three tiers: primary (clinics), secondary (hospitals and general hospitals) and tertiary care (general hospitals). Starting 1989, hospitals that met the criteria in terms of facilities, workforce, equipment, patient composition, etc, could apply to be designated as a tertiary care institution subject to demand for number of hospital beds from each health region.[7] There were 42 tertiary care institutions, and the composition of these hospitals did not change during the period of our analysis. During this period, there was little to no room for a new tertiary care entry. This is because the number of hospital beds provided by the then-tertiary care hospitals saturated the market for each health region, and tertiary hospitals were "renewed" every 3 years instead of being re-selected.

As opposed to public-sector dominant healthcare financing, healthcare delivery in South Korea is predominantly provided by the private sector: approximately 90% of hospitals are private institutions. Since the launch of the NHI program, private providers are not allowed to opt out from the program. Private health-care providers mainly supply health care services, and the fee schedule is established through annual negotiations between the NHIC and provider associations.[8] The fixed price schedule includes fees for each medical procedure, with adjustments for whether a hospital is

[7] There were 9 health regions during this period.

[8] The Korean Medical Association (KMA) and the Korean Hospital Association (KHA) are among the most important provider organizations.

a primary, secondary, or a tertiary care institution. Patients are responsible for any co-payments applicable to the medical services they receive, and the NHIC reimburses healthcare providers for the share of medical costs not borne directly by the patient on the basis of the fee schedule. Therefore, the price is exogenous to both the hospitals and patients. Fee regulation has been the subject of recurrent complaints by providers in South Korea, who claim that they are not adequately compensated for their services as a result of historically low levels of NHI fees.

Although the NHI service flow is designed to progress from primary to secondary to tertiary care, patients have the complete freedom to choose any healthcare provider at any level, with some financial incentives. To achieve an efficient distribution of limited healthcare resources, insurance coverage largely depends on the tier of the hospital. For example, the NHI insurance coverage for clinics is 70%, and it is 60% and 50% for hospitals and general hospitals, respectively. To receive treatment in tertiary hospitals, patients must be referred by primary or secondary care hospitals, in which case 40% of their bills are covered by insurance − otherwise, they can expect to pay 100% of the bill. The referral by a primary or secondary care physician is easy to obtain, so there is essentially no gatekeeping system. The insurance coverage is identical at all levels of hospitals for inpatient care, with patients being responsible for 20% of medical expenses.

2.2 Entry of the High-Speed Train

South Korea's HST system, Korea Train eXpress (KTX), began commercial operations on April 1st 2004 with the objective to alleviate (foreseeable) traffic congestion. Construction of the HST system occurred in two stages. [9] The first-stage construction involved building the Gyeongbu HST Line connecting Seoul to Daegu and electrifying the existing Gyeongbu Line connecting Daegu-Busan, as well as electrifying the existing Honam Line connecting Daejeon-Mokpo.[10] The second-stage HST system, which involved the construction of the new Gyeongbu HST line connecting Daegu to Busan replacing the existing electrified tracks, went into service in November of 2010. In this paper we only focus on the first-stage HST system. Although the launch of the second-stage HST system

[9] Note that here we are referring to the construction of the Gyeongbu HST system. The construction of additional HST systems was completed only after 2015. Additional electrified (existing) lines were added by the end of 2010.

[10] Newly constructed links included 51.6 miles of viaducts and 47.0 miles of tunnels. Electrification of the existing rail comprised of 82.5 miles across Daegu to Busan, 12.9 miles across Daejeon, and 164.3 miles from Daejeon to Mokpo and Gwangju. First stage Gyeongbu HST stations include Seoul Station, Gwangmyeong, Cheonan-Asan, Daejon, Dongdaegu stations, and the electrified Gyeongbu line connecting Dongdaegu and Busan includes Miryang, Gupo and Busan stations. Honam line includes Yongsan station, Seodaejeon, Dungyae, Nonsan, Iksan, Gimje, Jeongeub, Jangseong, Songjeongni, Gwangju, Naju, and Mokpo stations. There exists a depot for HST along the Gyeongui Line at Haengsin station. Thus some HST services continue beyond Seoul and Yongsan station and terminate at Haengsin station. For detailed information on HST services see Cho and Chung (2008).

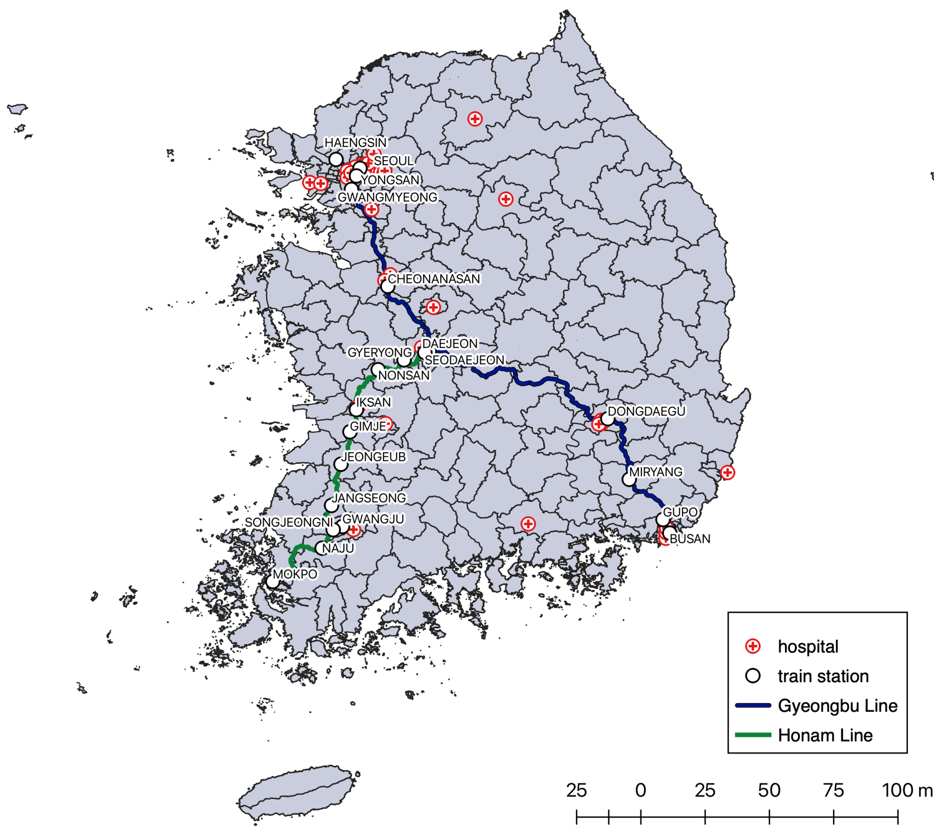

Figure 1: HST line and hospitals

enabled the HST to reach full speed through Daegu-Busan corridor, this shock was much smaller in magnitude compared to the shock generated by the first-stage HST system. Figure 1 displays two HST lines of the first-stage HST system, the Gyeongbu Line (blue) connecting Seoul-Busan and the Honam line (green) connecting Seoul-Mokpo. The figure also plots the hospitals that are included in our final sample (more on this in the next section).

At the time of its launch in 2004, the HST operated 128 times per day (94 times on Gyeongbu Line, and 34 times on the Honam Line), and the daily frequency increased to 163 in the following years, greatly reducing travel time. As an example, the HST system has reduced the travel time from Seoul to Busan from more than 5 hours by car to 2 hours and 40 minutes by train. HST fares were fixed and kept low, at approximately 55% of the corresponding air fares for the same routes, to encourage the use of the HST.[11]

3 Data

We rely on a number of data sources at a patient and hospital level. Our patient data comes from the National Health Insurance Services (NHIS) which is a health insurance claims dataset collected by the solo insurer system NHIS. Our data consists of a nationally representative random sample, which accounts for 2% of the entire South Korean population who received medical treatment at a hospital. The data contain anonymized patient-level information on medical procedures that a patient received at a hospital. Detailed information on patient demographics, diagnosis, patient's home location, the chosen hospital and the date of hospital admission are observed. In addition, if the patient died, we observe the month/year of the patient's death.[12] The geographic unit of our data (and hence patients' home location) is defined either at a city, county or at a district level depending on where a patient lives. This is because some counties are not populated enough to qualify for a city, while some smaller cities are not populated enough to be sub-divided into districts.[13] To

[11] In addition to low regular prices, various discounts (60% off the regular passes and 20% off the reserved tickets) were available to attract as many passengers as possible.

[12] We want to use 30-day mortality following a surgery as our measure of hospital quality as it is the most commonly used outcome-based measure. However, we do not observe the exact date of the surgery in our data. To complicate matters further, we only observe the year and month of patients' death instead of the exact date. Therefore our (proxy) measure of 30-day mortality rate is obtained as follows: We construct a dummy variable M whose element m_i takes value 1 if (i) patient i who was admitted to hospital in month mm_i day dd_i and year $yyyy_i$ dies either in month mm_i and year $yyyy_i$ or in month $mm_i + 1$ and year $yyyy_i$ for $mm_i = 1, ...11$ and (ii) length of hospital-stay does not exceed 30 days. If patient was admitted to hospital in $mm_i = 12$ and year $yyyy_i$, μ_i takes value 1 if patient dies in month mm_i and year $yyyy_i$ or in January of year $yyyy_i + 1$.

[13] South Korea is made up of 17 first-tier administrative divisions (province level). These are further subdivided into cities (si), counties (gun), districts (gu), towns (eup), townships (myeon), neighborhoods (dong) and villages (ri).

simplify the exposition, we will henceforth ignore the distinction between city/country/district and denote the smallest geographic unit that we observe in the data as a "district". The boundaries of each "district" are delineated in figure 1. Since patients' home location is at the district level, we use the coordinates of the centroid of each district as patients' location.

The hospital data also comes from the NHIS dataset. Hospitals in the NHIS dataset are anonymized and their location is observable only at the city-province level. To get a more precise location of the hospitals, which is essential for our analysis, we combine the NHIS dataset with that obtained from the HIRA (Health Insurance Review Assessment). Although the identity of the hospitals in the HIRA dataset is also anonymized, we are able to match this dataset to the NHIS dataset using hospital characteristics. In addition to the hospital characteristics such as number of hospital beds, number of nurses and hospital tier, the HIRA dataset contains hospital location at the district level, which in turn allows us to obtain exact coordinates for each hospital.

