RETURN ON MARKETING INVESTMENT: NEW PERSPECTIVES FROM EDUCATION AND HEALTHCARE

A Dissertation

by

YIXING CHEN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Shrihari Sridhar
Co-Chair of Committee,	Vikas Mittal
Committee Members,	Alina Sorescu
	Yonghong An
Head of Department,	Manjit S. Yadav

May 2020

Major Subject: Business Administration

ABSTRACT

There is a rich tradition in the marketing literature that develops models to help executives quantify the financial return on marketing-mix investment. My dissertation adds novel insights to this literature by examining the social return of marketing investment in economically large and societally important sectors that are under-researched in marketing: Education and Healthcare.

Essay 1 provides a comprehensive examination of benefits and costs of school district internet access spending (SDIAS). Compiling a dataset on SDIAS, school performance, and household internet access, I find that while a $1-million increase in SDIAS is associated with an improvement of academic performance indicators by .2 to 3 percentage points (amounting to $1.2 million to $2.5 million for a school district), it is also associated with a 7% increase in Part II offense-related disciplinary problems (amounting to a $38,700 to $80,160 annual cost for a school district). I also find that the benefits and costs of SDIAS are more pronounced in schools in regions with a higher level of household internet access, highlighting the need for school districts to tailor their supplementary initiatives by the level of pre-existing internet exposure in the neighborhoods.

Essay 2 examines the causal effect of outreach interventions on cancer screening completion, uncovers patient-level treatment effect heterogeneity, and assesses the return on these outreach interventions. Using a unique multi-period randomized field experiment among 1,800 at-risk patients for liver cancer, I find that: 1) compared to the

usual-care condition, outreach alone (outreach with patient navigation) increases screening completion rates by 10-20 (13-24) percentage points, and 2) patient-level treatment effects vary substantially across periods and by patients' demographics, health status, visit history, health system accessibility, and neighborhood socioeconomic status, thereby facilitating the implementation of the targeted outreach program. The simulation shows that the targeted outreach program improves the return on the randomized outreach program by 74%-96% (or $1.6 million to $2 million).

ACKNOWLEDGEMENTS

I would like to thank my dissertation chair Dr. Shrihari Sridhar for believing in me and making my life at the Texas A&M University a very fulfilling experience. Hari lifted me up when I had no direction early in my doctoral study and taught me valuable traits over the last five years: passion, motivation, persistence, and discipline.

I would like to appreciate my dissertation co-chair Dr. Vikas Mittal for his guidance and support. Vikas taught me unique attributes of a good person and an independent scholar: sincerity, responsiveness, precision, and deep listening. His tastes of pens and Reebok shoes will continue to enrich my intellectual curiosity.

I also thank my committee members Dr. Alina Sorescu and Dr. Yonghong An for helping me strive for academic excellence during the course of this research. Alina taught me the usefulness of maintaining a deep focus on relevance. I thank Yonghong for taking the time to teach me how to conduct high-quality research that attracts a broad appeal across different research areas.

Thanks also go to the marketing faculty, doctoral colleagues, and department staff for making my time at Texas A&M University a great experience. Special thanks to my close friends for always being there for me.

Finally, this work is dedicated to my family. I would like to express my gratitude to my mother Shirong Guan and father Li Chen for their effort, love, and friendship in bringing me up to be a better human being, and to my grandparents for their mentorship and encouragement.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Dr. Shrihari Sridhar [chair of dissertation committee], Dr. Vikas Mittal [co-chair of dissertation committee] at the Jones Graduate School of Business of Rice University, Dr. Alina Sorescu of the Department of Marketing, and Dr. Yonghong An of the Department of Economics.

The data analyzed for Essay 2 was provided by Dr. Amit G. Singal at the University of Texas Southwestern Medical Center. The sections entitled “Introduction” “Results and Discussion”, and “Concluding Remarks” in Essay 2 were drafted in part by Dr. Ju-Yeon Lee of the Department of Marketing at the Iowa State University.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a graduate assistantship, Mays Department of Marketing Fellowship, and PhD Enhancement Fellowship from Texas A&M University.

This work was also made possible in part by National Institutes of Health (NIH) under R01CA222900 and by Agency for Healthcare Research and Quality under (AHRQ) under R24HS022418. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or AHRQ.

NOMENCLATURE

ADA	Average Daily Attendance
AEIS	Academic Excellence Indicator System
ATE	Average Treatment Effect
CATE	Conditional Average Treatment Effect
CT	Computed Tomography
DAEP	Disciplinary Alternative Education Program
EMR	Electronic Medical Record
HCC	Hepatocellular Carcinoma
MRI	Magnetic Resonance Imaging
SDIAS	School District Internet Access Spending
TAKS	Texas Assessment of Knowledge and Skills
TAPR	Texas Academic Performance Report
TEA	Texas Education Agency
USAC	Universal Service Administrative Company

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Page

1. INTRODUCTION

There is a rich stream of literature that develops quantitative models to help researchers and executives quantify the financial return on marketing-mix investment. My dissertation adds novel insights to this literature by examining the social return of marketing investment in economically large and societally important sectors that are under-researched in marketing: Education and Healthcare.

School districts advertise the academic benefits of school internet access in their strategic communication to parents as a means to attract and retain students. Yet, the benefits and costs of school district internet access spending (SDIAS) are not clearly understood. Does SDIAS increase academic performance among students? What is the role of SDIAS in exacerbating or mitigating school disciplinary problems? Essay 1 provides a comprehensive examination of these questions. Using a longitudinal dataset that combines SDIAS of 1,166 school districts with school performance of 9,804 Texas public schools between 2000 and 2014, I find that a $1-million increase in SDIAS is associated with an improvement of 9 out of 10 academic performance indicators (effect sizes ranging from 1% to 10% of a standard deviation), amounting to an increase of cumulative present student income value of $1.2 million to $2.5 million for a school district. This finding provides academic evidence for school districts to communicate the tangible value of SDIAS to their parents, students, and funding agencies. Yet, a $1-million increase in SDIAS is associated with a 7% increase in Part II offense-related school disciplinary problems, amounting to a $38,700 to $80,160 yearly cost for a school district. This highlights the need for supplementary initiatives to monitor and mitigate the disciplinary consequences of SDIAS if school districts want to maximize the academic benefits. The positive and negative consequences of SDIAS are more pronounced in schools in regions with a higher level of household internet access, highlighting the need for school districts to tailor their supplementary initiatives by the level of pre-existing internet exposure in the neighborhoods where schools are located.

In a similar vein, healthcare organizations make substantial investments in direct-to-patient outreach interventions to encourage the use of regular cancer screening as undergoing regular screening enables early detection, potentially saving money and lives. In Essay 2, I use a multiperiod, randomized field experiment involving 1,800 at-risk patients for liver cancer to examine the causal effect of outreach interventions on screening completion, uncover patient-level treatment effect heterogeneity, and assess the return on these interventions. I find that 1) compared to the usual-care condition, outreach alone (outreach with patient navigation) increases screening completion rates by 10-20 (13-24) percentage points, and 2) patient-level treatment effects vary substantially across periods and by patients' demographics, health status, visit history, health system accessibility, and neighborhood socioeconomic status, thereby facilitating the implementation of the targeted outreach program. The simulation shows that the targeted outreach program improves the return on the randomized outreach program by 74%-96% or $1.6 million to $2 million. Thus, outreach marketing provides a substantial positive payoff to the healthcare system.

2. INVESTIGATING THE ACADEMIC PERFORMANCE AND DISCIPLINARY CONSEQUENCES OF SCHOOL DISTRICT INTERNET ACCESS SPENDING

2.1. Introduction

U.S. public schools spend more than any other developed nation on their students' education every year. For example, in 2010, the United States spent $11,000-$12,000 per student on education, whereas countries such as Australia, Belgium, and Denmark spent $9,313 per student[1]. Yet, a 2017 nation-wide survey of more than 7,200 parents showed that only 43% of parents with children in traditional public schools are "very satisfied" with their child's school, compared to 61% of parents with children in private schools[2]. This result echoes the deteriorating confidence that parents have in public schools—while 63% of parents indicated a "great deal" of confidence in public schools in 1973, the number was down to 44% in 2017. Parents' confidence in public schools comes when the school district can raise and maintain academic performance (Black 1999; Jacob and Lefgren 2007) while keeping students safe from physical and cyberbullying attacks (Rabovsky 2011). Against this backdrop, I examine how *school district internet access spending* (hereafter "SDIAS"), a quintessential marketing-mix investment, affects school academic performance and school disciplinary problems.[3]

---

[1] https://www.cbsnews.com/news/us-education-spending-tops-global-list-study-shows/
[2] http://ccubes.net/press-features/survey-public-school-parents-less-satisfied-with-engagement-opportunities/
[3] I have multiple sources of evidence to support that internet access spending is a school district-level strategic decision with a stated goal to enhance internet connectivity at the school level. First, in accordance with EducationSuperHighway, a leading nonprofit organization focused on upgrading the internet access in every U.S. public school classroom, internet access budget planning (e.g., broadband

In terms of marketing, school districts have traditionally invested in awareness campaigns (Kotler and Levy 1969), visual branding (DiMartino and Jessen 2016), open houses (Oplatka 2007), curricular innovation (Lubienski 2003), and direct mail using brochures and informational packets (Lubienski 2007). In the last two decades, SDIAS has rapidly evolved as a dominant marketing-mix investment. Whereas only 14% of the K-12 classrooms in U.S. had internet access in 1996, 98% of school districts meet the FCC's 100 kbps per student goal for internet access in 2018 (EducationSuperHighway 2018).

As a marketing strategy, school districts routinely publicize the potential of internet access as a means to improve school academic performance. For example, the Arkansas Public School Computer Network (APSCN) pledged a $11 million investment in internet access with the stated goals of "*supporting personalized and differentiated instruction for every student*[4]", the Brewton School District in Alabama recently publicized their collaboration with Mediacom to upgrade their classrooms with high-speed internet[5], and the Bellville independent school district in Texas documents that "*a major strategic goal of the district is to promote educational excellence by facilitating resource sharing, innovation, and communications by providing internet access to*

initiative, fiber, Wi-fi network upgrade) is a school district-level decision. Second, Public School Network Capabilities Study, a report to the Texas Legislature from the Texas Education Agency, is conducted at the school district level to assess the network capabilities of public schools. Third, as I will detail in the data section, the majority of internet access funds are requested by school districts instead of individual schools in Texas: while internet access funds requested from individual schools are $26.2 million, those from school districts are $483.7 million.

[4] https://cds.educationsuperhighway.org/wp-content/uploads/2015/10/Smarter-Spending-for-Smarter-Students-Arkansas-Report-December-2014.pdf

[5] https://www.ncta.com/whats-new/spotlight-mediacom-is-bringing-a-rural-school-in-alabama-up-to-speed

students, teachers and administrators in the district." Parents also serve as advocates for internet access in schools; a pilot study of 3,924 parents shows that parents' satisfaction with internet access in their child's school is associated with their overall satisfaction with their child's school (see Appendix A.1).

Despite the importance of SDIAS in the marketing and strategic positioning of school districts, there is little theoretical agreement on payoffs to SDIAS as summarized in Table 2.1. Studies in Table 2.1 show a positive (Dettling, Goodman, and Smith 2018), neutral (Faber, Sanchis-Guarner, and Weinhardt 2016), and negative (Belo, Ferreira, and Telang 2014) effect. The literature review suggests that this discrepancy in findings stems from three main gaps. First, several studies use household internet access in the geographic area (e.g., broadband coverage) to proxy SDIAS (Dettling, Goodman, and Smith 2018; Vigdor, Ladd, and Martinez 2014). However, household internet access in the geographic area may not accurately reflect a school district's strategic decision to invest in internet access (see Gap 1 in Table 2.1). For example, my data show that 67% of the school districts in regions with above-median broadband coverage spend less than the median on SDIAS in a certain year.

Second, extant research examines either household internet access or school internet access, but not both. As such it does not consider whether and how the effectiveness of SDIAS varies by changes in household internet access. The reality is that both matter given the growing use of one-to-one-computing, learning management systems in school curriculum, and digital textbooks in school districts. Public schools lament the digital-use gap (or the gap in internet access at school and at home)[6] for lack of student progress, encouraging and incentivizing parents to commit towards significant investments in high-speed internet access at home.[7] However, no study has put this argument to an empirical test. Thus, it is not known if the payoffs to SDIAS are related to household internet access in the neighborhood (see Gap 2 in Table 2.1).

Third, empirical evidence about the potential downside of SDIAS is lacking, although there is an ongoing debate that increasing SDIAS may be correlated with school disciplinary problems (see Gap 3 in Table 2.1). For example, the Hutto Independent School District in Texas advocates having open internet access in schools to avoid overprotecting children while monitoring their use of the internet.[8] In contrast, other schools and policy advocates believe that internet access has stoked bullying, self-esteem issues, and stalking.[9] Accordingly, public schools feel the burden of monitoring potentially malicious online activity in which their children may engage, with the luxury of SDIAS.

I address these three gaps by providing a comprehensive empirical examination of the payoffs to SDIAS on i) school academic performance and ii) school disciplinary problems, estimating the effect of household internet access along with SDIAS. I concatenate school district internet access spending data from Universal Service Administrative Company (USAC), school academic performance data from Texas

---

[6] https://www.govtech.com/education/news/home-internet-access.html
[7] https://www.aps.edu/news/internet-essentials-provides-affordable-high-speed-internet-access-to-many-aps-families
[8] https://www.educationdive.com/news/best-practices-to-protect-and-empower-students-online/444690/
[9] http://middleboroughtv.com/should-social-media-be-banned-in-school/

Table 2.1 Literature on the Impact of Internet on Academic Outcomes

Study	Focal independent variable	Effect Direction of internet access on academic performance	Gap 1: Focal independent variable proxies school district internet spending?	Gap 2: SDIAS effectiveness contingent on household internet access?	Gap 3: Considers potential downside (i.e. disciplinary problems) of SDIAS?
Dettling, Goodman, and Smith (2018)	A rural (urban) zip code has broadband coverage when there is at least 1 provider per 12 square miles (at least 1 provider for every 2,700 people)	+	No	No	No
Faber, Sanchis-Guarner, and Weinhardt (2016)	Dummy variable that denotes whether a discontinuous jump in internet connection that occurs (i.e. cross the boundary segment from the slower to faster side).	n.s.	No	No	No
Vigdor, Ladd, and Martinez (2014)	Number of broadband service providers in a zip code	−	No	No	No
Belo, Ferreira, and Telang (2014)	Mean value of total monthly school traffic over each year	−	No	No	No
Goolsbee and Guryan (2006)	Subsidy rate of total E-rate funds requested (internet access, telecommunications, and internal connections) for a school district	n.s	No	No	No
Hazlett, Schwall, and Wallsten (2019)	Per pupil amount of E-rate funding (internet access, telecommunications, and internal connections) committed to a school district	−	Yes	No	No
Violette (2017)	Amount of E-rate funding on internet access committed to a school district	n.s.	Yes	No	No
Current research	Amount of internet access spending that a school district pays its service providers	+	**Yes**	**Yes**	**Yes**

Note. SDIAS refers to school district internet access spending.

Education Agency (TEA), school discipline data from Public Education Information Management System (PEIMS), and household internet access data from Federal Communications Commission (FCC) Form 477. My yearly data cover 9,804 public schools over 2000-2014, including 10 academic performance indicators and 47 types of school disciplinary problems which are classified into serious criminal offenses such as aggravated kidnapping (Part I offenses) and relatively less serious criminal offenses such as possession of an illegal knife (Part II offenses). The econometric model controls for persistent unobservables (using school fixed effects), common time-varying unobservables (using year fixed effects), individual-specific time-varying unobservables (using a rich set of school-, school district-, and county-level covariates drawn from extant literature), and any remnant endogeneity using an instrumental variable approach that leverages the institutional nature of the E-rate funding reimbursement for internet access spending.

The findings are as follows. First, a $1-million increase in SDIAS is associated with an improvement in 9 out of 10 academic performance indicators: there is a .2 to 3 percentage points increase in all four college readiness indicators (e.g., SAT/ACT meet criterion rate) as well as 5 out of 6 commended performance indicators over grades 3 to 11 in Texas (e.g., reading, writing, social studies). The effect sizes are substantive: a $1-million increase in SDIAS generates cumulative present income value for high school graduates through an improvement in college readiness indicators, amounting to $1.2 million to $2.5 million for a school district.

Second, and very interestingly, the positive effects of SDIAS on school academic performance are higher among schools in regions with a higher level of household internet access. The positive effects of SDIAS on academic performance are on average 9% higher among schools in regions with a high level of household internet access than those in regions with the mean level of household internet access. This suggests that higher household internet access, which broadens students' off-campus internet usage, likely reinforces the supplementary online learning being added to regular school teachings as a result of SDIAS.

Third, whereas an increase in SDIAS is reassuringly not associated with Part I offense-related school disciplinary problems, it has a positive and statistically significant effect on Part II offense-related school disciplinary problems. A $1-million increase in SDIAS increases Part II offense-related school disciplinary problems by 7%, amounting to a $38,700 to $80,160 yearly cost for a school district. Interestingly, the deleterious effect of SDIAS on school disciplinary problems is higher in schools in regions with a higher level of household internet access; the negative effect of SDIAS is on average 19.6% higher among schools in regions with a high level of household internet access than those in regions with mean household internet access.

