THE ROLE OF INFORMATION SYSTEMS IN HEALTHCARE

by

Jianing Ding

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

School of Management

West Lafayette, Indiana

May 2023

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Susan Feng Lu, Co-Chair

Krannert School of Management, Purdue University

Dr. Jinyang Zheng, Co-Chair

Krannert School of Management, Purdue University

Dr. Karthik Kannan

Eller College of Management, University of Arizona

Dr. Zaiyan Wei

Krannert School of Management, Purdue University

Approved by:

Dr. Yanjun Li

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Fundamental changes have been happening in healthcare organizations and delivery in these decades, including more accessible physician information, the low-cost collection and sharing of clinical records, and decision support systems, among others. Emerging information systems and technologies play a signification role in these transformations. To extend the understanding and the implications of information systems on healthcare, my dissertation investigates the influence of information systems on enhancing healthcare operations. The findings reveal the practical value of digitalization in indicating healthcare providers' cognitive behaviors, responding to healthcare crises, and improving medical performance.

The first essay investigates the unrevealed value of a special type of user-generated content in healthcare operations. In today's social media world, individuals are willing to express themselves on various online platforms. This user-generated content posted online help readers get easy assess to individuals' features, including but not limited to personality traits. To study the impact of physicians' personality traits on medicine behaviours and performance, we take a view from the perspective of user generated content posted by their supplier side as well as using physician statements which have been made available in medical review websites. It has been found that a higher openness score leads to lower mortality rates, reduced lab test costs, shorter time usage in hospitals treated by physicians with greater openness scores. Furthermore, taking these personality traits into consideration in an optimization problem of ED scheduling, the estimation of counterfactual analysis shows an average of 11.4%, 18.4%, and 17.8% reduction in in-hospital mortality rates, lab test expenditures, and lengths of stay, respectively. In future operation of healthcare, physicians' personalities should be taken into account when healthcare resources are insufficient in times of healthcare pandemics like COVID-19, as our study indicates that health service providers personality is an actual influence on clinical quality.

In the second essay, we focus on the influences of the most severe healthcare pandemic in these decades, COVID-19, on digital goods consumption and examine whether digital goods consumption is resilient to an individuals physical restriction induced by the pandemic. Leveraging the enforced quarantine policy during the COVID-19 pandemic as a

quasi-experiment, we identify the influence of a specific factor, quarantine policy, on mobile app consumption in every Apple app store category in the short and long terms. In the perspective of better responding in the post-pandemic era, the quantitative findings provide managerial implications to the app industry as well as the stock market for accurately understanding the long-term impact of a significant intervention, quarantine, in the pandemic. Moreover, by using the conditional exogenous quarantine policy to instrument app users daily movement patterns, we are able to further investigate the digital resilience of physical mobility in different app categories and quantify the impact of an individuals physical mobility on human behavior in app usage. For results, we find that the reduction in 10% of ones physical mobility (measured in the radius of gyration) leads to a 2.68% increase in general app usage and a 5.44% rise in app usage time dispersion, suggesting practitioners should consider users physical mobility in future mobile app design, pricing, and marketing.

In the third essay, we investigate the role of an emerging AI-based clinical treatment method, robot-assisted surgery (RAS), in transforming the healthcare delivery. As an advanced technique to help diminish the human physical and intellectual limitations in surgeries, RAS is expected to but has not been empirically proven to improve clinical performance. In this work, we first investigate the effect of RAS on clinical outcomes, controlling physicians' self-selection behavior in choosing whether or not to use RAS treatment methods. In particular, we focus on the accessibility of RAS and explore how physician and patient heterogeneity affect the adoption of the RAS method, including learning RAS and using RAS. Investigating the decision-making process on RAS implementation in both the learning and using stages, we show the synergy of RAS implementation in alleviating healthcare racial disparity. Ultimately, the mechanism analysis will be conducted to reveal the underlying mechanism that induces the enhancement of surgical outcomes. For instance, the estimations tend to reveal that, more than surging clinical performance, RAS tends to increase standardization in time and steps when applying the treatment procedures.

1. INTRODUCTION

Healthcare largely influences individuals' lives and the functionality of our society. At the individual level, healthcare shortage or malpractice undermines the function of one's ordinary life. At the population level, the prevention and control of diseases are crucial for maintaining public healthcare. At the societal level, healthcare disparity, even the global health crisis, have severe consequences on productivity and social welfare. Considering the COVID-19 pandemic, which quickly spread worldwide and caused 6,625,121 cases of global death by 23 November 2022 based on Johns Hopkins University statistics[1], the role of Information systems (IS) has been highlighted not only in facilitating health service delivery but also in supporting normal living amid the pandemic. Moreover, IS has been shown the enormous potential to facilitate the pursuit of healthcare quality. For instance, past decades have observed a surge in clinical quality and efficiency tremendously through electronic health record systems [1], telemedicine [2], robot-assisted surgery [3], and the Internet of medical things [4], among others. Actually, IS has the power to establish a more effective design for providing health service delivery, even back to unpredictable healthcare interventions such as the COVID-19 crisis. In this perspective, this dissertation targets at investigating the even broader applications and implications of IS to the healthcare field from different perspectives, including data monitoring for hindsight, attribution analyzing for insight, prediction making for foresight, and eventually, decision-supporting as artificial intelligence.

The first essay investigates the unrevealed value of a special type of user-generated content in reflecting health service performance and improving healthcare operations. We look into the personal statements posted by physicians on the supply side and investigate their power to reflect physician personality and clinical performance. Moreover, we show leveraging personality traits provides a promising pathway toward understanding physicians' cognitive behavior and clinical performance. Specifically, we take advantage of physicians' personal statements collected from a healthcare online review system and we end up with 2,073 physicians' personality scores from the unstructured physician personal statements following the Big Five model of personality. To tackle with the two-way selection concern between

[1]↑COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). ArcGIS. Johns Hopkins University. Retrieved on 23 November 2022.

patients and service providers and test the individual-level instinctive behavior driven by personality traits, we leverage a special research design under emergency departments (EDs) as well as unexpected accidental encounters in Florida state. Furthermore, to alleviate the living location selection bias, we only retain encounters who live in a area different from the hospital's health service areas to simulate the exogenous matching between patients and service providers. The measurement errors of personality detection are corrected by using multiple detection algorithms and by repeated measurements, in which the personality scores detected more than once from difference statement sources. According to our analyses, patients treated by doctors who score higher in openness have lower rates of in-hospital mortality, lower costs for lab tests, and shorter hospital stays due to the implementation of new treatment methods. On the other hand, doctors who score higher in conscientiousness tend to have higher lab test costs and longer waiting times for major procedures. Additionally, doctors who score higher in agreeableness are more likely to help patients save on lab test costs, particularly for low-income patients, due to their empathetic nature. These findings deepen the understanding of the personality theory in the background of healthcare ED, reveal a huge hidden value of supply-side user-generated content in predicting service quality, and provide managerial implications regarding the future potential of leveraging physician personality in healthcare operations management.

The second essay investigates the digital resilience of digital goods consumption to the healthcare pandemic and physical restriction, respectively. Amid the COVID-19 pandemic, the quarantine policy, one of the most frequently applied and effective disease control methods, is strictly executed in many regions. As individuals are forced to stay at home due to the unpredictable and exogenous infection of unknown neighbors, the quarantine policy is conditionally exogenously imposed on individuals. From an empirical perspective, the implementation of quarantine policy provides us with a decent instrumental variable for individual-level physical mobility and allows us to examine whether individuals physical mobility has a causal effect on the consumption of digital goods, represented by app usage behavior in this study. Leveraging the natural experiment of being quarantined and the two-stage least-squares design, we first reveal the alteration of individual physical mobility and app usage behavior induced by the quarantine policy in the short term and the long run.

In terms of physical mobility, our estimation underpins the considerable decrease in visiting places and movement radius of gyration during quarantine and, surprisingly, reveals the compensatory behavior in the long-term by showing users tend to enhance physical mobility even more intensively after being released from quarantine. Second, using the quarantine policy to instrument physical mobility allows us to investigate and quantify the causal effect of physical mobility on app usage. The estimations show that 10% reduction in ones physical mobility leads to 2.68% increase in app usage in cellular data and 5.44% increase in app usage time dispersion, which sheds light on the resilience of digital goods consumption to physical restriction. Last, from the managerial implication perspective, we identify the app heterogeneity by specifying the effect of physical mobility on the app for each app category and high/low-ranked app groups. The business strategies for app designers, practitioners, and policy makers are developed to leverage the international travel quarantine and the long-term complementary effect of physical restriction.

AI-based treatments in medical care are being observed from various angles, indicating their increasing applicability. In the third essay, we investigate the influences and the decision-making process of implementing a particular AI-based treatment, robot-assisted surgery (RAS). RAS is designed to reduce the surgical limitations of human vision and avoid accidental movements. To provide with empirical evidence on the role of RAS in performance improvement, in this study, we study the RAS method and show it has the potential to alter medical service delivery processes and outcomes. To operationalize our research design, we focus on RAS in high-risk diseases due to its high mortality rate and high demand for surgery precision. And we only look for those diagnoses that RAS procedures can treat. The following three trending questions are studied by modeling physicians decision-making process when learning and choosing treatment methods. First and foremost, we examine the effect of RAS on clinical outcomes, such as whether RAS indeed enhances clinical performance, controlling physicians self-selection behavior in choosing their treatment methods. Second, we investigate the access to learn and use RAS considering physician and patient heterogeneity, especially ethnicity features. For instance, we investigate the concern of racial health inequality by investigating the potential racial bias in the accessibility of RAS resources and revealing the synergy of RAS implementation on alleviating healthcare racial inequality.

More than that, underlying mechanisms are proposed and tested to provide more insights explaining how RAS can improve clinical performance. Specifically, RAS tends to show the potential of enhancing standardization and team cooperation in the treatment process, which might act as moderators to further gradually transform the health service delivery process. Our research addresses this gap in the existing literature and proves a particular AI implementation in healthcare, RAS, improves overall clinical performance. Moreover, we advance the understanding of RAS implementation by identifying the physicians decision-making process of treatment selection and investigating the disparity in the access to learning RAS and access to using RAS. Additionally, we expect the later mechanism analysis could provide suggestive implications to practitioners and policymakers and enrich our understanding of the future of RAS in the healthcare industry.

Integrating the above, the three studies in my dissertation profile have investigated the value of IS in healthcare related topics from several perspectives. By conducting empirical analyses, all of the proposed chapters aim to reveal the main theme: the utilization of information systems and information technologies has the significant role in transforming healthcare delivery and is a good channel for responding to healthcare interventions. The findings from three chapters of studies together provide implications on how to better design the healthcare delivery and help achieve higher social welfare. The insights are versatile but uniformly stress the critical role of IS in healthcare topics. The first project reveals the considerable value of supply-side user-generated content in reflecting healthcare providers' personalities, cognitive behaviors, and clinical outcomes. The second work highlights the power of digital resilience with a special focus on physical restrictions amid the healthcare pandemic, which implies one advantage of the digital economy is less reliance on offline operations. The third study models how healthcare service providers make decisions on learning and using an advance AI-based treatment, RAS, and how the RAS treatment plays a role in transforming the healthcare service delivery to be more efficacy and standard. The first project is under revision at Information Systems Research. The second project is under polishing and will be submitted for review very soon. The third project is a work in progress and will be completed in later months. The remaining sections of the dissertation are organized as follows: Chapter 2, 3, and 4 delve into the first, second, and third essays, respectively, as previously described. Chapter 5 presents a summary of the primary findings and outlines the next steps to be taken.

2. WHAT CAN PERSONAL STATEMENT TELL US? INSIGHTS ABOUT PHYSICIANS' PERSONALITY TRAITS AND CLINICAL PERFORMANCE

2.1 Introduction

Healthcare services have been undergoing significant changes because of information technology. These changes have only accelerated amid COVID-19 pandemic, including the surge in use of telemedicine [5], [6], the reduced information asymmetry due to online healthcare platforms [7], [8], the app implementation on contagion control [9], among others. While these technology related services are taking hold, hospitals and clinics are generally attempting to be nimbler [10]. In this paper, we attempt to focus on one novel way for hospitals and clinics to be more nimble by understanding the personality traits of physicians and the induced impacts on their clinical performance.

Note the personality-performance model has been proposed by psychologists for decades [11], indicating one's achievement and career success can be largely explained by personality traits [12], [13], but not yet been rigorously tested so far. Specifically, existing psychology literature relies on questionnaires to infer personality traits for a limited number of subjects and document the rough correlations between personality features and cognitive behaviors [14]. To our knowledge, there are few empirical papers that systematically document the causal inference between personality and performance due to the following two reasons: (1) the lack of personality data at a large scale; and (2) the shortage of clean identification strategies.

Integrating the user-generated content (UGC) and natural language processing (NLP) techniques allows us to overcome the first challenge, achieving large-scale personality data. Notably, many healthcare platforms not only allow for demand side engagement in terms of posting reviews, but also allow the supply side expression, in the form of service providers posting personal statements. In light of demand-side UGC's predictive power of product quality, sales, among others [15]–[18], we seek to extract valuable information from supply-side UGC leveraging NLP. Existing literature has provided compelling evidence that human beings linguistics have tremendous psychological value [19]–[21]. Recently, the rise of social media and the development of NLP techniques help to break down the data barriers [22]–[24]. In addition, NLP approach fixes two deficiencies in the traditional methods. First, it mitigates the sample selection issue due to the low response rates to surveys or questionnaires. Social media accumulates substantial online word of mouth, which allows us to have a glimpse of individual personalities through their public postings and generates a representative sample of interest for personality analyses on a large scale with low cost [25]. Second, an interviewee could lie when filling in a survey. However, one's linguistic style is immune to the manipulated information, as what we analyze for personality measure is language using rather than the content itself. While one can report false information, the language using is not the focus of faking and not easily to be faked. For example, the usage of past or present tense verbs, first or third person pronouns, the percentage of article using, all of these writing styles are what we rely on for personality detection. Thus, in this study, we attempt to leveraging NLP techniques to extract physician personalities from UGC that physicians posted on online healthcare platform.

In particular, we describe physician personality traits using the widely recognized Big Five personality traits, which are also known as OCEAN: Openness to experience (O), Conscientiousness (C), Extraversion (E), Agreeableness (A), and Neuroticism (N). The NLP techniques is applied to physician personal statements for obtaining their personality trait measures. We then evaluate the impact of these traits on three well-accepted healthcare performance measures: in-hospital mortality rates, medical expenditures, and patients' length of stay (LOS). We obtained healthcare performance measures from the patient discharge data in Florida. Combining these datasets, we attempt to investigate whether physicians' Big Five personality traits indicate their clinical performance in terms of patient's health outcomes.

The other challenge comes from the endogeneity issues in establishing the personality-performance relationship. For example, patients and physicians are not randomly matched since patients usually select their physicians. However, data corresponding to patients with accidental injuries in ED is quasi-randomly matched and conducive for our analysis. Hence, we address this identification challenge by using those patients who were injured accidentally, sent to ED and assigned to on-call physicians. For another example, to correct for the measurement errors of the Big Five personality trait scores, we impose an error correction model which uses two types of statements for cross validation. The final sample involves 180 hospitals, 2,073 physicians, and 98,701 emergency visits made by patients encountering accidental injuries.

The findings of this study demonstrate a strong correlation between physicians' clinical performance and their personality traits, as predicted by the Big Five personality model. Specifically, physicians who score high on openness tend to be receptive to new ideas and innovations. Our analysis reveals that physicians with high scores on openness tend to adopt new treatment procedures earlier than their peers, resulting in lower mortality rates, lower lab test costs, and shorter lengths of stay for their patients. On the other hand, conscientious physicians tend to follow systematic protocols and take a thorough approach, which can result in longer waiting times for their patients to receive principal treatment and higher lab test costs. Agreeable physicians, who tend to be empathetic, help patients save on medical expenditures for lab tests, particularly low-income patients. We conducted additional tests to understand the underlying mechanisms, and the results were robust across different samples, specifications, and measures.

The contributions can be found in the following three aspects. First, our paper is the first to investigate and find the relationship between physicians personalities on their patients' health outcomes. In the perspective of psychological and healthcare literature, this work takes the first step to find that physicians personality matters in treatment and remarkably contributes to understanding the Big Five personality theory in healthcare ED context. Second, this work is the first to reveal the huge hidden value of supply-side UGC in indicating service performance. The findings imply that online platforms should provide channels to suppliers generating such content express themselves and indicate future research might explore more on the predictive ability of supply-side UGC on other dimensions. Particularly in healthcare field, noticing online information about physicians become an important informational source for patients to find the right doctors [7], it is crucial for us to understand and exploit supply-side UGC in healthcare delivery and operations. Third, we contribute to healthcare management in practical perspective by revealing the importance of service provider personality factors. Despite the numerous studies in the healthcare operations literature to enhance operational efficiency in hospitals [26]–[31], there is a limited amount of research that examines care delivery through the lens of psychology and investigates how physicians' personalities impact patient outcomes. Meanwhile, we validate the personality measures generated by the NPL method. Leveraging the physician posts on online review platform and social media, the extracted personality measures can be used for further improve user experience and healthcare delivery.

The rest of this paper is structured as follows. In Section 2.2, we provide the background, and the validation of online personal statements. Section 2.3 reviews the related literature with respect to healthcare operations and personality traits. In Section 2.4, we develop our hypotheses, which capture the important and relevant traits from the healthcare setting. In Section 2.5, we present our data structure. In Section 2.6, we discuss our research design and the econometric specifications of our study. In Section 2.7, we describe our main results and the robustness checks for this study. Finally, we discuss the potential mechanism in Section 2.8 and conclude our paper with actionable insights in Section 2.9.

2.2 Background on UGC by physician, and Validation of Personality Traits

This section introduces background related to physician generated contents and the validity of extracting personality traits from them.

2.2.1 UGC by Physician

The rise of social media alleviates information asymmetry between patients and physicians to some extent. The earlier versions of these online platforms focused on reviews that consumers contributed. In particular, patients and their family members can voluntarily share their experiences online, express satisfaction about their received medical services, and provide ratings on a specific physician. As reported [32], the introduction of online physician review platforms significantly changes the demand for individual physicians and motivates physicians to improve their performance. Recently, some innovative features have emerged in online physician review platforms. In addition to providing basic physician information including name, working address, phone number, educational background, and board certificates, more platforms have started encouraging physicians to express themselves online, so as to allow physicians to market and speak for themselves publicly.

One of the primary data sources for this study is the personal statements of physicians posted on one of the largest online physician review platforms. A typical personal statement includes two parts, care philosophy and biography. The biography statements generally contain a physicians educational background and specified expertise. Many physicians also share their altitudes to patients and lives in care philosophy. To provide direct understanding of how physician statements look like, we list two examples in Figure 2.1. And we illustrate additional examples of website interface and online personal statements in Appendix 2.10.1. It is noteworthy that the content and interface of this website kept changing in recent years [1], thus the personal statements we collected should be precisely described as the statement data in Jan, 2019.

2.2.2 Validation of Personality Detection and Data Generation Process

Given that linguistic features infer personality traits leveraging NLP methods [19]–[21], in this study, we facilitate personality extraction integrating NLP and physician personal statements provided on the platform. Specifically, for each physician, we conduct a text analysis to connect the unstructured personal statements with one's revealed personality traits through the Linguistic Inquiry and Word Count (LIWC) dictionary and calculate one's personality trait score along each of the Big Five traits. The processing method and the introduction of the LIWC dictionary are detailed in Appendix 2.10.2. Notably, physician personalities are captured not only by the content of their respective statements, but also by the words they select and their manner of speaking. Therefore, we assume it is difficult for individual to hide revealed personalities, despite one could strategically choose over the information to be posted on platforms.

Each personal trait score is measured by percentile in the physician sample and is scaled in the range of 0 to 1, with 0 indicating the lowest level and 1 indicating the highest level

[1]↑For example, we have observed the remove of some care philosophy; interface alteration of physician personal web page; and update of personal statement.

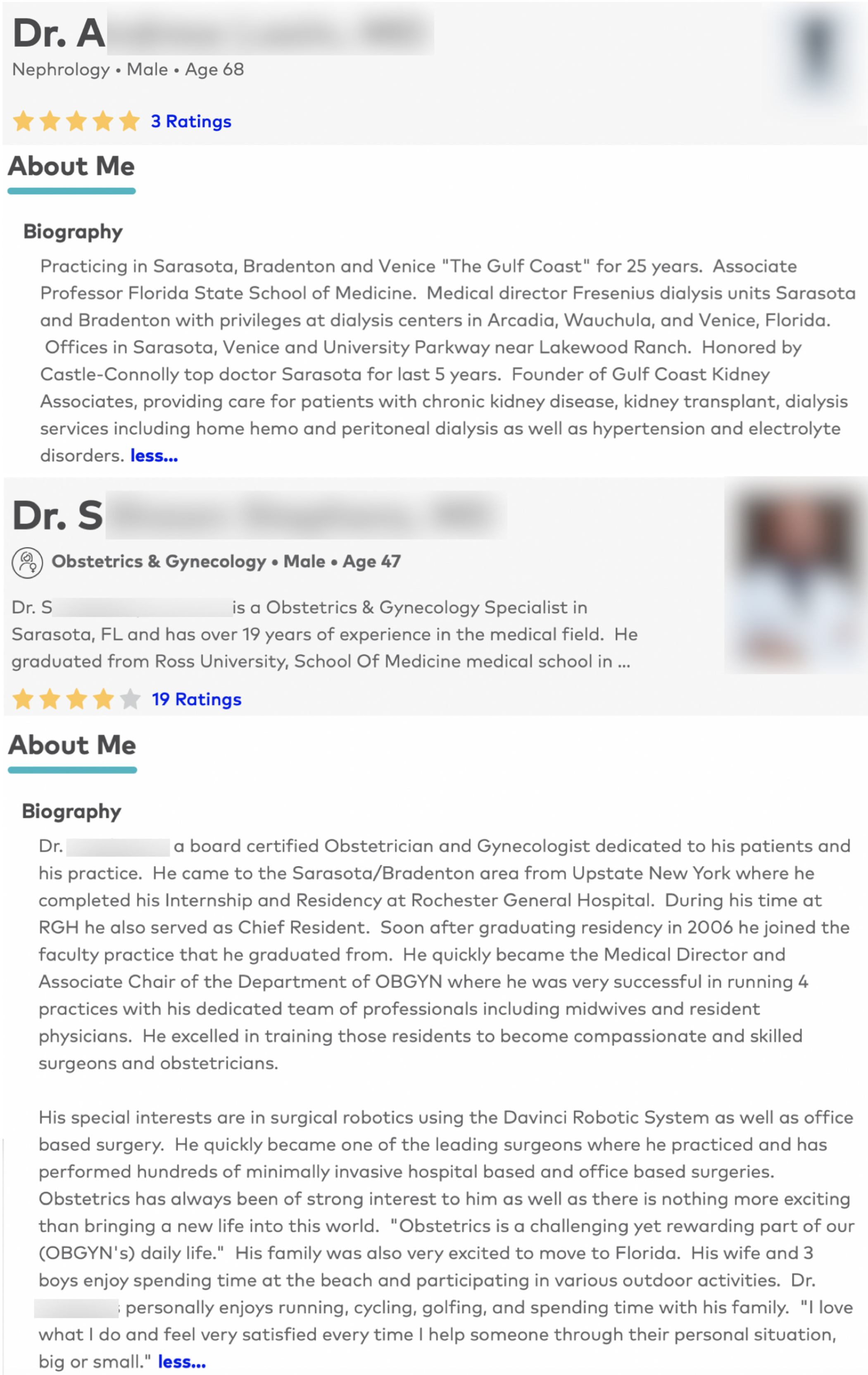
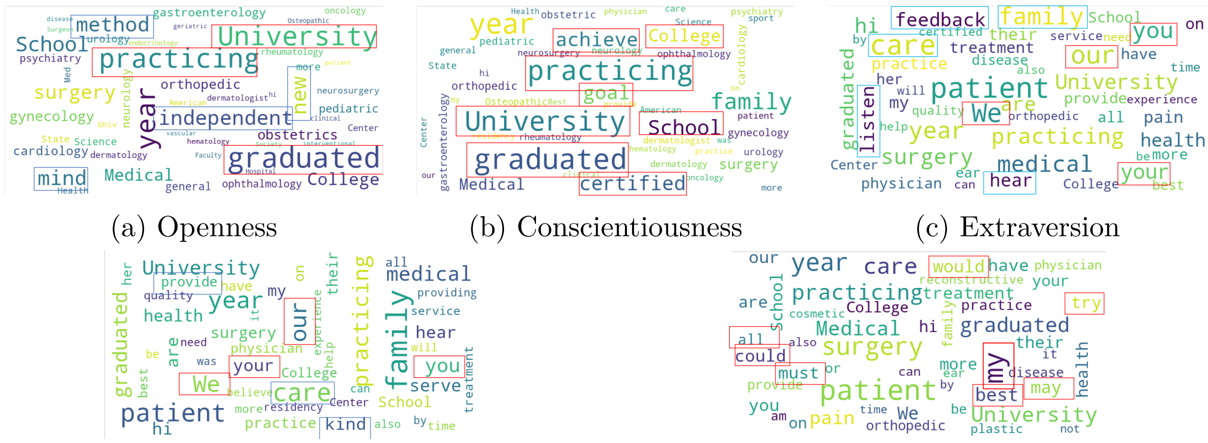


Figure 2.1. Examples of Physician Personal Statement

for the corresponding trait. To provide a straight-forward demonstration of each personality trait, we use the statements of physicians whose specific trait scores are above the median to make the word cloud. Figure 2.2 shows that the linguistic differences between different personality traits are quite distinct. For example, physicians with high openness scores tend to mention medical terms, methodologies, and future goals, while those with high conscientiousness scores place more emphasis on their past achievements, like education and certificates. Also, extroverts are likely to include words focusing on communication and social activities, while physicians with high agreeableness scores are likely to include words that illustrate kindness and care to their patients. In contrast to others, physicians with high neuroticism scores seldom use plural pronouns. Instead, they prefer using singular pronouns and superlative adjectives. To provide the clear structure, we summarize and provide guidelines for distinguishing personality traits according to the way people express themselves in Figure 2.2f, in terms of how they use pronouns, what kinds of topics they are likely to post, and whether there are any special word-using habits they have.

However, there are several concerns regarding extracting personal from physician statements via NLP methods. First, one may question about how much useful information can be contained in limited words with resembled structure. Notably, although some contents presented in physician statements are seemingly similar, the words they use help us detect their personalities. To illustrate what makes difference in personality detection, we elaborate how physician statement affects personality taking examples shown in Figure 2.1. Browsing personal statements, we can observe the differences in word usage and posted topics manifest the significant variations in Big Five personality scores, as shown in Table 2.1. For example, what Dr. A mentioned is related to honors and achievements, resulting in a high consciousness score. In contrast, Dr. S emphasizes more on personal lifestyle, family, and care to patients, thus he receives higher scores in extraversion and agreeableness compared to average. Moreover, the usage of uncertain words leads to a high score in neuroticism for Dr. S. Another potential concern is whether those online personal statements are written by the physicians themselves. Notably, although it is impossible to tease out the case that one's statement might be polished by other editors or administrators, we argue those changes are independent identically distributed and are captured in error term. Moreover, both physi-



(a) Openness (b) Conscientiousness (c) Extraversion

(d) Agreeableness (e) Neuroticism

	Pronouns	Topics	Typical Word Using
Openness	Rarely mentioned	Achievements; Future goals; Thought	Methods; Terms
Conscientiousness	Rarely mentioned	Achievements; Certificate	
Extraversion	Plural pron.	Care; Family; Feedback; Hear	
Agreeableness	All pron.	Care; Family; Sense	
Neuroticism	First single pron.	Care	(Un)Certainty; Superlative adj.

(f) The guidelines for distinguishing personalities

Figure 2.2. Word Cloud and Language Using Habits of Each Personality Trait

cians have served as on-call ED physicians with different specialties, suggesting the diversity of physician specialties in our sample.

Table 2.1. Big Five Scores Extracted from Personal Statements

Personality Traits	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Dr. A	0.84	0.79	0.35	0.31	0.22
Dr. S	0.22	0.29	0.83	0.95	0.75

To further alleviate the above concern of personality detection accuracy using small text and ghost-writing, we carefully examine the validity of personality measures extracted from physicians statement. First, we enlarge the number of detection words by collecting the Tweets data of a subset of physicians who can be matched with Florida data and examining the consistency of personality detection. After collecting physician's Tweets and filtering out the user account posting less than 5 Tweets in entire timeline, we have 233 physicians, including 109 physicians have been working in ED situation, successfully matched to our Florida physician statement data. On average, one account has 26 Tweets posted with the total word count per account equals to 506 words. By comparing the personality scores detected by physician personal statements (113 words on average) and the personality scores detected by physician Tweets (506 words on average), we present the consistency rate [2] and the Pearsons correlation between two measures of physician personality traits and summarize the results in Table 2.2. According to the generally reported accuracy of computational linguistic approach in binary classification task [24], while differing in corpus and approaches, method with binary test accuracy ranging between 64.38% and 70.39% is decent to be considered as reliable and efficient. Moreover, the Pearsons correlation between personal statement detected personality and Tweets detected personality supports the high correlation between the two sets of personality scores. Both imply the validity of personality measures detected by physician statements.

Table 2.2. Comparison between Personal Statements Personality and Twitter Personality

Personality Trait	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Detection Consistency	64.81%	66.09%	64.38%	70.39%	66.52%
Pearsons Correlation	0.3947***	0.4178***	0.4556***	0.4340***	0.4121***

P-value for Pearson's correlation:
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Moreover, we validate the statement-detected personality measures by applying advanced detecting method, Word Embedding with Gaussian Process method [23], which is proved to largely reduce the requirement of text size with about eight times fewer data for personality detection. Leveraging the Word Embedding with Gaussian Process method further helps us overcome the concern of small size of text in physician personal statements and provides us with a set of personality measures to validate the measures developed by LIWC (See results in Section 2.7.2).

[2] ↑Note that the consistency rate in Table 2.2 is calculated by the rate that personal statements and Tweets extract same output in binary classification task for each trait dimension.

2.3 Literature Review

This section reviews three streams of literature that are highly relevant to our work: personality traits, user-generated content (UGC), and healthcare operations.

2.3.1 Literature on Personality Traits

Psychologists have proposed various schemes for personality modeling such as 16PF [33], EPQ-R [34], MBTI [35], among others. Of the three streams of literature, the Big Five personality traits [11], also known as the OCEAN model or the five-factor model (FFM), is the most extensively employed due to its high consistency across age and gender. We adopt this personality framework in this study and provide the conceptual details of individual personality traits in Appendix 2.10.5.

While much of the initial work personality involve questionnaires or experimental subjects [36], [37], there is an increased attention paid to method development of inferring personality traits from language. Many find personality trait associations with linguistic cues, such as acoustic parameters [19], lexical categories [20] and more complex phrases [21]. Remarkably, Pennebaker and King [20] identified numerous linguistic characteristics linked to each of the Big Five personality traits based on LIWC. Similarly, Yarkoni [38] find that all Big Five personality traits are predominantly expressed in participants language use. Computer scientists have recently developed NLP methods to infer Big Five measures. Youyou et al. [39] show that personality traits can be achieved with higher accuracy. Harrison et al. [40] infer CEOs personality traits from their public speech using NLP. Adamopoulos et al. [25] study the influence of personality features induced through the channel of word-of-mouth and uses NLP to infer personality traits.

Prior research has extensively shown that personality in general influences an individuals adaptation to a specific job or organization [41]–[43]. In the focus of healthcare, Bulmer et al. [44] and Mullola et al. [45] demonstrate that career choices in the medical field differ depending on personality traits. Personalities also predict their performance as teachers in some hospitals. According to Scheepers et al. [46], conscientious, extraverted, and agreeable physicians are more engaged when teaching and also had better performance as teachers.