Our sample selection process is as follows. We define January 2003 to March 2004 as the pre-HST time period and then define January 2006 to March 2007 as the post-HST time period after allowing for some adjustment time.[14] We focus on patients who underwent a surgery at a tertiary hospital. We only consider tertiary hospitals in this paper for the following reason: Since we use 30-day mortality rates as measures of hospital quality, primary and secondary care hospitals are not suitable for the analysis because the majority of severely sick patients who are at risk of death receive treatment at tertiary hospitals. In addition, due to the fact that our data is a 2% sample, there are not enough observations per hospital for secondary care institutions. We consider all surgeries that were performed during the data period that resulted in at least xx% deaths within 30 days of admission to the hospital. Ideally we would look at patients suffering from one specific illness, or who underwent one specific type of surgery in order to minimize the contamination of hospital quality (impact on mortality rates) with patient selection.[15] Constraining our analysis to a single type of surgery, however, leaves us with too few observations (too few patients for each hospital). Limiting our attention to only one "category" of surgery (e.g., cardiovascular surgery) also leaves us with too few observations per hospital. To attenuate the contamination of hospital

Once a country attains a population of at least 150,000, it becomes a city. Cities with a population of over 500,000 are subdivided into districts. Districts are then further divided into neighborhoods (dong). Cities with a population of less than 500,000 are directly divided into neighborhoods (dong).

[14] We choose pre-HST period to start from year 2003 because patient mortality information is only available from 2003.

[15] Gowrisankaran and Town (2003) look at pneumonia patients, Kessler and McClellan (2000), Propper et al. (2004) look at acute myocardial infarction (AMI) patients, and Gaynor et al. (2016) look at patients receiving coronary artery bypass grafting (CABG) surgery.

quality from pooling patients across multiple types of surgeries, we control for the riskiness of each type of surgery in addition to patient demographics.

The key feature of our setting is that the entry of the HST enabled patients to exercise choice among alternatives with different travel distances. To take advantage of this feature, we drop the following patients who were less likely to exercise choice based on hospital location: First, patients who arrived at the hospital via ambulance because the emergency ambulance usually takes patients to a nearby hospital. Second, patients who arrived at the hospital via intra-hopital transfer as it is the physician who makes the choice of the hospital in this case. Next, we drop patients living on islands (Jeju and Ulleng Islands, as well as Shin-ahn and Ong-jin Gun) because it is difficult to calculate travel time to hospitals by car for these patients, a necessary component for estimating our demand model and performing counterfactuals. Our final sample consists 8,817 patients who went to 42 tertiary hospitals.

4 Empirical Strategy

The goal of our paper is to study the impact of increased hospital competition on hospital quality. Post-HST a hospital located closer to the HST station faces more competition than hospitals that are located further away from the HST station because the HST allows for greater substitutability between hospitals that are close to the HST. Therefore, to examine whether hospitals that are located closer to the HST experience an improvement in hospital quality after the entry of HST, we conduct our analysis using difference-in-differences (DiD) approach by exploiting the variation in distance from each hospital to the nearest train station. We identify the impact of competition from the interaction of a continuous treatment intensity variable (hospital's distance to the nearest HST station) with a dummy indicator for the post-HST period. This specification was employed by Gaynor et al. (2013) to study the impact of hospital competition.[16] Specifically, the DiD regression specification is given by

$$outcome_{jt} = b_{0j} + b_1 I(t=1) + b_2 I(t=1) \times dist_j^h + \varepsilon_{jt}. \tag{1}$$

We collapse time periods into pre- and post-HST periods so that $t=0$ denotes pre-HST and $t=1$ denotes post-HST. The variable $outcome_{jt}$ measures the quality of clinical care at hospital j in

[16] See Card (1992) and Acemoglu et al. (2004) for more about continuous treatment.

period t. As mentioned earlier, we use hospital-level 30-day mortality rates as the outcome variable after adjusting for patient selection; b_{0j} denotes a full set of hospital dummies; $I(\cdot)$ is an indicator function for the post-HST period which takes the value 1 for the post-HST period and 0 otherwise, and ε_{jt} is a random noise. The DiD coefficient of interest is b_2, which corresponds to the interaction term between a post-HST dummy and the distance from hospital j to the nearest train station, denoted as $dist_j^h$. This coefficient measures the change in the effect of distance to the nearest train station pre- and post-HST. If the outcome variable is hospital-level death rate, a positive value of b_2 implies that death rate is lower as hospitals are located closer to the HST station in the post-HST period. The identifying assumption is that without the entry of the HST, the trend in mortality rates would have been the same regardless of the distance to the train station. The entry of the HST induces a deviation from this parallel trend. We provide evidence supporting this assumption in Section 5.

4.1 Measure of Hospital Quality

Using raw mortality rates as a measure of quality is problematic due to patient selection bias: hospital selection is non-random. The existing literature address this issue by obtaining hospital quality measures that take into account this selection bias. Several papers control for hospital selection using a model in which distance between the patient's residence and alternative hospitals are used as exogenous variables (e.g., Gowrisankaran and Town 1999, Gowrisankaran and Town 2003, Kessler and McClellan 2000, Geweke et al. 2003, Tay 2003). For instance, Gowrisankaran and Town (1999) model mortality as a function of hospital choice dummy variables and patient characteristics, and apply linear instrumental variables approach using distance from each patient to *all* alternative hospitals.[17] The identifying assumption here is that where a patient chooses to live from alternative hospitals is uncorrelated to patient's severity of illness – an assumption that has been commonly used in empirical models of hospital choice, e.g. Capps et al. (2003), Gaynor and Vogt (2003), Ho (2009), Beckert et al. (2012). Geweke et al. (2003) develop a method to infer hospital quality using a binary probit model of mortality accompanied by a multinomial probit model of hospital choice. Here again, the distance is an exogenous variable that influences patient's hospital choice: the farther away a patient lives from a hospital, the less likely the patient is to choose that hospital, other things equal.

[17] While this approach is simple, there is no formal statistical model that rationalizes this approach, as discussed in Geweke et al. (2003).

In our setting, the HST may facilitate long-distance travel for certain type of patients, and hence the degree of patient selection may change as a result of the entry of the HST. For example, if severely ill patients take the HST to go to better hospitals, the degree of patient selection will be aggravated. To allow for this change in the degree of patient selection resulting from the reduction in travel time, we take into account changes in travel *time* between pre- and post-HST periods. We follow the approach developed by Geweke et al. (2003) which we describe below, but use travel *time* rather than travel distance from each patient to alternative hospitals as an instrument. Specifically, we define travel time for patient i to hospital j in period t as

$$\text{traveltime}_{ijt} = \begin{cases} \min(\text{cartime}_{ij}, \text{traintime}_{ij}) & \text{if } t = \text{post-HST } \& \text{ dist}_i^{pat} < 30 \;\&\; \text{dist}_j^h < 30 \\ \text{cartime}_{ij} & \text{otherwise,} \end{cases} \tag{2}$$

where cartime_{ij} denotes the drive time from patient i's location to hospital j by car and traintime_{ij} is the travel time from patient i's location to hospital j by HST. Driving times by car are obtained using the *georoute* routine developed by Weber and Péclat (2017) which calculates the driving time between two points under normal traffic conditions. Note that traintime_{ij} is obtained by summing the following three components: (i) drive time from i's location to i's nearest HST station h, (ii) travel time from station h to station k, which is the closest HST station to hospital j and (iii) drive time from station k to hospital j. The variables $dist_i^{pat}$ and $dist_j^h$ are, as described earlier, travel time from patient i to the closest train station and travel time from hospital j to the closest train station, respectively. While the effect of HST does not have physical boundaries, we nevertheless constrain the effect of the HST to patients and hospitals that are located within 30 minutes of the train station. This is to account for changes in travel time only for patients that live (and visit hospitals) sufficiently close to the HST station, and is reflective of the data which reveal that there are no significant differences in travel times between pre- and post- HST for patients living beyond 30 minutes of the HST station. Having explained our instrument, we next proceed to briefly explain the model developed by Geweke et al. (2003) that will be used to obtain selection-corrected mortality rates.

Model of Patient Mortality and Hospital Choice

Following Geweke et al. (2003), we obtain adjusted mortality rates by estimating a structural probit equation in which the death probability is a function of hospital choice and patient's observed characteristics. To account for patient selection problem, the mortality probit model is accompanied by a multinomial probit model of hospital choice in which the travel time affects choice.

Specifically, the mortality probit equation for patient i $(i = 1, ..., N)$ is given by

$$m_i^* = c_i'\beta + x_i'\gamma + \varepsilon_i, \tag{3}$$

where m_i^* is the latent outcome variable for the observed mortality indicator m_i that equals 1 if the patient dies 30 days following the admission to the hospital and 0 otherwise, c_i is a $J \times 1$ dimensional vector whose element c_{ij} equals 1 if i chooses hospital j and 0 otherwise, x_i is a $k \times 1$ vector of observed patient characteristics that can affect mortality, and ε_i is an independent and normally distributed error term with mean 0 and variance σ^2.[18] The probability of patient i's death from choosing hospital j can be written as $P(m_i = 1) = \Phi((\beta_j + x_i'\gamma)/\sigma)$. Here, the parameters β and σ are jointly unidentified. While conventional probit models resolve this identification issue by setting $\sigma = 1$, this creates an additional problem in the current setting: hospital choice c_i is likely to be correlated in part with ε_i because sicker patients may prefer better hospitals, which will lead to biased estimates of β. To resolve this selection issue, we supplement the mortality equation with a multinomial model of hospital choice as described below.

For the model of hospital choice, let $\tilde{Z}_i = [\tilde{z}_{i1}, \tilde{z}_{i2}, ..., \tilde{z}_{iJ-1} - \tilde{z}_{iJ}]'$ be the $J \times q$ dimensional matrix of characteristics specific to the combination of patient i and hospital j, such as travel time from patient's home to the hospital, and serves as an instrument. Specifically, we include travel time from the patient's home to each hospital (in hundreds of minutes), the square of travel time, and the product of travel time and patient's characteristics. The $J \times 1$ dimensional hospital choice latent vector $\tilde{c}_i^*$ is given by

$$\tilde{c}_i^* = \tilde{Z}_i\alpha + \tilde{\eta}_i, \tag{4}$$

where the observed hospital choice vector c_{ij} takes value 1 if $\tilde{c_{ij}^*} \geq \tilde{c_{ik}^*} (k = 1, ..., J)$ and 0 otherwise, and $\tilde{\eta}_i \sim N(0, \tilde{\Sigma})$ is a vector of independent error terms. Note that as in the mortality probability, the parameters α and $\tilde{\Sigma}$ are jointly unidentified.

[18] For patient characteristics x_i we include an indicator for gender, indicator for low income, age indicators (10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79 and 80 or older), riskiness of the surgery that the patient undergoes as well as the riskiness of the patients' diagnosed disease, and indicators for each year-quarter. We construct the riskiness of each surgery as the death rate of that specific surgery across all patients in our sample. Similarly, the riskiness of diagnosis is constructed as the death rate of that specific diagnosis across all patients in the sample.

Normalizing the utility of the J-th alternative to 0, we get

$$c_i^* = Z_i\alpha + \eta_i, \tag{5}$$

where $Z_i = [\tilde{z}_{i1} - \tilde{z}_{iJ}, \tilde{z}_{i2} - \tilde{z}_{iJ}, ..., \tilde{z}_{iJ-1} - \tilde{z}_{iJ}]'$, $c_i^* = [\tilde{c}_{i1}^* - \tilde{c}_{iJ}^*, ..., \tilde{c}_{iJ-1}^* - \tilde{c}_{iJ}^*]'$, and $\eta_i = [\tilde{\eta}_{i1} - \tilde{\eta}_{iJ}, ..., \tilde{\eta}_{iJ-1} - \tilde{\eta}_{iJ}]'$ with $\eta_i \sim N(0, \Sigma)$.