Together, the findings offer four contributions to the marketing discipline. First, I clearly document that increasing SDIAS improves school academic performance, which is linked to student's potential income value. To be comprehensive, I examine 10 indicators of academic performance that cover a) both state mandatory test scores and college readiness indicators (e.g., SAT) to compensate the opt-in nature of the college readiness indicators and b) all stages of primary and secondary education to avoid the focus on a single educational stage. This finding validates the results from the pilot study (Appendix A.1) which showed that parents' satisfaction with school internet access is positively associated with their overall satisfaction with their child's public school, which in turn, is positively related to their reenrollment intentions into the same public school and negatively associated with switching intention to competing private schools. For the education sector, I show how a marketing-mix investment such as SDIAS can be used to satisfy and retain customers. Substantively, for the education sector, I show that the effect size of SDIAS is comparable to other strategic investments such as investments in hardware (Fairlie and London 2012), software (Roschelle et al. 2016), class size (Angrist and Lavy 1999), and teacher experience (Hanushek 2011). Thus, school districts should carefully and judiciously allocate resources to SDIAS alongside other value-enhancing investments during their strategic planning cycles.

Second, by showing that the SDIAS effectiveness is contingent on household internet access, this study provides several insights. In terms of research, it resolves conflicting findings because studies have used household internet access and SDIAS as interchangeable measure of the same construct. The results show that investing in schools through SDIAS may not yield the intended effects unless the internet exposure that students get in their homes is taken into account. Thus, any policy designed to reap the benefits of internet usage on learning outcomes needs to incentivize both schools and households (Belo, Ferreira, and Telang 2016; Wei et al. 2011)—an issue that has not been examined or tested in past research.

Third, I uncover and document the disciplinary risks associated with increased

SDIAS. By showing that $1-million increase in internet access spending is associated with an 7% increase in the number of Part II offense-related school disciplinary problems, I add to extant literature on the antecedents of school disciplinary actions. This literature has focused on classroom and teacher characteristics, child's family situation and school racial composition but not internet access. The results can shed light on developing effective schooling and parenting strategies that mitigate the deleterious consequences of internet access. By addressing and mitigating disciplinary issues, school district administrators can utilize SDIAS as way to attract and retain customers.

Fourth, I add to the burgeoning literature that examines the link between internet access and crime (e.g., Bhuller et al. 2013; Chan, Ghose, and Seamans 2016; Lindo, Siminski, and Swensen 2018). The results show that both institutional access (i.e., through school) and personal access (i.e., through household) can have separate effects on crime. Perniciously, internet access at home may compound the deleterious effect of SDIAS on school disciplinary problems. Thus, increased SDIAS needs to be supplemented with adequate investments in monitoring and/or mitigating its negative repercussions, and these investments may be dependent on the types of households served by the school district.

In the next section, I discuss the data, institutional setting, and identification strategy. Following that, I present the empirical findings, and conclude with a discussion of theoretical and practical takeaways and potential limitations.

2.2. Data

I construct the dataset by concatenating information from a variety of sources. I collect SDIAS data from Universal Service Administrative Company (USAC), school academic performance data from Academic Excellence Indicator System (AEIS) and Texas Academic Performance Reports (TAPR), school discipline data from Public Education Information Management System (PEIMS), county-level population data from the National Cancer Institute's Surveillance Epidemiology and End Results (Cancer-SEER) program,[10] county-level unemployment rates from the Bureau of Labor Statistics Local Area Unemployment Statistics (BLS), and county-level median household income from U.S. Census Small Area Income and Poverty Estimates (SAIPE). Figure 2.1 summarizes these data sources, and Appendix A.2 describes the concatenation procedure. I next describe the data used in this study.

2.2.1. SDIAS Data

The Universal Service Program for Schools and Libraries (commonly known as E-rate) was established in 1996 with the goal of providing funding to schools, school districts, and libraries to obtain telecommunications and information services at an affordable rate. The program allows schools, school districts, and libraries to request funding to subsidize their costs under five service types: telecommunications (e.g., local and long distance wired telephone service), internet access (e.g., basic conduit access to

[10] As reported by Stevens et al. (2015), the Cancer-SEER population data are more accurate than data from the U.S. Census because they "are based on an algorithm that incorporates information from Vital statistics, IRS migration files, and the Social Security database".

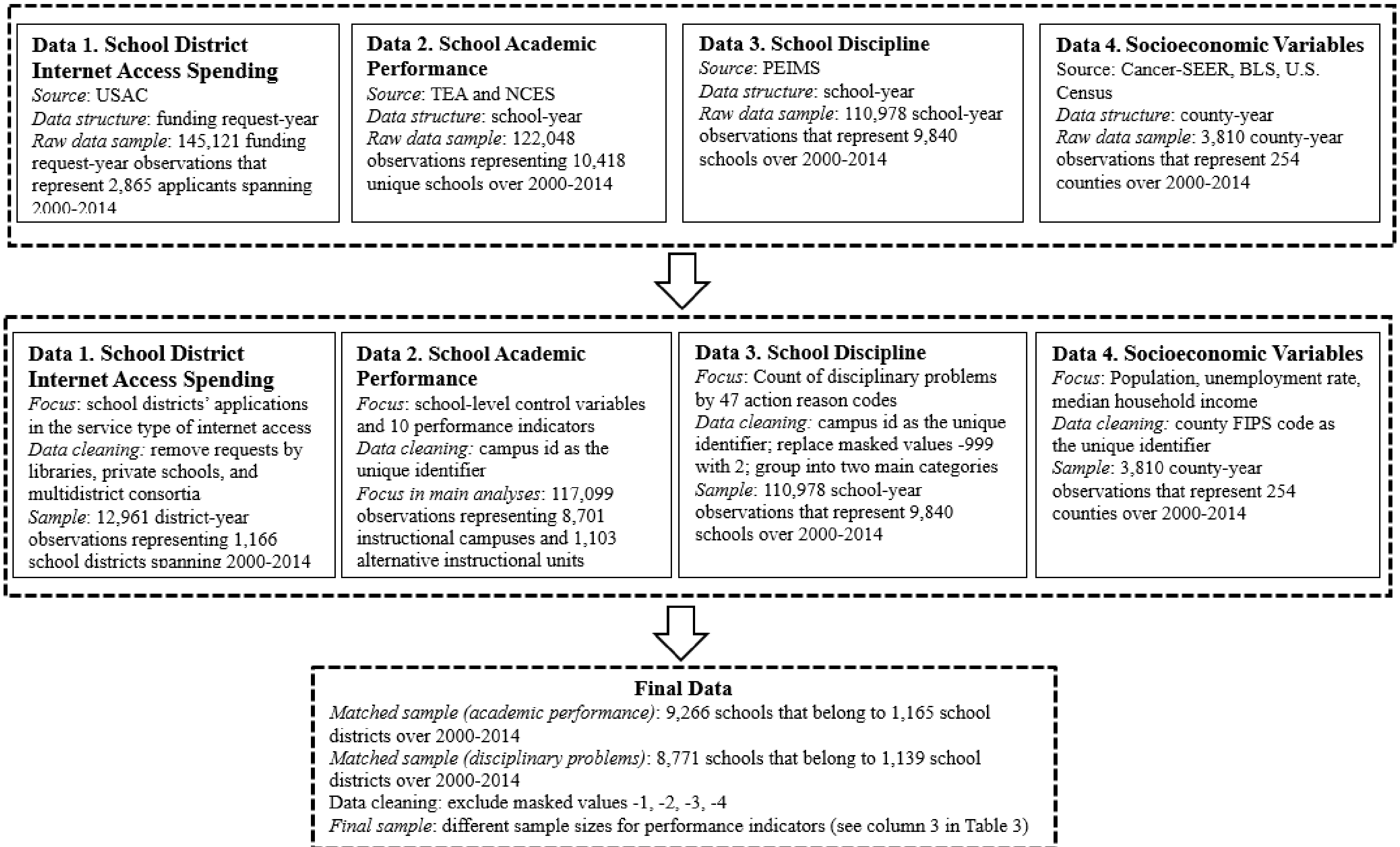


Figure 2.1 Summary of Data Collection and Concatenation

the internet, broadband connectivity, and Wi-Fi), internal connections (e.g., access points, routers, switches, hubs, and wiring), managed internal broadband services[11], and basic maintenance of internal connections (e.g., cable maintenance). The program officially began in 1998 by offering $2.4 billion annually to schools and libraries. The annual funding cap of the program was adjusted to $3.9 billion in December 2014 to further improve broadband connectivity as well as to expand Wi-Fi networks.

The E-rate funding application process involves multiple steps (see Appendix A3 for a summary). In Step 1, eligible schools[12], school districts, or libraries complete FCC Form 470 that describes the service requests (e.g., broadband connectivity), and USAC posts these requests for service providers' consideration. Subsequently, service providers offer the prices to compete for these service requests through a bidding process.[13] In Step 2, schools, school districts, or libraries select the service contract after evaluating the bids received. Schools and libraries may consider multiple factors in their evaluation, but the price of the eligible products and services must be the most heavily weighted factor under the rules of the E-rate program. In Step 3, schools, school districts, or libraries

---

[11] USAC did not provide funding support in the category of managed internal broadband services until 2015.

[12] Eligible schools must meet the statutory definition of elementary and secondary schools defined in the No Child Left Behind Act of 2001 (20 U.S.C. Section 7801(18) and (38)): An elementary school is a non-profit institutional day or residential school, including a public elementary charter school, that provides elementary education, as determined under state law. A secondary school is a non-profit institutional day or residential school, including a public secondary charter school, that provides secondary education, as determined under state law, except that such term does not include any education beyond grade 12. Schools operating as for-profit businesses or that have endowments exceeding $50 million are not eligible.

[13] The bidding process is open and fair as 1) all bidders are treated the same; 2) no bidder can have advance knowledge of the project information; 3) all bidders have common information and know what is required of them; 4) with limited exceptions, service providers and potential service providers cannot give gifts to applicants; and 5) the value of free services (e.g., promotional offers) must be deducted from the pre-discount cost of funding requests.

complete Form 471 to request funding in accordance with the service costs in the contract, and USAC reviews the request and determines the funding commitment based on a) the percentage of students eligible for the National School Lunch Program in the school district and b) the urban or rural status of the school district (Appendix A.3).[14] For example, a school district in rural area that has 40% of students eligible for the NSLP is expected to receive 70% (i.e. discount rate) of total internet access funds requested as the funding commitment. In Step 4, schools, school districts, or libraries file an FCC Form 486 to inform USAC that the delivery of approved services has started, and invoicing process can begin. In Step 5, applicants or service providers receive the reimbursement of the service costs after completing the invoicing process.

I obtain all E-rate requests of applicants in the state of Texas from 2000 to 2014 from USAC. This contains 145,121 funding requests by 2,865 applicants spanning 2000-2014. Each request includes service type (e.g., internet access), applicant type (i.e. school, school district, library, consortium), total funds requested (i.e. service costs), funding status (whether the funding request is approved by USAC), discount rate, funding commitment (total funds granted by USAC), and final reimbursement (actual amount received by applicants or service providers).

Appendix A.3 shows the breakdown of total funds requested by service type and over time: internal connections (51%), telecommunications (35%), and internet access

---

[14] Schools may request funding individually or as a school district. An individual school that is part of a school district does not calculate its discount rate based on its own student population, but instead uses the discount rate calculated for the school district.

(11%) are the three dominant service types where schools and libraries request funding. I examine internet access spending, i.e. cost figures listed in the service contract between schools/school districts and their service providers. Appendix A.3 shows the ratio of internet access spending to total funds requested in all categories grew from 7% in 2000 to 36% in 2014.

The data show that 91.4% of total funds requested on internet access come from individual schools ($26.2 million) and school districts ($483.7 million) in Texas, while the remaining 8.6% come from libraries and consortium. Since I focus on the SDIAS, I remove all requests by libraries, private schools, and multidistrict consortia, while retaining funding requests only on internet access. In addition, I aggregate all requests to the level of school district for two reasons: first, in Texas, the overwhelming majority of funds ($483.7 million) is requested by school districts rather than individual schools; second, I am able to retain funding information about individual schools' requests. Thus, I end up with 12,961 school district-year observations (2000-14) representing 1,166 school districts[15] that requested $491.6 million on internet access.[16] On average, school districts requested $213,188 per year, with a standard deviation of $681,015.

2.2.2. Texas School Academic Performance Data

I obtain school academic performance data from the Academic Excellence

[15] As shown in Appendix A3, these 1,166 school districts that requested E-rate funding (orange dots) represent 88% of Texas school districts, showing that E-rate is almost a universal funding source for school districts in Texas. School districts that did not request funding (blue dots) are scattered across Texas.

[16] Note that this amount accounts for 96.4% of all internet access funds requested from schools and school districts combined (i.e. 483.7+26.2=509.9 million dollars), and the remaining proportion (i.e. 3.6%) is requested by private schools.

Indicator System (AEIS) and Texas Academic Performance Reports (TAPR) maintained by Texas Education Agency (TEA). AEIS provides performance indicators for each public school in the state of Texas. These reports also provide extensive school-level profile information about student, staff, finances, and programs. TEA replaced AEIS with TARP to report performance information in the school year 2012-2013; thus, I collect the last two years of data from TARP. To be comprehensive, I examine a wide range of academic performance indicators based on two official performance standards that TEA adopts: (a) Texas state accountability system and (b) Gold performance acknowledgment system. TEA uses the former to evaluate whether districts and campuses are academically acceptable (i.e. accountability ratings) and the latter to acknowledge them for high performance on indicators other than those used to determine accountability ratings.

Texas public school students in grades 3-11 are evaluated by a comprehensive assessment program in accordance with the state-mandated curriculum. Texas first administered the Texas Assessment of Academic Skills (TAAS) test to all eligible students in grades 3 through 8 and grade 10 during the 2001–2002 school year. From 2002–3 to 2011-12, Texas administered the Texas Assessment of Knowledge and Skills (TAKS) test to all eligible students in grades 3-11. In Spring 2012, Texas students began taking the State of Texas Assessments of Academic Readiness (STAAR). The assessments are administered to students in grades 3–8 and high school courses with end-of-course assessments. Among these three state exams, I focus on TAKS test results since the exam period of TAKS (2003-2011) has the largest overlap with the data period (i.e. 2000-2014). I summarize the subjects evaluated at each grade level for TAKS in Appendix A.4. The key indicators of TAKS are commended performance on mathematics, reading/ELA, writing, science, and social studies, and the overall score. The indicator is reported as the percent of students who have shown a thorough understanding of the knowledge and skills for the subject across all grade levels at the school. However, TAKS does not require testing of all subjects in all grades. For example, only students in grades 4 and 7 are required to take the writing test, while students in grades 3-11 are required to take the mathematics test. I collect data on all the tests scores available by year, within the purview of the TAKS system.

In addition to the TAKS mandated exams, students in grades 9-12 can also voluntarily take exams or complete advanced courses to indicate their college readiness: i.e. whether students are able to perform college-level course work. College readiness indicators include SAT/ACT test results, AP/IB test results, advanced course/dual enrollment completion, and Recommended High School Program/Distinguished Achievement Program (RHSP/DAP) graduates. I also collect data on these indicators.

In total, I collect the data of 10 academic performance indicators, including six TAKS commended performance indicators (mathematics, reading/ELA, writing, science, and social studies, all subjects combined), SAT/ACT meet criterion rate, AP/IB meet criterion rate, advanced course completion rate, and RHSP/DAP graduates.

These data are collected at the school level. The raw data consists of 122,048 observations representing 10,418 unique Texas schools over 2000 to 2014, and I focus on instructional campuses (8,701 schools) and alternative instructional units (1,103 schools).[17] In my main analysis, two institutional factors affect the sample sizes of performance indicators. First, students in grades 3-11 are evaluated on different subjects at each tested grade level, and commended performance is measured across all grade levels at the school level (see Appendix A.4). Second, according to Family Educational Rights and Privacy Act requirements (FERPA), TEA is required to create a set of mask rules to conceal the performance indicators in the case of small samples. I summarize the mask rules in Appendix A.4. As I do not observe the true values of the masked observations, I exclude masked observations when conducting all analyses.

The second panel of Table 2.2 presents the definitions and summary statistics of school academic performance indicators, all of which are expressed in rates. For instance, on average, 13.5% of non-special education graduates scored at or above criterion score of SAT or ACT, and 10.2% (20.6%) of non-mobile students achieve commended performance in all subjects combined (science). As an illustrative example, Appendix A.4 shows the temporal variation in four performance indicators. Whereas SAT/ACT meet criterion rate remains relatively steady, commended performance indicators show an upward trend over time.

---

[17] In Texas, a school is an instructional campus (108,363 observations), alternative instructional unit (8,736 observations), juvenile justice alternative education program (JJAEP, 2,423 observations), disciplinary alternative education programs campus (DAEP, 2,515 observations), or budgeted campus (11 observations). My main analysis focuses on instructional campus and alternative instructional units as the goals of JJAEP and DAEP are distinct. For instance, JJAEP intends to reduce delinquency, increase offender accountability and rehabilitate offenders through a comprehensive, coordinated community-based juvenile probation system. To check this statement, I confirm that in my data, there are only 42 (163) observations for SAT/ACT meet criterion rate (TAKS commended performance all subjects) from JJAEP and DAEP.

