Summer 8-15-2021

Essays on Markets with Frictions

Hyesung Yoo
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST.LOUIS

Olin Business School

Dissertation Examination Committee:
Maria Ana Vitorino, Co-Chair
Song Yao, Co-Chair
Tat Chan
Yulia Nevskaya
Raphael Thomadsen

Essays on Markets with Frictions

by
Hyesung Yoo

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2021
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am indebted to my advisors, Maria Ana Vitorino and Song Yao, for their guidance, support and encouragement. Both of them were always there for me, starting from University of Minnesota, to Washington University, and through the entire COVID job market. They spent more time helping me than I could have ever hoped for.

I am also indebted to Mark Bergen for his guidance and endless support.

I thank Tat Chan, Yulia Nevskaya, and Raphael Thomadsen for serving as members of my dissertation committee and for their consistent help and kindness. I also owe gratitude to Amil Petrin and Linli Xu for serving as preliminary committee members.

I also thank my family for the constant support and encouragement.

Hyesung Yoo

Washington University in Saint Louis

August 2021

Dedicated to my parents and my sister

ABSTRACT OF THE DISSERTATION

Essays on Markets with Frictions

by

Hyesung Yoo

Doctor of Philosophy in Business Administration

Washington University in St. Louis, 2021

Professor Maria Ana Vitorino, Co-Chair

Professor Song Yao, Co-Chair

In this dissertation, I explore the implications of various forms of frictions on market outcomes. Specifically, I look at search frictions in two-sided markets, geographic frictions in a healthcare market, and the use of a machine learning approach in the presence of regulatory frictions.

In the first chapter, I leverage the entry of a high-speed train system in South Korea as a natural experiment to establish the causal effect of competition between hospitals on health care quality and consumer welfare. Using a difference-in-differences estimator, we examine the effects of competition on hospitals depending on their proximity to train stations, notably how increased competition impacts health outcomes as measured by 30-day mortality rates. Our results suggest that increased competition leads to an improvement in the quality of clinical care. To evaluate the overall impact of the HST on patient welfare, we estimate a structural model of hospital choice, allowing for a flexible formation of patients' consideration sets. We find that patients living near a HST station experience an improvement in welfare arising from the reduction in travel time as well as improvements in hospital quality. Patients living further away from HST stations also experience an improvement in welfare although they do not gain from the reduced travel time due to the improvement in the quality of treated hospitals. We also find that the HST can have a beneficial impact on patient health

by facilitating patients' sorting to better hospitals, even while holding quality of clinical care constant.

In the second chapter, I study the impact of search frictions and preferences on the formation of a match in two-sided markets. Since agents on both sides have private preferences regarding each others' characteristics, forming a match based on mutual compatibility requires extensive costly search. To better understand the relative impact of search frictions and preferences on match outcomes, we use data from a field experiment conducted on an online dating platform wherein randomly selected users are given the ability to know upfront a piece of information about the private preference of the opposite gender (information which otherwise should have been searched for). We find descriptive evidence suggesting that reducing search frictions through the provision of information may lead to less sorting between matched couples in terms of various characteristics such as race and education level. To investigate the relative contribution of search frictions and preferences on assortative matching, we develop and estimate a model that incorporates both costly search and preference heterogeneity across users. Identification of our model relies on the variation in information caused by the experiment as well as the exclusion restriction to separately identify preferences from costs. Our estimation results reveal that frictions play a significant role in shaping matching outcomes. Using model estimates, we simulate matches under various environments, including the Gale-Shapley protocol. We find that removing frictions leads to significantly less sorting between couples. We also find that frictions in our platform lead to significant departure from efficiency. These results highlight the importance of platform designs that aim to reduce search frictions. In addition, with one-third of the marriages in the U.S. beginning online, this paper shows how the design of an online platform can contribute to diversity, which can in turn alleviate persistent social inequality.

In the third chapter, I study how we can use machine learning methods to overcome challenges faced by firms in the presence of restrictive privacy regulations. The ever-increasing volume of consumer data provide unprecedented opportunities for firms to predict consumer behavior, target customers, and provide customized service. Recent trends of more restrictive privacy regulations worldwide, however, present great challenges for firms whose business activities rely on consumer data. We address these challenges by applying the recently developed federated learning approach - a privacy-preserving machine learning approach that uses a parallelized learning algorithm to train a model locally on each individual user's device. We apply this approach to data from an online retailer and train a Gated Recurrent Unit recurrent neural network to predict each consumer's click-stream. We show the firm can predict each consumer's activities with a high level of accuracy without the need to store, access, or analyze consumer data in a centralized location, thereby protecting their sensitive information.

Chapter 1

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

1.1 Introduction

It is important to understand how competition affects service quality in the health care industry. However, empirical evidence on this topic is mixed. Policies to improve the efficiency and the quality of health care have been introduced in several countries, but their effectiveness remains ambiguous. Difficulty in assessing the impact of competition is partly due to the fact that competition in health care markets is geographically based, as pointed out by Propper et al. (2008) and Gaynor et al. (2013a).

Many existing studies rely on cross-sectional and over time variation in hospital market structure to identify the impact of competition on service quality of hospitals. However, the market structure may be endogenous because the quality of incumbent hospitals and potential entrants may affect their strategic entry and exit decisions, hence the market structure. Other studies exploit changes in health-related policies, which are exogenous shocks that spur competition. Yet the analysis are often complicated by the fact that when policies are *health*-related, they may affect the incentives of the agents involved in ways unanticipated by researchers. If such incentive changes are not accounted for in the analysis, the conclusions may be biased.

In this article we exploit the entry of high-speed train (HST henceforth) system in South Korea to examine the effects of competition on the quality of health care. As of April 2004, Korea Train eXpress (KTX) started operating in South Korea, connecting most major cities by high-speed rail. An important aspect of the South Korean healthcare industry is that patients have the full freedom to go to any hospital of their choice and prices are fixed. 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 news 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, when Kim et al. (2008) randomly surveyed HST passengers arriving in Seoul and asked them: "Have you used the HST to seek treatment in hospitals located in Seoul at least once?" 36% (out of 561 passengers) responded "Yes". The news reports and the survey provide some evidence that patients indeed use the HST for medical purposes. Clearly the reduction in travel time facilitates access to better hospitals, implying that hospitals that previously competed locally are now competing with those located further away.

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

We use the fact that the HST does not extend to all regions, thereby increasing competition only for hospitals that are located sufficiently close to HST stations. In the current context, there are treated hospitals - hospitals that are located close to a HST station, as well as treated patients - patients that live close to a HST station (more discussion on this subject in the next section). Although our primary interest is to study the impact of competition on hospital quality, we distinguish patients in the treated group from those in the control group so as to provide descriptive evidence on patients' responses to the entry of the HST, as well as to investigate differential changes in patients' welfare and health outcomes based on where they live.

We begin by providing descriptive evidence to show that patients in the treated group traveled further distances to visit a hospital after the entry of the HST, whereas we do not see such a pattern for patients in the control group, suggesting that patients responded differently to the entry of the HST depending on the proximity from their home to the HST station. Using a difference-in-differences estimator, we then examine the impact of increased hospital competition on the hospital clinical quality, as measured by 30-day risk-adjusted mortality rates following admissions for a surgery. Specifically, we look at all surgeries that were conducted during this period where mortality rate can be used as a measure of quality of clinical care.We find that increased competition improves the clinical quality: Hospitals affected by the entry of the HST experience a decrease in adjusted mortality rates.

We then estimate a structural model of hospital choice and use the model estimates to quantify the impact of the entry of the HST on patient welfare. We find that patients living in treated regions experience an improvement in welfare due to both reduction in travel costs as well as enhanced clinical quality. Although patients living in control regions do not benefit from reduced travel costs (because there is no HST station near their homes), they also experience an increase in welfare because many of them choose to go to hospitals that are affected by HST. We further use the model estimates to measure the effect of patients' sorting to better hospitals (due to lower travel costs) on their health outcomes (survival from the surgery). This is implemented by comparing the number of death in the post-HST period to a counterfactual scenario when the HST is removed. From this analysis we find that a substantial number of lives can be saved annually with the HST as a result of patients sorting to better hospitals. Our research contributes to the literature on hospital competition and quality in health care. The most influential study of health care markets with fixed prices is Kessler and McClellan (2000), who examine the impact of market concentration on both costs and mortality rates for US Medicare Acute Myocardial Infarction (AMI) patients. They find that in the 1980s competition led to higher costs but lower mortality rates, but find that after 1990, competition resulted in both lower costs and lower mortality rates, and conclude that competition is unambiguously welfare improving post-1990.[2] Exploiting the 2006 English pro-competitive policy shift, Gaynor et al. (2013a) study the impact of the competition on quality (as measured by death rates from heart attack) as well as other measures of quality such as hospital productivity and expenditures (hospital operating expenditures and expenditures per admission) using a difference-in-differences research design. They find that increased competition improves the quality of clinical care without increasing expenditures.[3] Leveraging the same reform, Gaynor et al. (2016) find that patients became more responsive to clinical quality post-reform, and that hospitals responded to changes in demand by improving quality.

Some other papers that study the relationship between competition and healthcare quality, however, find opposite results. Using Medicare data for AMI and pneumonia patients, Gowrisankaran and Town (2003) also estimate the impact of hospital market structure on

[2]Other papers such as Shen (2003) finds mixed effects, and Shortell and Hughes (1988) find no effects of competition on quality using Medicare patient data.

[3]For measures of hospital productivity, they use simple measure of labor productivity- the number of admissions per clinical staff.

mortality rates for Medicare patients and find that mortality rate is worse for patients treated in hospitals with more intense competition. This is in contrast to the classical theoretical literature that increased competition under fixed prices results in improved quality. Gowrisankaran and Town (2003) provide a possible explanation: If the profit margin on Medicare patients is sufficiently low, then greater competition for these Medicare patients can cause the hospital to focus on more profitable HMO patients and give up on investing in Medicare patients. In fact, Brekke et al. (2011) show theoretically that under fixed prices, increasing competition through either lower transportation costs (increased substitutability) or a higher number of hospitals may have ambiguous effects on quality if profit margins are low or negative, or if hospitals deviate from profit-maximizing behavior. Several papers find further empirical evidence that support the results of Gowrisankaran and Town (2003) (Propper et al. (2004), Propper et al. (2008), Lewis and Pflum (2017), Colla et al. (2016)). Leveraging the 1991 health reform in the UK National Health Service, Propper et al. (2004) find that the relationship between competition and AMI mortality rates are negative. Propper et al. (2004) investigate the changes further and find that increased competition reduces waiting times, suggesting that hospitals facing more competition reduce services that affect mortality rates (that are unobserved) in order to increase other activities which are better observed by the health-care buyers. Findings of Lewis and Pflum (2017) also suggest that in response to competition, hospitals divert the resources away from investing in clinical quality, which is imperfectly observed, in order to increase investment in amenities that are better observed by the patient.[4]

Our research advances the existing literature in health economics by studying the effects of competition with fixed prices following an exogenous shock. Because the shock (entry of HST)

[4]In contrast to Medicare patients, however, both Gowrisankaran and Town (2003) and Lewis and Pflum (2017) find that competition improves clinical quality for HMO patients. Lewis and Pflum (2017) explain that because HMOs can better evaluate the clinical quality of hospitals than individual patients, hospitals have higher incentives to improve clinical quality levels when competing for inclusion in HMO provider networks.

that increases competition is orthogonal to hospital market structure or any other aspect of healthcare, our setting provides a unique and novel natural-experiment that helps answer our research question. Furthermore, because the HST only reaches certain regions of the country, not only can we do a pre-post analysis, but we are also able to explore the variations in the degree of treatment for more convincing insights. To the best of our knowledge, Gaynor et al. (2013a) and Propper et al. (2008) are the only papers that employ difference-in-differences approach to study this question.

Our research is also the first in which the competition is driven by a shock that reduces travel costs. In their theoretical model, Brekke et al. (2011) measure intensified competition in two ways; more hospitals in the market, and lower transportation costs (increased substitutability between hospitals). The importance of tradeoff between quality and travel time that patients face is highlighted in Tay (2003). This tradeoff between quality and travel time is what gives market power to hospitals. The entry of the HST reduces the travel time faced by the patients, thereby alleviating this tradeoff. As long as hospital quality remains unchanged, the introduction of the HST therefore should be unambiguously welfare improving. We show that this is indeed the case by decomposing the changes in patient welfare resulting from changes in travel time and changes in mortality rates.

Our research is also closely related to the literature on constrained choice sets. Ho (2006), Dafny et al. (2013) and Gaynor et al. (2016) also analyze the effects of removing choice constraints within the health care context. Our setting is more similar to that of Gaynor et al. (2016) in which the patients' choice sets are unobserved. To exploit the unique feature of our setting in which the entry of the HST increased the number of hospitals in a patient's consideration set, we adopt a modeling approach used in the geography/transportation literature. Specifically, we explicitly model the formation of consideration sets for the post-HST period when individuals have limited time resource, and evaluate the welfare effects of the removal of the HST in a counterfactual scenario.

Finally, our research adds to the fast growing literature on the economic impacts of transportation infrastructure (Benerjee et al. (2012); Qin (2016); Donaldson (2018); Heuermann and Schmieder (2018); Qin et al. (2018)). While these papers mainly study the impact on economic activities that are directly affected by the new transportation system, such as inter-regional trade, per capita GDP, co-opetition between transportation modes, and housing/commute decisions, our paper looks at the unexpected externality caused by the HST.

The rest of this paper is structured as follows. In the next section we describe the relevant aspects of the industry; Section 1.3 describes the data; In Section 1.4 we describe our estimation strategy and present the results. Section 1.5 outlines the structural demand model, and Section 1.6 presents the estimation results. Section 1.7 analyzes the welfare effects of the entry of the HST and Section 1.8 concludes.

1.2 Industry Details

1.2.1 Health Care Industry

National Health Insurance (NHI) program in South Korea is a compulsory single-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.

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. This is to ensure that private health-care providers respond to changes in demand which the public health insurance has brought about.

The NHIC negotiates the level of medical service fees annually with provider associations. The fee schedule includes fees for all medical services and materials including drugs, as well as remuneration of providers for the services they provide. Patients are responsible for any co-payments applicable to the medical services they received, and the NHIC reimburses healthcare providers the share of medical costs not borne directly by the patient on the basis of the fee schedule. 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, which did not keep pace with inflation until the mid-1990s.

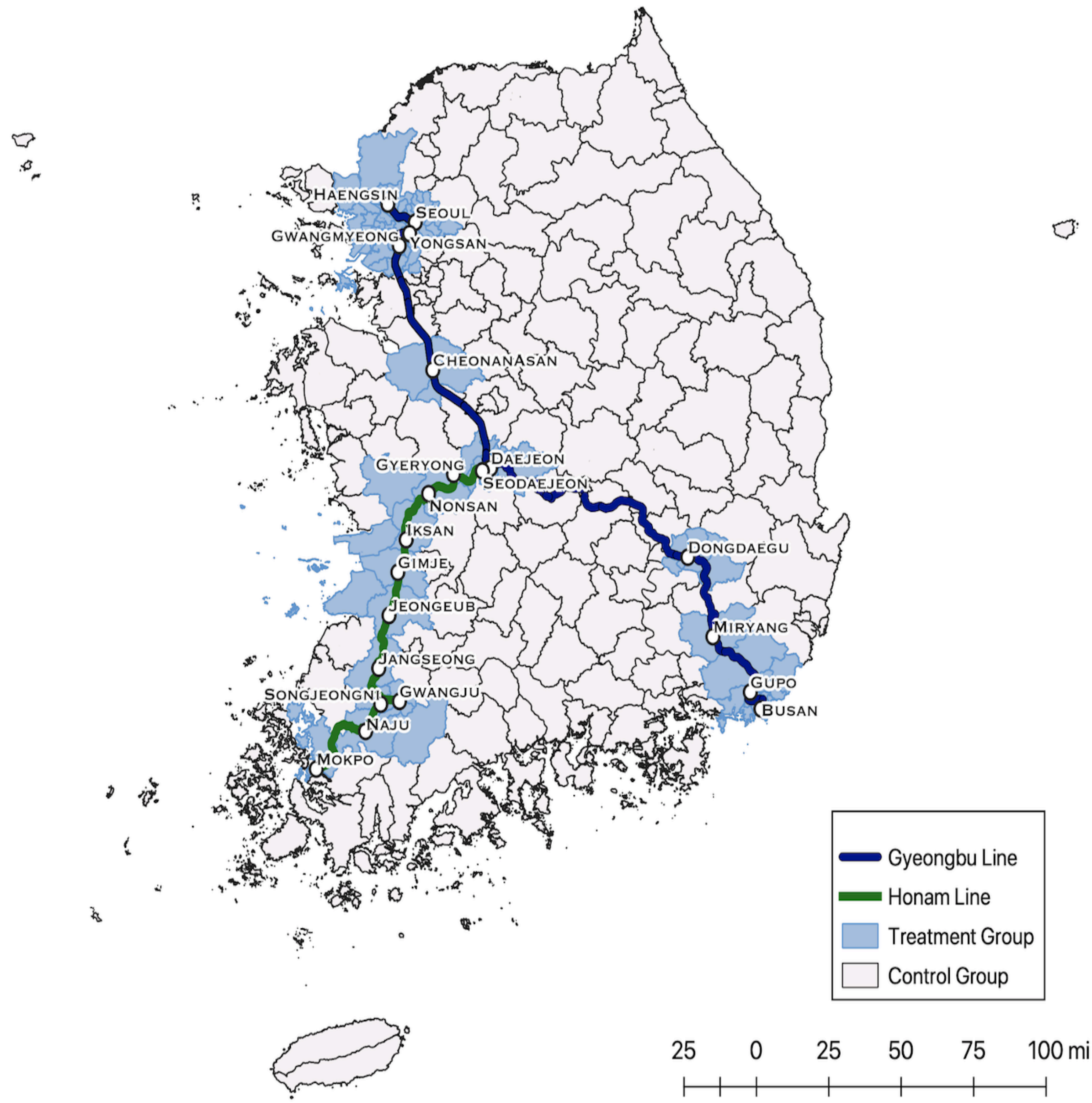
In fact, according to a report published by Health Insurance Review Assessment (HIRA) in 2006, the fixed fee schedule covers on average only 73.9% of the costs incurred by providers.[5] Differential margins for different medical services lead physicians to provide more of those services with higher margins. Specialties of which services are paid relatively generously attract a greater number of applicants for the residency training. Popular specialties include psychiatry, ophthalmology and dermatology, whereas radiology, thoracic surgery and anesthesiology are the least popular. Moreover, to mitigate the effects of negative margins, physicians encourage patients to receive uninsured medical services, for which hospitals have the full freedom to set their own price.

Healthcare delivery system in South Korea is classified into three tiers: primary, secondary, and tertiary care. Although NHI service flow is designed to progress from primary to secondary to tertiary care, patients have the complete freedom to choose a healthcare provider at any level within this system with some financial incentives. To achieve an efficient distribution of limited healthcare resources, outpatient insurance coverage largely depends on the tier of the hospital. Patients must be referred by primary or secondary care hospitals to receive outpatient treatment in tertiary hospitals, in which case 40% of their bills are covered by insurance (Otherwise, they can expect to pay 100% of the bill). The insurance coverage is identical at all levels of hospitals for inpatient care, where patients pay 20% of medical expenses. In our analysis we only focus on inpatient surgical treatments.

1.2.2 Entry of High-Speed Train

South Korea's HST system, Korea Train eXpress (KTX), commenced commercial operations on April 1st 2004, substantially altering patterns of long-distance travel. Construction of

[5]Source: http://www.medicaltimes.com/News/39629, accessed on November 30, 2018



Notes: This map displays first-stage HST lines. Shaded areas represent regions that are treated - districts whose centroids are located within 15 miles of the HST station.

Figure 1.1: Treated Regions

the HST system occurred in two stages.[6] The first-stage construction involved building 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.[7] 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 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, Gyeongbu Line (blue) connecting Seoul-Busan and Honam line (green) connecting Seoul-Mokpo. We define a "treatmented area" as an area located within 10 miles of the HST station. Shaded areas in Figure 1 represent treated areas whose centroids are within 10 miles of a HST station.[8]

At the time of the 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. 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.[9] The HST system has reduced

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

[7] Newly constructed links included 51.6 mi of viaducts and 47.0 mi of tunnels. Electrification of the existing rail comprised of 82.5 mi across Daegu to Busan, 12.9 mi across Daejeon, and 164.3 mi 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).

[8] The robustness of our 10-mile definition of treatment is discussed in Appendix C

[9] 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.

the travel time from Seoul to Busan from more than 5 hours by car to 2 hours 40 minutes by train.

1.3 Data

We rely on a number of data sources at the patient, hospital and city-county-district level.[10]

Our patient data comes from the National Health Insurance Services (NHIS) which is a health insurance claims dataset collected by the single insurer system NHI (NHIS-2018-2-139). Our data are of a nationally representative random sample, which accounts for approximately 2% of the entire South Korean population for years 2003 to 2007. The data contain patient-level information on medical procedures received at the hospitals. Detailed information on patient demographics, diagnosis, patients' location at the district level[11] and hospital choice are observed, as well as the date of admission, number of inpatient treatment days, and the month/year of the patient's death.

The identity of the hospitals in the NHIS dataset are anonymized and hospital location is observable only at the provincial 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 HIRA (Health Insurance Review Assessment) which, in addition to the hospital characteristics in the NHIS dataset, also provides hospital location at the district level[12]. The identity of the

[10] 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). 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 diveded into neighborhoods (dong).

[11] More precisely, patients' locations are at the city-county-district level because some counties are not populated enough to qualify for a city and hence are not sub-divided into districts.

[12] For the same reasons as discussed in footnote 11, hospitals' locations are at the city-county-district level

	Pre-HST				Post-HST			
	Mean	SD	Min	Max	Mean	SD	Min	Max
				Control Hospitals (N = 55)				
Total admissions	156	138.1	13	598	193.1	188.0	13	789
Hospital beds	567.3	296.3	99	1256	567.3	296.3	99	1256
Mortality rates	0.050	0.041	0	0.280	0.052	0.050	0	0.308
				Treated Hospitals (N = 112)				
Total admissions	243.9	256.4	17	1501	290	313	13	1943
Hospital beds	703.2	453.9	121	2993	703.2	453.9	121	2993
Mortality rates	0.043	0	0.176	0.040	0.040	0.029	0	0.176

Notes: hospital-treatment in this table is defined as being located within 15-miles of train stations

Table 1.1: Summary Statistics: Hospital Characteristics

hospitals in the HIRA dataset is also anonymized, but we are able to match this dataset to NHIS dataset using hospitals characteristics.

Our sample selection process is as follows: To study the causal impact of increased competition on the quality of clinical care, we define January 2003 to March 2004 as the pre-HST time period and January 2006 to March 2007 as the post-HST time period. The data are collapsed into pre- and post-HST period.

We focus on patients who underwent a surgery. Specifically, we consider all surgeries that were conducted during this period that resulted in at least one death. Since our data is a 2% sample of the entire population, 1 death in the data can be inferred as 50 deaths in the entire population. Ideally we want to 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 (mortality rates) with patient selection.[13] Constraining our analysis to a single type

[13] 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.

(Fractions of)	Control Patients				Treated Patients			
	Pre-HST Period		Post-HST Period		Pre-HST Period		Post-HST Period	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	0.490	0.450	0.486	0.500	0.497	0.500	0.483	0.500
Ages 0-19 Years	0.150	0.357	0.147	0.354	0.149	0.356	0.155	0.362
Ages 20-39 Years	0.188	0.391	0.167	0.373	0.213	0.410	0.177	0.382
Ages 40-59 Years	0.290	0.454	0.290	0.454	0.300	0.458	0.303	0.460
Ages 60-79 Years	0.328	0.470	0.344	0.475	0.299	0.458	0.320	0.466
Ages 80 Years +	0.044	0.204	0.052	0.221	0.039	0.194	0.044	0.206
Income Group 0-1	0.084	0.277	0.077	0.267	0.068	0.251	0.067	0.250
Income Group 2-4	0.198	0.399	0.188	0.391	0.186	0.389	0.179	0.383
Income Group 5-7	0.296	0.456	0.286	0.452	0.295	0.456	0.289	0.453
Income Group 8-10	0.422	0.494	0.449	0.497	0.451	0.498	0.465	0.499
Comorbidity	0.873	0.333	0.847	0.360	0.854	0.353	0.844	0.363
Nobs	18,639		22,431		17,252		20,664	

Notes: patient-treatment in this table is defined as living within 10-miles of train stations

Table 1.2: Summary Statistics: Patient Characteristics

of surgery, however, leaves us with too few observations (too few patients as well as too few hospitals). Limiting our attention to only one "category" of surgery (e.g. cardiovascular surgery) also leaves us with too few observations. To attenuate the contamination of hospital quality from pooling patients across multiple types of surgeries, we control for the riskiness of each type of surgery as well as the patients' diagnosed disease when obtaining the adjusted mortality rates. Details of this procedure are provided in the Appendix A.

The key feature of our setting is that the entry of HST enables 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 we are unable to calculate the travel time to hospitals by car and ferry for these patients, a necessary component for estimating our demand model and performing counterfactuals.

We drop outpatient admissions, where the patient stayed at the hospital less than 24 hours to ensure that the patients in our sample are sick enough. Following Tay (2003) and Ho (2006), we exclude hospitals with fewer than 10 admissions per period, and we only keep hospitals that appear in both pre- and post-HST periods to facilitate the comparison of hospital quality. Our final sample consists of 167 hospitals and 78,986 patients.

In our setting, there are "Treated Hospitals" and "Treated Patients". "Treated Hospitals" are hospitals that are located within 15 miles of the HST station, and "Treated Patients" are patients who live within 15 miles of the HST station. In our main analysis, we define Treated Hospitals as hospitals that are located within 15 miles of the HST station unless otherwise specified. For Treated Patients, we show descriptive statistics using both 10 mile and 15 mile definitions of treatment. In the Appendix, we show that our results hold consistently even if we change the definition of treatment as being located within 5 miles, 10 miles, 20 miles. Table 1 and Table 2 provides summary statistics of hospital characteristics and patient characteristics, respectively.[14]

In Table 1.3 we present descriptive evidence on changes in patients' travel patterns following the entry of HST. Panel A.1 reports the average travel distances (in miles) before-and after the introduction of HST, defining patients living within 10 miles of the HST station as treated. While there are no changes in travel distances for patient living in control regions, patients living in treated regions clearly traveled further distances after the entry of the HST (approximately 8 percent increase post HST). These differences become more salient when

[14]In our data we observe up to two diagnosis per patient, main diagnosis and sub-diagnosis. Since not all patients have a sub-diagnosis, we construct a comorbidity dummy variable which takes value 1 if a patient has a sub-diagnosis and 0 otherwise.

we only focus on patients living in non-Seoul areas (Panel A.2): while there is no difference in travel distance for patients living in control regions, the average travel distance increased by 11 percent for patients living in treated regions.[15] While the absolute difference in travel time for treated patients before and after the HST may seem small, this change account to approximately 10% increase in proportion of patients traveling beyond 50 miles (see Appendix D)

Our final sample excludes patients who transferred from other hospital 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 for patients living in treated regions 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 3 Panel A.3 reports the mean travel distances for patients who arrived at hospitals via transfer or ambulance. As expected, we do not see changes in travel distance for patients living in treated regions.

Panel B reports the average travel distances by period and region, defining patients living within 15 miles of the HST station as treated. The patterns reported in this table are consistent with those in Panel A.

[15] More precisely, we exclude Seoul and the surrounding metro area.

Table 1.3: Descriptive Evidence of Changes in Travel Distance

Distance Traveled	Control Patients			Treated Patients		
	Pre-HST	Post-HST		Pre-HST	Post-HST	
	Mean	Mean	%Δ	Mean	Mean	Δ
	(st.dev)	(st.dev)	t-stat	(st.dev)	(st.dev)	t-stat
Panel A: Patient Treatment: within 10 miles of train station						
			A1. Excluding ambulance and transfer patients			
	27.338	27.825	1.78%	13.713	14.781	7.79%
	(43.183)	(43.295)	t: 1.135	(36.365)	(37.602)	t: 2.797**
Nobs	18,639	22,431		17,252	20,664	
			A2. Excluding patients living in Seoul			
	29.601	29.704	0.35%	21.957	24.458	11.4%
	44.344	44.313	t: 0.224	48.372	50.821	t: 3.192***
Nobs	16,867	20,520		7,587	8,557	
			A3. Ambulance and transfer patients			
	22.112	22.363	1.14%	11.096	9.3890	-15.4%
	(36.525)	(35.887)	t: 0.114	(34.202)	(25.805)	t: 1.041
Nobs	387	853		478	901	
Panel B: Patient Treatment : within 15 miles of train station						
			B1. Excluding ambulance and transfer patients			
	33.071	33.472	1.21%	13.334	14.128	5.95%
	(46.240)	(46.425)	t: 0.747	(34.746)	(35.567)	t: 2.485**
Nobs	13,556	16,580		22,335	26,515	
			B2. Excluding patients living in Seoul			
	33.071	33.472	1.21%	19.963	21.113	5.76%
	(46.240)	(46.425)	t: 0.7471	(44.109)	(45.381)	t: 1.9593*
Nobs	13,556	16,580		10,898	12,497	
			B3. Ambulance and transfer patients			
	28.190	25.529	-9.44%	10.652	10.079	-5.38%
	(39.746)	(37.536)	t: 0.953	(32.310)	(26.379)	t: 0.396
Nobs	265	638		600	1,116	

1.4 Difference-in-Differences Estimation and Results

In this section we study the impact of hospital competition on the quality of clinical care using a difference-in-differences approach. Specifically, we compare pre- and post-HST mortality rates of hospitals that are located near the HST stations, using hospitals located further away from the HST stations as the control group. Figure 2 plots the hospital-level raw mortality rates by quarter for pre- and post-HST periods.[16] From this figure, we can see that mortality rates at the treated and control hospitals roughly follow a similar trend.

We first briefly describe our estimation strategy, and then proceed to describe the issue concerning the use of raw mortality rates as a measure of hospital clinical quality, followed by a description of how to resolve this problem. We then report our estimation results.

1.4.1 Difference-in-Differences

We analyze the impact of hospital competition on the quality of clinical care using difference-in-differences estimator. Specifically, we estimate the equation as below:

$$Y_{jt} = \beta_0 + \beta_1 Post_t + \beta_2 Treated_j \cdot Post_t + \mu_j + \varepsilon_{jt} \tag{1.1}$$

where Y_{jt} denotes the quality of clinical care at hospital j in period t, $Post_t$ is a dummy variable which equals 1 if post-HST period, $Treated_j$ is a dummy variable which takes value 1 if hospital j is located in a treated region, $Treated_j \cdot Post_t$ is an interaction term of $Treated_j$

[16] In the following analysis, we collapse all the pre-HST and post-HST quarters into a single pre-HST and post-HST period, respectively. If we calculate hospital-level mortality rates at the quarter level, many hospitals are left with too few admissions per period, which makes estimating adjusted mortality rates difficult.

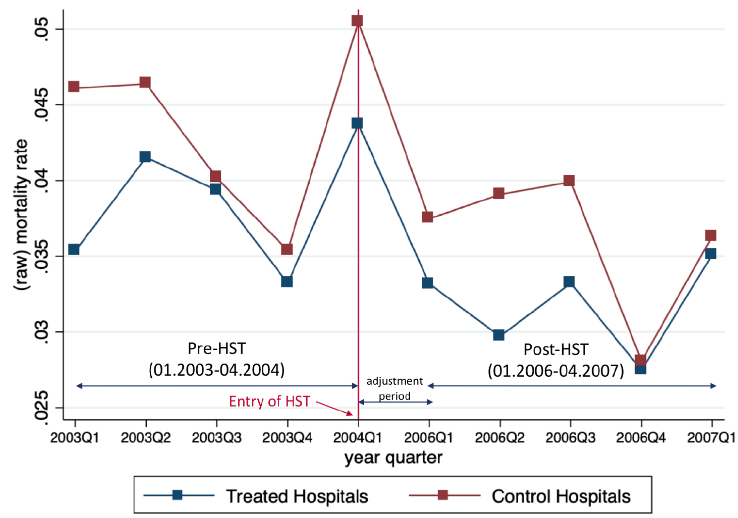


Figure 1.2: Trend of hospital-level mortality rates

and $Post_t$. We control for hospital-specific characteristics with hospital fixed effects, μ_j. Coefficient β_2 captures the impact of increased competition on Y_{jt} and is of primary interest.

1.4.2 Adjusted Mortality Rate

Using raw mortality rates as a measure of quality is problematic due to patient selection bias: severely ill patients may choose high quality hospitals. The existing literature address this selection bias by obtaining adjusted mortality rates (Gowrisankaran and Town (1999), Gowrisankaran and Town (2003), Kessler and McClellan (2000), Geweke et al. (2004), Tay (2003)). Specifically, Gowrisankaran and Town (1999) propose controlling for patients' severity of illness with an instrumental variables (IV) framework using geographic location data, i.e. distance from each patient to *all* hospitals. Although the distance to the *chosen* hospital will be correlated with the patient's severity of illness, and hence cannot be a valid instrument, where a patient chooses to live relative to *all* hospitals is uncorrelated to patient's severity of illness. This assumption is commonly used in empirical models of hospital choice, e.g. Kessler and McClellan (2000), Gowrisankaran and Town (1999), Capps et al. (2003), Gaynor and Vogt (2003), Ho (2009), Beckert et al. (2012).