Because hospital choice is likely to be correlated with patients' unobserved severity of illness, the mortality error term ε_i and choice error term η_i can be correlated. The mortality error term can be parameterized as

$$\varepsilon_i = \eta_i'\delta + \zeta_i; \qquad cov(\eta_i, \zeta_i) = 0, \tag{6}$$

where the scale of ε_i is normalized by $var(\zeta_i) = 1$ and δ is a $(J-1) \times 1$ parameter vector. The variance of ε_i of the mortality probit equation then becomes $\sigma^2 = \delta'\Sigma\delta + 1$. This parameterization resolves the identification problem in equation (3), and the probability of mortality of patient with observed characteristics x_i under hypothetical random assignment to hospital j becomes $P(m_i = 1) = \Phi[(\beta_j + x_i'\gamma)/(\delta'\Sigma\delta+1)^{1/2}]$. Denote $q_j = \beta_j/(\delta'\Sigma\delta+1)^{1/2}$ as the *hospital j quality probit*. Note that, in the conventional probit model where σ is normalized to 1, quality probit becomes $q_j = \beta_j$.

For the remainder of this paper, we use quality probits (from the selection model) as our measure of hospital quality. Because lower quality probit implies lower death rate and in turn higher hospital quality, for ease of exposition we henceforth interchangeably use the term "risk-adjusted mortality rate" to denote the quality probit.

The model is estimated using Bayesian inference with the GGT computer program as described in Marquardt et al. (2021).[19] Using the Markov Chain Monte Carlo method, the GGT program iteratively simulates latent variable values conditional on data and parameters, and parameters conditional on data and latent variables, to simultaneously recover the joint posterior distribution of parameters and latent variables. Because the model is estimated using Bayesian inference, the estimation of the model depends on the prior distributions. For prior distributions we use the default setting of the GGT program, but also show that our results are robust to alternative specifications of prior distributions. More details on the model, prior distributions and the estimation methodology can be found in Geweke et al. (2003) and Marquardt et al. (2021).

[19] Evaluating one parameter vector for one patient through the Maximum likelihood would require evaluating the joint density of the mortality and hospital choice outcomes for that patient. Given the number of endogenous variables and the correlation between the error terms in the mortality and hospital choice equations, evaluating the likelihood for just one parameter vector would be extremely computationally burdensome.

	pre-HST					post-HST				
	mean	median	sd	min	max	mean	median	sd	min	max
female	0.401	0	0.490	0	1	0.402	0	0.490	0	1
age 0-24	0.430	0	0.495	0	1	0.437	0	0.496	0	1
age 25-49	0.269	0	0.444	0	1	0.318	0	0.466	0	1
age 50-74	0.145	0	0.352	0	1	0.115	0	0.319	0	1
age 75 +	0.156	0	0.363	0	1	0.131	0	0.337	0	1
Seoul resident	0.308	0	0.462	0	1	0.303	0	0.460	0	1
low income	0.182	0	0.385	0	1	0.173	0	0.378	0	1
medium income	0.330	0	0.470	0	1	0.340	0	0.474	0	1
high income	0.487	0	0.499	0	1	0.487	0	0.500	0	1
diagnosis risk	0.045	0.024	0.055	0	0.75	0.050	0.031	0.058	0	0.75
surgery risk	0.053	0.017	0.088	0.003	0.625	0.052	0.017	0.088	0.003	0.625
death	0.044	0	0.205	0	1	0.051	0	0.219	0	1
Observations			4,055					4,762		

Notes: Most of the patient characteristics are binary variables, and therefore the mean represents the fraction. "Seoul resident" is a binary variable that equals 1 if a patient lives in Seoul and 0 otherwise. There are 11 (group 0 - group 10) income groups in our data. We classify groups 0-3 as low income, groups 4-7 as medium income, and groups 8-10 as high income. "death" is binary variable that equals 1 if patient dies within 30 days of admission to the hospital, and 0 otherwise.

Table 1: Patient Characteristics

5 Descriptive Statistics

We first proceed by providing descriptive evidence on patients' response to the entry of the HST with respect to their travel patterns. Then, we will provide some hospital-level summary statistics.

5.1 Patients' Response to the Entry of the HST

Table 1 provides summary statistics of patient characteristics. We first show that patients' travel patterns changed following the entry of the HST. If patients indeed responded to the entry of the HST, we expect patients living closer to the HST stations to choose hospitals that are located further away from their home. Figure 2 plots percent changes in average travel distance by district, following the entry of the HST, separately for patients who live in Seoul and patients who live in non-Seoul regions. Since patients in South Korea generally have a preference for hospitals that are located in Seoul, we expect the HST to have minimal effect on travel patterns of patients who already live in Seoul. From the plot on the left, proximity to the HST station doesn't seem to affect patients' travel distance. From the plot on the right, however, regions that are located very close to the HST station experienced a large increase in average travel distance following the entry of the

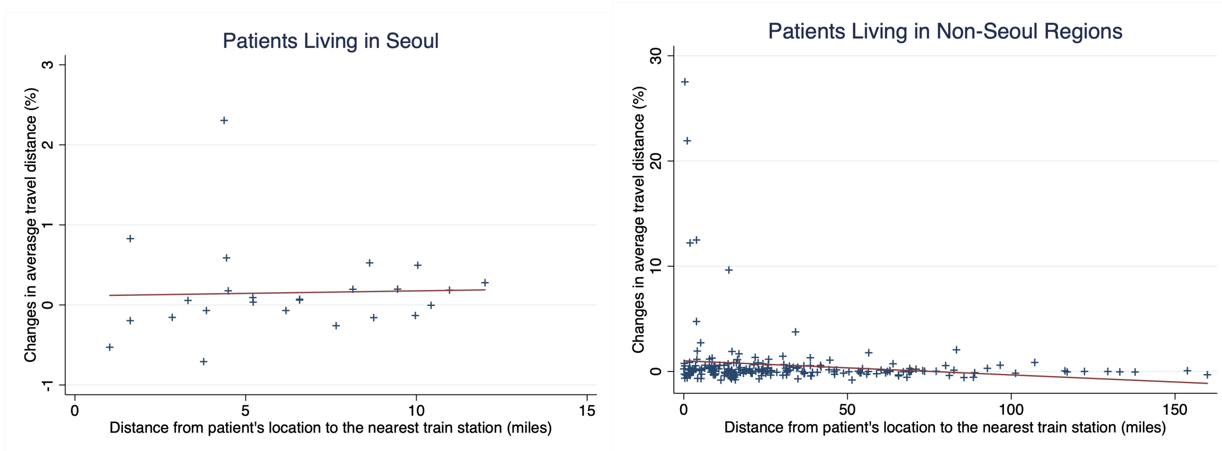

Figure 2: Proximity to the train station and changes in average travel distance by district

HST.

To get an estimate of the effect of the HST on patients' travel distance, we leverage the variation in distance from each patient to the nearest train station to examine whether patients that live closer to the train station travel longer distances after the entry of the HST. Specifically, we estimate the equation as below:

$$traveldist_{it} = a_0 + a_1 I(t=1) + a_2 I(t=1) \times dist_i^{pat} + a_3 X_{it} + \mu_i + \varepsilon_{it} \tag{7}$$

Here the dependent variable, $traveldist_{it}$, is the travel distance of patient i in period t; $I(\cdot)$ is an indicator function for the post-HST period, which takes the value 1 for the post-HST period and 0 otherwise; X_{it} denotes patient characteristics (age, gender, diagnosis type and surgery type dummy variables), μ_i denotes a full set of district dummy variables of where the patient's home is located in, and ε_{jt} is a random noise. The coefficient of interest is a_2, which corresponds to the interaction term between the post-HST dummy variable and $dist_i^{pat}$, which is the distance (in miles) from patient i's home to the nearest HST station. This coefficient measures whether patients who live closer to the train station traveled further distances following the entry of the HST.

In Table 2 column 1 we report OLS regression estimates for equation (7) using all the patients in our final sample. The result suggests that there was a marginally significant increase in distance traveled for patients that live closer to the train station after the entry of the HST - while the positive coefficient on the post dummy variable suggests that patients on average traveled 3.1 miles more following the entry of the HST, this effect decreases as patients are located further away from

	all patients	excluding Seoul residents	ambulance and transfer
	(1)	(2)	(3)
post	3.0781***	5.1842***	-1.5582
	(0.8639)	(1.3205)	(3.1554)
$post \times dist_i^{pat}$	-0.0594*	-0.1038***	0.0109
	(0.0316)	(0.0351)	(0.1147)
district FE	✓	✓	✓
surgery type FE	✓	✓	✓
diagnosis type FE	✓	✓	✓
R^2	0.4451	0.4196	0.5755
Observations	8,817	6,124	2,316

Notes: Models are estimated using OLS with standard errors (in parentheses under coefficients). Standard error clustered at the district level. In addition to diagnosis and surgery type, all regressions control for patient characteristics (age and gender).

*** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

Table 2: OLS Estimates of Patient's Distance on HST on Travel Distance

the train station (a 1 mile increase in distance from patient's home to the closest train station is associated with a 0.06 miles decrease in travel distance.

We also estimate equation (7) only using patients that live in non-Seoul regions and report the results in column 2. For these patients, the closer they live to the HST station, we find a significant increase in distance traveled. The positive coefficient on the post dummy variable suggests that patients living in non-Seoul regions traveled 5.2 miles more following the entry of the HST, but this effect significantly decreases as patients live further away from the train station. A one mile increase in distance from patient's home to the closest train station is associated with a 0.1 miles decrease in travel distance.

As mentioned earlier, our final sample excludes patients who transferred from other hospitals and who arrived at a hospital via ambulance because these patients are less likely (if any) to exercise choice. If the increase in travel distance is a consequence of the entry of the HST, we should not see changes in travel distance for patients who arrived at hospitals via transfer or ambulance because these patients did not take the HST. Table 2 column 3 reports the regression estimates for equation (7) using only those patients who transferred or took an ambulance (note that these patients are not included our final sample and are not used in further analysis). In this case, we find that HST has no effect on travel distance, suggesting that HST only affected patients who exercised their choice.

It would be difficult to attribute increased travel distance to HST if a patient traveled longer distance

	treated hospitals			control hospitals		
	pre-HST	post-HST	z-stat	pre-HST	post-HST	z-stat
control patients	0.124	0.135	0.943	0.107	0.121	1.306
	(0.008)	(0.008)		(0.007)	(0.007)	
Observations	1,719	1,971		1,719	1,971	
treated patients	0.060	0.084	2.281	0.076	0.088	1.078
	(0.007)	(0.004)		(0.008)	(0.008)	
Observations	1,088	1,346		1,088	1,346	

Notes: This table shows the changes in proportion of patients (excluding Seoul residents) who traveled more than 50 miles to arrive at the hospitals. z-statistic for test of proportions. Standard error in parentheses.