Table 2.2 Variables, Definitions, and Summary Statistics

Variable	Definitions	N	Mean	SD	Min	Max
Internet access spending	Total cost of internet access that school districts pay to service providers	100,842	213,188	681,015	0	5,725,525
Academic Performance						
SAT/ACT Meet Criterion Rate (%)*	Percent of non-special education graduates who scored at or above criterion (1110 on the total of math and critical reading of SAT, or 24 on the ACT)	18,421	13.5	11.3	0	91.7
AP/IB Meet Criterion Rate (%)*	Percent of non-special education 11th and 12th grade students with >= one AP or IB score at or above the criterion (>=3 on AP, and >=4 on IB)	12,252	9.2	10.3	0	98.3
Advanced Course Completion (%)	% of students who receive credit for >= one advanced course in grades 9–12.	27,643	19.2	17.3	0	100
RHSP/DAP Graduates (%)	Percent of graduates who were reported as having satisfied the course requirements for Recommended High School Program or Distinguished Achievement Program	21,960	64.1	28.6	0	100
TAKS Commended Performance: All subjects (%)*	Percent of non-mobile students who have shown a thorough understanding of the knowledge and skills across all grade levels.	61,566	10.2	9.4	0.1	75.2
TAKS Commended Performance: Science (%)*	Percent of non-mobile students who have shown a thorough understanding of the knowledge and skills on science across all grade levels.	46,802	20.6	17.2	0.0	91.1
TAKS Commended Performance: Math (%)*	Percent of non-mobile students who have shown a thorough understanding of the knowledge and skills on math across all grade levels.	62,704	21.3	14.6	0.1	87.1
TAKS Commended Performance: Reading/ELA (%)*	Percent of non-mobile students who have shown a thorough understanding of the knowledge and skills on ELA across all grade levels.	62,777	25.3	14.0	0.2	82.2
TAKS Commended Performance: Social Studies (%)*	Percent of non-mobile students who have shown a thorough understanding of the knowledge and skills on social studies across all grade levels.	24,288	26.5	16.7	0.4	89.1
TAKS Commended Performance: Writing (%)*	Percent of non-mobile students who have shown a thorough understanding of the knowledge and skills on writing across all grade levels.	44,877	23.1	14.3	0.4	93.1
Disciplinary Outcomes						
Part I disciplinary problems	Number of students who are under discipline for Part I crime offenses	110,384	1.2	2.9	0	109
Part II disciplinary problems	Number of students who are under discipline for Part II crime offenses	110,384	10.9	24.2	0	448
Control variables						
Percent of students in reduced/free lunch	Percent of students eligible for free/reduced-price lunch	122,048	58.5	27.7	0	100
Percent of African American students (%)	Percent of African American students enrolled	122,048	13.6	18.8	0	100
Percent of Hispanic students (%)	Percent of Hispanic students enrolled	122,048	45.2	31.5	0	100
Enrollment	Total number of students who were reported in membership	122,048	570.8	488.7	1	5,094
Average teacher experience	Sum of years of experience for teachers divided by the number of teachers	117,618	11.5	3.5	0	50
Student/teacher ratio	Number of students enrolled divided by the total number of teachers	117,571	14.3	3.7	0	62.0
Unemployment rate (%)	Unemployment rate in the Texas county	122,048	6.1	1.9	1.9	19.4
Population	Total population in the Texas county	122,048	1,016,257	1,276,613	258	4,441,928
Household income	Estimates of median household income in the Texas county	122,048	44,977.9	12,131.0	17,201	92,466
Instructional spending	Instruction expenditures per pupil each district	121,972	4,930.9	1,144.0	0	10,7688.9

*I scale these metrics by multiplying the percent of students who took the exams. E.g., SAT/ACT Meet Criterion Rate= [Number of examinees who scored at or above criterion divided by number of examinees] * [Number of examinees divided by number of non-special education graduates]; TAKS refers to Texas Assessment of Knowledge and Skills.

2.2.3. School Discipline Data

I collect school discipline data from Public Education Information Management System (PEIMS). The data contains the number of student classroom removals (i.e. counts of disciplinary problems) by student code of conduct violations (i.e. disciplinary action reason codes) at the school level over 2000-2014. Appendix A.5 provides the list of 47 different reason codes. They include serious crimes such as aggravated kidnapping (code 19) as well as misdemeanors such as possession of cigarette or tobacco products (code 33). I organize the data in two steps: first, in accordance with Family Educational Rights and Privacy Act requirements or FERPA, TEA is required to conceal the number of disciplinary problems with the symbol -999 to indicate that there are less than five counts of disciplinary problems within a school. As the data are counts, I replaced these values with 2 instead of 2.5. Second, given the larger number of reason codes and sparse observations within some of these reason codes, I group these reason codes into two categories using the Federal Bureau of Investigation's Crime Part I and II offense definitions.[18] Part I offense-related disciplinary problems (Part I) are generally most serious, and Part II offense-related disciplinary problems (Part II) are relatively less serious. For instance, aggravated assault (codes 29 and 30) is included in Part I, while assault (codes 27 and 28) in Part II. I have 110,978 school-year observations that represent 9,840 schools over 2000-2014. The third panel in Table 2.2 reports two aggregate measures of school disciplinary problems, all of which are expressed in

---

[18] https://ucr.fbi.gov/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/offense-definitions

counts. On average, there are 1.2 (10.9) classroom removals due to Part I (Part II).

2.2.4. Socioeconomic Data

I obtain county-level population data from the Cancer-SEER, county-level unemployment rates from the BLS, and county level income data from SAIPE. I have 3,810 county-year observations representing 254 counties over 2000-2014.

2.3. Empirical Analysis

2.3.1. Identification Strategy: School Academic Performance

I start with a linear model linking the dependent variable Y_{ijkt} of a school i of school district j in county k in year t (e.g., SAT/ACT meet criterion rate) to $SDIAS_{jkt}$ (i.e. SDIAS in school district j in county k in year t). The baseline specification is:

(1) $$Y_{ijkt} = \lambda + \alpha_1 SDIAS_{jkt} + \varepsilon_{ijkt},$$

The coefficient α_1 captures the effect of SDIAS on school academic performance, and the term ε_{ijkt} is the error component.

The identification of α_1 is complicated by several issues endemic to observational data. First, schools may perform differently on the basis of persistent unobserved characteristics that may also be correlated with their school districts' internet access spending behavior (i.e. unobserved cross-sectional differences). For instance, urban school districts may invest more in internet access, and schools in those districts may also perform better than those in rural school districts since they have better access to high quality internet infrastructure (e.g., Wi-Fi infrastructure) that may increase internet access spending, and also offer better access to amenities and infrastructure that attract high-performing families. In such a case, the estimates of SDIAS on school academic performance would be biased. Accordingly, I control for school-specific cross-sectional difference by decomposing the error term as $\varepsilon_{ijkt} = u_i + \epsilon_{ijkt}$, where u_i represents school fixed effects (Wooldridge 2010).

Second, temporal variation in SDIAS and school academic performance may be driven by common time-varying unobservables. For instance, schools may perform better and be motivated to spend more during periods particularly conducive to growth (e.g., economic boom) or when academic outcomes are more salient than usual in students' minds (e.g., periods around the statewide launch of STAAR assessment program in Texas in 2012). Accordingly, I include year fixed effects to capture such common time-varying shocks and any other state-specific common temporal shocks that may affect all schools. Thus, I have:

(2) $Y_{ijkt} = \lambda + \alpha_1 SDIAS_{jkt} + u_i + \eta_t + \epsilon_{ijkt}$,

where u_i and η_t represent school and year fixed effects respectively.

Third, even after employing this unobserved-effects approach, SDIAS could still be correlated with school-, school district-, or region-specific time-varying unobservables that also affect school academic performance. For example, school districts with a growing reputation (or a growth in instructional budget) may have a propensity to invest more financial resources in internet access, and its schools perform better than the average school’s trends. In addition, school districts located in areas where socioeconomic status is higher (e.g., income, population growth) may be more willing to pay for premium broadband internet subscription. These unobserved school-, school district-, or region-specific proxies of increased internet access spending are likely correlated with academic performance, e.g., schools in areas with a higher income level may also achieve a higher level of academic performance. Omitting these factors would induce an upward bias on the impact of internet access spending on school academic performance given a positive correlation between school academic performance and socioeconomic characteristics such as income.[19]

To alleviate such concerns, I include a rich set of proxies for omitted variables at school-, school district-, and county-level. At the school level, I control for variables identified in the literature, including enrollment, composition of economic disadvantaged students (peer group characteristics; Sacerdote 2001), racial composition (Skiba et.al 2014), teacher experience (teacher training; Angrist and Lavy 2001), and student-teacher ratio (class size; Angrist and Lavy 1999). At the school district level, I control for instructional spending per pupil as a proxy for school district instructional budget. At the county level, I control for three socioeconomic factors: population, unemployment rate, and median household income. Thus, I arrive at:

$$(3) \qquad \mathrm{Y_{ijkt}} = \lambda + \alpha_1 \mathrm{SDIAS_{jkt}} + \boldsymbol{\beta}_1' \mathbf{Z_{ijkt}} + \beta_2 \mathrm{M_{jkt}} + \boldsymbol{\beta}_3' \mathbf{X_{kt}} + \mathrm{u_i} + \eta_t + \epsilon_{\mathrm{ijkt}},$$

where $\mathbf{Z_{ijkt}}$, $\mathrm{M_{jkt}}$, and $\mathbf{X_{kt}}$ include school-, school district-, and county-level proxies respectively. The identifying assumption in Equation (3) is that accounting for time-varying covariates and fixed effects captures all confounding correlated unobservables.

[19] I checked the differences in socioeconomic characteristics between school districts with below-median and above-median internet access spending. In line with the endogeneity concerns, for example, household income is significantly higher in above-median internet access spending areas than that in below-median internet access spending areas (u_1=48,064, u_2=42,410, $p < .01$).

Finally, school districts might still strategically determine SDIAS with some expectation of better academic performance that in turn drives internet access spending. This strategic expectation is unobserved to the researcher but likely correlated with both internet access spending and academic performance and could result in endogeneity notwithstanding the controls discussed so far. To alleviate this concern, I supplement the model with an instrumental variable (IV) approach. I seek a source of variation in SDIAS that can be excluded from the equation determining school academic performance and thereby serve as an instrument for internet access spending. The source of exogenous variation in internet access spending comes from the invoicing process of the E-rate program, which generates spatial and temporal variation in reimbursement. Specifically, I use the reimbursed amount of internet access spending (i.e. actual funding received) as the instrument for internet access spending (i.e. actual service cost).

A valid instrument needs to meet relevance criterion (i.e., IV should be correlated with the endogenous variable) and exclusion restriction criterion (i.e., IV should relate to the dependent variable only through the endogenous variable). Reimbursement meets the relevance criterion because USAC can process the reimbursement only after (a) USAC has issued a decision letter with a positive funding commitment; (b) school internet access services have started/continued; and (c) applicants and service providers have completed certification forms. That is, reimbursement is highly correlated with funding commitment, which is by definition proportion to internet access spending. Thus, SDIAS should be positively correlated with reimbursement in the same school district. This intuition is supported by the positive correlation ($\rho = .86$, $p < .01$) and strong first-stage results (see Appendix A.6).

The exclusion restriction holds if SDIAS is the only channel through which reimbursement affects school academic performance. In accordance with the reimbursement rules[20], USAC determines the funding commitment only based on the percentage of students eligible for the National School Lunch Program in the school district and the urban or rural status of the school district (the rules are public knowledge), so that school districts can anticipate the funding commitment. However, there is a gap between funding commitment and reimbursement (see Appendix A.6). The process through which such a gap is generated is arguably exogenous to school districts, as it is not within school districts' control but within USAC's control. In addition, reimbursement is unlikely to directly influence academic performance, as student behaviors are largely induced by personal factors (e.g., background) and social-contextual factors (e.g., learning environment) but not by the reimbursement the school district receives, which is unobserved and uncontrolled by students. Finally, according to the reimbursement rules, reimbursement will not materialize until services have started (i.e. school district internet access spending is determined). However, some might argue that reimbursement may still influence academic performance through other channels. For example, school districts in rural areas are likely to get a larger amount of reimbursement, which is associated with better academic performance. However, socioeconomic variables in the first stage regressions account for these concerns (Chan,

[20] https://www.usac.org/sl/applicants/step03/discounts.aspx#school-district

Ghose, and Seamans 2016). Together, I argue that the only channel through which reimbursement impacts academic performance is SDIAS itself.

The IV model is given by the following two-equation system, where (4b) is the first stage and (4a) is the second stage:

(4a) $\mathrm{Y_{ijkt}} = \delta \mathrm{SDIAS_{jkt}} + \boldsymbol{\Theta}'_{\mathbf{11}}\mathbf{Z_{ijkt}} + \Theta_{12}\mathrm{M_{jkt}} + \boldsymbol{\Theta}'_{\mathbf{13}}\mathbf{X_{kt}} + \tau_{1i} + \pi_{1t} + \nu_{1ijkt},$

(4b) $\mathrm{SDIAS_{jkt}} = \gamma \mathrm{Reimbursment_{jkt}} + \boldsymbol{\Theta}'_{\mathbf{21}}\mathbf{Z_{ijkt}} + \Theta_{22}\mathrm{M_{jkt}} + \boldsymbol{\Theta}'_{\mathbf{23}}\mathbf{X_{kt}} + \tau_{2i} + \pi_{2t} + \nu_{2ijkt},$

where $\mathrm{Reimbursment_{jkt}}$ refers to the reimbursed amount that the school district j in county k received in year t. All other variables are as previously defined.

I estimate Equation (3) using robust standard errors that are clustered at the school level to allow for heteroskedasticity and correlated errors within schools over time (Bertrand, Duflo, and Mullainathan 2004), and Equation (4a) and (4b) using two-stage least square (2SLS).

2.3.2. Identification Strategy: School Disciplinary Problems

Given the discrete nature of disciplinary problems, and because many cells have zero counts, the estimates are based on a conditional fixed-effects Poisson specification with enrollment as the exposure variable. Again, I use clustered standard errors at the school level to allow for heteroskedasticity and correlated errors within schools over time (Bertrand, Duflo, and Mullainathan 2004). This model specification has two appealing features: first, like linear models, the Poisson model is not subject to inconsistency caused by the incidental parameters problem associated with fixed effects. Second, overdispersion is corrected by calculating clustered sandwiched standard errors (Cameron and Trivedi 2005). Given the same set of issues concerning identification of the school disciplinary problem equations, I maintain the use of school fixed effects, year fixed effects, and a rich set of school-, school district-, and county-level proxies. The specification is as follow:

(5a)
$$\mu_{ijkt} = \exp\begin{pmatrix} \omega + \gamma_1 SDIAS_{jkt} + \ln(Enrollment_{ijkt}) + \boldsymbol{\gamma_2'N_{ijkt}} \\ +\gamma_3 M_{jkt} + \boldsymbol{\gamma_4'X_{kt}} + u_i + \eta_t \end{pmatrix},$$

(5b)
$$\Pr(Y{=}Y_{ijkt} \mid \mu_{ijkt}, Enrollment_{ijkt}) = \frac{e^{-\mu_{ijkt}Enrollment_{ijkt}} \left(\mu_{ijkt}Enrollment_{ijkt}\right)^{Y_{ijkt}}}{Y_{ijkt}!},$$

where Y_{ijkt} is the number of Part I (Part II) offense-related disciplinary problems in school i in school district j in county k, taking place on year t. u_i and η_t are school fixed effects and year fixed effects, respectively. $\mathbf{N_{ijkt}}$ contain all control variables in $\mathbf{Z_{ijkt}}$ but not enrollment.

To correct for remnant endogeneity of SDIAS in (5a) and (5b), I use the control function approach to include the correction term (residuals) obtained from the first stage that uses reimbursement as the excluded variable. The identification assumptions of using reimbursement are identical to the school academic performance equations.

2.4. Results: School Academic Performance

2.4.1. Econometric Results

For academic performance, I present the results of fixed-effects model in Table 2.3 and those of IV model in Table 2.4. I discuss the results reported in Table 2.4 (which is consistent with the results in Table 2.3).

As shown in Table 2.4, there are significant positive effects of school district internet access spending on 9 out of 10 performance indicators of academic achievement.

Table 2.3 Impact of SDIAS on School Academic Performance (Fixed Effects)

	(1) SAT/ACT meet criterion	(2) AP/IB meet criterion	(3) Advanced course completion	(4) RHSP/DAP graduates	(5) CP all	(6) CP science	(7) CP math	(8) CP reading	(9) CP social studies	(10) CP writing
Internet access spending (millions)	.248***	.262**	.379	1.046***	.269***	.109	.113	.564***	.536***	.291***
	(.091)	(.115)	(.262)	(.327)	(.059)	(.138)	(.091)	(.078)	(.142)	(.112)
Percent of reduced/free lunch students	-.033***	-.022*	-.045***	-.011	-.043***	-.082***	-.057***	-.055***	.007	-.102***
	(.011)	(.013)	(.014)	(.028)	(.005)	(.012)	(.008)	(.007)	(.015)	(.012)
Percent of African American	-.140***	-.092*	-.081**	.013	-.041***	.011	-.148***	-.120***	-.041	-.106***
	(.024)	(.047)	(.037)	(.064)	(.013)	(.030)	(.019)	(.017)	(.040)	(.024)
Percent of Hispanic students	-.119***	-.027	-.041*	.092**	-.078***	-.083***	-.159***	-.152***	-.085***	-.105***
	(.019)	(.028)	(.023)	(.045)	(.009)	(.021)	(.013)	(.012)	(.028)	(.019)
Teacher experience	.181***	.225***	.034	-.114	.246***	.462***	.321***	.225***	.255***	.364***
	(.036)	(.048)	(.051)	(.092)	(.019)	(.046)	(.028)	(.025)	(.059)	(.041)
Student/teacher ratio	.133***	.239***	.111**	-.244**	.114***	.559***	.186***	-.058*	.032	.003
	(.049)	(.080)	(.051)	(.102)	(.021)	(.055)	(.032)	(.030)	(.065)	(.050)
Enrollment (hundreds)	.152***	.070	.316***	.122	.067**	-.184**	-.004	.139***	.256***	.300***
	(.035)	(.046)	(.073)	(.101)	(.028)	(.080)	(.039)	(.040)	(.059)	(.076)
Instructional spending (thousands)	-.134	.117	-.127	-.217	-.011	.305	-.044	-.060	-.142	.056
	(.164)	(.156)	(.181)	(.401)	(.071)	(.198)	(.118)	(.121)	(.220)	(.226)
Unemployment rate	.187**	.088	.024	-.002	.088**	.035	.055	.315***	.322**	.034
	(.082)	(.092)	(.160)	(.243)	(.042)	(.108)	(.070)	(.066)	(.135)	(.105)
Population (millions)	.310	6.299***	9.699***	-2.984	4.850***	14.440***	7.126***	4.675***	4.316***	3.181***
	(.611)	(1.138)	(1.542)	(2.687)	(.462)	(1.139)	(.660)	(.570)	(1.339)	(.851)
Household income (thousands)	-.020	-.017	-.067	.079	.037**	-.057	-.005	.034	.132***	-.004
	(.029)	(.028)	(.049)	(.084)	(.015)	(.036)	(.023)	(.021)	(.045)	(.034)
Intercept	15.027***	-3.156	11.813***	34.072***	-1.431	-18.75***	9.059***	15.539***	-1.529	14.012***
	(1.805)	(2.177)	(3.002)	(5.413)	(1.194)	(2.846)	(1.765)	(1.683)	(3.057)	(2.635)
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered standard error	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	.031	.092	.472	.416	.435	.584	.497	.478	.649	.327
N	14,644	10,366	19,402	16,555	51,547	39,469	52,355	52,316	19,635	38,112

Standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$.
RHSP/DAP refers to Recommended High School/Distinguished Achievement Program; CP refers to commended performance on Texas Assessment of Knowledge and Skills.