In our setting, the HST facilitates long-distance travel for severely ill patients, and hence the degree of patient selection may be aggravated as a result of the entry of the HST. To allow for this change in the degree of patient selection resulting from the reduction in travel time, we use different sets of instruments in pre- and post-HST periods. We follow Gowrisankaran and Town (1999) but use travel *time* rather than travel distance from each patient to all hospitals as instruments for hospital choice. This is to account for the changes in travel time for patients living sufficiently close to the HST station in post-HST era (because even with HST, the actual distance to the hospitals does not change - what changes in the post-HST period is the travel time).

Specifically, we obtain an adjusted mortality rate by estimating a linear probability model where we regress an indicator for whether a patient dies approximately 30 days following the admission (conditional on choosing hospital j) on a set of hospital/time period dummies and patient's observed characteristics.[17] The mortality of patient i in period t is given as

$$\mu_{it} = \boldsymbol{\psi}'\boldsymbol{c}_i + \boldsymbol{\gamma}'\boldsymbol{h}_i + s_{it} + \eta_{it} \tag{1.2}$$

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 $(c_{i1pre}, ..., c_{iJpre}, c_{i1post}, ...c_{iJpost})$ where c_{ijt} equals 1 if patient i chooses hospital j in period t, $\boldsymbol{h}_i$ 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

[17]The reason for why we use a linear probability model is because it is difficult to use non-linear models in the presence of endogenous variables. Detailed explanation on this is discussed in and Town (1999).

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. We will refer to the estimated measure of quality of care, $\hat{\boldsymbol{\psi}}$ as the adjusted mortality rate. Note that we are slighltly abusing the terminology as $\hat{\boldsymbol{\psi}}$ is not adjusted mortality probabilities per se, but is the hospital's impact on patients' mortality conditional on observed characteristics. Nevertheless we use this terminology for the simplicity. Because hospital choice is likely to be correlated with patients' unobserved severity of illness, estimating equation (1.2) using OLS will lead to biased estimates. For instance, if sicker patients are more likely to choose a certain hospital j, 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) a set of dummy variables indicating whether a hospital is the closest one to a patient's location. As mentioned before, this is to account for the changes in travel time for patients living sufficiently close to the HST station in the post-HST era, and is based on the assumption that where a patient chooses to live relative to *all* hospitals is uncorrelated to her severity of illness. We define travel time for patient i to hospital j in period t as

$$\text{traveltime}_{ijt} = \begin{cases} \min(\text{cartime}_{ij}, \text{traintime}_{ij}) & \text{if } t = \text{post-HST} \\ \text{cartime}_{ij} & \text{if } t = \text{pre-HST} \end{cases} \tag{1.3}$$

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.[18] Tests of validity

[18] 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. We obtain driving time by car by using *georoute* routine developed by Weber and Peclat (2017) which calculates the driving time between two points under normal traffic conditions.

of our IV strategy and further details on estimating adjusted mortality rates are provided in Appendix B.

Having obtained the adjusted mortality rates using the instrumental variable approach outlined above, we use this measure of clinical quality to look the impact of hospital competition on the quality of clinical care using a difference-in-differences estimator.

1.4.3 Estimation Results

As a starting point to this analysis, we first estimate equation (1.1) using hospital-level raw mortality rates as a dependent variable. Table 1.4 Panel A reports the results, defining hospitals located within 15 miles of HST as treated. We implement a simple difference regression (pre vs post) in Column 1 to analyze the changes in hospital quality after the entry of the HST. The coefficient on Post dummy variable is negative ($\beta_1 = -0.00078$, interpreted as decrease in mortality rates by 0.078 percentage points) but not significant. Column 2 reports the difference-in-differences estimates. Since hospitals located near the HST station are the ones that are most affected by the entry of the HST and hence are exposed to increased competition, the estimated diff-in-diff coefficient captures the impact of increased hospital competition. We can see that the diff-in-diff coefficient is negative ($\beta_2 = -0.0042$, interpreted as decrease in mortality rates by 0.42 percentage points) but not significant.[19]

Hospital-level raw mortality rates, however, 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 essential to take into account differences in patient

[19] In Appendix A, we estimate equation (1.1) using hospital-level raw mortality rates as a dependent variable while controlling for "hospital-level case-mix". There we show that the diff-in-diff coefficient is negative and significant.

case-mix across hospitals, especially since we are using patients undergoing various types of different surgeries. In order to control for the case-mix at the patient-level, we estimate equation (1.2) using OLS, and use estimated $\hat{\psi}$ as a measure of clinical quality to estimate equation (1.1). 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 1.4, Panel A Column 3. The diff-in-diff coefficient is negative and significant ($\beta_2 = -0.014$), i.e. increased competition due to the entry of HST decreased adjusted mortality rates by 1.4 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 (1.1). The results are reported in Table 1.4, Panel A Column 4. After controlling for unobserved severity of illness, we see that the (absolute) magnitude of diff-in-diff coefficient has become larger. The diff-in-diff coefficient is -0.082 and significant, suggesting that increased hospital competition leads to an improvement in the clinical quality by approximately 8 percentage points. However, since our estimated $\hat{\psi}$ does not necessarily lie within [0,1] interval, it is difficult to directly translate $\hat{\psi}$ to "mortality rates" (since mortality rates should lie within [0,1]). Therefore, to facilitate interpretation, we rescale/normalize the IV-estimated quality of clinical care so that estimates lie within the same bounds as our raw hospital-level mortality rates. Table 1.4, Panel A Column 5 reports the diff-in-diff estimation

Table 1.4: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care

	raw mortality	raw mortality	adjusted mortality OLS	adjusted mortality IV	adjusted mortality IV (rescaled)
	(1)	(2)	(3)	(4)	(5)
Panel A: Excluding ambulance and transfer patients					
Post	-0.00078	0.0020	0.004	0.025	0.004
	(0.0031)	(0.0077)	(0.007)	(0.030)	(0.004)
Treated×Post		-0.0042	**-0.014***	**-0.082****	**-0.012****
		(0.0082)	(0.007)	(0.038)	(0.006)
Panel B: Including ambulance and transfer patients					
Post	-0.00078	0.0020	0.003	0.029	
	(0.0031)	(0.0077)	0.007	0.033	
Treated×Post		-0.0042	**-0.014****	**-0.084****	
		(0.0082)	0.007	0.040	
R-squared	0.6885	0.6892	0.6011	0.6181	0.6181
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		55	55	55	55
Treated Hospitals		112	112	112	112
Observations	334	334	334	334	334

Notes: Models are estimated using OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants. Treatment is defined as being located within 15-miles of the HST station.

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

results using these "rescaled" measures of hospital quality. After rescaling, the magnitude of the diff-in-diff coefficient is similar to those in Columns 2 and 3.

As aforementioned, our final sample excludes patients who transferred from other hospital and who arrived at a hospital via ambulance. Including these patients in our sample should not change our results because the quality of clinical care should be independent from how patiens arrived at a hospital. We include these patients in our sample and estimate equation (1.1) and report the results in Table 1.4, Panel B. Our results hold consistently, and the DID estimates are similar to those in Panel A.

1.4.4 Discussion of the Results

The results in the previous subsection 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 counterfactuals.

1.5 The Model of Hospital Choice

To evaluate the impact of the HST on patient welfare we need to look at hospital choice that patients would have made had the HST not been launched. To do this, we estimate a hospital choice model, 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 number of hospitals in a choice set for patients living close to a HST station. To capture the changes in patients' choice set in our model, we extend the basic conditional logit model by imposing travel time constraints on patients, following literature in geography and transportation. We assume that travel time to each hospital determines whether that hospital is included in patients' choice set or not. If a hospital is located too far away from a patients' location, a patient with travel-time constraints 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.

1.5.1 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 our data. The indirect utility of patient i from choosing hospital j, $j = 1, ..., J$ is defined as

$$u_{ij} = \sum_{l=1}^{L} X_{j,k} \mathbf{Y}_i^{'} \beta_{\cdot,l}^{xy} + Z_j \mathbf{Y}_i^{'} \alpha^z + D_{ij} + \mathbf{X}_j^{'} \beta^x + \alpha Z_j + \varepsilon_{ij} \tag{1.4}$$

where $\mathbf{X}_j$ is a L vector of hospital characteristics; $\mathbf{Y}_i$ is a K vector of patient-specific demographics; D_{ij} is the travel time from patient i's home to hospital j; Z_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 equation (4) using logit 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.[20] However, since our measure of quality is controlled for patient case-mix, this issue does not arise (Gaynor et al. (2013b)).

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

1.5.2 Choice Set Formation

The entry of the HST has enlarged patients' consideration sets by reducing the travel cost. Hospitals that would not previously have been considered by the patient may now be considered. We model this change consideration sets by imposing a travel-time constraint on patients. We assume that time is a limited resource that constrains the 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 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

$$Pr(y_{ij} = 1) = \int_{t=0}^{\infty} Pr(y_{ij} = 1|J_i) dP_T(t;\theta) \tag{1.5}$$

where J_{it} is a choice set of individual i who has a travel-time threshold t. Hospitals are discrete and mutually exclusive alternatives. Hence, if the hospitals are sorted according to their travel time from patient's location in ascending order, equation (13) can be simplified to a summation over all the nested sets of hospitals defined by incremental travel-time thresholds, given as

$$Pr(y_{ij} = 1) = \sum_{r=1}^{J} Pr(y_{ij} = 1|J_i r) p_T(r;\theta) \tag{1.6}$$

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

$$p_T(r;\theta) = P_T(t_{r+1};\theta) - P_T(t_r;\theta). \tag{1.7}$$

The attractive 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.

Nevertheless, the computational complexity still remains due to large 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

$$\pi_{i,r'} = \frac{\exp(\gamma_r + \mathbf{Y}_i^{'}\phi_{r'})}{\sum_l^{R_T} \exp(\gamma_l + \mathbf{Y}_i^{'}\phi_{r'})} \tag{1.8}$$

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

$$Pr(y_{ij} = 1) = \sum_{r'=1}^{R_T} Pr(y_{ij} = 1 | J_{ir'}) \pi_{i,r'} \tag{1.9}$$

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:

$$LL = \sum_{i=1}^{N} \sum_{j=1}^{J} y_{ij} \log \left(\sum_{r'=1}^{R_T} Pr(y_{ij} = 1 | J_{ir'}) \pi_{i,r'} \right). \tag{1.10}$$

1.6 Demand Estimation Results

We estimate the conditional logit model of hospital choice under travel-time constraint (ANCS-DC). The covariates that enter the utility function are as follows: "TravelTime" refers to travel time (in minutes) between the patient and a hospital in the choice set, and is defined in units of 100 minutes. Age1 is a dummy variable that equals 1 if a patient is between 25 and 50 years of age and 0 otherwise; Age2 is a dummy variable that equals 1 if a patient is between 50 and 75 years of age and 0 otherwise; Age3 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 into the lowest income group (total 10 groups); HighSeverityMainSick is a dummy variable that equals 1 if a patient is diagnosed with a disease of mortality rate greater than 0.2 and 0 otherwise; SeverityMainSick is a dummy variable that equals 1 if a patient is diagnosed with a disease of mortality rate within the range [0.1, 0.2) and 0 otherwise; HighSeveritySubSick is a dummy variable that equals 1 if a patient is diagnosed with a comorbidity of mortality rate greater than 0.2 and 0 otherwise; SeveritySubSick is

Table 1.5: Demand Model Estimates

	(1) Multinomial Logit		(2) ANCS-DC	
	Coefficient	Standard error	Coefficient	Standard error
TravelTime	-4.9798***	0.0242	-3.6825***	0.0312
Mortality	-0.6850***	0.1258	-0.6643***	0.0601
Mortality^2	-5.7107***	0.1878	-5.5142***	0.1707
Mortality×Female	0.1241	0.0939	0.1258**	0.0633
Mortality×Age[25-50)	0.0393	0.1520	0.0856	0.0762
Mortality×Age[50-75)	-0.0087	0.1406	0.0858	0.0704
Mortality×Age[75+)	0.3312**	0.1626	0.3264***	0.1023
Mortality×LowIncome	0.6514***	0.1752	0.6565***	0.1084
Mortality×HighSeverityMainSick	0.6026*	0.3108	0.6262***	0.2045
Mortality×SeverityMainSick	-0.3047*	0.1660	-0.3007***	0.0838
Mortality×HighSeveritySubSick	-0.3024	0.2866	-0.3103	0.1944
Mortality×SeveritySubSick	-0.7323***	0.1731	-0.7429***	0.1232
Mortality×HighSeveritySurgery	-0.3195**	0.1345	-0.2745***	0.0898
Mortality×SeveritySurgery	-0.6331***	0.1411	-0.6161***	0.1122
Mortality×Disabled	0.6148*	0.3573	0.6091***	0.2288
HospitalBed	0.0995***	0.0019	0.1014***	0.0019
HospitalBed×Female	0.0006	0.0014	0.0008	0.0014
HospitalBed×Age[25-50)	-0.0015	0.0023	0.0004	0.0023
HospitalBed×Age[50-75)	0.0145***	0.0021	0.0121***	0.0021
HospitalBed×Age[75+)	-0.0066***	0.0025	-0.0185***	0.0030
HospitalBed×LowIncome	-0.0195***	0.0029	-0.0180***	0.0029
HospitalBed×HighSeverityMainSick	-0.0204***	0.0051	-0.0216***	0.0051
HospitalBed×SeverityMainSick	0.0266***	0.0022	0.0258***	0.0022
HospitalBed×HighSeveritySubSick	-0.0204***	0.0046	-0.0190***	0.0050
HospitalBed×SeveritySubSick	0.0250***	0.0023	0.0250***	0.0023
HospitalBed×HighSeveritySurgery	0.0125***	0.0019	0.0136***	0.0021
HospitalBed×SeveritySurgery	0.0148***	0.0021	0.0162***	0.0021
HospitalBed×Disabled	0.0044	0.0051	0.0021	0.0053
Log-Likelihood	-162,533.49		-1.60,575.7	

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

a dummy variable that equals 1 if a patient is diagnosed with a comorbidity of mortality rate within the range [0.1, 0.2) and 0 otherwise; HighSeveritySurgery is a dummy variable that equals 1 if a patient is undergoing a surgery of mortality rate greater than 0.2 and 0 otherwise; SeveritySubSurgery is a dummy variable that equals 1 if a patient is undergoing a surgery with a mortality rate within the range [0.1, 0.2) and 0 otherwise; Disabled is a dummy variable that equals 1 if disabled with kidney and other dysfunction and 0 otherwise; HospitalBed is number of beds in a hospital, and is defined in units of 100 beds.

The estimation results are reported in Column 2 of Table 1.5. 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 coefficients suggest that patients are less likely to go to hospitals that are located further away from their home.

Our estimates suggest that patients dislike hospitals with poor clinical quality (as measured by adjusted mortality rates) and hospital quality enters patients' utility nonlinearly. We find that patients with more severe comorbidities and patients who are undergoing a more risky surgery are more sensitive to the quality of clinical care. We do not find differences in sensitivity to mortality rates between patients of different genders, ages.

Patients generally prefer larger hospitals (as measured by the number of hospital beds). Lower income patients are less likely to choose larger hospitals. Sicker patients are generally also likely to choose larger hospitals.

Table 1.7presents the estimates of the parameters of travel-time threshold probabilities. We discretize travel-time threshold into 9 points: 30, 60, 90, 120, 180, 240, 300, 360, and 420 minutes.[21] To reflect decreasing marginal disutility of time spent traveling, threshold points are 30 minutes apart (instead of 60) below 120 minutes of travel time. Several of our estimates

[21] Travel time threshold of 420 minutes includes all the hospitals in our data.

show bimodality over time constraints which makes the interpretation complicated. Patients living in metro areas are likely to have choice set to be within 30 minutes or 360 minutes. Our estimates suggest that low income patients are more likely to be time constrained in their choice. This may 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 less likely to be time constrained within 30 minutes, but are also less likely to be constrained within 360 minutes. Since older patients are more likely to be sicker (and have more time if they have retired) they may be less time constrained than younger people, and are willing to travel longer distances. At the same time, since they are older, they may experience difficulty traveling too much, resulting in a bimodal distribution. Coefficients on disease, comorbidity are ambiguous.

We also estimate the hospital choice model using conventional multinomial logit model (without travel-time constraints). The estimates of the parameters are reported in Column 1 of Table 1.5. In most respects, the signs and magnitude of the estimates are very similar to those obtained using the ANCS-DC model. We prefer to use the ANCS-DC model 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 $-2 \times (-162,533 + 160,575) = 3,916$ with 89 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.

Table 1.7: Estimates of the Time Constraint Parameters

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	30 min	60 min	90 min	120 min	180 min	240 min	300 min	360 min	360+ min
Intercept	0.637***	-1.896***	-73.211***	-76.453***	-59.211***	-57.339***	-44.548***	-1.041***	
	(0.043)	(0.196)	(13.512)	(14.423)	(10.116)	(11.374)	(11.389)	(0.093)	
Metro	13.049***	-57.128***	-17.018***	4.010***	0.071	10.436***	-106.473***	11.589***	12.153***
	(0.601)	(10.394)	(3.559)	(1.385)	(1.031)	(1.150)	(22.278)	(0.573)	(0.637)
LowIncome	4.640***	5.951***	-2.454**	5.703***	0.401	8.728***	34.438***	4.400***	5.045***
	(0.493)	(0.398)	(1.000)	(1.916)	(1.081)	(1.465)	(6.148)	(0.489)	(0.619)
Female	-9.402***	-50.635***	-23.092***	-6.811***	-2.473**	-25.380***	-43.193***	-9.604***	-9.098***
	(0.509)	(8.397)	(4.415)	(1.369)	(1.003)	(3.807)	(7.090)	(0.510)	(0.506)
Age[25-50)	-2.286***	-19.842***	-8.735***	12.740***	-1.512	-14.962***	-7.782***	-0.065	-2.984***
	(0.098)	(3.892)	(1.466)	(3.438)	(1.359)	(2.079)	(1.569)	(0.085)	(0.156)
Age[50-75)	-10.141***	-8.388***	-9.457***	-25.552***	1.445	6.720*	0.454	-7.257***	-11.122***
	(0.371)	(0.411)	(1.651)	(4.014)	(1.160)	(3.583)	(2.292)	(0.410)	(0.862)
Age[75+]	-5.605***	-3.662***	-10.121***	-4.502***	-14.309***	-19.567***	51.174***	-18.061***	-4.272***
	(0.271)	(0.613)	(1.859)	(1.033)	(2.795)	(2.817)	(13.502)	(2.950)	(0.272)
MainSickRisk	-14.407***	-24.412***	-2.845***	-5.304***	-4.123***	21.723***	27.601***	-9.224***	-16.257***
	(0.484)	(2.086)	(1.081)	(1.408)	(1.320)	(4.603)	(8.836)	(0.665)	(0.918)
SubSickRisk	-15.429***	-22.642***	4.397***	-1.116	-4.819***	4.237***	138.542***	-13.390***	-17.398***
	(0.478)	(1.730)	(1.094)	(1.000)	(1.298)	(1.284)	(30.355)	(0.801)	(0.601)
SurgeryRisk	23.418***	24.473***	-4.698***	1.543	-5.775***	28.676***	-287.276***	22.400***	24.549***
	(0.844)	(1.013)	(1.297)	(1.134)	(1.275)	(5.470)	(60.461)	(1.173)	(0.743)

Metro area corresponds to 7 metropolitan cities consisting of Seoul, Busan, Daegu, Incheon, Gwangju, Daejeon, and Ulsan. *** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

1.7 Counterfactual Analysis

Using the estimates from the demand model we evaluate the impact of the entry 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: Using pre-HST travel times and pre-HST clinical quality as a baseline, we first 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.

We then evaluate the impact of the entry of the HST on patients' health outcomes through its effect on patients' sorting to better hospitals. In other words, we are only interested in quantifying the impact of patients' sorting to better hospitals on their health outcomes (ignoring hospitals' response to greater competition). We compare the number of deaths in post-HST period to a counterfactual scenario where the train is removed while keeping hospital quality constant.

1.7.1 Changes in Patient Welfare

We compute the changes in patient welfare from the advent of the HST: changes in travel time and changes in hospital quality. Using the parameter estimates from the demand model, we simulate a post-HST scenario where the HST is removed and travel time remains the same as that of the pre-HST level. 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 post-HST travel time can be expressed as

$$E(Surplus_i^{train}) = \sum_{r=1}^{R_T} E(Surplus_{i|r}^{train}) \cdot \pi_{ir} = \sum_{r=1}^{R_T} E\left(\max_{j\in J_{i|r}^{train}} (\bar{U}_{ij} + \varepsilon_{ij})\right) \cdot \pi_{ir} \tag{1.11}$$

while in with pre-HST travel time, it is expressed as

$$E(Surplus_i^{no\ train}) = \sum_{r=1}^{R_T} E(Surplus_{i|r}^{no\ train}) \cdot \pi_{ir} = \sum_{r=1}^{R_T} E\left(\max_{j\in J_{i|r}^{no\ train}} (\bar{U}_{ij} + \varepsilon_{ij})\right) \cdot \pi_{ir}. \tag{1.12}$$

For patients living in treated regions, the choice set $J_{i|r}^{train}$ can differs from $J_{i|r}^{no\ train}$ because changes in travel time changes 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

$$E(Surplus_i^{train}) = \sum_{r=1}^{R_T} \ln\left(\sum_{j\in J_{i|r}^{train}} \exp(\bar{U}_{ij})\right) \pi_{ir} \tag{1.13}$$

and

$$E(Surplus_i^{no\,train}) = \sum_{r=1}^{R_T} \ln\left(\sum_{j\in J_{i|r}^{\text{no train}}} \exp(\bar{U}_{ij})\right) \pi_{ir} \tag{1.14}$$

The average change in surplus per patient is given as

$$E(\Delta Surplus_i) = \frac{1}{N}\sum_{i=1}^{N} E(Surplus_i^{train}) - E(Surplus_i^{no\ train}) \tag{1.15}$$

Table 1.9: Counterfactual Analysis

Panel A. Changes in Patient Welfare

		Change in travel time No change in quality	No change in travel time Chage in quality	Change in travel time Chage in quality
Treated Patients	ΔUtility	0.2815	0.0170	0.2973
	Dollar Value	$1,269	$76.82	$1,336
Control Patients	ΔUtility	0	0.0197	0.0197
	Dollar Value	0	$88.51	$88.51

Panel B. Impact of Sorting on Patient Survival (number of lives saved)

	No change in quality	Chage in quality
Treated Patients	0.2510	12
Control Patients	0	4
Total	0.2510	16

Notes: Panel A reports the changes in patient welfare in terms of expected utility (unit in expected utility) and dollar value under various counterfactual scenarios. Panel B reports the number of patients that would survive as a result of sorting to better hospitals.

where N is the number of patients in post-HST period.

We first calculate the quantity in equation (1.15) 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. To obtain these quantities, we used pre-HST travel time as well as pre-HST quality of clinical care as our baseline. The results are reported in Table 1.9, panel A. Assuming the quality of clinical care did not change, patients living in treated regions experience an average increase of 0.2815 units in expected utility. This increase in welfare arises from reduction in travel time, and the resulting ability of patients to sort to better hospitals. There is no change in welfare for patients living in control regions as they do not benefit from the entry of HST. Since there is no price coefficient in the demand model due to the absence of 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.[22] Comparing the gains in utils to the preference over distance, we find that the welfare effect of the reduction in travel distance for the treated patients corresponds to 7.6 minutes reduction in travel time.[23] 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 \$2,071 ($167 \times 7.6 = 1,269$) per patient.[24]

Next, we calculate the changes the welfare arising from changes in quality of clinical care, holding the changes in travel time constant. Patients living in treated regions experience an average increase of 0.0170 units in expected utility. Patients living in control regions experience an average increase of 0.0197 units in expected utility. The increase in expected utility for patients living in control regions arises from the fact that they face higher clinical quality although they do not benefit from the new transportation system. Applying the same back of the envelope calculation as before to monetize the gains in utils, the improvement in clinical quality yields a welfare effect of approximately \$76.82 per patient for patients living in treated regions, and \$88.51 per patient for patients living in control regions.[25]

[22] Gowrisankaran et al. (2015) estimate that a one minute reduction in travel time to hospitals increases patient surplus by \$167.

[23] $0.2815/(-3.6825) = -0.0764$, where -3.6825 is the coefficient on travel time. Travel time in the regression is defined in units of 100 minutes.

[24] Due to the travel time constraint in our model, the number of hospitals that a patient considers changes when travel time to each hospital changes. Increased number of hospitals will affect the welfare gains because the term in parentheses in equations 1.13 and 1.14 is simply the denominator of the logit choice probability (which is simply the outcomes of the mathematical form of the extreme value distribution, and has no economic meaning (Train (2009)). Therefore, we also calculate changes in welfare arising from reduced travel time (holding hospital quality constant) while holding the number of hospitals in the consideration set constant.

[25] The welfare effect of the improvement in clinical care for the treated patients corresponds to approximately 0.46 minutes reduction in travel time, $0.0170/(-3.6825) = -0.0046$. Multiplying this by the value per minute reduction in time, we get $0.46 \times 167 = 76.82$. Similarly, the welfare effect of the improvement in clinical care for the control patients corresponds to approximately 0.53 minutes reduction in travel time. Multiplying this by the value per minute reduction in time, we get $0.53 \times 167 = 88.51$.

Finally, we calculate the changes the welfare arising from both, changes in travel time and changes in quality of clinical care. Patients living in treated regions experience an average increase of 0.2973 units in expected utility. Patients living in control regions experience an average increase of 0.0197 units in expected utility (identical to the case when quality of clinical care changes, holding the changes in travel time constant). This yields a welfare effect of approximately $1,336 per patient for patients living in treated regions, and $88.51 per patient for patients living in control regions.[26]

1.7.2 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:

$$E(\Delta Mortality) = \sum_i \left[E(Mortality_i)^{\text{train}} - E(Mortality_i)^{\text{no train}}\right] \tag{1.16}$$

[26]The welfare effect of the improvement in clinical care for the treated patients corresponds to approximately 8 minutes reduction in travel time, $0.2973/(-3.6825) = -0.0807$. Multiplying this by the value per minute reduction in time, we get $8 \times 167 = 1,336$.

where

$$E(Mortality_i)^{\text{train}} = \sum_j Pr_{ij}^{\text{train}} \cdot Prob(Mortality_i | choice = j, Health_i) \tag{1.17}$$

and

$$E(Mortality_i)^{\text{no train}} = \sum_j Pr_{ij}^{\text{no train}} \cdot Prob(Mortality_i | choice = j, Health_i). \tag{1.18}$$

Equations (1.17) and (1.18) denote the mortality probability with post-HST travel time and pre-HST travel time, respectively. $Mortality_i$ is an indicator variable which takes value 1 if the patient dies and 0 otherwise. Term $Prob(Mortality_i | choice = j, Health_i)$ that appears in both equations is the predicted probability of death conditional on choice of hospital and patient's health status. Since we estimated adjusted mortality rates and the coefficients on patient case-mix using linear probability model, the predicted mortality probability may not necessarily lie within the (0,1) interval. Therefore, we obtain the predicted probability of death using the Linear Discriminant Model (LDM) method which transforms the coefficients from the linear probability model into maximum likelihood estimates of the parameters of a linear discriminant model. [27] The LDM implies a logistic regression model for the dependence of the outcome on the predictors. This method ensures that the predicted probabilities lie within the (0,1) interval.

The results are reported in Table 1.9, panel B. Our estimates from this counterfactual analysis suggest that 0.25 more lives can be saved from patients' sorting. Since our data is a 2 percent random sample of the entire population, this translates to approximately 12 lives over the five quarters, which is equivalent to 10 lives on an annual basis.[28]

[27] https://statisticalhorizons.com/better-predicted-probabilities

[28] $0.25 \times 50 \times (4/5) = 10$

Next, we calculate how many more lives are saved due to patient sorting when the quality of clinical care also responds to the entry of HST. Our estimates suggest that 12 lives (480 lives on an annual basis) of patients living in treated regions and 4 lives (160 lives) of patients living in control regions can be saved.[29]

1.8 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 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 estimator to examine the impact of competition on health outcomes measured by 30-day mortality rates following admissions for cardiovascular or neurological 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 the reduced travel time, they benefit from the improvement in the quality of treated hospitals. We also

[29] $12 \times 50 \times (4/5) = 480$ and $4 \times 50 \times (4/5) = 160$

find that HST has 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 patient health by facilitating patients' sorting to better hospitals through lower travel costs.

1.9 References

Banerjee, A., E. Duflo, and N. Qian (2012): "On the Road: Access to Transportation Infrastructure and Economic Growth in China," *NBER Working Paper* No. 17897.

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.

Brekke, K. R., L. Siciliani, and O. R. Straume (2011): "Hospital Competition and Quality with Regulated Prices," *The Scandinavian Journal of Economics*, 113, 444–469.

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

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

Colla, C., J. Bynum, A. Austin, and J. Skinner (2016): "Hospital Competition, Quality, and Expenditures in the U.S. Medicare Population," *NBER Working Paper No.* 22826.

Dafny, L., K. Ho, and M. Varela (2013): "Let Them Have Choice: Gains from Shifting Away from Employer-Sponsored Health Insurance and toward an Individual Exchange," *American Economic Journal: Economic Policy*, 5, 32–58.

Donaldson, D. (2018): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 108, 899–934.

Gaynor, M., R. Moreno-Serra, and C. Propper (2013a): "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 (2013b): "Free to Choose? Reform and Demand Response in the English National Health Service," Working Paper.

——— (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 (2004): "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.

Heuermann, D. F. and J. F. Schmieder (2018): "The Effect of Infrastructure on Worker Mobility: Evidence from High-Speed Rail Expansion in Germany," *NBER Working Paper* No. 24507.

Ho, K. (2006): "The welfare effects of restricted hospital choice in the US medical care market," *Journal of Applied Econometrics*, 21, 1039–1079.

——— (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*.

Lewis, M. S. and K. E. Pflum (2017): “Competition and Quality Choice in Hospital Markets,” Working Paper.

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.

Qin, M., G. John, and M. A. Vitorino (2018): “Planes, Trains and Co-Opetition: Evidence from China,” Working Paper.

Qin, Y. (2016): “"No county left behind?" The distributional impact of high-speed rail upgrades in China,” *Journal of Economic Geography*, 17, 489–520.

Shen, Y. C. (2003): “The effect of financial pressure on the quality of care in hospitals,” *Journal of Health Economics*, 22, 243–269.

Shortell, S. M. and E. F. Hughes (1988): “The Effects of Regulation, Competition, and Ownership on Mortality Rates Among Hospital Inpatients,” *New England Journal of Medicine*, 318, 1100–07.

Silber, J. H., P. R. Rosenbaum, J. Tanguy, R. N. Ross, L. J. Bressler, O. EvenShoshan, 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.

Train, K. (2009): Discrete Choice Methods with Simulation, *Cambridge University Press*, 2 ed.

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

Chapter 2

Search Frictions, Sorting and Matching in Two-Sided Markets

2.1 Introduction

In many two-sided marketplaces agents search for potential partners to form a match based on a mutual agreement.[30] Since agents on both sides of the market have preferences regarding each others' characteristics, and because these preferences are often private, agents on each side do not know which counterparts are willing to match. Forming a match based on mutual compatibility in the presence of private preferences, therefore, generally requires extensive costly search.

[30]Examples can be found across a wide range of industries, such as a marriage and dating market, college admission, online labor service (Taskrabbit and Upwork) and hospitality (Airbnb). Fradkin (2015) shows that more than 40 percent of booking inquiries on AirBnB platform are rejected. Among those, approximately 14 percent of rejections are driven by hosts' preferences concerning the characteristics of the searcher or the trip.

The question of who matches with whom had been a central question in the matching literature. In the presence of costly search, both preferences and search frictions shape the formation of a match. Understanding the relative impacts of these two forces on match outcomes is not only theoretically important but also managerially relevant because the design of online two-sided platforms should vary depending on whether match outcomes are primarily a result of preferences or search frictions.

In this paper, we try to obtain a better understanding of the relative impact of preferences and search frictions on match outcomes using data from an online dating platform *MonCherie* (name disguised per the request of the data provider).[31] In particular, the goal of this paper is twofold: First, we disentangle the relative impact of preference and frictions on assortative matching (i.e., sorting), a widely observed phenomenon where couples display resemblance across various characteristics such as age, education level, ethnicity, and income, etc. Sorting in married couples has been an important topic of study as it may have long-term impact on economic development and inequality through its impact on the outcomes of children and accumulation of human capital (Raquel (2003); Raquel and Rogerson (2001)).[32] Second, we quantify the departure from efficiency caused by frictions in the platform. This will offer insights on the gains that users can achieve when search frictions are reduced through a better design of the platform.