Table 3: Proportion of Patients who Traveled to arrive at Hospitals

to arrive at a hospital which is located far away from the HST station. To provide further evidence of the effect of the HST on patients' travel, we next show that patients who live closer to train station traveled long distances only to visit hospitals that are *also* located close to the HST station. To facilitate this analysis, we first define "treated" as being located within 10 miles of the HST station and define "control" as being located beyond 10 miles from the HST station. This allows us to categorize patients into two groups: "treated patients" (patients who live within 10 miles of the HST station), and "control patients" (patients who live beyond 10 miles of the HST station). Similarly, we can categorize hospitals into two groups: "treated hospitals" (hospitals that are located within 10 miles of the HST station), and "control hospitals" (hospitals that are located beyond 10 miles of the HST station). For each group of patient who went to each group of hospitals, we then calculate the proportion of patients who traveled more than 50 mile to arrive at each type of hospital. The results reported in Table 3 show that there was a significant increase in proportion of treated patients who traveled more than 50 miles to arrive at treated hospitals (from 6 percent to 8.4 percent). We do not see significant difference in travel patterns for treated patients going to control hospitals. Likewise, we do not see any significant changes in travel patterns for control patients. These patterns suggest that patients didn't simply travel longer distances by driving longer hours, but instead provide some evidence that they took the HST to go to a hospital that is also closely located to the HST station.

Table 4 shows changes in patient care seeking post entry of HST by hospital quality. For our measure of quality we use the adjusted mortality rates obtained through the probit selection model in the previous section. If patients use the HST to sort to better hospitals, we should see better hospitals attracting more patients relative to worse hospitals. We define better hospitals as those in the bottom quartile of the adjusted mortality distribution and define worse hospitals as those

	Bottom quartile			Top quartile		
Share of patients	pre-HST	post-HST	z-stat	pre-HST	post-HST	z-stat
Traveled less than 50 mi.	0.2540	0.2396	-1.5645	0.1896	0.2148	2.9287
	(0.0068)	(0.0062)		(0.0062)	(0.0055)	
Traveled more than 50 mi.	0.0284	0.0569	6.5299	0.0210	0.0176	-1.1339
	(0.0026)	(0.0033)		(0.0022)	(0.0019)	
Non-Seoul patients who traveled more than 50 mi.	0.0259	0.0559	6.9848	0.0200	0.0166	-1.1870
	(0.0025)	(0.0033)		(0.0022)	(0.0019)	
Traveled more than 50 mi. & live within 10 mi. of HST	0.0042	0.0195	6.4665	0.0042	0.0055	0.8515
	(0.0010)	(0.0020)		(0.0010)	(0.0011)	
Traveled more than 50 mi. & who live beyond 50 mi. away from HST	0.0133	0.0162	1.1035	0.0079	0.0061	-1.0172
	(0.0018)	(0.0018)		(0.0014)	(0.0011)	
Observations	4,055	4,762		4,055	4,762	

Notes: This table shows the changes in proportion of patients who visited hospitals in top and bottom quartile of mortality distribution. z-statistic for test of proportions. Standard error in parentheses.

Table 4: Changes in Patient Care Seeking by Hospital Mortality

in the top quartile. While better hospitals experienced a significant increase in share of patients who traveled more than 50 miles to arrive at the hospital, there is no significant change for worse hospitals. This pattern continues to hold when we only look at non-Seoul residents who travel more than 50 miles, as well as share of patients who live within 10 miles of the HST station and travel more than 50 miles. For both better and worse hospitals, there are no changes in share of patients who live beyond 50 miles from the HST station and traveled more than 50 miles. For these patients, the entry the HST is likely to have small (if any) effect, and hence it doesn't change their sorting patterns.

5.2 Hospital Characteristics

Table 5 provides summary statistics of hospital characteristics. Figure 3 presents the relationship over the entire period of our analysis (including the adjustment period) between distance to the nearest train station and raw 30-day mortality rates. Due to data limitations, we do not have patient death information prior to 2003. Since the HST entered in April 2004, it is difficult to see (if any) pre-HST trends of mortality rates at the annual level. Therefore, for this analysis, we calculate 30-day mortality rates at the quarter level. Three time series lines are presented for the mean of the mortality rates, one for patients who visited hospitals in each quantile of the hospital's distance to the nearest train station.[20]The series is rather noisy, but we can see that all three series fluctuate together. Until the first quarter of 2004, mortality rates for all three quantiles display declining trends.

6 Difference-in-Differences Estimation Results

We use the measures of clinical quality obtained through the probit selection model to study the impact of increased hospital competition on hospital quality. As a starting point to this analysis, we first estimate equation (1) using hospital-level *raw* mortality rates as the outcome variable. Since hospitals located near the HST station are the ones that are most affected by the entry of the HST, the DiD coefficient on $d_{post} \cdot dist_j^h$ captures the impact of increased hospital competition. A positive value of the DiD coefficient implies that death rate is lower as hospitals are located closer to the HST station in the post-HST period. Column (1) in Table 6 reports the results. While marginally

[20]Since there are not enough patients at each hospital for each quarter, we calculate mortality rates at the patient level for each quantile of hospital's distance to the nearest train station.

	pre-HST					post-HST				
	mean	median	sd	min	max	mean	median	sd	min	max
number of admissions	96.5	81.5	75.6	14	456	113.4	93.5	93.3	11	535
number of beds	1,101	1,019	491.3	480	2,993	1,101	1,019	491.3	480	2,993
number of nurses	479.5	421	267.7	224	1,671	501.6	426	278.5	224	1,671
located in Seoul	0.45	0	0.5	0	1	0.45	0	0.5	0	1
mortality rate	0.045	0.038	0.027	0	0.143	0.053	0.052	0.038	0	0.238
Observations			42					42		

Notes: Variable "located in Seoul" is a binary variable that equals 1 if a hospital is located in Seoul and 0 otherwise. The mean of "located in Seoul" is a fraction hospitals that are located in Seoul.

Table 5: Hospital Characteristics

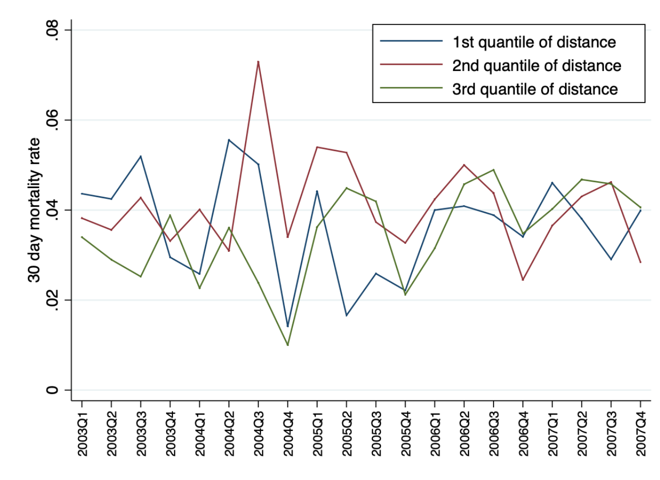

Figure 3: Trend of mortality rates (2003-2007)

significant, the DiD coefficient is positive: when the travel time from the hospital to its closest train station decreases by 1 minute (i.e., the hospital is closer to the train station), the hospital-level raw mortality rate decreases by 0.048 percentage points.

As discussed earlier, however, hospital-level raw mortality rates do not correctly reflect the true quality of clinical care due to differences in patients' health status across hospitals (referred to as hospital's "case-mix") i.e., hospitals with a larger number of sicker patients are more likely to have higher mortality rates. It is therefore necessary to take into account differences in patient case-mix across hospitals, for both observed and unobserved patient characteristics. We therefore use the adjusted mortality rates that we obtained through the selection probit model as the outcome variable to estimate equation (1). This measure of hospital quality was obtained through the model that incorporates patient's hospital choice, and therefore resolves the patient selection issue. The results are reported in Table 6, column (2). The DiD coefficient is positive and significant ($\beta = 0.006$), suggesting that more competition leads to improved hospital quality. Since adjusted morality rates are coefficients of the probit mortality equation (equation (3)), the magnitude of the DiD coefficient is difficult to interpret. Therefore, to interpret the magnitude of the DiD coefficient and results, we calculate the probability of death for an average (in terms of case-mix) patient for each hospital, and use the log of this measure as the dependent variable to estimate equation (1). Specifically, we estimate the equation

$$\log\left(\Phi[(\beta_j + \bar{x}'\gamma)/(\delta'\Sigma\delta + 1)^{1/2}]\right) = b_{0j} + b_1 I(t = post) + b_2 I(t = post) \times dist^h_j + \varepsilon_{jt},$$

	(1) raw mortality	(2) selection corrected	(3) selection uncorrected	(4) ambulance and transfer
$post$	0.0027	-0.0113	-0.0181	0.0937*
	(0.0066)	(0.0509)	(0.0346)	(0.0508)
$post \cdot dist_j^h$	0.0005*	0.0061**	0.0050***	0.0058**
	(0.0003)	(0.0028)	(0.0012)	(0.0026)
R^2	0.7709	0.5909	0.6271	0.6187
hospital FE	✓	✓	✓	✓
number hospitals	42	42	42	42
Observations	84	84	84	84

Notes: Models are estimated using OLS with standard errors (in parentheses under coefficients). All regressions include constants.

*** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

Table 6: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care

where $\bar{x} = \sum_i x_i/N$ denotes the average case-mix across all patients. The results from this regression suggests that a 1 minute reduction in travel time from the hospital to the nearest train station decreases the probability of death for an average patient by 1.98 percent (significant at the 5 percent level).

To see how accounting for patient selection affects the results, we also estimate equation (1) using adjusted mortality rates obtained through the conventional probit model that ignores the selection issue. As mentioned above, in the conventional probit model, σ is normalized to 1 and the quality probit is $q_j = \beta_j$. The results are reported in Table 6, column (3). While the DiD coefficient is positive and significant, which is consistent with our previous results, the magnitude of the DiD coefficient is smaller compared to the case where patient selection is accounted for (0.005 versus 0.0061). This suggests that ignoring selection may lead to misleading inferences about hospital quality.

Recall that our final sample excludes patients who transferred from other hospital and who arrived at a hospital via ambulance and transfer. Including these patients in our sample should not change our results because the quality of clinical care should be independent from how patients arrive to a hospital. To test the robustness of our results, we estimate equation (1) including transfer and ambulance patients into our sample. The results are reported in Table 6, column (4). The results from this analysis are consistent with the previous results, and the DiD estimates are similar to those in column (2) where we use the quality probits from the selection model.

To show the robustness of our findings to the specification of the prior distribution, we estimate the selection model in Section 4.1 using the following two variants of prior distribution specifications (this is the same approach used by Geweke et al. (2003) to test the sensitivity of their estimates to the specifications of the prior distribution): Variant 1 scales the prior standard deviations of α in the original selection model downward by a factor of 5 and τ^2 downward by a factor of 25. Variant 2 is the opposite of variant 1 in that the prior standard deviations are scaled up by a factor of 5 relative to the base model. We then use quality probits from these two specifications to estimate equation (1). Appendix B provides the results for each of these prior distributions. The results are consistent with the findings in the current section.