Table 2.4 Impact of SDIAS on School Academic Performance (Instrumental Variable)

	(1) SAT/ACT meet criterion	(2) AP/IB meet criterion	(3) Advanced course completion	(4) RHSP/DAP graduates	(5) CP all	(6) CP science	(7) CP math	(8) CP reading	(9) CP social studies	(10) CP writing
Internet access spending (millions)	.334***	.211**	.398**	2.972***	.224***	.247**	.034	.462***	.616***	.304***
	(.129)	(.085)	(.200)	(.412)	(.040)	(.105)	(.058)	(.055)	(.140)	(.093)
Percent of reduced/free lunch students	-.034***	-.021***	-.045***	-.016	-.042***	-.083***	-.057***	-.054***	.007	-.102***
	(.008)	(.007)	(.009)	(.017)	(.004)	(.010)	(.005)	(.005)	(.011)	(.010)
Percent of African American	-.140***	-.093***	-.081***	.011	-.040***	.010	-.147***	-.120***	-.042	-.106***
	(.019)	(.015)	(.021)	(.041)	(.008)	(.021)	(.012)	(.011)	(.026)	(.020)
Percent of Hispanic students	-.119***	-.027**	-.041***	.085***	-.077***	-.084***	-.159***	-.151***	-.086***	-.105***
	(.013)	(.012)	(.015)	(.028)	(.006)	(.015)	(.009)	(.008)	(.019)	(.015)
Teacher experience	.181***	.225***	.034	-.112*	.246***	.462***	.321***	.225***	.255***	.364***
	(.028)	(.028)	(.032)	(.060)	(.013)	(.034)	(.019)	(.018)	(.041)	(.033)
Student/teacher ratio	.132***	.239***	.111***	-.254***	.114***	.559***	.186***	-.057**	.032	.003
	(.042)	(.036)	(.034)	(.061)	(.016)	(.044)	(.024)	(.022)	(.047)	(.044)
Enrollment (hundreds)	.152***	.070***	.316***	.107	.067***	-.185***	-.003	.140***	.255***	.300***
	(.030)	(.021)	(.044)	(.086)	(.020)	(.051)	(.030)	(.029)	(.048)	(.064)
Instructional spending (thousands)	-.134	.115	-.127	-.194	-.014	.316**	-.050	-.067	-.140	.058
	(.112)	(.120)	(.141)	(.267)	(.060)	(.159)	(.089)	(.084)	(.168)	(.174)
Unemployment rate	.187**	.089	.024	-.030	.089**	.031	.057	.317***	.321***	.033
	(.076)	(.072)	(.116)	(.206)	(.035)	(.090)	(.053)	(.051)	(.105)	(.093)
Population (millions)	.371	6.255***	9.711***	-1.577	4.765***	14.692***	6.962***	4.461***	4.463***	3.210***
	(.513)	(.391)	(.745)	(1.447)	(.260)	(.725)	(.396)	(.378)	(.887)	(.656)
Household income (thousands)	-.018	-.019	-.067*	.107	.036***	-.053*	-.007	.032*	.133***	-.004
	(.024)	(.020)	(.036)	(.068)	(.011)	(.030)	(.017)	(.017)	(.036)	(.030)
Intercept	14.981***	-3.089**	11.798***	32.662***	-1.315	-19.09***	9.278***	15.823***	-1.652	13.966***
	(1.490)	(1.453)	(1.998)	(3.780)	(.829)	(2.273)	(1.242)	(1.178)	(2.315)	(2.257)
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	.031	.092	.472	.414	.435	.584	.497	.478	.649	.327
N	14,644	10,366	19,402	16,555	51,547	39,469	52,355	52,316	19,635	38,112

Standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$.

RHSP/DAP refers to Recommended High School/Distinguished Achievement Program; CP refers to commended performance on Texas Assessment of Knowledge and Skills.

First, internet access spending is positively associated with all four college readiness indicators. Specifically, as a school district spends $1 million on internet access, SAT/ACT meet criterion rate increases by .3 percentage points (b = .334, $p < .01$), AP/IB meet criterion rate by .2 percentage points (b = .211, $p < .05$), advanced course completion by .4 percentage points (b = .398, $p < .05$), and RHSP/DAP graduates by 3 percentage points (b = 2.972, $p < .01$).

Second, a $ 1-million increase in SDIAS increases commended performance (all subjects combined) by .2 percentage points (b = .224, $p < .01$), commended performance (science) by .2 percentage points (b = .247, $p < .05$), commended performance (reading) by .5 percentage points (b = .462, $p < .01$), commended performance (social studies) by .6 percentage points (b = .616, $p < .01$), and commended performance (writing) by .3 percentage points (b = .304, $p < .01$).

The estimates suggest that the effect sizes range from .014 to .104 of a standard deviation. In comparison, Dettling, Goodman, and Smith (2018) find that the positive effect of internet access on SAT scores is .003 of a standard deviation. Vigdor, Ladd, and Martinez (2014) find that internet access decreases math test scores by .027 of a standard deviation in North Carolina. The effect sizes here are larger than those documented in prior work. I attribute this result to the use of an investment measure of internet access in school districts (unlike past research that has used the number of broadband providers) and a more comprehensive set of performance indicators (10 in all), as well as a host of parametric and non-parametric control variables.

2.4.2. Substantive Implications: Student Income Value of SDIAS

From a school district's perspective, it is important to convey the effectiveness of SDIAS in a tangible way to demonstrate its value to its customers. Hanushek (2011) reviews studies showing a robust positive link between students' improvement in academic performance and increase in present value of income. I use the link between SDIAS and high school graduates' academic performance, and the link between the improvement in high school graduates' academic performance and increase in future income value shown in the literature (e.g., Hanushek 2011) to convey the financial impact of SDIAS in the form of student's present value of income.

According to the comprehensive empirical literature review in Hanushek (2011), two conclusions are clear: moving a standard deviation (i.e. effect size of 1) in academic performance increases a high school graduate's income by 10-20%, and the average present value of income for fulltime, full-year workers is $1.16 million (Hanushek 2011, p. 471). By implication, moving a standard deviation (i.e. effect size of 1) in academic performance increases a high school student's present income value by $116,000 (i.e. 1.16 million*10%) ~ $232,000 (i.e. 1.16 million*20%). As Table 2.5 shows, increasing SDIAS by $1 million increases SAT/ACT meet criterion rate by an effect size of .03 (obtained by taking the coefficient estimate of .334 and dividing by the standard deviation of SAT/ACT criterion which is 11.3). Increasing SAT/ACT meet criterion rates by .03 of a standard deviation increases a high school student's present income value by $3,480 (i.e. $116,000*.03) ~ $6,960 (i.e. $232,000*.03). Given that the average school district has 239 high school graduates, increasing SDIAS by $1 million increases

Table 2.5 Student Income Value Effect of SDIAS

			Present income value per student with 100% standard deviation change in academic performance		Present income value with the change of effect size in academic performance		Average number of graduates in a school district	Present income value per school district	
	Coefficient	Effect size	Lower bound	Higher bound	Lower bound	Higher bound	N	Lower bound	Higher bound
SAT/ACT meet criterion	.334	.030	$116,000	$232,000	$3,480	$6,960	239	$831,720	$1,663,440
AP/IB meet criterion	.211	.020	$116,000	$232,000	$2,320	$4,640	239	$554,480	$1,108,960
Advanced course completion	.398	.023	$116,000	$232,000	$2,668	$5,336	239	$637,652	$1,275,304
RHSP/DAP graduates	2.972	.104	$116,000	$232,000	$12,064	$24,128	239	$2,883,296	$5,766,592
Average Return								**$1,226,787**	**$2,453,574**

Note. Effect size is calculated using the coefficient obtained from the regression divided by the standard deviation of the corresponding dependent variable. According to a comprehensive empirical literature review in Hanushek (2011), moving a standard deviation (i.e. 100%) in academic performance yields 10 (lower bound)-20 (upper bound) percent higher income per student. Also, the average present value of income for fulltime, full-year workers age 25-70 is $1.16 million, resulting in the present value per student with a standard deviation change in academic performance between $116,000 (i.e. 1.16 million*10%) and $232,000 (i.e. 1.16 million*20%). Applying the effect size of internet access spending, I calculate columns 5-6 (e.g., $116000*0.03=$3,480).

a school district's high school population cumulative present value of income by $831,720 ~ $1,663,440 through an increase in SAT/ACT meet criterion rates.

Applying the same intuition, I calculate how a $1-million increase in SDIAS increases a school district's cumulative present income value among the population of high school graduates through an increase in AB/IB meet criterion rates, advanced course completion rates, and RHSP/DAP graduates (i.e. college readiness indicators relevant to high school graduates). On average, increasing SDIAS by $1 million increases a school district's high school graduate population cumulative present value of income by $1,226,787 ~ $2,453,574 through the improvement in college readiness indicators. Thus, increased SDIAS generates positive payoffs to school districts through the ability to increase school academic performance.

2.4.3. Substantive Implications: Comparing SDIAS with Other School Investments

Though not perfectly comparable (as costs of interventions might differ), the effectiveness of internet access spending is comparable to the effectiveness of other information and communications technology interventions employed by schools such as hardware (e.g., .14 of a standard deviation; Fairlie and London 2012), software (.18 of a standard deviation; Roschelle et al. 2016), class size (.048-.18 of a standard deviation; Angrist and Lavy 1999), and teacher quality (.13 of a standard deviation; Hanushek 2011). Thus, SDIAS is a fairly effective school resource investment when placed alongside other school resources.

2.4.4. Contingent Effect of Household Internet Access on SDIAS Effectiveness

I assess how the effects of SDIAS on school performance are contingent on technology exposure at home, captured by household internet access, which is strongly related to students' internet use at home. Household internet access increases "*the ability to use the internet technology*" not just at home, but also at school (Dewan and Riggins 2005, p. 301; Wei et al. 2011). In other words, home access increases a student's internet use capability, which can be leveraged at school (Malamud and Pop-Eleches 2011). Thus, school internet access should have a stronger effect among students whose internet use capability is heightened due to internet access at home. Internet access within a household also facilitates learning when children are able to reinforce internet-based learning at school and at home on a daily basis. Furthermore, educational institutions serve as a teaching role in the diffusion of the internet (Goldfarb 2006); Thus, the positive effects of school district internet access spending on academic performance are likely to be higher in regions where the level of household internet access is higher.

To test this argument, I use the number of broadband internet service providers in a county to measure household internet access. Kolko (2010) shows that household high-speed internet availability increases monotonically with the number of broadband providers, supporting the use of the number of broadband providers as a proxy for household internet access and usage in other studies in the literature (Dettling, Goodman, and Smith 2018; Sen and Tucker 2019; Vigdor et al. 2014).

Figure 2.2 shows considerable variation in household internet access across counties and over time. Cross-sectional variation in household internet access, conditional on observables, is often driven by exogenous supply-side factors such as weather, terrain, pre-existing infrastructure (see Belo, Ferreira, and Telang 2014, Bhuller et al. 2013; Kolko 2010; Sen and Tucker 2019). School fixed effects should capture much of this variation. Temporal variation in household internet access is arguably exogenous to school districts' decisions as school districts have little control over whether and when providers enter their zip code and little impact on aggregate usage (Dettling, Goodman, and Smith 2018). Indeed, Appendix A.6 shows SDIAS is not driven by household internet access.

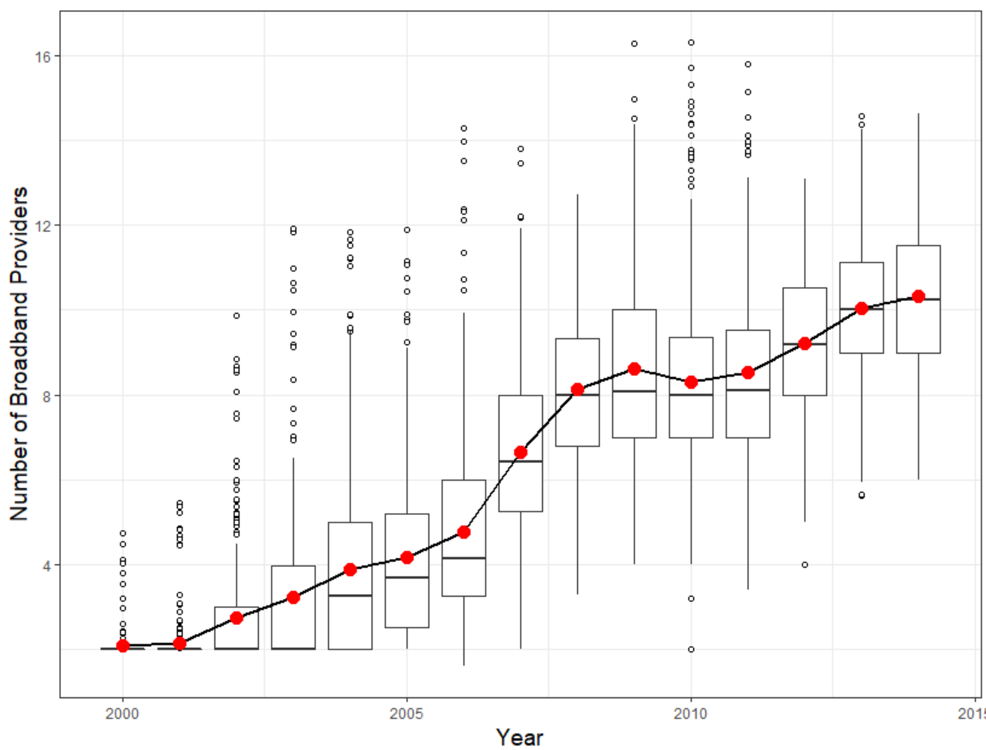


Figure 2.2 Variation in Household Internet Access

Note: The figure shows (a) the overall mean (red dots) and median (black dash lines) distribution of broadband internet service providers for each year during the period 2000–2014, and (b) the boxplot of distribution of broadband Internet service providers each year.

Next, I test the association of SDIAS on school academic performance in sub-samples of high, medium, and low household internet access. Results are reported in Table 2.6. First, there are significant positive effects of SDIAS on 7 out of 10 performance indicators of academic achievement in the sub-sample with a high level of household internet access. Indeed, In Table 2.7, I show that based on the effect sizes across academic performance indicators, the positive effects of SDIAS on academic performance are 9% higher among schools in regions with a high level of household internet access than those in regions with a mean level of household internet access.

Table 2.6 Effects of SDIAS on School Academic Performance: Contingent Role of Household Internet Access

	High household internet penetration			Medium household internet penetration			Low household internet penetration		
Dependent variable	Estimates	SE	N	Estimates	SE	N	Estimates	SE	N
SAT/ACT meet criterion	.257**	(.128)	3908	-.018	(.222)	5221	-.018	(1.864)	5397
AP/IB meet criterion	.336**	(.136)	3716	.041	(.121)	3670	.028	(.217)	2950
Advanced course completion	.424	(.324)	5538	.495	(.321)	6826	2.179	(1.887)	6892
RHSP/DAP graduates	4.341***	(1.040)	4094	-.524	(.752)	5838	9.607***	(2.011)	6483
CP all	.255***	(.0700)	18171	.152***	(.057)	20179	.746	(.795)	13144
CP science	.0645	(.225)	15378	.290*	(.154)	15462	.806	(2.482)	8600
CP math	.115	(.0976)	18209	.071	(.087)	20391	1.445	(1.264)	13701
CP reading	.551***	(.0909)	18220	.276***	(.086)	20425	-1.074	(1.298)	13626
CP social studies	1.066***	(.232)	5708	.412**	(.209)	7579	-4.293	(2.946)	6323
CP writing	.296*	(.154)	14211	.923***	(.147)	14879	2.665	(2.277)	9000
School fixed effects	Yes			Yes			Yes		
Year fixed effects	Yes			Yes			Yes		
Clustered standard errors	Yes			Yes			Yes		

* $p < .1$, ** $p < .05$, *** $p < .01$; Note. The total number of observations of high, medium, and low samples is slightly smaller than those reported in Table 2.4 (for example, 3,908+5,221+5,397=14,526<14,644) as I lose observations when merging my main sample with data on household internet access.