More specifically to assortative matching, there are two distinct explanations for assortative matching (Hitsch et al. (2010a)). One explanation is that sorting is an equilibrium outcome

[31] Although the primary reason for the existence of online dating platforms is to make the search for a partner as easy as possible, search frictions nevertheless still exist in these platforms: First, due to a large number of participants on the platform, users cannot consider all available profiles. Second, private preferences of agents on the other side of the market create uncertainty about whether a match will be achieved which may lead to misdirected efforts and sub-optimal matches. Therefore, an online dating platform provides an ideal environment for us to study the relative contribution of search frictions and preferenes on matching outcomes.

[32] With one-third of marriages in the United States beginning online (Cacioppo et al. (2013)), online dating has become the most popular way for couples in U.S. to meet.

driven by agents' preferences. For example, if mate preferences are "horizontal", people may prefer to match with a similar partner, which in turn results in sorting. Alternatively, if mate preferences are "vertical" in the sense that everyone ranks potential partners using the same criterion, then the ranks of matched partners will be positively correlated. In this case, couples will display sorting along attributes that are monotonically related with these ranks. Alternative explanation is that search frictions influence how couples meet, irrespective of preferences. For example, in *offline* dating markets, same ethnicity in couples might not necessarily be due to the preference regarding ethnicity, but instead might reflect the fact that people are constrained in their choices due to where they live. In reality, *both* preferences and search frictions affect how couples meet.

In an online dating context, the type of search cost can vary depending on how the platform is designed or how users conduct search. Typically, users on online dating platforms decide to make offers (e.g. send a message) based on the expected utility of a match given the probability that the offer will be accepted. Therefore, if two potential partners who differ in their characteristics yield identical utility, offer is more likely to be made to the one who is ex-ante more likely to accept it. As discussed earlier, since preferences do affect partner choice to a certain degree, partners who are similar to the focal user are ex-ante more likely to accept the offer than those who are very dissimilar. Therefore, unless a user searches for additional information that helps to make a more accurate prediction about the match probability, the offer is more likely to be sent to the one who is more similar, which in turn will result in sorting. Therefore, while sorting on online platforms may partly be due to users' preference for similar others, it is also influenced by search frictions because users with high search costs will not search for additional information, making decisions only based on the limited information that is provided in the default setting.

The unique feature of our data that helps us to answer our research questions is that they were generated from a field experiment conducted by MonCherie. The treatment of the field experiment is to provide a piece of information about the preference of the opposite side to randomly selected users, thereby reducing search friction for these users. Many online dating platforms typically give users the option to ℓike or not ℓike a profile.[33] In a default setting (control group), users do not know whether or not the person in the profile had $\ell iked$ them. The only way for the user to find out is by $\ell iking$ the profile: If the focal user $\ell ikes$ a profile, and if the person in the profile had also already $\ell iked$ him, both users get a notification about the mutual $\ell iking$. If no notification appears upon $\ell iking$ a profile, it implies that the person in the profile "did not ℓike" the focal user.[34]

$\ell iking$ a profile is costly. While the simple act of clicking a button or swiping the screen may be costless, a user faces the risk of having one's offer of affection rejected if the other doesn't ℓike back, which can hurt one's ego (for a comprehensive review on rejection, see Baumesiter and Dhave (2001) and Baumeister et al. (1993)). Since users in the control group have to click the ℓike button to find out if they were $\ell iked$ by the person in the profile, they have to engage in costly search to find out this information.

The experiment allowed randomly selected users (treatment group) to know *upfront* whether the person in the profile had $\ell iked$ them, without the need to ℓike a profile in order to find out this information. With the majority of initiated messages not receiving a response, knowing whether someone had $\ell iked$ them allows users to gauge the likelihood of getting a match more precisely. Since users in the treatment group know this piece of information

[33]Or swipe right (like) or left (not like) on a mobile device

[34]"did not ℓike" can happen if (i) the person in the profile had browsed the focal user's profile, but decided not to ℓike, or (ii) the person in the profile has not browsed the focal user's profile yet (i.e. the algorithm of MonCherie has not yet displayed the focal user's profile) and therefore had not had an oportunity to decide whether or not to ℓike the focal user. The focal user is not able to distinguish between these two causes.

upfront without having to take any further action (unlike users in the control group who have to engage in costly search to find out this information, the treatment reduces search frictions.

In the data, we find descriptive evidence suggesting that the treatment leads to less sorting between matched couples across various dimensions, i.e. treatment makes users to match with partners who are more dissimilar from themselves. Specifically, we find that users in the treatment group who matched with those who had *liked* them display significantly lower correlation in attributes (age, education level, Body-Mass-Index (BMI), ethnicity and attractiveness) with their matched partners compared to their control counterparts.[35] Since the treatment reduces search frictions by letting users know the information about who *liked* them without having to search for it, the descriptive patterns we see in the data suggest that reducing search frictions may lead to lower sorting between couples. The mechanism behind this pattern, according to our data, is that the treatment encourages users to initiate a conversation with potential partners who had *liked* them *and* who are dissimilar from themselves. That is, when *liked* by a dissimilar partner, users in the control group know this and hence initiate a conversation with them. Users in the control group, on the other hand, do now know this information (unless they engage in costly search), and therefore do not initiate a conversation with them due to the ex-ante low probability of receiving a reply.

Despite the reduction in uncertainty about the preferences of the other side, users still face uncertainty about whether a match will be formed because the *like* from a potential partner does not guarantee a match.[36] If the cost of initiating a contact is non-negligible, the decision to initiate a contact depends on the probability of a match. A user may decide to forgo contacting a desirable partner and save on the costs if the expected probability of a match is sufficiently low. However, if the cost of initiating a contact were non-existent, the decision to

[35] We only consider matches that were *initiated* by the focal user.

[36] i may have *liked* j because the utility from matching with j is higher than the utility of staying single. Nevertheless, i may reject j's offer if someone more preferable than j makes him an offer.

contact a potential partner would only depend on the utility of a match. Hence, in the latter case, only preferences shape the formation of a match.

To achieve our research objective, we model the decision process of a user who is considering whether to search for and to contact a potential partner of the opposite gender. The model incorporates preference heterogeneity across users and also allows for costly search as well as costly initial contacting. Browsing through each profile to obtain detailed information about a potential partner is costly because it takes time and cognitive effort to process the information. Thus, the user browses only a fixed number of profiles that maximizes the sum of the expected utility net of browsing costs, which represents his consideration set. For each profile in his consideration set, the user first has to decide whether to take a costly action of $\ell iking$ a profile. In our model, some users know upfront whether the potential partner had $\ell iked$ them (treatment group). The rest of the users do not know whether the potential partner had $\ell iked$ them (control group) and therefore have to search for this information through $\ell iking$ a profile. In addition to revealing the information about the $\ell ikes$ for the control group, $\ell iking$ a profile serves an additional purpose for users in both groups, namely signaling interest to the potential partner.[37] Upon his decision to ℓike (or not ℓike), the user then has to decide whether to initiate a conversation by sending a costly message. The resulting framework allows us to model users' decision to search for- and to contact potential partners and how these decisions are related to their preferences and costs.

Identification of search models is difficult due to the interdependence between search costs and preferences. Correspondingly, we rely on variation in information sets caused by the experiment as well as the exclusion restriction to separately identify preference from costs.

[37]The potential partner j may find out whether the focal user i had $\ell iked$ him by $\ell iking$ i's profile. Upon finding out that i had $\ell iked$ him, j may choose to message i.

Our estimation results reveal that search costs play a significant role in shaping the outcomes on the platform.

Based on the preference and cost estimates, we predict who matches with whom in a frictionless environment where only preferences shape the matching outcomes. Predicted matches in a frictionless environment are simulated using the deferred-acceptance algorithm of Gale and Shapley (1962). Note that Gale-Shapley mechanism assumes the presence of a central matchmaker who matches the agents given individual preferences over potential partners, and hence does not describe the decentralized search process of the online dating platform. Adachi (2003) shows, however, that as frictions disappear, the set of equilibrium outcomes in a decentralized search model reduces to the set of stable matchings in a corresponding Gale-Shapley problem. Moreover, repeated rounds of offer-making and corresponding rejections of the deferred-acceptance algorithm resemble the behavior of the users on the dating platform.

Our results reveal that frictions play a significant role in shaping assortative matching patterns: complete removal of frictions leads to a significantly lower level of sorting (lower attribute correlations between matched couples). Specifically, we find that removing frictions reduce age correlation by approximately 14 percent, education correlation level by 42 percent, and attractiveness correlation by 30 percent, and the proportion of users who match with partners of other ethnicity increases by roughly 13 percent, compared to the outcomes achieved in the default setting.

We then turn to the question of efficiency. Since the stable matching predicted by the Gale-Shapley algorithm is also Pareto-optimal, it serves as an efficiency benchmark. Using this benchmark, we quantify the departure from efficiency caused by frictions on the platform. We assign ordinal rankings to each matched partner (for both men and women) based on estimated preference parameters and compare the average rankings achieved across different protocols. We find that removing frictions improves the average ranking of the partner by approximately 7 percent compared to the default (control) setting. These numbers suggest that removing frictions significantly improves on the outcomes achieved in an environment with frictions,.

The rest of this paper is structured as follows. In the following section we review the related literature and its relevance to this paper. Section 2.3 describes the institutional details of our online dating platform. Section 2.4 describes the details of the experimental design. Section 2.6 summarizes the data and presents descriptive evidence suggesting that reducing frictions may lead to less sorting. In Section 2.7 we propose a model of search and choice for a partner. Estimation details and results are discussed in Section 2.8 and Section 2.9, respectively. Sections 2.10 and 2.11 present counterfactual exercises. Finally, Section 2.12 concludes.

2.2 Related Literature

This paper is closely related to recent literature that have studied frictions in online markets. Fradkin (2015) looks at search frictions in an online apartment rental market, Airbnb. He shows that on Airbnb, even after a buyer identifies an apartment of interest, many of the transactions can fail because the seller may reject the buyer, or because multiple buyers may contact the seller at the same time.[38] He studies how ranking algorithms can increase the efficiency of the platform in terms of greater number of matches. Horton (2014) shows that failed transactions due to information frictions is also common in online labor markets wherein employers inefficiently pursue oversubscribed workers, and studies how the platform can optimally allocate employers' attention to workers. Dinerstein et al (2018) compare shopping behavior and price competition on eBay under an alternative platform design, and

[38]This phenomenon is termed as " congestion".

show that guiding consumers toward a price ranking can lead to a higher surplus. While these papers are similar to ours in that they study frictions in two-sided online markets, they do not study assortative matching which is the main focus of our paper.[39]. In many markets such as a market for new physicians/law graduates, and college admissions, early transactions (i.e. "unraveling") can be problematic. Unraveling may, at least ex-post, be inefficient if information that is important for determining match quality evolves over time. In such markets, transactions arranged before critical information becomes available will not be able to achieve matchings that are as efficient as those that could be achieved after the information becomes available. Frechette et al. (2008) show that unraveling can create inefficient matching outcomes in the market for post-season college football games.

Our paper is also related to empirical work that estimates mate preferences using data on marriages or romantic relationships (Wong (2003); Choo and Siow (2006); Flinn and Del Boca (2012); Chan et al. (2015); Richards Shubik (2015)). These papers fit a structural model of equilibrium match formation in which preferences are parameterized to observed match outcomes. Wong (2003) estimates an equilibrium two-sided search model to explain marriage outcomes in the Panel Study of Income Dynamics (PSID). Choo and Siow (2006) estimate a frictionless tranferrable utility matching model. Arcidiacono et al. (2016) estimate a two-sided directed search model of romantic relationship formation and show that individuals direct their search for a partner based on partners' characteristics, endogenously determined probability of matching, and the terms of a relationship (i.e. whether sex is included in a relationship). While these paper use observed matches to estimate mate preferences, our data allows us to observe each user's entire search process. This enables us to estimate preferences based on users' decision to initiate a conversation, independent of whether an actual match was formed or not.

[39]Related research on information frictions and market inefficiency is a paper by Frechette et al. (2008)

More recently, there has been an increase in number of papers using data on speed-dating and online dating. Kurzban and Weeden (2005), Fisman et al. (2006) and Fisman et al. (2008) use data from "speed-dating" to study preferences for mates. Lee (2015) finds that online dating promotes marriages that exhibit weaker sorting along occupation and geographic proximity, but stronger sorting along education and other demographic traits. Lee and Niederle (2015) study the effect of preference signaling by attaching one of a limited number of virtual roses to a date request in a major Korean online dating website, and find that users are more likely to accept a date request when a virtual rose is attached.[40] Bapna et al. (2016) study how anonymous browsing affects user behavior and Bapna et al. (2019) study how the vote identity revelation affects user behavior in an online dating platform. More recently, Fong (2018) studied search and matching behavior in an online dating application, focusing on how users respond to market thickness. Bojd and Yoganarasimhan (2019) study the causal effect of popularity information in online dating and find evidence of strategic shading due to fear of rejection. Research in this area that is most close to ours is that by Hitsch et al. (2010a) and Banerjee et al. (2013). Using data from an online dating website, Hitsch et al. (2010b) study the efficiency of matches obtained in online dating markets and find that the matches predicted by the economic model (Gale-Shapley deferred-acceptance algorithm) are similar to the actual matches achieved in the dating website, suggesting that the matches achieved in their dating website are approximately efficient. They are also able to largely predict the assortative matching patterns observed in the matches, which suggests that assortative matching can arise in the absence of search frictions, primarily due to preference and market mechanism. Using similar approach, Banerjee et al., (2013) study how preferences for caste can affect equilibrium patterns of matching. They find that there is a very strong

[40] In several online dating websites, participants are allowed to send signals to potential partners. For example, in an online dating platform Tinder, users can send one "super like" each day to signal interest. In a matchmaking engine of the rating website "Hot or Not", participants can send each other costly virtual flowers which, according to the website, increase the chances of receiving a positive response.

preference for within-caste marriages, and also show that in equilibrium, ignoring caste-related preferences does not alter the matching patterns on non-caste attributes. Compared to these papers, our experiment enables us to detect a source of search friction which motivates us to disentangle the contribution of friction and preferences on the equilibrium sorting.

2.3 Institutional Details

MonCherie is a typical freemium community where most of the users sign up for a free account which allows them to use all the basic features to interact with other members of the platform. These basic features include browsing, clicking the profile for more details (henceforth click), *liking* and messaging. By paying a monthly subscription fee, users can also upgrade to a premium account which consists of a fixed bundle of premium features. Among other incremental features, the premium bundle includes the ability to know whether the person in the profile had *liked* the focal user.[41]

MonCherie is operated on both a website and a mobile app. The design of the website is somewhat different from the mobile app. Although the experiment was conducted on randomly selected users who use either the website or the mobile app, in this paper we only focus on the mobile app users because the design of the mobile app is much simpler, allowing the analysis to be more tractable. We first describe how the mobile app works for a user in a default setting (i.e. user in the control group with a non-premium account), and then proceed to describe how the experiment changes the operation of the app for users in the treatment group in the following section.

When a non-premium user in the control group opens the mobile app, a random profile is displayed to him based on the internal algorithm of MonCherie. Figure 1 illustrates what is

[41] We only sample from the users with a non-premium account to avoid selection bias.

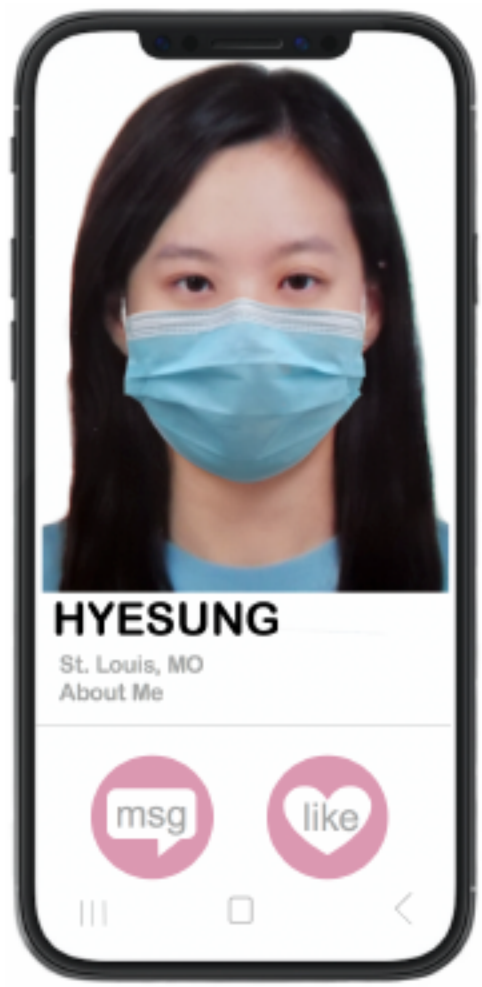


Figure 2.1: (Example) Profile displayed to the control group

(a) She did not *ℓike* you (b) She *ℓiked* you

Notes. This figure shows how *ℓiking* a profile reveals the information about whether the potential partner had *ℓiked* the focal user or not. (For illustrative purpose only. The image may be different on the actual app). The *ℓike* button turns red when the user chooses to *ℓike*. If the potential partner had *ℓiked* the focal user, both users will receive a notification about the mutual *ℓiking*. In addition to a notification, a heart icon appears on the top right-hand corner. If the potential partner had not *ℓiked* the focal user, neither a notification nor a heart icon appears.

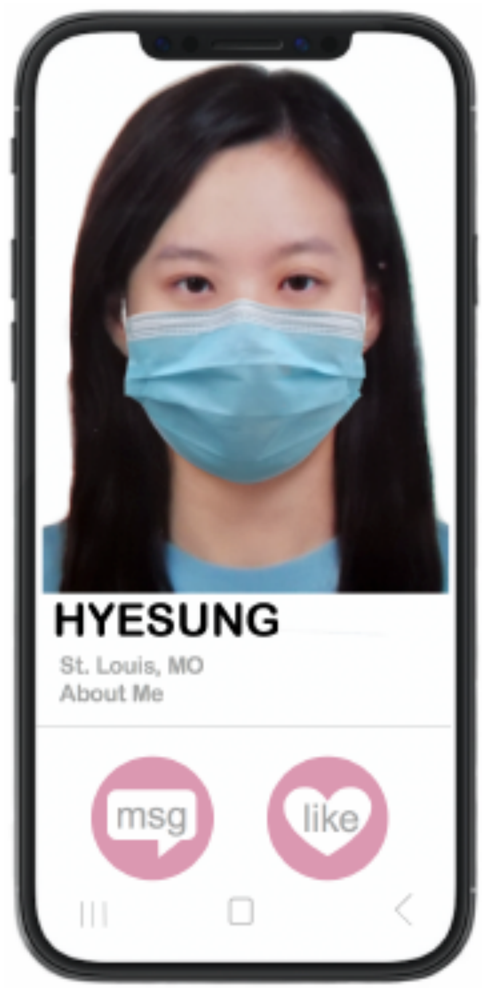
Figure 2.2: How *ℓiking* a profile reveals information

displayed to a user. A user is able to see the profile picture and detailed characteristics such as age, ethnicity, education level, etc. Upon browsing the profile, a user can choose to *ℓike* and/or message.

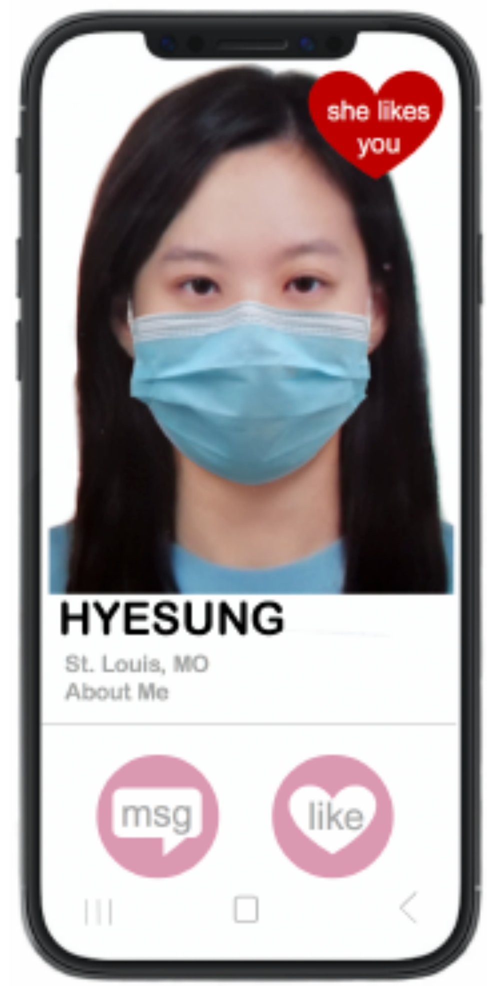
A user can choose to *ℓike* by either clicking a *ℓike* button (or by swiping right). Similarly, a user can choose to not *ℓike* by simply not clicking the *ℓike* button (or by swiping left). If the user *ℓikes* a profile, and if the person in the profile had already *ℓiked* the focal user, then both users receive a pop-up notification about the mutual *ℓiking* (Figure 2b). In addition to the pop-up notification about the mutual *ℓiking*, an icon appears next to the profile picture (heart icon on the upper right-hand corner) indicating that the person in the profile had *ℓiked* the focal user.[42] Therefore, *ℓiking* a profile reveals whether or not the person in the profile had *ℓiked* the focal user. If no notification nor a heart icon appears upon choosing to *ℓike*, it implies that the person in the profile had "not *ℓiked*" the focal user (Figure 2a). This can happen if either (i) the person in the profile had browsed the focal user's profile but decided not to *ℓike*, or (ii) the person in the profile hasn't browsed the focal user's profile yet (i.e. the algorithm of MonCherie hasn't yet displayed the focal user's profile to the person in the profile) and therefore hadn't had an opportunity to decide whether or not to *ℓike* the focal user. The focal user is not able to distinguish between these two causes of "not *ℓikes*".

A new profile is displayed to a user immediately after a user sends a message or clicks the "back" button on the phone. Also, if a user *swipes* in either direction as opposed to *clicking* the *ℓike* button, a new profile will be displayed to him unless a mutual *ℓiking* is reached, in which case Figure 2b is displayed and the user can decide whether to send a message. On the other hand, if a user clicks the *ℓike* button instead of *swiping* right, he can continue browsing the current profile (in this case, user sees Figure 2a if he hadn't been *ℓiked* and sees Figure 2b if he had been *ℓiked*) and has to decide whether to send a message.

[42] The heart icon is for illustrative purpose only. A different icon may appear in the actual app.



(a) she did not *like* you



(b) she *liked* you

Figure 2.3: Profile displayed to the treatment group

2.4 The Experiment

In this section we describe the design of the experiment and how it changed the operation of the app for the treatment group. The experiment was conducted on 100,000 newly registered random users of MonCherie over three consecutive months.[43] The three months are referred to

[43]The target population of 100,000 experimental users was randomly selected among newly registered users during a seven-day period in 2016. These users account for approximately less than 1 percent of the entire population of users of MonCherie as of 2019.

as the pre-treatment period (1st month), treatment period (2nd month), and post-treatment period (3rd month).[44] On the first day of the treatment period (2nd month), randomly selected 50,000 users received the following email:

> Hey username,
> You have been randomly selected to receive a super power - for the next 30 days, we're giving you the ability to know whether someone had liked you! Normally this feature is restricted to paid premium users only. Enjoy![45]

The remaining 50,000 users who serve as our control group received the following email:

> Hey username,
> It's a good time to visit MonCherie! We've got a bunch of matches lined up just for you. Enjoy!

Because the treatment was endowed on users by the platform without any required action from the users' side, users were unaware of being part of an experiment. Therefore, the observer bias is not applicable.

Figure 3 illustrates what is displayed to a user in the treatment group when he opens the app. When a profile is displayed to a user, he can immediately see whether the person in the profile had *liked* him or not. If the person in the profile had *liked* the focal user, a heart icon appears on the upper right-hand corner (Figure 3b). If the person in the profile had not *liked* the focal user, a heart icon is absent (Figure 3a). As aforementioned, a user is unable to distinguish whether the person in the profile had browsed his profile and decided not to *like*, or whether his profile was not browsed yet.

[44] The gift of treatment expired after 30 days.

[45] To disguise the identity of MonCherie (and the terminologies specific to the platform), the messages presented in the paper are slighly modified from the actual messages that were sent to users.

2.5 Data Description

For each of the 100,000 users in our experiment, we observe time-stamped actions (browsing, clicking, $\ell iking$, and messaging) over the three months. As explained earlier, the three months are pre-treatment period (1st month), treatment period (2nd month), and post-treatment period (3rd month). For each user, we have several self-reported demographic variables, such as gender, age, education level, ethnicity, body type. We also have information on users' sexual orientation (heterosexual, homosexual, bisexual).[46] We also observe time-stamped actions and demographics for all users who have interacted with the experimental users in any way (henceforth, correspondent users).[47] The data on correspondent users' $\ell iking$ behavior allows us to observe who had $\ell iked$ the focal experimental user. In addition, we observe as to whether a user was using a desktop or a mobile app, whether a user has a premium account and whether the account is valid (whether or not the user is a spammer/bot as determined by internal algorithm of MonCherie).

Out of our initial sample of 100,000 experimental users, we limit our sample to heterosexual users who browsed at least one profile during the treatment period.[48] We also limit our sample to mobile app users because the design of the mobile app is much simpler than the desktop version, which renders the analysis more trackable. We drop users with a premium account to avoid selection bias and we also drop users with a non-valid account. This leaves us with 16,119 male (7,930 treated, 8,189 control) and 7,112 female (3,470 treated, 3,642 control) experimental users. In our data, only 0.59 percent of the experimental users have interacted with other members of the experiment, and only 0.4 percent of the users in the

[46] Only a few demographic variables were provided to us due to privacy concerns.

[47] Two users i and j have interacted with each other if either i or j (or both i and j) browsed and/or $\ell iked$ and/or messaged the other person

[48] A large number of users became inactive a few days after their first use of the app.

treatment group have interacted with other members in the treatment group. Hence, we do not worry about the contamination bias.

Table 2.1 summarizes characteristics of users in our sample, separately for men and women.[49] We have self-reported information on users' age, education level, body type and ethnicity. Consistent with existing research that use data from other online dating services, there are more men than women (16,119 men, 7,112 women). Men are on average 31 years old and women are on average 34 years old. Among men and women who reported their education level, approximately 55 percent received their final degree from a university, and less than 20 percent have a postgraduate degree. The majority of users on the platform are White (approximately $60 - 66$ percent), followed by Hispanic (13 percent), Asian ($10 - 15$ percent) and Black ($10 - 12$ percent). Test of randomization of the treatment is reported in Appendix F.

[49]This table summarizes demographic characteristics of the experimental users only, not correspondent users.

Variable	Men				Women			
	Mean	SD	Median	Obs	Mean	SD	Median	Obs
Age	31.3	9.6	29	16,119	34.0	11.3	31	7,112
HighSchool	0.13	0.3	0	5,178	0.09	0.3	0	2,482
TwoYear	0.18	0.4	0	5,178	0.16	0.4	0	2,482
University	0.54	0.5	1	5,178	0.57	0.5	1	2,482
PostGrad	0.15	0.4	0	5,178	0.18	0.4	0	2,482
Skinny	0.14	0.4	0	4,541	0.25	0.4	0	1,648
Average	0.67	0.5	1	4,541	0.54	0.5	1	1,648
LittleExtra	0.15	0.4	0	4,541	0.16	0.4	0	1,648
Overweight	0.04	0.2	0	4,541	0.05	0.2	0	1,648
Asian	0.10	0.3	0	7,137	0.15	0.4	0	3,638
White	0.66	0.5	1	7,137	0.60	0.5	1	3,638
Black	0.10	0.3	0	7,137	0.12	0.3	0	3,638
Indian	0.04	0.2	0	7,137	0.02	0.1	0	3,638
MidEastern	0.03	0.2	0	7,137	0.01	0.1	0	3,638
Hispanic	0.13	0.3	0	7,137	0.13	0.3	0	3,638
NativeAmerican	0.02	0.2	0	7,137	0.02	0.1	0	3,638
PacificIslander	0.01	0.1	0	7,137	0.01	0.1	0	3,638

Notes. Many users choose not to report some of their demographic information, which leads to different number of observations for each demographic variable. In the data (prior to selecting our sample), we also have users with the following education levels: LawSchool, MedSchool, and PhD. In our final sample, although we do have correspondent users with these education levels, we do not have experimental users with these education levels. This is because experimental users with these education levels were dropped during our sample selection process (browsed at least one profile, heterosexual, valid, non-premium, mobile app users). Note that Indian and Asian are separately listed in this table. Although Indian is Asian, the data we received from MonCherie had listed them separately. Therefore, we also list them separately.

Table 2.1: Summary Statistics of Users Characteristics

2.6 Descriptive Statistics

2.6.1 Impact of the Treatment on User Activities

We first proceed by showing the effect of the treatment on user activities. From the perspective of the focal user, there are two types of potential partners: (i) potential partners who had *ℓiked* the focal user (henceforth "Likers"), and (ii) potential partners who did not *ℓike* the focal user (henceforth "NotLikers"). Since the treatment allows users to know whether the potential partner had *ℓiked* him or not without taking any further action, it is natural to think that the treatment would have differential impact on users' activities depending on whether the person in the profile is a Liker or a NotLiker.

In Table 2.2 we present summary statistics of user activities towards Likers, separately for men and women. Column 1a (2a) summarizes activities of men (women) in the control group and column 1b (2b) summarizes activities of men (women) in the treatment group. We also report t-statistics (columns 1c and 2c) to show if there are any significant differences between the two groups. For men, the treatment increases the profile clicks by 42.5 percent, increases *ℓikes* sent by 26 percent, and increases messages sent by 30 percent. We see a similar pattern for women: treatment increases profile clicks by 55.5 percent, increases *ℓikes* sent by 69.6 percent and increases messages sent by 37.9 percent.

We also test whether the treatment has an impact on the number of successful matches achieved by our experimental users. While we do not observe whether users actually went on an offline date, nor do we observe the actual content of the messages exchanged between users, we do observe the number of messages exchanged. Prior research by Bapna2016 and anecdotal evidence from the online dating industry has pointed out that exchange of three messages between potential couples is a good predictor of an actual online match, in which phone numbers are exchanged or mates are asked out for an offline date. In fact, senior executives of MonCherie revealed that they strongly believe that this measure of a match is an accurate predictor of an offline date. Moreover, despite knowing the exact content of users' messages, MonCherie uses this metric as a measure of a successful match for their own internal recommendation engine. We take a more conservative stance and define a successful match as an exchange of at least six messages. [50]Here we only consider "initiated" matches, where at least 6 messages were exchanged upon the experimental user starting a conversation. We find that the treatment increases men's initiated matches with Likers by 50 percent and increases women's initiated matches with Likers by 33.3 percent.

Table 2.3 presents summary statistics of user activities towards NotLikers. Except for the reduction in the number of *ℓikes* sent by men in the treatment group, the treatment does not lead to a significant difference in user activities towards NotLikers.

Summary statistics of user activities towards all potential partners (the sum of Likers and NotLikers) are reported in Appendix G. There we show that due to a large number of users in our data who received very small number of *ℓikes* (more NotLikers than Likers), the impact of the treatment on user behavior towards Likers is buried by user behavior towards NotLikers.

2.6.2 Frictions and Sorting

The correlation in mate attributes has been widely documented and studied in previous research across various disciplines. Consistent with the existing literature, matched couples

[50]Hitsch et al. (2010a) who had access to the actual content of the messages exchanged by users in their online dating website report that it took a median number of 6 messages to reveal their phone number, email address, or to say a key phrase like "get's together" or "let's meet".

	Men			Women		
	control (1a)	treated (1b)	t-stat (1c)	control (2a)	treated (2b)	t-stat (2c)
Number of users	7,930	8,189		3,470	3,642	
Profiles clicked						
Mean	4.0	5.7	**13.245**	11.0	17.1	**10.674**
Median	2	3		5	7	
SD	7.3	9.3		17.0	29.9	
Likes sent						
Mean	1.100	1.387	**6.319**	1.935	3.282	**6.902**
Median	0	0		0	0	
SD	2.696	3.054		6.295	9.706	
Initiated messages						
Mean	1.0	1.3	**7.133**	0.87	1.2	**5.019**
Median	0	0		0	0	
SD	2.5	2.8		2.0	2.8	
Initiated messages that led to match						
Mean	0.2	0.3	**4.211**	0.3	0.4	**3.855**
Median	0	0		0	0	
SD	0.8	0.9		0.9	1.2	

Table 2.2: Summary statistics of user activities towards Likers

Notes. If the difference between the treatment and control group is significant (at the 5 percent level), the t-statistics are in bold.

on MonCherie also display positive sorting patterns along various characteristics. In our data, we find that age is strongly correlated across men and women (Pearson correlation coefficient $\rho = 0.71$). Although small, years of education ($\rho = 0.12$), BMI ($\rho = 0.11$) and attractiveness ($\rho = 0.14$) also display positive correlations. The numbers reported here are very similar to those reported by Hitsch et al. (2010a). In Hitsch et al. (2010a), the correlation between matched couples in the data is 0.7 in age, 0.12 in education, and 0.3 in attractiveness.