Moreover, existing studies suggest that physician personality traits associate to service quality as a healthcare provider, including Teng et al. [47] show nurse personalities are relevant to perceived care quality, Phillips et al. [48] find personalities of orthopedic resident are relevant to perceptions of performance, Gur et al. [49] and Dillon et al. [50] show healthcare providers personality traits are related to teamwork and communication, which further affect the perceptions of service quality. Note that, while there are many studies that have associated physician personality traits with healthcare providers service quality directly or indirectly, the measures of service quality is either self-reported or other rated, which is highly biased. Different from other studies, we are the first to reveal the association between physician's personality traits with clinical performance measured by patient's health outcomes in terms of in-hospital mortality rate.

Drawing upon established theoretical concepts (as described in Section 2.4), our focus is on how the underlying personality traits of physicians can influence their clinical performance. In this respect, this work remarkably contributes to understanding the Big Five personality theory in healthcare ED context. Our research significantly deviates from prior studies on personality and healthcare performance, as we are the first to uncover the influence of physicians' personality traits on their patients' actual mortality. Specifically, we investigate how various personality dimensions impact patient clinical outcomes in terms of in-hospital mortality rate, lab test expenditure, and length of stay in the hospital, and how these personality traits tend to prompt physicians to behave instinctively and influence their clinical performance. Moreover, extracting the personality traits from physician personal statement via NLP method in large scale [25], we find physician in ED situation exactly behaves in accordance with Big Five personality traits theory, which validates the personality measures detected from linguistic features theoretically and empirically and supports the feasibility of leveraging personality measures as physician characteristics in future research.

2.3.2 Literature on UGC

Online review platforms have been extensively studied in the information system literature and marketing work [51]–[54]. Existing studies mainly address two types of questions.

A representative question of the first type, in the objective of increasing UGC production, explores what factors motivate users to create and share UGC. Studies that have already been conducted suggest different reasons why consumers may engage in UGC, including but not limited to the need for social interaction, building their reputation, receiving economic benefits, showing concern for other users, and fulfilling their desire for self-actualization [55]–[61]. The representative question of the second type aims to address how UGC affects sales [15]–[18] and firm marketing strategies [62], [63].

Realizing the prevalence of firm-initiated social media, researchers began to focus on the engagement of marketers on social media and study the effect of marketer-generated content (MGC). A substream of research explores the relative functions of MGC and UGC on online social platforms and reports UGC is more effective than MGC in increasing consumers purchase [64]–[66]. Some aim to leverage the joint effect of MGC and UGC on predicting product sales [67], [68]. Moreover, in practical perspective, marketing works focus more on demonstrating and utilizing the role of MGC in driving customer engagement [69], [70].

Despite these research efforts in studying UGC on demand side and MGC, there is few study in literature exploring UGC on supply side. Different from studies on MGC or demand-side UGC, this work is the first to reveal the huge hidden value of supply-side UGC in indicating service performance. The findings imply that online platforms should provide channels to suppliers generating such content express themselves and indicate future research might explore more on the predictive ability of supply-side UGC on other dimensions. Notably, we can easily differentiate demand and supply-side UGC by whether the content is generated by consumer or producer. Meanwhile, supply-side UGC and MGC follow the distinct principal in terms of entity and objective. First, MGC is owned by brands, rather than a single user, and managed by a hired professional team. Second, MGC is used for purely advertising purpose by sharing promotions and new releases. In contrast, retailer or service provider, as a social media user, should has more complicated incentives for posting. Therefore, our study fills in the blank in UGC and MGC literature by identifying the potential of supply-side UGC on predicting service quality.

2.3.3 Literature on Healthcare Operations

The existing studies in healthcare operations management primarily focuses on the impact of operational factors or economic influences, including workload [26]–[30], queue design [71], [72], and schedules [73] on healthcare outcomes. Some works consider how individual-level features of service providers might affect the process of operations, including temporary workers [74], peers [75], gender concordance [76], and experience [77].

Literature most relevant to our study investigating the effect of online review systems on healthcare is small but growing. While Gray et al. [78] and Saifee et al. [79] report no salient correlation between clinical outcomes and physicians online reviews, other evidence reveals valuable information contained in online ratings. Meanwhile, Gao et al. [80] find a positive association between customer's reviews and service provider experience even many of other characteristics. Gao et al. [81] further reveal the association between generated scores and patient perceptions of quality. Moreover, Lu and Rui [7] show that consumers can trust online physician ratings as highly-rated physicians induce lower mortality rate in CABG surgeries.

Different from above literature, our study takes the first step to find that physicians' personality matters in healthcare service delivery and suggest that physicians personality traits could be exploited practically in healthcare management in the following aspects. In the perspective of patients, this study sheds light on leveraging physician personal statement on online rating platform and considering implied personality as well as clinical performance in physician selection. In the perspective of healthcare facilities, we incorporate empirical results of physician personality effects into a scheduling problem and provide a prescription to improve the efficiency in hospital EDs. In the perspective of platform developers, we show possibility that app developers can improve the recommendation system and user experience through better matching patients with physicians taking personality factors into consideration. Therefore, this study is not limited to offer the empirical contributions to the literature, but also provides the practical and managerial insights for patients, app designers, and healthcare industry.

2.4 Hypotheses Development

The specific role of physician's personality traits in clinical performance can be investigated in healthcare emergency context, in which design physician behaves instinctively and makes treatment strategies highly following inherent personality [82]. Key variables of interest in this study are the most concerned three dimensions in healthcare performance, including medical quality, cost, and access [83]. Prior research has measured medical quality using mortality rate [7], [84], [85], and measured individual-level cost using medical expenditure information [86]. As regards accessibility, even though access is a metric mainly considered at the hospital level, measures such as individual LOS can indirectly reflect the integrated hospital accessibility [2]. For example, reducing the length of patient stay can help to increase the efficiency of emergency care by freeing up beds more quickly. More additional details about these measures are provided in Section 2.5.2.

Of the Big Five measures, the one with most wildly accepted conception is extraversion, which captures one's sociability [87]. While extroverts are observed to have higher performance in job with a social component [88], considering social interactions are rarely involved in emergency medical contexts, we conjecture extraversion won't project any salient effect on ED performance. Moreover, prior research has shown neuroticism, as a factor reflects one's emotional stability, appears to have fairly weak relationship with job performance [88], [89]. Consequently, we mainly focus on other three personality traits when developing hypotheses in ED context. In Section 2.7, our results show the insignificant relationships between these two personality traits and clinical performance, which support this decision. We hypothesize that the other three personality traits have an impact on one or more of the physician performance measures, and elaborate them in the following subsections.

2.4.1 Openness

The personality trait of openness has been consistently associated with better job performance leveraging the outstanding problem-solving and innovation capabilities. The literature guides our thinking about how openness trait might influence medical performance in urgent treatment situations. Prior studies consistently report individuals who have a high level of openness to experience tend to be enthusiastic about participating in new educational activities [19], [90], be more likely to trust and use advanced technologies [91], [92], and tend to accept job-related technologies [93], [94]. Notably, emerging technologies have been observed improves clinical quality and time efficiency tremendously through electronic health record systems [1], telemedicine [2], robot-assisted surgery [3], and the Internet of Medical Things in the healthcare dimension [4]. Therefore, leveraging the usage of healthcare technologies, we conjecture that open-minded physicians tends to treat patients with higher medical quality with the shorter time spent.

Hypothesis 1A (H1A): Higher physician openness is associated with lower mortality rate for patient.

Hypothesis 1B (H1B): Higher physician openness is associated with a shorter LOS for a patient.

While emerging methodologies might be more expensive than traditional ones and tend to increase surgical cost, they generally have the potential to lower costs in other aspects, particularly post-surgery expenditures. Specially, open-minded physician implementing new medical technology would improve a patient's health outcome due to more efficient treatments, which leads to fewer unnecessary diagnostics in recovery and lower costs on corresponding lab tests. Moreover, as open-minded physicians are seen as intelligent, flexible, and risk-taking [95], [96], which especially be exposed as instinct when facing emergencies, we expect openness trait won't induce increased lab test prior surgery. Therefore, integrated with reducing expenditure amid recovery, we anticipate that openness will be associated with lower lab test expenditures.

Hypothesis 1C (H1C): Higher physician openness is associated with lower lab test expenditures for a patient.

2.4.2 Conscientiousness

Conscientiousness is a personality trait that indicates a person's inclination towards being accountable, methodical, and orderly. This reflects an inherent drive to act with prudence and structure [97], [98]. Browsing literature on job performance and personality, prior studies consistently reveal the conscientious people tend to outperform in both work and study [87], [99]–[101]. To further identify the mechanisms through which conscientiousness factor affects the work quality, Jackson et al.[102], Tough [103], and Roberts et al. [104] reveal the role of being organized, well-planned, and achieving goals following to-do lists of fundamental tasks as a crucial channel to mediate the conscientiousnessperformance relationship. Considering the nature of unexpected ED encounters could not allow ahead plan nor full preparation in detail, we conjecture the effect of conscientiousness on job performance that measured by patient mortality rate will diminish.

Hypothesis 2A (H2A): Physician conscientiousness is not significantly associated with mortality rate for a patient.

Motivated by the elementary instinct to be responsible and cautious, a conscientious individual tends to act conservatively in terms of spending more time and effort preparing and grasping thorough information before taking actions, especially when preparing for the critical steps [87], [95], [105]–[107]. In ED context, that would manifest as a physician tends to keep a patient staying longer in hospital for recovery and get a well comprehensive understanding and capture of a patient-side health condition through systematically laboratory tests before patient discharges, which cause the longer stay in hospital and higher lab test expenditures for a patient. Hence, we hypothesize:

Hypothesis 2B (H2B): Physician conscientiousness is associated with longer LOS for a patient.

Hypothesis 2C (H2C): Higher physician conscientiousness is associated with higher lab test charges for a patient.

2.4.3 Agreeableness

Agreeableness manifests itself in the form of individuals exhibiting kind, sympathetic, cooperative, and warm attitudes toward others [95]. Prior studies find while agreeableness is highly correlated to team performance, it is typically a weak predictors among others when forecasting one's work performance [108]. As study mainly cares about and captures the physician-level performance in econometrics specification, we follow the literature and hypothesize agreeableness won't see salient changes in physician clinical performance.

Hypothesis 3A (H3A): Physician Agreeableness is not significantly associated with mortality rate for a patient.

How agreeableness plays a role in influencing LOS is ambiguous as well. In rationale, driving by the nature of being kind and cooperate, an agreeable physician tends to understand and satisfy patients and colleagues. However, due to the variety of patients' and colleagues' preferences, it is reasonable to assume the preferences are distributed randomly and the effects are pretty much canceled out. Hence, we hypothesize:

Hypothesis 3B (H3B): Physician agreeableness is not significantly associated with LOS for a patient.

According to Cohen et al. [109], physicians do control their test-ordering behaviors when they take patients' expenditures into account. Because agreeable individuals are empathetic and easily understand others' perspectives [99], [110], such physicians appreciate the high cost of healthcare services for these patients and, in turn, try to reduce medical expenses if possible. Therefore, agreeable physicians are likely to consider patients' affordability when writing out test prescriptions, which results in lower lab test expenditures. Moreover, if agreeable physicians help patients save expenditures out of empathy, we expect to see more salient effect especially for patients in need, such as those with low income or those request for money-saving options. Generally, we propose this hypothesis:

Hypothesis 3C (H3C): Higher physician agreeableness is associated with lower lab test expenditures for a patient.

2.5 Data

We use the 2010-2018 Florida patient discharge data including the hospital discharge data and the ambulatory/outpatient data. Specifically, 75.6% of observations are obtained from the hospital discharge data and the remaining 24.4% comes from the ambulatory/outpatient data. Both of the datasets are maintained by the Florida Department of Health and provides detailed information on patient characteristics, diagnosis and procedure codes, in-hospital mortality rates, medical expenditures, and attending physicians' national provider identifiers (NPI). In addition to above treatment details, the discharge data also contains priority of admission and ED hour of arrival, which facilitate identifying patients who are sent to ED and need immediate healthcare treatment due to the urgent conditions. The second dataset we use is collected from a healthcare online review platform. It includes physicians' personal statements and corresponding characteristics, including name, working address, and other attributes such as NPI. We merge the two datasets at the individual physician level using Floridas physician license verification data that includes NPI, name, and working address.

2.5.1 Data Preparation

When studying the personality impacts on clinical outcomes, note it is important to account for non-random matching that can occur between patients and physicians. For instance, patients may prefer an empathetic physician. To alleviate those concerns, we consider data involving ED arrivals with accidental injuries. In emergency situations, attending physicians are prescheduled, and patients do not have sufficient time to select physicians. Such an institutional arrangement provides a quasi-random assignment between physicians and patients. We use priority of admission codes and the particular injury records in the data to identify those patients who experienced accidents and were sent to ED between 2010 and 2018. We also address other identification concerns in detail in Section 2.5.3.

We merge the FL discharged data with the online physician profiles at review platform and obtain 5,566 physicians with clinical performance records. Among them, 3,272 (58.8%) physicians served ED patients with accidental injuries during 2010-2018. Further, 2,073 physicians, accounting for 63.4% of physicians in the ED-injury sample, have posted their personal statements with complete information online. The remaining physicians either miss more than one demographic characteristics, or have little information presented in their online personal statements. To conclude, our final sample includes 98,701 ED patients treatment records of 2,073 physicians.

2.5.2 Variables Used for Analysis

We obtain the clinical performance measures from the Florida discharge data, and extract the personality traits from the physicians' personal statements that are posted at review platform.

The Florida discharge data provides patient information at the level of an individual hospital visit. It tracks the attributes associated with the visit such as the attending physician, nature of the visit (i.e., emergency or not), whether the patient was discharged successfully or died, LOS, and medical expenditure such as lab tests, among other factors. The data allows us to compute healthcare performance metrics for each physician.

As a proxy for medical quality, we use information about mortality, which is a binary variable that equals one if a patient died after the treatment and zero otherwise. Notably, such a measure is robust and commonly used in the health management studies [7], [84], [85].

We capture treatment efficiency using LOS for inpatient cases[3]. LOS refers to the number of days a patient stays in a hospital for each visit. The average length of stay for a patient in the emergency department is about 3.4 days. This is a crucial indicator used to assess the effectiveness of a healthcare provider as well as the efficiency of hospital management [2], [111].

The medical expenditure is measured by lab test charges recorded in the patient discharge data. As mentioned earlier, it is directly available from the patient discharge data. In our analysis, we adopt lab test charges at the patient level [112].

Physicians' personality traits are extracted from their personal statements posted on review platform. Based on the physician statements, we conduct a analysis through the LIWC dictionary. The processing method and the introduction of the LIWC dictionary are detailed in Appendix 2.10.2. In short, for each personality trait, we sort the scores of all physicians in the sample in an ascending order and transform these personality scores using

[3]↑We study LOS only in inpatient cases and exclude the outpatient ones due to the difference in definition of LOS. Specifically, inpatient LOS measures the number of days elapsed from the admission date to the discharge date from hospital. Instead, outpatient LOS represents the number of days between visit beginning date and visit ending date, which covers the whole visit period and usually not be counted as efficiency factor for treatment.

percentiles[25]. Such transformation of the personality scores facilitate us to interpret the results in an understandable perspective. We also present the robustness results using the original scores, rather than transformed percentiles, in Appendix 2.10.6.

2.5.3 Issues on Selection Bias

There are two potential selection bias issues in our setting, which we explain below. We also discuss how we overcome these selection issues.

Our main concern is the aforementioned two-way selection between physicians and patients. For instance, a patient may use a physician's background, statement, and other relevant information in an online rating platform to select a physician based on her preference. Similarly, a physician can also shun away from a patient for various reasons [113]. Therefore, the randomness assumption of the matching between patients and physicians is violated. To address this two-way selection concern, we only consider the cases of ED arrivals with accidental injuries filtered by external causes of injury codes[4]. By doing this, we avoid accidental cases that may leave enough time for patients to select physicians.

The assignment of patients to physicians for ED arrivals with accidental injuries is quasi-random. From the patient's perspective, the U.S. Emergency Medical Treatment and Active Labor Act mandates that an ambulance must transport a patient with acute injury or illness to the closest hospital equipped to provide the necessary medical treatment, both for legal and clinical reasons [114]. Additionally, patients who sustain accidental injuries, particularly those with high mortality rates, are often in critical condition and have limited opportunities to research physician information online and select a doctor for treatment.

From the physician's perspective, the allocation of a seriously injured patient to an available physician largely depends on the prearranged schedules of physicians, taking into account both their workdays and shifts, as well as the timing of the accident occurrence. Because physicians' schedules are fixed apriori without any knowledge of the accident, it is reasonable to assume that the matching between healthcare providers and patients under the urgent situation is random.

[4]↑The external cause of injury codes include car accidents, gun shots, burns, and poisoning, among others, which can be identified using the ICD coding systems.

Table 2.3. Descriptive Statistics of Variables

Variable	No. of obs.	Mean	SD	Definition
Treatment procedure measure (at the treatment level)				
Mortality	98,701	0.012	0.107	Equals 1 if a patient died before being discharged
Lab charges	98,701	5504.27	13232.46	Charges for the laboratory tests
Length of stay	65,007	4.860	8.532	Number of days from admission to discharge
Waiting time	31,209	2.227	4.399	Number of days from admission to principal procedure
Physician characteristics (at the physician level)				
Openness	2,073	0.492	0.289	Level of openness in the personality of a physician
Conscientiousness	2,073	0.497	0.290	Level of conscientiousness of a physician
Extraversion	2,073	0.516	0.289	Level of extraversion in the personality of a physician
Agreeableness	2,073	0.513	0.289	Level of agreeableness in the personality of a physician
Neuroticism	2,073	0.511	0.290	Level of neuroticism in the personality of a physician
Age	2,073	54.96	9.48	Age of a physician
Female	2,073	0.201	0.401	Equals 1 if a physician is female
Education rank	2,073	127.75	33.07	Rank of the physician graduated medical school
Experience	2,073	24.59	9.86	Number of years since a physician graduated
Rating	2,073	4.003	0.699	Online rating of a physician
Number of review	2,073	23.61	15.81	Total number of reviews of a physician
Patient characteristics (at the patient level)				
Charlindex	98,701	0.825	1.463	Risk index measured by Charlson Comorbidity Index
Female	98,701	0.500	0.500	Equals 1 if a patient is female
White	98,701	0.836	0.371	Equals 1 if a patient is white
Black	98,701	0.088	0.284	Equals 1 if a patient is black
Other race	98,701	0.076	0.265	Equals 1 if a patient is neither white nor black
Age	98,701	58.29	24.41	Age of a patient
Medicare	98,701	0.498	0.500	Equals 1 if a patient is covered by Medicare
Medicaid	98,701	0.082	0.275	Equals 1 if a patient is covered by Medicaid
Private insurance	98,701	0.283	0.451	Equals 1 if a patient is covered by private insurance
Other insurance	98,701	0.137	0.343	Equals 1 if the patient is not covered by above
Income	98,701	53,805	17,842	Median household income at the zip code level
Hospital characteristics (at the hospital level)				
Beds	180	356.41	308.41	The number of beds in a hospital
Highway fatality	180	12.84	4.09	Motor vehicle crash death rate per 100,000

One may argue that patients could select locations based on their financial conditions, preferences, and other personal reasons. Also, their injury types might be highly correlated with where they live. For example, patients living in a region with high criminal rates may be more likely to incur gunshot injuries than those in relatively safer regions. As another example, patients living near highways might be more likely to encounter car accidents than those who do not drive on highways often. Such patients' location selection may lead to the non-random matching between patients and hospitals where affiliated physicians serve. To alleviate this concern, we only retain patients who live in a location different from the hospital's Health Service Areas (HSA).[5] This procedure allows us to focus on those cases with patients who live in one location but become injured in another location. Such a setting satisfies the assumption of random assignment.

After accounting for the aforementioned issues, our final sample used for analyses includes 2,073 physicians and 98,701 ED patient visits between 2010 and 2018. The summary statistics of all relevant variables are reported in Table 4.1.

2.6 Empirical Methods

The institutional design of ED arrivals with accidental injuries helps to alleviate concerns on the non-random matching between patients and physicians with different personality traits. In this section, we first use this quasi-random setting to propose a baseline specification. We then add an error correction model to correct for the measurement errors of the personality traits scores.

2.6.1 Econometric Model Identification

To examine the causal impacts of physician personality traits, we start with the following baseline specification:

$$Y_{\mathrm{ij}kt} = \beta_0 + \sum_{m=1}^{5} \beta_m S_{\mathrm{j}}^{m} + \beta_6 V_{\mathrm{it}} + \beta_7 P_{\mathrm{jt}} + \beta_8 H_{kt} + \beta_k + \beta_t + \beta_{\mathrm{j}} + \epsilon_{\mathrm{ij}kt}, \tag{2.1}$$

[5]↑HSAs were defined by the National Center for Health Statistics, a part of the Centers for Disease Control and Prevention, to represent a single county or cluster of contiguous counties that are relatively self-contained with respect to hospital care.

in which the subscripts of $Y_{\mathrm{ij}kt}$ correspond to patient i who received treatment from physician j in hospital k at time t. The dependent variable $Y_{\mathrm{ij}kt}$ is in-hospital mortality, lab test charges, or LOS, depending on the model. Physician j's Big Five personality scores are S_{j}^{m} in which $m = 1, \ldots, 5$ correspond to the individual measures of extraversion, neuroticism, agreeableness, conscientiousness, and openness; $V_{\mathrm{i}t}$ is a vector of patient characteristics including patient gender, age, race, diagnosis risk index, insurance payer type, and median household income; $P_{\mathrm{j}t}$ is a vector of observable physician characteristics including physician gender, age, educational background, work experience, online ratings, and number of reviews; and H_{kt} is the vector of hospital characteristics, including beds and nearby highway fatalities. We include hospital fixed effects β_k, time fixed effects β_t, and medical specialty fixed effects β_{j} in the specification. Because physician personality traits are time-invariant and stable across patients, controlling for physician fixed effects would absorb their personality trait measures. Instead, we control for unobserved physician characteristics which includes physician random effects. $\epsilon_{\mathrm{ij}kt}$ is the error term. We also cluster the standard deviations on the physician level.

Although the physician fixed effects are not applicable here, we argue that is not problematic in our analysis in two perspectives. First, leveraging the quasi-random research design, the effects of unobserved physician characteristics exerting on patients health outcomes follow independently identical distribution as random errors, which only bias the estimates if correlated to variables of interest. Since physician personality traits are inherent and barely changed, it is reasonable to assume personality is isolated from other unobserved features. Moreover, the control for physicians age, education, online rating, work experience, medical specialty, and especially, physician random effects, could further alleviate the concerns of unobserved physician heterogeneity.

2.6.2 Measurement Error Correction

Measuring personality traits from the language used is likely to be subject to measurement errors. Most prior studies using personality traits have used written or speech samples from experiment participants but have been done so in specific situations. For example, Fast and

Funder [115] use language collected from one-hour interviews of one's personal history; Hirsh and Peterson [116] use data from interviewees recalling their past and their plans for the future; and Baddeley and Singer [117] use daily diaries, writing assignments, and journal abstracts, as well as recent bereavement narratives. While the language that people use can reveal their social and psychological styles, there is still ambiguity regarding whether words generated in specific situations or regarding particular topics are robust enough to extend to other cases [118].

There are several concerns about the reliability of those online personal statements. One may argue that the review platform imposes restrictions on topics for physicians to post, and these topic-restricted self-expressions fail to comprehensively represent one's personalities. Some personal statements may be posted by administrative staff instead of physicians themselves. Additionally, the number of detected words per physician in our study is relatively small, which could further exacerbate any measurement errors. Thus, we are aware of these constraints as we interpret performance from personality traits inferred from language used in personality traits. We also attempt to address these two issues by taking advantage of repeated measurements from two different statement topics, as described in the following.

Generally, measurement errors can be corrected by using validation samples (in which the true scores of personality traits are observed), or by repeated measurements (in which the real personality scores are unobserved but are detected more than once). Although validation sample is an ideal choice, it is not feasible in our study to collect the true personality of a physician. Therefore, we take advantage of repeated measurements from different kinds of statements and correct the measurement errors. The use of additional information to correct measurement errors is commonly used in the econometric literature [119]. In our case, we treat information from a physician' biographical sketches as well as her care philosophy as two different statements for our analysis.

Consider the condition that we have 5 error-prone variables to explore their impacts. Suppose the vector of the true personality scores for physician j is $X_{\mathrm{j}} = (X_{\mathrm{j}}^1, X_{\mathrm{j}}^2, ..., X_{\mathrm{j}}^5)'$. The vector of detected personality scores is $S_{\mathrm{j}l} = (S_{\mathrm{j}l}^1, S_{\mathrm{j}l}^2, ..., S_{\mathrm{j}l}^5)'$, in which $l = 1, 2$ corresponds to the manner in which those personality traits were measured. Thus, $S_{\mathrm{j}1}$ is obtained from physician statement 1 (the biographical sketch) and $S_{\mathrm{j}2}$ from statement 2 (the care philosophy). Under the classical measurement error model, measurements (personality scores) of personality traits can be described as:

$$S_{\mathrm{j}l}^{m} = X_{\mathrm{j}}^{m} + \epsilon_{\mathrm{j}l}^{m}, \qquad l = 1, 2;\ m = 1, 2, ..., 5, \tag{2.2}$$

in which the $\epsilon_{\mathrm{j}l}^{m}$ are error terms with mean 0, variance σ_m^2, and are independent of each other and of X_{j}^{m}, $S_{\mathrm{j}l}^{m}$, and dependent variable $Y_{\mathrm{ij}kt}$.

Note that our goal is to estimate the corrected parameters $\beta = (\beta_1, \beta_2, ..., \beta_5)'$ in the linear model shown in Equation 2.1. To achieve that, we perform the correction by applying regression calibration, which is the most commonly used approach to fix measurement errors. According to Rosner et al. [120], if we let Z_{j} denote the physician-level control variables included in Equation 2.1, then the corrected β_m can be estimated by using $E[X_{\mathrm{j}}^{m} \mid S_{\mathrm{j}1}, Z_{\mathrm{j}}]$ in place of S_{j}^{m}. Thus, in the classical measurement error model presented in Equation 2.2, the correction factor λ can be computed through a linear estimation of $S_{\mathrm{j}2}$ on $S_{\mathrm{j}1}$ and other variables used for control[6], which are measured without errors, and we can achieve the expectations leveraging the linear regression models:

$$S_{\mathrm{j}2}^{m} = \beta_0^{m} + \sum_{m=1}^{5} \lambda_m S_{\mathrm{j}1}^{m} + \phi^{m} Z_{\mathrm{j}}^{m} + \epsilon_{\mathrm{j}m}, \qquad m = 1, 2, ..., 5 \tag{2.3}$$

We can then easily apply correction factor λ_m to $\beta_m^{*} = \beta_m / \lambda_m$ for each $m = 1, 2, ..., 5$ and obtain unbiased β_m^{*} after measurement error corrections. In the rest of the paper, our results are corrected for measurement errors. And the estimation results using uncorrected personality scores are also listed as robustness checks in Section 2.7.2.

[6]↑Some physicians posted their biographies only and did not provide information on care philosophy online. Hence, we cannot apply the measurement error structure and have to rely on the uncorrected scores for these cases. In the robustness checks, we provide the results using scores without error correction or excluding these cases without care philosophy.

2.7 Empirical Results

This section reports our estimation results and their corresponding robustness checks. Measurement errors have been corrected in all the reported results. The results are robust to alternative specifications, measures, and models.

2.7.1 Main Results for Clinical Performance

Columns (1)-(2) in Table 2.4 present the effect of personality traits on in-hospital mortality. We include physician characteristics, patient characteristics, hospital characteristics, and hospital fixed effects in all our regressions and correct for measurement errors. Columns (1) shows results with controlling for physician random effects. Column (2) shows our results when we estimate a sample that excludes those physician statements that contain fewer than 150 strings, in addition to adding physician random effects. These two specifications shown in Columns (1)-(2) will be applied in all of the remaining tests without additional clarification.

The results in Columns (1)-(2) of Table 2.4 show that openness is negatively associated with patient mortality at the five percent significance level. Translating the coefficient into magnitude, a 10-percentile increase in physician openness rank induces a 4.2% reduction in patient in-hospital mortality. The results are robust to different specifications, which supports H1A.

Columns (3)-(4) of Table 2.4 presents the impact of physician personality traits on LOS. The results show that, all else being equal, patients being treated by physicians with higher openness scores experience shorter hospital stays. A 10-percentile increase in physician openness rank results in a 4.1% decrease in LOS. These results support our hypotheses H1B.

Columns (5)-(6) of Table 2.4 present the effect of personality traits on medical expenditures, measured by laboratory test charges. The findings indicate that patients treated by physicians who score high on agreeableness or openness tend to have lower costs for laboratory tests. By contrast, those being treated by physicians with high conscientiousness scores are positively correlated with their lab test charges. These results support our hypotheses H1C, H2 and H3.

Table 2.4. The Effect of Personality Traits on Physician Clinical Performance

	In-Hospital Mortality		LOS		Lab Test Expenditure	
	Random Effect	More than 150 strings	Random Effect	More than 150 strings	Random Effect	More than 150 strings
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Openness	-0.005**	-0.005*	-1.399***	-1.684***	-0.358***	-0.424***
	(0.003)	(0.003)	(0.406)	(0.430)	(0.085)	(0.088)
Conscientiousness	0.001	0.001	0.592*	0.679*	0.189***	0.223***
	(0.002)	(0.002)	(0.327)	(0.346)	(0.063)	(0.066)
Extraversion	0.003	0.002	-0.366	-0.835	-0.001	-0.073
	(0.003)	(0.003)	(0.485)	(0.516)	(0.078)	(0.083)
Agreeableness	-0.002	-0.002	-0.421	0.007	-0.200***	-0.167**
	(0.002)	(0.003)	(0.369)	(0.334)	(0.076)	(0.081)
Neuroticism	-0.002	-0.003	-0.174	-0.266	-0.055	-0.052
	(0.002)	(0.002)	(0.243)	(0.248)	(0.056)	(0.057)
Hospital Characteristics	Y	Y	Y	Y	Y	Y
Patient Characteristics	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y
Hospital Fixed Effect	Y	Y	Y	Y	Y	Y
Physician Random Effect	Y	Y	Y	Y	Y	Y
Observations	98,701	91,094	98,701	91,094	98,701	91,094

Standard errors in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Translating the coefficients into magnitude, a 10-percentile increase in the agreeable or openness ranks is associated with an average of 3.0% or 4.2% reductions in lab test charges, respectively. On the contrary, a 10-percentile increase in conscientiousness score ranks leads to an average increase of 3.1% in lab test charges. Our results suggest that agreeable and open-minded physicians tend to incur fewer lab test costs while conscientious physicians are likely to prescribe more lab tests for diagnoses and treatments.

In summary, the results in Table 2.4 show the impact of physicians' personal traits on quality, cost, and efficiency. Our analysis reveals that patients treated by conscientious physicians tend to have higher expenses on laboratory tests. In contrast, physicians with higher scores on openness are linked with lower rates of in-hospital mortality, reduced charges for laboratory tests, and shorter length of hospital stay. Moreover, being treated by agreeable physicians is connected to lower medical costs in terms of laboratory test expenditure.

2.7.2 Robustness Checks

To examine the robustness of our analysis, we replicate our analysis under different specifications. In the previous tables, we have presented robustness checks using different length of words for NLP. In this subsection, we do additional robustness tests and provide corresponding results.

First, to alleviate the concern of personality detection accuracy using small text, we conduct the robustness checks for our main estimation applying the set of personality measures generated by Word Embedding with Gaussian Process method [23]. According to Arnoux et al. [23], the proposed method is proved to largely reduce the requirement of text size, with about eight times fewer data, for the same accuracy of personality detection. Thus, leveraging the Word Embedding with Gaussian Process method in personality detection helps us to overcome the limitation of small size of text in physician personal statements and provides us with a set of personality measures, which is expected to have higher accuracy. Table 2.5 shows the results are robust to different specifications under this set of personality measures generated by Arnoux et al. [23] method, which strongly supports our hypotheses and, in the meantime, validates the LIWC-based personality measures detected by physician statements.