To test the robustness of our findings to alternative measures of hospital quality, we employ the linear instrumental variables approach as in Gowrisankaran and Town (1999) and Gowrisankaran and Town (2003). Results are reported in Appendix A, and are consistent to what we find in this section.

The results in in this section suggest that increased competition leads to an improvement in hospital clinical quality. To evaluate the impact on patient welfare, we next estimate a demand model of hospital choice and use the model estimates to perform welfare analysis and various counterfactuals.

7 Model of Hospital Choice

To evaluate the impact of the HST on patient welfare we need to consider the hospital choice that patients would have made had the HST not been launched. To do this, we estimate a structural model of hospital choice, and conduct a reverse counterfactual analysis by switching off the impact of the HST. The entry of the HST reduces travel time and thereby increases the number of hospitals in the choice set of patients living close to a HST station. To capture the changes in patients' choice sets in our model, we extend the traditional conditional logit model by imposing travel-time constraints on patients, an approach that has been used in the geography and transportation literatures. We assume that the travel time to each hospital determines whether that hospital is included in a patient's choice set or not. If a hospital is located too far from a patient's location, a patient with a travel-time constraint will exclude it from his choice set. This translates to a decrease in the size of the choice set for patients living close to a HST station once the HST is removed.

7.1 Utility and Demand

Each patient i chooses from $J_i \subseteq J$ hospitals in his choice set, indexed $j = 1, ...J_i$ where J is the total number of hospitals in the data. The indirect utility of patient i choosing hospital j, $j = 1, ..., J$ is defined as

$$u_{ij} = \sum_{l=1}^{L} X_{j,k}\mathbf{Y}_i'\beta^{xy}_{\cdot,l} + Q_j\mathbf{Y}_i'\alpha^z + D_{ij} + \mathbf{X}_j'\beta^x + \alpha Q_j + \varepsilon_{ij}, \tag{8}$$

where $\mathbf{X}_j$ is a vector of hospital characteristics with length L; $\mathbf{Y}_i$ is a K vector of patient-specific demographics; D_{ij} is the travel time from patient i's home to hospital j; Q_j denotes the quality of clinical care at hospital j; ε_{ij} is an idiosyncratic taste shock that is distributed i.i.d. type I extreme value. β^{xy}, α^z and β^x are $K \times L$, $K \times 1$, and $L \times 1$ matrices of coefficients, respectively. Following previous literature on hospital choice, we assume that all patients are admitted to some hospital, and hence there is no outside option in our model.

We estimate the parameters in equation (8) using a maximum likelihood approach. One might be concerned about the endogeneity of quality of clinical care in the utility function. Previous literature has found that treating a larger number of cases is associated with better outcomes. Hospitals with higher unobserved quality will attract larger volume of patients, and this will in turn lead to higher quality of clinical care.[21] To address this concern, following Gaynor et al. (2016), we include an entire set of hospital indicator variables to estimate hospital fixed effects.

7.2 Choice Set Formation

The entry of the HST enlarged patients' consideration sets by reducing travel costs. Hospitals that would not previously have been considered by the patient may now be considered. We model this change in consideration sets by imposing a travel-time constraint on patients. We assume that time is a limited resource that constrains choice options from being evaluated. This assumption is consistent with theoretical and empirical literature in geography and regional science where a relationship between the available time budget and individuals' destination choice has been established. Our modeling approach follows the Approximate Nested Choice-Set Destination Choice (ANCS-DC) model developed by Thill and Horowitz (1997) which explicitly models the formation of choice sets when individuals have limited time resources.

Each patient has a travel-time threshold T_i which confines his choice set. We let T_i to be a random

[21] For more literature on volume-quality relationship, see Birkmeyer et al. (2002), Silber et al. (2010), and Halm et al. (2002).

variable with cumulative distribution $P_T(t;\theta)$, where parameterization by θ allows $P_T(t;\theta)$ to depend on observable patient characteristics. Then, the unconditional probability of patient i choosing hospital j is given as

$$Pr(y_{ij}=1)=\int_{t=0}^{\infty} Pr(y_{ij}=1|J_{it})dP_T(t;\theta) \tag{9}$$

where J_{it} is a choice set of individual i who has a travel-time threshold t and $Pr(y_{ij}=1|J_{it})$ is the probability of choosing hospital j conditional on facing choice set J_{it}. Since hospitals are discrete and mutually exclusive alternatives, hospitals can be sorted according to their travel time from a patient's location in ascending order. Then, equation 9 can be simplified to a summation over all the nested sets of hospitals defined by incremental travel-time thresholds, given as

$$Pr(y_{ij}=1)=\sum_{r=1}^{J} Pr(y_{ij}=1|J_i r)p_T(r;\theta), \tag{10}$$

where $p_T(r;\theta)$ is the probability that travel time threshold is between travel times to destinations r and $r+1$, i.e.,

$$p_T(r;\theta)=P_T(t_{r+1};\theta)-P_T(t_r;\theta). \tag{11}$$

The appealing feature of this modeling approach is that it enables us to avoid considering all subset combinations of hospitals which would result in 2^{J-1} choice sets for each patient. The number of possible choice sets is substantially reduced by exploiting the non-random ordering of hospitals based on their travel time from patients' location and travel-time constraints. Therefore, all hospitals that are located closer than any hospital that satisfies the inclusion criterion set by the travel-time threshold are also included in the choice set, and all hospitals that are located further than any hospital that does not satisfy the inclusion criterion are excluded. Despite this simplification, the computational complexity still remains due to the number of hospitals in our data.

To further reduce the computational burden, we reduce the support of p_T by restricting the entire series of travel-time thresholds to take only a few discrete values. Specifically, let $T_{r'}$ denote the travel-time threshold with $r^{'}=1,...,R_T$, where R_T is the number of possible travel-time thresholds after the number of discrete thresholds has been approximated to a few manageable points. We denote the probability that patient i's threshold is $T_{r'}$ as $\pi_{i,r'}$. Let $\pi_{i,r'}$ be a function of concomitant (demographic) variables, defined as

$$\pi_{i,r'}=\frac{\exp(\gamma_r+\mathbf{Y}_i^{'}\phi_{r'})}{\sum_l^{R_T}\exp(\gamma_l+\mathbf{Y}_i^{'}\phi_{r'})}, \tag{12}$$

	(1) ANCS-DC		(2) Multinomial Logit	
	Coefficient	Standard error	Coefficient	Standard error
TravelTime	-1.7097***	0.0235	-3.2683***	0.0140
AdjustedMortality	-0.3241**	0.1367	-0.2499***	0.0245
NursePerBed	7.6387***	0.2799	6.1856***	0.1511
AdjustedMortality×Female	-0.1385	0.1629	-1.0186***	0.0634
AdjustedMortality×Old	-1.6288***	0.4288	-0.9847***	0.2619
AdjustedMortality×LowIncome	-0.0105	0.1937	0.1205	0.1219
AdjustedMortality×HighRiskSurgery	-0.0619***	0.0188	-0.0295	0.0308
NursePerBed×Female	-0.7758***	0.1182	-0.7564***	0.1206
NursePerBed×Old	2.6883***	0.2283	2.4353***	0.1290
NursePerBed×LowIncome	-0.2806***	0.0527	-0.4650***	0.0714
NursePerBed×HighRiskSurgery	-0.0003	0.0381	0.0808***	0.0320
Log Likelihood	-22,257 (38.2370)		-22,806 (3.3610)	

Notes: *** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

To account for the standard errors of the mortality rates, we employ bootstrapping, and report the means and standard deviations of the parameter estimates across the bootstrap replications. We also report the mean of the log-likelihood across all bootstrap replications and the standard deviation in parentheses.

Table 8: Demand Model Estimates

where $\mathbf{Y}_i$ is a $K \times 1$ vector of patient demographics (Gupta and Chintagunta 1994). Then the probability that hospital j is chosen is

$$Pr(y_{ij} = 1) = \sum_{r'=1}^{R_T} Pr(y_{ij} = 1 | J_{ir'}) \pi_{i,r'}, \tag{13}$$

where $J_{ir'}$ is the set of all hospitals h such that $D_{ih} \leq T_{r'}$. The model is estimated by maximizing the following log likelihood function:

$$LL = \sum_{i=1}^{N} \sum_{j=1}^{J} y_{ij} \log \left(\sum_{r'=1}^{R_T} Pr(y_{ij} = 1 | J_{ir'}) \pi_{i,r'} \right). \tag{14}$$

8 Demand Model Estimation Results

We estimate the conditional logit model of hospital choice under travel-time constraints (ANCS-DC). The covariates that enter the utility function are as follows. AdjustedMortality is the hospital quality probit from the selection model that we obtained in Section 4.1. TravelTime refers to travel time between the patient and a hospital in his choice set, and is defined in units of 100 minutes.

NursePerBed refers to ratio of number of nurses to number of hospital beds. We also interact the following patient characteristics with AdjustedMortality and NursePerBed variables: Female indicator variable; Old is a dummy variable that equals 1 if a patient is above 75 years of age and 0 otherwise; LowIncome is a dummy variable that equals 1 if a patient falls in the bottom 20 percent of the income distribution; HighRiskSurgery is a dummy variable that equals 1 if a patient undergoes a surgery for which the risk of death belongs in the 75th percentile and 0 otherwise.

The estimation results are reported in column (1) of Table 8. The results are, for the most part, intuitive. Travel time to the hospital plays an important role in patients' decisions when choosing a hospital – the negative coefficient suggests that patients are less likely to go to hospitals that are located further away from their home. Our estimates also suggest that patients prefer hospitals with higher nurse-to-bed ratio and dislike hospitals with poor clinical quality. We find that older patients are more sensitive to quality of care and the number of nurses per bed. Female patients and patients with lower income prefer hospitals with a smaller nurse-to-bed ratio. Patients who undergo a more risky surgery are generally more sensitive to the quality of clinical care.

To estimate the parameters of the travel-time threshold probabilities, we discretize the travel-time thresholds into 9 points: 30, 60, 90, 120, 150, 180, 210, 240 and 240+ minutes.[22] Concomitant variables that enter the time threshold probability are as follows: Metro is an indicator variable that equals 1 if a patient lives in a metropolitan area other than Seoul and 0 otherwise.[23] Seoul is an indicator variable that equals 1 if a patient lives in Seoul. We also include the variables Female, Old, LowIncome and HighRiskSurgery. Table 10 presents the parameter estimates of the parameters of the travel-time threshold probabilities. Several of our estimates show bi-modality over time constraints which makes complicates the interpretation of several of the coefficients.