Our experiment reduces the search friction present in our platform for the treatment group by revealing the information about *likes*. By looking at how the treatment affects sorting patterns between matched couples, we can get insights on the impact of frictions on assortative matching. To compare the sorting patterns between the two groups, we first construct a

	Men			Women		
	control (1a)	treated (1b)	t-stat (1c)	control (2a)	treated (2b)	t-stat (2c)
Number of users	7,930	8,189		3,470	3,642	
Profiles clicked						
Mean	80.1	78.9	-0.420	28.9	29.5	0.388
Median	21	21		10	9	
SD	183.3	177.5		131.6	68.1	
Likes sent						
Mean	94.5	82.5	**-2.403**	13.4	14.0	0.319
Median	8	9		0	0	
SD	3.8	3.2		77.4	70.5	
Initiated messages						
Mean	21.4	21.0	-0.332	2.3	2.2	-0.497
Median	2	2		0	0	
SD	65.3	65.7		8.1	8.0	
Initiated Messages that led to match						
Mean	1.36	1.38	0.178	0.45	0.46	0.12
Median	0	0		0	0	
SD	4.7	4.9		1.6	1.7	

Table 2.3: Summary statistics of user activities towards NotLikers

Notes. If the difference between the treatment and control group is significant (at the 5 percent level), the t-statistics are in bold.

measure of attribute difference (henceforth "attribute difference") that can be used to test whether the treatment leads to a significantly different sorting pattern between a matched man and a woman. Specifically, the attribute difference between a man m and a woman w is obtained as $\Delta = |X_m - X_w|$ where X_m and X_w are m and w's characteristics, respectively. We obtain attribute difference between couples for age, education level, BMI, and attractiveness.[51] For ethnicity, we construct the attribute difference as a dummy variable that takes value 1 when m and w are of different ethnicity, and 0 otherwise.[52]

[51]Education level in our data is categorical. We transform education level into years of education in the following way: University is normalized to 0; High school = −4; Two-year college = −2; Masters = 2; Law School= 3; MedSchool = 3; PhD = 6 (6 years PhD is the new black). For BMI, we obtain the mean BMI for the following categories: skinny, average, heavier, and overweight from World Health Organization (WHO) Classification for Obesity Corresponding to Body Mass Index (BMI).

[52]We use Asian, Black, White, Indian, Hispanic, MidEastern, NativeAmerican and Pacific Islander as ethnic categories. Although Indian is Asian, MonCherie puts them into a different category. Users of MonCherie

We additionally create a synthetic variable that measures the attractiveness of a user, which is the total count of $\ell ikes$ received divided by the sum of $\ell ikes$ and "not $\ell ikes$" received. More specifically, the measure of attractiveness is obtained for each user i as follows:

$$\text{attractiveness}_i = \frac{\#\ell ikes\text{Received}_i}{\#\ell ikes\text{Received}_i + \#\text{Not}\ell ikes\text{Received}_i}$$

This measure is then converted to a decile, separately for men and women (with 1 being least attractive and 10 being most attractive).

Table 2.4, columns (1a) and (1b) display mean attribute differences (standard deviation in parentheses and number of observations in brackets) of couples *who matched with Likers*, for the control and treatment group, respectively. Column (1c) reports the difference between the two groups. Interestingly, attribute differences of the treatment group are significantly larger compared to those of the control group, across all dimensions. The age difference between couples in the treatment group ($\Delta = 4.6$) is on average 0.3 years (or 7.2 percent) greater than the age difference between couples in the control group ($\Delta = 4.3$); The difference in years of education between couples in the treatment group ($\Delta = 1.9$) is on average 0.18 years (or 10.4 percent) greater than that of couples in the control group ($\Delta = 1.7$); Approximately 58 percent of the users in the treatment group matched with partners of different ethnicity, which is 5 percent greater than that of the control group (53 percent); The BMI difference between couples in the treatment group ($\Delta = 4.3$) is 7.5 percent greater than the BMI difference between couples in the control group ($\Delta = 4.0$); The attractiveness difference

can choose mutiple ethnicities to describe themselves (e.g. Asian & White). We classify all users who choose more than one ethnicity as "other ethnicity". If both m and w fall into other-ethnicity category, the attribute difference equals 1.

between couples in the treatment group ($\Delta = 2.87$) is roughly 5 percent greater than the attractiveness difference between couples in the control group ($\Delta = 2.73$).[53]

We also use the Pearson correlation coefficient as an alternative meaure of sorting and report the correlation patterns in matches with Likers in Table 2.4 columns 2a−2b. Fisher's z-statistic (Column 2c) is used to test whether correlation coefficients between the two groups are statistically different. We find that our previous findings holds for age and attractiveness: the correlation in age for the treatment group ($\rho = 0.77$) is 4 percent less compared to the control group ($\rho = 0.8$); the correlation in attractiveness for the treatment group ($\rho = 0.11$) is approximately 50 percent less compared to the control group ($\rho = 0.16$). The correlation in education level for the treatment group is also slightly less than the control group, but the difference between the two groups is not significant.

Tables 2.5 shows differences in degree of sorting between the treatment and control group along various attributes when they matched with NotLikers. Since the focal user is not able to distinguish whether the NotLiker choose to not *like* the focal user, or whether they have not yet seen the focal user's profile, the impact of treatment on assortative matching is ambiguous.

Our results provide evidence suggesting that the treatment reduces the degree of sorting between couples. One possible explanation for this is that the treatment triggers users to initiate a match with potential partners whom they otherwise wouldn't in the absence of the treatment. Figure 2.5 plots the mean probability of initiating a conversation through messaging against the probability of receiving a reply (henceforth "match probability"). We can see that the probability of messaging is increasing in match probability. Table 2.6 reports

[53]When we do not convert the attractivenes measure into a decile, this result is reversed. We decided to convert attractiveness into a decile separately for men and women given vastly different patterns of attractiveness scores between the two genders.

	Matches between m and Liker w					
	Attribute Difference			Pearson Correlation Coefficient		
	control	treatment	Difference	control	treatment	Fisher's z
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Age	4.280	4.588	0.308***	0.800	0.770	3.342***
	(4.303)	(4.670)		(0.009)	(0.009)	
	[3,271]	[4,289]		[3,271]	[4,289]	
Education	1.709	1.887	0.179***	0.121	0.118	0.082
	(1.487)	(1.533)		(0.025)	(0.023)	
	[1,287]	[1,631]		[1,287]	[1,631]	
Ethnicity	0.531	0.577	0.046***			
	(0.499)	(0.494)				
	[1,586]	[2,141]				
BMI	4.017	4.323	0.306**	0.017	0.054	-0.884
	(3.490)	(3.654)		(0.028)	(0.033)	
	[994]	[1,224]		[994]	[1,224]	
Attractiveness	2.730	2.872	0.139***	0.161	0.108	2.301***
	(2.090)	(2.198)		(0.019)	(0.016)	
	[3,196]	[4,182]		[3,196]	[4,182]	

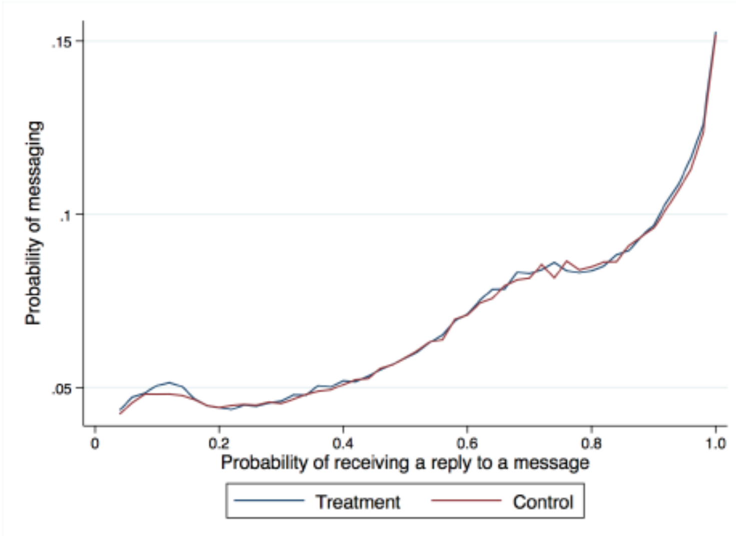
Notes. Standard deviation (for absolute diff) in parentheses. Number of observations in brackets. When exchanging at least 6 messages, pairs sent at least 3 messages each.

*Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Table 2.4: Attributes Differences with Initiated Matches

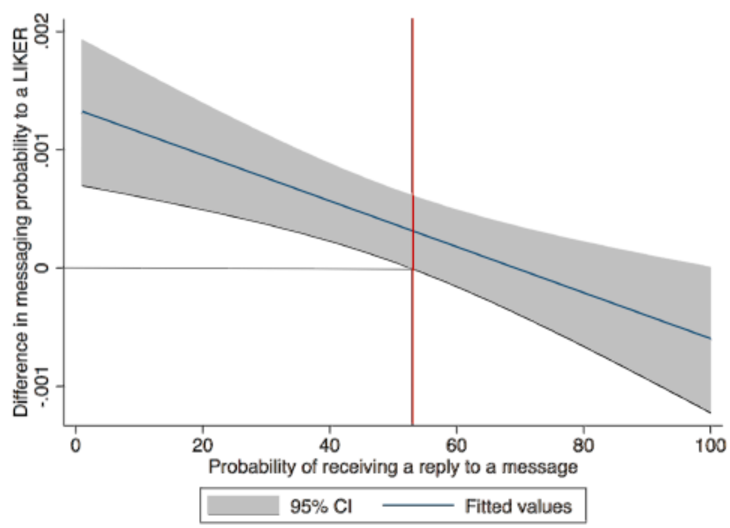
the summary statistics of match probabilities conditional on sending a message. The table shows that there is a greater probability of receiving a reply from a Liker than from a NotLiker (0.46 versus 0.13). A user who might be discouraged from sending a message under the control condition due to the (ex-ante) low probability of receiving a reply, may decide to send a message under the treatment condition once he sees that the other party had $\ell iked$ him.

To test this, we need to compare the two groups' likelihood of messaging to Likers for (ex-ante) low and high values of match probabilities. However, the expected match probabilities for the two groups are different. All else identical, the expected match probability with a Liker will be greater for a user in the treatment group than a user in the control group (unless the user in the control group $\ell ikes$ a profile and find out that he was $\ell iked$). In order to



Notes. Predicted match probabilities (probability of receiving a reply to a message) are converted into 50 quantiles.

Figure 2.4: Probability of messaging with respect to match probabilities



Notes. Predicted match probabilities (probability of receiving a reply to a message) are converted into 50 quantiles.

Figure 2.5: Probability of messaging to a LIker with respect to match probabilities

	Matches between m and NotLiker w		
	control	treatment	Difference
	(a)	(b)	(c)
Age	5.404	5.403	-0.000
	(5.303)	(5.326)	
	[12,888]	[13,505]	
Education	1.788	1.910	0.122***
	(1.423)	(1.492)	
	[4,808]	[4,717]	
Ethnicity	0.559	0.556	-0.003
	(0.497)	(0.497)	
	[6,138]	[6,319]	
BMI	4.311	4.218	-0.092
	(3.830)	(3.654)	
	[3,694]	[3,923]	
Attractiveness	3.051	2.942	-0.110***
	(2.337)	(2.265)	
	[12,355]	[12,974]	

Notes. Standard deviation in parentheses. Number of observations in brackets. When exchanging at least 6 messages, pairs sent at least 3 messages each.

*Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Table 2.5: Attributes Differences with Initiated Matches

	Mean	SD
Liker	0.463	0.499
NotLiker	0.131	0.338

Table 2.6: Probability of receiving a reply to a message

make a correct comparison, we need to use identical expected match probabilities for the two groups. Hence, we estimate expected match probabilities using logistic regression under the assumption that *all* users are in the control group. Assuming that all users are in the treatment group will not change the inference that we derive from this exercise. Figure **??** plots the difference (between the treatment and control group) in probability of messaging to Likers against the match probabilities. Specifically, this difference is obtained by subtracting the control group's mean probability of messaging to a Liker from that of the treatment group, separately for each quantile of match probability. The positive value of this difference implies that the treatment group is more likely to send a message compared to the control group, and vice versa. The difference in the figure is downward sloping and is always positive for sufficiently low match probabilities (the cutoff value of the match probability is approximately at 0.5). The positive difference for lower values of match probabilities suggest that users in the treatment group are more likely (compared to the control group) to message those who have *liked* them who ex-ante are less likely to respond.

So far, we have shown that the treatment leads to less sorting for matches achieved with Likers, suggesting that reducing search frictions may lower sorting between couples. The patterns observed in this section, however, are merely suggestive of the impact that frictions play on sorting and are not conclusive: First, the experiment reduces the search friction for the treatment group, but it does not reduce the search friction for the correspondent users who have interacted with our treatment group. Therefore, the limitation of our experiment is that it reduces the search friction for users on only one side of the market. Second, the treatment does not remove *all* frictions. Although the treatment reduces the uncertainty regarding preferences of potential partners to some extent, users nevertheless still face uncertainty about whether they will achieve a match. If the cost of initiating a contact is non-negligible, the decision to initiate a contact would depend on the probability of a match. Therefore, a user may decide to forgo a desirable partner if the expected probability of a match is sufficiently low. In an environment where only preferences shape the formation of a match, the cost of initiating a contact should be non-existent.

In the following section we develop a model of costly search (as well as costly initial contacting) that incorporates preference heterogeneity across users. The estimates from the model allows us to simulate equilibrium matches in a frictionless environment where only preferences shape the matching outcomes.

2.7 Model

2.7.1 Overview

We first present a brief overview of the model that summarizes the actions that a user can take at each stage of the search process. Here we describe the model only from the perspective of a male user m, but it is identical for a female user w. The details of each stage are described in the following subsections.

At the beginning of each exogenously given session τ, m chooses to browse k_τ number of profiles that maximizes his net benefit of browsing. The cost of browsing k_τ profiles is $c_m^{\text{browse}}(k_\tau)$. m then browses through each of the profiles $w \in \{w_1, w_2, ..., w_{k_\tau}\}$ that appear randomly until the number of profiles that he browsed reaches k_τ, at which point he stops.

- For each of the k_τ profiles, m goes through the *liking stage* (LS) followed by the *message stage* (MS):

1. Liking stage (LS): m chooses action $d_{mw} \in \{\text{like, nlike}\}$ which indexes his decision to ℓike or not ℓike w's profile ($d_{mw} = \text{like}$ if ℓike, and $d_{mw} = \text{nlike}$ if not ℓike). $\ell iking$ w's profile is an *indirect* way of sending the signal of interest, which leads to a positive probability of a match. If m chooses $d_{mw} = \text{like}$, he has to pay a cost c_m^{like}.
   - Two types of users: treated ($h_m = \text{treated}$) and control ($h_m = \text{control}$):
     (a) If $h_m = \text{treated}$, m can observe at the beginning of the liking stage (prior to choosing action d_{mw}), whether w had $\ell iked$ him or not.
     (b) If $h_m = \text{control}$, m cannot observe whether w had ℓiked him or not. This information will be revealed to him only if and after he chooses $d_{mw} = \text{like}$.

2. Message stage (MS): m has to choose $\mu_{mw} \in \{\text{msg}, \text{nmsg}\}$, which indexes his decision whether or not to message w ($\mu_{mw} = \text{msg}$ if message, and $\mu_{mw} = \text{nmsg}$ if not message). Sending a message is a *direct* way contacting w. If $\mu_{mw} = \text{msg}$, m has to pay a cost c_m^{msg}. If m had chosen $d_{mw} = \text{like}$ during the liking stage, there is a positive probability of matching with w even if m chooses to not message her (through indirect signaling, i.e. $\ell iking$, at the liking stage.).

2.7.2 Model Details

We consider an online dating platform where in each period, N_M men and N_W women are searching for a partner. Time is discrete, and we assume that discounting across time is negligible, i.e. time discount factor $\rho \approx 1$. This assumption had been used in existing research that use data from online dating platforms (Hitsch et al. (2010a); Fong (2018)). Each man is indexed by $m \in \mathcal{M} = \{1, 2, ..., N_M\}$. Similarly, each woman is indexed by $w \in \mathcal{W} = \{1, 2, ..., N_W\}$. In each period, a random profile of a woman is displayed to a man. The utility to a man of matching with a woman depends on own and woman's characteristics. A match occurs if both man and woman agree to match. We assume that users do not agree to match if the net expected utility from matching is lower than each user's reservation utility.

In each period, women's profiles are drawn randomly from a distribution F_W. We assume that the distribution of single users' profiles is exogenously given and is stationary over time. To guarantee stationarity, we assume that users who are matched exit the market and are immediately replaced by their "clones" as in McNamara and Collins (1990), Burdett and Coles (1997), Bloch and Ryder (2000) and Adachi (2003).[54]

[54] A clone of a man m (woman w) has identical characteristics as man m (woman w).

Latent Utility

We assume that partner preferences depend on observed own and partner attributes, and idiosyncratic preference shock e_{mw} which follows i.i.d logistic distribution:

$$U_M(m, w) = U_M(X_m, X_w; \Theta_M) + e_{mw} \tag{2.1}$$

where X_M and X_W are m and w's observed characteristics which follow a distribution F_M and F_W, respectively. Θ_M is a (column) vector that represents men's preferences (similarly, Θ_W represents women's preferences).

The latent utility that m gets if he matches with w is parameterized as

$$\begin{aligned} U_M(X_m, X_w; \Theta_M) = \quad & x'_w\beta_M + (|x_w - x_m|'_+)\beta^+_M + (|x_w - x_m|'_-)\beta^-_M \\ & + \textstyle\sum_{r,s=1}^{N_{eth}} 1\{x^{eth}_{mr} = 1 \text{ and } x^{eth}_{ws} = 1\} \cdot \beta^{eth}_{M,rs} \end{aligned} \tag{2.2}$$

where x_m and x_w are m and w's characteristics which have continuous values, respectively. $|x_w - x_m|_+$ is the difference between m and w's attributes if this difference is positive, and $|x_w - x_m|_-$ is the absolute value of this difference if this difference is negative. More formally, $|x_w - x_m|_+ = \max(x_w - x_m, 0)$ and $|x_w - x_m|_- = \max(x_m - x_w, 0)$. x^{eth}_m and x^{eth}_w are sets of dummy variables indicating m and w's ethnicity, respectively. For example, $x^{eth}_{m,\text{asian}} = 1$ if m is Asian and 0 otherwise. The set of preference parameters to be estimated is $\Theta_M = (\beta_M, \beta^+_M, \beta^-_M, \beta^{eth}_M)$ for men and $\Theta_W = (\beta_W, \beta^+_W, \beta^-_W, \beta^{eth}_W)$ for women.

Information Structure

Before proceeding to describe the actions that a user can take at each stage of the search process, we first lay out users' information structure for each profile that he is browsing. This information structure depends on the focal user's type (treated or control), the stage of the search process (liking stage or message stage), and the actions taken at the liking stage.

Let $\Omega_{mw}^{full} = \{X_w, e_{mw}, \ell_{wm}\}$ be the information set of m when he has full information about w. Under full information, m observes X_w, e_{mw} as well as ℓ_{wm}, which is a dummy variable that takes value 1 if w had *liked* m and 0 otherwise. Denote Ω_{mw}^{stage,h_m} as the information that m of type h_m has about w at each stage, $stage \in \{LS, MS\}$.

The information set of m about w at the liking stage (LS) is given by

$$\Omega_{mw}^{LS,h_m} = \begin{cases} X_w, e_{mw}, \mathbb{E}[\ell_{wm}|X_m, X_w] & \text{if } h_m = \text{control} \\ X_w, e_{mw}, \ell_{wm} & \text{if } h_m = \text{treated} \end{cases} \tag{2.3}$$

At the liking stage, unlike treated users who have full information about w, users in the control group do not observe the true value of ℓ_{wm}, and therefore form an expectation about ℓ_{wm} conditional on own and w's characteristics.

As explained earlier, the exact value of ℓ_{wm} will be revealed to users in the control group if they choose to *like* a profile ($d_{mw} =$ like) during the liking stage. Therefore, at the message stage, all users *except* those in the control group who have chosen to not *like* a profile will have full information about w. The information set at the message stage can then be summarized as follows:

$$\Omega_{mw}^{MS,h_m} = \begin{cases} X_w, \mathbb{E}[\ell_{wm}|X_m, X_w], e_{mw} & \text{if } h_m = \text{control } \& \ d_{mw} = \text{nlike} \\ X_w, e_{mw}, \ell_{wm} & \text{if } h_m = \text{treated or } (h_m = \text{control } \& \ d_{mw} = \text{like}) \end{cases} \tag{2.4}$$

Formation of consideration set

When a user opens the app, profiles are displayed to him sequentially in a random fashion. User must engage in search to browse the pool of potential partners. Because browsing is costly, he browses only a limited number of profiles, which makes up his consideration set. The characterization of a consideration set requires an assumption on the type of search method that is employed by the user (simultaneous vs sequential search). In simultaneous search, once a user decides *how many* and *which* alternatives to consider, he searches and gathers information from all alternatives in his consideration set to resolve uncertainty (Stigler (2961); Roberts and Lattin (1991); Mehta et al. (2003); Honka (2014); Pires (2016)). Uses then chooses an alternative from his consideration set that gives the highest utility. In sequential search on the other hand, user calculates each alternative's reservation utility, ranks them in a decreasing order, and then starts searching with the top-ranked alternative, and work his way down (Weitzman (1979); Kim et al. (2010); Chen and Yao (2017)).

The traditional simultaneous search model, in which a user samples a fixed number of alternatives and purchase the alternative that gives the highest utility, is not directly applicable in our setting. This is because (i) there is a positive probability of being rejected by the other party, and (ii) profiles are displayed to a user in a random fashion and hence the user cannot choose *which* profiles to consider. For the same reason, we are not able to apply the sequential search model. We therefore consider a variant of a simultaneous search model, where a user chooses *how many* profiles to search, but does not decide on *which* alternatives to consider. We assume that the user chooses a fixed number (a number of profiles to browse) that maximizes the sum of the expected utility minus the total cost of browsing those profiles.

While we observe the number of profiles browsed by each user during one month treatment period of our data, we do not know how many profiles the user had browsed (or will browse) during his entire course of search at the platform. We cannot simply estimate the cost of browsing based on the total number of profiles browsed during one month because users will continue browsing more profiles in following months. We therefore assume that users choose a fixed number (of profiles to browse) at the beginning of each exogenously given "session". While we observe time-stamped actions for each user, we do not observe session IDs (nor do we know whether a user opened/closed the app). Therefore, we define sessions based on minutes elapsed between user's activities. Specifically, we define a new session if 180 minutes (3 hours) elapse without a user taking any action.

Specifically, m's *net* expected utility of a random profile (the expected utility minus the expected utility of staying single minus the expected costs of messaging and/or liking), denoted as z, is an i.i.d. draw from a distribution F_Z. At the beginning of each session τ, m chooses to search k_τ profiles that maximizes his net benefit of browsing, denoted as $\Gamma_{m\tau}(k_\tau)$, i.e. the sum of the net expected utility among the browsed profiles minus the total cost of browsing

$$\Gamma_{m\tau}(k_\tau) = k_\tau \cdot \int z f_z(z) dz - c_m^{\text{browse}}(k_\tau) \tag{2.5}$$

where $c_m^{\text{browse}}(k_\tau)$ is the total cost of browsing k_τ profiles and is a convex function of k_τ. This cost can be interpreted as time and cognitive effort spent on browsing k_τ profiles. A user picks the number k_τ which maximizes his net benefit of browsing. If a user chooses to browse k_τ profiles, he has to pay a total cost $c_m^{\text{browse}}(k_\tau)$, but will have k_τ profiles to browse from.

This describes how a user in our model forms his consideration set for each session. Once a user has chosen the size of his optimal consideration set, he goes through the liking and message stage for each randomly appearing profile $w \in \{w_{1,} w_2, ..., w_{k_\tau}\}$ for the first k_τ random profiles that are displayed to him during session τ. In what follows, we remove the subscript τ to simplify the exposition. Next, we first describe the message stage and then describe the screening stage in a backwards induction manner.

Message Stage

Let $V_M(m)$ denote man m's expected utility of remaining single and continuing the search. Similarly, let $V_W(w)$ denote woman w's expected utility of remaining single and continuing the search. We assume that $V_M(m)$ and $V_W(w)$ do not change across the browses. This is because while the user only browses k_τ profiles during session τ, he can continue browsing additional profiles in the next session. In addition, the user also has an outside option of searching for a partner offline.[55] For the moment, let us suppose that these expected utilities are given. Conditional on his choice $d_{mw} \in \{\text{like, nlike}\}$ at the liking stage, the expected utility at the message stage from choosing $\mu_{mw} \in \{\text{msg, nmsg}\}$ is given by

$$EU_{mw}^{MS}(\text{msg}|d_{mw}) = \begin{cases} -c_m^{\text{msg}} + U_M(m,w) \cdot \tilde{\pi}_{mw}^{MS}(\text{msg}|d_{mw}) + V_M(m) \cdot \left(1 - \tilde{\pi}_{mw}^{MS}(\text{msg}|d_{mw})\right) & \text{if } \mu_{mw} = \text{msg} \\ U_M(m,w) \cdot \tilde{\pi}_{mw}^{MS}(\text{nmsg}|d_{mw}) + V_M(m) \cdot \left(1 - \tilde{\pi}_{mw}^{MS}(\text{nmsg}|d_{mw})\right) & \text{if } \mu_{mw} = \text{nmsg} \end{cases}$$

where c_m^{msg} is the cost of sending a message, and can be interpreted as (i) time and effort to compose a message, and/or (ii) aversion towards experiencing negative emotion in case w does not respond.

$\tilde{\pi}_{mw}^{MS}(\mu_{mw}|d_{mw})$ is the expectation (formed at the message stage) about the match probability, which depends on both μ_{mw} and d_{mw}. Specifically, $\tilde{\pi}_{mw}^{MS}(\mu_{mw}|d_{mw})$ can be written as

$$\tilde{\pi}_{mw}^{MS}(\mu_{mw}|d_{mw}) = \mathbb{E}_{\Delta_{hm}^{MS}}\left[\pi_{mw}(d_{mw}, \mu_{mw}|X_m, X_w, \ell_{wm})|\Omega_{mw}^{MS}, d_{mw}\right]$$

where $\pi_{mw}(d_{mw}, \mu_{mw}|X_m, X_w, \ell_{wm})$ is the match probability when m chooses a sequence of actions (d_{mw}, μ_{mw}) conditional on X_m, X_w and ℓ_{wm}. The expectation is taken with

[55] This assumption may be the limitation of the current version of this paper which we hope to address in the future.

respect to $\Delta_{h_m}^{MS} = \Omega_{mw}^{full} \backslash \Omega_{mw}^{MS,h_m}$, which is the discrepancy between the full information set and the information that type h_m has at the beginning of the message stage.[56] If d_{mw} = like , then $\tilde{\pi}_{mw}^{MS}(\text{nmsg}|d_{mw}) > 0$ due to the indirect signaling effect of interest. We assume that if m neither $\ell ikes$ not messages w, the match probability is zero, i.e. $\pi_{mw}(\text{nlike, nmsg}|X_m, X_w, \ell_{wm}) = 0$. We also assume that the match probability is independent of d_{mw} if m messages w, i.e. $\pi_{mw}(\text{like, msg}|X_m, X_w, \ell_{wm}) = \pi_{mw}(\text{nlike, msg}|X_m, X_w, \ell_{wm})$.

m will choose to message w if and only if

$$EU_{mw}^{MS}(\text{msg}|d_{mw}) \geq EU_{mw}^{MS}(\text{nmsg}|d_{mw}). \tag{2.6}$$

Denote $\Delta\tilde{\pi}_{mw}^{MS}$ as the expected difference in match probability from messaging and not messaging

$$\Delta\tilde{\pi}_{mw}^{MS} = \begin{cases} \tilde{\pi}_{mw}^{MS}(\text{msg}|\text{like}) - \tilde{\pi}_{mw}^{MS}(\text{nmsg}|\text{like}) & \text{if } d_{mw} = \text{like} \\ \tilde{\pi}_{mw}^{MS}(\text{msg}|\text{nlike}) - \tilde{\pi}_{mw}^{MS}(\text{nmsg}|\text{nlike}) & \text{if } d_{mw} = \text{nlike} \end{cases} \tag{2.7}$$

Since $\tilde{\pi}_{mw}^{MS}(\text{nmsg}|\text{nlike}) = \mathbb{E}_{\Delta_{h_m}^{MS}}\Big[\pi_{mw}(\text{nlike,nmsg}|X_m, X_w, \ell_{wm})\Big] = 0$, with a bit of algebra one can easily see that condition 2.6 can be rewritten as

$$U_M(m, w) - V_M(m) \geq \frac{c_m^{\text{msg}}}{\Delta\tilde{\pi}_{mw}^{MS}} \tag{2.8}$$

Since e_{mw} follows i.i.d logistic distribution, the probability of sending a message to w can be written as

$$\Pr(\mu_{mw} = \text{msg}|d_{mw}) = \frac{\exp\Big(U_M(X_m, X_w; \Theta_M) - V_M(m) - c_m^{\text{msg}} \cdot \Delta\tilde{\pi}_{mw}^{MS}(\text{msg})^{-1}\Big)}{1 + \exp\Big(U_M(X_m, X_w; \Theta_M) - V_M(m) - c_m^{\text{msg}} \cdot \Delta\tilde{\pi}_{mw}^{MS}(\text{msg})^{-1}\Big)} \tag{2.9}$$

[56] If $\Omega_{mw}^{full} \backslash \Omega_{mw}^{MS,h_m} = \emptyset$, then $\tilde{\pi}_{mw}^{MS}(\mu_{mw}|d_{mw}) = \pi_{mw}(d_{mw}, \mu_{mw}|X_m, X_w, \ell_{wm})$.

Liking Stage

The utility at the liking stage from choosing $d_{mw} \in \{\text{like,nlike}\}$ is

$$U^{LS}_{mw}(d_{mw}) = \begin{cases} -c^{\text{like}}_{m} + \varepsilon^{\text{like}}_{mw} & \text{if } d_{mw} = \text{like} \\ \varepsilon^{\text{nlike}}_{mw} & if d_{mw} = \text{nlike} \end{cases} \tag{2.10}$$

where c^{like}_{m} is the psychological cost that m incurs if he chooses to ℓike w. This cost can be incurred due to several reasons, some of which can be: (i) m simply does not find w attractive enough to ℓike her, (ii) aversion towards the negative emotion that will be incurred in case w does not ℓike him back (Baumeister and Dhavale (2001)), (iii) negative emotion associated with rejecting w after having sent the ℓike signal to trigger w's response. Research has shown that the object of unwanted affection may experience annoyance, frustration, and that rejecting the other's overtures may cause guilt, discomfort, and other distress (Baumeister et al. (1993)). $\varepsilon^{\text{like}}_{mw}$ and $\varepsilon^{\text{nlike}}_{mw}$ are error terms observed by m (but unobserved by the researcher) that affects m's decision to ℓike w. We assume that $\varepsilon^{\text{like}}_{mw}$ and $\varepsilon^{\text{nlike}}_{mw}$ are distributed i.i.d Type I EV.