Second, note that some physicians, accounting for 45.9% of 2,073 physicians, choose to only post biography without posting care philosophy, which prevent us from correcting the measurement errors for these physicians. To alleviate the potential concerns that only part of sample using measurement error corrected personality measures, we provide two set of robustness checks. In the first test, we estimate the sample on Equation 2.1 without correcting measurement errors in personality scores. In the second test, we exclude these physicians who post biography only and limit the sample to those physicians who posted both biography and care philosophy. Note that sample size in each robustness checks could be different, thus, the percentile personality scores for physicians may not be same and need

Table 2.5. Hypotheses Test Using WEGP Detected Scores

	In-Hospital Mortality		LOS		Lab Test Expenditure	
VARIABLES	Random Effect (1)	More than 150 strings (2)	Random Effect (3)	More than 150 strings (4)	Random Effect (5)	More than 150 strings (6)
Openness	-0.000236**	-0.000238**	-0.0674***	-0.0720***	-0.0156***	-0.0174***
	(0.000102)	(0.000107)	(0.0140)	(0.0142)	(0.00325)	(0.00333)
Conscientiousness	4.41e-05	4.12e-05	0.0207**	0.0245**	0.00535***	0.00633***
	(6.57e-05)	(6.94e-05)	(0.00969)	(0.00989)	(0.00189)	(0.00195)
Extraversion	0.000147	0.000112	-0.00893	-0.0243	0.00270	0.000994
	(0.000111)	(0.000127)	(0.0155)	(0.0169)	(0.00313)	(0.00327)
Agreeableness	-4.96e-05	-2.25e-05	-0.0117	0.000811	-0.00606**	-0.00509*
	(8.22e-05)	(9.46e-05)	(0.0130)	(0.0126)	(0.00254)	(0.00272)
Neuroticism	-0.000120	-0.000129	0.000659	0.00192	-0.00172	-0.000743
	(8.79e-05)	(9.45e-05)	(0.0117)	(0.0125)	(0.00294)	(0.00303)
Hospital Characteristics	Y	Y	Y	Y	Y	Y
Patient Characteristics	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y
Hospital Fixed Effect	Y	Y	Y	Y	Y	Y
Physician Random Effect	Y	Y	Y	Y	Y	Y
Observations	98,701	91,094	98,701	91,094	98,701	91,094

Standard errors in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

to be reassigned in each test. Specifically, for each personality trait under each test, we sort the scores in an ascending order and transform these personality scores using percentiles as was originally done in Adamopoulos et al. [25]. The results shown in Table 2.6 are robust.

Furthermore, in our main sample, about three quarters of observations are obtained from the inpatient data from Florida dataset and one quarter from the ambulatory/outpatient data. We concern that the two datasets may calibrate the LOS in different ways. To relieve the concerns due to the data generating process, we limit our estimation to the inpatient sample and present the robust results in the first three columns of Table 2.7. Besides, we create a dummy variable which equals 1 if an observation is recorded in the inpatient data and 0 otherwise, and include it in Equation 2.1. The results, shown in the second three columns of Table 2.7, are robust. More robustness checks are conducted and the corresponding results are included in Appendix 2.10.6.

Table 2.6. Robustness Check: Concern for Physicians Posting Single Statement

	No Measurement Error Correction			Physicians with Both Statements		
	Mortality	LOS	Expenditure	Mortality	LOS	Expenditure
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Openness	-0.005*	-1.244***	-0.348***	-0.008*	-1.762***	-0.424***
	(0.002)	(0.372)	(0.078)	(0.004)	(0.460)	(0.136)
Conscientiousness	0.001	0.426	0.179***	0.002	0.878**	0.381***
	(0.002)	(0.295)	(0.058)	(0.003)	(0.370)	(0.093)
Extraversion	0.002	-0.427	-0.004	0.0027	-0.661	0.170
	(0.003)	(0.415)	(0.071)	(0.004)	(0.453)	(0.145)
Agreeableness	-0.001	-0.275	-0.190***	-0.0018	-0.372	-0.366***
	(0.002)	(0.337)	(0.072)	(0.004)	(0.454)	(0.131)
Neuroticism	-0.002	-0.187	-0.047	-0.004	-0.367	-0.070
	(0.002)	(0.233)	(0.054)	(0.002)	(0.316)	(0.087)
Hospital Characteristics	Y	Y	Y	Y	Y	Y
Patient Characteristics	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y
Hospital Fixed Effect	Y	Y	Y	Y	Y	Y
Physician Random Effect	Y	Y	Y	Y	Y	Y
Observations	98,701	98,701	98,701	43,555	43,555	43,555

Standard errors in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

2.8 Mechanism Discussion and Heterogeneous Analysis

Section 2.4 hypothesized how the physician personal traits affect clinical performances. The empirical analyses in the previous section evaluated those hypotheses. In this section, we further investigate the possible mechanism that explains the impact of the personality traits on clinical performances. In doing so, we seek to develop a deeper understanding of the mechanism.

2.8.1 Openness Impact on New Technology Adoption

The personality literature suggests that open-minded physicians are more likely to learn and apply new methodologies and advanced technology [91], [92], [121]. We therefore test

Table 2.7. Robustness Check: Control for Inpatient Cases

	Inpatient Cases Only			With Inpatient Dummy		
	Mortality	LOS	Expenditure	Mortality	LOS	Expenditure
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Openness	-0.006*	-0.887**	-0.389***	-0.005*	-1.180***	-0.312***
	(0.004)	(0.396)	(0.090)	(0.003)	(0.399)	(0.079)
Conscientiousness	0.001	0.440	0.204**	0.001	0.488	0.178***
	(0.003)	(0.316)	(0.067)	(0.002)	(0.310)	(0.058)
Extraversion	0.002	0.082	-0.047	0.002	-0.354	0.004
	(0.004)	(0.411)	(0.084)	(0.003)	(0.468)	(0.073)
Agreeableness	-0.003	-0.541	-0.196**	-0.00180	-0.322	-0.187***
	(0.003)	(0.356)	(0.080)	(0.002)	(0.337)	(0.071)
Neuroticism	-0.003	-0.067	-0.060	-0.002	-0.136	-0.052
	(0.003)	(0.243)	(0.059)	(0.002)	(0.228)	(0.052)
Hospital Characteristics	Y	Y	Y	Y	Y	Y
Patient Characteristics	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y
Hospital Fixed Effect	Y	Y	Y	Y	Y	Y
Physician Random Effect	Y	Y	Y	Y	Y	Y
Observations	65,007	65,007	65,007	98,701	98,701	98,701

Standard errors in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

this assumption that the innovative nature of open-minded physicians makes them quick learners and thus high performers.

To verify this underlying assumption, we examine whether physicians with high openness scores are more likely to adopt new treatment techniques. To study the new methodology adoption, we do not restrict our sample to be injury cases in EDs. By matching Florida inpatient data ranging from 2015, quarter 4 to 2018, quarter 4 with physician data,[7] we obtain a merged data set containing 1,125,366 treatment-level observations of 4,041 physicians. For robustness, we also report the results using the physicians who encountered with the ED patients with accidental injuries. We then identify those treatment procedures that

[7]↑The CMS new medical methodology section has been available since the 4^{th} quarter of 2015 when the ICD-10 codes were introduced.

are classified as the New Methodology[8] according to the definition of ICD-10-PCS New Technology Section Codes.

We use a survival model as the main statistical approach to analyze the data and verify this assumption. We define the first time when a physician j implements a new treatment procedure i as the time of adoption and denote it as T_{ij}. If technology i is adopted by physician j during the sample period, T_{ij} equals the relative time in quarters of new technology i adopted by physician j since that technology is released; if not, T_{ij} equal the end of the observed period. The function for the hazard rate is listed as follows:

$$h_{\text{ij}}(t) = \lim_{dt \to 0} \frac{Pr(t \leq T_{\text{ij}} < t + dt \mid t \leq T_{\text{ij}})}{dt} \tag{2.4}$$

We use the Cox model and assume:

$$h_{\text{ij}}(t) = h_{\text{i0}}(t)\text{e}xp\left(\sum_{m=1}^{5} \beta_m S_{\text{j}}^{m} + \beta_6 P_{\text{j}}\right) \tag{2.5}$$

in which, S_{j}^{m} corresponds to the Big Five personality scores with measurement errors corrected. P_{j} denote physician characteristics including gender, physician specialty, education rank, working experience, online rating, and number of reviews. Also, $h_{\text{i0}}(t)$ is the baseline hazard function of new technology i. It is noteworthy that the personality scores are transformed by percentiles in a sample and hence could vary across different samples. In the robustness checks, we provide the results without such variable transformation.

The estimated results are listed in Table 2.8. Columns (1) and (2) show the results using the sample of all physicians and a subsample of physicians who treated ED patients with accidental injuries only, respectively. These results show that high openness scores are associated with the early adoption of new treatment procedures while other personality traits are not. To alleviate the concern of low adoption rates of new technology, we conduct the survival model using the new procedure with the highest adoption rate in Columns (3)-(4) and the procedures with the top 5 adoption rates[9] in Columns (5)-(6). These results are

[8]↑https://www.cms.gov/Medicare/Coding/ICD10/2016-ICD-10-PCS-and-GEMs

[9]↑Including (1) extirpation of matter from coronary artery via percutaneous approach; (2) replacement of aortic valve using zooplastic tissue via open approach; (3) replacement of aortic valve using zooplastic tissue via open approach via percutaneous approach; (4) replacement of skin using porcine liver derived skin

Table 2.8. The Effect of Physician Personality Traits on Technology Adoption

	Full Sample		Top 1 Tech		Top 5 Tech	
	All Phy	ED Phy	All Phy	ED Phy	All Phy	ED Phy
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Openness	1.602**	1.496**	3.481**	2.966**	1.509**	1.267*
	(0.676)	(0.723)	(1.374)	(1.431)	(0.741)	(0.746)
Conscientiousness	-0.619	-0.616	-1.177	-2.329**	-0.451	-0.694
	(0.475)	(0.467)	(0.904)	(1.032)	(0.551)	(0.545)
Extraversion	1.130*	0.980	1.496	0.900	0.524	0.292
	(0.582)	(0.635)	(0.993)	(1.061)	(0.638)	(0.659)
Agreeableness	-0.081	0.000	2.291	1.886	1.129	0.964
	(0.600)	(0.616)	(1.438)	(1.465)	(0.763)	(0.756)
Neuroticism	0.147	0.330	-0.655	-0.060	-0.156	0.047
	(0.440)	(0.467)	(0.875)	(0.944)	(0.469)	(0.484)
Failures	97	89	22	18	60	56
Observations	37,840	31,152	1,720	1,416	8,600	7,080

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

consistent across alternative samples. The evidence endorses the conjecture that physicians with higher openness scores are associated with better clinical performance because they are open to new treatment methods and could take advantage of advanced technologies.

2.8.2 Conscientiousness Impact on Time Usage Preparing for Principle Procedure

Conscientious physicians are very cautious and prefer to pay efforts in investigating the comprehensive information of the problem before making diagnoses as well as applying treatments. As a result, they tend to conduct extra lab tests. Following the same logic, we expect that patients served by a conscientious physician have to wait longer to receive a main treatment procedure. To conduct this test, we define waiting time as the number of days elapsed

substitute via external approach; (5) fusion of 2 or more cervical vertebral joints using nanotextured surface interbody fusion device via open approach.

Table 2.9. The Effect of Physician Personality Traits on Waiting Time

	Waiting Time		
	No Random Effect	Random Effect	More than 150 Strings
VARIABLES	(1)	(2)	(3)
Openness	0.0346	0.0346	-0.106
	(0.199)	(0.199)	(0.222)
Conscientiousness	0.590***	0.590***	0.617***
	(0.199)	(0.199)	(0.212)
Extraversion	0.316	0.316	0.215
	(0.209)	(0.209)	(0.240)
Agreeableness	-0.010	-0.010	0.021
	(0.182)	(0.182)	(0.205)
Neuroticism	-0.046	-0.046	-0.066
	(0.128)	(0.128)	(0.140)
Hospital Characteristics	Y	Y	Y
Patient Characteristics	Y	Y	Y
Physician Characteristics	Y	Y	Y
Hospital Fixed Effect	Y	Y	Y
Physician Random Effect	N	Y	Y
Observations	31,205	31,205	27,714

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

from the admission date to the date that a patient undergoes a principal procedure.[10] This measure is defined to reflect how a physician arranges the sequence of procedures.

Table 2.9 presents the estimates using Equation 2.1-2.3 in which the dependent variable is waiting time. The results indicate that being treated by a conscientious physician seems to have longer waiting times. Specifically, a one-standard-deviation increase in conscientious percentile leads to 7.55% or a 0.2 day spent in waiting to receive principle procedures. This evidence combined with the result of the extra expenditure in lab tests shown on Table 2.4 suggests that conscientious physicians make efforts to thoroughly understand a patient's condition before principal treatment, which results in more lab test cost and longer waiting time.

[10]↑This is the procedure performed for definitive treatments, rather than for diagnostic or exploratory purposes, or that is necessary to take care of a complication. In our main sample, only 31,209 ED patients took the definitive treatments.

2.8.3 Agreeableness Differential Impact on Expenditure for Patients in Need

The aforementioned evidence shows that being treated by an agreeable physician is associated with low charges on lab tests. The personality literature suggests that agreeable individuals tend to put themselves in others' positions [122]. Applying this argument to the healthcare setting, we conjecture that highly agreeable physicians tend to help low-income objectives reduce unnecessary laboratory tests and save medical expenditure because such patients are more sensitive to medical costs than those with higher incomes.

Table 2.10. The Effect of Personality Traits on Expenditures for Low Income Group

	Lab Test Expenditure		
VARIABLES	No RE (1)	Random Effect (2)	More than 150 Strings (3)
Openness	-0.312***	-0.337***	-0.409***
	(0.101)	(0.087)	(0.091)
Conscientiousness	0.176**	0.189***	0.228***
	(0.075)	(0.065)	(0.068)
Extraversion	0.014	-0.030	-0.106
	(0.095)	(0.083)	(0.089)
Agreeableness	-0.187*	-0.144*	-0.103
	(0.098)	(0.082)	(0.088)
Neuroticism	-0.080	-0.065	-0.063
	(0.067)	(0.059)	(0.060)
Agreeableness*LowIncome	-0.151**	-0.119**	-0.135**
	(0.072)	(0.053)	(0.059)
Ag+Ag*LowIncome	-0.338***	-0.263***	-0.238**
Prob >F (Ag+Ag*LowIncome)	0.0061	0.0061	0.0280
Openness*LowIncome	Y	Y	Y
Conscientiousness*LowIncome	Y	Y	Y
Extraversion*LowIncome	Y	Y	Y
Neuroticism*LowIncome	Y	Y	Y
Hospital Characteristics	Y	Y	Y
Patient Characteristics	Y	Y	Y
Physician Characteristics	Y	Y	Y
Hospital Fixed Effect	Y	Y	Y
Physician Random Effect	N	Y	Y
Observations	98,701	98,701	91,094

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

To verify this conjecture, we divide our sample into two groups using the cutoff of the mean patient income at the zipcode level. We then examine the differential effect of agreeableness on lab test expenditure by income group. The underlying rationale for this test is that highly agreeable physicians are inclined to do favors for low-income patients who are concerned about expenditures and affordability. Table 2.10 shows that agreeable physicians with higher agreeableness scores are associated with lower lab test charges and that such effects are more pronounced for patients in the low income group.

2.9 Conclusion

This study examines the impact of physicians personality features on their medical behaviors and service performance. We apply NLP techniques to physicians personal statements from an online review platform and extract 2,073 physicians' personality traits. To overcome endogeneity issues, we take advantage of the institutional design of ED arrivals to achieve the random assignments between patients and physicians and adopt a structural model to correct the measurement errors of the Big Five personality traits.

Using Florida discharge data from 2010 to 2018, we find that physicians with high openness scores tend to have patients with lower mortality rates, lower charges for lab tests, and shorter LOS. Patients being treated by conscientious physicians tend to have larger expenditure on lab tests. By contrast, patients with agreeable physicians tend to have lower medical expenditure on lab tests. Our findings indicate that the personality traits of physicians do have an impact on their clinical performance. To be emphasized, we also uncover possible mechanisms by which these traits influence the health outcomes. We find that physicians with openness trait tend to earlier learn and adopt new treatment techniques, ones with conscientiousness make the patients wait longer for undergoing principle procedures, and those with agreeableness traits tend to help low-income patients. All these findings are consistent with the personality features captured by Big Five Model and thus validate the personality measures generated by this NPL method.

Our analysis of physicians personality impacts on clinical performance yields several important insights for hospital administrators. Hospitals can improve patient outcomes by taking physician personality traits into consideration under critical situations like during the pandemic. For example, the COVID-19 pandemic in recent years caused in severe physician backlogs in hospitals [123]. Since hospitals cannot easily increase the number of physicians over a short time span, one of the hospital systems that talked to us was interested in taking the personality traits of the existing physician into scheduling consideration and help to improve clinical performance and efficiency.

More importantly, hospitals can implement our NLP method with respect to physicians personal statements, so they may inform physicians about their own personalities, train them to promote strength and avoid weakness in their personalities and minimize the mismatch between patients and physicians. Doing so can eventually reduce mortality rates, lower medical expenditures, and benefit the public.

So far as we know, there are few papers which use the online personal statements to extract personality traits of individual physicians. Hence, few physicians have incentives to fake on their online personal statements for unknown benefits and uncertain goals. By contrast, our text analysis approach overcomes the shortcomings in the traditional methods using surveys or questionnaires to know about personality traits of individuals. First, individuals could lie or hide information when answering survey questions. Second, it is costly to implement a traditional method using surveys or questionnaires and the response rates to these traditional methods are very low. Our method allows us to obtain physicians' personalities at a large scale across many hospitals with low cost.

This study is not without limitations. First, the measures of personality traits are time invariant for individual physicians, which yields the difficulties in controlling for physician fixed effects. We cannot account for certain unobserved characteristics of physicians. Nevertheless, the quasi-random assignment of patients to different physicians may greatly alleviate this concern. Further, we also control for physicians age, education, rating, experience, physician specialty, and their random effects. Second, the measurement errors of personality traits are inevitable given the nature of text analysis [124]. Future work can assess physician personalities with even more sophisticated techniques in order to update the linguistic-personality dictionary and improve accuracy in detecting personalities.

2.10 Appendix

2.10.1 Physician Personal Statement

In the following figures, we present the interface of a physician's personal website. We first collect the basic information of physicians, such as name, specialty, gender, age, working address(es), and phone number. Figure 2.3 is an example. By clicking the six buttons featured in Figure 2.3, a user can easily access any corresponding section, such as "Reviews," "About Me," and "Locations."

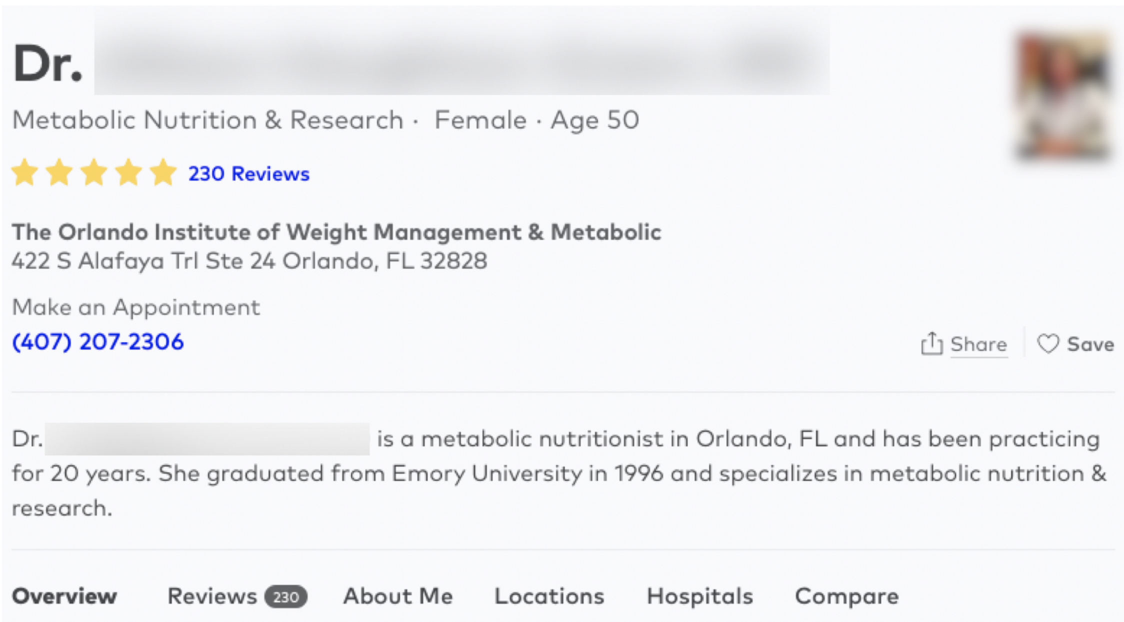


Figure 2.3. Interface of Physician's Personal Website

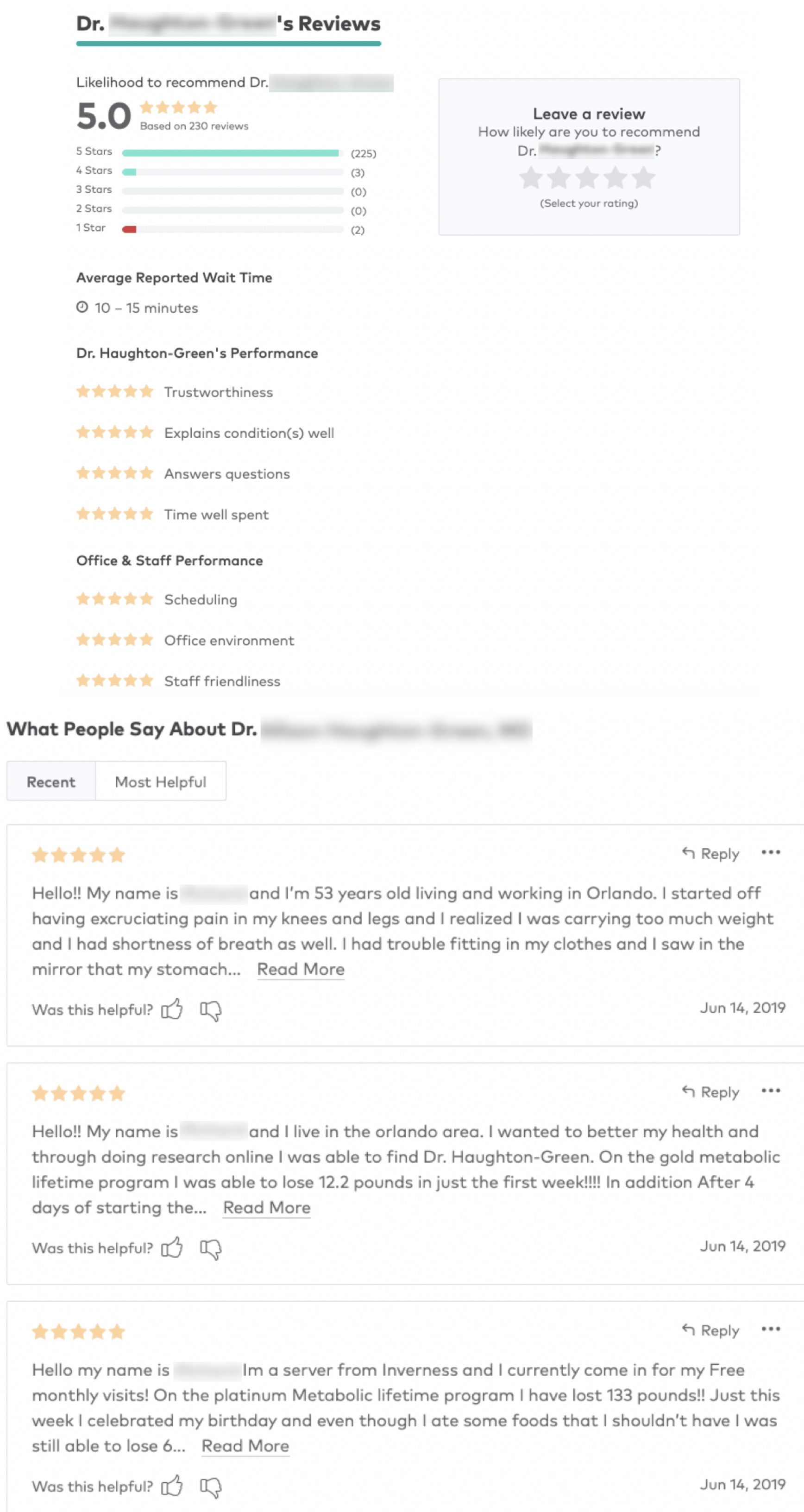
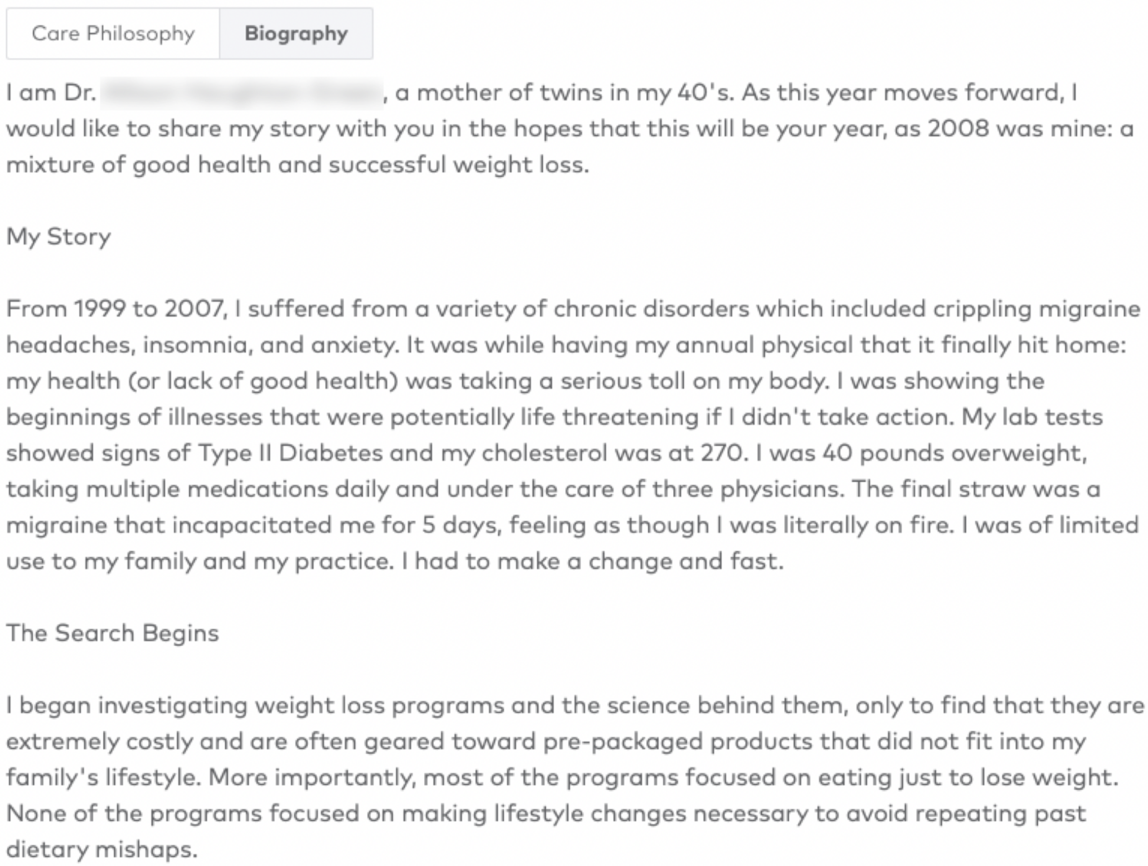


Figure 2.4. Ratings and Reviews



Care Philosophy | **Biography**

I am Dr. [redacted], a mother of twins in my 40's. As this year moves forward, I would like to share my story with you in the hopes that this will be your year, as 2008 was mine: a mixture of good health and successful weight loss.

My Story

From 1999 to 2007, I suffered from a variety of chronic disorders which included crippling migraine headaches, insomnia, and anxiety. It was while having my annual physical that it finally hit home: my health (or lack of good health) was taking a serious toll on my body. I was showing the beginnings of illnesses that were potentially life threatening if I didn't take action. My lab tests showed signs of Type II Diabetes and my cholesterol was at 270. I was 40 pounds overweight, taking multiple medications daily and under the care of three physicians. The final straw was a migraine that incapacitated me for 5 days, feeling as though I was literally on fire. I was of limited use to my family and my practice. I had to make a change and fast.

The Search Begins

I began investigating weight loss programs and the science behind them, only to find that they are extremely costly and are often geared toward pre-packaged products that did not fit into my family's lifestyle. More importantly, most of the programs focused on eating just to lose weight. None of the programs focused on making lifestyle changes necessary to avoid repeating past dietary mishaps.

Figure 2.5. Physician's Personal Statement

In addition to the basic information of a physician, we focus more on the care philosophy as well as the biography shown in Figure 2.5, and we refer to them as personal statements since these paragraphs offer optimal self-expressing content to implement personality detection, especially as these paragraphs are written by physicians about unconstrained topics. Since physicians generally post different content under tag care philosophy and tag biography, we record them as statement 1 and statement 2, respectively. In order to increase the number of detecting words and further improve our estimation accuracy, both statements 1 and 2 will be applied for personality traits detection, and the details will be introduced in Section 2.10.2. The statement always contains a physician's educational background and specified expertise. Furthermore, some physicians choose to share their experiences and the reason why they became doctors. The main descriptions we collected about a physician's basic information, personal statement, and care philosophy are all contained in the tag "About Me" of the six tags that we showed in Figure 2.3.

Figure 2.4 illustrates how a review process is conducted and how the overall reviews are presented to consumers on a website. As we can see, the rating of the physician herself contains four dimensions. Additionally, the rating part can also reflect the performance of office and staff. Except for ratings, reviewers are able to provide optional text comments that provide feedback to a physician and offer more detail that she wants to share with other potential patients.

2.10.2 Detection of Big Five Personality Scores

Below we outline the NLP techniques we utilized to identify the Big Five personality traits from physicians' personal statements. This includes explaining our methodology and reasoning behind it. The personality traits of a physician can be inferred from analyzing the linguistic features present in their unstructured personal statement. Specifically, we extract the physician's score for each of the Big Five traits from two sections of their personal statement: the care philosophy and biography.

For each physician, the statements of care philosophy and biography are treated separately in order to increase the number of measurement as well as improve the estimating accuracy (details are included in Section 2.6.2). Additionally, we preprocess the statements by eliminating stop words, non-English words, and reducing each word to its stem. Following this preprocessing of physician personal statements, the average number of words per statement was 113. Although this figure may be lower than the average number of words used in other studies that utilize user personality attributes [22], our method has demonstrated consistent and accurate outcomes, as presented in Section 2.6.2.