Patients living in metro areas and Seoul are more likely to have a choice set to within 90 minutes or beyond 210 minutes compared to patients who live outside these regions. Our estimates also suggest that low income patients are more likely to be time constrained in their choice, and are more likely to have a 30-minute time constraint. This can be due to the monetary cost of traveling long distances. For example, low income patients may not have a car, which is not uncommon given the public transportation infrastructure in South Korea. Older patients are more likely to be time constrained within 60 minutes than younger patients. Meanwhile, older patients are also more likely to travel beyond 240 minutes. This could be because older patients who are more sick may

[22]The travel time threshold of 240+ includes hospitals that are located 240 minutes or more from the train station.

[23]Metro area corresponds to 6 metropolitan cities excluding Seoul consisting of Busan, Daegu, Incheon, Gwangju, Daejeon and Ulsan.

be willing to travel longer distances. Coefficients on the riskiness of the surgery is ambiguous.

We also estimate the hospital choice model using a conventional multinomial logit model (without travel-time constraints). The estimates of the parameters are reported in column (2) of Table 8. The sign and magnitude of the estimates obtained using the traditional multinomial logit model are very similar to those obtained using the ANCS-DC model. We prefer to use the ANCS-DC model, however, because the general theory of choice behavior postulates that individuals follow a two-stage decision process in which the alternatives are reduced to a smaller set (consideration set). The construction of these choice sets depend on factors such as the individual's awareness, feasibility, saliency or accessibility of the alternatives, and mis-specifying the considerations sets may lead to inconsistent parameter estimates. In our setting, we are not able to use an ad-hoc rule such as "15 miles within a patients' home" to define a choice set because a substantial number of patients travel very long distances (even prior to the entry of the HST) to seek better health care services. The ANCS-DC model that we employ is flexible in this manner because it allows the travel time thresholds to be probabilistic, and also to depend on patients' demographic characteristics. We also use the likelihood ratio test to test whether modeling of the choice set incorporated in the formulation of the ANCS-DC model enhances the representation of the observed hospital choice over the conventional multinomial logit model. The χ^2 statistic for this test is $-2\times(-22,806+22,258) = -1,096$ with 112 degrees of freedom, leading to significance at the 0.01 level. This establishes the relevance of travel-time constraints in modeling the hospital choice problem.

9 Counterfactual Analyses

Using the estimates from the demand model we evaluate the impact of the HST on patient welfare. We decompose changes in patient welfare arising from (i) the reduced travel time and (ii) changes in hospital quality. We implement this using the following steps. First, we compute a counterfactual level of clinical care that would arise if the HST is removed (details of this procedure described below), denoted as q_0. Then, using this counterfactual hospital quality and travel-time by car as a baseline, we calculate changes in patient welfare arising from reduced travel time, assuming hospital quality did not change. Next, using the same baseline, we calculate changes in welfare arising from improved hospital quality, assuming that travel time did not change. Finally, we calculate changes in welfare arising from both, reduced travel time and changes in clinical quality.

Next, we evaluate the impact of the entry of the HST on patients' health outcomes. We compare

	30 min	60 min	90 min	120 min	150 min	180 min	210 min	240 min
Intercept	17.2615***	17.8519***	17.4003***	-17.9937***	-21.7797***	-19.8414***	-5.9396***	17.5877***
	(2.6046)	(2.6332)	(2.6423)	(3.0533)	(3.5549)	(3.3340)	(1.1551)	(2.6195)
Metro	2.5793***	2.7020***	2.4578***	-3.0308***	-4.5932***	-2.2823***	2.3009***	2.6175***
	(0.3385)	(0.3566)	(0.3299)	(0.5821)	(0.8663)	(0.3922)	(0.3190)	(0.3420)
Seoul	3.6153***	3.1813***	3.4342***	-3.5962***	-4.5445***	-5.3691***	3.7832***	3.4126***
	(0.5099)	(0.4682)	(0.4629)	(0.7118)	(0.8255)	(1.0097)	(0.4987)	(0.4867)
Female	1.5244***	1.0969***	1.067***	-2.9823***	-4.4934***	1.4734***	4.0788***	2.2359***
	(0.2256)	(0.2489)	(0.2335)	(0.5697)	(0.8306)	(0.2172)	(0.2507)	(0.2421)
Old	16.2797***	24.8265***	-21.6369***	-1.5494***	0.1003***	-0.2240	-37.0428***	23.4989***
	(3.4413)	(4.0175)	(3.7991)	(0.3200)	(0.0390)	(0.1463)	(6.7372)	(4.0247)
LowIncome	35.6560***	-28.7569***	-13.7390***	-6.6266***	-4.3358***	-6.0640***	58.3984***	-31.8741***
	(6.5406)	(4.7814)	(2.3112)	(1.1733)	(0.7264)	(1.0735)	(10.2546)	(6.9005)
HighRiskSurgery	2.1671***	1.5883***	2.1471***	-2.0483***	-3.5532***	-2.0894***	2.1166***	1.9475***
	(0.2488)	(0.2044)	(0.2223)	(0.3294)	(0.6355)	(0.3979)	(0.2504)	(0.2114)

Notes: *** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

Table 10: Estimates of the Time Constraint Parameters

the number of deaths in post-HST period to number of deaths in a counterfactual scenario where the train is removed, and decompose these effects into those caused by changes in travel time and changes in hospital quality.

9.1 Changes in Patient Welfare

We compute the changes in patient welfare from the entry of the HST as follows. Using the parameter estimates from the demand model, we simulate a post-HST scenario where the HST is removed. Recall from our demand model that when travel-time becomes longer (i.e., if the travel time is that of the pre-HST level), constraints imposed on patients' travel-time will force them to remove further-located hospitals (which are included in the choice set if the travel time is that of the post-HST level) from the consideration set.

The expected patient surplus (in utils) for patient i with HST can be expressed as

$$E\left[Surplus_i(t_1,q_1)\right]=\sum_{r=1}^{R_T}E\left[Surplus_{i|r}(t_1,q_1)\right]\cdot\pi_{ir}=\sum_{r=1}^{R_T}E\left[\max_{j\in J_{i|r}^{t_1}}(\bar{U}_{ij}+\varepsilon_{ij})\right]\cdot\pi_{ir}, \tag{15}$$

where t_1 and $q_1$denote travel time and hospital quality with HST, respectively. Similarly, the expected patient surplus when the HST is removed can be expressed as

$$E\left[Surplus_i(t_0,q_0)\right]=\sum_{r=1}^{R_T}E\left[Surplus_{i|r}(t_0,q_0)\right]\cdot\pi_{ir}=\sum_{r=1}^{R_T}E\left[\max_{j\in J_{i|r}^{t_0}}(\bar{U}_{ij}+\varepsilon_{ij})\right]\cdot\pi_{ir}, \tag{16}$$

where t_0 and q_0 denote travel time and hospital quality when the HST is removed. The choice set $J_{i|r}^{t_1}$ differs from the choice set $J_{i|r}^{t_0}$ because changes in travel time change the composition of hospitals in a choice set. Assuming that ε_{ij} is distributed i.i.d extreme value, the above expression can be rewritten as a logit-inclusive value

$$E\left[Surplus_i(t_1,q_1)\right]=\sum_{r=1}^{R_T}\ln\left(\sum_{j\in J_{i|r}^{t_1}}\exp(\bar{U}_{ij})\right)\pi_{ir}, \tag{17}$$

and

$$E\left[Surplus_i(t_0,q_0)\right]=\sum_{r=1}^{R_T}\ln\left(\sum_{j\in J_{i|r}^{t_0}}\exp(\bar{U}_{ij})\right)\pi_{ir}. \tag{18}$$

The average change in surplus is then given by

$$E\left[\Delta Surplus\right]=\frac{1}{N}\sum_{i=1}^{N}\left[E\left[Surplus_i(t_1,q_1)\right]-E\left[Surplus_i(t_0,q_0)\right]\right], \tag{19}$$

where N is the number of patients in post-HST period. This change in welfare can further be decomposed into two parts, the change in welfare from (i) reduction in travel time and fron (ii) improved hospital quality:

$$\begin{aligned}E\left[\Delta Surplus_i\right]=&\tfrac{1}{N}\textstyle\sum_{i=1}^{N}\left[E\left[Surplus_i(t_1,q_0)\right]-E\left[Surplus_i(t_0,q_0)\right]\right]\\&+\tfrac{1}{N}\textstyle\sum_{i=1}^{N}\left[E\left[Surplus_i(t_1,q_1)\right]-E\left[Surplus_i(t_1,q_0)\right]\right].\end{aligned} \tag{20}$$

The first part of equation (20) captures the changes in welfare derived from the changes in travel time while the second part measures the welfare changes caused by the change in hospital quality.

To obtain the counterfactual level of hospital quality that would occur when the HST is removed, we use the fitted values from equation (1) and given by

$$outcome_{jt}=b_{0j}+b_1I(t=post)+b_2I(t=post)\times dist_j^h+\varepsilon_{jt}.$$

Specifically, since the DiD coefficient captures how much the distance to the train station affects the quality in the post-HST period, we can use the above equation to compute the fitted values of hospital quality in a counterfactual scenario when the HST is removed. The details of this procedure are as follows. Denote the hospital which is located furthest away from the HST station in our data as j_{max}. Since this hospital is located far from the train station, we assume that the entry and thus the removal of the HST does not affect this hospital in any way. Denote the distance from j_{max} to its nearest HST station as $dist_{max}$. Since we are removing the HST in the post-HST period (i.e., 2006Q1-2007Q1), we set $d_{post}=1$ and set $dist_j^h=dist_{max}\ \forall j$. In other words, if all hospitals are relocated to be very far from the train station to the degree that they are unaffected by the entry of the train, i.e., $dist_j^h=dist_{max}$, then this effect on hospitals is equivalent to removing the HST.

		time	quality	time & quality
total	ΔUtility	0.2264	0.0174	0.2437
	Dollar Value	$2,211	$170	$2,380
first quartile	ΔUtility	0.3613	0.0160	0.3773
	Dollar Value	$3,507	$156	$3,685
second quartile	ΔUtility	0.1578	0.0231	0.1809
	Dollar Value	$1,503	$225	$1,767
third quartile	ΔUtility	0.1193	0.0166	0.1359
	Dollar Value	$1,169	$162	$1,327
fourth quartile	ΔUtility	0	0.0089	0.0089
	Dollar Value	0	$87	$87

Table 11: Changes in Patient Welfare

We obtain q_0 as the following fitted value

$$\hat{q}_{0,j} = \hat{b}_{0j} + \hat{b}_1 + \hat{b}_2 \cdot dist_{max},$$

and use this as the adjusted mortality rate that would occur in the counterfactual scenario when the HST is removed.

While we could use the q_j obtained in section 4.1 as the hospital quality in a scenario with HST, we instead choose to use the following fitted values for a better comparison[24]

$$\hat{q}_{1,j} = \hat{b}_{0j} + \hat{b}_1 + \hat{b}_2 \cdot dist^h_j.$$

Note that we use the actual distance $dist^h_j$ such that hospitals are affected by the HST.