The expected utility from both stages (liking stage & message stage) is

$$EU^{BothStages}_{mw}(d_{mw}) = U^{LS}_{mw}(d_{mw}) + \max\left\{\mathbb{E}_{\Delta^{LS}_{h_m}}\left[EU^{MS}_{mw}(\text{msg}|d_{mw})|\Omega^{LS}\right], \mathbb{E}_{\Delta^{LS}_{h_m}}\left[EU^{MS}_{mw}(\text{nmsg}|d_{mw})|\Omega^{LS}\right]\right\} \tag{2.11}$$

where the expectation is taken with respect to $\Delta^{LS}_{h_m} = \Omega^{full} \backslash \Omega^{LS,h_m}$, the discrepancy between the full information set and the information that type h_m knows at the beginning of the liking stage.[57] Since $\pi_{mw}(\text{nlike, nmsg}|X_m, X_w, \ell_{wm}) = 0$, we have $\mathbb{E}_{\Delta^{LS}_{h_m}}\left[EU^{MS}_{mw}(\text{nmsg}|\text{nlike})|\Omega^{LS}\right] =$

[57]Note that the expectation about the match probability at the liking stage is $\tilde{\pi}^{LS}_{mw}(d_{mw}, \mu_{mw}) = E_{\Delta^{LS}_{h_m}}\left[\pi_{mw}(d_{mw}, \mu_{mw}|X_m, X_w, \ell_{wm})|\Omega^{LS}_{mw}\right]$

$V_M(m)$. The choice-specific expected utility from both stages can then be written as

$$EU_{mw}^{BothStages}(d_{mw}) = \begin{cases} -c_m^{\text{like}} + \varepsilon_{mw}^{\text{like}} & \\ +\max\left\{\mathbb{E}_{\Delta_{h_m}^{LS}}\left[EU_{mw}^{MS}(\text{msg}|\text{like})|\Omega^{LS}\right], \mathbb{E}_{\Delta_{h_m}^{LS}}\left[EU_{mw}^{MS}(\text{nmsg}|\text{like})|\Omega^{LS}\right]\right\} & \text{if } d_{mw} = \text{like} \\ \varepsilon_{mw}^{\text{nlike}} + \max\left\{\mathbb{E}_{\Delta_{h_m}^{LS}}\left[EU_{mw}^{MS}(\text{msg}|\text{nlike})|\Omega^{LS}\right], V_M(m)\right\} & \text{if } d_{mw} = \text{nlike} \end{cases}$$

m will choose $d_{mw} = \text{like}$ if and only if

$$EU_{mw}^{BothStages}(like) \geq EU_{mw}^{BothStages}(nlike)$$

Let $EU_{mw}^{BothStages}(d_{mw}) = \bar{EU}_{mw}^{BothStages}(d_{mw}) + e_{mw}^{d_{mw}}$. Since $\varepsilon_{mw}^{\text{like}}$ and $\varepsilon_{mw}^{\text{nlike}}$ follow i.i.d Type I EV distribution, the probability of choosing $d_{mw} = \text{like}$ is

$$\Pr(d_{mw} = \text{like}) = \frac{\exp\Big(\bar{EU}_{mw}^{BothStages}(\text{like}) - \bar{EU}_{mw}^{BothStages}(\text{nlike})\Big)}{1 + \exp\Big(\bar{EU}_{mw}^{BothStages}(\text{like}) - \bar{EU}_{mw}^{BothStages}(\text{nlike})\Big)} \tag{2.12}$$

2.7.3 Removing frictions and the Gale-Shapley problem

Suppose that all costs are removed, i.e. $c_m^{\text{browse}} = c_m^{\text{like}} = c_m^{\text{msg}} = 0$ (similarly, $c_w^{\text{browse}} = c_w^{\text{like}} = c_w^{\text{msg}} = 0$). Consider m's decision problem of whether to send a message to w. The condition in equation 2.8 reduces to

$$U_M(m, w) \geq V_M(m). \tag{2.13}$$

The match probability no longer appears in the threshold condition in equation 2.13: m messages as long as the utility from a match is greater than the expected value of remaining single. Now consider m's decision to ℓike w. The decision to ℓike no longer affects the threshold condition at the message stage through its impact on match probability. $\varepsilon_{mw}^{\text{like}}$ and $\varepsilon_{mw}^{\text{nlike}}$ are the only factors that affect m's decision to ℓike w. To simplify the analysis, we remove the redundancy of deciding whether to ℓike when all the costs are set to zero. That is, we assume that messaging is the only way of contacting a potential mate when all the costs are removed.

In a frictionless environment in which all the costs are removed, the expected value of remaining single can be characterized as a system of Bellman equations for man m and woman w

$$\begin{aligned} V_M(m) = &\quad \rho \int U_M(m,w)\Pr(\text{match}) + V_M(m)\Big(1 - \Pr(\text{match})\Big) dF_W(w) \\ V_W(w) = &\quad \rho \int U_M(m,w)\Pr(\text{match}) + V_M(m)\Big(1 - \Pr(\text{match})\Big) dF_M(m) \end{aligned} \tag{2.14}$$

where $\Pr(\text{match}) = \Pr\left(U_M(m,w) \geq V_M(m) \,\&\, U_W(w,m) \geq V_W(w)\right)$. The system of equations above defines a monotone iterative mapping that converges to a profile of reservation utilities $\left(V_M^{GS}(m), V_W^{GS}(w)\right)$ solving the system, and thus characterizing the stationary equilibrium in this market. The equilibrium reservation utilities $\left(V_M^{GS}(m), V_W^{GS}(w)\right)$ can be thought of as person-specific prices that clear the demand and the supply of that person.

Adachi (2003) shows that as $\rho \to 1$, the set of equilibrium outcomes in a decentralized search model reduces to the set of stable matchings in a corresponding Gale-Shapley marriage problem. A stable match is defined, following Gale and Shapley (1962), as a pairing where there are no pairs (m,w) who are willing to abandon their partners and match with each other. Specifically, Adachi (2003) shows that the system of Bellman equations in (2.14) coincide with the following system of equations characterizing the set of stable matchings in a Gale-Shapley marriage problem:

$$\begin{aligned} V_M^{GS}(m) &= \max_{\mathcal{W}\cup\{m\}} \left\{U_M(m,w) | U_W(w,m) \geq V_W^{GS}(w)\right\} \\ V_W^{GS}(w) &= \max_{\mathcal{M}\cup\{w\}} \left\{U_W(w,m) | U_M(m,w) \geq V_M^{GS}(m)\right\} \end{aligned} \tag{2.15}$$

This result is intuitive: If time is not discounted, and if there are no search costs (costs of browsing, as well as the costs of $\ell iking$ and messaging) each man continues the search process until he finds a woman such that $U_M(m,w) \geq V_M^{GS}(m)$ and $U_W(w,m) \geq V_W^{GS}(w)$. Then a man will be matched with the best woman who is willing to match with him, and vice versa. However, this is how the set of stable matchings are characterized in Gale-Shapley problem. Henceforth we denote $V_M^{GS}(m)$ and $V_W^{GS}(w)$ as the expected utility of staying single in a frictionless environment for man m and woman w, respectively.

Let us first describe how we estimate $V_M(m)$ and $V_W(w)$ in the presence of frictions. As mentioned earlier, we assume that $V_M(m)$ and $V_W(w)$ remain constant across browses. Ideally, we want to estimate these reservation values using user-specific fixed effects following Hitsch et al. (2010a) and Banerjee et al. (2013). However, not only is this approach computationally burdensome due to large number of users in our sample, but it is also unsuitable in our setting due to selection issue: in order to include user-specific fixed effects, we need to drop users who haven't $\ell iked$ and/or messaged any profiles. However, since treatment affects the way users ℓike and message, dropping these individuals may bias our results.

Therefore, we estimate the reservation values of remaining single using K-means clustering, where we classify users into groups based on their observed characteristics. K-means is an unsupervised learning approach that partitions the dataset into K pre-defined non-overlapping clusters. Each user is assigned to a cluster such that the sum of the squared distance between the users' characteristics and the cluster's centroid (arithmetic mean of all users' characteristics that belong to that cluster) is at the minimum. In this paper we use $K = 10$. Ex ante, the expected value of remaining single should be identical for the control and treatment group. This is because prior to actually browsing a specific profile, users in both groups do now know the true value of ℓ_{wm} and have to form expectations about it. Therefore, we do not include the treatment status as a feature when partitioning users into clusters.

Then the question is how to obtain the expected value of remaining single in the absence of frictions? When frictions are present, users search K_m profiles on the platform, while also having the option of searching for partners offline. When frictions are not present, users can search on the site forever, but also have the option of searching offline. Since we cannot estimate $V_M^{GS}(m)$ from the data, we use $V_M(m)$ and $V_W(w)$ estimated from the model as the expected value of remaining single when simulating matches in a frictionless environment. Due to costs incurred while searching K profiles, we have $V_M^{GS}(m) \geq V_M(m)$. Hence, by using $V_M(m)$ we are underestimating the expected value of remaining single in a frictionless environment, making users "less selective".[58] However, we do not think that this will affect our results for the following reason: when simulating matches using deferred-acceptance algorithm (which we describe in more detail later in the text), each man ranks all women from the most preferred to the least preferred. Then, each man makes an offer to his most preferred woman and the woman either accepts or declines the offer. Men whose offers have been declined then make an offer to his next most-preferred woman, and this process continues until all men exhaust the list of women that give utility greater than their reservation utility. Hence the underestimation of $V_M^{GS}(m)$ is a problem only to the extent that users match with partners whose utility is close to $V_M(m)$. Our simulation results show that the average utility of a matched partner is approximately 780 percent greater than $V_M(m)$ and 95 percent of users

[58] Whether this will bias will create more sorting or less sorting is an empirical question: If preferences are the main determinants of sorting, then underestimating $V_M(m)$ will lead to less sorting. On the other hand, if search frictions are the main determinants of sorting, then underestimating $V_M(m)$ will lead to more sorting.

matched with a partner whose utility is *at least* 34 percent greater than $V_M(m)$.[59] Since the majority of matches are with partners whose utility is far above from $V_M(m)$, underestimation of the reservation value should not affect our results.

2.8 Estimation

2.8.1 Probability of a match

We estimate m's beliefs about the match probability at each stage directly from the data. In particular, we use a binary logit specification to model the probability of receiving a reply conditional on own and potential partner's attributes. In addition, the probability of a match depends on whether or not w had $\ell iked$ m. Since $\ell iking$ is costly, w will ℓike m only if $U_W(w, m)$ is sufficiently high. When $U_W(w, m)$ is sufficiently high, w is more likely to accept m's offer.

Let $\text{Respond}_{wm}^{\text{like}}$ be a binary variable that equals 1 if w responds to m's ℓike by sending a message to m and 0 otherwise.[60] We estimate the following equation using logistic regression:

$$\text{Respond}_{wm}^{\text{like}} = U_W(X_w, X_m; \Theta_W) + \psi^{\text{like}} L_{wm}^{\text{like}} + e_{wm} \tag{2.16}$$

where $U_W(\cdot)$ is defined similarly as in equation 2.1. L_{wm}^{ℓ} is a variable that equals ℓ_{wm} if m is in the treatment group and equals $\mathbb{E}[\ell_{wm}|X_m, X_w]$ otherwise:

$$L_{wm}^{\ell} = {\ell_{wm}}^{h_m} \cdot \mathbb{E}[\ell_{wm}|X_m, X_w]^{1-h_m}. \tag{2.17}$$

[59] Approximately 99 percent of users matched with a partner whose utility is *at least* 6 percent greater than $V_M(m)$.

[60] Note that here we index the subscript as wm as opposed to mw. This is to reflect w's preferences and decisions towards m

The parameter ψ^{like} is the effect of L^{ℓ}_{wm} on the probability of responding to m's ℓike. We use the predicted values from equation 2.1 as the probability of a match from only sending a ℓike.

Similarly, let $\text{Reply}^{\text{msg}}_{wm}$ be a dummy variable that takes value 1 if w replies to m's message and 0 otherwise. We estimate the following equation using logistic regression:

$$\text{Reply}^{\text{msg}}_{wm} = U_W(X_m, X_w; \Theta_W) + \psi^{\text{msg}}_{stg} L^{\text{msg,stg}}_{wm} + e_{wm} \tag{2.18}$$

where

$$L^{\text{msg,stg}}_{wm} = \begin{cases} {\ell_{wm}}^{h_m} \cdot \mathbb{E}[\ell_{wm}|X_m, X_w]^{1-h_m} & \text{if } stg = LS \\ {\ell_{wm}}^{h_m} \times \Big(\ell_{wm} \cdot 1\{d_{mw} = \text{like}\} + \mathbb{E}[\ell_{wm}|X_m, X_w] \cdot 1\{d_{mw} = \text{nlike}\}\Big)^{1-h_m} & \text{if } stg = MS \end{cases} \tag{2.19}$$

and ψ^{msg}_{stg} is the effect of $L^{\text{msg,stg}}_{wm}$ on the probability of replying to m's message. Note that the value of $L^{\text{msg,stg}}_{wm}$ depends on whether m is at the liking stage or at the message stage. If m is at the liking stage, $L^{\text{msg,LS}}_{wm}$ equals ℓ_{wm} if m is in the treatment group and $\mathbb{E}[\ell_{wm}|X_m, X_w]$ otherwise. At the message stage, users in the control group who have chosen $d_{mw} = \text{like}$ will find out the true value of ℓ_{wm}. The value of $L^{\text{msg,MS}}_{wm}$ for these users, as well as users in the treatment group, equals ℓ_{wm}. For users in the control group who have chosen $d_{wm} = \text{nlike}$, $L^{\text{msg,MS}}_{wm}$ equals $\mathbb{E}[\ell_{wm}|X_m, X_w]$. We use the predicted values from equation 2.18 as the probability of a match from sending a message.

2.8.2 Likelihood

We maximize the joint likelihood of m's decisions in consideration, liking and message stages, for each of the profiles browsed. The likelihood of our model is given by

$$L = \prod_{m=1}^{N} \prod_{\tau=1}^{T_m} \mathrm{Pr}_{m\kappa\tau}{}^{\zeta_{m\kappa t}} \prod_{w=1}^{J_w} \left(\mathrm{Pr}_{mwd} \cdot \mathrm{Pr}_{mw\mu|mwd}^{\delta_{mw}} \left(1 - \mathrm{Pr}_{mw\mu|mwd}\right)^{1-\delta_{mw}} \right)^{\vartheta_{mw}}$$
$$\times \left((1 - \mathrm{Pr}_{mwd}) \cdot \mathrm{Pr}_{mw\mu|mwd}^{\delta_{mw}} \left(1 - \mathrm{Pr}_{mw\mu|mwd}\right)^{1-\delta_{mw}} \right)^{1-\vartheta_{mw}} \quad (2.20)$$

where $\mathrm{Pr}_{m\kappa\tau}$ is the probability that m chooses to browse κ number of profiles in session τ, Pr_{mwd} is the probability that m *likes* w, and $\mathrm{Pr}_{mw\mu}$ is the probability that m messages w. $\zeta_{m\kappa\tau}$ is a binary variable indicating the number of profiles browsed in session τ; ϑ_{mw} indicates the decision made at the liking stage ($\vartheta_{mw} = 1$ if $d_{mw} =$ like, $\vartheta_{mw} = 0$ otherwise) and δ_{mw} indicates the decision chosen at the message stage ($\delta_{mw} = 1$ if $\mu_{mw} =$ msg, $\delta_{mw} = 0$ otherwise).

Note that the probability to browse κ profiles in session τ, $\mathrm{Pr}_{m\kappa t}$, does not have a closed form. We therefore follow Honka (2014) and Honka et al. (2017) and use a simulation approach to calculate them. We use a kernel-smoothed frequency simulator (McFadden, 1989) in the estimation and smooth the probabilities using a multivariate scaled logistic CDF (Gumbel, 1961):

$$F(\omega_1, ..., \omega_\Lambda; s_1, ..., s_\Lambda) = \frac{1}{1 + \sum_{\lambda=1}^{\Lambda} exp(-s_\lambda \omega_\lambda)} \qquad \forall \lambda = 1, ..., \Lambda, \quad (2.21)$$

where $s_1, ..., s_\Lambda$ are scaling parameters.

The probability of choosing to search k profiles at session t is obtained as follows:

1. Draw Q draws of e_{mw}, $\varepsilon_{mw}^{\text{like}}$, and $\varepsilon_{mw}^{\text{nlike}}$ for each m and w combination

2. For each draw e_{mw}^{q}, $\varepsilon_{mw}^{\text{like},q}$, and $\varepsilon_{mw}^{\text{nlike},q}$, compute

$$EU_{mw}^{BothStages}(\text{like}) = -c_m^{\text{like}} + \max\left\{\mathbb{E}_{\Delta_{hm}^{LS}}\left[EU_{mw}^{MS}(\text{msg}|\text{like})|\Omega^{LS}\right], \mathbb{E}_{\Delta_{hm}^{LS}}\left[EU_{mw}^{MS}(\text{nmsg}|\text{like})|\Omega^{LS}\right]\right\} + \varepsilon_{mw}^{\text{like},q}$$

and

$$EU_{mw}^{BothStages}(\text{nlike}) = \max\left\{\mathbb{E}_{\Delta_{hm}^{LS}}\left[EU_{mw}^{MS}(\text{msg}|\text{nlike})|\Omega^{LS}\right], V_M(m)\right\} + \varepsilon_{mw}^{\text{nlike},q}$$

Using $EU_{mw}^{BothStages}(\text{like})$ and $EU_{mw}^{BothStages}(\text{nlike})$, obtain z_{mw}^{q}:

$$z_{mw}^{q} = \begin{cases} EU_{mw}^{BothStages}(like) - V_M(m) & if\ EU_{mw}^{BothStages}(like) \geq EU_{mw}^{BothStages}(nlike) \\ EU_{mw}^{BothStages}(nlike) - V_M(m) & otherwise \end{cases}$$

3. Calculate $\mathbb{E}[z_m] = \frac{1}{Q}\sum_{q=1}^{Q}\mathbb{E}[z_{mw}^{q}]$ where the expectation is taken with respect to all w's that m has browsed

Compute the probability of choosing k

1. Draw Q draws of $\eta_{1,m\tau}$ and $\eta_{2,m\tau}$ for each m and τ combination from its distribution

2. For each m and τ, let $\mathit{\Gamma}_{m\tau,k}^{q} = k\mathbb{E}[z_m] - (\phi_1 + \eta_{1,m\tau}^{q})k - \phi_2 k^2$

3. Define the optimality conditions for observed user search behavior

$$\mathit{\Gamma}_{m\tau,k}^{q} \geq \max\left(\mathit{\Gamma}_{m\tau,k'}^{q}\right) \qquad \forall k' \neq k$$

and calculate the following difference

$$x^q_{m\tau k} = \Gamma^q_{m\tau,k} - \max\left(\Gamma^q_{m\tau,k'}\right) \qquad\qquad \forall k' \neq k$$

4. Calculate the smoothed probability using

$$Pr^q_{m\tau k} = \frac{1}{1+\exp(-s_1 \cdot x^q_{m\tau k})}$$

where s_1 is the tuning parameter

5. Finally, average the choice probability across all Q draws for η_{mt}

$$Pr_{m\tau k} = \frac{1}{Q}\sum_{q=1}^{Q} Pr^q_{m\tau k}$$

6. We use a scaling factor of $s_1 = 5$ and take $Q = 50$ draws.

2.8.3 Identification

Identification of search models is difficult due to the interdependence between search costs and preferences. Correspondingly, we rely on "exclusion restriction" and the variation in information sets caused by the experiment to separately identify preference from costs. When we choose different sets of covariates to enter the utility and the cost function, covariates that enter the cost function (but not the utility function) serve as exclusion restriction for identification.

Specifically, we choose the following functional form for c^{like}_m

$$c^{\text{like}}_m = \gamma^{\text{like}}_{1,m} + \gamma^{\text{like}}_{2,m} \cdot L^{\text{LS}}_{wm} \tag{2.22}$$

where

$$L_{wm}^{\text{LS}} = \begin{cases} 1 - \ell_{wm} & \text{if } h_m = \text{treated} \\ \mathbb{E}[1 - \ell_{wm} | X_m, X_w] & \text{if } h_m = \text{control} \end{cases} \tag{2.23}$$

L_{wm}^{LS} represents whether w had "not *ℓiked*" m: for users in the treatment group, L_{wm}^{LS} equals 1 when $\ell_{wm} = 0$ (w did not *ℓike* m) and 0 otherwise; for users in the control group, $L_{wm}^{\text{LS}} \in [0, 1]$ and becomes closer to 1 as the expected value of ℓ_{wm} becomes smaller. The first component of the messaging cost, $\gamma_{1,m}^{\text{like}}$, is the fixed cost of *$\ell$iking*. Although there is no reason to believe that the act of swiping a phone or clicking a button is costly, we let the data speak for its magnitude. $\gamma_{2,m}^{\text{like}}$ is the psychological burden of liking someone who hasn't indicated interest to a focal user through *ℓiking* him.

In our model, we have assumed that ℓ_{wm} affects users' *expected* utility (excluding costs) only through its impact on match probability (this match probability is estimated directly from the data), and does not directly enter the utility function. Since ℓ_{wm} (and $\mathbb{E}[1-\ell_{wm}|X_m, X_w]$) enters c_m^{like} but not the utility function, L_{wm}^{LS} serves as an exclusion restriction for identification.

One may argue, however, that ℓ_{wm} should enter the utility function, i.e. knowing that w had *ℓiked* m ($\ell_{mw} = 1$) may shift m's utility upwards. In fact, research in psychology has demonstrated that receiving information that another is attracted to you is a powerful determinant of attraction, a phenomenon often referred to as the "reciprocity of liking" (Backman and Secord (1959); Eastwick and Finkel (2009)). However, it is unclear whether the effect of reciprocity of liking shifts the *utility* of a match per se. In addition, there is another stream of research (e.g., unrequited love, Baumeister et al. (1993)) suggesting that liking is not *always* reciprocated. For these reasons we choose to not include ℓ_{wm} in our utility specification. Even if we were to include ℓ_{wm} in our utility specification, since there is no reason to believe that $\mathbb{E}[\ell_{wm}]$ should enter the utility function, the variation in L_{wm}^{LS} created by the experiment allows us to use it as an exclusion restriction.

We parameterize $\gamma_{1,m}^{\text{like}}$ and $\gamma_{2,m}^{\text{like}}$ as

$$\gamma_{1,m}^{\text{like}} = \tilde{\gamma}_1^{\text{like}} \cdot \frac{\exp(z_m'\lambda_1^{\text{like}})}{1+\exp(z_m'\lambda_1^{\text{like}})} \qquad \text{and} \qquad \gamma_{2,m}^{\text{like}} = \tilde{\gamma}_2^{\text{like}} \cdot \frac{\exp(z_m'\lambda_2^{\text{like}})}{1+\exp(z_m'\lambda_2^{\text{like}})} \tag{2.24}$$

where z_m are a set of m's characteristics and $\tilde{\gamma}_1^{\text{like}}$, $\tilde{\gamma}_2^{\text{like}}$, λ_1^{like} and λ_2^{like} are parameters to be estimated. The functional form makes the second term in both expressions to be positive for all values of z_m. Therefore, as long as $\tilde{\gamma}_1^{\text{like}} > 0$, we have $\gamma_{1,m}^{\text{like}} > 0$. Similarly, we have $\gamma_{2,m}^{\text{like}} > 0$ as long as $\tilde{\gamma}_2^{\text{like}} > 0$. For z_m, we use m's attractiveness level.

We assume a similar functional form for the cost of messaging, c_m^{msg} :

$$c_m^{\text{msg}} = \gamma_{1,m}^{\text{msg}} + \gamma_{2,m}^{\text{msg}} \cdot L_{wm}^{\text{MS}} \tag{2.25}$$

where

$$L_{wm}^{\text{MS}} = \begin{cases} 1-\ell_{wm} & \text{if } h_m = \text{treated} \\ 1-\ell_{wm} & \text{if } h_m = \text{control} \quad \& \quad d_{wm} = \text{like} \\ \mathbb{E}[1-\ell_{wm}|X_m, X_w] & \text{otherwise} \end{cases} \tag{2.26}$$

The interpretation of L_{wm}^{MS} is similar to L_{wm}^{LS} - it represents whether w had "not ℓ*iked*" m : for users in the treatment group, as well as users in the control group who have chosen $d_{wm} = \text{like}$, L_{wm}^{MS} takes value 1 when $\ell_{wm} = 0$ (w did not ℓ*ike* m) and 0 otherwise; for users in the control group who have chosen $d_{wm} = \text{nlike}$, $L_{wm}^{\text{MS}} \in [0,1]$ and becomes closer to 1 as the expected value of ℓ_{wm} becomes smaller. Users' choice of d_{wm} during the liking stage makes L_{wm}^{MS} different from L_{wm}^{LS}, creating additional variation that helps with identification.

The first component of the messaging cost, $\gamma_{1,m}^{\text{msg}}$, is time and effort to compose a message. The second component of this cost, $\gamma_{2,m}^{\text{msg}}$, is the psychological burden of messaging someone who hasn't indicated interest to a user through *liking* him. We parameterize $\gamma_{1,m}^{\text{msg}}$ and $\gamma_{2,m}^{\text{msg}}$ as

$$\gamma_{1,m}^{\text{msg}} = \tilde{\gamma}_1^{\text{msg}} \cdot \frac{\exp(z_m'\lambda_1^{\text{msg}})}{1+\exp(z_m'\lambda_1^{\text{msg}})} \qquad \text{and} \qquad \gamma_{2,m}^{\text{msg}} = \tilde{\gamma}_2^{\text{msg}} \cdot \frac{\exp(z_m'\lambda_2^{\text{msg}})}{1+\exp(z_m'\lambda_2^{\text{msg}})} \tag{2.27}$$

where $\tilde{\gamma}_1^{\text{msg}}$, $\tilde{\gamma}_2^{\text{msg}}$, λ_1^{msg} and λ_2^{msg} are parameters to be estimated.

Finally, the cost of browsing, c_m^{browse}, is convex in k_τ and is given as

$$c_m^{browse}(k_\tau) = (\phi_1 + \eta_{m\tau}) \cdot k_\tau + \phi_2 \cdot k_\tau^2 \tag{2.28}$$

where $\eta_{m\tau}$ is a cost shock that is unobserved by the researcher and is independently distributed across users and sessions as a mean zero normal random variable. ϕ_1 and ϕ_2 are parameters to be estimated. The parameters of the browsing cost are identified by the number of profiles that are browsed by each user. Since it is utility-maximizing for all users with browsing costs in a given range to browse a specific number of profiles, the variation in our data allows us to identify only a "range" of browsing costs which rationalizes a specific number of browses, and not the point estimates. The point estimates of the parameters of the browsing cost are identified by the functional form of the utility function and the distributional assumption of the unobserved part of the utility and the cost.

2.9 Estimation Results

Table 2.7 reports the maximum likelihood estimates of preference parameters (cost parameters are presented in Table 2.9). Columns 1 and 2 shows the results for men and women, respectively. Our estimation results are, generally, consistent with the findings of existing research (Fisman et al. (2006); Hitsch et al. (2010a); Hitsch et al. (2010b); Kurzban and Weeden (2005)):

Men place greater emphasis on partners' physical attributes (measured by age, body type and attractiveness) than do women. On the other hand, women place more emphasis on characteristics that reflects partners' social status and earning potential (which we measure using education level).

While users of both genders prefer younger partners, men place about four times as much weight on partners' younger age than do women. Regarding age difference, women prefer men who are older than themselves. Unexpectedly, however, male users of MonCherie also prefer female partners who are older than themselves.

We also find that attractiveness and BMI are important determinants of preference for both genders. Both men and women prefer more attractive partners. As expected, men place about four times as much weight on partners' attractiveness than women do. While both men and women prefer a partner who is more attractive than themselves, women display strong aversion towards men who are less attractive than themselves. Regarding BMI, men display aversion towards women with a large BMI, while women tend to prefer heavier men. In terms of differences in BMI, men avoid partners who have a larger BMI than themselves, whereas women do not show this tendency.

With regards education level, men's preferences over education level are opposite of women's. Men shy away from women with higher education level, while women prefer men with more years of education.[61] Negative signs on coefficients of education difference terms suggest that users prefer partners of similar educational level, but they are not statistically significant.[62]

[61] Although some coefficients on categorical education level variables are insignificant, we can see a clear pattern that men shy away from women with higher education level, while women prefer men with more years of education. Specifically, our results show that men prefer women with highschool and two-year college degrees to women with an undergraduate and a postgraduate degree (masters, law, medical school and a PhD). In contrast, women show strong preference towards men with a postgraduate degree.

[62] The coefficient on Education Difference (-) for men is positive, but it is not statistically significant. The coefficient on Education Difference (+) for men is statistically significant, suggesting that men do not like women who are less educated than them (this is consistent with the results in Hitsch et al. (2010a))

Both men and women generally have a relative distaste for a partner of a different ethnicity. However, some coefficients on ethnicity are positive (but most are not statistically significant).

Table 2.9 reports estimates of cost parameters. All cost parameter estimates, except for the fixed cost of messaging for men ($\tilde{\gamma}_1^{\text{msg}}$), are positive and significant, suggesting that these costs have important implications for user behavior. Our estimates suggest that women have higher costs of browsing, $\ell iking$, and messaging compared to men.

	Preference of men (1)		Preference of women (2)	
	Estimate	SE	Estimates	SE
Age	-2.542***	0.229	-0.590***	0.028
Age Difference (+)	-2.5352***	0.061	-1.334***	0.074
Age Difference (-)	3.253***	0.052	0.715***	0.059
HighSchool	2.657***	0.000	-2.719***	1.065
TwoYear	2.275***	0.980	-1.566	1.024
University	-0.044	0.623	0.210	0.577
Masters	-1.839*	1.033	0.468	0.891
Law	-11.740**	5.776	10.564***	3.473
Med	-13.133**	6.679	1.3667	4.431
PhD	-7.733	4.826	7.702***	2.758
Education Difference (+)	-2.491***	0.386	-0.461	0.376
Education Difference (-)	0.671	0.480	-0.561	0.321
Skinny	2.861***	0.754	1.543***	0.568
Average	1.657***	0.664	1.890***	0.668
Heavier	-0.079	1.134	12.505***	1.417
Overweight	-0.357	2.706	-0.328	3.301
BMI Difference (+)	0.045	0.634	0.623***	0.035
BMI Difference (-)	-0.077***	0.032	-0.029	0.033
Attractiveness	11.904***	0.146	3.0270***	0.090
Attractiveness Difference (+)	0.079	0.204	-13.185***	1.553
Attractiveness Difference (-)	3.637***	0.219	6.741***	0.169
Asian; mate Indian	5.992	9.848	-14.584	8.487
Asian; mate white	-25.995***	8.981	-0.850	1.368
Asian; mate black	-5.894	32.566	-2.294	3.963
Asian; mate Hispanic	-10.476	16.776	-5.187***	2.406
Asian; mate Native	-3.897	9.317	13.736	8.757
Asian; mate Pacific	-1e-05	2.601	-15.683	9.239
Asian; mate Mideast	-0.746	8.790	6.138	9.274
Asian; mate other	-33.073***	3.216	-6.812***	1.047
White; mate Asian	-11.912***	2.122	-15.947	10.406
White; mate Indian	-4.686	14.370	-5.878	6.918
White; mate black	-60.058***	6.987	-9.549***	3.354
White; mate Hispanic	-4.902***	2.035	-2.577	2.944
White; mate Native	6.012	4.989	-7.608	5.007
White; mate Pacific	-1.423	5.220	-6.952	14.388
White; mate Mideast	-2.753	8.542	-11.835***	4.551
White; mate other	-9.462***	0.827	0.602	0.801

Notes. For estimation, we used all users, but only 25% of the randomly selected sessions.
*Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Table 2.7: Model Estimates

	Preference of men (1)		Preference of women (2)	
	Estimate	SE	Estimates	SE
Black; mate Asian	2.865	11.137	0.869	8.52
Black; mate Indian	1.882	5.229	9.781	14.624
Black; mate white	-7.256	4.601	6.077	3.934
Black; mate Hispanic	12.184	7.931	-4.410	9.110
Black; mate Native	-0.560	4.330	0.064	0.436
Black; mate Pacific	1.343	10.568	-0.391	18.653
Black; mate Mideast	1.446	13.826	-2.285	19.551
Black; mate other	0.490	2.278	-1.204	2.333
Hispanic; mate Asian	-9.066	15.028	0.937	12.986
Hispanic; mate Indian	-3.161	19.708	-0.130	14.709
Hispanic; mate white	-27.178***	3.287	12.155***	2.4056
Hispanic; mate black	-19.299***	9.533	2.415	12.936
Hispanic; mate Native	1.360	2.499	1.657	13.755
Hispanic; mate Pacific	0.057	6.322	-1.159	8.613
Hispanic; mate Mideast	-0.387	1.797	-3.518	6.123
Hispanic; mate other	-15.565***	2.820	3.835***	1.827
Native; mate Asian	5.312	19.024	-0.300	10.1500
Native; mate Indian	0.302	7.681	-0.836	19.528
Native; mate white	9.197	5.377	15.833***	4.322
Native; mate black	3.630	15.788	-24.790***	9.041
Native; mate Hispanic	-2.561	22.392	7.615	7.536
Native; mate Pacific	-0.116	4.637	–	–
Native; mate Mideast	0.638	5.368	-25.913***	10.983
Native; mate other	1.871	6.821	24.430***	7.052
Pacific; mate Asian	1.027	3.400	-0.014	10.528
Pacific; mate Indian	-0.176	14.473	–	–
Pacific; mate white	-4.843	19.624	4.260	18.345
Pacific; mate black	-0.579	2.102	–	–
Pacific; mate Hispanic	4.041	13.910	-4.233	12.899
Pacific; mate Native	-0.685	2.552	-0.919	9.757
Pacific; mate Mideast	0.503	3.641	–	–
Pacific; mate other	-4.821	12.137	-3.052	14.489
Mideast; mate Asian	2.247	5.814	-0.476	8.708
Mideast; mate Indian	0.798	3.699	0.002	1.465
Mideast; mate white	23.906	3.827	-7.353	7.408
Mideast; mate black	-4.326	9.967	-0.005	10.931
Mideast; mate Hispanic	1.092	5.861	-3.709	16.842
Mideast; mate Native	1.678	6.747	–	–
Mideast; mate Pacific	-0.683	3.147	–	–
Mideast; mate other	-4.416	3.314	-5.447	9.577

Notes. For estimation, we used all users, but only 25% of the randomly selected sessions.
*Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Table 2.8: Model Estimates (Continued)

	Men		Women	
	Estimates	SE	Estimates	SE
	Cost of Browsing			
ϕ_1	0.223***	0.012	0.933***	0.002
ϕ_2	0.036***	0.006	0.074***	0.000
	Cost of Liking			
$\tilde{\gamma}_1^{\text{like}}$	0.718***	0.028	2.684***	0.007
$\tilde{\gamma}_2^{\text{like}}$	0.161***	0.028	0.543***	0.263
λ_1^{like}	0.183***	0.005	3.158***	0.028
λ_2^{like}	0.174***	0.010	-2.059***	0.067
	Cost of Messaging			
$\tilde{\gamma}_1^{\text{msg}}$	0.911	0.620	3.379***	0.029
$\tilde{\gamma}_2^{\text{msg}}$	1.390***	0.444	1.100***	0.042
λ_1^{msg}	2.263***	0.075	3.241***	0.050
λ_2^{msg}	23.075***	0.876	0.340***	0.055
LL	-826181.0244		-304963.5235	

Notes. For estimation, we used all users, but only 25% of the randomly selected sessions.
*Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Table 2.9: Cost Estimates

2.10 Predicted Matching Patterns

We want to compare equilibrium matches under different protocols to quantify the relative impact of frictions and preferences on assortative matching. Specifically, we want to compare the following sets of equilibrium matches: (i) equilibrium matches in a default setting (when everyone is in the control group), (ii) equilibrium matches when both sides of the market (both men and women) are gifted with the treatment, and (iii) equilibrium matches in a frictionless environment. Note that we are simulating equilibrium matches for cases (i) and (ii) instead of using the actual matches that are observed in the data. This is because due to the existence of search costs, a user browses only a limited number of profiles, which make up his consideration set. Since observed matches in the data are matches achieved selectively among the profiles within a user's consideration set, the pool of matched partners is not identical to the initial pool of available potential partners in the market. To make a correct comparison across different protocols, the initial pool of available potential partners must be identical across different protocols. Therefore, in the following sections we assume that the pool of men and women attempting to find a partner are all the users who are part of the experiment (16,119 men and 7,112 women). Because the experiment was conducted on a randomly selected sample of newly registered users, it is safe to assume that they are representative of the entire population of users of MonCherie.