The latent personality traits of physicians are derived using linguistic analytic. To be specified, we analyze the texts of physician personal statement by applying the following three-step method according to Adamopoulos et al. [25]: in step one, for each statement, we use the LIWC dictionary to map each word to one or more dictionary categories and compute a weighted category score for each LIWC category [124]; in step two, we estimate a weighted score for each personality trait, which is determined by the correlation between the cumulative word category scores and personalities [38], [125]. In the final step, we rank the weighted scores in ascending order for each personality trait, and then we use the percentile value (from 1% to 100%) associated with that specific trait as the corresponding personality score. The summary statistic of personality scores is listed in Table 4.1 under physician characteristics. We classify individuals into broad categories for each personality trait only when their respective score surpasses the 50th percentile. The word cloud for each high-level group statement is listed in Figure 2.2 and the speeching styles are totally different among personality groups. Intuitively, we determine wording-using styles of each personality trait and list the chart 2.2f which introduces the guidelines for distinguishing personalities. As we can see, physicians with high openness scores mention lots of medical terms, methodologies, and abstract topics, while those with high conscientiousness scores stress past achievement more (e.g., education, certificates). Extroverts are more likely to include words related to communication and social activity. Moreover, physicians in the high agreeableness group show more kindness and care to their patients. In contrast to others, physicians with high neuroticism seldom use plural pronouns, and instead prefer using singular pronouns and superlative adjectives.

2.10.3 Hospital Discharges Data

The data we use for hospital discharges are compiled by the Office of Data Collection and Quality Assurance (DCQA) from all licensed acute care hospitals, including psychiatric and comprehensive rehabilitation units, comprehensive rehabilitation hospitals, ambulatory surgical centers, and emergency departments. Because the government requires mandatory data collection, our study's sample is unbiased.

In detail, hospital inpatient and outpatient discharge data contain a patient's basic characteristics, including gender, age, ethnicity, race, principle insurance payer, and address, among others. As for the hospital information, our data set contain American Health Care Association (AHCA) facility numbers, which allows us to match any outsourced information as required, such as the distance from a high mortality rate accident place to a hospital and the bed resource of a facility. In addition to hospital information, each inpatient or outpatient discharge record lists an operating physician's NPI, which can be used as a unique label to link with a physician's online features. After matching hospital information via AHCA facility number and matching physician profile via NPI, our full sample covers 831,112 accidental injury encounters (the details of encounter selection will be covered in Section 2.5.3) from 2010 to 2018 on 2,019 physicians who served at least one patient in the entire period in Florida. The summary statistics after the subsampling is shown in Table 4.1 listed in Section 2.5.3.

The data also capture sufficient details of a patient's symptoms, including up to 30 diagnosis ICD-9-CM (or ICD-10-CM) codes that describe diagnoses established to be responsible for occasioning the admission and conditions that are related to treatment services; details of a physician's operations, including up to 30 procedure ICD-9-CM (or ICD-10-PCS) codes representing main procedures during treatment and other necessary procedures provided during the hospitalization; and details of charges in different categories, such as pharmacy charges, laboratory test charges, and emergency charges, among others.

2.10.4 List of Variables in Main Specification

To show more details of what specific characteristics are included in Equation 2.1, we list all variables we used in the following table.

Table 2.11. Table of Variables

Dependent Variables		Mortality; Lab Test Charges; LOS
Independent Variables	Physician Characteristics	Extraversion Score Neuroticism Score Agreeableness Score Conscientiousness Score Openness Score Physician Age Physician Gender Education Rank Working Experience Online Rating Number of Reviews
	Patient Characteristics	Patient Gender Patient Race Patient Age Zip code level Income Insurance Type Charlindex
	Hospital Characteristics	Number of Beds Highway Fatality
	Fixed/Random Effects	Year FE Quarter FE Facility FE Physician Specialty FE Physician RE

2.10.5 Descriptions of Big Five Personality Traits

Openness (or Openness to Experience): The openness to experience dimension encompasses one's level of creativity, curiosity, and an array of interests, and mainly involves traits such as a vivid imagination, intellectual curiosity, novelty-seeking, and variety. Essentially, openness to experience signifies a disposition to have a lively imagination, explore intellectually, and remain receptive to novel ideas and experiences [126]. Moreover, openness is associated with originality, inquisitiveness, and inventiveness, and indicates an individual's inclination towards self-cultivation and novel experiences. Put simply, people who are more open-minded are typically more innovative, imaginative, and insightful [127]. Notably, research suggests that individuals with high levels of openness have an increased capacity for independent judgment when making decisions, especially in pressing circumstances [128].

Conscientiousness: The conscientiousness dimension pertains to one's level of organization and responsibility [126]. Specifically, high levels of conscientiousness are marked by thoughtfulness, self-control, and a goal-oriented demeanor. Individuals with high conscientiousness scores tend to be planners and prefer organized behavior to impulsive actions. Furthermore, conscientiousness is characterized by diligence, persistence, strong organizational skills, responsibility, and a focus on achieving objectives [128]. Costa and McCrae [126] and McCrae and Costa Jr [95] have associated conscientiousness with self-discipline, accomplishment, and competency.

Extraversion: The extraversion dimension centers on one's level of sociability and enthusiasm. High extraversion is indicative of excitability, talkativeness, assertiveness, and emotional expressiveness. Moreover, extraversion is characterized by self-confidence, assertiveness, activity, and a tendency to seek excitement. Extraverts typically display positive emotions, have more frequent and intense personal interactions, and a greater need for stimulation [129]. According to some studies, extraversion encompasses traits such as assertiveness, activity, enthusiasm, talkativeness, warmth, energy, and dominance [126].

Agreeableness: Agreeableness refers to a person's level of friendliness and kindness, and is often associated with behaviors such as altruism, compassion, and cooperation towards others. High agreeableness is also linked with a tendency to trust and a desire to benefit others. In contrast, individuals with low agreeableness scores may be more suspicious or antagonistic towards others. People who score high in agreeableness often have satisfying relationships with others, value close relationships, and report higher levels of happiness and life satisfaction [130]–[132].

Neuroticism: This dimension is known as emotional stability or sometimes referred to as low neuroticism. Emotional stability characterizes individuals who experience less negative emotions such as anxiety, sadness, and moodiness, and are less reactive to stressors. They tend to be more resilient, adaptable, and even-tempered in the face of challenges and setbacks. People who score high in emotional stability are more likely to handle stress effectively, have a positive outlook on life, and experience more positive emotions such as happiness and contentment [126], [128].

2.10.6 Robustness Checks

More robustness checks results are posted in this section. We first test whether multi-collinearity issues, especially for Big Five personality scores, occur in our analysis. To do so, we conduct the test through Variance Inflation Factor (VIF) for every independent variable of interest and list the results for Big Five personality scores in Table 2.12. None of VIFs exceeds 5, which indicates the multi-collinearity is weak and is acceptable to be ignored. Additionally, we test whether results on physician performance, lab test expenditures, and LOS, are consistent by adding each personality trait separately in our model. And the results list in Table 2.13, 2.15, and 2.14 support our statement there is no significant multi-collinearity issue.

Table 2.12. VIF Results for Big Five Personality Scores

	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
VIF	4.58	2.33	3.95	3.32	1.43

To alleviate the concern that transforming the absolute personality scores into the percentile score may distort the results, we also test by using physicians' original personality scores, which is unchanged regardless of sample in each scenario. And we list the results in Table 2.16. Same as what we expected, the transformation of personality scores only tune absolute scores into the relative values following a uniform distribution, which makes the results explanatory without changing the sign. To conclude, these results listed below reveal that the impact of personalities are robust.

Table 2.13. Robustness Check: Multi-Collinearity Test on Physician Performance

	In-Hospital Mortality					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Openness	-0.00542**	-0.00471***				
	(0.00266)	(0.00176)				
Conscientiousness	0.00148		-0.00128			
	(0.00213)		(0.00176)			
Extraversion	0.00266			0.00202		
	(0.00274)			(0.00181)		
Agreeableness	-0.00239				0.00232	
	(0.00243)				(0.00175)	
Neuroticism	-0.00215					-0.00177
	(0.00175)					(0.00173)
Hospital Characteristics	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y
Hospital Fixed Effect	Y	Y	Y	Y	Y	Y
Physician Random Effect	Y	Y	Y	Y	Y	Y
Observations	98,701	98,701	98,701	98,701	98,701	98,701

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

Table 2.14. Robustness Check: Multi-Collinearity Test on LOS

	LOS					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Openness	-1.399***	-0.661***				
	(0.406)	(0.219)				
Conscientiousness	0.592*		0.320			
	(0.327)		(0.259)			
Extraversion	-0.366			0.0480		
	(0.485)			(0.263)		
Agreeableness	-0.421				0.242	
	(0.369)				(0.269)	
Neuroticism	-0.174					-0.115
	(0.243)					(0.235)
Hospital Characteristics	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y
Hospital Fixed Effect	Y	Y	Y	Y	Y	Y
Physician Random Effect	Y	Y	Y	Y	Y	Y
Observations	98,701	98,701	98,701	98,701	98,701	98,701

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

Table 2.15. Robustness Check: Multi-Collinearity Test on Lab Test Expenditure

	Lab Test Expenditure					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Openness	-2,257***	-1732.5***				
	(714.1)	(456.6)				
Conscientiousness	1,692**		1861.2***			
	(730.8)		(649.5)			
Extraversion	350.0			157.4		
	(614.6)			(527.0)		
Agreeableness	-1,608**				-1000.5**	
	(688.4)				(505.8)	
Neuroticism	-303.0					-363.8
	(508.6)					(509.9)
Hospital Characteristics	Y	Y	Y	Y	Y	Y
Physician Characteristics	Y	Y	Y	Y	Y	Y
Hospital Fixed Effect	Y	Y	Y	Y	Y	Y
Physician Random Effect	Y	Y	Y	Y	Y	Y
Observations	98,701	98,701	98,701	98,701	98,701	98,701

Standard errors in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Table 2.16. Robustness Check: Using Original Personality Scores

	Mortality	LOS	Expenditure
VARIABLES	(1)	(2)	(3)
Openness	-0.0005	-0.0479	-226.6***
	(0.0003)	(0.0413)	(82.80)
Conscientiousness	0.0002	0.0925	383.3**
	(0.0004)	(0.0645)	(166.9)
Extraversion	0.0004	0.0513	75.35
	(0.0004)	(0.0567)	(87.58)
Agreeableness	-0.0003	-0.0835	-244.5**
	(0.0004)	(0.0565)	(112.6)
Neuroticism	-0.0006	-0.0114	-59.45
	(0.0004)	(0.0603)	(122.8)
Hospital Characteristics	Y	Y	Y
Physician Characteristics	Y	Y	Y
Hospital Fixed Effect	Y	Y	Y
Physician Random Effect	Y	Y	Y
Observations	98,701	65,007	98,701

Standard errors in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

3. IS DIGITAL GOODS CONSUMPTION RESILIENT TO PHYSICAL RESTRICTION?

3.1 Introduction

In December 2019, an outbreak of pneumonia has been first reported in Wuhan, China, and been recorded as the first coronavirus (COVID-19) case. In March 2020, the COVID-19 pandemic rapidly spread worldwide and was officially recognized as a global pandemic by the World Health Organization (WHO). As one of the most severe disasters in recent years, the ongoing COVID-19 pandemic, tremendously reduces individual's physical mobility [133], which has been reported further lead to profound interventions for numerous industries including construction, hotel and restaurant, manufacturing, service, and e-commerce [134], [135]. Mobile app industry is not an exception. Surprisingly, different from majority of industry severely disrupted by people's reducing daily movement, anecdotal evidences report app industry peaked along with mobility reduction, including the surge in activity on dating apps[1], launch of health apps [136], installation of entertainment apps[2], registration in sharing apps[3], and time spent on general apps usage[4]. These phenomena are not counterintuitive as the mobility reduction blocks the channels to contact with the outer world and motivates people to live and connect through digital approaches, containing mobile apps. On the other hand, mobility reduction induced by COVID-19 pandemic unleashes the possibility of using tremendous amount of substitutes of mobile apps. As an example, New York Times reports the switch of users from apps to webs[5], since one of the most crucial advantages of mobile apps, portability, is eliminated during staying at home. Moreover, for entertainment purpose, compared to apps based on mobile phone, other advanced platforms (such as STEAM and Omni One) based on PC, console, or virtual reality device are more compelling due to comprehensive functions and better user experience. In this sense, the app usage would

[1]↑https://fortune.com/2021/02/12/covid-pandemic-online-dating-apps-usage-tinder-okcupid-bumble-meet-group

[2]↑https://www.inmobi.com/blog/2020/10/08/americans-are-turning-to-apps-for-entertainment-during-covid-19

[3]↑https://www.bbc.com/news/business-57981598

[4]↑https://www.forbes.com/sites/johnkoetsier/2020/08/17/weve-spent-16-trillion-hours-on-mobile-so-far-in-2020/?sh=4000eaee6d61

[5]↑https://www.nytimes.com/interactive/2020/04/07/technology/coronavirus-internet-use.html

decrease because of the substitute effect. To detangle the conflict rationales and fill in the blank of research regarding the impact of mobility restriction on mobile app demand, in this study, we first aim to investigate how COVID-19 induced physical mobility restrictions influence the app industry considering the app category heterogeneity.

Actually, in the practical perspective, a more important task is to estimate the long-term effect of mobility restriction on app usage. As the app industry and the corresponding stock market are largely determined by the long-term evaluation of app economy in post-pandemic era, the ignorance of long-term impact of COVID-19 quarantine will under/overvalue the app market, which leads to the biased business strategy. Nevertheless, how quarantine affects app usage behavior in the long-term is even more ambiguous compared to the short-term one. Notably, there could be a complementary effect of app usage in post-pandemic period due to the learning and/or addicting behavior of app users after quarantine induced app exploration. In the meantime, it is possible to observe a substitute effect of long-term app usage due to the tediousness of app overuse amid quarantine and/or the compensatory behavior of reducing app usage and increasing physical movement after release. To better understand the influencing mechanism of app usage behaviors, leverage the long-term economic impacts of mobility on app industry, and provide the clear identified implications for better responding in post-pandemic era, it is appealing and crucial to identify not only the short, but also the long-term influence on app usage.

To operationalize our research agenda, we use the app usage data from a leading telecommunication company in China and take advantage of a natural experiment research design. Specifically, during COVID-19 pandemic, local government administrations in China publish the enforced quarantine policy announced that: Once any Coronavirus cases are found, the whole neighborhood, which includes hundreds or thousands of residents, has to be quarantined for two weeks for the purpose of avoiding the risk of virus spread. Notably, the policy is restrictedly implemented by local government to guarantee the safety of unquarantined region for citizens' regular living. While being doubted due to ethical issues of enforced mobility restriction, the policy is extremely effective with zero local infection COVID case reported after policy implementation[6]. As the residents are enforced to stay at home due

[6]↑https://www.nytimes.com/2020/03/18/world/asia/china-coronavirus-zero-infections.html

to the unpredictable infection of unknown neighbors, this quarantine treatment is randomly assigned to individuals and largely avoids endogeneity concerns. In an empirical perspective, COVID-19 quarantine policy in China provides us with a clear identification for claiming the causal effect on app user behavior changes. We assign mobile users who have been quarantined in observed period to treatment group, others to control group. Meanwhile, app usage behavior alter in two weeks under quarantine is identified by short-term effect and the behavior alter in the next two weeks after quarantine release is identified by long-term effect. We depict user's app usage behavior in two aspects, including usage volume measured by logarithm of cellular data and usage time dispersion measured by odds ratio of occupied time in 24-hour grid.

Form the supplemental perspective, the natural experiment of COVID-19 quarantine provides us with an unique opportunity to examine whether people's physical mobility has the causal effect on their app usage. Notably, to better understand human behavior on mobile apps and leverage digitization, researchers in the past decades have paid special attention to leverage physical mobility predicting app usage [137]–[139], and many have shown strong correlation between physical mobility and app usage [139]–[142]. However, none of them successes to claim the casual relationships and the mechanism is even more ambiguous due to the infeasibility of experiment, the lack of large-scale data in real-world situation, the ethical issues of trace monitoring, among others. In rationale, there could the following three alternative causal effects, separately or jointly, execute between physical mobility and app usage: (1) One's physical mobility alters one's app usage behavior; (2) One's app usage behavior changes one's physical mobility; (3) Other factor(s) simultaneously affect one's physical mobility and app usage. As the quarantine policy arbitrary exerted on app user level and enormously restrict one's physical mobility, which further leads to individual behavior change in digitalization. In this sense, the two-stage least-squares (2SLS) design allows us generalizing from a particular quarantine effect to physical mobility effect, and shedding light on the causal effect of mobility on app usage.

Leveraging the natural experiment and the 2SLS research design, we first estimate the short-term effects and long-term effects of quarantine on individual physical mobility and app usage, respectively. Our analyses reveal several notable findings. First, we reveal the alteration of individual physical mobility and app usage behavior induced by the quarantine policy in both short and long-term. As for the physical mobility, our estimation underpins the huge decrease of visiting places and movement radius of gyration during quarantine and, surprisingly, reveals the compensatory behavior in the long-term as users tend to enhance physical movement even intensively than before after being released from quarantine. In the perspective of app usage, we reveal the quarantine policy effect on app usage following the categorization criterion of Apple app store and show in most of categories, quarantine induces users increasing app usage in terms of data usage and time dispersion, except for travel apps and mobile games. While the short-term impacts of quarantine are not counter-intuitive, we observe the continuous surge in app usage even after the user being released from quarantine, which implies the complementary effect of quarantine in the long-term and indicates the lasting prospect of app economy in post-pandemic era. Second, the special 2SLS design based on quarantine policy allows us to investigate and quantify the effect of physical mobility on app usage behaviors. Therefore, the study is not limited in an unusual crisis situation induced by COVID-19. Most importantly, the research could be extended to more general circumstances revealing the casual effect of physical mobility on app usage. Specifically, the estimations show that 10% reduction in one's physical mobility leads to 2.68% increase of app usage in cellular data and 5.44% increase in app usage time dispersion, which indicates the substitute effect between user-level app usage and physical mobility. Last, in managerial implication perspective, we identify the app heterogeneity by specifying the effect of physical mobility on app for each app category and for top/low-ranked app groups. In detail, creativity apps, social apps, and financial apps hold the top three usage in quarantine implying humans instinct to knowledge creating and sharing, connecting, and trading when isolated. Additionally, the shrinking demand of travel and the availability of PC games during staying-at-home observes the decreasing usage of travel apps and insignificant change of usage in game apps. While we find substitute effect between physical mobility and app usage of head apps, no salient effect is observed when looking for tail apps usage, indicating head apps dominates human connecting and exploring desire in virtual world. Additionally, we develop feasible business strategies for app designer, practitioner, and policy maker that leverage the international travel quarantine in post-pandemic and the long-term complementary effect of quarantine with special focus on apps with high mobile-adaptable to alleviate the competition with substitute devices.

Our study contributes to the literature in the following three ways. First, this work explores and quantifies the impact of COVID-19 quarantine policy on comprehensive digital services through individual level app usage in both the short term and the long term. Meanwhile, we identify one causal factor, quarantine policy, of explored app usage induced by COVID-19 pandemic among others. Second, this study quantifies the marginal effect of individual's physical mobility on app usage behavior. Specifically, the application of 2SLS design based on COVID-19 quarantine policy provides us a unique opportunity to analyze the causal relationship between individual's physical mobility and virtual exploration activities via app usage. To our knowledge, this study is the first to identify the causal impact of physical mobility on one's app usage in large scale natural experimental setting. Additionally, our analyses show the substitute relationship between physical mobility and app usage. Third, this study helps further understand influencing mechanism of app usage behaviors and provides managerial insights with app industry in the perspective of app categories and app ranks. Specifically, we show possibility that app developers can improve app demand by utilizing the substitute relationship and embedding relative features under consideration of physical mobility in post-pandemic era. Thus, our paper is not limited to offer the empirical contributions to the literature, but also provides the practical and managerial insights for app designer and industry.

The remaining of this paper is organized as follows: Section 3.2 provides a literature review that emphasizes the connections between physical activities and digital consumption, the correlation between physical mobility and app usage patterns, and the demand for apps. Section 3.3 outlines the data structure and identification techniques employed in the study. Section 3.4 introduces the econometric model that was utilized. Section 3.5 discusses the primary findings. Finally, in Section 3.6, we conclude the study by discussing the implications of our results and their relevance for management.

3.2 Literature Review

3.2.1 Digital Resilience to COVID-19

The ongoing COVID-19 pandemic could lead to catastrophic interruptions for various industry sectors [134], including the dining [143], [144], the retailer and wholesales [145], [146], supply chain and logistics [147], tourism industry [148], and creative industry [149]. Among them, while restriction in mobility is implied as one of the most influential factors amid pandemic such that causes the decrease in offline activities [144], [146], [149], none of above study clearly isolates the effect of physical mobility from other COVID-19 induced impacts.

Unlike those traditional economy, digital economy is expected to observe the digital resilience due to the backing of online activities through information technologies and less subjects to physical mobilities [150]. However, the observed impacts of COVID-19 on digital economy are mixed. On the one hand, telemedicine [5], [151], social media and digital education implications [152]–[155], online communication digital platforms including Microsoft Teams, Zoom, and others [156], [157], COVID-19 motivated contact tracing apps [158], and app development jobs [159] have been growing tremendously amid pandemic. On the other hand, research shows the decreasing use of music streaming services [160] and the drop in e-commerce with later recovery [135]. Given that the current studies are all tangent to a particular type of digital products or services, there is a call to research to comprehensive analysis that covers the full spectrum of digital economy. Also, while Chu et al. [5], Carroll and Conboy [152], Chen et al. [153], Sarkady et al. [148], and Sim et al. [160] have implied the reduction in physical mobility as a reason for the changes in digital goods usage amid pandemic, none of work has attributed the changes of digital activities to physical mobility yet.

Our study aims to fill in the research gap in above two streams. First, we provide the empirical evidence on digital resilience in a much boarder spectrum by containing thirteen mainstream types of digital services, which largely extend the understanding of the power of digital resilience in various realms. Moreover, different from prior works those haven't identified any mechanisms inducing the digital resilience, the quasi-experiment design allows us to clearly isolate the impact of one crucial element, the restriction in physical mobility, among other factors amid the COVID-19 pandemic. To our knowledge, this study is the pointer that investigates and quantifies how COVID-19 quarantine affects user's app usage behavior in all categories of digital services and identify a specific factor that induce the surge in app usage. In this regard, our results can provide practical guidelines for app industry to better respond to physical restriction across app categories and across app ranks, respectively, and take advantage of the long-term economic influences in the post-pandemic crisis.

3.2.2 App Economy and Physical Mobility

Previous research on the demand for mobile apps can be broadly categorized into three main substreams. The first substream focuses on app attributes, which includes the different price characteristics of an app [161], the diversity in categories [162], and the leverage of platform synergy to increase app performance [163], [164]. The second research substream focuses on various market factors. For instance, the effect of best-seller rank [165], [166], the advertising traffic [167], the membership overlap between social apps [168], and the release of copycat apps [169]. The third stream further looks into the impact of the environmental factors such as the community behavior [170], visiting locations [139], [141], even air pollution [171].

One of the environmental factors that is important to consider is physical mobility, which has been shown to be closely linked to app usage behavior in previous studies. Qiao et al. [172] were among the first to reveal the strong correlation between app usage and human mobility, considering factors such as individual mobility characteristics, location, and travel patterns. Yang et al. [173] also found significant correlations between human mobility and app usage, and noted that these relationships varied across different app categories. Other studies have focused on how apps interact with mobility and physical activities. For example, Zhu et al. [141] proposed a new location-based probabilistic mechanism for recommending mobile apps, while Lu et al. [142] used users' moving speed as a key metric of mobility and found that although low-speed users switched apps more frequently, they tended to use a narrower range of app categories. In order to better understand the reasons behind app usage and improve app performance and demand, De Nadai et al. [139] conducted a six-month study of 400,000 individuals and found that human behavior in virtual spaces was similar to that in physical spaces, in terms of the capacity to explore new apps or visit new locations.

However, the prior studies investigate in the perspective of claiming association relationship or facilitating mobility for prediction propose. To our knowledge, there is no literature proving the causal relationship between physical mobility and app usage in any directions. Such an absence has been mainly caused by the limitation of experimental settings. Although above mechanisms are possible be tested by experiments, the results and induced conclusions of lab or field experiment are hard to be generalized and extrapolated to real-world scenario due to identification challenges from the confounding bias. In addition to being vulnerable to confounders, limited generalization is another challenge for experimental method. Specifically, while experiments could examine the short-term impact of physical mobility restriction, it is impractical and unethical to exposure a large number of individuals to mobility restriction and test the long-run effect in a large scalar. Meanwhile, the ethical issues regarding restricting one's movement is inevitable and eager to be further discussed. Our research setting, however, considers long-term scenario and link physical and virtual activities without above mentioned concerns by facilitating the quasi-random experiment based research design. Although several industrial articles mention a number of apps show rapidly increasing usage[7] [8], these app reports simply present the monthly trend of app usage on app level amid pandemic lockdown, which fails to identify the specific mechanism leading to app usage surge and cannot imply any causal effects due to confounders. Remarkably, to our knowledge, our study is the first one that evaluates the causal effect of the individual-level physical mobility on app usage behaviors in large-scale setting.

[7]↑https://clevertap.com/blog/q1-data-impact-of-covid-19/

[8]↑https://techcrunch.com/2022/01/10/how-the-mobile-app-ecosystem-adapted-to-the-covid-19-pandemic-in-2021/

3.3 Research Context and Data

3.3.1 COVID-19 Pandemic and Quarantine Policy in China

This study leverages the enforced two-week quarantine policy in China amid the COVID-19 pandemic for the empirical strategy. According to the quarantine policy, once any coronavirus cases are found, the whole neighborhood that contains hundreds of even more citizens is enforced to be quarantined at home for two weeks in order to avoid the risk of virus spread. While in quarantine, health authorities will test residents as often as daily for COVID-19 and will not permit residents to leave their rooms. In accordance with regulation, the local government has formed special teams to distribute living supplies to individual families. Meanwhile, to avoid healthcare issues during quarantine, medical teams are on call for an emergency to occur in policy-affected neighborhoods.

Two advantages are attached to this research context. The first key reason supporting our identification strategy is the power of execution in implementing quarantine policy in China, as overall Chinese citizens are more collectivism. Although many cities around the world are locked down amid pandemic as physical isolation is acknowledged as the best way to fully prevent and largely control the COVID-19 spread, most governments haven't forcibly banned people from going out. Unlike other countries, Chinese governments insist to implement forcible but effective rules controlling people's physical movement to strictly prevent COVID-19 spread. Remarkably, the policy is compulsorily implemented by local government to guarantee the safety of those unquarantined regions for citizens' regular living. While being doubted due to ethical issues of enforced mobility restriction, the compulsory policy is extremely successful with zero local infection COVID-19 case reported after policy implementation and effectively protects unquarantined citizens' physical activities[9]. Moreover, such a treatment is scarcely available from other natural contexts, nonetheless an experiment whose intervention is typically transitory to short-term.

Second, the quarantine policy we investigate is nearly exogenous to most affected individual citizens[10]. For all the individuals (except the one who is diagnosed positive for

[9]↑https://www.nytimes.com/2020/03/18/world/asia/china-coronavirus-zero-infections.html
[10]↑Except the infected individual who triggers this policy.

coronavirus and triggers quarantine), the launch of policy is imposed on them exogenously. As a result, the set of quarantined individuals can be theoretically exchangeable in regard to their heterogeneity with the rest, unquarantined individuals, at least partially. Despite the app users with various app-using habits may be unevenly distributed across regions (e.g., It is possible that app users living in rural areas may rely less on mobile apps and be less impacted by quarantines, when compared to those residing in urban areas, due to the sparse population and delayed development of telecommunications in developed regions), resulting in correlation among location, being quarantined, and app usage, such an issue can be conveniently alleviated by matching the quarantined user to the unquarantined ones with similar historical app usage habits as well as locations, as we have a broad and large set of unquantized users. For example, the quarantined individuals, theoretically, are nearly exchangeable with nearby but unquarantined individuals. In addition, the implementation provides us with a clear cutoff between quarantined and unquarantined individuals, which allows us to perfectly isolate two groups of users. Therefore, the variation of individual-level app usage behavior on quarantine after matching is unlikely to be confounded by unobservable socioeconomic factors.

3.3.2 Digital Goods Consumption

To comprehensively measure the digital economy consumption, we quantify the app usage across different categories to reflect digital goods consumption for corresponding groups of services. Measuring such mobile app consumption by category-aggregated app usage has three advantages compared to other methods of depicting digital goods consumption. First, mobile apps are a group of representative digital goods and have one of the strongest presences in digital goods economy [174]. Particularly, these decades observe the emerging mobile app industry and its profound influences on our society and economy, which reveals the dominative role of app economy in ongoing digital revolution [175]. Meanwhile, as mobile app is based on a typical digital environment, smartphone, it shares consistent background factors with other digital goods. In this sense, as an experiential good, mobile app consumption process (i.e., user experience) is also fairly representative of most digital goods.

Second, unlike most digital goods that focus on a particular service, mobile apps cover a very board range of digital goods, which include but not limited to online social activities, information search, digital reading, e-commerce, among others. Therefore, leveraging mobile app usage allows us to conveniently measure the consumptions of digital goods in a much broader spectrum, which is critical for our research objective but not easily feasible for other digital goods with limited functions. It also allows us to detect potential differences of digital resilience in multiplicity of digital realms.

Third, the accessibility to individual time-specific app usage record allows us to achieve the more objective and detailed measure of digital consumption. Unlike other studies measured app demand by in-app purchases or one-time app downloads [161], [162], [165], which can reflect the app demand but are somehow biased for capturing the real digital goods consumption, this study has the access to more objective measures to reflect a user's real-time app consumption behaviors. By decoding the log files of anonymized individual app users, we easily find, at a certain time, which apps were consumed, how much data is consumed, and so forth. In addition to app features, due to the global positioning system embedded in smartphones, each particular app usage record is associated with geographic location of user, which is crucial for our research design in depicting user's physical mobility. Overall, such a rich set of information is hardly recoverable in other measurement of digital goods consumption.

3.3.3 Data Description

The data are compiled from two sources. First, through the cooperation of a research lab under a leading telecommunication company in China, we are granted to have all the mobile app consumption record of a full sample set of anonymous users in the city of Harbin. This part of dataset contains a total of 39,179,302 app hourly usage records of 73,590 users in the period from March 1, 2020 to May 31, 2020. In specific, for each user in particular hour slot, we can gauge user's cellular data usage for each mobile app. Due to the huge dimensionality of mobile apps, we further divide apps into 11 categories based on the categorization criterion of Apple app store and aggregate app usage on each category. Particularly, 11 app categories include: (1) Social; (2) Productivity; (3) Finance; (4) Utilities; (5) Creativity; (6) Shopping and Food; (7) Travel; (8) Entertainment; (9) Games; (10) Information and Reading; and (11) Others. And our sample lies in the categories of apps that the user has used. As a result, we achieve a panel date set, including user id, time stamp, hourly data usage, app category, and importantly, the coordinate of user's location when using apps which is recorded by GPS.

Our second data comes from the government announcement of the same city. It is publicly released and contains information about locked down neighborhood during observation window period. Such announcement allows us to specify the coordinate locations of quarantined area in terms of longitude and latitude, and the time period of quarantine implementation. The above data is joined with the GPS coordinates of users in the app usage data to identify whether and when a user has been quarantined. A user is deemed to be quarantined if (1) one lives within or near the locked-down neighborhoods, and (2) the variation of her GPS trajectory calculated from the decoded GPS position is nearly zero, verifying nearly no physical movement in distance for two weeks. A time-varying binary variable $Quarantine$ for an app user who is under quarantine at a time equals to 1, others 0, to capture the real-time quarantine state of a user. In addition to investigating the immediate effect of quarantine, we are interested in whether the quarantine influence on users is persistent in the long run even after the quarantine ends. Thus, we further define a variable, $LongTerm$, indicating whether an app user has been quarantined and then released in recent two weeks to capture the potential persistence effect of quarantine. These two variables are our major independent variables under interest.