Table 2.7 Comparison of Effects of SDIAS on School Academic Performance: Contingent Role of Household Internet Access

	Mean household internet penetration		High household internet penetration	
Dependent variable	Estimates	Effect size	Estimates	Effect size
SAT/ACT meet criterion	.334	.030	.257	.022
AP/IB meet criterion	.211	.020	.336	.033
Advanced course completion	.398	.023	.424	.022
RHSP/DAP graduates	2.972	.104	4.341	.150
CP all	.224	.024	.255	.027
CP science	.247	.014	.0645	.007
CP math	.034	0	.115	0
CP reading	.462	.033	.551	.036
CP social studies	.616	.037	1.066	.060
CP writing	.304	.021	.296	.017

Note. Effect size is calculated using the coefficient obtained from the regression divided by the standard deviation of the corresponding dependent variable.

In addition, there are significant positive effects of SDIAS on 5 out of 10 performance indicators in the sub-sample with medium household internet access, and significant positive effects of SDIAS on only 1 out of 10 performance indicators in the sub-sample with low household internet access. Evidently, only when household internet access meets a threshold limit, it does complement effective learning at schools investing in internet access.

2.4.5. Robustness Checks

2.4.5.1. Alternative dependent variables

In terms of additional outcomes, I also test the association of SDIAS on a) four-year completion rate (graduation) and b) attendance rate, which are also documented in these two standards. The results are consistent in showing that SDIAS is associated with an improvement in school academic performance (see Table 2.8).

Table 2.8 Summary of Robustness Check: School Academic Performance

	Model 1		Model 2		Model 3		Model 4		Model 5	
Dependent variables	Estimates	SE	Estimates	SE	Estimates	SE	Estimates	SE	Estimates	SE
SAT/ACT meet criterion	.175*	(.099)	.171*	(.099)	.248***	(.091)	.334***	(.129)	.334***	(.129)
AP/IB meet criterion	.244**	(.120)	.242**	(.120)	.262**	(.115)	.211**	(.085)	.226***	(.085)
Advanced course completion	.463*	(.265)	.455*	(.265)	.379	(.262)	.398**	(.200)	.428**	(.199)
RHSP/DAP graduates	1.030***	(.338)	1.027***	(.338)	1.046***	(.327)	2.972***	(.412)	2.993***	(.412)
CP all	.171***	(.061)	.167***	(.061)	.269***	(.059)	.224***	(.040)	.215***	(.039)
CP science	.027	(.137)	.018	(.137)	.109	(.138)	.247**	(.105)	.238**	(.105)
CP math	-.062	(.093)	-.069	(.093)	.113	(.091)	.034	(.058)	.035	(.058)
CP reading	.390***	(.077)	.389***	(.077)	.564***	(.078)	.462***	(.055)	.453***	(.055)
CP social studies	.490***	(.143)	.484***	(.143)	.536***	(.142)	.616***	(.140)	.616***	(.140)
CP writing	.121	(.111)	.122	(.112)	.291***	(.112)	.304***	(.093)	.293***	(.093)
Robustness checks										
Completion rate (graduation)	.704***	(.212)	.709***	(.212)	.743***	(.215)	.723***	(.197)	.733***	(.197)
Attendance rate	.016*	(.008)	.016*	(.008)	.019**	(.008)	.025***	(.009)	.024***	(.009)
School fixed effects	✓		✓		✓		✓		✓	
Year fixed effects	✓		✓		✓		✓		✓	
Grade type	✓		✓		✓		✓		✓	
County-level controls	✓		✓		✓		✓		✓	
Instructional spending per pupil			✓		✓		✓		✓	
School-level controls					✓		✓		✓	
Instrument							✓		✓	
Control for strategic application									✓	

* $p < .1$, ** $p < .05$, *** $p < .01$.

Note that results in Model 4 are the same as those reported in Table 2.4, I present these estimates to present the stepwise procedure.

2.4.5.2. Stepwise inclusion of control variables

I include control variables in a stepwise manner to ascertain the extent to which the estimates are sensitive to the inclusion of time-varying proxies (models 1-3, Table 2.8). I add county-level controls in model 1, school district instructional spending in model 2, school-level controls in model 3. The results are consistent across these stepwise models.

2.4.5.3. Modeling strategic application decision

School districts may choose to apply for funding at a specific stage in their life cycle, so solely focusing on the applicants may produce biased estimates of the internet access spending effect. For instance, if school districts choose to apply for funding following an improvement in academic achievement, then post-funding performance may reflect reversion to the mean and thereby not reflect the impact of internet access spending. To address this concern, I use the number of peer applicants in the same county in the same year as an excluded variable for school districts' application decision (i.e. Heckman selection model). On the one hand, the number of peer applicants in the same county represent the school district's neighbor districts and should be positively correlated with the school district's application decision (social influence). On the other hand, it should only impact academic outcomes through school districts' application decision as the peer districts' application decision is out of control of students in the corresponding school district. Results are consistent with my main specification (model 5, Table 2.8).

2.5. Results: School Disciplinary Problems

2.5.1. Econometric Results

Focusing on school disciplinary problems, I present the results of fixed-effects model (columns 1-4) and those with correct terms (columns 5-8) in Table 2.9. Table 2.9 shows that the effect of SDIAS on Part I offense-related disciplinary problems is not significant (b = .025, *n.s.*), but the effect of SDIAS on Part II offense-related school disciplinary problems is statistically significant and positive (b = .073, $p < .01$). The interpretation is an additional $1-million increase in internet access spending is associated with a 7% increase in Part II disciplinary problems. To get a more granular understanding, I sub-categorize Part II into (a) illegal usage, possession, or exchange and (b) physical violence. The effect of SDIAS on Part II offense-related disciplinary problems are driven by both sub-categories (b = .056, $p < .01$; b = .064, $p < .01$).

2.5.2. Substantive Implications: Disciplinary Costs of SDIAS

From a school district's perspective, it is important to recognize the cost of SDIAS, so the school district can refine its communication to its customers. As Part II offense-related school disciplinary problems represent negative behaviors (e.g., under the influence of drugs, possessing weapons, or assaults), students who engaged in such behaviors receive expulsions and are placed in alternative educational programs such as Disciplinary Alternative Education Programs (DAEPs). Expulsions incur both administrative costs and average daily attendance (ADA) losses: i.e. since a school district's ADA is used to calculate the amount of state aid received, the school district stands to lose money when students miss school days due to expulsions. DAEPs incur

Table 2.9 Impact of SDIAS on School Disciplinary Problems

	(1) Part I	(2) Part II	(3) Part II Type I	(4) Part II Type II	(5) Part I	(6) Part II	(7) Part II Type I	(8) Part II Type II
Internet access spending (millions)	.018	.053***	.033***	.059***	.025	.073***	.056***	.064***
	(.013)	(.006)	(.008)	(.008)	(.016)	(.008)	(.012)	(.010)
Percent of reduced/free lunch	.001	.004***	.001	.007***	.001	.004***	.001	.007***
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Percent of African American	.010***	.001	-.008***	.005*	.010***	.001	-.008***	.005*
	(.003)	(.002)	(.002)	(.002)	(.003)	(.002)	(.002)	(.002)
Percent of Hispanic students	.007***	-.002	-.000	-.004*	.007***	-.002	-.000	-.004*
	(.003)	(.002)	(.002)	(.002)	(.003)	(.002)	(.002)	(.002)
Teacher experience	-.018***	-.019***	-.009**	-.025***	-.018***	-.019***	-.008**	-.025***
	(.005)	(.004)	(.004)	(.005)	(.005)	(.004)	(.004)	(.005)
Student/teacher ratio	-.026***	-.011***	-.017***	-.002	-.026***	-.011***	-.017***	-.002
	(.006)	(.004)	(.004)	(.005)	(.006)	(.004)	(.004)	(.005)
Instructional spending (thousands)	-.001	-.088***	-.030	-.155***	-.001	-.087***	-.029	-.155***
	(.030)	(.022)	(.020)	(.030)	(.030)	(.022)	(.020)	(.030)
Unemployment rate	-.023	.013	.034***	-.007	-.023	.013	.034***	-.007
	(.015)	(.012)	(.012)	(.015)	(.015)	(.012)	(.012)	(.015)
Population (millions)	.003	-.081	-.015	-.123*	.007	-.053	.000	-.112
	(.091)	(.056)	(.065)	(.074)	(.091)	(.057)	(.065)	(.076)
Household income (thousands)	-.001	-.007**	-.005*	-.005	-.001	-.006**	-.005	-.005
	(.004)	(.003)	(.003)	(.004)	(.004)	(.003)	(.003)	(.004)
Residuals					-.017	-.061***	-.062***	-.017
					(.027)	(.013)	(.018)	(.016)
Endogeneity correction	No	No	No	No	Yes	Yes	Yes	Yes
Poisson specification	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Conditional fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered standard errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	58,874	81,919	73,300	75,788	58,874	81,919	73,300	75,788

Standard errors in parentheses * $p < .1$, ** $p < .05$, *** $p < .01$.

operation costs to maintain the student education. Next, I use the link between SDIAS and school disciplinary problems, and information on ADA loss, administrative cost of expulsions, and operation costs of DAEPs to calculate the disciplinary costs of SDIAS.

The first panel of Table 2.10 shows the calculation of costs associated with expulsions; according to Philips (2010), the average cost of expulsions is $170. As with ADA loss, assuming the midpoint of the school year as the average expulsion date, each student loses 90 school days, amounting to $3,780 ADA loss per student (90*$42 ADA daily loss). Combining the $170 administrative cost with the $3,780 ADA loss leads to an average negative cost of $3,950 per student. In addition, increasing SDIAS by $1 million increases Part II offense-related disciplinary problems by 7%, which is equal to approximately 6 more students in a school district having Part II offense-related disciplinary problems per year, amounting to the loss of $23,700.

As with operation costs of DAEPs, the average cost per seat for Dallas Independent School District is $9,410 (Texas Appleseed 2012), while that for Clear Creek Independent School District is $2,500 (Phillips 2010). I use these two numbers as the upper and lower bound of the operation cost per seat. Substantively, the estimates indicate that a $1-million increase in SDIAS is associated with approximately 6 more students having Part II offense-related disciplinary problems per year for the school district. Given that cost per student per year in a disciplinary alternative education program ranging from $2,500 to $9,410, the operation costs of 6 more students having Part II offense-related disciplinary problems per year range from $15,000 (i.e. $2,500*6) ~ $56,460 (i.e. $9,410*6).

Table 2.10 Disciplinary Costs of SDIAS

	Coefficient	Average number of Part II per school	Estimated incremental number of Part II per school district	Calculation of unit cost per expulsion: Number of school days due to expulsions*	ADA rate**	Administrative costs per expulsion*	Unit cost per expulsion	Revenue loss per year for a school district
Part II	.073	10.9	6	90	$42	$170	$3,950	**$23,700**

Note. * Information obtained from Philips (2010) "The Financial Cost of Bullying, Violence, and Vandalism".

**ADA rate is calculated as $\frac{\text{Texas state average revenue per WADA at the compressed tax rate} \times \text{ADA to WADA ratio}}{\text{Number of School Days per Year}}$ = ($5,368*1.4)/180 = $42

	Coefficient	Average number of Part II per school	Estimated incremental number of Part II per school district	Operation cost per student per year in a disciplinary alternative education program		Operation costs per year for a school district	
				Lower bound (Clear Creek ISD)*	Higher bound (Dallas ISD)*	Lower bound	Higher bound
Part II	.073	10.9	6	$2,500	$9,410	**$15,000**	**$56,460**

Note. * the value of lower bound is obtained from Philips (2010) "The Financial Cost of Bullying, Violence, and Vandalism", and the upper bound value is obtained from "Breaking Rules, Breaking Budgets: The Cost of Exclusionary Discipline in Dallas ISD".

To summarize, increased SDIAS generates a total yearly cost of $38,700 to $80,160 for an average school district through an increased number of school disciplinary problems. Note that this is a highly conservative estimate as I did not consider any indirect costs associated with Part II criminal offenses (e.g., enrollment loss) as well as other downstream consequences (e.g., costs due to lawsuits). Indeed, McCollister, French, and Fang (2010) estimate that depending on the severity of crime, the unit crime cost ranges from $3,532 (larceny/theft) to $1,278,424 (murder).

2.5.3. Contingent Effect of Household Internet Access on SDIAS Effectiveness

Increased household internet access has been shown to increase neighborhood crime by facilitating anonymous social interactions and reinforcing negative behaviors (Glaser, Dixit, and Green 2002), accelerating exposure to media violence (Anderson and Bushman 2001), and others' criminal activities (Bhuller et al. 2013).

I present the impact of SDIAS on school disciplinary problems in sub-samples of high, medium, and low household internet access. Table 2.11 shows that the effect of SDIAS on Part I offense-related disciplinary problems is not significant in sub-samples of high, medium, and low household internet access. However, the effect of SDIAS on Part II offense-related disciplinary problems is driven by schools in regions with a high level of household internet access (b = .057, $p < .01$) and a medium level of household internet access (b = .033, $p < .01$), and become insignificant for schools in areas with low household internet access (b = -.035, *n.s.*). Indeed, the negative effects of SDIAS on school disciplinary problems are 19.6% higher among schools in regions with a high level of household internet access than those in regions with a mean level of household internet access (Table 2.12). By implication, providing SDIAS in school districts high household internet penetration needs to be supplemented by investments in monitoring and mitigating any negative repercussions thereof.

Table 2.11 Effects of SDIAS on School Disciplinary Problems: Contingent Role of Household Internet Access

	High household internet penetration			Medium household internet penetration			Low household internet penetration		
Dependent variable	Estimates	SE	N	Estimates	SE	N	Estimates	SE	N
Part I	-.019	(.027)	16,324	.017	(.028)	15,000	-.005	(.026)	14,523
Part II	.057***	(.011)	27,984	.033**	(.015)	24,682	-.035	(.023)	19,439
Poisson specification	Yes			Yes			Yes		
Conditional fixed effects	Yes			Yes			Yes		
Year fixed effects	Yes			Yes			Yes		
Clustered standard errors	Yes			Yes			Yes		

* $p < .1$, ** $p < .05$, *** $p < .01$

Note. The total number of observations of high, medium, and low samples is smaller than those reported in Table 2.9 (e.g., 16,324+15,000+14,523=45,847<58,874) for two reasons: (1) I lose observations when merging my main sample with data on household internet access, and (2) conditional fixed-effects model dropped all groups with only one observation per group and those with all zero outcomes.

Table 2.12 Comparison of Effects of SDIAS on School Disciplinary Problems: Contingent Role of Household Internet Access

	High household internet penetration		Mean household internet penetration	
Dependent variable	Estimates	Estimated effect	Estimates	Estimated effect
Part II	.057	.952	.073	.796

Note. The estimated effect is calculated by multiplying the coefficient by the corresponding sample mean (e.g., 17 (mean of Part II in the high sample)*.057=.952).

2.5.4. Robustness Checks

2.5.4.1. Stepwise inclusion of control variables

Similar to what I have discussed earlier, I include control variables in a stepwise manner and verify that the substantive effects of SDIAS on school disciplinary problems are consistent across models (models 1-3, Table 2.13).

2.5.4.2. Alternative model specification

As observations of all zero outcomes are dropped in the conditional fixed effects specification, I use the random effects Poisson model as an alternative specification when dependent variables are counts of school disciplinary problems. The results are consistent (models 4-5 of Table 2.13).

2.5.4.3. Modeling strategic application decision

I use the number of peer applicants in the same county in the same year as an excluded variable for school districts' application decision (i.e. Heckman selection). The results are consistent (model 8, Table 2.13).

2.5.4.4. Crime in the neighborhood

There might be a concern that the number of criminal activities in the area may simultaneously drive school district internet access spending and school disciplinary problems. Thus, I collect the number of reports of violent crime and poverty crime incidents from National Incident Based Reporting System (NIBRS) collected by the Federal Bureau of Investigation (FBI) at the county level since 2004. Results are robust to the inclusion of this control (model 9, Table 2.13).

Table 2.13 Summary of Robustness Check: School Disciplinary Problems

	Model 1		Model 2		Model 3		Model 4		Model 5	
Dependent variables	Estimates	SE	Estimates	SE	Estimates	SE	Estimates	SE	Estimates	SE
Part I	.014	(.012)	.014	(.012)	.011	(.012)	.011	(.014)	.011	(.014)
Part II	.062***	(.006)	.060***	(.006)	.056***	(.006)	.076***	(.008)	.071***	(.008)
Part II Type I	.036***	(.009)	.036***	(.009)	.033***	(.008)	.054***	(.012)	.053***	(.012)
Part II Type II	.074***	(.008)	.070***	(.008)	.063***	(.008)	.069***	(.010)	.066***	(.010)
Year fixed effects	✓		✓		✓		✓		✓	
Grade type	✓		✓		✓		✓		✓	
County-level controls	✓		✓		✓		✓		✓	
Instructional spending per pupil			✓		✓		✓		✓	
School-level controls					✓		✓		✓	
Instrument							✓		✓	
Control for strategic application									✓	
Conditional fixed effects Poisson										
Neighborhood crime										

	Model 6		Model 7		Model 8		Model 9	
Dependent variables	Estimates	SE	Estimates	SE	Estimates	SE	Estimates	SE
Part I	.018	(.013)	.025	(.016)	.025	(.016)	-.001	(.018)
Part II	.053***	(.006)	.073***	(.009)	.070***	(.008)	.055***	(.009)
Part II Type I	.033***	(.008)	.056***	(.012)	.056***	(.012)	.034***	(.013)
Part II Type II	.059***	(.008)	.064***	(.010)	.063***	(.010)	.060***	(.011)
Year fixed effects	✓		✓		✓		✓	
Grade type	✓		✓		✓		✓	
County-level controls	✓		✓		✓		✓	
Instructional spending per pupil	✓		✓		✓		✓	
School-level controls	✓		✓		✓		✓	
Instrument			✓		✓		✓	
Control for strategic application					✓		✓	
Conditional fixed effects Poisson	✓		✓		✓		✓	
Neighborhood crime							✓	

* $p < .1$, ** $p < .05$, *** $p < .01$.