We simulate equilibrium matches in a frictionless environment using the Gale-Shapley deferred-acceptance algorithm. The Gale-Shapley marriage problem assumes the presence of a central matchmaker which recommends a matching to agents given individuals' preferences over potential partners, and hence does not describe the online matching process wherein agents have to incur costs to find a partner in the absence of a central matchmaker. Adachi (2003) shows, however, that as search costs become negligible, the set of equilibrium matches obtained in a two-sided search and matching model is identical to the set of stable equilibrium matches predicted by the Gale-Shapley algorithm. Moreover, repeated rounds of offer-making and corresponding rejections of the deferred-acceptance algorithm resemble the search and messaging behavior of our users. This not only allows us to simulate equilibrium matches under a frictionless environment, but also provides a theoretical efficiency benchmark that can be used to evaluate whether frictions in our platform lead to significant departures from efficiency. Predicted matches under the control and treatment settings are obtained by introducing frictions to the deferred-acceptance algorithm.[63] introduce ad-hoc constraints to the deferred-acceptance algorithm to account for search frictions.

In what follows, we first explain how we compute predicted matches in a frictionless environment using the Gale-Shapley deferred-acceptance algorithm. Then, we describe how we compute equilibrium matches when users engage in costly search (and costly *ℓiking* and

[63] Banerjee et al. (2013)

messaging). Predicted matches from these simulations will then be used to answer questions regarding the relative impact of frictions and preferences.

2.10.1 Empirical Strategy

Before we compute the stable matches using the Gale-Shapley deferred-acceptance algorithm, we first need to construct ordinal preferences (rankings) over the entire set of women (men) for each man (woman) using estimated preference parameters from the model. Specifically, estimated preference parameters from the model are used to construct the predicted utility that each man would get from matching with each woman in the sample (and vice versa for women) using the following equation:[64]

$$\hat{U}_M(m, w) = U_M(X_m, X_w; \hat{\Theta}_M)$$

Predicted utility $\hat{U}_M(m, w)$ is then transformed into an ordinal ranking $R_m(w)$ of user m with respect to woman w as

$$R_m(w) = n \qquad \text{if} \qquad \begin{cases} \hat{U}_M(m, w') > \hat{U}_M(m, w) > \hat{U}_M(m, w'') \\ \text{and } R_m(w') = n - 1 \text{ and } R_m(w'') = n + 1 \end{cases}$$

where n is an integer. We apply this methodology to all users in the sample to obtain a full set of ordinal preferences for each user with respect to all users of the opposite gender.

The man-optimal stable matching using deferred-acceptance algorithm is executed as follows:[65]

[64]Note that ordinal preferences are constructed based on the predicted "utility" as opposed to "expected utility".

[65]We obtain predicted utility values for each draw of the parameter, calculate the average utility and the corresponding average ranking, and run the deferred-acceptance algorithm once. The woman-optimal stable matching is obtained similarly.

1. All men first propose to their most highly-ranked woman, as long as $\hat{U}_M(m, w) \geq V_M(m)$

2. Among all the offers that each woman receives, she selects the most highly-ranked man, as long as $\hat{U}_W(w, m) \geq V_W(w)$

3. All men who haven't been selected then propose to their second most-highly ranked woman

4. If a woman receives a new offer that is higher-ranked than the one she is currently holding, the woman releases the old offer and keeps the new offer. Released man then has to propose to the next woman in his ranking list

5. This process continues until all men go through all women such that $\hat{U}_M(m, w) \geq V_M(m)$

Ties are broken randomly. The above process describes how we obtain a set of stable matches implied by the estimated preferences when frictions are negligible.

We next describe how we incorporate frictions to this algorithm when: (i) everyone is in the control group, and (ii) users on both sides (both men and women) are gifted with the treatment. Note that ordinal preferences of users in the presence of frictions are constructed using the "expected" utility that each user will get with each potential partner. Since a match probability depends on the value of ℓ_{wm}, the expected utility will also depend on ℓ_{wm}. However, we do not observe the value of ℓ_{wm} because users in our experiment did not interacted with each other (except a few who did). We solve this problem by simulating the values of ℓ_{wm}. That is, we simulate the $\ell ikes$ that a user receives from the opposite gender *before* he engages in search, for each pair of m and w. Specifically, the simulation of ℓ_{wm} and matching in an environment with frictions is executed as follows:

1. Construct ℓ_{wm} for each pair of users (w, m) as follows:

(a) Draw random utility terms, e_{wm}, for each pair of w and m[66] who assume that the noise in the utility function comes from a measurement error, we follow Hitsch et al. (2010a) who assume that the error term is a "structural noise". This is because several important dimensions of the profile (such as picture or income) that affect choice and are observed to users are unobserved by the econometrician, whereas researchers in Banerjee et al. (2013) observe everything that is observed by the agent.

(b) For each w, using the draw and estimated preference parameters, construct $\mathbb{E}_{\Delta_{hw}^{LS}}\left[\hat{EU}_{wm}^{MS}(\text{msg}|\text{like})|\Omega^{LS}\right]$, $\mathbb{E}_{\Delta_{hw}^{LS}}\left[\hat{EU}_{wm}^{MS}(\text{nmsg}|\text{like})|\Omega^{LS}\right]$, $\mathbb{E}_{\Delta_{hw}^{LS}}\left[\hat{EU}_{wm}^{MS}(\text{msg}|\text{nlike})|\Omega^{LS}\right]$ and $V_W(w)$

(c) w decides to *like* m if the expected utility from *liking* is greater than its cost:

$$\ell_{wm} = \begin{cases} 1 & \text{if } \max\left\{\mathbb{E}_{\Delta_{hw}^{LS}}\left[\hat{EU}_{wm}^{MS}(\text{msg}|\text{like})|\Omega^{LS}\right], \mathbb{E}_{\Delta_{hw}^{LS}}\left[\hat{EU}_{wm}^{MS}(\text{nmsg}|\text{like})|\Omega^{LS}\right]\right\} \\ & \quad - \max\left\{\mathbb{E}_{\Delta_{hw}^{LS}}\left[\hat{EU}_{wm}^{MS}(\text{msg}|\text{nlike})|\Omega^{LS}\right], V_W(w)\right\} + \varepsilon_{wm}^{\text{like}} - \varepsilon_{wm}^{\text{nilke}} > c_w^{\text{like}} \\ 0 & \text{otherwise} \end{cases}$$

2. Taking ℓ_{wm} constructed in the previous step as given, construct a consideration set for each m

(a) Draw random utility term, e_{mw}, for each pair of m and w.

(b) Create a grid $k = 0, 1, 2, ..., k_{max}$ for each m[67]

(c) For each m

i. Sort profiles by z_{mw} in a descending order, and calculate $\mathbb{E}[z_{mw}]$ for profiles that haven't been browsed yet

[66] As opposed to Banerjee et al. (2013)

[67] The marjority of users browsed less than 100 profiles during each session. Hence we set $k_{max} = 100$.

ii. Draw cost shock $\eta_{m\tau}$, and for each k in the grid calculate $\Gamma_{m\tau,k} = k \cdot \mathbb{E}[z_{mw}] - (\phi_1 + \eta^q_{1,m\tau}) \cdot k - (\phi_2) \cdot k^2$

iii. The optimal $k_{m\tau} = k^*$ is such that $\Gamma^q_{m\tau,k^*} \geq \max\left(\Gamma^q_{m\tau,k}\right) \quad \forall k \neq k^*$

iv. Repeat steps ii−iv until $k^* = 0$ or until m goes through all the profiles

v. Calculate $K_m = \sum k_{m\tau}$ which is the entire consideration set that m will search across all sessions

3. Create ordinal preferences for each user for all users of opposite gender

(a) For each pair m and w, m first chooses $d_{mw} \in \{\text{like,nlike}\}$. m chooses $d_{mw} = \text{like}$ iff

$$\max\left\{\mathbb{E}_{\Delta^{LS}_{h_m}}\left[\hat{EU}^{MS}_{mw}(\text{msg}|\ell\text{ike})|\Omega^{LS}\right], \mathbb{E}_{\Delta^{LS}_{h_m}}\left[\hat{EU}^{MS}_{mw}(\text{nmsg}|\text{like})|\Omega^{LS}\right]\right\}$$
$$-\max\left\{\mathbb{E}_{\Delta^{LS}_{h_m}}\left[\hat{EU}^{MS}_{mw}(\text{msg}|\text{nlike})|\Omega^{LS}\right], V_M(m)\right\} + \varepsilon^{\text{like}}_{mw} - \varepsilon^{\text{nlike}}_{mw} > c^{\text{like}}_m$$

(b) Conditional on d_{mw}, m then chooses $\mu_{mw} \in \{\text{msg,nmsg}\}$. m chooses $\mu_{mw} = \text{msg}$ iff

$$\hat{EU}^{MS}_{mw}(\text{msg}|d_{mw}) \geq \hat{EU}^{MS}_{mw}(\text{nmsg}|d_{mw})$$

(c) For each m, compute the predicted expected utility $E\hat{U}_{mw}$

$$E\hat{U}_{mw} = U_M(m,w) \cdot \tilde{\pi}^{MS}_{mw}(d_{mw}, \mu_{mw}) + V_M(m) \cdot [1 - \tilde{\pi}^{MS}_{mw}(d_{mw}, \mu_{mw})]$$
$$- c^{\text{like}}_m 1\{d_{mw} = \text{like}\} - c^{\text{msg}}_m 1\{\mu_{mw} = \text{msg}\}$$

(d) Transform $E\hat{U}_{mw}$ into ordinal ranking $R_m(w)$ such that

$$R_m(w) = n \qquad \text{if} \qquad \begin{cases} E\hat{U}_{mw'} > E\hat{U}_{mw} > E\hat{U}_{mw''} \\ \text{and } R_m(w') = n-1 \text{ and } R_m(w'') = n+1 \end{cases}$$

where n is an integer

4. Compute equilibrium matches:

   (a) Define C_m as the set of all profiles that m either *ℓiked* or messaged and let $I_m = C_m \cap K_m$

   (b) All men first propose (either message or *ℓike*) to their most highly-ranked woman within the set I_m

   (c) Women considers all offers they receive. If a woman received a message from a man, the net utility from selecting this man is $U_W(w, m) - V_W(w)$. On the other hand, if a woman only receives a *ℓike* from a man but did not receive a message, the net utility from selecting this man is $U_W(w, m) - c_w^{msg} - V_W(w)$. This is because if a woman only receives a *ℓike*, she must initiate a conversation to a man, which is costly. Woman selects a man that gives highest net utility as long as it is greater than zero.

   (d) All men who haven't been chosen by women then propose to the next best woman within the set I_m

   (e) If a woman receives a new offer that is preferable to the one she is currently holding, she releases the previous offer. The released man then has to propose to the next woman on his list within his set I_m

   (f) This continues until each man m exhausts the list of women in his set I_m

Ties are broken randomly.[68]

	Man-optimal			
	control (1)	treated (2)	Gale-Shapley (3)	Fisher's z (4)
Age	0.5986	0.5541	0.5146	**-5.585**
	(0.0178)	(0.0146)	(0.0146)	
	[3,079]	[3,359]	[6,554]	
Education	0.2521	0.2094	0.1444	**-2.143**
	(0.0376)	(0.0397)	(0.0382)	
	[533]	[555]	1,173	
Attractiveness	0.7789	0.7920	0.5418	**-30.057**
	(0.0079)	(0.0075)	(0.0098)	
	[2,917]	[3,028]	[6,336]	
BMI	-0.1430	-0.1277	-0.0701	0.525
	(0.0913)	(0.0724)	(0.0821)	
	[103]	[88]	[106]	
Ethnicity	0.3947	0.4141	0.4419	**2.2962**
	(0.4891)	(0.4929)	(0.4967)	
	[793]	[838]	[2,159]	

Notes. Pearson correlation coefficient. Bootstrap standard errors in parentheses. Number of observations in brackets. If the difference between the treatment and control group is significant (at the 5 percent level), Fisher's z-statistics are in bold.

Table 2.10: Attribute Correlations in Predicted Matches

2.10.2 Predicted Matching Patterns

Table 2.10 shows attribute correlation patterns between couples for predicted male-optimal matches. Columns (1), (2) and (3) show correlations in user attributes for matches under the control, treatment and frictionless (Gale-Shapley) settings, respectively. Column (4) reports Fisher's z-statistic that tests the difference in correlations between the control (Column 1) and frictionless (Column 3) protocols. Number of matches increase as we gradually reduce frictions, from control (3,079) to treated (3,359), and from treated to frictionless (6,554) setting. Some users who got matched under one protocol did not a get a match under an alternative protocol

[68] We obtain predicted utility values for each draw of the parameter, calculate the average utility and the corresponding average ranking, and run the deferred-acceptance algorithm once.

(and vice versa). As expected, frictions play a significant role in assortative matching. The magnitude of predicted age correlation ($\rho = 0.60$ under control setting) is roughly similar to the actual correlation in the data ($\rho = 0.71$). Age correlation under the treatment ($\rho = 0.55$) is approximately 7 percent less than that of the control ($\rho = 0.60$). Completely removing frictions further reduces age correlation ($\rho = 0.51$) to approximately 14 percent less compared to the control setting, and this difference is statistically significant. The magnitude of the predicted correlation in education ($\rho = 0.25$ under control setting) is greater than the actual education level correlation in the data ($\rho = 0.12$). The correlation in education level under the treatment ($\rho = 0.21$) and frictionless (0.15) protocols are both lower than that of the control ($\rho = 0.25$). Education correlation level under the frictionless setting is approximately 42 percent lower compared to the control setting. Our model overpredicts the correlation in attractiveness level. Nevertheless we see a consistent pattern: completely removing frictions reduces attractiveness correlation ($\rho = 0.54$) by approximately 30 percent compared to that of the control setting ($\rho = 0.78$), and this difference is statistically significant. Our model does not perform well in predicting correlation patterns in BMI, and predicts negative correlations in BMI (however, the correlations are not statistically significant). Since we are not able to calculate correlation coefficients for ethnicity due to its categorical nature, we compare the mean of a binary variable that equals 0 if matched couples are of same ethnicity and 1 otherwise. The mean of this binary variable represents the proportion of users who matched with a partner of identical ethnicity. Approximately 44 percent of the users got matched with a partner of different ethnicity under the frictionless protocol, which is roughly 13 percent more than that of the control protocol (39 percent), and this difference is statistically significant.

Unlike certain attributes for which preferences may be horizontal, it is almost universally agreed that preference for attractiveness is vertical, i.e. everyone ranks attractiveness using the same criterion. Under vertical preferences, theoretically, reduction of friction to lead to more sorting. This is because, in the absence of frictions, the most attractive man would match with the most attractive woman who is willing to accept him and vice versa. However, this may not hold in our case since different number of matches arise under different protocols in our simulations, and some individuals get matches in one but not in other protocols. To see whether reducing friction allows users to match with more attractive (vertically preferred) partners, we sum the attractiveness of a man and a woman and compare the average sum of attractiveness across protocols. We find that the average sum of attractiveness under the control setting is 13.36 (sd:5.29), 12.73 (sd: 5.29) under the treatment setting, and 13.75 (sd: 4.0) under the frictionless setting. It is unclear why the average attractiveness of a partner declined from the control to treatment condition, but we do find that the average attractiveness of a partner has increased from the control to frictionless setting, and this improvement is statistically significant.

We also use the absolute value of attribute difference ($\Delta = |X_m - X_w|$) as an alternative measure of sorting and find similar patterns. Table 2.11 reports the results. Columns (1), (2) and (3) display the means of attribute differences between matched couples, under alternative protocols: control, treatment and frictionless protocols, respectively. Column (4) reports the t-statistic that tests the significance of differences in correlations between the control (Column 1) and the frictionless (Column 3) protocols. Consistent with our previous results, frictions play a significant role in assortative matching. The magnitude of our predicted age difference ($\Delta = 4.30$ under control setting) is approximately similar to the actual age difference in the data. Age difference under the treatment ($\Delta = 4.8$) is approximately 12 percent greater than under the control ($\Delta = 4.30$) setting. Age difference under the frictionless setting ($\Delta = 6.46$) is approximately 50 percent greater than under the control setting, and this difference is statistically significant. Differences in years of education under the frictionless

	Mean Attribute Difference			
	control (1)	treated (2)	Gale-Shapley (3)	t-stat (4)
Age	4.2959	4.7681	6.4599	**15.453**
	(3.7529)	(3.6124)	(7.3655)	
	[3,079]	[3,359]	[6,554]	
Education	1.2420	1.1928	0.9327	**-4.1699**
	(1.4180)	(1.4360)	(1.4213)	
	[533]	[555]	1,173	
Attractiveness	1.1628	1.1853	2.7112	**33.080**
	(1.5008)	(1.4188)	(2.3137)	
	[2,917]	[3,028]	[6,336]	
BMI	4.1748	4.375	3.2453	-1.648
	(4.2781)	(4.0693)	(3.8688)	
	[103]	[88]	[106]	

Notes. Pearson correlation coefficient. Standard deviation in parentheses. Number of observations in brackets. If the difference between the treatment and control group is significant (at the 5 percent level), the t-statistics are in bold.

Table 2.11: Attribute Differences in Predicted Matches

setting ($\Delta = 0.93$) is approximately 25 percent lower than under the control ($\Delta = 1.24$) setting. Attractiveness difference under the treatment ($\Delta = 1.19$) is approximately 2.6 percent greater compared to the control ($\Delta = 1.16$) setting. Completely removing frictions further increases this difference ($\Delta = 2.71$), to approximately 130 percent greater than that under the control setting, and this difference is statistically significant.

2.11 Welfare Analysis

2.11.1 Rank Differences of Matched Outcomes

To study whether reducing frictions makes users better off, we want to see whether reducing frictions results in matches with more preferred partners. Closely following Hitsch et al.

	Mean	SD	t-stat	Nobs
Panel A: Total				
ΔR^{tr-ct}	-0.7962	14.3811	**-8.439**	23,235
ΔR^{GS-tr}	-6.1846	18.5691	**-50.769**	23,235
ΔR^{GS-ct}	-6.9808	17.9721	**-59.208**	23,235
Panel B: Match in either protocol				
ΔR^{tr-ct}	-2.0905	23.2459	**-8.460**	8,849
ΔR^{GS-tr}	-9.8182	22.6215	**-52.508**	14,636
ΔR^{GS-ct}	-11.3228	21.7886	**-62.197**	14,325

Notes. The table shows summary statistics on the predicted rank differences across alternative protocols.

Table 2.12: Rank Differences

(2010b), we implement this as follows: For each user, we assign the ordinal ranking to all potential partners based on the predicted utility. The most desirable partner will be ranked 1st, the second desirable partner will be ranked 2nd, and so forth. For each user $i \in \{m, w\}$, let R_i^{ct} be the rank of i's matched partner predicted under the control setting, R_i^{tr} be the rank of i's matched partner predicted under the treatment setting, and R_i^{GS} be the rank of i's matched partner predicted under the frictionless setting. For users who did not get a match, we assign the ranking of his/her reservation value $V_I(i)$. Denote $\Delta R_i^{tr-ct} = 100 \times (R_i^{tr} - R_i^{ct})/N_J$ as the difference between the ranks achieved under the treatment and control protocol, expressed in terms of percent (of number of users in the opposite gender), where N_J is the number of potential partners. Likewise, let $\Delta R_i^{GS-ct} = 100 \times (R_i^{GS} - R_i^{ct})/N_J$ be the difference in ranks achieved under frictionless and control protocols, expressed in terms of percent (of number of users in the opposite gender). In table 2.12 Panel A, we report means, standard deviations, and t-statistics of predicted average rank differences across protocols. The mean is computed as $\Delta \bar{R}^{tr-ct} = (N_I + N_J)^{-1} \times \left(\sum_{i=1}^{N_I} \Delta R_i^{tr-ct} + \sum_{j=1}^{N_J} \Delta R_j^{tr-ct} \right)$. If $\Delta \bar{R}^{tr-ct}$ is negative, the treatment could have improved, on average, on the allocation achieved under control protocol. Likewise, let $\Delta \bar{R}^{GS-ct} = (N_I + N_J)^{-1} \times \left(\sum_{i=1}^{N_I} \Delta R_i^{GS-ct} + \sum_{j=1}^{N_J} \Delta R_j^{GS-ct} \right)$ be the average difference in ranks achieved under frictionless and control protocols. If

$\Delta\bar{R}^{GS-ct}$ is negative, frictionless setting could have improved, on average, on the allocation achieved under control protocol. Similar interpretation can be applied to $\Delta\bar{R}^{GS-tr} = (N_I + N_J)^{-1} \times \left(\sum_{i=1}^{N_I} \Delta R_i^{GS-tr} + \sum_{j=1}^{N_J} \Delta R_j^{GS-tr}\right)$. We find that the treatment improves average ranks achieved by approximately 0.8 percent relative to control setting. Although the magnitude of this improvement is small, we find that it is statistically significant. The Gale-Shapley protocol further improves average ranks by approximately 7 percent relative to control setting. These numbers suggest that removing frictions improves on the outcomes (in terms of rankings) achieved under control setting.

A large number of users did not get a match in any of the alternative protocols, which results in a ranking difference of zero. A large number of users with zero ranking difference masks the average rank difference. Therefore, in table 2.12 Panel B, we report the statistics for users who got a match in at least one of the protocols under comparison. For example, ΔR^{tr-ct} computes the difference in ranks achieved under the frictionless and control protocols (expressed in terms of percent) for only those users who got a match in either treatment, control, or both protocols. Consistent with our previous findings, we find that the treatment improves average ranks achieved by approximately 2 percent relative to the control setting, and the Gale-Shapley protocol further improves average ranks by approximately 11 percent relative to the control setting.

The results presented in Table 2.12 consider both, users who got a match and users who did not get a match. Since the removal of the friction results in more number of successful matches, the efficiency gain from the removal of frictions may be a mere outcome of more number of matches.[69] To explore this issue further, we also compare the average achieved ranking (expressed in terms of percent as before) only for users who did get a match. The

[69] If a match is achieved, it means that the utility from this match is greater than the expected value of remaining single.

average ranking under the control setting is 0.343 (sd: 0.259, N=6,158) and 0.229 (sd: 0.204, N=13,106) under the Gale-Shapley protocol, which is a significant improvement from that of the control protocol by 33 percent. This suggests that the improvement in efficiency presented in Table 2.12 is not solely driven by increase in number of matches.

2.12 Conclusion

This paper investigates the impact of search frictions on the formation of a match in two-sided markets. With agents on both sides having private preferences regarding each others' chracteristics, finding a match based on mutual agreement requires extensive costly search. Using data from an online dating platform, we estimate a model of costly search that incorporates preference heterogeneity across users. Our estimation results reveal that frictions play a significant role in shaping matching outcomes. Our predicted matches suggest that matches achieved in a frictionless environment display significantly lower attribute correlations between couples across various dimensions (age, education level, ethnicity, attractiveness), compared to matches achieved in a market with frictions. We also find that removing frictions lead to significant gains in terms of partner rankings.

Our findings can provide important managerial implications for the pricing of premium features, in how much users are willing to pay for an additional piece of information about the preferences of the other side. In addition, our findings can shed light on what type of information should be displayed on users' profile. Information that is helpful in gauging the preferences of other users can greatly improve consumer experience. Furthermore, with one-third of the marriages in the U.S. happening online, our paper shows how the design of an online platform can contribute to diversity, which can in turn alleviate persistent social inequality.

2.13 References

Adachi, H. (2003): "A search model of two-sided matching under nontransferable utility," *Journal of Economic Theory*, 113, 182–198.

Arcidiacono, P., A. Beauchamp, and M. McElroy (2016): "Terms of endearment: An equilibrium model of sex and matching," *Quantitative Economics*, 7, 117–156.

Backman, C. W. and P. F. Secord (1959): "The Effect of Perceived Liking on Interpersonal Attraction," *Human Relations*, 12, 379–384.

Banerjee, A., E. Duflo, M. Ghatak, and J. Lafortune (2013): "Marry for What? Caste and Mate Selection in Modern India," *American Economic Journal: Microeconomics*, 5, 33–72.

Bapna, R., P. Mojumder, J. Ramaprasad, and A. Umyarov (2019): "StrongSignaling and Identity-Revelation in Online Dating:Evidence from a Randomized Field Experiment," Working Paper.

Bapna, R., J. Ramaprasad, G. Shmueli, and A. Umyarov (2016): "One-Way Mirrors in Online Dating: A Randomized Field Experiment," *Management Science*, 62, 3085–3391.

Baumeister, R. F. and D. Dhavale (2001): Interpersonal Rejection, *Oxford University Press*.

Baumeister, R. F., S. R. Wotman, and A. M. Stillwell (1993): "Unrequited Love: On Heartbreak, Anger, Guilt, Scriptlessness, and Humiliation," *Journal of Personality and Social Psychology*, 64, 377–394.

Bloch, F. and H. Ryder (2000): "Two-Sided Search, Marriages, and Matchmakers," *International Economic Review*, 41, 93–115.

Bojd, B. and H. Yoganarasimhan (2019): "Star-Cursed Lovers: Role of Popularity Information in Online Dating," Working Paper.

Burdett, K. and M. G. Coles (1997): "Marriage and Class," *The Quarterly Journal of Economics*, 112, 141–168.

Cacioppo, J. T., S. Cacioppo, G. Gonzaga, and E. a. V. Ogburn (2013): "Marital Satisfaction and Break-ups Differ Across On-line and Off-line Meeting Venues," *Proceedings of the National Academy of Sciences*, 110, 10135–10140.

Chan, T. Y., B. H. Hamilton, and N. W. Papageorge (2015): "Health, Risky Behavior and the Value of Medical Innovation for Infectious Disease," *Review of Economic Studies*, 83, 1456–1510.

Chen, Y. and S. Yao (2017): "Sequential Search with Refinement: Model and Application with Click-Stream Data," *Management Science*, 63, 4345–4365.

Choo, E. and A. Siow (2006): "Who Marries Whom and Why," *Journal of Political Economy*, 114, 175–201.

Dinerstein, M., L. Einav, J. Levin, and N. Sundaresan (2018): "Consumer Price Search and Platform Design in Internet Commerce," *American Economic Review*, 108, 1820–1859.

Eastwick, P. W. and E. J. Finkel (2009): "Reciprocity of Liking," *Encyclopedia of Human Relationships*, 3, 1333–1336.

Fisman, R., S. S. Iyengar, E. Kamenica, and I. Simonson (2006): "Gender Differences in Mate Selection: Evidence From a Speed Dating Experiment," *The Quarterly Journal of Economics*, 121, 673–697.

——— (2008): "Racial Preferences in Dating," *Review of Economic Studies*, 75, 117–132.

Flinn, C. J. and D. Del Boca (2012): "Endogenous household interaction," *Journal of Econometrics*, 166, 49–65.

Fong, J. (2018): "Search, Selectivity, and Market Thickness in Two-Sided Markets," working paper.

Fradkin, A. (2015): "Search, Matching, and the Role of Digital Marketplace Design in Enabling Trade: Evidence from Airbnb," Working Paper.

Fréchette, G. R., A. E. Roth, and M. U. Ünver (2008): "Unraveling yields inefficient matchings: evidence from post-season college football bowls," *The RAND Journal of Economic*, 38, 967–982.

Gale, D. and L. S. Shapley (1962): "College Admissions and the Stability of Marriage," *The American Mathematical Monthly*, 69, 9–15.

Hitsch, G. J., A. Hortaçsu, and D. Ariely (2010a): "Matching and Sorting in Online Dating," *American Economic Review*, 100, 130–63.

——— (2010b): "What makes you click?-Mate preferences in online dating," *Quantitative Marketing and Economics*, 8, 393–427.

Honka, E. (2014): “Quantifying search and switching costs in the US auto insurance industry,” *The RAND Journal of Economics*, 45, 847–884.

Honka, E., A. Hortaçsu, and M. A. Vitorino (2017): “Advertising, consumer awareness, and choice: evidence from the U.S. banking industry,” *The RAND Journal of Economics*, 48, 611–646.

Horton, J. J. (2014): “Misdirected Search Effort in a Matching Market: Causes, Consequences and a Partial Solution,” *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, 357–357.

Kim, J. B., P. Albuguerque, and B. J. Bronnenberg (2010): “Online Demand Under Limited Consumer Search,” *Marketing Science*, 29, 1001–1023.

Kurzban, R. and J. Weeden (2005): “HurryDate: Mate preferences in action,” *Evolution and Human Behavior*, 26, 227–244.

Lee, S. (2015): “Effect of Online Dating on Assortative Mating: Evidence from South Korea,” *Journal of Applied Econometrics*, 31, 1120–1139.

Lee, S. and M. Niederle (2015): “Propose with a rose? Signaling in internet dating markets,” *Experimental Economics*, 18, 731–755.

McNamara, J. M. and E. J. Collins (1990): “The job search problem as an employerc andidate game,” *Journal of Applied Probability*, 27, 815–827.

Mehta, N., S. Rajiv, and K. Srinivasan (2003): “Price Uncertainty and Consumer Search: a Structural Model of Consideration Set Formation,” *Marketing Science*, 22, 58– 84.

Pires, T. (2016): “Costly search and consideration sets in storable goods markets,” *Quantitative Marketing and Economics*, 14, 157–193.

Raquel, F. (2003): “Household formation, inequality, and the macroeconomy,” *Journal of the European Economic Association*, 1, 683–697.

Raquel, F. and R. Rogerson (2001): “Sorting and Long-Run Inequality,” *Quarterly Journal of Economics*, 116, 1305–1341.

Richards Shubik, S. (2015): “Peer effects in sexual initiation: Separating demand and supply mechanisms,” *Quantitative Economics*, 14, 663–702.

Roberts, J. H. and J. M. Lattin (1991): “Development and Testing of a Model of Consideration Set Composition,” *Journal of Marketing Research*, 28, 429–440.

Stigler, G. (1961): "The Economics of Information," *Journal of Political Economics*, 69, 213–225.

Weitzman, M. L. (1979): "Optimal Search for the Best Alternative," *Econometrica*, 47, 641–564.

Wong, L. Y. (2003): "Structural Estimation of Marriage Models," *Journal of Labor Economics*, 21, 699–727.

Chapter 3

Using Machine Learning to Address Customer Privacy Concerns: An Application with Click-stream Data

3.1 Introduction

As consumers constantly generate massive amounts of data, unprecedented opportunities exist for firms to harness the power of individual-level consumer data to predict their behavior and to target and customize service to consumers. The rapid growth of the use of consumer data, however, has also increased debate surrounding the protection of consumers' privacy. The Privacy Rights Clearinghouse reports that 8,909 data-breach incidents have been made public since 2005, compromising billions of sensitive personal records.[70] The scope and the extent of data breaches are alarming. For instance, millions of users were affected by the 2017

[70] Source: https://www.privacyrights.org/data-breaches, accessed on December 1, 2018.