Our two key dependent variables describing user's app usage behaviors are quantity of app consumption, and time dispersion of digital consumption in app category. For $Data\ Usage$, we decode each user's daily app consumption per category and measure the app consumption in terms of cellular data usage induced by using such category of apps. For $Time\ Dispersion$ of app usage behavior, we depict whether an individual uses a category of apps in concentrated or decentralized time slots in a day. We leverage the hourly app usage records and further define $Time\ Dispersion$ of app usage by odds ratio of hourly-based app occupation per day. Specifically, $Time\ Dispersion$ is calculated by $ln(\frac{n}{24-n+\epsilon})$, where $n \in \{0, 1, 2, ..., 24\}$ and indicates the count of hours during which the app is used within a day. In particular, ϵ

is a very small number, nearly 0[11], to prevent the infinity of time dispersion when $n = 24$. As we see that both above variables, $Data\ Usage$ and $Time\ Dispersion$, tend to be left-skewed and their logarithmic distribution better satisfies the Gaussian-distributed error term assumption, in the latter analysis we use the logarithmic $Data\ Usage$ and $Time\ Dispersion$ as the dependent variables.

To reveal user behaviors in physical mobility, we decoded user physical movement information, $Radius\ of\ Gyration$ and $Number\ of\ Locations$, as two supplemental variables. Specifically, leveraging user's longitudes and latitudes while app usage, we quantified each user's daily radial distance of movement range by $Radius\ of\ Gyration$, which is commonly used as a crucial factor describing individuals' mobility. By definition, $Radius\ of\ Gyration$ is calculated by $\sqrt{\frac{1}{n}\sum_{\mathrm{i}=1}^{n}(x_{\mathrm{i}} - \frac{1}{n}\sum_{\mathrm{i}=1}^{n} x_{\mathrm{i}})^2}$, where x_{i} denotes the longitude and latitude of geographical observation i and can be simply understood as the geographical radius of the area that individuals traveled and activated in. In addition, we gauge the number of visited locations for each user on each observation day and denote it by variable $Number\ of\ Locations$. The detailed definition and the statistical summary of all the above variables are listed in Table 3.1 as follows[12].

Note that our data set contains a sample of citizens in an entire city containing all of 9 districts, including both suburban and urban areas, from very diverse backgrounds. In this regard, the research design ensures representativeness for the neighborhood. Also, the focal city is an average-scale and developed city in China. Compared to the contexts in the smaller scope of subjects, the large magnitude of research context in this study suggests the potential for enhanced representativeness for the general Chinese population, considering that the app users with various app using habits may unevenly distributed across regions.

[11]↑Without loss of generality, ϵ is set as 8.928e-43 such that $Time\ Dispersion$ equals to 100 when $n = 24$.

[12]↑The table presents statistical summary of the main sample after propensity score matching described under Section 3.4, which contains 1,820 app users with 167,440 app usage records in each one of 11 app categories.

Table 3.1. Descriptive Statistics

	Variable	Mean	Std. Dev.	Min	Max	Description
Data	*All Categories*	9.771	4.753	-6.908	17.023	Logarithm valued data usage in mobile app by user i on day t
Usage	*Social*	6.225	4.511	-6.908	14.706	Data usage of social apps
	Productivity	1.976	4.553	-6.908	14.606	Data usage of productive apps
	Financial	-0.173	6.123	-6.908	13.535	Data usage of financial apps
	Utilities	3.559	4.321	-6.908	15.002	Data usage of utility apps
	Creativity	3.297	8.583	-6.908	15.974	Data usage of creativity apps
	Shopping & Food	1.577	5.637	-6.908	15.403	Data usage of shopping and delivery apps
	Travel	-5.987	2.880	-6.908	11.479	Data usage of apps for trip
	Entertainment	0.814	6.290	-6.908	16.494	Data usage of entertainment apps
	Game	-1.212	5.591	-6.908	13.708	Data usage of gaming apps
	Information & Reading	4.055	5.428	-6.908	15.700	Data usage of information consuming apps
	Other Apps	2.239	5.335	-6.908	16.871	Data usage of other apps
Time	*All Categories*	12.542	33.151	-6.908	100	Logarithm valued app time usage of app by user i on day t
Dispersion	*Social*	1.878	17.251	-6.908	100	Data usage of social apps
	Productivity	-2.414	4.462	-6.908	100	Data usage of productive apps
	Financial	-4.041	2.773	-6.908	100	Data usage of financial apps
	Utilities	-1.674	5.168	-6.908	100	Data usage of utility apps
	Creativity	-3.865	2.766	-6.908	100	Data usage of creativity apps
	Shopping & Food	-3.085	2.917	-6.908	100	Data usage of shopping and delivery apps
	Travel	-6.394	2.425	-6.908	100	Data usage of apps for trip
	Entertainment	-3.677	3.239	-6.908	100	Data usage of entertainment apps
	Game	-4.149	3.500	-6.908	100	Data usage of gaming apps
	Information & Reading	-2.312	3.680	-6.908	100	Data usage of information consuming apps
	Other Apps	-2.784	4.576	-6.908	100	Data usage of other apps
Radius of Gyration		10.898	33.486	0	818.274	The radial distance to average coordinate in kilometers by user i in day t
Num. Locations		27.697	32.644	1	442	Number of locations user i visits on day t
Quarantine		0.075	0.264	0	1	Dummy indicator of the individual is under the quarantine
LongTerm		0.072	0.258	0	1	Dummy indicator of the individual is in 14 days after quarantine

3.4 Model

3.4.1 Empirical Design

To investigate the impact of physical mobility on app usage behaviors, the straightforward way should be simply regressing app usage on physical mobility. However, it is prone to two main concerns. First, the coefficient of physical mobility on app usage captured through naive two-way fixed effect regression can be confounded by unobserved individual characteristics, such as user gender, wealth, and living location, among others. Although the fixed effects and the control variables can adjust the meaningful portion of the confounding bias, we still fail to rule out other unobserved confounders. Second, the simple two-way fixed effect scheme is prone to reverse causality. Overcoming the above limitations calls for a clear empirical design, such as a valid instrument variable (IV) that does not suffer from confounders and could clearly restrict the direction of causality.

To address the above two caveats, we enhance the empirical design by incorporating the COVID-19 quarantine policy to instrument individual-level physical mobility. Notably, using the user-level quarantine status (induced by the policy) as the IV is both valid and strong in predictive power. First, given that the quarantine treatment is triggered exogenously conditional on the user's features and only influences individual-level app consumption behaviors via its influence on the reduction of physical mobility, quarantine policy is a decent IV for our research purpose. Second, the policy strictly imposes a prohibition on individuals from going outside of the home, which is forcedly supervised by local governments and monitored through the quarantined individuals' GPS. In this regard, the quarantine policy is an extremely effective IV that obligatorily and explicitly reduces the treated individual's physical mobility to almost zero once the quarantined policy is imposed. To operationalize, we apply user-level quarantine status as IV following a standard 2SLS procedure, in which the first stage captures the quarantine policy effect on an individual's physical mobility and the second stage estimates the marginal effect of physical mobility on an individual's app usage.

To further alleviate the concern of exchangeability between quarantined users and unquarantined ones, we enhance the model specification by leveraging the propensity score matching (PSM) technique. To be specified, the PSM technique is applied to create a group of app users that experience quarantine during the policy execution period ("treatment group") and a group of users who is the most identical to the treated counterpart in the perspective of the probability of getting quarantined but do not experience quarantine ("control group"). As a result, the matching provides us with two groups of quarantined and non-quarantined users that are relatively comparable and exchangeable in terms of their background features, which further improves the validity of applying quarantine policy as IV for individual-level physical mobility.

3.4.2 Matching

To operationalize the empirical scheme of matching, we apply the logistic function leveraging a vector of aggregated numerical characteristics to approximate the propensity scores. In particular, we include the total count of app consumption records, app diversity in terms of app categories, number of physical location visits, and radius of gyration, from the first ten days before the implementation of quarantine policy. Using the propensity scores, we employ one-to-one nearest neighbor matching without replacement to identify the closest-matched pairs of controlled-and-treatment app users. As a result, we obtain 910 pairs of app users matched, a totally of 1,820 users. It is clear to find that systematic differences in user background characteristics, such as app usage and physical mobility, between the quarantined and non-quarantined groups are substantially reduced after conducting the matching.

Figure 3.1 displays a comparison of the overall distributions of propensity scores for the matched and unmatched samples. The distribution of propensity scores for the control and treatment groups are highly overlapped, even without matching. After matching, the discrepancy of propensity scores between app users in the treatment and control groups is largely reduced, and the distributions of propensity scores become almost identical across both groups. Table 3.2 tests the difference of means for all covariates used in the matching process and suggests that, after matching, none of the variables show significant differences between the treatment and control groups. These pieces of evidence support the exchange-

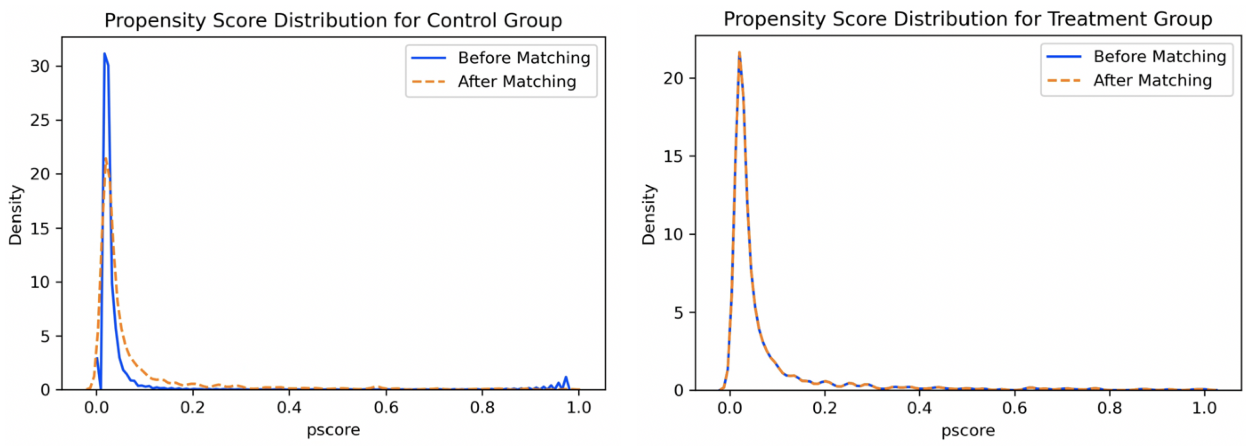


Figure 3.1. Propensity Scores Distribution before and after Matching

Table 3.2. Results of Balance Tests on User's Features after Matching

	Treatment		Control		Difference
Variables	Mean	Std. Dev	Mean	Std. Dev	p-value
Num. App Usage Records	149.670	123.053	151.870	131.049	0.712
App Usage Deversity	10.777	1.649	10.812	1.606	0.645
Num. Locations	2624.033	2857.786	2612.866	2858.155	0.934
Radius of Gyration	7.700	24.036	6.375	18.514	0.188

ability between the treatment and control users, indicating that any sharp differences in app usage after quarantine, if they exist, cannot be explained by individual attributes.

Notably, after the above PSM procedure, we further control for user-level fixed effects to correct for the impact of unobserved time-invariant individual features and control for time-fixed effects to correct for time trends.

3.4.3 Model Specification

After matching, we build up the empirical design by leveraging the quarantine policy as the IV for individual-level physical mobility. Using this design, only the policy intervention, i.e., the variation of physical mobility and app usage due to quarantine, is used to identify the impact of mobility on app consumption.

The consequent estimation follows a standard 2SLS-IV estimator, whose first stage estimates the policy effect on an individual's physical mobility and second stage attributes the change in app usage to the policy effect on mobility. In practice, the first stage of 2SLS when applying quarantine policy as IV exactly follows the DID specification. Specifically, once we obtained two comparable groups of users in the perspective of user regular activities and behaviors, we uncover the causal effects of quarantine policy on an individual's physical mobility and app usage behavior, respectively. Then, in the second stage of the 2SLS scheme, we apply DID-predicted physical mobility to estimate the marginal effect of physical mobility on app usage behavior.

To be more specified, in the first stage, we rely on the equations listed below to estimate the influences of the quarantine on an individual's physical mobility and app usage:

$$PhysicalMobility_{\mathrm{it}} = \beta_0 + \beta_1 Quarantine_{\mathrm{it}} + \beta_2 LongTerm_{\mathrm{it}} + Time_t + User_{\mathrm{i}} + \epsilon_{\mathrm{it}} \quad (3.1)$$

$$Y_{\mathrm{it}} = \gamma_0 + \gamma_1 Quarantine_{\mathrm{it}} + \gamma_2 LongTerm_{\mathrm{it}} + Time_t + User_{\mathrm{i}} + \epsilon'_{\mathrm{it}} \quad (3.2)$$

where the dependent variable in Equations 3.1 $PhysicalMobility_{\mathrm{it}}$ denotes the logarithm value of the visited location count for user i in day t; the dependent variable in Equations 3.2 Y_{it} refers to the logarithm value of data consumption or the logarithm value of usage time distribution formatted in 24-hour grid according to the specification; $Quarantine_{\mathrm{it}}$ is an indicator of whether an individual i in day t being quarantined; $LongTerm_{\mathrm{it}}$ is an indicator of whether an individual i in day t was quarantined in the past 14 days (Note the time span of quarantine is government-regulated as two weeks in studied city. For instance, if a user is in quarantine from April 12 to April 25, then variable $Quarantine_{\mathrm{it}}$ equals 1 for interval April 12 to April 25 and equals 0 otherwise, and variable $LongTerm_{\mathrm{it}}$ equals 1 for interval April 26 to May 9 and equals 0 otherwise); $User_{\mathrm{i}}$ is individual fixed effect; $Time_t$ denotes time (day-specific) fixed effect that capture the common seasonality and time trend across individuals, and ϵ_{it} and ϵ'_{it} are the noise terms. In Equations 3.1 (3.2), coefficient β_1 (γ_1) measures the impact in physical mobility (or the app usage according to the model) induced by policy restriction executed on individuals, if we compare it with the corresponding activities of individuals without quarantine restriction. Coefficient β_2 (γ_2) measures the persistent effect for quarantine that have been imposed and ended within the recent 14 days.

In the second stage, with the assumption that the quarantine restriction affects app consumption indirectly and only through the channel of reducing a user's physical mobility, we estimate the marginal impact of physical mobility on app usage behavior by the following equation:

$$Y_{\mathrm{it}} = \theta_0 + \theta_1 Physical\hat{M}obility_{\mathrm{it}} + Time_t + User_{\mathrm{i}} + \epsilon''_{\mathrm{it}} \tag{3.3}$$

where θ_1 in Equation 3.3 measures the impact of predicted physical mobility obtained in stage 1 on user app usage. As for two dependent variables $Y_{\mathrm{i}t}$, we model the data usage, which is measured by the data usage in kilobytes aggregated on each category, and the time dispersion, which is measured by the logarithm results of odds ratios depicting the division between app usage intervals and spare intervals, respectively. Particularly, we apply the 2SLS design with the correction for comparable treatment and control group members via the PSW method.

Notably, such special design is valuable because it provides us an opportunity to quantify the causal impact of physical mobility on mobile app consumption patterns through the unique regulation during the pandemic, which makes the results applicable to a wider range of situations. Specifically, this study derives the marginal effect of individual-level physical mobility on app consumption, rather than merely estimates the event effect of the quarantine policy. Therefore, the findings provide more generalized insights and can be applicable in many other settings regarding how physical mobility affects digital resilience. Moreover, this 2SLS approach also offers the solution of alleviating the confounding bias associated with the estimation of the physical mobility effect leveraging quarantine policy as IV. In addition to providing marginal effect of physical mobility on full categorical app usage, we further apply the above specification on each app category and estimate the marginal effect in every particular category, respectively, for providing finer analysis in diverse sectors.

3.5 Empirical Results

3.5.1 Results for Policy Effect

We first apply DID method in regression stage one following Equation 3.1 and present the impact of quarantine on the user's physical mobility patterns, including the number of locations one visits per day and the radius of gyration of one's physical movements. The estimations are listed in Table 3.3. Intuitively, the quarantine policy significantly decreases an individual's physical mobility during the treatment. In the meantime, what interests us is the compensatory impact of quarantine, even after being released. To be more specific, the results suggest that after being released from quarantine, individuals who had been quarantined during the pandemic are more likely to exhibit increased physical mobility compared to those who were not quarantined. This increase takes the form of visiting more places and expanding the radius of their activity area. This indicates that individuals tend to compensate for their lack of physical mobility during quarantine by increasing their activity range and visiting a greater variety of places when they are once again able to move freely.

Table 3.3. DID Results for Physical Mobility

VARIABLES	Num. Locations		Radius of Gyration	
	(1)	(2)	(3)	(4)
Quarantine	-1.391***	-1.184***	-0.273***	-0.233***
	(0.0147)	(0.0154)	(0.0262)	(0.0269)
LongTerm		0.737***		0.139***
		(0.0176)		(0.0217)
User Fixed Effect	Yes			
Time Fixed Effect	Yes			
Number of Observations	167,440			
Number of Users	1,820			
R-Squared	0.201	0.210	0.054	0.054

Standard errors in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Similarly, we then report the effects of quarantine's immediate effect and long-term effect on app usage behaviors in Table 3.4, following Equation 3.2. What interests us most is the long-term impact of quarantine. We find that impact of policy is lasting after treatment. Specifically, individuals who have been locked down amid the pandemic but then released

Table 3.4. DID Results for Cellular Data Usage

VARIABLES	Data Usage		Time Dispersion	
	(1)	(2)	(3)	(4)
Quarantine	0.550***	0.652***	0.746***	0.678**
	(0.0368)	(0.0388)	(0.282)	(0.0269)
LongTerm		0.363***		-0.244
		(0.0442)		(0.339)
User Fixed Effect	Yes			
Time Fixed Effect	Yes			
Number of Observations	167,440			
Number of Users	1,820			
R-Squared	0.036	0.037	0.039	0.039

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

are more like to be addicted to mobile apps and consume more in both data consumption and time dispersion compared with those who haven't.

3.5.2 2SLS-IV Results

Columns (1) and (3) in Table 3.5 reports on the main estimations of 2SLS estimates, the marginal effect of physical mobility on app usage. The results shown in the first column reveal the main findings on aggregated full app groups, which indicates a significant negative treatment effect of physical mobility on app consumption. Moreover, the substitute relationship between physical and virtual space activities is implied.

Generally speaking, the 10% reduction in one's physical mobility leads to a 2.68% increase in one's overall app consumption. The results for app usage time dispersion show the similar pattern, which indicate the reduce in physical mobility causes widely dispersed time slots when using general types of apps. More specifically, 10% reduction in one's physical movement leads to 5.44% increase in app usage time dispersion. However, considering the opposite significant impacts of long-term quarantine event on physical mobility and app data usage, we conjecture the mobility effect exerted on app data usage might not restricted to linear relationship. Thus, in Columns (2) and (4), we add the square term of physical

Table 3.5. 2SLS Results for the Impact of Physical Mobility on App Usage Behavior

VARIABLES	Data Usage		Time Dispersion	
	(1)	(2)	(3)	(4)
Physical Mobility	-0.251***	-0.882***	-0.503***	0.0396
	(0.0242)	(0.0869)	(0.186)	(0.666)
Mobility Square		0.166***		-0.143
		(0.0220)		(0.169)
User Fixed Effect	Yes			
Time Fixed Effect	Yes			
Number of Observations	167,440			
Number of Users	1,820			
R-Squared	0.036	0.037	0.038	0.038

Standard errors in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

mobility to better fit the association between physical mobility and mobile app demand. The estimations are consistent with results in Table 3.3 and Table 3.4. Specifically, the data usage follows U-shape along physical mobility, while time dispersion follows linear form.

3.5.3 App Heterogeneity

Heterogeneity in App Category

We conduct external analysis to explore heterogeneity among app categories. In particular, the apps are further divided into 11 categories, representing the various functions to satisfy user's demand. Such categorization provides us a way to estimate the physical mobility influence of various types of virtual exploration desire separately.

According to the results in Table 3.6, the reduce in physical space induces users to increase app usage in most categories, except for travel apps and games. The study findings indicate that the use of certain types of apps increases during periods of physical isolation during a quarantine. For example, there is a noticeable increase in demand for online shopping and delivery service apps, which suggests that individuals rely on these apps to deal with the inconvenience amid the quarantine. Additionally, there is a significant increase in the use of apps that facilitate social interaction and online content consumption during quarantine,

Table 3.6. 2SLS Results for the Impact of Physical Mobility on Cellular Data Usage

VARIABLES	Data Usage					
	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	Social	Productivity	Financial	Utilities	Creativity
Physical Mobility	-0.882***	-0.515***	-0.181*	-0.727***	-0.430***	-1.303***
	(0.0869)	(0.0903)	(0.0939)	(0.117)	(0.0839)	(0.169)
Mobility Square	0.166***	0.0976***	0.0279	0.158***	0.0870***	0.275***
	(0.0220)	(0.0229)	(0.0238)	(0.0296)	(0.0212)	(0.0428)
User FE	Yes					
Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.037	0.129	0.075	0.159	0.216	0.015

VARIABLES	Data Usage					
	(7)	(8)	(9)	(10)	(11)	(12)
	Shopping	Travel	Entertainment	Games	Reading	Others
Physical Mobility	-0.357***	0.0665	-0.568***	-0.0187	-0.219*	-0.364***
	(0.114)	(0.0569)	(0.128)	(0.102)	(0.113)	(0.113)
Mobility Square	0.0729**	-0.0042	0.125***	-0.0057	0.0348	0.0899***
	(0.0288)	(0.0144)	(0.0325)	(0.0259)	(0.0285)	(0.0286)
User FE	Yes					
Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.078	0.074	0.050	0.358	0.072	0.074

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

which supports the idea that people turn to digital approach to achieve the social interactions as they may have had in person. On the other hand, the use of travel apps decreases during periods of physical isolation, which is not surprising as these apps are designed for convenient travel and are less relevant when outdoor physical movement is restricted. Unlike other app categories, mobile game apps are not saliently affected by physical mobility restriction. This result implies that game players either be immune to physical isolation, or transfer to other gaming channels, such as computer and console, due to the diminished convenience advantages of mobile games for users quarantined at home.

Table 3.7. 2SLS Results for the Impact of Physical Mobility on Usage Time Dispersion

VARIABLES	Time Dispersion					
	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	Social	Productivity	Financial	Utilities	Creativity
Physical Mobility	-0.503***	-0.166	0.0047	-0.0566***	-0.100***	-0.0798***
	(0.186)	(0.113)	(0.0283)	(0.0151)	(0.0347)	(0.0156)
User FE	Yes					
Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.038	0.018	0.019	0.137	0.031	0.015
VARIABLES	Time Dispersion					
	(7)	(8)	(9)	(10)	(11)	(12)
	Shopping	Travel	Entertainment	Games	Reading	Others
Physical Mobility	-0.0394**	0.0219	-0.0109	-0.0085	-0.0722***	-0.0104
	(0.0175)	(0.0146)	(0.0195)	(0.0201)	(0.0241)	(0.0293)
User FE	Yes					
Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.037	0.030	0.022	0.229	0.021	0.029

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

Looking in another dimension of app usage behavior, we further examine whether physical mobility affects app usage in occupied time grid and list in Table 73.7. The estimations show a reduction in physical mobility causes widely dispersed time slots when using some categories of apps, including financial, utilities, creativity, shopping, and information acquisition apps. More specifically, 10% reduction in one's physical movement leads to 5.44% increase in app usage time dispersion. Notably, time dispersion, according to its definition, captures hour-based frequency of app usage and somehow manifests app dependency in time basis. The scattered time distribution of app usage induced by reduced physical mobility indicates there is substitute impact of physical mobility on virtual exploration frequency, especially with regards to the creation and acquisition of physical items (e.g., obtain financial products via investment; obtain physical products via online shopping) and virtual items (e.g., explore and download more apps through utility apps; generate information in short video; acquire information via news and reading apps). Meanwhile, we consider the effect of policy itself, and find when individuals are forced to quarantine, their work and leisure activities may start to blend together, making it harder to distinguish between the two. For example, individual without physical mobility restriction tends to leave away from all kinds of apps in working time. In contrast, one under quarantine is inclined to click in mobile apps during working hours, leading to higher app dependency.

Heterogeneity in Head and Tail Apps

In addition to exploring heterogeneity across app functions, we further consider potential differences over app prevalence by separately testing the effect on popular apps and niche apps. Figure 3.2 plots a distribution of app usage measured by total count in records per app, which suggests significant long tail phenomena. Thus, it is possible that the above tested effects differ between head and tail apps, and necessary to examine whether the demand expansion during pandemic and the alteration of app usage pattern occurs in both head and tail apps.

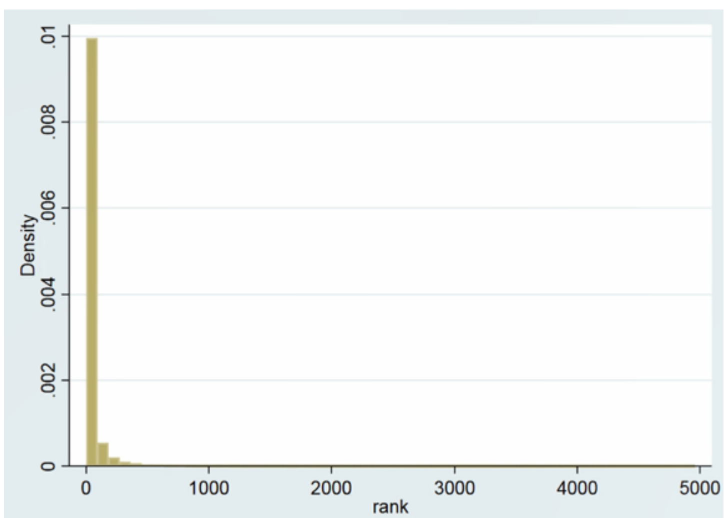


Figure 3.2. Distribution of App Usage over App Rank

To achieve it, we do same analysis testing physical mobility effect on app data usage and time dispersion for head apps, which ranked higher than or equal to 50, and tail apps, which ranked lower than 50, respectively. As the results shown in Table 3.8 and 3.9, while the usage

Table 3.8. Marginal Impact of Physical Mobility on Data Usage Head Apps

VARIABLES	Data Usage					
	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	Social	Productivity	Financial	Utilities	Creativity
Physical Mobility	-0.919***	-0.502***	-0.222***	-0.583***	-0.517***	-1.107***
	(0.0869)	(0.0903)	(0.0953)	(0.121)	(0.0862)	(0.176)
Mobility Square	0.175***	0.0967***	0.0387	0.127***	0.109***	0.225***
	(0.0220)	(0.0229)	(0.0241)	(0.0306)	(0.0218)	(0.0445)
User FE, Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.037	0.136	0.074	0.135	0.228	0.014

VARIABLES	Data Usage					
	(7)	(8)	(9)	(10)	(11)	(12)
	Shopping	Travel	Entertainment	Games	Reading	Others
Physical Mobility	-0.388**	0.0785	-0.760***	0.554***	-0.213*	0.0631
	(0.115)	(0.0493)	(0.131)	(0.0871)	(0.114)	(0.125)
Mobility Square	0.0835***	-0.0107	0.171***	-0.119***	0.0335	-0.0152
	(0.0291)	(0.0125)	(0.0331)	0.0221	(0.0288)	(0.0317)
User FE, Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.074	0.075	0.051	0.555	0.071	0.092

Table 3.9. Marginal Impact of Physical Mobility on Time Dispersion of Head Apps

VARIABLES	Time Dispersion					
	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	Social	Productivity	Financial	Utilities	Creativity
Physical Mobility	-0.599***	-0.451***	-0.141**	-0.288***	-0.0475	-0.292***
	(0.125)	(0.0893)	(0.0583)	(0.101)	(0.0939)	(0.0817)
User FE, Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.162	0.112	0.016	0.029	0.018	0.007

VARIABLES	Time Dispersion					
	(7)	(8)	(9)	(10)	(11)	(12)
	Shopping	Travel	Entertainment	Games	Reading	Others
Physical Mobility	0.0169	-0.0069	0.129	-0.352***	-0.0041	-0.520***
	(0.0785)	(0.0325)	(0.0982)	(0.0837)	(0.0889)	(0.129)
User FE, Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.010	0.013	0.024	0.032	0.016	0.082

for top-ranked apps changes somehow in the similar trend as overall apps, most categories of low-ranked apps are not benefited from demand expansion in terms of data usage. In other words, the decrease in physical mobility significantly surges the usage of top-ranked apps, but slightly influences niche apps, except for mobile games. Notably, the usage of top mobile game apps is boosted by physical mobility, while the usage of niche mobile games is restricted by physical mobility. That points out the unique feature of game: since playing games is much more time consuming compared with using other mobile apps, users tend to stick with most-played games in commuting, and spend time exploring niche games when they are free from travel or moving physically.

Analogously, we conduct same analysis for head apps and tail apps on time dispersion and present estimations in Table 3.10 and 3.11. The results are consistent with the one in data usage. Specifically, for top apps, we find the reduce of mobility in physical space leads to increase of digital exploration frequency in some ways, including the creation and acquisition of both physical items (wealth and products) and digital goods (information and new apps); while the tail apps are immune to physical mobility alteration. The frequency of games playing also follows the same trend as data usage shown in Table 3.8 and 3.9, also suggesting less physical mobility induces less frequency of playing popular games and instead, more exploration of niche games.

3.5.4 Robustness Checks

Leads and Lags

We first test the assumption of parallel trend by including both leads and lags, as shown in Equation 3.4[13]. Specifically, we include the period from two weeks before to two weeks behind of the last period of quarantine triggers, and the estimated leads and lags are plotted in Figure 3.3. The estimates show no effects in the two weeks before the app user is quarantined. Meanwhile, in the several days after being treated, we observe the dramatically increasing

[13]↑Without loss of generality, the effect before 14 days ahead to event day is integrated and captured by β_{-15} and the effect after 14 days behind of event day is integrated and captured by β_{15}.