Note that results in Models 6 and 7 are the same as those reported in Table 2.9, I present these estimates to present the stepwise procedure.

2.6. Validating Effect Heterogeneity: Technology Exposure Measure

On the positive side, SDIAS could increase student exposure to new learning tools to monitor their learning progress and facilitate social transmission of new knowledge (Dabbagh and Kitsantas 2012). On the negative side, SDIAS could accelerate exposure to age-restricted content and media violence, which increases the propensity to behave aggressively in school (Anderson and Bushman 2001). Accordingly, I argue that SDIAS increases both school academic performance and school disciplinary problems through increasing students' *exposure to internet-based information* at school.

To further validate this source of effect heterogeneity, I collect school-level internet download speed data in 2016 from USAC open platform and matched 6,986 schools with the main sample. The average school internet download speed is 2,810 Mbps with a standard deviation 5,206 Mbps. When I split the sample into high and low download speed (i.e., internet download speed is above and below the median 1,024 Mbps respectively), I find most of positive effects of SDIAS on school academic performance and those effects on Part II are concentrated among schools with a high internet download speed (see Table 2.14). These results verify that both the return (school academic performance) and risk (school disciplinary problems) are magnified among schools with a higher internet download speed, which bring faster access to online information and enable more usage.

Table 2.14 Validating Results: School Internet Download Speed

	High Download Speed			Low Download Speed		
Dependent variable	Estimates	SE	N	Estimates	SE	N
SAT/ACT meet criterion	.371***	(.107)	6,324	.351	(.485)	6,543
AP/IB meet criterion	.169*	(.103)	5,853	-.225	(.243)	3,439
Advanced course completion	.188	(.239)	7,870	.801	(.621)	7,551
RHSP/DAP graduates	2.967***	(.421)	6,665	-.815	(1.247)	6,788
CP all	.218***	(.044)	30,150	.630	(.469)	13,817
CP science	.145	(.116)	23,631	3.654***	(1.170)	10,570
CP math	-.0282	(.063)	30,361	2.289***	(.778)	14,153
CP reading	.436***	(.059)	30,275	-.616	(.756)	14,035
CP social studies	.532***	(.141)	9,758	1.647	(1.856)	6,623
CP writing	.152	(.099)	23,841	1.879	(1.306)	8,988
School fixed effects	Yes			Yes		
Year fixed effects	Yes			Yes		
Clustered standard errors	Yes			Yes		

	High Download Speed			Low Download Speed		
Dependent variable	Estimates	SE	N	Estimates	SE	N
Part I	.011	(.019)	33,737	.067	(.044)	16,408
Part II	.073***	(.009)	48,676	.028	(.049)	20,996
Part II Type I	.053***	(.014)	43,940	.018	(.039)	18,820
Part I Type II	.063***	(.011)	45,896	.043	(.076)	18,870
Poisson specification	Yes			Yes		
Conditional fixed effects	Yes			Yes		
Clustered standard errors	Yes			Yes		
Year fixed effects	Yes			Yes		

* $p < .1$, ** $p < .05$, *** $p < .01$.

Note. The total number of observations of high sample and low sample is smaller than those reported (e.g., 6,324+6,543=12,867<14,644) as I lose observations when merging my main sample with data on school internet download speed.

2.7. Conclusion

Even though many parents and school district administrators advocate investing in internet access to improve academic outcomes, the contribution of SDIAS to school performance is *ex-ante* ambiguous. I quantify the effect of SDIAS on school academic performance and school disciplinary problems in Texas. Increased school district internet access spending simultaneously enhances school performance in 9 out of 10 performance indicators, and drives a 7% increase in the number of school disciplinary problems related to Part II offenses. These effects are exacerbated in regions where households have better internet access. I establish that these effects by a combination of different identification strategies and a rich set of robustness checks to rule out unobservable factors that might potentially drive the results.

Future research can investigate the degree to which resource reallocation to safety technology (e.g., in Indiana, Social Net Watcher detects dangerous words posted on social media and alert the school administrators) or school policies and initiatives (e.g., digital citizenship training) reduce disciplinary problems. Also, while I present suggestive evidence of the heterogeneity, future research can provide more direct evidence on the school-level heterogeneity by using information on internet search behavior or internet traffic. Third, I only examine internet access spending as an aggregate metric, so further research could assess the impact of internet access spending at a more disaggregated level by gathering information on functional areas where the spending is allocated (e.g., Wi-Fi) to guide the optimal allocation of internet access spending. Finally, researchers could generate empirical generalizations for other regions.

2.8. References

Anderson, Craig A., and Brad J. Bushman (2001), "Effects of Violent Video Games on Aggressive Behavior, Aggressive Cognition, Aggressive Affect, Physiological Arousal, and Prosocial Behavior: A Meta-Analytic Review of the Scientific Literature," *Psychological Science*, 12 (5), 353–59.

Angrist, Joshua D., and Victor Lavy (1999), "Using Maimonides' Rule to Estimate the

Effect of Class Size on Scholastic Achievement," *The Quarterly Journal of Economics*, 114 (2), 533–75.

——— and ——— (2001), "Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools," *Journal of Labor Economics*, 19 (2), 343–69.

Belo, Rodrigo, Pedro Ferreira, and Rahul Telang (2014), "Broadband in School: Impact on Student Performance," *Management Science*, 60 (2), 265–82.

———, ———, and ——— (2016), "Spillovers from Wiring Schools with Broadband: The Critical Role of Children," *Management Science* 62 (12), 3450-71.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004), "How Much Should We Trust Differences-in-Differences Estimates?," *The Quarterly Journal of Economics*, 119 (1), 249–75.

Bhuller, Manudeep, Tarjei Havnes, Edwin Leuven, and Magne Mogstad (2013), "Broadband Internet: An Information Superhighway to Sex Crime?," *Review of Economic Studies*, 80 (4), 1237–66.

Black Sandra E. (1999), "Do Better Schools Matter? Parental Valuation of Elementary Education," *The Quarterly Journal of Economics*, 114 (2), 577-99.

Cameron, A. Colin, and Pravin K. Trivedi (2005), "Microeconometrics: Methods and Applications," Cambridge: Cambridge University Press.

Chan, Jason, Anindya Ghose, and Robert Seamans (2016), "The Internet and Racial Hate Crime: Offline Spillovers from Online Access," *MIS Quarterly*, 40 (2), 381–403.

Dabbagh, Nada and Anastasia Kitsantas (2012), “Personal Learning Environments, Social Media, and Self-Regulated learning: A Natural Formula for Connecting Formal and Informal Learning,” *Internet and Higher Education*, 15 (1), 3–8.

Dettling, Lisa J., Sarena Goodman, and Jonathan Smith (2018), “Every Little Bit Counts: The Impact of High-Speed Internet on the Transition to College," Review of Economics and Statistics, 100 (2), 260-73.

Dewan, Sanjeev, and Frederick J. Riggins (2005), “The Digital Divide: Current and Future Research Directions,” *Journal of the Association for Information Systems*, 6 (12), 298-337.

DiMartino, Catherine, and Sarah Butler Jessen (2016), “School Brand Management: The Policies, Practices, and Perceptions of Branding and Marketing in New York City’s Public High Schools,” *Urban Education* 51 (5), 447-75.

EducationSuperHighway (2018), “2018 State of the States”.

Faber, Benjamin, Rosa Sanchis-Guarner, and Felix Weinhardt (2016), “Faster Broadband: Are There Any Educational Benefits?,” Working paper No. 480, Centre for Economic Performance, London School of Business.

Fairlie, Robert W., and Rebecca A. London (2011), “The effects of home computers on educational outcomes: Evidence from a field experiment with community college students,” *The Economic Journal*, 122 (561), 727-53.

Glaser, Jack, Jay Dixit, and Donald P. Green (2002), “Studying Hate Crime with the Internet: What Makes Racists Advocate Racial Violence?,” *Journal of Social Issues*, 58 (1), 177–93.

Goldfarb, Avi (2006), “The (Teaching) Role of Universities in the Diffusion of the Internet,” *International Journal of Industrial Organization*, 24 (2), 203-25.

Goolsbee, Austan, and Jonathan Guryan (2006), “The Impact of Internet Subsidies in Public Schools,” *Review of Economics and Statistics*, 88 (2), 336–47.

Hanushek, Eric A (2011), “The Economic Value of Higher Teacher Quality,” *Economics of Education Review*, 30 (3), 466-79.

Hazlett, Thomas W., Ben Schwall, and, Scott Wallsten (2019), “The Educational Impact of Broadband Subsidies for Schools under E-rate, *Economics of Innovation and New Technology*, 28 (5), 483-97.

Jacob, Brian A., and Lars Lefgren (2007), “What Do Parents Value in Education? An Empirical Investigation of Parents' Revealed Preferences for Teachers,” *The Quarterly Journal of Economics*, 122 (4), 1603–37.

Kolko, Jed (2010), “A New Measure of US Residential Broadband Availability,” *Telecommunications Policy*, 34 (3), 132–43.

Kotler, Philip, and Sidney J. Levy (1969), “Broadening the Concept of Marketing." *Journal of Marketing*, 33 (1), 10-15.

Lindo, Jason M., Peter Siminski, and Isaac D. Swensen (2018), “College Party Culture and Sexual Assault,” *American Economic Journal: Applied Economics*, 10 (1), 236-65.

Lubienski, Christopher (2007), “Marketing schools: Consumer Goods and Competitive Incentives for Consumer Information,” *Education and Urban Society*, 40 (1), 118-41.

——— (2003), “Innovation in Education Markets: Theory and Evidence on the Impact of Competition and Choice in Charter Schools,” *American Educational Research Journal*, 40 (2), 395-443.

Malamud, Ofer and Cristian Pop-Eleches (2011), “Home Computer Use and the Development of Human Capital,” *Quarterly Journal of Economics*, 126 (2), 987–1027.

McCollister, Kathryn E., Michael T. French, and Hai Fang (2010), “The cost of crime to society: New crime-specific estimates for policy and program evaluation,” *Drug and Alcohol Dependence*, 108 (1-2), 98-109.

Oplatka, Izhar (2007), “The Place of the “Open House” in the School-Choice Process: Insights from Canadian Parents, Children, and Teachers,” *Urban Education*, 42, 163-84.

Phillips, Rick (2010), “The Financial Costs of Bullying, Violence, and Vandalism,” Reston, VA: NASSP, 28-29.

Rabovsky, Thomas (2011), “Deconstructing School Choice: Problem Schools or Problem Students?,” *Public Administration Review*, 71 (1), 87-95.

Roschelle, Jeremy, Mingyu Feng, Robert F. Murphy, and Craig A. Mason (2016), “Online Mathematics Homework Increases Student Achievement,” AERA Open, 2 (4), 1–12.

Sacerdote, Bruce (2001), “Peer Effects with Random Assignment: Results for Dartmouth Roommates,” *The Quarterly Journal of Economics*, 116 (2), 681–704.

Sen, Ananya, and Catherine E. Tucker (2019), “Information Shocks and Internet Silos:

Evidence from Creationist Friendly Curriculum," *Working Paper*, 1–56.

Skiba, Russell J., Choong-Geun Chung, Megan Trachok, Timberly L. Baker, Adam Sheya, and Robin L. Hughes (2014), "Parsing Disciplinary Disproportionality: Contributions of Infraction, Student, and School Characteristics to Out-of-School Suspension and Expulsion," *American Educational Research Journal*, 51 (4), 640–70.

Stevens, Ann H., Douglas L. Miller, Marianne E. Page, and Mateusz Filipski (2015), "The Best of Times, the Worst of Times: Understanding Pro-Cyclical Mortality," *American Economic Journals: Economic Policy*, 7 (4), 279–311.

Texas Appleseed (2012), "Breaking Rules, Breaking Budgets: Cost of Exclusionary Discipline in 11 Texas School Districts," 1-87.

Vigdor, Jacob L., Helen F. Ladd, and Erika Martinez (2014), "Scaling the Digital Divide: Home Computer Technology and Student Achievement," *Economic Inquiry*, 52 (3), 1103-19.

Violette, Diana (2017), "A Study of Internet Spending and Graduation Rates: A Correlational Study," Doctoral dissertation, University of Central Florida.

Wei, Kwok-Kee, Hock-Hai Teo, Hock Chuan Chan, and Bernard C. Y. Tan (2011), "Conceptualizing and Testing a Social Cognitive Model of the Digital Divide," *Information Systems Research*, 22 (1), 170–87.

Wooldridge, Jeffrey M. (2010), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press.

3. IMPROVING CANCER OUTREACH EFFECTIVENESS THROUGH TARGETING AND ECONOMIC ASSESSMENTS: INSIGHTS FROM A RANDOMIZED FIELD EXPERIMENT

3.1. Introduction

In 2018, over 1.7 million new cases of cancer were diagnosed in the United States, and the cost of cancer care surpassed $147 billion (National Cancer Institute 2018). Following the guidelines of the National Comprehensive Cancer Network[21], healthcare institutions encourage at-risk patients to undergo regular screening as it opens the door for early detection, more cost-effective treatment options, and better recovery prognosis. Regular screening reduces mortality rates for lung (28% drop; McMahon et al. 2008), breast (24% drop; Nyström et al. 1993), and liver (37% drop; Zhang, Yang, and Tang 2004) cancers. Moreover, cancer screening can reduce annual treatment costs by nearly $5,000 (Benoit, Grönberg, and Naslund 2001).

Accordingly, healthcare institutions invest heavily in direct-to-patient outreach interventions to increase screening completion among at-risk patients. For example, Johns Hopkins Hospital's cancer center uses e-mails, letters, seminars, and community events to encourage screening completion among patients (Johns Hopkins Medicine 2019). Nascent evidence shows that outreach interventions can increase cancer screening completion rates (e.g., Singal et al. 2019). Yet, more needs to be done to improve the

[21] https://www.nccn.org/patients/guidelines/content/PDF/hepatobiliary-patient.pdf

effectiveness of outreach interventions. With 1.7 million outreach interventions launched in 2015, and $123 million spent on prevention and education efforts,[22] only 8% of US adults over 35 utilize preventive services (Borsky et al. 2018). This percentage is too low. Hence, a critical research priority is to improve outreach effectiveness (Marketing Science Institute 2016).

Healthcare institutions face three major challenges in improving outreach effectiveness. First, most studies examine only the main effects of medical interventions (e.g., Singal, Pillai, and Tiro 2014), neglecting variation due to patient demographics, health status, visit history, health system accessibility, and neighborhood socioeconomic status (Figueroa and Jha 2018). Examining heterogeneity in customer response to marketing interventions (Allenby and Rossi 1998; Ascarza 2018), if suitably applied, can help healthcare institutions implement "personalized healthcare marketing" to boost outreach effectiveness.

Second, medical scholars mainly evaluate the outreach effectiveness (or compare the relative efficacy of outreach interventions) in a single period (e.g., Basch et al. 2006). Given the importance of regular screening compliance over multiple periods (Chubak and Hubbard 2016), it is critical to incorporate each patient's prior screening compliance. Understanding patient dynamics can help healthcare institutions improve outreach effectiveness through "dynamic personalization" over multiple periods.

Third, quantifying the return on outreach interventions to incorporate the health

---

[22] https://www.cancer.org/content/dam/cancer-org/online-documents/en/pdf/infographics/where-does-your-money-go-infographic-print.pdf.

benefits and financial costs of these interventions will help healthcare institutions communicate the tangible value they bring to the community and enable funding agencies to sustain these interventions (Andersson et al. 2008). As the director of cancer education for the Stanford Cancer Center notes, "durable, long-term solutions will require a substantial investment in academic/community partnerships to improve cancer education."[23]

To addresses these three challenges, I use a multi-period pragmatic randomized field experiment conducted at a large hospital system with at-risk patients for hepatocellular carcinoma (HCC), the most common type of primary liver cancer (Singal et al. 2019). Patients were randomly assigned (1:1:1) to three different conditions—usual care, outreach alone, or outreach with patient navigation. *Usual care* is the baseline condition where physicians offer preventive care recommendations at their discretion during a patient's usual care visits. As described in detail later, *outreach alone* and *outreach with patient navigation* provide two different levels of direct marketing efforts based on outreach mails, outreach calls, and customized motivational education by trained patient navigators. The focal outcome is the patient's screening completion status within 6 months (Period 1), 6-12 months (Period 2), and 12-18 months (Period 3) of the initial randomization. This enables an investigation of the impact of outreach interventions on regular screening compliance. I evaluated screening completion status every 6 months, as this interval has been demonstrated to increase early detection and

[23] http://med.stanford.edu/news/all-news/2010/03/rhoads-named-director-of-cancer-outreach-programs-for-stanford-cancer-center-cancer-prevention-institute-of-california.html

survival compared to longer screening intervals (Santi et al. 2010). To incorporate patient heterogeneity, I iteratively construct the focal covariates based on the extant medical literature and pragmatic considerations that the study design affords, including patients' demographics, health status, visit history, health system accessibility, neighborhood socioeconomic status, and prior screening compliance.