We first calculate the value given by equation (19) assuming the quality of clinical care did not change. This allow us to evaluate the changes in welfare from the reduction in travel time only. The results are reported in Table 11. Assuming the quality of clinical care did not change, patients on average experience an increase of 0.2264 units in expected utility. This increase in welfare arises from a reduction in travel time, and the resulting ability of patients to sort into better hospitals. Since there is no price coefficient in the demand model due to the absence of a price mechanism in this market, we cannot directly convert the welfare change from utils into a dollar value. Therefore, following Gaynor et al. (2016), we first translate the gains in terms of the preference over distance, and then convert the welfare estimates into a dollar value using additional data from other sources.[25]

[24]The pearson correlation coefficient of the predicted value of $\hat{q}_{1,j}$ and the actual $q_{1,j}$ in the post period is greater than 0.8, with a p-value less than 0.000.

[25]Gowrisankaran et al. (2015) estimate that a one minute reduction in travel time to hospitals increases patient

Comparing the gains in utils to the preference over distance, we find that the welfare effect of the reduction in travel distance for patients belonging to the first quartile corresponds to a 13-minute reduction in travel time.[26] Applying a \$167 value per minute reduction in travel time (Gaynor et al. 2016; Gowrisankaran et al. 2015), the reduction in travel time yields a welfare effect of approximately \$2,211 ($167 \times 13 = 2,211$) per patient.

To look at the gains in welfare based on how close patients live from the train station, we divide the distance to the nearest train station into quartiles, and calculate changes in welfare separately for patients belonging to each quartile. Patients who belong to the first quartile experience an average increase of 0.3613 units in expected utility. This increase in welfare arises from a reduction in travel time, and the resulting ability of patients to sort to better hospitals. As patients' are located further away from the train station, the benefit from the entry of the HST becomes smaller (0.1578 units increase for patients belonging to the second quartile; 0.1193 units increase for patients belonging to the third quartile; no change for patients belonging to the fourth quartile). The welfare effect of the reduction in travel distance for patients who belong to the first quartile corresponds to approximately \$3,507 ($167 \times 21 = 3,507$) per patient. Similarly, patients who belong to the second and third quartiles experience a welfare gain corresponding to approximately \$1,503 and \$1,169, respectively.

Next, we calculate the changes in welfare arising from changes in quality of clinical care, holding the changes in travel time constant. Holding travel time constant, patients experience, on average, an increase of 0.0174 units in expected utility. Applying the same back of the envelope calculations as before to monetize the gains in utils, the improvement in clinical quality yields a welfare gain of approximately \$170 per patient.

Patients who belong to the first quartile in terms of distance to the nearest train experience an average increase of 0.0160 units in expected utility; patients who belong to second and third quartiles experience an increase of 0.0231 and 0.0166, respectively; patients who belong to the fourth quartile experience an increase of 0.0089 utils. The increase in expected utility for patients belonging to the fourth quartile arises from the fact that they face higher clinical quality even though they do not benefit from the new transportation system. This corresponds to monetary gains of approximately \$156 for patient in the first quartile, \$225 for patients in the second quartile, \$161.99 for patient in the third quartile and \$86.84 for patients in the fourth quartile.

surplus by \$167.

[26] $0.2264/(-1.7097) = -0.1324$, where -1.7097 is the coefficient on travel time. Travel time in the regression is defined in units of 100 minutes.

	time	quality	time & quality
total	-0.2003	1.6575	1.8576
first quartile	0.1426	0.9155	1.0580
second quartile	0.0520	0.5895	0.6415
third quartile	0.0057	0.1105	-0.1161
fourth quartile	0	0.0420	0.0420

Table 12: Impact of the HST on Patient Survival

Finally, we calculate the changes in welfare arising from both changes in travel time and changes in quality of clinical care. On average, the increase in expected utility is 0.2437, corresponding to \$2,380 per patient. Patients who belong to the first quartile experience an average increase of 0.3773 units in expected utility; patients who belong to the second and third quartiles experience an increase of 0.1809 and 0.1359, respectively; patients who belong to the fourth quartile experience an increase of 0.0089 (identical to the case when quality of clinical care changes, holding changes in travel time constant). This yields a welfare gain of approximately \$3,685 per patient for patients in the first quartile, \$1,767 in the second quartile, \$1,327 in the third quartile, and \$86.84 for patients in the fourth quartile.

9.2 The Impact of Patients' Sorting on Survival

The HST has enabled patients to choose hospitals that were previously difficult to consider due to long travel distances. Therefore the HST has not only improved the quality of clinical care through increased competition among hospitals, but has also increased the size of the choice set for the patients which in turn has resulted in patients' sorting to better hospitals. One way to directly measure the benefits generated by the HST through its impact on patient sorting is to calculate how many patients would have died in the post-HST period if the HST were to be removed, i.e. post-HST period patients are faced with the pre-HST level travel time to the hospitals.

To implement this, we closely follow Gaynor et al. (2016) and calculate the expected differences in mortality across all patients:

$$E(\Delta Mortality) = \sum_i \left[E\left[Mortality_i(t_1, q_1)\right] - E\left[Mortality_i(t_0, q_0)\right]\right], \tag{21}$$

where

$$E\left[Mortality_i(t_1, q_1)\right] = \sum_j \Pr_{ij}(t_1, q_1) \cdot Prob(Mortality_i | choice = j, Health_i), \tag{22}$$

and

$$E\left[Mortality_i(t_0, q_0)\right] = \sum_j \Pr_{ij}(t_0, q_0) \cdot Prob(Mortality_i | choice = j, Health_i). \tag{23}$$

The probability of patient i choosing hospital j is denoted by $\Pr_{ij}(t, q)$. Equations (22) and (23) denote the mortality probability with HST and without HST, respectively. The variable $Mortality_i$ is an indicator variable which takes value 1 if the patient dies and 0 otherwise. As in the previous subsection, we also decompose the differences in mortality caused by reduced travel time and improved clinical quality.

The results are reported in Table 12. We first assume the quality of clinical care does not change. Our estimates from this counterfactual analysis suggest that 0.2003 lives of patients can be saved from patients' sorting. Since our data corresponds to a 2-percent random sample of the entire population, this translates to approximately 10 lives over the five quarters, which is equivalent to 8 lives on an annual basis.[27]

As before, we divide the distance to the nearest train station into quartiles, and calculate the number of lives saved separately for patients belonging to each quartile. Our calculations show that 0.1426 lives of patients in the first quartile (5.7 lives on an annual basis), 0.0520 lives of patients in the second quartile (2.08 lives on an annual basis), 0.0057 lives of patient in the third quartile (0.228 lives on an annual basis), and 0 lives of patients in the fourth quartile can be saved due to patients' sorting.

Next, we calculate how many lives are saved due to patient sorting when the quality of clinical care also responds to the entry of HST. Our estimates suggest that 1.6575 lives (66.3 lives on an annual basis) of patients can be saved. When we look at the patients by quartile, 0.9155 lives (36.62 lives on an annual basis) of patients in the first quartile, 0.5895 lives of patients in the second quartile (23.58 lives on an annual basis), 0.1105 lives of patient in the third quartile (4.42 lives on an annual basis), and 0.0420 lives of patients in the fourth quartile can be saved (1.68 lives on an annual basis).

Finally we calculate how many lives are saved due to both the reduction in travel time and the improvements in clinical quality. We find that 1.8576 lives (74 lives on an annual basis) of patients

[27] $0.2003 \times 50 \times (4/5) = 8.012$

can be saved. When we look at the patients by quartile, 1.0580 lives (42.32 lives on an annual basis) of patients in the first quartile, 0.6415 lives of patients in the second quartile (25.7 lives on an annual basis), 0.1161 lives of patient in the third quartile (4.6 lives on an annual basis), and 0.0420 lives of patients in the fourth quartile can be saved (1.68 lives on an annual basis).

10 Conclusion

This paper exploits the entry of HST in South Korea, which reduced patients' travel costs, increasing substitutability among hospitals and thereby increasing hospital competition. This exogenous shock allows us to look at the impact of reduced travel time on patient behavior as well as to study the causal impact of increased competition on hospital quality. Taking advantage of the differential effects of the entry of the HST on hospitals located in different regions of the country, we use a difference-in-differences approach to examine the impact of competition on health outcomes measured by 30-day mortality rates following admissions for surgeries. On the methodological side, we utilize the heterogeneous effects of the entry of the HST on patients living in different areas of the country to obtain a reliable measure of hospital-level quality of clinical care.

We find that the entry of the HST improves patient mobility, and that intensified hospital competition leads to an improvement in clinical quality. To evaluate the overall impact of HST on patient welfare, we estimate a structural model of hospital choice, allowing for a flexible formation of patients' consideration set. We find that patients living near a HST station experience an improvement in welfare arising from reduction in travel time as well as improvement in hospital quality. Patients living further away from HST stations also experience an improvement in welfare because, while they do not benefit from reduced travel time, they benefit from the improvement in the quality of hospitals that affected by the entry of the HST. We also find that HST led to a substantial improvement on the probability of patient survival through its effect on patient sorting, even while holding hospital quality constant.

Overall, our paper suggests that increased hospital competition can lead to beneficial health outcomes and that an improvement in transportation infrastructure can have a beneficial impact on patients' health by facilitating patients' sorting to better hospitals through lower travel costs.

References

Acemoglu, D., D. Autor, and D. Lyle (2004): "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury," *Journal of Political Economy*, 112, 497–551.

Beckert, W., M. Christensen, and K. Collyer (2012): "Choice of NHS-funded Hospital Services in England," *The Economic Journal*, 122, 400–417.

Birkmeyer, J. D., A. E. Siewers, E. V. Finlayson, T. A. Stukel, F. L. Lucas, I. Batista, H. G. Welch, and D. E. Wennberg (2002): "Hospital Volume and Surgical Mortality in the United States," *The New England Journal of Medicine*, 346, 1128–37.

Bloom, N., C. Propper, S. Seiler, and J. Van Reenen (2015): "The Impact of Competition on Management Quality: Evidence from Public Hospitals," *The Review of Economic Studies*, 82, 457–489.

Capps, C., D. Dranove, and M. Satterthwaite (2003): "Competition and Market Power in Option Demand Markets," *The RAND Journal of Economics*, 34, 737–763.

Card, D. (1992): "Using regional variation in wages to measure the effects of the federal minimum wage," *Industrial and Labor Relations Review*, 46, 22–37.

Cho, N.-G. and J.-K. Chung (2008): "High speed rail construction of Korea and its impact," *Korea Research Institute for Human Settlement*, 12.

Cooper, Z., S. Gibbons, S. Jones, and A. McGuire (2011): "Does Hospital Competition Save Lives? Evidence from the English NHS Patient Choice Reforms," *Economic Journal*, 121, F228–F260.

Gaynor, M., R. Moreno-Serra, and C. Propper (2013): "Death by Market Power: Reform, Competition, and Patient Outcomes in the National Health Service," *American Economic Journal: Economic Policy*, 5, 134–66.

Gaynor, M., C. Propper, and S. Seiler (2016): "Free to Choose? Reform, Choice, and Consideration Sets in the English National Health Service," *American Economic Review*, 106, 3521–57.