Table 3.10. Marginal Impact of Physical Mobility on Data Usage Tail Apps

VARIABLES	Data Usage					
	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	Social	Productivity	Financial	Utilities	Creativity
Physical Mobility	-0.535	-0.614	-0.225**	-0.255***	-0.0974	-0.307***
	(0.659)	(0.405)	(0.102)	(0.0560)	(0.118)	(0.0571)
Mobility Square	0.0682	-0.203**	0.0584**	0.0544***	0.0020	0.0600***
	(0.167)	(0.103)	(0.0257)	(0.0142)	(0.0298)	(0.0145)
User FE, Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.037	0.018	0.019	0.120	0.039	0.015

VARIABLES	Data Usage					
	(7)	(8)	(9)	(10)	(11)	(12)
	Shopping	Travel	Entertainment	Games	Reading	Others
Physical Mobility	-0.214***	-0.0107	-0.298***	0.408**	-0.0024	-0.238**
	(0.0583)	(0.0510)	(0.0613)	(0.0695)	(0.0868)	(0.0945)
Mobility Square	0.0478***	0.0073	0.0740***	-0.0949***	-0.0189	0.0671***
	(0.0148)	(0.0129)	(0.0155)	(0.0176)	(0.0220)	(0.0239)
User FE, Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.041	0.026	0.026	0.327	0.022	0.039

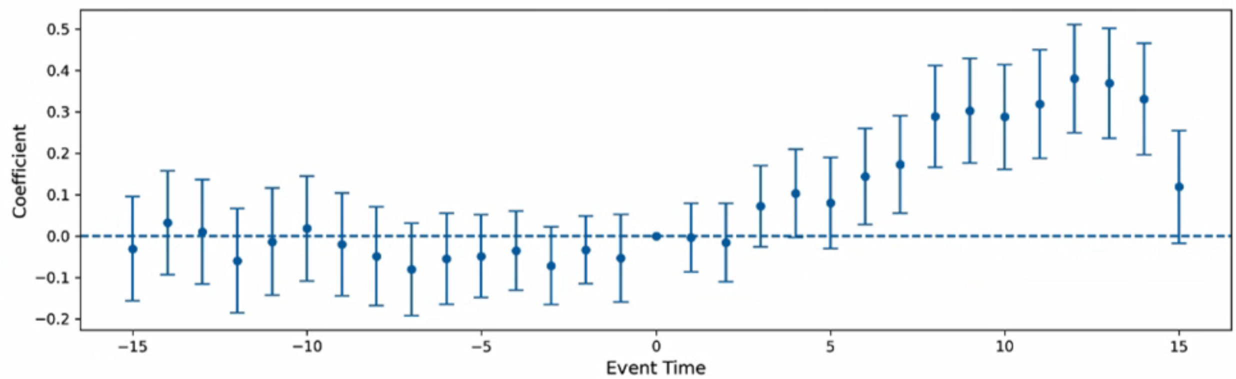
Table 3.11. Marginal Impact of Physical Mobility on Time Dispersion of Tail Apps

VARIABLES	Time Dispersion					
	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	Social	Productivity	Financial	Utilities	Creativity
Physical Mobility	-0.0037	-0.0830	0.0422	-0.141**	-0.109	-0.0667
	(0.199)	(0.0540)	(0.0429)	(0.0560)	(0.0754)	(0.0437)
User FE, Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.035	0.154	0.082	0.044	0.034	0.043

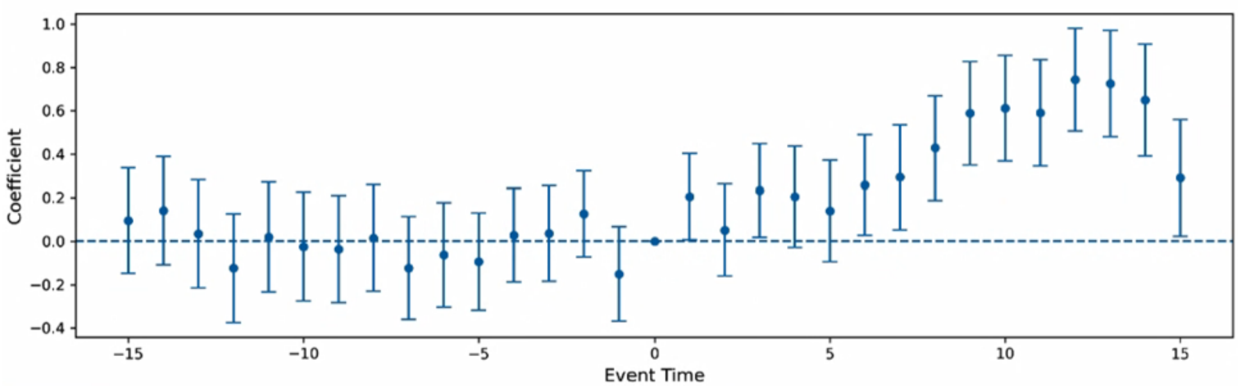
VARIABLES	Time Dispersion					
	(7)	(8)	(9)	(10)	(11)	(12)
	Shopping	Travel	Entertainment	Games	Reading	Others
Physical Mobility	-0.0529	0.0256	0.0323	-0.249***	-0.0384	-0.0702
	(0.0571)	(0.0210)	(0.0748)	(0.0501)	(0.0511)	(0.132)
User FE, Time FE	Yes					
Observations	167,440					
Num. Users	1,820					
R-Squared	0.038	0.015	0.026	0.034	0.060	0.053

effects on both data usage, plotted in blue error bars, and time dispersion, plotted in orange error bars. This pattern implies the consistency with the parallel trend assumption.

$$Y_{\mathrm{it}} = \beta_0 + \sum_{\tau=1}^{15} \beta_{-\tau} D_{\mathrm{i},-\tau} + \sum_{\tau=1}^{15} \beta_{\tau} D_{\mathrm{i},\tau} + Time_t + User_{\mathrm{i}} + \eta_{\mathrm{it}} \tag{3.4}$$



(a) Leads and Lag Test on Data Usage



(b) Leads and Lag Test on Time Dispersion

Figure 3.3. Leads and Lags Test on Data Usage and Time Dispersion

Placebo Test

Additionally, we conduct a series of placebo tests and show the validation in applying a DID strategy. Leveraging placebo falsification, we simply use data for an alternative type of app users whose app usage behavior would not be affected by the quarantine policy amid pandemic. Specifically, we first randomly assign 910 untreated app users into two groups,

Table 3.12. Placebo Test on Cellular Data Usage

VARIABLES	Data Usage									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Quarantine	0.0321	-0.185	-0.064	0.153	-0.071	-0.096	0.152	0.046	-0.156	0.108
	(0.141)	(0.140)	(0.134)	(0.119)	(0.131)	(0.118)	(0.133)	(0.132)	(0.132)	(0.116)
LongTerm	0.0169	-0.166	-0.008	-0.130	-0.025	-0.024	-0.063	-0.043	-0.165	0.271
	(0.156)	(0.170)	(0.164)	(0.160)	(0.161)	(0.157)	(0.148)	(0.162)	(0.147)	(0.144)
User FE	Yes									
Time FE	Yes									
Observations	83,720									
Num. Users	910									

Standard errors in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

Table 3.13. Placebo Test on App Usage Time Dispersion

VARIABLES	Time Dispersion									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Quarantine	-0.268	-1.469*	-0.422	-0.699	0.317	-0.007	0.561	0.075	-0.458	0.656
	(0.758)	(0.756)	(0.732)	(0.810)	(0.710)	(0.763)	(0.760)	(0.743)	(0.709)	(0.764)
LongTerm	-0.461	-0.928	-0.656	-0.708	0.688	0.238	-0.346	-0.043	-1.580*	0.413
	(0.863)	(0.944)	(0.928)	(0.916)	(0.916)	(0.909)	(0.880)	(0.887)	(0.904)	(0.892)
User FE	Yes									
Time FE	Yes									
Observations	83,720									
Num. Users	910									

Standard errors in parentheses
*** p< 0.01, ** p< 0.05, * p< 0.1

treatment group and control group, and then examine the quarantine policy effect on their app usage behavior containing data usage as well as time dispersion, respectively. Columns (1)-(10) in Table 3.12 and 3.13 present the 10-round placebo falsification results of short and long-term quarantine effect on data usage and time dispersion, respectively. As the estimation shows, we cannot observe any significant faked treatment effect, which provides the credibility to our original identification and corresponding estimation results.

Regression Discontinuity Design

Considering citizens might spontaneously reduce their physical mobility during the quarantine period, especially those who live close to the quarantined regions, the quarantine policy might induce continuous impact near the quarantined regions. To alleviate such concern, we further conduct the supplemental test for identifying the discontinuity change near the boundary of the quarantined area. Specifically, the regression discontinuity design (RDD) is used in our study to examine the effect of the quarantine policy on the physical mobility and app consumption patterns of users. Since the RDD was structured for merely capturing the discontinuities by teasing out all the continuous changes across the boundary, it is a decent approach to help us examine whether there is a significant impact of quarantine itself on the boundary of the quarantined area. Assuming that any unobserved determinants of physical mobility and app usage both vary smoothly when passing the quarantine boundary, we investigate if there is a discontinuous changes, which is induced by quarantine, in physical mobility and app usage behavior, respectively, to the quarantined region. To better capture all the continuous changes across the boundary, we adjust for a high flexibility degree of polynomial function in the interval from the boundary on either side of the boundary to largely alleviate all possible bias across the boundary.

To operationalize the above design, we apply specifications listed below to examine for the discontinuous effects of the quarantine on the boundary:

$$PhysicalMobility_{\text{it}} = \delta_0 + \delta_1 Q_{\text{it}} + f(Distance_{\text{it}}) + Q_{\text{it}} f(Distance_{\text{it}}) + Time_t + u_{\text{it}} \tag{3.5}$$

$$Y_{\text{it}} = \omega_0 + \omega_1 Q_{\text{it}} + f(Distance_{\text{it}}) + Q_{\text{it}} f(Distance_{\text{it}}) + Time_t + u'_{\text{it}} \tag{3.6}$$

Where we use Q_{it} denote the indicator of whether the user i is located in the quarantined area at time t. Meanwhile, the polynomial function of user's distance to the quarantine boundary $f(Distance_{it})$ is added on both sides, inside and outside the quarantine regions. Table 3.14 presents the results on physical mobility patterns derived from an RDD based on distance from the quarantine boundary to further enlarge the flexibility. In general, the findings prove that quarantine policy produces sharp differences in individual-level physical

Table 3.14. RDD Results for the Impact of Quarantine on Physical Mobility

VARIABLES	Physical Mobility					
	(1)	(2)	(3)	(4)	(5)	(6)
	Num. Locations			Radius of Gyration		
Quarantine	-0.933***	-1.250***	-1.249***	-0.199***	-0.229***	-0.226***
	(0.0444)	(0.0426)	(0.0426)	(0.0507)	(0.0489)	(0.0489)
Polynomial Degree	Cubic	Square	Linear	Cubic	Square	Linear
Bandwidth	1km					
Time FE	Yes					
Observations	29,876					
Num. Users	2,134					
R-Squared	0.417	0.406	0.406	0.365	0.348	0.347

VARIABLES	Physical Mobility					
	(7)	(8)	(9)	(10)	(11)	(12)
	Num. Locations			Radius of Gyration		
Quarantine	-0.933***	-1.002***	-1.213***	-0.199***	-0.124***	-0.294***
	(0.0444)	(0.0488)	(0.0570)	(0.0507)	(0.0548)	(0.0625)
Bandwidth	1km	0.75km	0.5km	1km	0.75km	0.5km
Polynomial Degree	Cubic					
Time FE	Yes					
Observations	29,876	25,942	21,616	29,876	25,942	21,616
Num. Users	2,134	1,853	1,544	2,134	1,853	1,544
R-Squared	0.417	0.438	0.455	0.365	0.391	0.415

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

mobility controlling the smoothly varying observable determinants at the boundary. To be more robust, in Columns (1) - (6), we test the different degrees of polynomial function in the distance to the quarantine boundary, and in Columns (7) - (12), we test the different bandwidths when defining the nearby unquarantined individuals. Among the different specifications, we find that estimations are highly consistent in both significance and magnitude.

Following the same principle, we show the RDD estimation on app usage behaviors in Table 3.15. Consistent with our main results as well, the findings suggest a significant increase in both app data usage and app usage time dispersion for the quarantined areas, compared with the nearby but unquarantined area. The results are robust across the polynomial functions in Columns (1) - (6) and across the bandwidths in Columns (7) - (12). Combining the

Table 3.15. RDD Results for the Impact of Quarantine on App Usage

VARIABLES	App Usage					
	(1)	(2)	(3)	(4)	(5)	(6)
	Cellular Data Usage			Time Dispersion		
Quarantine	0.779***	0.714***	0.715***	0.618***	0.544***	0.544***
	(0.0658)	(0.0626)	(0.0622)	(0.0326)	(0.0310)	(0.0310)
Polynomial Degree	Cubic	Square	Linear	Cubic	Square	Linear
Bandwidth	1km					
Time FE	Yes					
Observations	29,876					
Num. Users	2,134					
R-Squared	0.061	0.060	0.058	0.103	0.101	0.100

VARIABLES	App Usage					
	(7)	(8)	(9)	(10)	(11)	(12)
	Cellular Data Usage			Time Dispersion		
Quarantine	0.779***	0.784***	0.643***	0.618***	0.614***	0.555***
	(0.0658)	(0.0704)	(0.0796)	(0.0326)	(0.0350)	(0.0395)
Bandwidth	1km	0.75km	0.5km	1km	0.75km	0.5km
Polynomial Degree	Cubic					
Time FE	Yes					
Observations	29,876	25,942	21,616	29,876	25,942	21,616
Num. Users	2,134	1,853	1,544	2,134	1,853	1,544
R-Squared	0.061	0.059	0.061	0.103	0.101	0.102

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

results in Table 3.14 and 3.15, the app user's behavior changes during the quarantine period cannot be fully explained by the continuous variables across the boundary. In other words, the discontinuous changes in both physical mobility and app usage are induced by the quarantine policy itself rather than the user's distance to the quarantined neighborhood. Thus, the conjecture that discontinuous impact caused by quarantine policy is fully supported by RDD estimations.

3.6 Discussion and Conclusion

Using individual-level app usage data in all app categories, this study first interested in and investigate the impact of trending quarantine policy amid one of the most destructive crisis in decades, the COVID-19 pandemic, in both short term and long period. Our results on immediate effect reveal the power of digital resilience to interventions, covering a very board range of digital goods, which include but not limited to online social activities, digital reading, e-commerce, among others. Most importantly, the long-term effects of quarantine policy are also quantified and indicate the persistent impact on app usage even after the user is physically released from quarantine, enabling a mining of the app economic value of in post-pandemic period.

Moreover, applying a highly recommended source of natural, strict, and exogenous variation in physical mobility generated by quarantine policy amid COVID-19 in China, this study then systematically examines whether digital goods consumption in a comprehensive spectrum is resilient to individual's physical restriction and provides a more generalized results in terms of marginal effect of physical mobility. Empirically, we quantify that 10% reduction in one's physical mobility leads to 2.68% increase of app data usage and 5.44% increase in app usage time dispersion, implying the substitute relationship in people's digital goods consumption and physical mobility. The magnitudes are salient if further considering the remarkably large market size of app industry. The estimations are consistent in various measures of app usage. Meanwhile, the assumptions of empirical design are carefully examined and shown hold in different tests.

This research contributes in three main perspectives. First, we provide the empirical evidence on digital resilience in a much boarder spectrum by containing thirteen mainstream types of digital services, which largely extend the understanding of the power of digital resilience in various realms. Moreover, different from prior works those haven't identified any mechanisms inducing the digital resilience, the quasi-experiment design gives us a chance to clearly isolate the influence of one crucial element, the restriction in physical mobility, among other factors amid the COVID-19 pandemic. On such point, this is the first study that investigates and quantifies how COVID-19 quarantine affects user's app usage behavior in all categories of digital services and identify a specific factor that induce the surge in app usage.

Most importantly, the research could be extended to more general circumstances revealing the casual effect of physical mobility on app usage. Instrumenting physical mobility by the execution of quarantine policy, this study shows the quantitative role of physical mobility in reshaping consumer behaviors in various digital categories from economic and behavioral perspectives. Specifically, the application of 2SLS design based on COVID-19 quarantine policy provides us a unique chance to tease out possible confounders and be able to analyze the causal relationship between individual's physical mobility and digital goods consumption particularly via app usage. So far as we known, our study is the first to evaluate the causal effect of the user's physical mobility on app usage in large-scale setting.

Third, this study helps further understand influencing mechanism of app usage behaviors and provides tailored managerial insights with app industry in the perspective of comprehensive app categories as well as app ranks or roles on market. Specifically, we show possibility that app developers can improve app demand by utilizing the substitute relationship and embedding relative features under consideration of physical mobility in post-pandemic era. Thus, our paper is not limited to offer the empirical contributions to the literature, but also provides the practical and managerial insights for app designers and policy makers.

This study also has several limitations that can be addressed in the future study. First, the estimates we obtained are theoretically apply to one Chinese city in this study. Although we cautiously claimed the external validity of large populations in such city and contexts that could account for a sufficiently large portion of population all over the world, extrapolating them to vastly different subpopulations (e.g., from citizens in developing countries to citizens in developed countries) could bias the results and induce misleading inferences. Second, while this study reveals and quantifies the causal effect of physical mobility on digital goods consumption, implying the substitute effect between physical activities and digital exploration, the underlying mechanism explaining such effect is unclear. Thus, future research with supplemental data availability is encouraged to conduct mechanism analyses and uncover the veil.

4. HEALTH SERVICE DELIVERY: DOES ROBOT-ASSISTED SURGERY PLAY A ROLE?

4.1 Introduction

In the past decades, individuals from different backgrounds, social groups, and countries continuously receive divergent levels of health services [176]. Commonly, inequalities in health are attributed to unavoidable inequalities caused by biological differences [176]–[178] and unjust preventable inequities caused by environmental factors, poverty, behaviors, etc [179], [180]. According to centers for disease control and prevention, such disparity is considered as one of the most severe threats to the publics health as the lack of standardized health service experience over a huge group of populations.

Remarkably, with the surge in development for artificial intelligence (AI) technology, the increasing applicability of AI-based treatment in medical care is revealed in multiple perspectives, including but not limited to diagnosis and treatment recommendations, patient engagement and adherence, and surgery assistance. Among other AI-based applications in healthcare, robot-assisted surgery (RAS), with its great precision, flexible movement, and high-definition monitor, is invented to assist diminish the concerns of hand tremors, reduce several surgical limitations of human eyesight, and alleviate the unexpected and mistaken procedures of physicians [181]. More than that, RAS tends to show the potential of enhancing standardization and team cooperation in the treatment process, which might act as moderators to further gradually transform the health service delivery process [182]. Notably, even though AI-related technologies are increasingly prevalent and expected to see growth in various surgeries in the next few years, the evidence on practical healthcare performance improvement is limited [183]–[185]. Therefore, in this work, we expect to provide empirical evidence on RAS method may alter medical service delivery processes as well as outcomes. Specifically, this study aims to identify AIs potential causal impacts on the healthcare delivery process and outcomes, particularly on the performance of physicians, and investigate the specific physicians, patients, or facilitys characteristics that lead to more implementation of AI technology.

To operationalize our research design, we focus on RAS in high-risk diseases due to its high mortality rate and high demand for surgery precision. And we only look for diagnoses that can be treated by RAS procedures. By modeling physicians decision-making process when learning and choosing treatment methods, we study the following three trending questions: First and foremost, we examine the effect of RAS on clinical outcomes, such as whether RAS indeed enhances clinical performance, controlling physicians self-selection behavior in choosing their treatment methods. Second, we focus on the access to learn and use RAS and explore how physician and patient heterogeneity, especially ethnicity features, affect the adoption of the RAS method, including learning RAS and using RAS. For instance, we investigate the concern of racial health inequality by investigating the potential racial bias in the accessibility of RAS resources and revealing the synergy of RAS implementation on alleviating healthcare racial inequality. Ultimately, underlying mechanisms are proposed and tested to provide more insights explaining through which way RAS is able to improve clinical performance.

This work contributes to the literature in the following three important aspects. First, while the potential of RAS in healthcare has been discussed for decades, the effectiveness and safety of RAS in the treatment of patients have yet to be clearly established using practical data in clear design. Our study fills this gap by proving a particular AI implementation in healthcare, RAS, improves overall clinical performance controlling for treatment method selection bias. Thus, our paper sheds light on empirically testing and proving the relative strengths of RAS in enhancing clinical performance. Second, this study advances the understanding of RAS implementation by identifying the physician's decision-making process of treatment selection. Moreover, we uncover the disparity in access to RAS due to race features. By segmenting the access to RAS into two stages, the access to learning RAS and the access to using RAS, we further investigate racial disparities in each stage. And to our knowledge, this research is the first to reveal the RAS selection process and propose a potential way to alleviate healthcare disparity. In the meantime, we are able to emphasize an unanticipated benefit of promoting RAS adoption in practice. Third, the pattern we found and the mechanism analysis could provide suggestive implications to practitioners and policymakers and enrich our understanding of the future of RAS in the healthcare industry. For instance, the potential of RAS to magnify the standardizing capability of RAS in uniform the time and procedures through health service delivery.

Our paper is structured in the following manner. The literature review is presented in Section 4.2. Our model-free findings and research context are included in Section 4.3. A detailed, top-down demonstration of our model is provided in Section 4.4. Our primary findings and robustness checks are presented in Section 4.5. Finally, Section 4.7 concludes our work and includes the future plan.

4.2 Literature Review

4.2.1 Information Technology and Health Service Delivery

Healthcare operations work has focused on the influence of traditional operation processes or economic indices such as workload [26]–[30], queue design [71], [72], and schedules [73] on health service quality. Other researchers also have pointed out the salient role of physician characteristics on healthcare operations, including temporary workers [74], peers [75], gender concordance [76], and experience [77].

Notably, emerging information technologies have been observed transformations in both healthcare delivery procedures and clinical performance. A substream of literature investigates the impact of online platform on healthcare operations. The findings are kind of mixed and highly depend on the research context. Specifically, some studies report not observing any significant relationship between healthcare performance and service provider's online reviews [78], [79]. In contrast, other studies uncover the positive effect of online platforms on healthcare delivery. For example, the positive relationship is revealed between ratings and physician background [80], [81], indicating the predictive power of online ratings on patient perceptions of received service quality. Such findings are further empirically proved by Lu and Rui [7], in which they show the causal effect of online ratings on practical performance of surgeons. More than previewing the quality of healthcare service, what even more appealing is information technologies' potential to transform health service delivery process as well as outcome. To be more general, prior research shows electronic health record systems enhance the health service quality [186], save the annual cost of facilities [187], alleviate the hospitalization rate in emergency department [188], reduce the patient-level readmission rate of heart attach encounters [189], and decrease the length of stay in hospital [190], among others. Furthermore, in the perspective of healthcare operating process, literature reveals the power of health information technology interoperability to facilitate health information exchange between disparate providers [191] and alleviate the healthcare inequality in terms of service quality [1]. More fields are emerging in recent decades, including robot-assisted surgery [3], telemedicine [2], and the Internet of Medical Things with special emphasize on medical technologies [4]. Despite the effects are varying across diagnostic procedures and implemented situations, the value of information technologies on transform healthcare delivery is generally established and waits to be further mined.

4.2.2 RAS Usage and Efficacy

The RAS-related techniques are growing at both academic and corporate levels to overcome the healthcare difficulties, such as shortage in high-end resource, performance disparity, healthcare inequality, among others [192]. Three main merits are attached to RAS compared to traditional laparoscopic or other types of minimally invasive surgical procedures. First, leveraging virtual reality and augmented reality techniques, RAS is able to provide a high-resolution and realistic three-dimensional vision of an immersive environment rather than a conventional two-dimensional one [193]–[195], which intensively enhances the representation of the live operating area and deepens the understanding of a patient's diagnosis [192], [196], [197]. Second, robotic technologies have been mentioned to provide an augmented surgical hand that enables minimally invasive surgery through enhanced telemanipulation. In particular, with their multiple artificial operating arms, surgical robots are great for precision, flexible movement, and multi-processing [198], [199]. The third and most intelligent progress is the data collecting and processing capability of the RAS technique. Embedded with historical medical records and integrated to many other technologies developed for operating surgery, RAS is powered to recognize typical attributes, conduct simulations, and generate surgical strategies [192], [200], [201].

In practical, RAS seems to have reduced mortality and complications with ambiguous evidence [183]. For instance, Seco et al. [202] review 16 relative studies and find that minimally invasive mitral valve surgery observes higher clinical outcomes compared to the non-robotic one. Moreover, Yanagawa et al. [185] control for the patient background and report largely decreased mortality rate in patients taken care by the RAS treatment. Chandra et al. [203] find RAS increases access to partial nephrectomy, which further leads to lower mortality rate. However, most RAS studies are conducted and present evidence on statistical level without any controlling for patient, physician, facility characterise, which highly determine the clinical performance. Even others have attempted to alleviated selection bias, as pointed out by themselves, such efforts only prove the association relationship between RAS implementation and healthcare outcomes [185]. Therefore, taking advantage of a sufficient data resource and the appropriate clinical context, this study aims to tease out the intervention of unobserved factors and identify the causality impact of the RAS technique on clinical performance.

4.3 Data and Research Context

This section elaborates the research context regarding RAS implementation, introduces the data structure and variables of interest, and show some preliminary evidences on the potential effect of RAS on clinical performance.

4.3.1 Data

We utilize inpatient and outpatient encounter data collected from the Florida Department of Health and physician demographics provided by data vendors. Specifically, hospital encounter data are collected from all licensed health service facilities.

In detail, the Florida encounter data set contains the basic patient characteristics, including the patient's gender, age, ethnicity, race, principal insurance payer, address, etc. In addition to the patient information, we combine operating physician characteristics, including physician gender, ethnicity, work experience, educational background, etc., through the National Provider Identifier, a unique label to identify each registered physician. After matching the physician profile, our full sample covers 114,120 encounters from 2015, quarter 3, to 2018. quarter 4, on 2,164 physicians who served at least one patient in the entire period in Florida.

The data also capture sufficient details of patient's symptoms, including up to 30 diagnosis ICD-9-CM (or ICD-10-CM) code describing diagnosis established to be responsible for occasioning the admission and conditions related to the major health services; details of physician's operations, including up to 30 procedure ICD-9-CM (or ICD-10-PCS) code representing main treatment as well as other necessary procedures provided during the hospitalization; details of charges in different categories; clinical outcomes; etc. Particularly, through procedure ICD-9-CM (or ICD-10-PCS) code, we identify whether a patient is received a traditional treatment or a RAS treatment. We report the summary statistics in Table 4.1 in the following.

4.3.2 Research Context

Similar to most of advanced technological, prerequisite knowledge and practicing is required for a surgeon to apply RAS in real cases. In order to further ensure the safe and sustained growth of RAS techniques, experts still continues working on developing objective-based curriculum and advanced training for RAS implementation [204]–[206]. Although RAS training for residents has been increasing with the expansion of RAS, a survey among residents regarding obstetrics and gynaecology specialty reported to have 79% of service providers reveal their willingness to get RAS training in their residency programs, but only 38% of the residents truly have the access to RAS training [207].

Considering the limited access to RAS training and the prerequisite of relative training for the RAS implementation, in this study, we model registered surgeons who participate in three stages of decisions with various level of involvement in terms of RAS method.

First, the physician decides on participation in RAS training, including either participating in training or not participating in training. This will be inferred by historical treatment methods for each physician. Specifically, if a physician has applied RAS treatment at least

Table 4.1. Descriptive Statistics of Variables

Variable	No. of obs.	Mean	SD	Definition
Treatment procedure measure				
RAS	114,120	0.011	0.105	Equals 1 if a encounter is treated by RAS
Mortality	114,120	0.032	0.176	Equals 1 if a encounter died before being discharged
Physician characteristics				
Experience	2,164	24.673	11.333	Number of years since a physician graduated
Female	2,164	0.192	0.394	Equals 1 if a physician is female
White	2,164	0.445	0.497	Equals 1 if a encounter is white
Black	2,164	0.034	0.180	Equals 1 if a encounter is black
Other race	2,164	0.521	0.500	Equals 1 if a encounter is neither white nor black
Education rank	2,164	33.713	30.226	Rank of the physician graduated medical school
Patient characteristics				
Charlindex	114,120	0.686	1.334	Risk index measured by Charlson Comorbidity Index
Female	114,120	0.380	0.485	Equals 1 if a encounter is female
White	114,120	0.794	0.404	Equals 1 if a encounter is white
Black	114,120	0.119	0.324	Equals 1 if a encounter is black
Other race	114,120	0.086	0.280	Equals 1 if a encounter is neither white nor black
Age	114,120	67.917	16.512	Age of a encounter
Medicare	114,120	0.633	0.482	Equals 1 if a encounter is covered by Medicare
Medicaid	114,120	0.070	0.255	Equals 1 if a encounter is covered by Medicaid
Private insurance	114,120	0.189	0.391	Equals 1 if a encounter is covered by private insurance
Other insurance	114,120	0.108	0.311	Equals 1 if a encounter is not covered by above

once, we mark her as a participant in RAS training; otherwise, we treat her as a non-participant.

Second, the physician decides on the treatment method facing a specific patient. Specifically, a physician who is non-participant can only use traditional treatment that not requires any additional training. In contrast, a physician who is a participant can either choose to use RAS or traditional treatment according to the patient's condition and particular circumstance.

Third, the physician performs the surgery through the chosen treatment, RAS or traditional treatment, and generates clinical outcomes. Based on our research objective, we measure the clinical outcome directly by the patient's live condition, saved or death after surgery, because this is a commonly applied proxy in healthcare studies to reflect the service performance [7], [84], [85].

During the first two decisions making processes, physicians' choices are influenced by their work experience, preference, and professional capability. With the introduction of RAS treatment, practitioners have access to the more advanced method, along with additional costs in both learning and applying. Meanwhile, the introduction of RAS treatment largely affects the decision-making process of practitioners who are selecting treatment methods. More specifically, practitioners rationally choose treatment methods under the circumstance to optimize their utility by consistently learning and applying RAS treatment. Leveraging the model of decision-making process, this study not only addresses the endogeneity of treatment selection when exploring the impact of RAS on healthcare performance but also reveals physicians' preference for both RAS learning and using originated from demographic features.

4.3.3 Preliminary Evidences

We now show some preliminary evidences from the simple reduced form approach. Specifically, we simply regress the in-hospital mortality on patient characteristics, physician fixed effects, time fixed effects, and facility fixed effects, in which the binary variable *RAS* indicates whether the treatment is RAS method or traditional treatment. The results are shown in Table 4.2.

Table 4.2. OLS Estimation of RAS Effect on In-Hospital Mortality

	In-Hospital Mortality			
VARIABLES	(1)	(2)	(3)	(4)
RAS	-0.0245***	-0.0154***	-0.0248***	-0.0150***
	(0.00599)	(0.00246)	(0.00612)	(0.00249)
Charlindex	0.00717***	0.00768***	0.00711***	0.00706***
	(0.000425)	(0.000400)	(0.000491)	(0.000435)
Patient - Male	0.00720***	0.00614***	0.00716***	0.00573***
	(0.00141)	(0.00102)	(0.00151)	(0.00108)
Patient Age	0.000793***	0.000725***	0.000770***	0.000639***
	(5.47e-05)	(4.53e-05)	(5.82e-05)	(4.46e-05)
Constant	-0.0158	-0.00243	-0.00478	0.0290
	(0.0542)	(0.0483)	(0.0672)	(0.0556)
Patient Characteristics	Yes	Yes	Yes	Yes
Physician Fixed Effect	Yes	Yes	Yes	Yes
Year, Quarter Fixed Effect	Yes	Yes	Yes	Yes
Emergency Cases	Yes	No	Yes	No
Apply for RAS-Equipped Hospital	No	No	Yes	Yes
Number of NPI	2,164	2,317	1,971	2,096
Observations	114,120	216,037	90,081	174,278

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

Column (1) shows the estimation on our main sample, which include only emergency encounters to alleviate the bias induced by treatment method selection. Column (2) releases such constraint and examines both the emergent and non-emergent encounters. Column (3) considers a more rigorous conditions by restricting samples in facilities that the RAS treatment is applied at least once in prior period, to further ensure the accessibility of RAS equipment in corresponding hospitals. Column (4) in turn tests both the emergent and non-emergent encounters but only in RAS-equipped hospitals. As a result, while the endogenous concerns induce us fail to conclude on causal impact, the results of reduced form approach reveal the strong evidence showing the significant negative association between RAS usage and after-surgery mortality in every specification. Next we effort on develop the model to capture the decision-making process, thus address the RAS selection bias.

4.4 Econometric Method

Our model is proposed based on estimated endogenous treatment procedure following Bratti and Miranda [208] and Rabe-Hesketh and Skronda [209]. Although sharing the similar principal as the treatment effect approach developed by Heckman [210], the method we proposed is more generalized and flexible to fit two-stage decision-making process. Specifically, in our main analysis, we treat the RAS treatment choice as an endogenous factor as it is decided based on physician's features and patient's unobserved conditions. Rather than those approach of separately estimating the method choice and outcome delivery process in two separate steps, we estimate the integrated decision processes simultaneously.

In detail, in our model, a physician first makes two-stage decision regarding RAS learning and RAS using based on the observed physician characteristics and unobserved individual heterogeneity, measured by latent factors. Then, we model a patient's clinical outcome as a function of physician-selected treatment method, patient characteristics, physician fixed effect, and time fixed effect. In the rest of this section, we explain the general setup of a decision-making model in Section 4.4.1 and model the clinical outcome in Section 4.4.2.