Equifax data-breach incident that exposed sensitive personal information such as driver's license numbers, credit history, and even social security numbers.[71]

Consumers have expressed serious concerns pertaining to how firms handle consumers' data and protect their privacy. According to an online survey conducted by IBM in 2018, 78% of U.S. consumers said that a company's ability to keep consumer data private is "extremely important," and only 20% responded that they "completely trust" the companies they interact with to keep their private data safe.[72] Another survey by Consumer Reports finds that in the aftermath of Facebook's Cambridge Analytica scandal in 2018, in which the British consulting company deceitfully acquired and used millions of Facebook users' data, 70% of Facebook users have changed their behavior, taking more precautions with their posts, revising privacy settings, and turning off location tracking.[73] These examples show that with growing concerns over privacy issues, consumers have become skeptical of firms' promises about the use and protection of consumers' personal data. As a result, firms are now facing a crisis of trust and confidence from their consumers.

Governments are also concerned with the adequacy of data security and protection of consumer privacy implemented by companies. Accordingly, governments in many countries are considering regulations that greatly restrict firms' access, use, and sharing of consumer data. One noteworthy privacy legislation is the European Union's General Data Protection Regulation (GDPR). With the goal of creating more consistent protection of consumer personal data across all EU nations, the GDPR went into effect on May 25, 2018, as the

[71]Source: `https://arstechnica.com/information-technology/2018/05/equifax-breach-exposed-millions-of-drivers-licenses-phone-numbers-emails/`, accessed on December 1, 2018.

[72]Source: `http://analytics-magazine.org/survey-finds-deep-consumer-anxiety-over-data-privacy-and-security/`, accessed on November 25, 2018.

[73]Source: `https://www.cmswire.com/information-management/how-facebooks-cambridge-analytica-scandal-impacted-the-intersection-of-privacy-and-regulation/`, accessed on November 21, 2018.

primary law regulating how companies protect EU citizen's personal data.[74] Under GDPR, organizations must obtain explicit consent from users in order to store users' personal data, and also have a legal obligation to inform users of the purpose of data collection and processing, as well as of the identities of third parties with whom the data will be shared.[75] Companies that fail to comply with the GDPR are subject to costly penalties of up to €20m, or 4% of a firm's global turnover of the previous year (whichever is greater). Furthermore, note that in addition to EU members, any company, regardless of its location, must comply with the regulation if it markets goods and services to EU residents (known as "extra-territoriality"). The impact of the GDPR thus exceeds the boundaries of EU and changes data-protection requirements globally.

The GDPR is just the beginning - recent high-profile data breaches have further triggered calls for more urgent and strict data-protection measures worldwide. For example, modeled after the GDPR, the California Consumer Privacy Act of 2018 (CCPA) was recently passed (June 2018) and will become effective in 2020. Much like the GDPR, the CCPA provides consumers more control over their personal information by requiring California-based organizations to obtain explicit consent from users before sharing or selling consumer data to third parties. India is also one step closer to having its own data-protection law. In July 2018, the Indian government published the draft of the Personal Data Protection Bill, which proposes a comprehensive data-protection framework and is similar to the GDPR in terms of extra-territoriality and global-turnover-based penalties.

[74] According to GDPR directive, "personal data" are defined as "any information relating to an identifiable person who can be directly or indirectly identified by reference to an identifier. This definition provides for a wide range of personal identifiers to constitute personal data, including name, identification number, location data or online identifier, reflecting changes in technology and the way organizations collect information about people."

[75] Source: `https://eugdpr.org/the-regulation/gdpr-faqs/`, accessed on November 21, 2018.

While offering more rights and protection to consumers, such strict regulations will inevitably limit firms' ability to tailor their marketing activities and services to each individual consumer. Not only will these regulations negatively affect the profitability of firms that rely heavily on individual consumer data for prediction and targeting, but their impact on consumer welfare is also ambiguous. Note that firms' targeting activities often provide additional value to consumers, for instance, through lower search costs or through a better match with a product (e.g., Yao and Mela (2011), Anderson and Simester (2013)). Consequently, it is unclear whether consumers will eventually be better off if firms stop exploring consumer data under these new privacy policies. Therefore, it is imperative to find solutions that can alleviate the potential negative side effects of restrictive privacy regulations, while preserving data security.

In this paper, we show how machine learning approaches can achieve such objectives by enabling firms to continue benefiting from the abundance of consumer data without the need to store or access the data, hence mitigating the privacy concerns. In particular, we demonstrate how firms may achieve accurate targeting without centralized storage or access to the data, by building a Gated Recurrent Unit (Cho et al. (2014)) recurrent neural network (RNN) under the Federated Learning algorithm (McMahan et al. (2017)). The Gated Recurrent Unit (GRU, henceforth) recurrent neural network algorithm can achieve a highly accurate prediction about a consumer's next action conditional on what she has done or experienced in the past (e.g., a firm can predict which movie a consumer is more likely to watch based on her watch history and which movies are recommended to her). The Federated Learning (FL, henceforth) algorithm stores the private data locally on each user's device, while the model parameters are also updated locally on that device using those data. During the training, the firm does not need to access the private data directly, thereby keeping them safe. Only those locally updated parameters from consumers' devices are communicated to the central server (firm). Upon receiving those updated parameters from consumers, the firm aggregates them to update a "shared" model.[76]

The FL approach has a distinct advantage over other methods devised to protect privacy. Even if mostly anonymized, datasets that are stored and accessible at the firm's data center can still put consumer privacy at risk (Sweeney (2000)). For instance, consider the Differential Privacy algorithm (Dwork et al. (2006)) that Apple has deployed since 2016 as a key feature to protect consumer identity. When Apple collects and stores user data, it adds statistical noise to a user's profile and activities to mask the user's identity. A study by Tang et al. (2017) finds, however, that Apple's privacy-breach risk still exceeds the level that the research community typically considers acceptable. By contrast, the FL trains the model on each consumer's device locally, and therefore greatly reduces such risks because the firm never transfers, accesses, or stores consumers' personal data. The only information that is transmitted between the firm and consumers is the locally updated parameters that are necessary to improve the shared model.

Another attractive property of FL, which also distinguishes it from other distributed learning algorithms, is that it is robust to non-IID and highly unbalanced datasets. The data stored on any given consumer's device are almost certainly not representative of the population distribution, and the amount of data stored will vary substantially based on the consumer's usage of the device and the firm's service. While much of the previous research on distributed learning does not consider unbalanced and non-IID datasets, the FL approach works relatively well on these types of data by repeatedly averaging locally updated parameters.

[76] In the machine learning literature, "parameters" and "weights" are often used interchangeably. We use "parameters" to distinguish our meaning from "weights" used in weighted-averaging calculations, which appear later in the paper.

Furthermore, the FL is communication efficient. One major constraint in the design of large-scale distributed learning algorithms is the communication cost. In a typical distributed learning setting where the data are stored in a decentralized manner over a cluster of devices (nodes), communication costs are considerable. The development of an efficient distributed learning algorithm that can minimize the number of communication iterations among nodes is therefore an important issue. In the FL setting, the network and power connection of a consumer's device make communication costs the principal constraint. McMahan et al. (2017) demonstrate how two components of the FL approach can substantially reduce the number of communication rounds necessary for achieving a target accuracy level. The two components are (1) increasing parallelism, so that more consumers do computation independently during each communication round, and (2) increasing computation on each consumer's device, so that multiple updates are performed at the consumer level during each communication round.

To demonstrate the applicability of the proposed approach in a general marketing setting, we train the GRU with the FL algorithm using a highly unbalanced and non-IID consumer browsing dataset at an online retailer, with the objective to predict a consumer's click-stream. To establish a benchmark, we also train the GRU using a standard centralized learning approach. In contrast to the FL algorithm, the centralized learning requires the firm to store, access, and train all consumers' data collectively at a data center. We show the prediction accuracy of the proposed approach is comparable to that of the centralized learning method. Consequently, this approach allows firms to target consumers with high accuracy without compromising the security of personal data.

The rest of this paper is structured as follows: In the following section, we briefly discuss related literature. Section 3.3 gives a brief overview of the FL algorithm, as well as the GRU. In Section 3.4, we apply the FL algorithm with the GRU to a practical marketing problem, training a model to predict each consumer's next-clicked item using an online retailer's click-stream data. Section 3.5 concludes.

3.2 Related Literature

This paper adds to a stream of literature on consumer privacy.Han et al. (2003) study the trade-off that consumers face between the benefits and costs of providing personal information. They find the benefits such as monetary rewards and future convenience significantly affect consumers' preferences over websites with various privacy policies. They also quantify individuals' valuation of protection of personal information, and find it is worth between \$30.49 and \$44.62. Tucker (2014) shows that increasing users' perception of more control over their private information increases the effectiveness of behavioral targeting. Leveraging the implementation of European Union's opt-in tracking policy as a natural experiment, Goldfarb and Tucker (2011) demonstrate that display advertising becomes far less effective (65% reduction in effectiveness on average) in terms of stated purchase intent as a result of the privacy regulation. In the context of the online display ad industry, Johnson (2013) finds that reduced targeting due to stricter privacy policies decreases advertiser surplus, and that publishers' revenues also decrease as a result. More recently, Rafieian and Yoganarasimhan (2018) use machine learning techniques to quantify the value of targeting information, specifically, the relative importance of contextual information (based on the content of the website and hence privacy preserving) versus behavioral information (based on user-tracking and thereby jeopardizing privacy). They find that targeting consumers based on behavioral information is more effective than targeting based on contextual information, and that strict privacy regulations that ban user-tracking substantially reduce the value of behavioral targeting. For a more comprehensive review and discussion on big data and consumer privacy, see Jin (2018).

This paper also relates to literature on privacy-preserving machine learning (Barni et al. (2011), Xie et al. (2014), Rubinstein et al. (2012), Sarwate and Chaudhuri (2013), Duchi et al. (2012), Mohassel and Zhang (2017)). Privacy-preserving deep learning has been an active research area in recent years. The most relevant study in this domain is Shokri and Shmatikov (2015), who propose a method based on Differential Privacy (DP) for collaborative deep learning, where each party asynchronously trains a neural network locally and selectively shares only a subset of parameters with other parties. They do not, however, take into account the non-IID and unbalanced properties of the data. McMahan et al (2017) advance this literature by developing the FL algorithm that is robust to unbalanced and non-IID data distributions that are the defining characteristics of data stored in each consumer's device. This distributed learning technique offers the firm as well as consumers the benefits of the shared model trained from rich data, without having to compromise the security of personal data. In our paper, we further combine the FL approach with the GRU approach. We use the model to predict consumer click-streams and demonstrate its accuracy and applicability in marketing.

This paper also belongs to the literature that explores path-tracking and click-stream data to study consumers' decision-making along the purchase funnel (e.g., Moe and Fadder (2004), Montgomery et al. (2004), Park and Fader (2004), Hui et al. (2009)). Unlike typical brick-and-mortar data, which only record consumers' final transactional events, path-tracking and click-stream data can accurately capture the entire shopping path of a consumer in a complete and timely manner. As shown in recent studies, insights obtained from such data can provide a better understanding of consumers' search behavior and market competition, as well as enable managers to optimize their marketing efforts (e.g., Bronnenberg et al. (2016), Chen and Yao (2017), Seiler and Yao (2017), Yao et al. (2017)). Tracking and storing path and click-stream information, however, also intensifies privacy concerns. Even after the data are anonymized, the empirical patterns embedded in the data can reveal a substantial amount of personal information (Valentino-DeVries et al (2018)). Our paper demonstrates the possibility of analyzing path-tracking and click-stream data without jeopardizing consumers' privacy.

3.3 Model

In this section, we provide a brief description of the FL as well as the GRU algorithms. We first describe the FL's process of model distribution and aggregation executed by the central server, and then proceed to describe the GRU algorithm that trains the model locally at each individual consumer's device using personal data.

3.3.1 Server

The data are partitioned over K consumers, with n_k number of observations for consumer k, $k = 1, ..., K$. Let $\mathcal{P}_k = \{1, ..., i, ..., n_k\}$ be the set of indices for consumer k's data points; that is, $n_k = |\mathcal{P}_k|$. At round τ of communication between consumer devices and the central server, a fraction $C \in (0, 1]$ of all consumers are randomly selected to form a set S_τ (i.e., only a fraction C of consumers are selected during each communication round for computational efficiency). The model parameters of the current round, Θ_τ, are distributed from the central server to all consumers who have been selected to be included in this set. Next, each consumer k's device computes the average gradient g_k on her local data at the current parameters Θ_τ. The average gradient g_k can be written as $g_k(\Theta_\tau) = \nabla L_k(\Theta_\tau)$, where $L_k(\Theta_\tau) = \frac{1}{n_k}\sum_{i\in\mathcal{P}_k} l_i(\Theta_\tau)$, and $l_i(\Theta_\tau)$ is the loss function of the prediction on observation i.

The parameters are locally updated as

$$\Theta^k_{\tau+1} \leftarrow \Theta_\tau - \eta g_k, \tag{3.1}$$

where η is a learning rate. In other words, in parallel, each consumer locally takes one step of gradient descent at the current parameters using her local data. The resulting parameters, $\Theta_{\tau+1}^{k}$, $\forall k \in S_{\tau}$, are sent to the central server. The central server then takes a weighted average of parameters received from the consumers and updates the shared central model $\Theta_{\tau+1} \leftarrow \sum_{k \in S_{\tau}} \frac{n_k}{n_{S_{\tau}}} \Theta_{\tau+1}^{k}$, where $n_{S_{\tau}}$ is the total number of observations across all consumers in S_{τ}. This process repeats until convergence.

Sometimes, increasing local training epochs may further improve the communication efficiency (i.e., reduce the number of communication rounds necessary for convergence).[77] Specifically, during round τ of communication, instead of updating the local parameters only once at each consumer's device, it is possible to modify the procedure by increasing the number of local training epochs to $E > 1$ times before communicating to the central server. Let e be the index of local training epochs. Then consumer k's parameters at round τ are updated as

$$\Theta_{\tau+1}^{k,e+1} \leftarrow \Theta_{\tau+1}^{k,e} - \eta g_k(\Theta_{\tau+1}^{k,e})$$

$$e = 1, 2, ..., E$$

$$\text{with } \Theta_{\tau+1}^{k,1} = \Theta_{\tau} \text{ and } \Theta_{\tau+1}^{k} = \Theta_{\tau+1}^{k,E+1}.$$

The improvement in communication efficiency through this additional step, however, is not guaranteed. The improvement in efficiency may depend on characteristics of data that are stored on each consumer's device (e.g., sparsity). As we show in our application in Section 3.4 (as well as shown in McMahan et al (2017)), the additional local training epochs may not necessarily enhance the speed of neural network convergence. Accordingly, in practice,

[77]In the machine learning literature, an "epoch" is defined as one round of passing *all* data forward and backward through the network. Because the training happens on each individual consumer's device and all her data are passed through the local neural network, each training iteration can be viewed as one local epoch.

firms need to fine-tune the number of local epochs to achieve a high level of communication efficiency.

3.3.2 Consumer k

At each consumer's node, we employ the GRU to predict each consumer's next-clicked item during a browsing session. The GRU solves the vanishing gradient problem of the vanilla RNN using an "update gate" vector and a "reset gate" vector. These two gates determine how much information from a consumer's previous clicks needs to be passed along to make predictions about future clicks. They can be trained to retain information from multiple steps back or to ignore the information that is irrelevant for the prediction. For notational simplicity, we omit the indices k and τ that index a specific consumer and a communication round, respectively.

During a specific browsing session, a consumer makes $T \geq 2$ clicks. Suppose J alternative products are available at each session. At step t $(t = 1, 2, ..., T)$ of the browsing session, the consumer can choose one product to click. Let matrix $X = [x_1, x_2, ..., x_T]$ be the sequence of vectors representing the consumer's click-stream in a given browsing session. $x_t \in \mathbb{R}^{J \times 1}$ is a J-dimensional vector whose j-th element equals 1 if a consumer clicks on product j at step t, and 0 otherwise.

Given the sequence $[x_{1,}x_2..., x_t]$ up to step t, $t = 1, 2, ..., T-1$, our objective is to predict x_{t+1}, the click vector at step $t+1$. At each t, $t = 1, 2, ..., T-1$, the hidden state of the previous step, $h_{t-1} \in \mathbb{R}^{D \times 1}$,[78] and the input x_t are passed to the gated recurrent unit.[79] The gated recurrent unit in turn updates the current hidden state h_t $(t = 1, 2, ..., T-1)$ using the

[78] $D \times 1$ is the dimension of the hidden state vector.
[79] Note h_0 is a vector with all elements equal 0.

following architecture

$$z_t = \sigma([W_z x_t + U_z h_{t-1} + b_z) \quad (3.2)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (3.3)$$

$$\hat{h}_t = tanh(W_h x_t + U_h(h_{t-1} \odot r_t) + b_h) \quad (3.4)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \hat{h}_t, \quad (3.5)$$

where z_t and r_t are update and reset gates, respectively; $\hat{h}_t$ and h_t are the current memory and the hidden state, respectively; $\sigma(\cdot)$ is the sigmoid function; $tanh(\cdot)$ is a hyperbolic tangent function; and $\odot$ denotes an element-wise multiplication. $W_z, W_r, W_h, U_z, U_r, U_h$ are the matrices, and b_z, b_r, b_h are the vectors of parameters to be learned. The intuition of the GRU is as follows:

Update gate (equation 3.2): The update gate z_t allows the model to control how much of the information from previous steps (which is summarized in h_{t-1}) should be carried forward to the current hidden state h_t. The update gate helps the model remember long-term information.

Reset gate (equation 3.3): Despite their identical formula, the reset gate r_t is different from the update gate z_t. The difference comes from the parameter matrices and vectors, and more importantly, the gate's usage. The reset gate r_t allows the model to drop any previous information that is irrelevant for future predictions.

Current memory (equation 3.4): The current memory $\hat{h}_t$ consolidates the new input x_t (the click vector in step t) with the previous hidden state h_{t-1}. The latter holds information from the consumer's click activities in previous $t - 1$ steps.

Hidden state (equation 3.5): The hidden state h_t uses the update gate as the weight to store relevant information from the previous hidden state h_{t-1} and the current memory $\hat{h_t}$.

The hidden state h_t is then used to calculate the prediction of the click vector of step $t+1$, $\hat{x}_{t+1}$. The prediction $\hat{x}_{t+1}$ takes the form of a J-dimensional vector, whose j-th element is the probability of the consumer clicking product j. Specifically,

$$\hat{x}_{t+1} = \left[\frac{exp(o_{t,1})}{\sum_{j=1}^{J} exp(o_{t,j})}, \cdots, \frac{exp(o_{t,J})}{\sum_{j=1}^{J} exp(o_{t,j})} \right]' \tag{3.6}$$

$$o_t = [o_{t,1}, o_{t,2}, ..., o_{t,J}]' \tag{3.7}$$

$$= Vh_t + b_v, \tag{3.8}$$

where V and b_v are another set of matrix and vector of parameters to be learned.

Finally, we use the cross-entropy error as the loss function, which is defined as

$$L = \frac{1}{T-1} \sum_{t=1}^{T-1} x_{t+1} \cdot log(\hat{x}_{t+1}). \tag{3.9}$$

The full set of model parameters to be learned are

$$\Theta_\tau = \{W_u, U_u, b_u, W_r, U_r, b_r, W_h, U_h, b_h, V, b_v\}. \tag{3.10}$$

3.4 Application: Click-stream Prediction

We apply the FL algorithm to a click-stream dataset from an online retailer, and train the GRU locally at each consumer's node using only that consumer's personal data. Our goal is

Table 3.1: Summary Statistics of Training Dataset

	Training Set				
	Mean	SD	Med	Min	Max
Number of sessions per customer	1.89	2.55	1	1	52
Number of clicks per customer	28.43	137.89	9	2	7,332
Number of clicks per session	15.03	49.68	5	2	1,844
Number of unique products clicked per customer	6.34	9.85	4	1	217
Number of unique products clicked per session	4.40	4.55	3	1	88
Number of customers	3,632				
Number of sessions	6,873				
Number of clicks	103,270				

to show how the FL algorithm can fit into a broad marketing framework. In particular, we combine the FL with the GRU to test the performance of the prediction of each consumer's click-stream within a browsing session. As discussed in Hidasi et al (2016), the prediction of the next-clicked product or a set of products in a customer's click-stream often become the basis for a website's recommendation system. A well-calibrated recommendation system in turn may enhance the conversion rate of the online retailer. Consequently, accurately predicting a consumer's click-stream within a browsing session has substantial managerial implications. To evaluate the accuracy of the prediction, we focus on the predicted probability on the next-clicked product. In particular, we use "Recall@K" averaged over all clicks of all consumers as our evaluation metric of prediction accuracy. Recall@K is widely used in the machine learning literature for predicting click-through rates (Hidasi and Tikk (2016)). For our application, the consumer clicks on only *one* product at each step. In this case, Recall@K is a dummy variable. More specifically, for a given prediction at step t, Recall@K equals 1 if the list of K products with the highest predicted click probabilities includes the product that the consumer actually clicks. Recall@K equals 0 if the actually clicked product does not appear in the K-product list.

Table 3.2: Summary Statistics of Test Dataset

	Test Set				
	Mean	SD	Med	Min	Max
Number of sessions per customer	1.85	2.12	1	1	28
Number of clicks per customer	26.93	73.44	9	2	1,212
Number of clicks per session	14.58	42.35	5	2	1,008
Number of unique products clicked per customer	6.47	8.91	4	1	125
Number of unique products clicked per session	4.39	4.23	3	1	43
Number of customers			908		
Number of sessions			1,677		
Number of clicks			24,454		

We use a dataset from a large Chinese online liquor retailer, which contains a set of 5,711 randomly selected customers shopping in the wine category on the website during July 2016. For each customer, we observe her individual-level click-stream at the website. During the observation window, these 5,711 customers initiate 13,154 browsing sessions on the website.[80] During these sessions, they make 132,328 clicks on 1,660 products.

On average, each product appears in approximately 275 sessions, but with a large variance. Some unpopular products only appear once in browsing sessions across customers, while the most popular product appears in 1,177 sessions. We aggregate unpopular products that appear in less than five sessions into one composite good. There are 798 such products, and they constitute only 2.04% of total clicks in the data. We also drop sessions in which a customer makes only one click, because our objective is to predict the next-clicked item during the browsing session. As a result, we lose 4,604 observations (clicks) after dropping those sessions.

Our final sample consists of 127,724 clicks, 4,540 customers, 8,550 sessions, and 863 products (862 products and one composite good). On average, we have 1.89 sessions per customer,

[80] A session ends when the customer closes the website's browser window/tab.

each session consists of 15 clicks, each customer has 28 clicks, and each customer clicks on 4.1 unique products per session and 6.37 unique products in total.

We randomly select 80% of unique customers for training and use the remaining 20% as a test set to calibrate the out-of-sample prediction accuracy. Our training dataset contains 103,270 clicks, 3,632 customers, and 6,873 sessions. Our testing data consist of 24,454 clicks, 908 customers, and 1,677 sessions. Note that due to our random assignment of consumers into training/test sets, the total number of unique products clicked vary across the two groups of consumers, even though they face the same set of alternative products. Summary statistics of the training set and the test set are reported in Table 3.1 and Table 3.2, respectively.

Each shopping session of a consumer forms a separate sequence. That is, if any of the consumer's sessions ends, we reset the appropriate hidden state. We fix the size of the hidden states to 100 and let each session of a consumer constitute a minibatch (6,783 sessions/minibatches in total in the training set). The full model has 376,363 parameters to learn. For optimization of the loss function, we use the Adam algorithm with squared-root decay of learning rates.[81] To establish a benchmark, we also train and test the GRU using the centralized learning approach, that is, standard stochastic gradient descent on the full training set, where we use the same train/test split as in the FL setting, again with each session forming a minibatch. For computational efficiency, we choose $C = 0.2$; that is, 20% of randomly selected consumers work independently during each communication round. We also show the results obtained from setting $C = 0.1$ for comparison. We also vary the level of E, the number of local training epochs on each consumer's device using her local data before communicating to the central server.

[81]For centralized learning, we set the learning rate η to $3 \times 1e - 4$. For FL, η is set to 5 when the sampling rate $C = 0.2$ and the number of local training epoch $E = 1$; η is set to 3.3 when $C = 0.2$ and $E = 2$; η is set to 1.2 when $C = 0.1$ and $E = 1$. We trained over a wide range of learning rates, and these performed the best in terms of speed of convergence.

Table 3.3: Prediction Accuracy and Communication Rounds

Model	C	E	Recall@1	Communication rounds
Centralized Learning-GRU	-	-	0.60	-
Federated Learning-GRU	0.1	1	0.43	8,287
Federated Learning-GRU	**0.2**	**1**	**0.53**	**555**
Federated Learning-GRU	0.2	2	0.52	620

Table 3.3 reports the out-of-sample prediction accuracy as measured by Recall@1 averaged over all predicted clicks. We present the prediction accuracy levels for FL-GRU with various sampling rate C and local training epoch E. We also report the prediction accuracy obtained through the centralized learning approach as a baseline. When the sampling rate $C = 0.2$ and local training epoch $E = 1$, the FL achieves a prediction accuracy of 53%. That is, when we train the GRU using the FL approach, with 53% probability, the product with the highest predicted click probability is the actual product the consumer has clicked (out of 863 alternative products). The prediction accuracy obtained via the FL approach is comparable to that of the centralized approach, with the FL approach achieving 88% of the baseline prediction accuracy of the centralized approach.[82] We want to emphasize that with the FL approach, the central server/firm has never stored, accessed, or directly analyzed individual consumer data. Hence, the accuracy level of the FL is fairly impressive.

Computational costs are minimal in the FL setting because the size of the dataset stored in any single device is small while modern devices have fast processors. By contrast, communication costs are of major concern in distributed optimization settings such as the FL, because information needs to be passed back and forth between the nodes and the central server during the model optimization. In particular, limited upload bandwidth, network connection (3G, 4G, WiFi), and power plug-ins (battery) hinder unlimited communication. McMahan et al (2017) show the following two elements of the FL may substantially reduce the number of communication rounds necessary for convergence:

[82] i.e., $0.53/0.60 = 0.88$.

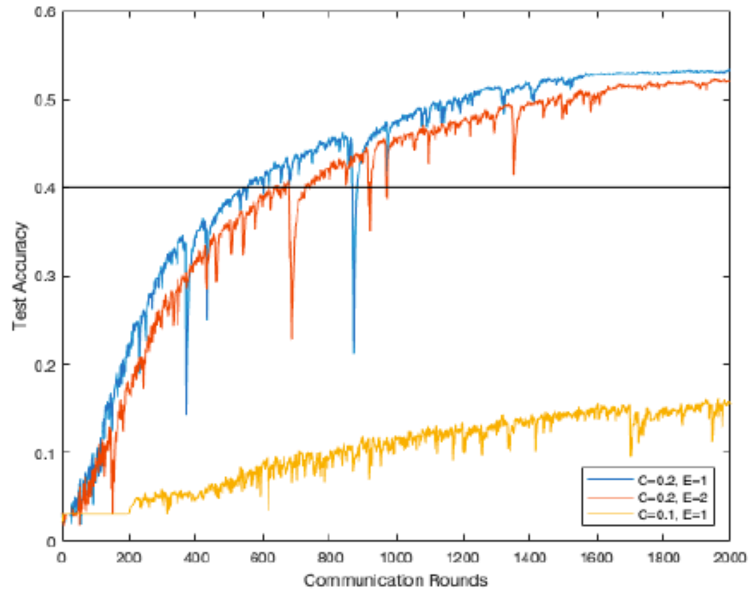


Figure 3.1: Test Accuracy for (i) $C = 0.1, E = 1$, (ii) $C = 0.2, E = 1$, and (iii) $C = 0.2, E = 2$. Plot for $C = 0.1, E = 1$ is only shown up to 2,000 communication rounds in order to compare the communication efficiency with the baseline of $C = 0.2$

1. Increasing parallelism by increasing sampling rate: More consumers do computation independently during each communication round, and

2. Increasing computation at each consumer's node: Multiple updates are performed at the consumer level during each communication round.

We report in Table 3.3 the minimum number of communication rounds necessary to achieve a target accuracy of 40%. Figure 3.1 shows the learning curves, where the horizontal line represents the target 40% accuracy level. The target accuracy is reached after $8,287$ communication rounds when $C = 0.1, E = 1$. Increasing parallelism by setting C to 0.2 while maintaining $E = 1$ drastically reduces the number of rounds to only 555.[83] This figure also includes results for $C = 0.2, E = 2$, which performs slightly worse than $E = 1$.[84]

3.5 Conclusion

Massive amounts of data generated by consumers provide a wealth of opportunities for firms to accurately predict consumer behavior and to target and provide customized services, thereby improving profitability as well as enhancing consumer experience. However, the rapid growth of the use of consumer data, along with recent data-breach incidents, has raised concerns regarding the protection of consumers' privacy. Governments in several countries are introducing regulations that greatly restrict firms' access, use, and sharing of consumer data. These regulations greatly restrict business activities of firms that rely heavily on consumer data for their business activities. Therefore, firms must find solutions to mitigate the impact of restrictive privacy regulations while keeping consumers' private data safe.

In this paper, we show how machine learning approaches allow firms to continue benefiting from vast amounts of consumer data without compromising consumers' privacy. Specifically, we discuss a recently developed FL approach, which uses a parallelized deep learning algorithm to train a model locally on each individual consumer's device. As an instantiation to demonstrate the applicability of this approach in a marketing setting, we build a session-based GRU recurrent neural network that predicts each consumer's click-stream under the FL framework. We show the prediction accuracy of the trained neural network via the FL approach is

[83] For the centralized learning, the GRU is trained on the full training set. The model parameters are updated iteratively and sequentially for each minibatch (i.e., simple stochastic gradient descent). For the centralized learning to achieve the 40% accuracy, 440 training epochs are necessary. One interesting analogy about communication is that if each minibatch update is counted as a communication round, the total number of communication rounds is $440 \times 6,873 = 3,024,120$. This number is much higher than the 555 rounds needed for the FL approach, implying a substantial computational burden.

[84] As discussed in section 3.3, increasing local training epochs may not necessarily enhance communication efficiency. In et al (2017), the authors draw the same conclusion.

comparable to that of the benchmark centralized approach. Through this application, we demonstrate how firms can continue targeting consumers with a high level of accuracy without having to store, access, or analyze consumer data in centralized locations, thereby preserving consumers' sensitive information.

3.6 References

Anderson, E. T. and D. Simester (2013): "Advertising in a Competitive Market: The Role of Product Standards, Customer Learning, and Switching Costs," *Journal of Marketing Research*, 50.

Barni, M., P. Failla, R. Lazzeretti, A.-R. Sadeghi, and T. Schneider (2011):

"Privacy-Preserving ECG Classification With Branching Programs and Neural Networks," *IEEE Transactions on Information Forensics and Security*, 6, 452–468.

Bronnenberg, B. J., J. B. Kim, and C. F. Mela (2016): "Zooming In on Choice: How Do Consumers Search for Cameras Online?" *Marketing Science*, 35, 693–829.

Chen, Y. and S. Yao (2017): "Sequential Search with Refinement: Model and Application with Click-Stream Data," *Management Science*, 63, 4345–4365.

Cho, K., D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio (2014): "Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation," *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1724–1734.

Duchi, J. C., M. I. Jordan, and M. J. Wainwright (2012): "Privacy Aware Learning," *arXiv:1210.2085*.

Dwork, C., F. McSherry, K. Nissim, and A. Smith (2006): "Calibrating Noise to Sensitivity in Private Data Analysis," *In Theory of Cryptography Conference (TCC)*, 265284.

Goldfarb, A. and C. E. Tucker (2011): "Privacy Regulation and Online Advertising," *Management Science*, 57, 57–71.

Hann, I.-H., K.-L. Hui, T. S. Lee, and I. Png (2003): "The Value of Online Information Privacy: An Empirical Investigation," Unpublished Manuscript.

Hidasi, B., A. Karatzoglou, L. Baltrunas, and D. Tikk (2016): "Session-Based Recommendations with Recurrent Neural Networks," Published as a conference paper at ICLR.

Hidasi, B. and D. Tikk (2016): "General Factorization Framework for Context-Aware Recommendations," *Data Mining and Knowledge Discovery*, 30, 342–371.

Hui, S. K., P. S. Fader, and E. T. Bradlow (2009): "Path Data in Marketing: An Integrative Framework and Prospectus for Model Building," *Marketing Science,* 28, 320– 335.

Jin, G. Z. (2018): "Artificial Intelligence and Consumer Privacy," *NBER WORKING PAPER SERIES.*

Johnson, G. (2013): "The Impact of Privacy Policy on the Auction Market for Online Display Advertising," Working Paper.

McMahan, H., E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas (2017): "Communication-Efficient Learning on Deep Networks from Decentralized Data," *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 54.

Moe, W. W. and P. S. Fader (2004): "Dynamic Conversion Behavior at E-Commerce Sites," *Management Science*, 50, 326–335.