Gaynor, M. and W. B. Vogt (2003): "Competition among Hospitals," *The RAND Journal of Economics*, 34, 764–785.

Geweke, J., G. Gowrisankaran, and R. J. Town (2003): "Bayesian Inference for Hospital Quality in a Selection Model," *Econometrica*, 71, 1215–1238.

Gowrisankaran, G., A. Nevo, and R. Town (2015): "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry," *American Economic Review*, 105, 172–203.

Gowrisankaran, G. and R. J. Town (1999): "Estimating the quality of care in hospitals using instrumental variables," *Journal of Health Economics*, 18, 747–767.

——— (2003): "Competition, Payers, and Hospital Quality," *Health Services Research*, 38, 1403–1422.

Gupta, S. and P. K. Chintagunta (1994): “On Using Demographic Variables to Determine Segment Membership in Logit Mixture Models,” *Journal of Marketing Research*, 31, 128–136.

Halm, E., C. Lee, and M. Chassin (2002): “Is volume related to outcome in health care? A systematic review and methodologic critique of the literature,” *Annals of Internal Medicine*, 137, 511–20.

Ho, K. (2009): “Insurer-Provider Networks in the Medical Care Market,” *American Economic Review*, 99, 393–430.

Kessler, D. P. and M. B. McClellan (2000): “Is Hospital Competition Socially Wasteful?” *The Quarterly Journal of Economics*, 115, 577–615.

Kim, J., J. Lee, W. Yoo, and S. Park (2008): “KTX ui GungangYunghyangPyungka,” *Korea Institute for Health and Social Affairs.*

Kosecoff, J., K. L. Kahn, W. H. Rogers, E. J. Reinisch, M. J. Sherwood, L. V. Rubenstein, D. Draper, C. P. Roth, C. Chew, and R. H. Brook (1990): “Prospective Payment System and Impairment at Discharge: The ’Quicker-and-Sicker’ Story Revisited,” *Journal of the American Medican Association*, 264, 1980–1983.

Marquardt, K., G. Gowrisankaran, and R. Town (2021): “GGT: Stata module to implement Geweke, Gowrisankaran, and Town Model Quality Estimator,” *https://EconPapers.repec.org/RePEc:boc:bocode:s458713*.

Moscelli, G., H. Gravelle, and L. Siciliani (2021): “Hospital competitionand Quality for Non-Emergency Patients in the English NHS,” *RAND Journal of Economics*, 52, 382–414.

Propper, C., S. Burgess, and D. Gossage (2008): “Competition and Quality: Evidence from the NHS Internal Market 1991-9,” *The Economic Journal*, 118, 138–170.

Propper, C., S. Burgess, and K. Green (2004): “Does competition between hospitals improve the quality of care?: Hospital death rates and the NHS internal market,” *Journal of Public Economics*, 88, 1247–1272.

Silber, J. H., P. R. Rosenbaum, J. Tanguy, R. N. Ross, L. J. Bressler, O. Even-Shoshan, S. A. Lorch, and K. G. Volpp (2010): “The Hospital Compare Mortality Model and the Volume-Outcome Relationship,” *Health Services Research*, 45, 1148–6773.

Tay, A. (2003): “Assessing Competition in Hospital Care Markets: The Importance of Accounting for Quality Differentiation,” *The RAND Journal of Economics*, 34, 786–814.

Thill, J.-C. and J. L. Horowitz (1997): “Modelling Non-Work Destination Choices with Choice Sets Defined by Travel-Time Constraints,” *Recent Developments in Spatial Analysis*, 186–208.

Weber, S. and M. Péclat (2017): “A simple command to calculate traveldistance and travel time,” *The STATA Journal*, 17, 962–971.

Appendix A: Alternative measure of hospital quality

In this section, we present estimation results of equation 1 using an alternative measure of hospital quality as in Gowrisankaran and Town (1999). Specifically, we obtain measure of hospital quality by estimating a linear probability model where we regress m_i on a set of hospital dummies and patient's observed characteristics. The mortality of patient i is given as

$$\mu_{it} = \boldsymbol{\psi}'\boldsymbol{c}_i + \boldsymbol{\gamma}'\boldsymbol{h}_i + s_{it} + \eta_{it} \tag{24}$$

where μ_{it} is a dummy variable that denotes the death of patient i within 30 days of the admission, $\boldsymbol{c}_i$ is a vector of dummy variables where c_{ijt} equals 1 if patient i $(i = 1, ..., N)$ chooses hospital j $(j = 1, ..., J)$, $\boldsymbol{h}_i$ is a vector of patient characteristics that can affect mortality, s_{it} is unobserved (by the researcher) severity of illness, and η_{it} is an i.i.d. normal error term. The parameter vectors to estimate are $\boldsymbol{\psi}$ and $\boldsymbol{\gamma}$. With the linear probability model, the elements of estimated fixed effects $\hat{\boldsymbol{\psi}}$ are interpreted as the incremental probability of death from choosing a particular hospital conditional on observed health status, and is used as our measure of quality of care. The coefficient vector $\boldsymbol{\gamma}$ captures the impact of patients' observed health status on the probability of death. Following section 4.1, we will refer to the estimated measure of quality of care, $\hat{\boldsymbol{\psi}}$ as the adjusted mortality rate. Because hospital choice is likely to be correlated with patients' unobserved severity of illness, estimating equation (24) using OLS will lead to biased estimates. For instance, if sicker patients are more likely to choose a certain hospital j, then s_{it} and c_{ijt} will be positively correlated, and hence $\hat{\boldsymbol{\psi}}_{\boldsymbol{j}}$ will be overestimated.

To address the endogeneity of hospital choice, we use two sets of instrumental variables for hospital choice dummy variables $(\boldsymbol{c}_i)$: (i) the travel time to each hospital, and (ii) and instruments of the form $exp(-\phi \times \text{traveltime}_{ijt})$, where we define travel time for patient i to hospital j in period t as

$$\text{traveltime}_{ijt} = \begin{cases} \min(\text{cartime}_{ij}, \text{traintime}_{ij}) & \text{if } t = \text{post-HST \& dist}_i^{pat}{<}30 \text{ \& dist}_j^h{<}30 \\ \text{cartime}_{ij} & \text{otherwise} \end{cases} \tag{25}$$

Here cartime_{ij} denotes the drive time from patient i's location to hospital j by car, and traintime_{ij} is the travel time from patient i's location to hospital j by HST. [28] $dist_i^{pat}$ is the travel time from

[28]Note that traintime_{ij} is obtained by summing the following three components: (i) drive time from i's location to i's nearest HST station h, (ii) travel time from station h to station k, which is the closest HST station to hospital j

patient i to the closest train station and $dist^h_j$ is the travel time from hospital j to the closest train station. We constrain the effect of the HST to patients and hospitals living 30 minutes within the train station. This is to account for the changes in travel time only for patients living sufficiently close to the HST station in the post-HST era, and is based on the pattern in the data where there are no significant differences in travel times in pre- and post- HST for patients living beyond 30 minutes of the HST station.

Formal specification tests for the validity of our instruments are provided in Table A.1.[29] Our overidentifying restrictions are valid as we fail to reject the null of the Sargan-Hansen overidentification test. We reject the null hypothesis of the Hausman Endogeneity test which means that our OLS and IV estimates are statistically different. We also perform the Wald-Test of Weak Instruments and reject the hypothesis that our instruments are weak. These tests provide support for the validity of our IV specification.

Sargan-Hansen	χ^2	60.5670
Overidentification Test	p-value	0.9792
Hausman	χ^2	8,600
Endogeneity Test	p-value	0.0000
Wald-Test of	χ^2	934.9818
Weak Instruments	p-value	0.0000

Table A.1: Tests for Validity of Instruments

	(1) OLS	(2) IV
d_{post}	-0.0061	-.01088***
	(0.0082)	(0.0285)
$d_{post} \cdot dist^h_j$	0.0005 **	0.0015**
	(0.0002)	(0.0007)
R^2	0.5746	0.6502
hospital FE	✓	✓
number hospitals	42	42
Observations	84	84

Notes: Models are estimated using OLS with standard errors (in parentheses under coefficients). All regressions include constants.

*** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

Table A.2: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care

and (iii) drive time from station k to hospital j. We obtain driving time by car by using *georoute* routine developed by Weber and Péclat (2017) which calculates the driving time between two points under normal traffic conditions.

[29]Note that we perform the specification tests for the data pooled across pre- and post- HST periods.

In order to control for the observed case-mix at the patient-level, we first estimate equation (24) using OLS, and use estimated $\hat{\psi}$ as a measure of clinical quality to estimate equation (1). Note that although this measure of quality controls for observed health status at the individual patient-level, it does not control for unobserved (to the researcher) severity of illness which may be correlated with patients' hospital choice, and hence may be biased. The results are reported Table A.2, column 1. The DiD coefficient is positive and significant ($\beta_2 = 0.0005$), i.e. when a hospital's travel time to a closest train station decreases by 1 minute (i.e. hospital is closer to the train station), (adjusted) mortality rate decreases by 0.05 percentage points. As already mentioned, however, simply controlling for observed patient case-mix is not sufficient to correctly measure the quality of clinical care. Patients' unobserved (to the researcher) severity of illness, which may be correlated with hospital choice, may contaminate the quality of clinical care. We further control for patients' unobserved severity of illness by instrumenting hospital choice dummy variables for each period with travel time to each hospital, and use thus (using IV) obtained adjusted mortality rates as the dependent variable to estimate equation (1). The results are reported in Table A.2, column 2. After controlling for unobserved severity of illness, we see that the (absolute) magnitude of the DiD coefficient has become larger. The DiD coefficient is 0.0015 and significant, suggesting that when a hospital's travel time to a closest train station decreases by 1 minute, (adjusted) mortality rate decreases by 0.15 percentage points.

Appendix B

To test to robustness of our findings to the specification of the prior distribution, we estimate the selection model of section 4.1 using the following two variants of prior distribution specifications: Variant 1 scales the prior standard deviations of α in the original selection model downward by a factor of 5 and τ^2 downward by a factor of 25. Variant 2 is the opposite of variant 1 in that the prior standard deviations are scaled up by a factor of 5 relative to the base model. This is the same approach used by Geweke et al. (2003) to test the sensitivity of their estimates to specifications of the prior distribution. We then use measures of hospital quality from these two specifications to estimation equation 1. Table B.1 provide the estimation results of equation 1 using the measures of hospital quality obtained using these alternative prior distributions. The results are consistent with the findings in section 6.

	(1) prior variant 1	(2) prior variant 2
d_{post}	0.0194	0.8249***
	(0.0254)	(0.0897)
$d_{post} \cdot dist_j^h$	0.0032***	0.0093**
	(0.0015)	(0.0045)
R^2	0.6241	0.6891
hospital FE	✓	✓
number hospitals	42	42
Observations	84	84

Notes: Models are estimated using OLS with standard errors (in parentheses under coefficients). All regressions include constants.

*** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

Table B.1: Robustness to Prior Specification