Relative to the baseline condition, outreach alone (outreach with patient navigation) increases screening completion rates by 10-20 (13-24) percentage points, but the effectiveness of two outreach interventions does not significantly differ. Yet, central to this article, these main effects mask considerable heterogeneity in outreach effectiveness and precluding patient-level targeting (Blanchard et al. 2012; Hutchinson, Kamakura, and Lynch 2000). I uncover patient-level treatment effects of these two interventions using causal forests, a state-of-the-art development in the machine learning and economics literature (Wager and Athey 2018). I find that: (1) compared to outreach alone, outreach with patient navigation induces a higher proportion of patients with significant positive heterogeneous treatment effects in Periods 2 (9%) and 3 (23%); (2) the increase in screening completion as a result of either outreach alone or outreach with patient navigation is higher for patients who are female, minority, in better health status, with a more frequent visit history, covered by medical-assistance insurance, reside in closer proximity to clinics, and reside in a more populated neighborhood; (3) the increase in screening completion as a result of outreach alone is higher for patients who are younger, commute faster, and reside in a neighborhood with more public insurance coverage; in contrast, screening completion as a result of outreach with patient navigation is higher for patients who are older and reside in a higher-income neighborhood.

Incorporating these patient-level differences in their responsiveness to outreach interventions, and a well-established scheme of cost-benefit calculation that quantifies health benefits and financial costs associated with outreach interventions (e.g., Goossens et al. 2017), I assign patients to the baseline, outreach alone, or outreach-with-patient-navigation condition in each period based on their predicted treatment effect and predicted net return. As a result, the commensurate return on the patient-level targeted outreach program is \$3,704,270-\$4,167,419 when extrapolated to 3,217 eligible patients in the hospital's database. The targeted outreach program improves the return on the randomized outreach program (\$2,130,921) by 74%-96%.

I make several contributions to theory and practice. First, the literature on marketing interventions in healthcare typically relies on experimentally-manipulated and theory-driven moderators, such as test accuracy (Luce and Kahn 1999) or consumer goals (Wang, Keh, and Bolton 2010), which are not only difficult to measure but are also impractical for healthcare institutions to implement. In contrast, healthcare institutions can readily utilize *observable* patient characteristics—such as ethnicity, visit history, and insurance coverage—that are of theoretical relevance. The bulk of the marketing literature has focused on attitudinal consequences of health messages and communications using self-reported measures such as behavioral intentions (e.g., Bolton et al. 2008), risk perceptions (e.g., Menon, Block, and Ramanathan 2002), and attitudes (e.g., Basil and Brown 1997) in a lab setting. While insightful, they are of little practical relevance to addressing actual behaviors.

Second, I contribute to the medical literature on cancer outreach effectiveness, which has focused primarily on the main effects of cancer outreach interventions from randomized field studies. For marketing, the causal-forests approach provides a practical way to improve the efficacy of experimental studies by systematically exploring the treatment-effect heterogeneity across intervention types, across patient subgroups, and over time without pre-specifying the sources of heterogeneity. The application of causal forests helps alleviate the marketing field's concern about external validity due to factors that are not explicitly manipulated or modeled, limiting the generalizability of findings beyond what is being studied (Cook and Campbell 1979; Lynch 1982).

Third, I provide insights into what patient subgroup benefits more (less) from outreach interventions, offer ways to customize the interventions, and can help practitioners allocate limited financial resources to those with the largest potential gains. For example, while outreach programs typically target diverse, socioeconomically difficult-to-reach disadvantaged patient populations to improve their health outcomes (Singal et al. 2019), patients more responsive to outreach interventions tend to be female, minority, in good health status, with more frequent visit history, covered by medical-assistance insurance, reside in closer proximity to clinics, and reside in more populated neighborhoods. Thus, simply targeting one or two patient characteristics may not maximize the gains from the outreach interventions.

Fourth, my approach provides a roadmap for implementing personalized healthcare marketing by customizing outreach interventions and quantifying the return on such interventions. Using patient-level treatment effect estimates with valid confidence intervals, I provide not only a tool that can recommend the most suitable intervention for each patient given their profile but also an individual-level cost-benefit analysis to measure the return on personalized healthcare marketing investments.

3.2. Institutional Setting, Study Design, and Data

3.2.1. Institutional Setting: Cancer Outreach and Importance of Regular Screening

The field experiment is based on the cancer outreach efforts of a large hospital system to increase regular screening completion for early detection of hepatocellular carcinoma (HCC), the most common form of liver cancer, among patients with higher risk of HCC. Most patients with liver cancer do not display symptoms until it reaches an advanced stage. As such, they often miss the time window during which treatment options, such as transplant and surgical resection, are available. The 5-year survival rate for early-stage liver cancer patients who undergo surgical therapy is 60%-70%, while the 5-year relative survival rate for liver cancer is 18% (American Cancer Society 2019). Yet, the utilization rate of HCC screening is below 20% in the general cirrhotic population, and even lower among low socioeconomic-status and non-Caucasian patients (Singal et al. 2012).

Because underuse of HCC screening is one of the most common causes for late-stage tumor detection, contributing to poor overall survival (Singal, Marrero, and Yopp 2014), the outreach program was designed to promote *regular screening*. To guide the best clinical practice, American Association for the Study of Liver Diseases and National Comprehensive Cancer Network have issued evidence-based recommendations for regular screening in high-risk populations such that regular screening is recommended to be performed every 6 months (e.g., Marrero et al. 2018; Tzartzeva et al. 2018). This screening interval was initially based on tumor doubling times, but it has been shown to 1) be superior for early detection than longer intervals (e.g., 12 months; Santi et al. 2010) but not inferior to shorter intervals (e.g., 3 months; Trinchet et al. 2011), and 2) minimize patient and provider burden (Bruix and Sherman 2011).

The hospital system conducted a randomized trial between December 2014 and March 2017. The study was approved by the University of Texas Southwestern Medical Center Institutional Review Board. The trial protocol is available on clinicaltrials.gov (NCT02312817), where the study is registered. The random assignment (1:1:1) consisted of one baseline condition (no outreach) and two conditions with outreach interventions (outreach alone and outreach with patient navigation) with the outcome being HCC screening completion status.[24]

3.2.2. Study Design

The eligibility criteria for patient inclusion using established norms have been developed in the medical field (see Appendix B.1 for details). From the 3,217 eligible patients in the hospital's database, 1,800 patients were randomly selected for the study.[25]

Focal independent variable: intervention type. As summarized in Figure 3.1, each patient was randomly assigned to one of three conditions in a 1:1:1 ratio:

---

[24] The hospital system is the sole safety-net provider for Dallas County, which minimizes omitted variable bias that could emanate from competitive efforts by other organizations in the area.
[25] The hospital system obtained a waiver of informed consent to minimize volunteer bias.

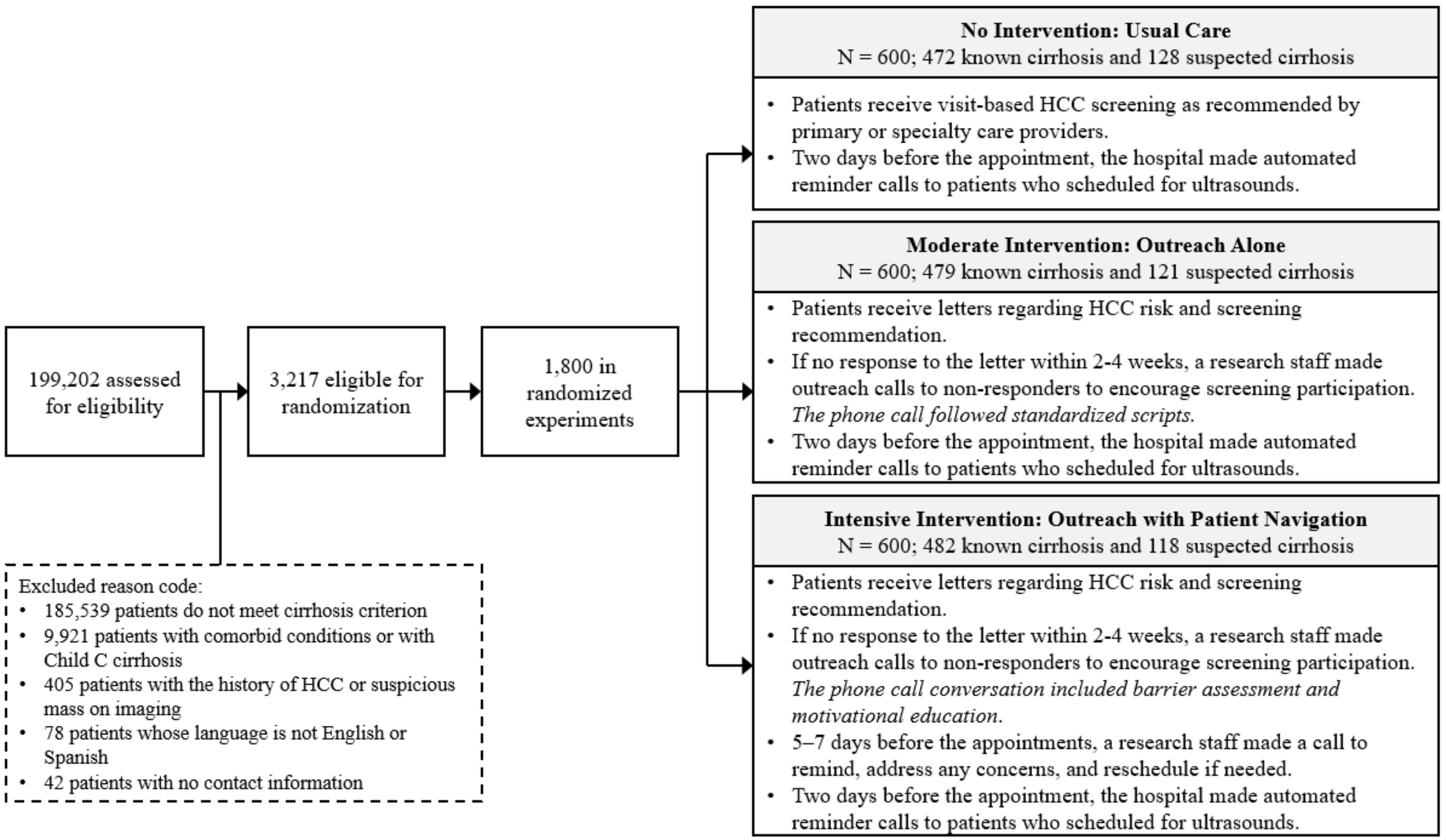


Figure 3.1 Study Design: Cancer Outreach Interventions

Notes. Patients in the outreach-alone and outreach-with-patient-navigation conditions could receive usual care.

- No outreach or usual care (baseline condition): Patients received visit-based HCC screening as recommended by primary or specialty care providers and were not contacted by the outreach marketing team. For patients who scheduled ultrasounds, the hospital system placed automated reminder telephone calls two days before the ultrasound appointments.
- Outreach-alone intervention: Like in the baseline condition, patients were eligible for usual care, as offered through their usual outpatient encounters. Patients were also mailed a one-page letter, which contained information on the risk of HCC in patients with cirrhosis and the benefits and risks of HCC screening, a brief summary of the screening procedure, and a recommendation to the patient to make an appointment for an ultrasound (see Appendix B.1 for details). In order to increase participation, the staff then made outreach calls to non-responders (i.e., patients with returned mail, and those who did not respond to mailed invitations within 2-4 weeks). During telephone calls, trained research staff followed standardized scripts. Mails and telephone calls were in English or Spanish, depending on patients’ preferences. Also, the hospital system placed automated reminder telephone calls two days before appointments for patients who scheduled ultrasounds.
- Outreach-with-patient-navigation intervention: Like patients in the baseline condition and those in the outreach-alone condition, patients in this condition were eligible for care as offered through their usual outpatient encounters. Patients in this condition had an experience identical to those in the outreach-alone condition, with two additions: (1) the way the research staff communicated during outreach

telephone calls, and (2) an additional reminder call from the research staff. During telephone calls, if patients in this condition declined to make an appointment for screening, the research staff used a standardized telephone script to identify potential barriers and then provided customized motivational messages to encourage screening participation. Examples of barriers include preparation involved, pain during the test, etc. (see Appendix B.1 for details). For instance, if a patient is concerned about "prep involved" with the screening, the research staff alleviates this concern by stating, "a liver ultrasound is a quick procedure. The ultrasound usually takes less than 30 minutes and the appointment should take around 1 hour from start to finish." For scheduled ultrasounds, the hospital system's research staff called the patients 5-7 days before the appointments to provide a reminder, address any concerns, and reschedule the appointment if needed. For these patients, the hospital system also placed automated reminder telephone calls two days before the ultrasound appointments. Overall, as shown in Figure 3.1, this condition is the most intense and comprehensive intervention in the study.

Multi-period study design and sample sizes in Periods 2 and 3. To achieve the goal of encouraging regular screening completion, the study repeated the outreach-alone and outreach-with-patient-navigation interventions after 6 months and then after 12 months of the initial randomization. I define Period 1 as the time within 6 months of the first randomization, Period 2 as the time between month 6 and month 12 since the first randomization, and Period 3 as the time between month 12 and month 18 since the first randomization. In summary, the hospital system undertook the outreach interventions in all three periods, each period being six months apart, and each patient belonging to the same condition across the three periods.

Once a patient has completed the screening in the first period, the patient does not exit the pool and will still be contacted in the next six-month period. There are two exceptions to the repeated interventions: (1) if the patient completes the screening and is diagnosed with HCC during the experiment, the patient exits the pool as the providers must refer the patient for HCC treatment instead of routine screening; (2) if the patient completes the screening and is deceased during the experiment, the patient cannot complete the screening in later periods. As a result, the sample size is 1,800 for Period 1, 1,772 for Period 2, and 1,743 for Period 3[26]. The sample size in the baseline, outreach-alone, and outreach-with-patient-navigation condition is (600, 600, 600) for Period 1, (591, 592, 589) for Period 2, and (577, 584, 582) for Period 3.

Dependent variable: screening completion status. Screening completion status is measured as a patient getting an abdominal imaging screening test (1) or not (0). I observe the dependent variable for each patient in Periods 1, 2, and 3.

3.2.3. Constructing Focal Covariates: An Iterative Approach

Taking theoretical and pragmatic considerations into account, I followed a four-step iterative approach to determine the focal covariates that inform the patient heterogeneity in response to outreach interventions. This process of including covariates

[26] Out of 28 patients excluded in Period 2, 12 were excluded as they were diagnosed with HCC in Period 1, and 16 were deceased. Out of 57 patients excluded in Period 3, 23 were excluded due to being diagnosed with HCC in Period 2, and 38 were deceased (4 of them are both diagnosed with HCC and deceased).

starts from original yet tentative variables that researchers have in hand, and calls for additional data collection guided by a multifaceted understanding of theory models, past studies, research questions, and practice. The approach resembles a theory-in-use process (e.g., Zeithaml et al. 2020) whereby I used an exploratory and confirmatory stance to represent how heterogeneity interacts with treatment, and it requires extensive and systematic iterations.

Step 1: Utilize original variables. I begin with the variables that are available in the electronic medical record system (EMR) and are relevant to practitioners and inspired by "real-world phenomena"; see MacInnis et al. forthcoming) and are well-documented in the medical research (thus relevant to academic scholars). EMRs store and track key patients' information to better serve their needs[27]. The system provides patient demographics (e.g., Wetherell et al. 2013), health status (e.g., Ferrante, Chen, and Kim 2008), and visit history record (Skolnik 2011). As shown in Table 1, past studies in healthcare have analyzed these "ready-for-use" variables in EMR (e.g., Humiston et al. 2011; McCarthy et al. 2018).

Step 2: Construct theoretically-relevant variables. I use the information available in the EMR to construct new variables that are not captured by the raw unrefined data but draw upon theories such as health belief model and protection motivation theory (e.g., Moorman and Matulich 1993; Lisjak and Lee 2014). I use a patient's health insurance and location information to construct variables that proxy a patient's insurance

[27] In 2017, 85.9% of office-based physicians in the US used an EMR system.

coverage (financial access to care) and proximity to clinics (geographical access to care). This is consistent with the research showing that health system accessibility and "improving health system accessibility across the socio-economic spectrum" (p. 19)[28] is a strategic priority for policy makers.[29]

Step 3: Explore external secondary data sources. To supplement the above, I also gather additional data from external secondary sources. In particular, socioeconomic factors can help marketing researchers develop a better understanding of under-studied and underserved consumers (MacInnis et al. forthcoming). I collect data on each patient's neighborhood socioeconomic status—including educational attainment, income, commute time, private/public health insurance coverage, employment status, and population—by collecting zip-code level data from American Community Survey.

Step 4: Incorporate contextually-relevant variables. Along with variables that are static in nature, I include each patient's screening compliance in the prior periods. Incorporating screening compliance across multiple rounds of screening (Chubak and Hubbard 2016) captures the temporal variation in screening completion. It also informs how outreach effectiveness might vary due to patients' prior behavioral pattern.