4.4.1 Two-Stage Decision-Making Process

Notably, estimating the performance of the RAS treatment is challenging due to the potential endogeneity of the treatment method. This is induced by the existing important confounder for both the implementation of RAS method (that physician and patient can choose) and the clinical outcome of patient. Such shared unobserved heterogeneity could affect both the variable of interest (the adoption of the RAS method) and dependent variable (clinical outcomes, such as patient mortality and medical expenditure), which causes confounding bias. For instance, patients with severe cases may be more likely to choose the RAS treatment but also have a higher risk of post-surgery mortality compared to patients with mild health conditions. To address this issue, our specification leverages two latent factors η_{it1} and $\eta_{\text{ij}t2}$ to account for potential endogeneity in treatment adoption. The parametric approach is applied and allows us to simulate and tease out the dependence between stages of decisions and achieve the impact of RAS usage controlling for the multi-level of endogenous decision.

To be specified, we model physician's choices of treatment selection that is driven by both physician and patient specific variables in two stages. In the first stage, the decision is mainly driven by physician characteristics indicative of physician's inclinations to pursue RAS techniques. The decision is given the maximum expected utility from whether a physician learn RAS techniques or not. Two expected utilities represent the expected utilities for two possible selections: learn RAS; or do not learn RAS. A physician will choose the selection that has the highest utility responding to RAS learning.

In the second stage, the decision is mainly driven by patient characteristics indicative of physician's propensity to apply RAS treatment in a certain medical situation. Specifically, facing a certain patient, the physician decide on which treatment method is better to be used: RAS method; or non-RAS traditional method. Remarkably, two-stage decisions are dependent in the sense that only physicians select on learning RAS have the right to select using RAS in treatment.

As for the notations in empirical models, we let dummy variable P denote a physician's RAS participation and model a physician who decides whether to learn RAS or not denoted by participation indicator of either $P = 1$ or $P = 0$. Similarly, we let dummy variable T denote a physician's RAS treatment selection and model a physician who decides whether to use RAS or not by the treatment indicator of either $T = 1$ or $T = 0$.

Both physician's participation and treatment selection are modeled following a continuous latent variable approach and indicator function. And equations for probability of deciding on learning RAS and probability of deciding on using RAS are estimated simultaneously. Below we show the two-stage treatment selection equations:

$$\begin{aligned}
&Stage\ 1: \\
&\qquad P^*_{\mathrm{it}} = Z_{\mathrm{it}}\gamma + \lambda_1\eta_{\mathrm{it}1} + \epsilon'_{\mathrm{it}}\ ,\ P_{\mathrm{it}} = \mathbb{1}(P^*_{\mathrm{it}} > 0) \\
&Stage\ 2: \\
&\qquad T^*_{\mathrm{ij}t} = (R_{\mathrm{ij}t}\theta + \lambda_2\eta_{\mathrm{ij}t2} + \epsilon''_{\mathrm{ij}t})P_{\mathrm{it}}\ ,\ T_{\mathrm{ijt}} = \mathbb{1}(T^*_{\mathrm{ij}t} > 0)
\end{aligned} \tag{4.1}$$

where vector Z_{i} represents a set of explanatory variables of physician i, including physician work experience, medical school rank, physician gender, physician race, the constant term; and vector $R_{\mathrm{ij}t}$ represents a set of explanatory variables of patient j treated by physician i at time t, including patient age, patient gender, patient race, risk index, insurance type, year fixed time, and the constant term; $\eta_{\mathrm{i}t1}$ and $\eta_{\mathrm{ij}t2}$ are latent factors representing unobserved shared individual heterogeneity in which $\eta_{\mathrm{it}1} \sim \mathcal{N}(0, \sigma_1^2)$ and $\eta_{\mathrm{ij}t2} \sim \mathcal{N}(0, \sigma_2^2)$. For ease of computation, we restrict $\sigma_1^2 = \sigma_2^2 = 1$. And ϵ'_{i} and $\epsilon''_{\mathrm{ij}t}$ are idiosyncratic error terms.

4.4.2 Clinical Outcomes with Latent Factor

We further model two dimensions of clinical outcomes, after-surgery mortality and medical expenditure. We postulate that RAS treatment affect clinical outcomes and outcomes across patient race, with the basic assumption that each dimension of clinical outcome is a function of endogenous treatment method, patient characteristics, physician fixed effect, and time fixed effect. In particular, functions are various for different clinical measurements. Specifically, we assume that patient mortality indicator is generated according to conditional Logit model shown as follows:

$$Pr(Mortality_{\mathrm{ijt}} = 1 \mid X_{\mathrm{ijt}}, \beta_1, T_{\mathrm{ijt}}, \eta_{\mathrm{ijt2}}) = Pr(X_{\mathrm{ij}t}\beta_1 + \delta_1 T_{\mathrm{ijt}} + \eta_{\mathrm{ijt2}} + \epsilon_{\mathrm{ijt}} > 0) \tag{4.2}$$

$$= \frac{\exp(X_{\mathrm{ij}t}\beta_1 + \delta_1 T_{\mathrm{ijt}} + \eta_{\mathrm{ijt2}})}{1 + \exp(X_{\mathrm{ijt}}\beta_1 + \delta_1 T_{\mathrm{ijt}} + \eta_{\mathrm{ijt2}})} \tag{4.3}$$

where $Mortality_{\mathrm{ij}t}$ is a binary variable that equals zero if patient j served by physician i at time t is alive after the treatment and equals to one otherwise; vector $X_{\mathrm{ij}t}$ represents explanatory variables for patient j served by physician i at time t, including patient age, patient gender, patient race, patient risk index, year-quarter fixed effect, physician fixed effect, hospital fixed effect, and the constant term; σ_3 denotes the variance of logarithm patient expenditure; and $\eta_{\mathrm{ij}t2}$ is latent factor representing unobserved individual heterogeneity sharing with Equation 2.2.

We assume the clinical outcomes are independent conditional on determined covariates and coefficients. Given the probability of categorical clinical outcome, we list the log likelihood function as follows:

$$\begin{aligned} llk(\beta,\theta,\gamma,\delta,\lambda) &= \sum_{\mathrm{j}} \sum_{p\in\{0,1\}} \sum_{t\in\{0,1\}} \sum_{y\in\{0,1\}} \mathbb{1}_{P=p,T=t,Y=y} \\ ln &\left[\int Pr(p \mid \eta_1) Pr(t \mid \eta_2) Pr(Mortality \mid X, T, \eta_2) f(\eta_1) f(\eta_2) d\eta_1 d\eta_2 \right] \end{aligned} \tag{4.4}$$

The estimation can be obtained by maximizing the log-likelihood function. And we apply a maximum simulated likelihood (MSL) method to achieve it [211].

4.5 Empirical Results

The general estimation scheme follows the MSL estimation method. Specifically, we apply MSL to recover the parameters numerically by using a Newton-type algorithm.

We first show the estimation results of physician's decision-making process in Table 4.3 and Table 4.4. Recall a physician facing the two-stage decisions in which she decides on whether to participate in RAS learning in each time period (if she has not learned RAS yet), then decides on whether to use RAS only if she was trained before. In Table 4.3, the coefficients of physician characteristics estimate the RAS learning preferences. Focusing on race factors, the estimates show that compared to White physicians, physicians in Black are less likely to get RAS training. Similarly, physician in other minority races are less likely to be trained and have the ability of using RAS either. Both imply the racial disparity in learning RAS techniques.

Additionally, in Table 4.4, the coefficients of patient characteristics estimate the RAS using preferences when the insurance type and physician fixed effect are controlled. Specifically,

Table 4.3. Decision-Making Process regarding RAS Learning

	RAS Learning	
VARIABLES	(1)	(2)
Teaching Hospital	0.7021***	0.2682***
	(0.3779)	(0.1774)
Work Experience	-0.4097***	-0.9349***
	(0.0317)	(0.0899)
School Rank	0.3925***	0.5002***
	(0.2297)	(0.3602)
Physician - Black	-2.5325***	-2.1924***
	(1.0391)	(2.008)
Physician - Others	-6.6142***	-3.2478***
	(2.7872)	(2.5506)
Physician - Male	1.6822***	2.3318***
	(1.7421)	(1.4706)
Constant	-1.9199***	1.9070***
	(0.1811)	(0.8914)
Year, Quarter Fixed Effect	Yes	Yes
Apply for RAS-Equipped Hospital	No	Yes
Observations	2,164	1,971

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

the insignificance of patient race in RAS using decision-making process show that patients' race won't affect whether they are treated by RAS or not. Taken together, our results imply that even though we find racial bias in terms of physician's race when learning RAS method, patients are indeed treated equally across race during treatment method selection.

We then report the estimation results of clinical outcomes with individual heterogeneity in Table 4.5. We show apparent evidence for RAS improves clinical performance in terms of reducing the in-hospital mortality rate of patients. To be specified, the coefficient of RAS on mortality is significantly negative, which indicates the implementation of RAS compared to traditional treatment leads to lower mortality rate after surgery. Remarkably, the results

Table 4.4. Decision-Making Process regarding RAS Using

	RAS Using	
VARIABLES	(1)	(2)
Charlindex	-0.7476***	-0.1684***
	(0.2522)	(0.0636)
Patient Age	-0.0454***	-0.0641***
	(0.0048)	(0.0261)
Patient - Black	0.1551	-0.3746
	(0.3910)	(0.5699)
Patient - Others	0.1202	-0.2660
	(0.3668)	(0.3706)
Patient - Male	-0.1537	-0.1475
	(0.3047)	(0.3516)
Constant	0.3357***	0.7594***
	(0.1283)	(0.2758)
Insurance Dummy	Yes	Yes
Physician Fixed Effect	Yes	Yes
Hospital Fixed Effect	Yes	Yes
Year, Quarter Fixed Effect	Yes	Yes
Apply for RAS-Equipped Hospital	No	Yes
Observations	114,120	90,081

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

also reveal RAS function of alleviating health racial inequality. In particular, the significant effect of patient in Black and other races shows salient racial inequality in scarcity of learning the RAS treatment method. Interestingly, during the RAS using stage, we observe the insignificant disparity in terms of applying RAS over difference patient races, which indicates the differences of clinical improvement across patient races turn to be indistinguishable once physician have accessed to RAS method. To conclude, the estimations are consistent with the intuition from RAS usage, suggesting that implementation of RAS is inclined to causally improve the practical outcomes and tends to be applied to patient with different

Table 4.5. Latent Factor Model Estimation of RAS Effect on Mortality

	In-Hospital Mortality	
VARIABLES	(1)	(2)
RAS	-0.0246***	-0.0469***
	(0.0112)	(0.0251)
Charlindex	0.0529***	0.0249***
	(0.0027)	(0.0055)
Patient Age	0.0024***	0.0032***
	(0.0009)	(0.0018)
Patient - Male	0.0008**	0.0029*
	(0.0032)	(0.0059)
Patient - Black	-0.0070	0.0020
	(0.0084)	(0.0167)
Patient - Others	-0.0025	-0.0113
	(0.0030)	(0.0218)
Constant	-0.1416***	9.1899***
	(0.0265)	(0.2185)
Insurance Dummy	Yes	Yes
Physician Fixed Effect	Yes	Yes
Hospital Fixed Effect	Yes	Yes
Year, Quarter Fixed Effect	Yes	Yes
Apply for RAS-Equipped Hospital	No	Yes
Observations	114,120	90,081

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

races equally. This finding remarkably uncovers one of unique benefits of RAS application. Taking advantage of this effect and promoting RAS implementation might decrease healthcare disparity and generate more equal healthcare environment.

Columns (1) and (3) of Table 4.6 present the effect of RAS usage on degree of standardization in patient-level health outcomes, measured by the focal-normalized deviation. Columns (2) and (4) of Table 4.6 present the effect of RAS usage on physician-level clinical performance measured by the coefficient of variation aggregated on each physician. We measure the treatment outcomes in two aspects. First, whether the patient is alive after the treatment, which is the most common measurement for clinical performance. In addition, considering the extremely low mortality rate, we also apply the indicator of whether

Table 4.6. RAS Effect on Variation in Treatment Outcomes

	Variation in Patient Health Outcomes			
VARIABLES	Mortality		Routine Discharge	
	(1)	(2)	(3)	(4)
RAS	-0.504***	-0.564***	-0.0461***	-0.104***
	(0.104)	(0.188)	(0.0148)	(0.0333)
Charlindex	0.122***		0.00586***	
	(0.00997)		(0.00164)	
Patient Age	0.0109***		0.0008***	
	(0.000924)		(0.000289)	
Patient - Male	0.101***		0.00415	
	(0.0256)		(0.00445)	
Patient Characteristics	Yes	No	Yes	No
Physician Fixed Effect	Yes	Yes	Yes	Yes
Hospital Fixed Effect	Yes	Yes	Yes	Yes
Year, Quarter Fixed Effect	Yes	Yes	Yes	Yes
Observations	114,120	25,599	114,120	25,599

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

the patient is directly discharged back to home, rather than be sent to other healthcare facilities after the treatment. And this routine discharge indicator is a better and milder index suggesting the clinical outcomes. The results show that, no matter we measure the clinical outcomes by subject indicator or mild indicator, the RAS method not only increases the clinical performance, but also shows its potentials to reduce the variations in clinical outcomes. In other words, the patients treated by RAS are more likely to have the consistent surgical results regardless of their characteristics, compared to those treated by the traditional surgery method.

4.6 Mechanism Exploration

To further understand how RAS implementation induces the enhancement of clinical delivery, we explore and test the alternative mechanisms that might explain our findings regarding RAS usage. For example, RAS tends to show the potential of enhancing standardization in the treatment procedures, which might act as moderators to further gradually transform the health service delivery process and achieve better health outcomes.

To investigate the possible impact of RAS method in standardizing the treatment procedures, we measure the degree of standardization by the negative coefficient of variation aggregated on physician level. The larger degree of standardization is, the larger similarity of procedures have.

Table 4.7. RAS Effect on Physician-Level Procedure Standardization

	Standardization in Procedures			
VARIABLES	(1)	(2)	(3)	(4)
	Waiting Time	Charges	Procedures Num.	Categories Num.
RAS	0.483***	0.0688***	0.682***	0.995***
	(0.0233)	(0.0061)	(0.0286)	(0.0493)
Physician Fixed Effect	Yes	Yes	Yes	Yes
Hospital Fixed Effect	Yes	Yes	Yes	Yes
Year, Quarter Fixed Effect	Yes	Yes	Yes	Yes
RAS-Equipped Hospital	Yes	Yes	Yes	Yes
Observations	18,620	18,620	18,620	18,620

Standard errors in parentheses
*** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$

The standardization of treatment procedures are captured in the following four perspectives: (1) Waiting time indicates the time elapsed from patient being admitted to receiving the surgery, which suggests the time spent for surgery preparation; (2) Lab test charges are applied for suggesting the conduction of laboratory tests attached to the principal surgery, including the preparation and recovery period; (3) We count the total number of recorded procedures through the ICD-10 procedure code, which is capable for describing the detailed steps associated with the entire treatment; (4) In order to further depict categorical dimension of procedure steps, we measure the number of procedure categories. Unlike the number of procedures that indicating the fineness of treatment steps, the categorical number denotes the broadness of treatment of steps.

Columns (1)-(4) of Table 4.7 present the effect of RAS usage on degree of standardization in treatment procedures, measured by the negative coefficient of variation. The results show that being treated by RAS method are associated with high degree of standardization in all above four dimensions of treatment procedures. In other words, those patients being treated by RAS are positively correlated with more uniform procedures in terms of time control, associated lab tests, detailed procedures, and procedure categories. These results support our conjecture that RAS improves the standardization when delivering the healthcare service.

4.7 Conclusions and Managerial Insights

In this study, we reveal the impact of RAS on healthcare deliveries by modeling physicians decision-making processes when learning and using the RAS method. Our findings extend the understanding of the RAS effect and implementation. Specifically, we show RAS not only increases clinical performance but also alleviates health service inequality and standardize the service delivery process compared with receiving traditional treatment. This work contributes to the stream of literature by adding evidence of the effectiveness of RAS in clinical performance. Then, we are the first to uncover RASs effect on alleviating racial health bias. Moreover, this study advances the understanding of RAS implementation by identifying the physicians decision-making process of treatment selection, which facilitates managerial implication to further RAS promotion.

This study is still going on. And our next plan focuses on demonstrating the managerial insights of how we promote RAS through accessibility on physician learning level and patient using level. In addition to promotion perspective, we are interest in developing and conducting the counterfactual analysis in order to further interrupted how we could better apply RAS in terms of precisely identify target patient and how much we will be benefited from RAS implementations.

5. CONCLUSIONS AND FUTURE WORK

This dissertation focuses on the value and utilization of information systems and information technologies in transforming healthcare delivery. In the first study, NLP techniques were leveraged to analyze physicians' personal statements from a healthcare platform. The objective was to investigate the impact of personality traits extracted through NLP on physicians' medical behaviors even the outcomes. The study uncovers that the personalities of service providers have a significant influence on their clinical performance. Specifically, using Florida discharge data from 2010 to 2018, we find that physicians with high openness scores is more likely to have patients with lower mortality rates, lower charges for lab tests, and shorter LOS by leveraging advanced treatment techniques. Patients being treated by conscientious physicians tend to have a more considerable expenditure on lab tests. By contrast, patients with agreeable physicians tend to have lower medical expenses for lab tests. All these findings are consistent with the personality traits developed in theory and thus validate the personality measures generated by this NPL method. To be emphasized, we also uncover possible mechanisms by which these traits influence patients' health outcomes. Our analysis of physicians' personality impacts on clinical performance yields several important insights for hospital administrators. For instance, hospitals can improve patient outcomes by considering physician personality traits in scheduling. Moreover, from the perspective of benefiting the public, physician-generated content on a large scale can be utilized to detect service providers' personalities and trained physicians to promote strength in characters and minimize mismatches between patients and physicians.

The second study investigates the resilience of digital goods consumption to COVID-19 pandemic and identifies physical mobility as a causal factor affecting digital goods consumption. Applying a natural, strict, and exogenous variation in individual-level physical mobility generated by quarantine policy in China, we systematically examine whether digital goods consumption in a comprehensive spectrum is resilient to individual-level physical restriction. The generalized results in terms of the marginal effect of physical mobility are derived, which suggest the substitute effect between an individuals physical activities and digital consumption. The estimations are consistent in various measures of app usage. As a byproduct, the impacts of trending quarantine policy in the immediate and long period are estimated. The results on immediate effect reveal the power of digital resilience to interventions, covering a broad range of digital goods, which contains but not limited to online social activities, digital reading, and e-commerce, among others. Most importantly, the long-term effects of the quarantine policy indicate the persistent impact on app usage even after the user is physically released from quarantine, enabling further mining of the apps economic value in the post-pandemic period.

In the third essay, we reveal the impact of RAS on health service deliveries by modeling physicians decision-making processes when learning and using the RAS method. Our findings extend the comprehension of the RAS effects and implementations. Specifically, we show RAS not only increases clinical performance but also alleviates healthcare inequality and variation of health service deliveries compared with receiving traditional treatment. To understand how RAS transforms health service delivery and largely improves the clinical performance and reduces the service variation, we explored the underlying mechanisms from the perspective of standardization in treatment time and procedures that help explain the transformation. This work adds to the literature by offering empirical supports of the effectiveness of RAS in clinical performance and influencing the understanding of RAS implementation by identifying the physicians decision-making process of treatment selection, which facilitates managerial implication to further RAS promotion. As the work is still going on, the counterfactual analysis will be proposed and conducted to further interrupted how we could better apply RAS in terms of precisely identify target patient and how much we will be benefited from RAS implementations.

REFERENCES

[1] Y.-K. Lin, M. Lin, and H. Chen, "Do electronic health records affect quality of care? evidence from the hitech act," *Information Systems Research*, vol. 30, no. 1, pp. 306–318, 2019.

[2] S. Sun, S. F. Lu, and H. Rui, "Does telemedicine reduce emergency room congestion? evidence from new york state," *Information Systems Research*, vol. 31, no. 3, pp. 972–986, 2020.

[3] R. B. Gala, R. Margulies, A. Steinberg, *et al.*, "Systematic review of robotic surgery in gynecology: Robotic techniques compared with laparoscopy and laparotomy," *Journal of minimally invasive gynecology*, vol. 21, no. 3, pp. 353–361, 2014.

[4] S. Razdan and S. Sharma, "Internet of medical things (iomt): Overview, emerging technologies, and case studies," *IETE Technical Review*, pp. 1–14, 2021.

[5] C. Chu, P. Cram, A. Pang, V. Stamenova, M. Tadrous, R. S. Bhatia, *et al.*, "Rural telemedicine use before and during the covid-19 pandemic: Repeated cross-sectional study," *Journal of medical Internet research*, vol. 23, no. 4, e26960, 2021.

[6] E. H. Hwang, X. Guo, Y. Tan, and Y. Dang, "Delivering healthcare through teleconsultations: Implications for offline healthcare disparity," *Information Systems Research*, 2022.

[7] S. F. Lu and H. Rui, "Can we trust online physician ratings? evidence from cardiac surgeons in florida," *Management Science*, vol. 64, no. 6, pp. 2557–2573, 2018.

[8] S. Khurana, L. Qiu, and S. Kumar, "When a doctor knows, it shows: An empirical analysis of doctors responses in a q&a forum of an online healthcare portal," *Information Systems Research*, vol. 30, no. 3, pp. 872–891, 2019.

[9] A. Urbaczewski and Y. J. Lee, "Information technology anev the pandemic: A preliminary multinational analysis of the impact of mobile tracking technology on the covid-19 contagion control," *European Journal of Information Systems*, vol. 29, no. 4, pp. 405–414, 2020.

[10] E. Kweilin, G. Rahul, and S. Julian, "Building the vital skills for the future of work in operations," *Healthcare System & Services*, vol. McKinsey & Company, August, 2020, 2020.

[11] O. P. John, S. Srivastava, *et al.*, "The big five trait taxonomy: History, measurement, and theoretical perspectives," *Handbook of personality: Theory and research*, vol. 2, no. 1999, pp. 102–138, 1999.

[12] T. A. Judge and J. D. Kammeyer-Mueller, "Personality and career success," *Handbook of career studies*, no. 9, p. 59, 2007.

[13] S. E. Seibert and M. L. Kraimer, "The five-factor model of personality and career success," *Journal of vocational behavior*, vol. 58, no. 1, pp. 1–21, 2001.

[14] N. Li, M. R. Barrick, R. D. Zimmerman, and D. S. Chiaburu, "Retaining the productive employee: The role of personality," *Academy of Management Annals*, vol. 8, no. 1, pp. 347–395, 2014.

[15] W. Duan, B. Gu, and A. B. Whinston, "Do online reviews matter?an empirical investigation of panel data," *Decision support systems*, vol. 45, no. 4, pp. 1007–1016, 2008.

[16] A. Ghose and P. G. Ipeirotis, "Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics," *IEEE transactions on knowledge and data engineering*, vol. 23, no. 10, pp. 1498–1512, 2010.

[17] P. K. Chintagunta, S. Gopinath, and S. Venkataraman, "The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets," *Marketing science*, vol. 29, no. 5, pp. 944–957, 2010.

[18] Y.-Y. Wang, C. Guo, A. Susarla, and V. Sambamurthy, "Online to offline: The impact of social media on offline sales in the automobile industry," *Information Systems Research*, vol. 32, no. 2, pp. 582–604, 2021.

[19] G. M. Smith, "Usefulness of peer ratings of personality in educational research," *Educational and Psychological measurement*, vol. 27, no. 4, pp. 967–984, 1967.

[20] J. W. Pennebaker and L. A. King, "Linguistic styles: Language use as an individual difference.," *Journal of personality and social psychology*, vol. 77, no. 6, p. 1296, 1999.

[21] A. J. Gill and J. Oberlander, "Taking care of the linguistic features of extraversion," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 24, 2002.

[22] J. Golbeck, C. Robles, M. Edmondson, and K. Turner, "Predicting personality from twitter," in *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*, IEEE, 2011, pp. 149–156.

[23] P.-H. Arnoux, A. Xu, N. Boyette, J. Mahmud, R. Akkiraju, and V. Sinha, "25 tweets to know you: A new model to predict personality with social media," in *Eleventh international AAAI conference on web and social media*, 2017.

[24] S. tajner, S. Yenikent, and M. Franco-Salvador, "Five psycholinguistic characteristics for better interaction with users," in *2021 8th International Conference on Behavioral and Social Computing (BESC)*, IEEE, 2021, pp. 1–7.

[25] P. Adamopoulos, A. Ghose, and V. Todri, "The impact of user personality traits on word of mouth: Text-mining social media platforms," *Information Systems Research*, vol. 29, no. 3, pp. 612–640, 2018.

[26] D. S. Kc and C. Terwiesch, "Impact of workload on service time and patient safety: An econometric analysis of hospital operations," *Management science*, vol. 55, no. 9, pp. 1486–1498, 2009.

[27] A. Powell, S. Savin, and N. Savva, "Physician workload and hospital reimbursement: Overworked physicians generate less revenue per patient," *Manufacturing & Service Operations Management*, vol. 14, no. 4, pp. 512–528, 2012.

[28] L. Kuntz, R. Mennicken, and S. Scholtes, "Stress on the ward: Evidence of safety tipping points in hospitals," *Management Science*, vol. 61, no. 4, pp. 754–771, 2015.

[29] J. A. Berry Jaeker and A. L. Tucker, "Past the point of speeding up: The negative effects of workload saturation on efficiency and patient severity," *Management Science*, vol. 63, no. 4, pp. 1042–1062, 2017.

[30] M. Freeman, N. Savva, and S. Scholtes, "Gatekeepers at work: An empirical analysis of a maternity unit," *Management Science*, vol. 63, no. 10, pp. 3147–3167, 2017.

[31] C. W. Chan, L. V. Green, S. Lekwijit, L. Lu, and G. Escobar, "Assessing the impact of service level when customer needs are uncertain: An empirical investigation of hospital step-down units," *Management Science*, vol. 65, no. 2, pp. 751–775, 2019.

[32] A. D. Shukla, G. Gao, and R. Agarwal, "How digital word-of-mouth affects consumer decision making: Evidence from doctor appointment booking," *Management Science*, vol. Forthcoming, Forthcoming, 2020.

[33] H. E. Cattell and A. D. Mead, "The sixteen personality factor questionnaire (16pf)," *The SAGE handbook of personality theory and assessment*, vol. 2, pp. 135–178, 2008.

[34] H. J. Eysenck and S. B. G. Eysenck, *Manual of the Eysenck Personality Scales (EPS Adult): Comprising the EPQ-Revised (EPQ-R), EPQ-R Short Scale, Impulsiveness (IVE) Questionnaire...* Hodder & Stoughton, 1991.

[35] I. Myers, M. McCaulley, N. Quenk, and A. Hammer, *Mbti handbook: A guide to the development and use of the myers–briggs type indicator (3rd edn). sunnyvale*, 1998.

[36] C. R. Cloninger, T. R. Przybeck, and D. M. Svrakic, "The tridimensional personality questionnaire: Us normative data," *Psychological reports*, vol. 69, no. 3, pp. 1047–1057, 1991.

[37] C. J. Patrick, J. J. Curtin, and A. Tellegen, "Development and validation of a brief form of the multidimensional personality questionnaire.," *Psychological assessment*, vol. 14, no. 2, p. 150, 2002.

[38] T. Yarkoni, "Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers," *Journal of research in personality*, vol. 44, no. 3, pp. 363–373, 2010.

[39] W. Youyou, M. Kosinski, and D. Stillwell, "Computer-based personality judgments are more accurate than those made by humans," *Proceedings of the National Academy of Sciences*, vol. 112, no. 4, pp. 1036–1040, 2015.

[40] J. S. Harrison, G. R. Thurgood, S. Boivie, and M. D. Pfarrer, "Measuring ceo personality: Developing, validating, and testing a linguistic tool," *Strategic Management Journal*, vol. 40, no. 8, pp. 1316–1330, 2019.

[41] T. A. Judge, C. A. Higgins, C. J. Thoresen, and M. R. Barrick, "The big five personality traits, general mental ability, and career success across the life span," *Personnel psychology*, vol. 52, no. 3, pp. 621–652, 1999.

[42] J. A. LePine, "Team adaptation and postchange performance: Effects of team composition in terms of members' cognitive ability and personality.," *Journal of applied psychology*, vol. 88, no. 1, p. 27, 2003.

[43] H. Zhao and S. E. Seibert, "The big five personality dimensions and entrepreneurial status: A meta-analytical review.," *Journal of applied psychology*, vol. 91, no. 2, p. 259, 2006.

[44] M. Bulmer, J. R. Bohnke, and G. J. Lewis, "Predicting moral sentiment towards physician-assisted suicide: The role of religion, conservatism, authoritarianism, and big five personality," *Personality and Individual Differences*, vol. 105, pp. 244–251, 2017.

[45] S. Mullola, C. Hakulinen, J. Presseau, *et al.*, "Personality traits and career choices among physicians in finland: Employment sector, clinical patient contact, specialty and change of specialty," *BMC medical education*, vol. 18, no. 1, p. 52, 2018.

[46] R. A. Scheepers, O. A. Arah, M. J. Heineman, and K. M. Lombarts, "How personality traits affect clinician-supervisors work engagement and subsequently their teaching performance in residency training," *Medical teacher*, vol. 38, no. 11, pp. 1105–1111, 2016.

[47] C.-I. Teng, K.-H. Hsu, R.-C. Chien, and H.-Y. Chang, "Influence of personality on care quality of hospital nurses," *Journal of nursing care quality*, vol. 22, no. 4, pp. 358–364, 2007.

[48] D. Phillips, K. A. Egol, M. C. Maculatis, *et al.*, "Personality factors associated with resident performance: Results from 12 accreditation council for graduate medical education accredited orthopaedic surgery programs," *Journal of surgical education*, vol. 75, no. 1, pp. 122–131, 2018.

[49] A. Gur, D. Weimann Saks, and L. Stavi, "Healthcare employee's personality traits and service quality: Do team relationships matter?" *European Journal of Public Health*, vol. 30, no. Supplement_5, ckaa165–310, 2020.

[50] S. J. Dillon, W. Kleinmann, A. Seasely, *et al.*, "How personality affects teamwork: A study in multidisciplinary obstetrical simulation," *American journal of obstetrics & gynecology MFM*, vol. 3, no. 2, p. 100 303, 2021.

[51] A. Ghose and P. G. Ipeirotis, "Designing ranking systems for consumer reviews: The impact of review subjectivity on product sales and review quality," in *Proceedings of the 16th annual workshop on information technology and systems*, 2006, pp. 303–310.

[52] M. M. Susan and S. David, "What makes a helpful online review? a study of customer reviews on amazon. com," *MIS Quarterly*, vol. 34, no. 1, pp. 185–200, 2010.

[53] D. H. Park and S. Kim, "The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews," *Electronic commerce research and applications*, vol. 7, no. 4, pp. 399–410, 2008.

[54] M. Luca and G. Zervas, "Fake it till you make it: Reputation, competition, and yelp review fraud," *Management Science*, vol. 62, no. 12, pp. 3412–3427, 2016.

[55] M. Ma and R. Agarwal, "Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities," *Information systems research*, vol. 18, no. 1, pp. 42–67, 2007.

[56] A. Cheema and A. M. Kaikati, "The effect of need for uniqueness on word of mouth," *Journal of Marketing research*, vol. 47, no. 3, pp. 553–563, 2010.

[57] P. B. Goes, M. Lin, and C.-m. Au Yeung, "popularity effect in user-generated content: Evidence from online product reviews," *Information Systems Research*, vol. 25, no. 2, pp. 222–238, 2014.

[58] W. Khern-am-nuai, K. Kannan, and H. Ghasemkhani, "Extrinsic versus intrinsic rewards for contributing reviews in an online platform," *Information Systems Research*, vol. 29, no. 4, pp. 871–892, 2018.

[59] Y. Huang, P. V. Singh, and A. Ghose, "A structural model of employee behavioral dynamics in enterprise social media," *Management Science*, vol. 61, no. 12, pp. 2825–2844, 2015.