Mohassel, P. and Y. Zhang (2017): "SecureML: A System for Scalable PrivacyPreserving Mache Learning," Working Paper.

Montgomery, A. L., S. Li, K. Srinivasan, and J. C. Liechty (2004): "Modeling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science*, 23, 579–595.

Park, Y.-H. and P. S. Fader (2004): "Modeling Browsing Behavior at Multiple Websites," *Marketing Science*, 23, 280–303.

Rafieian, O. and H. Yoganarasimhan (2018): "Targeting and Privacy in Mobile Advertising," Working Paper.

Rubinstein, B. I. P., P. L. Bartlett, L. Huang, and N. Taft (2012): "Learning in a Large Function Space: Privacy-Preserving Mechanisms for SVM Learning," *J. Privacy and Confidentiality*, 4.

Sarwate, A. D. and K. Chaudhuri (2013): "Signal Processing and Machine Learning with Differential Privacy: Algorithms and Challenges for Continuous Data," *Signal Processing Magazine*, 30, 86–94.

Seiler, S. and S. Yao (2017): "The Impact of Advertising Along the Conversion Funnel," *Quantitative Marketing and Economics*, 15.

Shokri, R. and V. Shmatikov (2015): "Privacy-Preserving Deep Learning," *Proceedings of the 22nd ACM SIGSAC Conferences on Computer and Communications Security, CCS' 15.*

Sweeney, L. (2000): "Simple Demographics Often Identify People Uniquely," *Carnegie Mellon University Data Privacy Working Paper*.

Tang, J., A. Korolova, X. Bai, X. Wang, and X. Wang (2017): "Privacy Loss in Apple's Implementation of Dierential Privacy on MacOS 10.12," available at https://arxiv.org/abs/1709.02753.

Tucker, C. E. (2014): "Social Networks, Personalized Advertising, and Privacy Controls," *Journal of Marketing Research*, 51, 546–562.

Valentino-DeVries, J., N. Singer, M. H. Keller, and A. Krolik (2018): "Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret," *New York Times*, Dec. 10.

Xie, P., M. Bilenko, T. Finley, R. Gilad-Bachrach, K. Lauter, and M. Naehrig (2014): "Crypto-Nets: Neural Networks over Encrypted Data," *arXiv:1412.6181*.

Yao, S. and C. F. Mela (2011): "A Dynamic Model of Sponsored Search Advertising," *Marketing Science*, 30, 447–468.

Yao, S., W. Wang, and Y. Chen (2017): "TV Channel Search and Commercial Breaks," *Journal of Marketing Research*, 54, 671–686.

Appendix A

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 essential to take into account differences in patient case-mix across hospitals, especially since we are using patients undergoing various types of different surgeries. Specifically, we include the following hospital-level case-mix as control variables: Above70 (fraction of patients older than 70 years of age), SurgeryRisk (average deathrate of all surgeries conducted in each hospital, where deathrate of a surgery is calculate as the death rate of each surgery over our entire sample), DiseaseRisk (average deathrate of patients' diagnosed disease, where deathrate of a disease is calculate as the death rate of each disease over our entire sample), ComorbidityRisk (average deathrate of patients' diagnosed comorbidity, where deathrate of a comorbidity is calculate as the death rate of each comorbidity over our entire sample. If a patient does not have a comorbidity, this variable equals 0), DisabledFrac (fraction of patients with a kidney and other dysfunction), DisabilitySeverity (severity of disability, 1 mild, 2 severe), LowIncomeFrac (fraction of patients with low income).[85]

[85] In the data there are various categories of disabilities, such as intelectual disorder, mental disorder, hearing disability, etc. Since some of these disabilities are not likely to affect the mortality of a patient, we only consider Kidney Dysfunction and "Other Dysfuction". Other dysfunction includes (but does not distinguish between) speech disability, austistic disorder, cardiac dysfunction, respiratory dysfunction, liver dysfunction, facial disfigurement, intestinal fistular/urinary fistular.

Table A.1 reports the Diff-in-Diff estimates of the impact of competition on raw mortality rates controlling for hospital-level case mix. As expected, hospitals with more riskier diseases, riskier comorbidities, and more severe disabilities have higher mortality rates. Income and Age do not seem to affect hospital level mortality rates. After controlling for DisabilitySeverity, the coefficient on Disabled becomes negative. In Table A.2 we check the robustness of our results using only a subset of the control variables. Our results remain unchanged.

Although speech disability and autism may be unrelated to deathrate, we are not able to distinguish these disabilities from more critical ones such as cardiac and liver dysfunction.

Table A.1: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care

	Pre-Post	5-mile Treatment	10-mile Treatment	15-mile Treatment	20-mile Treatment
Raw Mortality Rates	(1)	(2)	(3)	(4)	(4)
Post	-0.002	0.001	0.004	0.007	0.008
	(0.003)	(0.003)	(0.004)	(0.005)	(0.006)
Treated×Post		**-0.008***	**-0.011****	**-0.014****	**-0.014****
		(0.005)	(0.005)	(0.006)	(0.006)
Above70	0.001	0.004	0.007	0.006	-0.000
	(0.048)	(0.047)	(0.047)	(0.046)	(0.047)
SurgeryRisk	-0.039	-0.026	-0.024	-0.027	-0.029
	(0.112)	(0.110)	(0.108)	(0.108)	(0.108)
DiseaseRisk	0.697***	0.703***	0.702***	0.725***	0.721***
	(0.234)	(0.229)	(0.228)	(0.224)	(0.226)
ComorbidityRisk	1.113***	1.122***	1.126***	1.136***	1.128***
	(0.392)	(0.388)	(0.379)	(0.373)	(0.376)
DisabledFrac	-1.939***	-2.187***	-2.280***	-2.274***	-2.106***
	(0.707)	(0.766)	(0.718)	(0.710)	(0.691)
DisabilitySeverity	2.310***	2.562***	2.660***	2.654***	2.471***
	(0.820)	(0.879)	(0.835)	(0.825)	(0.801)
LowIncomeFrac	-0.065	-0.071	-0.069	-0.069	-0.057
	(0.141)	(0.138)	(0.140)	(0.138)	(0.137)
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		105	69	55	43
Treated Hospitals		62	98	112	124
Observations	334	334	334	334	334
R-squared	0.8321	0.8351	0.8372	0.8398	0.8386

Notes: Models are estimated by OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. 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

height	Pre-Post	5-mile Treatment	10-mile Treatment	15-mile Treatment	20-mile Treatment
Raw Mortality Rates	(1)	(2)	(3)	(4)	(4)
Post	-0.003	0.000	0.004	0.006	0.008
	(0.002)	(0.003)	(0.004)	(0.005)	(0.006)
Treated×Post		**-0.008***	**-0.011****	**-0.014****	**-0.014****
		(0.005)	(0.005)	(0.006)	(0.006)
DiseaseRisk	0.662**	0.683***	0.687***	0.705***	0.693***
	(0.261)	(0.253)	(0.245)	(0.239)	(0.244)
ComorbidityRisk	1.0751***	1.092***	1.097***	1.106***	1.100***
	(0.349)	(0.346)	(0.334)	(0.330)	(0.333)
DisabledFrac	-2.019***	-2.259***	-2.341***	-2.342***	-2.177***
	(0.751)	(0.804)	(0.757)	(0.746)	(0.727)
DisabilitySeverity	2.393***	2.636***	2.723***	2.725***	2.545***
	(0.872)	(0.925)	(0.880)	(0.867)	(0.844)
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		105	69	55	43
Treated Hospitals		62	98	112	124
Observations	334	334	334	334	334
R-squared	0.8315	0.8344	0.8366	0.8391	0.8382

Notes: Models are estimated by OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants.

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

Table A.3: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care

	undadjusted raw mortality	unadjusted raw mortality	adjusted mortality OLS	adjusted mortality IV	adjusted mortality IV (rescaled)
	(1)	(2)	(3)	(4)	(5)
Post	-0.00078	0.0020	0.004	0.025	0.004
	(0.0031)	(0.0077)	(0.007)	(0.030)	(0.004)
Treated×Post		-0.0042	**-0.014***	**-0.082****	**-0.012****
		(0.0082)	(0.007)	(0.038)	(0.006)
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		55	55	55	55
Treated Hospitals		112	112	112	112
Observations	334	334	334	334	334
R-squared	0.6885	0.6892	0.6011	0.6181	0.6181

Notes: This table compares the DID estimates using raw mortality rates which do not control for hospital-level case mix. Models are estimated using OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants. Treatment is defined as being located within 15-miles of the HST station.

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

Table A.4: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care (unadjusted raw mortality)

unadjusted raw mortality	Pre-Post	5-mile Treatment	10-mile Treatment	15-mile Treatment	20-mile Treatment
	(1)	(2)	(3)	(4)	(4)
Post	-0.00078	-0.00044	0.001	0.0020	0.0048
	(0.0031)	(0.0043)	(0.006)	(0.0077)	(0.0096)
Treated×Post		-0.00091	-0.0031	-0.0042	-0.0075
		(0.0061)	(0.0069)	(0.0082)	(0.010)
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		55	55	55	55
Treated Hospitals		112	112	112	112
Observations	334	334	334	334	334
R-squared	0.6885	0.6885	0.6889	0.6892	0.6905

Notes: This table compares the DID estimates using raw mortality rates which do not control for hospital-level case mix. Models are estimated using OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants. Treatment is defined as being located within 15-miles of the HST station.

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

Appendix B

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 μ_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 Janurary of year $yyyy_i + 1$.

We then use this mortality dummy variable M to obtain the case-mix adjusted mortality rate by estimating the following linear probability model pooled across both pre- and post-HST periods:

$$M = C\boldsymbol{\psi} + H\boldsymbol{\gamma} + (S + \eta) \tag{B.1}$$

where M is a vector of dummy variable whose elements are switched on if a patient died. C is a matrix of hospital-time period dummy variables, and H and S are patients' observed and unobserved health status, respectively.[86] The estimated hospital fixed effects parameter, $\boldsymbol{\psi}$, is the case-mix adjusted mortality rate that will be used in our difference-in-differences estimation as well as in our structural model of hospital choice. The corresponding expression for an individual observation is as follows:

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

Following Gaynor et al (2013) hospital dummies are stacked in a block-diagonal matrix where each block represents each period. Along with patients' observed case-mix, the data are arranged as

$$X = \begin{bmatrix} C_{pre} & & H_{pre} \\ & C_{post} & H_{post} \end{bmatrix}$$

where all elements in the matrix other than C_t and H_t are equal to zero. C_t is given by

$$C_t = \begin{bmatrix} c^t_{11} & \cdots & c^t_{1J-1} \\ \vdots & \ddots & \vdots \\ c^t_{n_t 1} & \cdots & c^t_{n_t J-1} \end{bmatrix}$$

where n_t is the number of patients in period t, and the elements c^t_{ij} takes value one if patient i chooses hospital j among J alternatives in period t, and zero otherwise.

Allowing the hospital fixed effects to vary for each period, we need to instrument $(2 \cdot J - 1)$ hospital choice dummies for each period, requiring us of at least as many number of instruments.

[86] In patient characteristics matrix H, we include female dummy, age, income group, riskiness of the patient's surgery, riskiness of the patient's disease, riskiness of the patient's comorbidity and a disability dummy variable.

We use travel time to each of the J hospitals and additional J set of a dummy variables which equals 1 if a given hospital is the closest one to the patient, which gives us a total of $2 \cdot J$ instruments for each period. Specifically, we define travel time for patient i to hospital j in period t as

$$\text{TravelTime}_{ijt} = \begin{cases} \min(\text{cartime}_{ij}, \text{traintime}_{ij}) & \text{if } i \text{ lives in treated region in } t = post \\ \text{cartime}_{ij} & \text{otherwise} \end{cases}$$

The matrix of instrumental variables is constructed as

$$Z = \begin{bmatrix} Z_{pre} & & H_{pre} \\ & Z_{post} & H_{post} \end{bmatrix}$$

where Z_t is a matrix of $2 \cdot J$ instruments which is given by

$$Z_t = \begin{bmatrix} z_{11}^t & \dots & z_{1K}^t \\ \vdots & \ddots & \vdots \\ z_{n_t 1}^t & \dots & z_{n_t K}^t \end{bmatrix}$$

and $K = 2 \cdot J$ denotes the number of instruments.

Formal specification tests for the validity of our instruments are provided in Table B.1. 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.

In Table B.2 we report the estimates of the effect of patients' observed case-mix on patient mortality from OLS and IV methods.

Sargan-Hansen	χ^2	274.0181
Overidentification Test	P-value	0.9936
Hausman	χ^2	3,068
Endogeneity Test	P-value	0.0001
Wald-Test of	χ^2	31,610
Weak Instruments	P-value	0.0001

Table B.1: Tests for Validity of Instruments

Table B.2: Estimates of the effect of patient characteristics on mortality from OLS and IV methods (standard errors in parentheses)

	OLS Coefficients	IV Coefficients
Female	0.003**	0.002
	(0.001)	(0.003)
MediumIncome	0.008***	0.009**
	(0.002)	(0.004)
HighIncome	0.007***	0.006
	(0.002)	(0.004)
Age[20-40)	-0.002***	-0.003
	(0.002)	(0.004)
Age[40-60)	-0.011	-0.010**
	(0.002)	(0.004)
Age[60-80)	0.0064***	0.008*
	(0.002)	(0.005)
Age[80+)	0.079***	0.072***
	(0.005)	(0.010)
MainsickRisk	0.518***	0.512***
	(0.010)	(0.020)
SubsickRisk	0.523***	0.537***
	(0.010)	(0.017)
SurgeryRisk	0.383***	0.382***
	(0.005)	(0.012)
Disabled	-0.022***	-0.032***
	(0.004)	(0.008)
DisabilitySevere	0.003***	0.006
	(0.012)	(0.016)

Notes: Models are estimated by OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants.

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

Appendix C

Table C.1: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care

OLS Estimated	Pre-Post	5-mile Treatment	10-mile Treatment	15-mile Treatment	20-mile Treatment
Adjusted Mortality Rates	(1)	(2)	(3)	(4)	(5)
Post	-0.005**	-0.002	0.001	0.004	0.007
	0.003	0.004	0.005	0.007	0.008
Treated×Post		**-0.001****	**-0.010***	**-0.014***	**-0.016***
		0.005	0.006	0.007	0.008
	(p: 0.045)	(p:0.047)	(p:0.076)	(p:0.052)	(p:0.056)
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		105	69	55	43
Treated Hospitals		62	98	112	124
Observations	334	334	334	334	334
R-squared	0.5868	0.5947	0.5961	0.6011	0.6040

Notes: This table shows diff-in-diff estimates for case-mix adjusted mortality rates obtained via OLS for various definitions of hospital treatment (hospitals located within 5 miles, 10 miles, 15 miles, and 20 miles). For various definitions of treatment, we can consistenly see that the diff-in-diff coefficient is negative and significant at the 10 percent level. Models are estimated by OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants. *** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

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

1*IV Estimated	Pre-Post	5-mile Treatment	10-mile Treatment	15-mile Treatment	20-mile Treatment
Adjusted Mortality Rates	(1)	(2)	(3)	(4)	(4)
Post	-.0296	-0.025	0.016	0.025	0.026
	(0.019)	(0.026)	(0.027)	(0.030)	(0.036)
Treated×Post		**-0.014**	**-0.078****	**-0.082****	**-0.075***
		(0.037)	(0.038)	(0.038)	(0.043)
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		105	69	55	43
Treated Hospitals		62	98	112	124
Observations	334	334	334	334	334
R-squared	0.6086	0.6088	0.6181	0.6181	0.6156

Notes: This table shows diff-in-diff estimates for case-mix adjusted mortality rates obtained via IV method for various definitions of hospital treatment (hospitals located within 5 miles, 10 miles, 15 miles, and 20 miles). For various definitions of treatment (except for 5 mile treatment), we can consistenly see that the diff-in-diff coefficient is negative and significant at either 5 percent or 10 percent level. Models are estimated by OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants. *** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

Table C.3: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care [Rescaled]

1*IV Estimated	Pre-Post	5-mile Treatment	10-mile Treatment	15-mile Treatment	20-mile Treatment
Adjusted Mortality Rates	(1)	(2)	(3)	(4)	(4)
Post	-0.004	-0.004	0.002	0.004	0.004
	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)
Treated×Post		**-0.002**	**-0.011****	**-0.012****	**-0.011***
		(0.005)	(0.006)	(0.006)	(0.006)
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		105	69	55	43
Treated Hospitals		62	98	112	124
Observations	334	334	334	334	334
R-squared	0.6086	0.6088	0.6181	0.6181	0.6156

Notes: This table shows diff-in-diff estimates for case-mix adjusted mortality rates obtained via IV method (rescaled for interpretation purpose) for various definitions of hospital treatment (hospitals located within 5 miles, 10 miles, 15 miles, and 20 miles). For various definitions of treatment (except for 5 mile treatment), we can consistently see that the diff-in-diff coefficient is negative and significant at either 5 percent or 10 percent level. Models are estimated by OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants. *** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

Appendix D

Table D.1: Proportion of Patients who Traveled to arrive at Hospitals

	Treated Hospitals			Control Hospitals		
	Pre-HST	Post-HST	t-stat	Pre-HST	Post-HST	t-stat
			Patient Treatment: 10 miles			
Control Patients	0.288	0.292	t:0.7130	0.073	0.076	t:0.8973
	(0.453)	(0.455)		(0.260)	(0.266)	
N	7,202	8,065	p:0.4758	7,865	9,552	p:0.3696
Treated Patients	0.132	0.151	**t:3.3867**	0.335	0.376	t:1.2550
	(0.338)	(0.358)		(0.473)	(0.485)	
N	7,202	8,065	p:0.0007	385	492	p:0.2098
			Patient Treatment: 15 miles			
Control Patients	0.478	0.476	t:0.2291	0.067	0.072	t:1.1415
	(0.500)	(0.500)		(0.251)	(0.258)	
N	6,075	7,496	p:0.8188	7,487	9,084	p:0.2537
Treated Patients	0.110	0.120	**t:2.0532**	0.254	0.263	t:0.3915
	(0.313)	(0.324)		(0.436)	(0.440)	
N	10,185	11,617	p:0.0401	713	880	p:0.2537

Notes: This table shows the changes in proportion of patients (excluding Seoul and surrounding area) who traveled more than 50 miles to arrive at the hospitals. There is a significant increase in proportion of treated patients traveling more than 50 miles to arrive at treated hospitas. Standard deviation in parentheses.

Table D.2: Changes in Hospital-Level (Raw) Mortality Rates

Control Hospitals				Treated Hospitals			
Hospital Treatment: 10 mile radius							
Pre-Train	Post-Train	Δ in Means	t-stat	Pre-Train	Post-Train	Δ in Means	t-stat
0.051	0.052	0.001	0.1342	0.041	0.039	-0.002	-0.4939
(0.039)	(0.048)			(0.031)	(0.027)		
[69]	[69]			[98]	[98]		
Hospital Treatment: 15 mile radius							
Pre-Train	Post-Train	Δ in Means	t-stat	Pre-Train	Post-Train	Δ in Means	t-stat
0.050	0.052	0.002	0.2319	0.043	0.040	-0.002	-0.5331
(0.041)	(0.050)			(0.031)	(0.029)		
[55]	[55]			[112]	[112]		

Notes: This table shows the mean changes in raw mortality rates at the hospital level. Using raw mortality rates, we do not see any changes pre and post HST for both, control and treated hospitals. Standard deviation in parentheses. Number of hospitals in brackets.

Table D.3: Changes in Hospital-Level (IV adjusted) Mortality Rates

Control Hospitals				Treated Hospitals			
Hospital Treatment: 10 mile radius							
Pre-Train	Post-Train	Δ in Means	t-stat	Pre-Train	Post-Train	Δ in Means	t-stat
-0.030	-0.014	0.016	0.4523	0.002	-0.059	**-0.062**	-2.346
(0.192)	(0.229)	t: 0.4523		(0.203)	(0.164)	t:-2.346	
[69]	[69]			[98]	[98]		
Hospital Treatment: 15 mile radius							
Pre-Train	Post-Train	Δ in Means	t-stat	Pre-Train	Post-Train	Δ in Means	t-stat
-0.029	-0.004	0.025	0.5870	-0.002	-0.059	**-0.056**	-2.345
(0.212)	(0.237)			(0.192)	(0.168)		
[55]	[55]			[112]	[112]		

Notes: This table shows the mean changes in case-mix adjusted mortality rates at the hospital level (obtained using IV method). Adjusted mortality rates decrease post-HST for treated hospitals whereas there is no difference for the control hospitals. Standard deviation in parentheses. Number of hospitals in brackets.

Table D.4: Changes in Patient Level Mortality Rates by Destination

Control Hospitals				Treated Hospitals			
Hospital Treatment: 10 mile radius							
Pre-Train	Post-Train	Δ in Means	t-stat	Pre-Train	Post-Train	Δ in Means	t-stat
0.044	0.036	**-0.008**	-3.341	0.039	0.032	**-0.007**	-4.13
(0.205)	(0.187)			(0.193)	(0.176)		
[13,648]	[16,385]			[22,243]	[26,710]		
Hospital Treatment: 15 mile radius							
Pre-Train	Post-Train	Δ in Means	t-stat	Pre-Train	Post-Train	Δ in Means	t-stat
0.047	0.037	**-0.009**	-3.231	0.039	0.032	**-0.006**	-4.293
(0.211)	(0.189)			(0.193)	(0.177)		
[8,577]	[10,620]			[27,314]	[32,475]		

Notes: This table shows the mean changes in patient level (raw) mortality rates by destination (whether patients went to control or treated hospitals). Standard deviation in parentheses. Number of patients in brackets.

Table D.5: Patients's expected mortality rate at the hospital of his choice (Treated Hospitals)

	Pre-HST	Post-HST	Δ	t-stat	t-stat of diff in Δ
Patients who took the train	0.037	0.02	-0.017	1.7362*	1.191
to arrive at the hospital	(0.188)	(0.149)			
Number of patients	792	970			
Patients who did not take the train	0.039	0.032	-0.007	3.8864***	
	(0.193)	(0.176)			
to arrive at the hospital					
Number of patients	21,451	25,740			

Notes: We define “patients who took the train” as patients who traveled more than 50 miles. Difference between the changes in means is not statistically significant ($t = -1.1914$). Standard deviation in parentheses.

Appendix E

Table E.1: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care (including emergency admissions)

	Pre-Post	5-mile Treatment	10-mile Treatment	15-mile Treatment	20-mile Treatment
Raw Mortality Rates	(1)	(2)	(3)	(4)	(4)
Post	-0.0007	-0.00067	0.0028	0.0029	0.0058
	(0.0032)	(0.0045)	(0.0061)	(0.0075)	(0.0095)
Treated×Post		-0.000064	-0.0060	-0.0054	-0.0088
		(0.0062)	(0.0070)	(0.0081)	(0.0099)
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		105	69	55	43
Treated Hospitals		62	98	112	124
Observations	334	334	334	334	334
R-squared	0.6680	0.6680	0.6697	0.6693	0.6709

Notes: Models are estimated by OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants. *** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

Table E.2: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care (including emergency admissions)

	Pre-Post	5-mile Treatment	10-mile Treatment	15-mile Treatment	20-mile Treatment
OLS	(1)	(2)	(3)	(4)	(4)
Post	-0.007**	-0.003	0.0003	0.003	0.005
	0.003	0.004	0.006	0.007	0.008
Treated×Post		-0.009*	-0.012*	-0.014*	-0.016*
		0.005	0.006	0.007	0.009
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		105	69	55	43
Treated Hospitals		62	98	112	124
Observations	334	334	334	334	334
R-squared	0.5823	0.5888	0.5935	0.5967	0.5988

Notes: Models are estimated by OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants. *** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

Table E.3: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care (including emergency admissions)

	Pre-Post	5-mile Treatment	10-mile Treatment	15-mile Treatment	20-mile Treatment
IV	(1)	(2)	(3)	(4)	(4)
Post	-0.027	-0.025	0.026	0.029	0.027
	0.019	0.026	0.029	0.033	0.039
Treated×Post		-0.006	-0.091**	-0.084**	-0.072
		0.038	0.039	0.040	0.045
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		105	69	55	43
Treated Hospitals		62	98	112	124
Observations	334	334	334	334	334
R-squared	0.6052	0.6052	0.6181	0.6152	0.6115

Notes: Models are estimated by OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants. *** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

Table E.4: Diff-in-Diff Estimates of the Impact of Competition on Quality of Clinical Care (including emergency admissions)

	Pre-Post	5-mile Treatment	10-mile Treatment	15-mile Treatment	20-mile Treatment
Raw Mortality Rates	(1)	(2)	(3)	(4)	(4)
Post	-0.00427	-0.00140	0.00283	0.00513	0.00587
	(0.00271)	(0.00345)	(0.00436)	(0.00536)	(0.00623)
Treated×Post		**-0.00819***	**-0.0128****	**-0.0144****	**-0.0136****
		(0.00464)	(0.00500)	(0.00574)	(0.00659)
Above70	0.0130	0.0138	0.0200	0.0165	0.0105
	(0.0536)	(0.0531)	(0.0516)	(0.0515)	(0.0526)
SurgeryRisk	0.0459	0.0558	0.0558	0.0505	0.0495
	(0.122)	(0.118)	(0.117)	(0.117)	(0.118)
DiseaseRisk	0.481*	0.492*	0.495*	0.519*	0.510*
	(0.280)	(0.275)	(0.271)	(0.267)	(0.270)
ComorbidityRisk	1.184***	1.198***	1.191***	1.204***	1.198***
	(0.415)	(0.411)	(0.402)	(0.400)	(0.404)
DisabledFrac	-2.047***	-2.285***	-2.459***	-2.399***	-2.208***
	(0.724)	(0.789)	(0.738)	(0.732)	(0.704)
DisabilitySeverity	2.504***	2.746***	2.929***	2.871***	2.663***
	(0.877)	(0.941)	(0.891)	(0.884)	(0.851)
LowIncomeFrac	-0.00920	-0.0160	-0.0185	-0.0193	-0.00678
	(0.138)	(0.134)	(0.134)	(0.134)	(0.132)
Hospital FE	YES	YES	YES	YES	YES
Number Hospitals	167	167	167	167	167
Control Hospitals		105	69	55	43
Treated Hospitals		62	98	112	124
Observations	334	334	334	334	334
R-squared	0.813	0.816	0.820	0.822	0.820

Notes: Models are estimated by OLS with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. All regressions include constants. *** Significant at the 1 percent level; ** Significant at the 5 percent level;* Significant at the 10 percent level.

Table E.5: Descriptive Evidence of Changes in Travel Distance (Patients that appear in both periods)

	Control Patients			Treated Patients		
	Pre-HST	Post-HST		Pre-HST	Post-HST	
Distance Traveled	Mean	Mean	%Δ	Mean	Mean	Δ
	(st.dev)	(st.dev)	(t-stat)	(st.dev)	(st.dev)	Δ(t-stat)
Patients that appear in both periods (Patient Treatment: 15 miles)						
			Panel A. Distance Traveled (miles)			
	30.965	33.288		13.636	15.876	
	(42.155)	(45.306)	t:1.413	(35.687)	(38.381)	**t:2.045**
Nobs	1,333	1,530	p:0.1577	2,158	2,461	p:0.0409
			Panel B. Traveled to arrive at treated hospitals			
	0.483	0.485		0.066	0.084	
	(0.500)	(0.500)	t:0.092	(0.249)	(0.277)	**t:2.186**
Nobs	605	742	p:0.9264	2,081	2,354	p:0.0289
			Panel C. Traveled to arrive at control hospitals			
	0.074	0.086		0.260	0.318	
	(0.262)	(0.281)	t:0.866	(0.441)	(0.468)	t:0.850
Nobs	728	788	p:0.3865	77	107	p:0.3967

Notes: This table shows summary statistics for patients who appear in both, pre- and post-HST periods. Patient treatment is defined as living within 15 miles of the HST station. Hospital treatment is defined as being located within 15 mile of the HST station. Panel A shows changes in travel distance (miles) in each period for treated and control patients. Panel B shows changes in proportion of patients that arrived at the treated hospitals via traveling. Panel C shows the changes in proportion of patients that arrived at the control hospitals via traveling.

Appendix F

Table F.1: Randomization Check-Men

	Men					
	Control		Treatment			
Variable	Mean	SD	Mean	SD	t-stat	p-value
Age	31.317	9.572	31.281	9.651	0.239	0.811
HighSchool	0.126	0.332	0.127	0.333	-0.112	0.911
TwoYear	0.184	0.388	0.185	0.388	-0.064	0.949
University	0.550	0.498	0.531	0.499	1.367	0.172
PostGrad	0.140	0.347	0.157	0.364	-1.741	0.082
Skinny	0.141	0.348	0.144	0.007	-0.267	0.789
Average	0.674	0.469	0.662	0.473	0.853	0.394
LittleExtra	0.152	0.359	0.156	0.363	-0.366	0.714
Overweight	0.032	0.177	0.038	0.190	-0.956	0.339
Asian	0.099	0.298	0.095	0.293	0.526	0.599
White	0.634	0.482	0.617	0.486	1.514	0.130
Black	0.096	0.295	0.102	0.303	-0.866	0.387
Indian	0.043	0.204	0.042	0.201	0.258	0.797
MidEastern	0.024	0.153	0.026	0.160	-0.652	0.514
Hispanic	0.119	0.324	0.123	0.328	-0.518	0.604
NativeAmerican	0.025	0.156	0.022	0.148	0.767	0.443
PacificIslander	0.012	0.111	0.012	0.110	0.061	0.952

Table F.2: Randomization Check-Women

	Women					
	Control		Treatment			
Variable	Mean	SD	Mean	SD	t-stat	p-value
Age	34.065	11.377	33.855	11.236	0.784	0.433
HighSchool	0.098	0.298	0.085	0.280	1.113	0.266
TwoYear	0.161	0.368	0.153	0.360	0.521	0.602
University	0.562	0.496	0.577	0.494	-0.749	0.451
PostGrad	0.179	0.383	0.185	0.388	-0.364	0.716
Skinny	0.264	0.441	0.238	0.426	1.209	0.227
Average	0.539	0.499	0.537	0.499	0.052	0.959
LittleExtra	0.147	0.354	0.174	0.379	-1.503	0.133
Overweight	0.050	0.220	0.051	0.220	0.010	0.992
Asian	0.136	0.343	0.159	0.365	-1.924	0.054
White	0.594	0.491	0.584	0.493	0.649	0.516
Black	0.096	0.295	0.102	0.303	-0.866	0.387
Indian	0.013	0.114	0.018	0.133	-1.216	0.224
MidEastern	0.008	0.089	0.008	0.089	-0.007	0.995
Hispanic	0.123	0.328	0.125	0.331	-0.201	0.841
NativeAmerican	0.020	0.142	0.022	0.148	-0.390	0.697
PacificIslander	0.012	0.111	0.012	0.110	0.061	0.952

Appendix G

Table G.1: User Acitivites (All correspondent users)

	Men			Women		
	control (1a)	treated (1b)	t-stat (1c)	control (2a)	treated (2b)	t-stat (2c)
Number of users	7,930	8,189		3,470	3,642	
Profiles Browsed						
Mean	259.4	250.4	-0.894	190.7	210.6	1.385
Median	59	58		27	31	
SD	641.8	627.4		567.1	639.9	
Profiles viewed						
Mean	84.1	84.7	0.190	39.8	46.6	**3.630**
Median	24	26		15	17	
SD	186.0	180.8		69.9	86.5	
Likes sent						
Mean	95.6	83.9	**-2.334**	15.3	17.2	1.0338
Median	9	10		1	1	
SD	341.3	292.4		80.1	75.3	
Initiated messages						
Mean	22.4	22.4	-0.058	3.1	3.3	0.860
Median	3	3		1	1	
SD	69.5	70.5		9.2	9.6	
Initiated Messages that led to match						
Mean	1.6	1.7	0.891	0.7	0.8	1.789
Median	0	0		0	0	
SD	5.1	5.3		2.2	2.5	

	Men			Women		
	control	treated	t-stat	control	treated	t-stat
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Number of users	7,930	8,189		3,470	3,642	
Profiles Browsed						
Mean	259.4	250.4	-0.894	190.7	210.6	1.385
Median	59	58		27	31	
SD	641.8	627.4		567.1	639.9	
Profiles viewed						
Mean	84.1	84.7	0.190	39.8	46.6	**3.630**
Median	24	26		15	17	
SD	186.0	180.8		69.9	86.5	
Likes sent						
Mean	95.6	83.9	**-2.334**	15.3	17.2	1.0338
Median	9	10		1	1	
SD	341.3	292.4		80.1	75.3	
Initiated messages						
Mean	22.4	22.4	-0.058	3.1	3.3	0.860
Median	3	3		1	1	
SD	69.5	70.5		9.2	9.6	
Initiated Messages that led to match						
Mean	1.6	1.7	0.891	0.7	0.8	1.789
Median	0	0		0	0	
SD	5.1	5.3		2.2	2.5	