[60] W. Shen, Y. (Hu, and J. R. Ulmer, "Competing for attention: An empirical study of online reviewers' strategic behavior.," *MIS Q.*, vol. 39, no. 3, pp. 683–696, 2015.

[61] Y. Liu and J. Feng, "Does money talk? the impact of monetary incentives on user-generated content contributions," *Information Systems Research*, vol. 32, no. 2, pp. 394–409, 2021.

[62] Y. Chen and J. Xie, "Online consumer review: Word-of-mouth as a new element of marketing communication mix," *Management science*, vol. 54, no. 3, pp. 477–491, 2008.

[63] C. Yi, Z. Jiang, X. Li, and X. Lu, "Leveraging user-generated content for product promotion: The effects of firm-highlighted reviews," *Information Systems Research*, vol. 30, no. 3, pp. 711–725, 2019.

[64] P. Albuquerque, P. Pavlidis, U. Chatow, K.-Y. Chen, and Z. Jamal, "Evaluating promotional activities in an online two-sided market of user-generated content," *Marketing Science*, vol. 31, no. 3, pp. 406–432, 2012.

[65] K.-Y. Goh, C.-S. Heng, and Z. Lin, "Social media brand community and consumer behavior: Quantifying the relative impact of user-and marketer-generated content," *Information systems research*, vol. 24, no. 1, pp. 88–107, 2013.

[66] S. Gong, J. Zhang, P. Zhao, and X. Jiang, "Tweeting as a marketing tool: A field experiment in the tv industry," *Journal of Marketing Research*, vol. 54, no. 6, pp. 833–850, 2017.

[67] T. Song, J. Huang, Y. Tan, and Y. Yu, "Using user-and marketer-generated content for box office revenue prediction: Differences between microblogging and third-party platforms," *Information Systems Research*, vol. 30, no. 1, pp. 191–203, 2019.

[68] K. Zhao, P. Zhang, and H.-M. Lee, "Understanding the impacts of user-and marketer-generated content on free digital content consumption," *Decision Support Systems*, vol. 154, p. 113 684, 2022.

[69] A. Pansari and V. Kumar, "Customer engagement: The construct, antecedents, and consequences," *Journal of the Academy of Marketing Science*, vol. 45, no. 3, pp. 294–311, 2017.

[70] M. Meire, K. Hewett, M. Ballings, V. Kumar, and D. Van den Poel, "The role of marketer-generated content in customer engagement marketing," *Journal of Marketing*, vol. 83, no. 6, pp. 21–42, 2019.

[71] H. Song, A. L. Tucker, and K. L. Murrell, "The diseconomies of queue pooling: An empirical investigation of emergency department length of stay," *Management Science*, vol. 61, no. 12, pp. 3032–3053, 2015.

[72] M. R. Ibanez, J. R. Clark, R. S. Huckman, and B. R. Staats, "Discretionary task ordering: Queue management in radiological services," *Management Science*, vol. 64, no. 9, pp. 4389–4407, 2018.

[73] C. Zacharias and T. Yunes, "Multimodularity in the stochastic appointment scheduling problem with discrete arrival epochs," *Management Science*, vol. 66, no. 2, pp. 744–763, 2020.

[74] S. F. Lu and L. X. Lu, “Do mandatory overtime laws improve quality? staffing decisions and operational flexibility of nursing homes,” *Management Science*, vol. 63, no. 11, pp. 3566–3585, 2017.

[75] T. F. Tan and S. Netessine, “When you work with a superman, will you also fly? an empirical study of the impact of coworkers on performance,” *Management Science*, vol. 65, no. 8, pp. 3495–3517, 2019.

[76] B. N. Greenwood, S. Carnahan, and L. Huang, “Patient–physician gender concordance and increased mortality among female heart attack patients,” *Proceedings of the National Academy of Sciences*, vol. 115, no. 34, pp. 8569–8574, 2018.

[77] B. R. Staats, D. S. Kc, and F. Gino, “Maintaining beliefs in the face of negative news: The moderating role of experience,” *Management Science*, vol. 64, no. 2, pp. 804–824, 2018.

[78] B. M. Gray, J. L. Vandergrift, G. G. Gao, J. S. McCullough, and R. S. Lipner, “Website ratings of physicians and their quality of care,” *JAMA internal medicine*, vol. 175, no. 2, pp. 291–293, 2015.

[79] D. H. Saifee, I. R. Bardhan, A. Lahiri, and Z. Zheng, “Adherence to clinical guidelines, electronic health record use, and online reviews,” *Journal of Management Information Systems*, vol. 36, no. 4, pp. 1071–1104, 2019.

[80] G. G. Gao, J. S. McCullough, R. Agarwal, and A. K. Jha, “A changing landscape of physician quality reporting: Analysis of patients online ratings of their physicians over a 5-year period,” *Journal of medical Internet research*, vol. 14, no. 1, e38, 2012.

[81] G. G. Gao, B. N. Greenwood, R. Agarwal, and J. McCullough, “Vocal minority and silent majority: How do online ratings reflect population perceptions of quality?” *MIS quarterly*, vol. 39, no. 3, pp. 565–589, 2015.

[82] D. Kahneman, *Thinking, fast and slow*. Macmillan, 2011.

[83] W. L. Kissick, *Medicine’s dilemmas: infinite needs versus finite resources*. Yale University Press, 1994.

[84] R. S. Huckman and G. P. Pisano, “The firm specificity of individual performance: Evidence from cardiac surgery,” *Management Science*, vol. 52, no. 4, pp. 473–488, 2006.

[85] P. W. Serruys, M.-C. Morice, A. P. Kappetein, *et al.*, "Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery disease," *New England journal of medicine*, vol. 360, no. 10, pp. 961–972, 2009.

[86] K. B. DeSalvo, T. M. Jones, J. Peabody, *et al.*, "Health care expenditure prediction with a single item, self-rated health measure," *Medical care*, vol. 47, no. 4, pp. 440–447, 2009.

[87] M. R. Barrick and M. K. Mount, "The big five personality dimensions and job performance: A meta-analysis," *Personnel psychology*, vol. 44, no. 1, pp. 1–26, 1991.

[88] M. R. Barrick, G. L. Stewart, and M. Piotrowski, "Personality and job performance: Test of the mediating effects of motivation among sales representatives.," *Journal of applied psychology*, vol. 87, no. 1, p. 43, 2002.

[89] J. E. Slaughter and E. E. Kausel, "The neurotic employee: Theoretical analysis of the influence of narrow facets of neuroticism on cognitive, social, and behavioral processes relevant to job performance," in *Research in personnel and human resources management*, Emerald Group Publishing Limited, 2009.

[90] J. M. Digman and N. K. Takemoto-Chock, "Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies," *Multivariate behavioral research*, vol. 16, no. 2, pp. 149–170, 1981.

[91] O. Nov and C. Ye, "Personality and technology acceptance: Personal innovativeness in it, openness and resistance to change," in *Proceedings of the 41st annual Hawaii international conference on system sciences (HICSS 2008)*, IEEE, 2008, pp. 448–448.

[92] G. B. Svendsen, J.-A. K. Johnsen, L. Almås-Sørensen, and J. Vittersø, "Personality and technology acceptance: The influence of personality factors on the core constructs of the technology acceptance model," *Behaviour & Information Technology*, vol. 32, no. 4, pp. 323–334, 2013.

[93] S. Devaraj, R. F. Easley, and J. M. Crant, "Research notehow does personality matter? relating the five-factor model to technology acceptance and use," *Information systems research*, vol. 19, no. 1, pp. 93–105, 2008.

[94] A. Y.-L. Chong, M. J. Liu, J. Luo, and O. Keng-Boon, "Predicting rfid adoption in healthcare supply chain from the perspectives of users," *International Journal of Production Economics*, vol. 159, pp. 66–75, 2015.

[95] R. R. McCrae and P. T. Costa Jr, "Personality, coping, and coping effectiveness in an adult sample," *Journal of personality*, vol. 54, no. 2, pp. 385–404, 1986.

[96] S. Nadkarni and V. K. Narayanan, "Strategic schemas, strategic flexibility, and firm performance: The moderating role of industry clockspeed," *Strategic management journal*, vol. 28, no. 3, pp. 243–270, 2007.

[97] R. R. McCrae and O. P. John, "An introduction to the five-factor model and its applications," *Journal of personality*, vol. 60, no. 2, pp. 175–215, 1992.

[98] P. R. Shaver and K. A. Brennan, "Attachment styles and the" big five" personality traits: Their connections with each other and with romantic relationship outcomes," *Personality and Social Psychology Bulletin*, vol. 18, no. 5, pp. 536–545, 1992.

[99] M. K. Mount, M. R. Barrick, and G. L. Stewart, "Five-factor model of personality and performance in jobs involving interpersonal interactions," *Human performance*, vol. 11, no. 2-3, pp. 145–165, 1998.

[100] F. L. Schmidt and J. E. Hunter, "The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings.," *Psychological bulletin*, vol. 124, no. 2, p. 262, 1998.

[101] G. M. Hurtz and J. J. Donovan, "Personality and job performance: The big five revisited.," *Journal of applied psychology*, vol. 85, no. 6, p. 869, 2000.

[102] J. J. Jackson, D. Wood, T. Bogg, K. E. Walton, P. D. Harms, and B. W. Roberts, "What do conscientious people do? development and validation of the behavioral indicators of conscientiousness (bic)," *Journal of research in personality*, vol. 44, no. 4, pp. 501–511, 2010.

[103] P. Tough, *How children succeed: Grit, curiosity, and the hidden power of character*. Houghton Mifflin Harcourt, 2012.

[104] Z. Roberts, A. Rogers, C. L. Thomas, and C. Spitzmueller, "Effects of proactive personality and conscientiousness on training motivation," *International Journal of Training and Development*, vol. 22, no. 2, pp. 126–143, 2018.

[105] W. Kelly and J. Johnson, "Time use efficiency and the five-factor model of personality," *Education*, vol. 125, p. 511, Sep. 2005.

[106] M. D. Biderman, N. T. Nguyen, and J. Sebren, "Time-on-task mediates the conscientiousness–performance relationship," *Personality and Individual Differences*, vol. 44, no. 4, pp. 887–897, 2008.

[107] F. Tabak, N. Nguyen, T. Basuray, and W. Darrow, "Exploring the impact of personality on performance: How time-on-task moderates the mediation by self-efficacy," *Personality and Individual Differences*, vol. 47, no. 8, pp. 823–828, 2009.

[108] B. H. Bradley, J. E. Baur, C. G. Banford, and B. E. Postlethwaite, "Team players and collective performance: How agreeableness affects team performance over time," *Small Group Research*, vol. 44, no. 6, pp. 680–711, 2013.

[109] D. I. Cohen, P. Jones, B. Littenberg, and D. Neuhauser, "Does cost information availability reduce physician test usage?: A randomized clinical trial with unexpected findings," *Medical care*, pp. 286–292, 1982.

[110] Y. Song and M. Shi, "Associations between empathy and big five personality traits among chinese undergraduate medical students," *PloS one*, vol. 12, no. 2, e0171665, 2017.

[111] A. P. Bartel, C. W. Chan, and S.-H. Kim, "Should hospitals keep their patients longer? the role of inpatient care in reducing postdischarge mortality," *Management Science*, vol. 66, no. 6, pp. 2326–2346, 2020. DOI: 10.1287/mnsc.2019.3325.

[112] R. J. Batt and C. Terwiesch, "Early task initiation and other load-adaptive mechanisms in the emergency department," *Management Science*, vol. 63, no. 11, pp. 3531–3551, 2017.

[113] D. Dranove, D. Kessler, M. McClellan, and M. Satterthwaite, "Is more information better? the effects of report cards on health care providers," *Journal of political Economy*, vol. 111, no. 3, pp. 555–588, 2003.

[114] L. X. Lu and S. F. Lu, "Distance, quality, or relationship? interhospital transfer of heart attack patients," *Production and Operations Management*, vol. 27, no. 12, pp. 2251–2269, 2018.

[115] L. A. Fast and D. C. Funder, "Personality as manifest in word use: Correlations with self-report, acquaintance report, and behavior.," *Journal of personality and social psychology*, vol. 94, no. 2, p. 334, 2008.

[116] J. B. Hirsh and J. B. Peterson, “Personality and language use in self-narratives,” *Journal of research in personality*, vol. 43, no. 3, pp. 524–527, 2009.

[117] J. L. Baddeley and J. A. Singer, “Telling losses: Personality correlates and functions of bereavement narratives,” *Journal of Research in Personality*, vol. 42, no. 2, pp. 421–438, 2008.

[118] J. W. Pennebaker, M. R. Mehl, and K. G. Niederhoffer, “Psychological aspects of natural language use: Our words, our selves,” *Annual review of psychology*, vol. 54, no. 1, pp. 547–577, 2003.

[119] R. H. Keogh and I. R. White, “A toolkit for measurement error correction, with a focus on nutritional epidemiology,” *Statistics in medicine*, vol. 33, no. 12, pp. 2137–2155, 2014.

[120] B. Rosner, D. Spiegelman, and W. Willett, “Correction of logistic regression relative risk estimates and confidence intervals for measurement error: The case of multiple covariates measured with error,” *American journal of epidemiology*, vol. 132, no. 4, pp. 734–745, 1990.

[121] K. E. Keeton, *An extension of the UTAUT model: How organizational factors and individual differences influence technology acceptance*. University of Houston, 2008.

[122] W. G. Graziano, M. M. Habashi, B. E. Sheese, and R. M. Tobin, “Agreeableness, empathy, and helping: A person× situation perspective.,” *Journal of personality and social psychology*, vol. 93, no. 4, p. 583, 2007.

[123] A. Jain, T. Dai, K. Bibee, and C. Myers, “Covid-19 created an elective surgery backlog. how can hospitals get back on track?” *Harvard Business Review*, 2020.

[124] J. W. Pennebaker, R. J. Booth, and M. E. Francis, “Linguistic inquiry and word count: Liwc,” *Austin, TX: liwc. net*, vol. 135, p. 2007, 2007.

[125] Y. R. Tausczik and J. W. Pennebaker, “The psychological meaning of words: Liwc and computerized text analysis methods,” *Journal of language and social psychology*, vol. 29, no. 1, pp. 24–54, 2010.

[126] P. T. Costa and R. R. McCrea, *Revised neo personality inventory (neo pi-r) and neo five-factor inventory (neo-ffi)*. Psychological Assessment Resources, 1992.

[127] V. Berglund, I. Johansson Sevä, and M. Strandh, "Subjective well-being and job satisfaction among self-employed and regular employees: Does personality matter differently?" *Journal of Small Business & Entrepreneurship*, vol. 28, no. 1, pp. 55–73, 2016.

[128] J. Camps, J. Stouten, and M. Euwema, "The relation between supervisors big five personality traits and employees experiences of abusive supervision," *Frontiers in psychology*, vol. 7, p. 112, 2016.

[129] A. B. Bakker, K. I. Van Der Zee, K. A. Lewig, and M. F. Dollard, "The relationship between the big five personality factors and burnout: A study among volunteer counselors," *The Journal of social psychology*, vol. 146, no. 1, pp. 31–50, 2006.

[130] D. W. Organ and A. Lingl, "Personality, satisfaction, and organizational citizenship behavior," *The journal of social psychology*, vol. 135, no. 3, pp. 339–350, 1995.

[131] P. T. Costa Jr, R. R. McCrae, and D. A. Dye, "Facet scales for agreeableness and conscientiousness: A revision of the neo personality inventory," *Personality and individual Differences*, vol. 12, no. 9, pp. 887–898, 1991.

[132] C. Therasa and C. Vijayabanu, "The impact of big five personality traits and positive psychological strengths towards job satisfaction: A review," *Periodica Polytechnica Social and Management Sciences*, vol. 23, no. 2, pp. 142–150, 2015.

[133] L. Lucchini, S. Centellegher, L. Pappalardo, *et al.*, "Living in a pandemic: Changes in mobility routines, social activity and adherence to covid-19 protective measures," *Scientific reports*, vol. 11, no. 1, pp. 1–12, 2021.

[134] G. Alekseev, S. Amer, M. Gopal, *et al.*, "The effects of covid-19 on us small businesses: Evidence from owners, managers, and employees," *Management Science*, 2022.

[135] B. R. Han, T. Sun, L. Y. Chu, and L. Wu, "Covid-19 and e-commerce operations: Evidence from alibaba," *Manufacturing & Service Operations Management*, 2022.

[136] L. C. Ming, N. Untong, N. A. Aliudin, *et al.*, "Mobile health apps on covid-19 launched in the early days of the pandemic: Content analysis and review," *JMIR mHealth and uHealth*, vol. 8, no. 9, e19796, 2020.

[137] E. Peltonen, E. Lagerspetz, J. Hamberg, *et al.*, "The hidden image of mobile apps: Geographic, demographic, and cultural factors in mobile usage," in *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 2018, pp. 1–12.

[138] D. Yu, Y. Li, F. Xu, P. Zhang, and V. Kostakos, "Smartphone app usage prediction using points of interest," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 4, pp. 1–21, 2018.

[139] M. De Nadai, A. Cardoso, A. Lima, B. Lepri, and N. Oliver, "Apps, places and people: Strategies, limitations and trade-offs in the physical and digital worlds," *arXiv preprint arXiv:1904.09350*, 2019.

[140] C. Shin, J.-H. Hong, and A. K. Dey, "Understanding and prediction of mobile application usage for smart phones," in *proceedings of the 2012 ACM conference on ubiquitous computing*, 2012, pp. 173–182.

[141] K. Zhu, L. Zhang, and A. Pattavina, "Learning geographical and mobility factors for mobile application recommendation," *IEEE Intelligent Systems*, vol. 32, no. 3, pp. 36–44, 2017.

[142] Z. Lu, Y. Feng, W. Zhou, X. Li, and Q. Cao, "Inferring correlation between user mobility and app usage in massive coarse-grained data traces," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 4, pp. 1–21, 2018.

[143] C. E. Tucker and S. Yu, "The early effects of coronavirus-related social distancing restrictions on brands," *Available at SSRN 3566612*, 2020.

[144] E. L. Glaeser, G. Z. Jin, B. T. Leyden, and M. Luca, "Learning from deregulation: The asymmetric impact of lockdown and reopening on risky behavior during covid-19," *Journal of regional science*, vol. 61, no. 4, pp. 696–709, 2021.

[145] A. Goldfarb and C. Tucker, "Which retail outlets generate the most physical interactions?" National Bureau of Economic Research, Tech. Rep., 2020.

[146] C. Prentice, L. Altinay, and A. G. Woodside, *Transformative service research and covid-19*, 2021.

[147] D. Ivanov, "Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (covid-19/sars-cov-2) case," *Transportation Research Part E: Logistics and Transportation Review*, vol. 136, p. 101 922, 2020.

[148] D. Sarkady, L. Neuburger, and R. Egger, "Virtual reality as a travel substitution tool during covid-19," in *Information and communication technologies in tourism 2021*, Springer, 2021, pp. 452–463.

[149] O. Khlystova, Y. Kalyuzhnova, and M. Belitski, "The impact of the covid-19 pandemic on the creative industries: A literature review and future research agenda," *Journal of Business Research*, vol. 139, pp. 1192–1210, 2022.

[150] R. Bukht and R. Heeks, "Defining, conceptualising and measuring the digital economy," *Development Informatics working paper*, no. 68, 2017.

[151] H. Leite, C. Lindsay, and M. Kumar, "Covid-19 outbreak: Implications on healthcare operations," *The TQM Journal*, 2020.

[152] N. Carroll and K. Conboy, "Normalising the new normal: Changing tech-driven work practices under pandemic time pressure," *International Journal of Information Management*, vol. 55, p. 102 186, 2020.

[153] T. Chen, L. Peng, X. Yin, J. Rong, J. Yang, and G. Cong, "Analysis of user satisfaction with online education platforms in china during the covid-19 pandemic," in *Healthcare*, MDPI, vol. 8, 2020, p. 200.

[154] V. Venkatesh, "Impacts of covid-19: A research agenda to support people in their fight," *International journal of information management*, vol. 55, p. 102 197, 2020.

[155] M. Saleem, S. Kamarudin, H. M. Shoaib, and A. Nasar, "Influence of augmented reality app on intention towards e-learning amidst covid-19 pandemic," *Interactive Learning Environments*, pp. 1–15, 2021.

[156] Y. K. Dwivedi, D. L. Hughes, C. Coombs, *et al.*, "Impact of covid-19 pandemic on information management research and practice: Transforming education, work and life," *International journal of information management*, vol. 55, p. 102 211, 2020.

[157] M. Marabelli, E. Vaast, and J. L. Li, "Preventing the digital scars of covid-19," *European Journal of Information Systems*, vol. 30, no. 2, pp. 176–192, 2021.

[158] H. Cho, D. Ippolito, and Y. W. Yu, "Contact tracing mobile apps for covid-19: Privacy considerations and related trade-offs," *arXiv preprint arXiv:2003.11511*, 2020.

[159] M. Mandel and E. Long, "The australian app economy, 2021 update," *Retrieved April*, vol. 2, p. 2021, 2021.

[160] J. Sim, D. Cho, Y. Hwang, and R. Telang, "Frontiers: Virus shook the streaming star: Estimating the covid-19 impact on music consumption," *Marketing Science*, vol. 41, no. 1, pp. 19–32, 2022.

[161] A. Ghose and S. P. Han, "Estimating demand for mobile applications in the new economy," *Management Science*, vol. 60, no. 6, pp. 1470–1488, 2014.

[162] G. Lee and T. S. Raghu, "Determinants of mobile apps' success: Evidence from the app store market," *Journal of Management Information Systems*, vol. 31, no. 2, pp. 133–170, 2014.

[163] T. Bresnahan, J. Orsini, and P.-L. Yin, "Platform choice by mobile app developers," *NBER working paper*, 2014.

[164] A. Tiwana, "Platform synergy: Architectural origins and competitive consequences," *Information Systems Research*, vol. 29, no. 4, pp. 829–848, 2018.

[165] O. Carare, "The impact of bestseller rank on demand: Evidence from the app market," *International Economic Review*, vol. 53, no. 3, pp. 717–742, 2012.

[166] R. Garg and R. Telang, "Inferring app demand from publicly available data," *MIS quarterly*, pp. 1253–1264, 2013.

[167] A. Tongaonkar, S. Dai, A. Nucci, and D. Song, "Understanding mobile app usage patterns using in-app advertisements," in *International Conference on Passive and Active Network Measurement*, Springer, 2013, pp. 63–72.

[168] H. E. Kwon, W. Oh, and T.-H. Kim, "One-sided competition in two-sided social platform markets? an organizational ecology perspective," 2015.

[169] Q. Wang, B. Li, and P. V. Singh, "Copycats vs. original mobile apps: A machine learning copycat-detection method and empirical analysis," *Information Systems Research*, vol. 29, no. 2, pp. 273–291, 2018.

[170] Y. Xu, M. Lin, H. Lu, *et al.*, "Preference, context and communities: A multi-faceted approach to predicting smartphone app usage patterns," in *Proceedings of the 2013 International Symposium on Wearable Computers*, 2013, pp. 69–76.

[171] L. Wang, J. Zheng, Y. Li, and Y. Tan, "Is digital goods consumption resilient to air pollution?" *Available at SSRN 4182042*, 2022.

[172] Y. Qiao, X. Zhao, J. Yang, and J. Liu, "Mobile big-data-driven rating framework: Measuring the relationship between human mobility and app usage behavior," *IEEE Network*, vol. 30, no. 3, pp. 14–21, 2016.

[173] L. Yang, M. Yuan, W. Wang, Q. Zhang, and J. Zeng, "Apps on the move: A fine-grained analysis of usage behavior of mobile apps," in *IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications*, IEEE, 2016, pp. 1–9.

[174] S. Al-Natour, H. Cavusoglu, I. Benbasat, and U. Aleem, "An empirical investigation of the antecedents and consequences of privacy uncertainty in the context of mobile apps," *Information Systems Research*, vol. 31, no. 4, pp. 1037–1063, 2020.

[175] P. Stone, R. Brooks, E. Brynjolfsson, *et al.*, "Artificial intelligence and life in 2030: The one hundred year study on artificial intelligence," 2016.

[176] D. R. Williams, N. Priest, and N. B. Anderson, "Understanding associations among race, socioeconomic status, and health: Patterns and prospects.," *Health Psychology*, vol. 35, no. 4, p. 407, 2016.

[177] R. A. Hummer, "Black-white differences in health and mortality: A review and conceptual model," *The Sociological Quarterly*, vol. 37, no. 1, pp. 105–125, 1996.

[178] N. E. Adler and D. H. Rehkopf, "Us disparities in health: Descriptions, causes, and mechanisms," *Annu. Rev. Public Health*, vol. 29, pp. 235–252, 2008.

[179] M. C. Arcaya, A. L. Arcaya, and S. V. Subramanian, "Inequalities in health: Definitions, concepts, and theories," *Global health action*, vol. 8, no. 1, p. 27 106, 2015.

[180] K. F. Ferraro, B. R. Kemp, and M. M. Williams, "Diverse aging and health inequality by race and ethnicity," *Innovation in aging*, vol. 1, no. 1, 2017.

[181] N. Enayati, E. De Momi, and G. Ferrigno, "Haptics in robot-assisted surgery: Challenges and benefits," *IEEE reviews in biomedical engineering*, vol. 9, pp. 49–65, 2016.

[182] G. I. Barbash, "New technology and health care costs–the case of robot-assisted surgery," *The New England journal of medicine*, vol. 363, no. 8, p. 701, 2010.

[183] T. J. Wilt, R. MacDonald, I. Rutks, T. A. Shamliyan, B. C. Taylor, and R. L. Kane, "Systematic review: Comparative effectiveness and harms of treatments for clinically localized prostate cancer," *Annals of internal medicine*, vol. 148, no. 6, pp. 435–448, 2008.

[184] A. Ng and P. Tam, "Current status of robot-assisted surgery," *Hong Kong Med J*, vol. 20, no. 3, pp. 241–250, 2014.

[185] F. Yanagawa, M. Perez, T. Bell, R. Grim, J. Martin, and V. Ahuja, "Critical outcomes in nonrobotic vs robotic-assisted cardiac surgery," *JAMA surgery*, vol. 150, no. 8, pp. 771–777, 2015.

[186] M. Magnus, J. Herwehe, D. Gruber, *et al.*, "Improved hiv-related outcomes associated with implementation of a novel public health information exchange," *International journal of medical informatics*, vol. 81, no. 10, e30–e38, 2012.

[187] M. E. Frisse, K. B. Johnson, H. Nian, *et al.*, "The financial impact of health information exchange on emergency department care," *Journal of the American Medical Informatics Association*, vol. 19, no. 3, pp. 328–333, 2012.

[188] J. Vest, L. Kern, T. Campion Jr, M. Silver, and R. Kaushal, "Association between use of a health information exchange system and hospital admissions," *Applied clinical informatics*, vol. 5, no. 01, pp. 219–231, 2014.

[189] M. Chen, S. Guo, and X. Tan, "Does health information exchange improve patient outcomes? empirical evidence from florida hospitals," *Health Affairs*, vol. 38, no. 2, pp. 197–204, 2019.

[190] T. Ayer, M. U. Ayvaci, Z. Karaca, and J. Vlachy, "The impact of health information exchanges on emergency department length of stay," *Production and operations management*, vol. 28, no. 3, pp. 740–758, 2019.

[191] Y. Li, L. X. Lu, S. F. Lu, and J. Chen, "The value of health information technology interoperability: Evidence from interhospital transfer of heart attack patients," *Manufacturing & Service Operations Management*, vol. 24, no. 2, pp. 827–845, 2022.

[192] M. Diana and J. Marescaux, “Robotic surgery,” *Journal of British Surgery*, vol. 102, no. 2, e15–e28, 2015.

[193] J. Marescaux, F. Rubino, M. Arenas, D. Mutter, and L. Soler, “Augmented-reality–assisted laparoscopic adrenalectomy,” *Jama*, vol. 292, no. 18, pp. 2211–2215, 2004.

[194] A. R. Lanfranco, A. E. Castellanos, J. P. Desai, and W. C. Meyers, “Robotic surgery: A current perspective,” *Annals of surgery*, vol. 239, no. 1, p. 14, 2004.

[195] S. Nicolau, L. Soler, D. Mutter, and J. Marescaux, “Augmented reality in laparoscopic surgical oncology,” *Surgical oncology*, vol. 20, no. 3, pp. 189–201, 2011.

[196] C. W. Kennedy, T. Hu, J. P. Desai, A. S. Wechsler, and J. Y. Kresh, “A novel approach to robotic cardiac surgery using haptics and vision,” *Cardiovascular Engineering: An International Journal*, vol. 2, no. 1, pp. 15–22, 2002.

[197] J. D’Agostino, M. Diana, M. Vix, L. Soler, and J. Marescaux, “Three-dimensional virtual neck exploration before parathyroidectomy,” *New England Journal of Medicine*, vol. 367, no. 11, pp. 1072–1073, 2012.

[198] T. Hu, A. E. Castellanos, G. Tholey, and J. P. Desai, “Real-time haptic feedback in laparoscopic tools for use in gastro-intestinal surgery,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2002, pp. 66–74.

[199] G. Tholey, T. Chanthasopeephan, T. Hu, J. P. Desai, and A. Lau, “Measuring grasping and cutting forces for reality-based haptic modeling,” in *International Congress Series*, Elsevier, vol. 1256, 2003, pp. 794–800.

[200] R. Satava, J. Bowersox, M. Mack, and T. Krummel, “Robotic surgery: State of the art and future trends,” *Contemp Surg*, vol. 57, no. 10, pp. 489–99, 2001.

[201] S. M. Prasad, C. T. Ducko, E. R. Stephenson, C. E. Chambers, and R. J. Damiano Jr, “Prospective clinical trial of robotically assisted endoscopic coronary grafting with 1-year follow-up,” *Annals of surgery*, vol. 233, no. 6, p. 725, 2001.

[202] M. Seco, C. Cao, P. Modi, *et al.*, “Systematic review of robotic minimally invasive mitral valve surgery,” *Annals of cardiothoracic surgery*, vol. 2, no. 6, p. 704, 2013.

[203] A. Chandra, J. T. Snider, Y. Wu, A. Jena, and D. P. Goldman, "Robot-assisted surgery for kidney cancer increased access to a procedure that can reduce mortality and renal failure," *Health Affairs*, vol. 34, no. 2, pp. 220–228, 2015.

[204] W. R. Chitwood Jr, L. W. Nifong, W. H. Chapman, *et al.*, "Robotic surgical training in an academic institution," *Annals of surgery*, vol. 234, no. 4, p. 475, 2001.

[205] L. W. Nifong and W. R. Chitwood Jr, "Building a surgical robotics program," *The American journal of surgery*, vol. 188, no. 4, pp. 16–18, 2004.

[206] A. N. Sridhar, T. P. Briggs, J. D. Kelly, and S. Nathan, "Training in robotic surgeryan overview," *Current urology reports*, vol. 18, no. 8, pp. 1–8, 2017.

[207] A. L. Smith, K. M. Schneider, and P. D. Berens, "Survey of obstetrics and gynecology residents training and opinions on robotic surgery," *Journal of robotic surgery*, vol. 4, no. 1, pp. 23–27, 2010.

[208] M. Bratti and A. Miranda, "Endogenous treatment effects for count data models with endogenous participation or sample selection," *Health economics*, vol. 20, no. 9, pp. 1090–1109, 2011.

[209] S. Rabe-Hesketh and A. Skrondal, "Understanding variability in multilevel models for categorical responses," in *AERA Annual Meeting, Vancouver, Canada*, Citeseer, 2012.

[210] J. J. Heckman, "Sample selection bias as a specification error," *Econometrica: Journal of the econometric society*, pp. 153–161, 1979.

[211] A. R. Hole, "Fitting mixed logit models by using maximum simulated likelihood," *The stata journal*, vol. 7, no. 3, pp. 388–401, 2007.