In summary, I include six sets of patient characteristics: (1) demographics that include age, gender (coded as 1 if a patient is female, 0 otherwise), ethnicity (non-Hispanic Caucasian, Hispanic, Non-Hispanic African American, or other/unknown), and

---

[28] https://conversation.digitalhealth.gov.au/sites/default/files/adha-strategy-doc-2ndaug_0_1.pdf

[29] http://www.bccdc.ca/pop-public-health/Documents/TowardsReducingHealthInequitiesFinalDiscussionPape.pdf

primary language (English, Spanish, or other); (2) health status that includes Child-Pugh B (coded as 1 if Child-Pugh score is higher than 6, 0 otherwise), Charlson Comorbidity Index, presence of documented cirrhosis (coded as 1 if yes), etiology of liver disease (hepatitis C, hepatitis B, alcohol, nonalcoholic steatohepatitis, or other); (3) visit history that includes the number of primary care visits in the year prior to cohort entry and receipt of hepatology care (coded as 1 if the patient received the hepatology care prior to cohort entry, 0 otherwise); (4) health system accessibility that includes insurance coverage (commercial, Medicaid, medical assistance/charity, Medicare, self-pay, or unknown), and proximity to clinics (coded as 1 if there are more than three clinics in the zip code that matches the first three digits of the zip code[30] where the patient resides, 0 otherwise); (5) neighborhood socioeconomic status that includes three-digit zip-code[31] level educational attainment (% with a bachelor's degree or higher), income (per capita income), average commute time, insurance coverage (% with a private or public health insurance plan), unemployment rate, and population; and (6) screening completion status in the prior period(s) (coded as 1 if a patient completes the screening test in Period 1 (2), 0 otherwise). Table 3.1 describes the details. Appendix B.2 compares the means of all variables across three conditions. Differences are not statistically significant, so the random assignment was successful.

---

[30] According to Health Insurance Portability and Accountability Act, I am not allowed to obtain patients' identifiable location information such as address and zip code. Thus, I obtain the de-identified version (i.e., first 3 digits of the zip code).

[31] As I only observe the first 3 digits of patients' zip code, all zip-code level covariates are aggregated to the 3-digit level by calculating the sum (i.e., population) or mean (% with Bachelor's degree or higher, mean travel time to work, and per capita income) across all 5-digit zip codes that share the same first three digits.

Table 3.1 Summary Statistics

Variable	Definition	Mean	SD	Min	Max
Dependent variable					
Completion in Period 1	Whether a patient underwent an abdominal imaging screening test, which includes ultrasound, MRI, and CT, in Period 1 (0-6 months after cohort entry). Coded as 1 if the patient completed, and 0 otherwise.	39.3%	-	0	1
Completion in Period 2	Whether a patient underwent an abdominal imaging screening test, which includes ultrasound, MRI, and CT, in Period 2 (6 months and 1 day-12 months after cohort entry). Coded as 1 if the patient completed, and 0 otherwise.	38.4%	-	0	1
Completion in Period 3	Whether a patient underwent an abdominal imaging screening test, which includes ultrasound, MRI, and CT, in Period 3 (12 months and 1 day- 18 months after cohort entry). Coded as 1 if the patient completed, and 0 otherwise.	34.9%	-	0	1
Independent variable					
Outreach intervention	A baseline condition and two different outreach intervention types:				
	(1) no outreach (usual care)	33.3%	-	0	1
	(2) moderate outreach (outreach alone)	33.3%	-	0	1
	(3) intensive outreach (outreach with patient navigation)	33.3%	-	0	1
Demographics					
Age	Age of the patient at cohort entry	55.3	10.5	21	90
Gender	Gender of the patient (0 = male, 1 = female)	40.6%	-	0	1
Ethnicity					
Non-Hispanic Caucasian	Non-Hispanic Caucasian = 1; otherwise = 0	28.3%	-	0	1
Hispanic	Hispanic = 1; otherwise = 0	37.8%	-	0	1
Non-Hispanic African American	Non-Hispanic African American = 1; otherwise = 0	32.1%	-	0	1
Other/unknown	Other/unknown = 1; otherwise = 0	1.7%	-	0	1
Language					
English	English = 1; otherwise = 0	76.9%	-	0	1
Spanish	Spanish = 1; otherwise = 0	22.7%	-	0	1
Other	Other = 1; otherwise = 0	.3%	-	0	1
Health Status					
Child Pugh B	Whether a patient is Child Pugh B, coded as 1 if Child Pugh Score >6; 0 otherwise.	28.3%	-	0	1
Charlson Comorbidity Index	Charlson Comorbidity Index Score	2.9	2.4	0	12
Documented cirrhosis	Cirrhosis diagnosis (0= Suspected Cirrhosis; 1=Known Cirrhosis)	79.6%	-	0	1
Etiology of liver disease					
Hepatitis C	Hepatitis C = 1; otherwise = 0	51.0%	-	0	1
Hepatitis B	Hepatitis B = 1; otherwise = 0	3.4%	-	0	1
Alcohol	Alcohol = 1; otherwise = 0	17.6%	-	0	1
Nonalcoholic steatohepatitis	Nonalcoholic steatohepatitis = 1; otherwise = 0	16.6%	-	0	1
Other	Other = 1; otherwise = 0	11.3%	-	0	1
Visit History					
Number of prior primary care visits	Number of primary care visits in the year prior to cohort entry	5.2	4.8	0	38
Receipt of hepatology care	History of hepatology care in the year prior to cohort entry, coded as 1 if the patient received the care, 0 otherwise.	25.7%	-	0	1
Health System Accessibility					
Insurance coverage					
Commercial	Yes = 1; otherwise = 0	3.0%	-	0	1
Medicaid	Yes = 1; otherwise = 0	20.6%	-	0	1
Medical assistance/charity	Yes = 1; otherwise = 0	41.0%	-	0	1
Medicare	Yes = 1; otherwise = 0	24.7%	-	0	1
Self-pay	Yes = 1; otherwise = 0	2.0%	-	0	1
Unknown	Yes = 1; otherwise = 0	8.7%	-	0	1
Proximity to clinics	Whether a patient has a close geographical proximity to clinics. Coded as 1 if there are more than 3 clinics in the zip code that matches the first three digits of the zip code where the patient resides, 0 otherwise.	66.7%	-	0	1

Table 3.1 Continued

Variable	Definition	Mean	SD	Min	Max
Neighborhood Socioeconomic Status					
Educational attainment (%)	Percent of people who is 18 years and over and received a bachelor's, master's, professional, or doctorate degree.	33.6	6.6	13.8	52.9
Income ($)	Per capita income: mean income computed for every man, woman, and child in the same zip code.	35,223.8	4,117.7	15,839.6	42,925.3
Average commute time (minutes)	Mean travel time to work from home during the reference week	27.1	2.1	19.1	33.6
Private health insurance coverage (%)	Percent of civilian noninstitutionalized population with the insurance coverage provided through an employer or union, a plan purchased by an individual from a private company, or military health care.	58.6	7.2	35.3	74.8
Public health insurance coverage (%)	Percent of civilian noninstitutionalized population with the insurance coverage provided through the federal programs Medicare, Medicaid, and VA Health Care, as well as the Children's Health Insurance Program and individual state health plans.	28.6	4.2	19.9	39.7
Unemployment rate (%)	Percent of civilians 16 years old and over classified as unemployed	4.0	0.4	2.5	5.7
Population	Number of people 16 years and over.	1,148,050	371,086	10,824	1,791,015

Notes. After excluding patients who were diagnosed with HCC or deceased, the screening completion rate is 38.5% in Period 2 and 35.4% in Period 3.

3.2.4. Model-Free Evidence

Figure 3.2 shows the variation in the number of periods where patients have completed screening. Whereas 435 patients (24%) completed the screening only once during the three periods, 660 patients (37%) did so more than once. Furthermore, outreach with patient navigation tends to increase the number of patients who completed screening in all three periods. Figure 3.3 shows screening completion rates in each condition in each period after excluding the patients who were deceased or diagnosed with HCC in the previous period(s). In Period 1, 25% in the no-outreach condition, 45% in the outreach-alone condition, and 48% in the outreach-with-patient-navigation condition underwent screening. The screening completion rate in the outreach-alone condition (difference = .198, $p < .01$) and that in the outreach-with-patient-navigation condition (difference = .232, $p < .01$) are significantly higher than the no-outreach

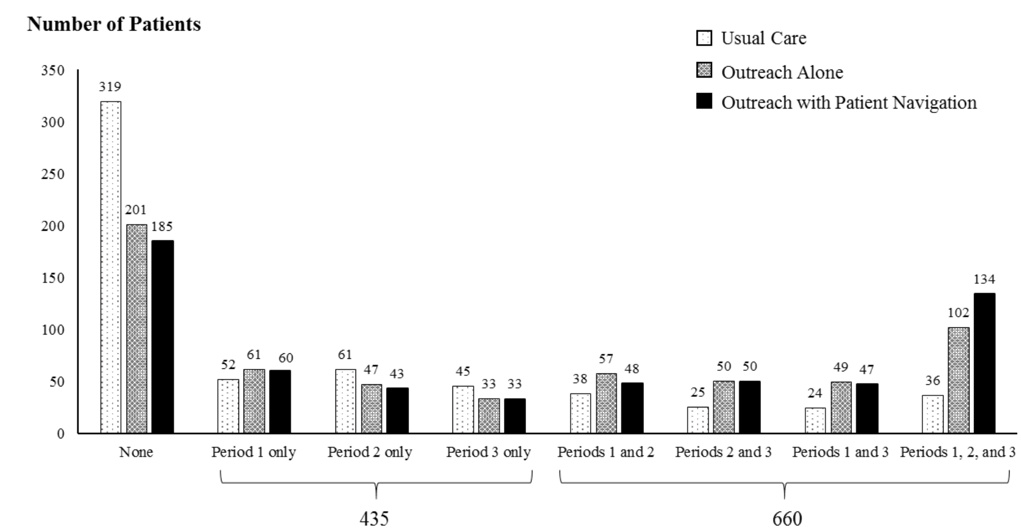


Figure 3.2 Frequency of Screening Completion in Each Study Condition across Periods

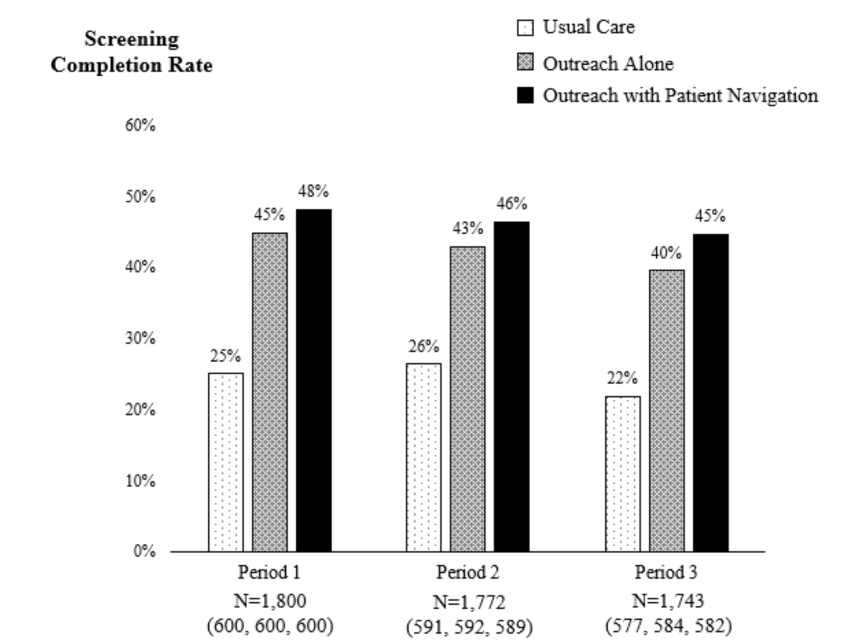


Figure 3.3 Screening Completion Rates in Each Study Condition

condition. Results in Periods 2 and 3 follow a similar pattern. Importantly, the screening completion rate in the outreach-with-patient-navigation condition is not statistically higher than that in the outreach-alone condition in Period 1 (difference = .033, *n.s.*) or Period 2 (difference = .033, *n.s.*), but is statistically higher in Period 3 (difference = .051, $p < .10$). The model-free evidence suggests that both outreach conditions outperform the baseline condition, but outreach with patient navigation is slightly more effective than outreach alone only in Period 3.

3.3. Empirical Strategy

3.3.1. Causal Forest Estimation of Patient-Level Treatment Effects

To draw inferences about the causal effect of different interventions, researchers typically estimate and compare the average treatment effects (i.e., main effects) of randomized interventions. Such a comparison may not consider that treatment effects across subgroups within and across treatment conditions. Moreover, to avoid searching for particularly responsive subgroups, medical researchers must register pre-analysis protocols for clinical trials to specify which subgroups will be analyzed. Such protocols may fail to identify strong but unexpected treatment-effect heterogeneity, especially in emergent fields where moderators are *ex-ante* ambiguous. I use causal forests to address these two challenges (Wager and Athey 2018). Causal forests enable nonparametric estimation of patient-level treatment effects with valid asymptotic confidence intervals, without (a) restrictions on the number of covariates, and (b) the need for a larger number of experimental conditions or repeated measures. Causal forests also alleviate concerns regarding spurious treatment-effect heterogeneity due to searching for particularly responsive subgroups (Appendix B.3 compares causal forests with several established approaches). Next, I outline the potential outcome framework, followed by an overview of causal forests.

3.3.1.1. Potential outcome framework

For illustration purposes, I consider the case of one period and the outreach-alone intervention (treatment condition) compared against no outreach/usual care (control condition). For a set of independent and identically distributed patients i = 1, …, n, I observe the outcome of interest Y_i (screening completion), treatment assignment W_i (i.e., whether the patient is assigned to the outreach-alone or no-outreach condition), and vector of patient characteristics X_i (e,g., patient demographics, health status). Following the potential outcome framework (Rubin 1974), for each patient i, there are two potential outcomes: if a patient is assigned to the treatment condition, I observe the outcome $Y_i = Y_{i1}$, and if the patient is assigned to the control condition, I observe $Y_i = Y_{i0}$. I define the conditional average treatment effect (CATE) (i.e., treatment effect at x) to assess whether the treatment effect is heterogeneous among subgroups:

(1) $\tau(x)=E[Y_{i1}-Y_{i0}|X_i=x]$.

The fundamental challenge to identify the CATE is that I only observe one of the two potential outcomes: Y_{i1} and Y_{i0}. Thus, I must invoke the assumption of unconfoundedness to estimate the CATE (Rosenbaum and Rubin 1983). As patients are randomly assigned to one of the experimental conditions, the treatment assignment W_i is independent of the potential outcomes conditional on X_i (i.e., $\{Y_{i1},Y_{i0}\}\perp W_i \mid X_i$). This assumption implies that the treatment is as good as random within each subpopulation indexed by X_i=x. Thus, given the data (X_i, Y_i, W_i), I can revise Equation (1) to be:

$$(2) \qquad \tau(x) = E[Y_{i1} - Y_{i0} | X_i = x] = E[Y_i | W_i = 1, X_i = x] - E[Y_i | W_i = 0, X_i = x].$$

Common approaches to estimate the function $\tau(x)$ include nearest neighbor matching and kernel methods, but these methods do not perform well in environments with many covariates or complex interactions among covariates (Wager and Athey 2018).

3.3.1.2. Causal forests

Causal forests combine causal inference in economics with random forests in machine learning. Random forests (Breiman 2001) deploy supervised machine learning algorithms to achieve high out-of-sample prediction accuracy with very little tuning, particularly with high dimensional data with underlying nonlinear relationships (Hastie, Tibshirani, and Friedman 2009). Random forests (1) build a large collection of individual decision trees such that each tree predicts the outcome variable given the vector of covariates and (2) average the predictions from those trees. First, each tree is trained on a bootstrap training sample (not on the original sample) with a randomly chosen subset of covariates (not with all the covariates), and it is built by recursively partitioning the chosen covariate space into splits, determining each split by minimizing the mean squared error of the prediction of outcomes in the case of regression trees. Given the tree split, each tree clusters the most similar observations into a terminal node known as a leaf. To predict the outcome of an observation outside of the estimation sample, each tree makes a prediction using the mean of outcomes in the leaf where this new observation belongs. Finally, a random forest averages the prediction from the trees.

Researchers have recently adapted random forests to draw inferences. The technique known as causal forests utilizes an algorithm for flexible modeling of interactions in high dimensions by building many causal trees and averaging their predictions to estimate the treatment effect function τ(x). Causal forests provide valid asymptotic confidence intervals for the treatment effects (Wager and Athey 2018). Given a profile of patient characteristics x, tree-based models help identify the most similar patients locally in the patient characteristics space with an adaptive neighborhood metric (i.e., similar patients are in the same leaf). Wager and Athey (2018) adapt the regression tree to estimate the within-leaf treatment effects by taking the difference between the mean outcomes of treated and control units in the same leaf:

$$\hat{\tau}(x)=\frac{1}{|\{i:W_i=1,\ X_i\in L\}|}\sum_{\{i:W_i=1,\ X_i\in L\}} Y_i-\frac{1}{|\{i:W_i=0,\ X_i\in L\}|}\sum_{\{i:W_i=0,\ X_i\in L\}} Y_i. \tag{3}$$

To ensure consistency and asymptotic normality, Wager and Athey (2018) prove a bias-reducing condition called honesty: a tree achieves honesty if each bootstrap training sample only uses the outcome of interest Y_i to estimate the within-leaf treatment effect based on Equation (3) or to determine where to split the covariate space, but not both. In other words, the bootstrap training sample is further split into two subsamples: one used to build the tree (i.e., understand where the treatment heterogeneity is given the covariates),[32] and the other used to estimate the treatment effects given the tree structure.

Using this process, causal forests produce an ensemble of B such trees (Breiman

---

[32] Unlike regression trees where the splits are determined by minimizing mean-squared error (MSE) of the prediction of *outcomes*, causal trees are built by minimizing the expected MSE of predicted *treatment effects*, which is equivalent to maximizing the variance of treatment effects across leaves minus a penalty for within-leaf variance (Athey and Imbens 2016).

2001; Wager and Athey 2018), each of which outputs an estimate $\hat{\tau}_b(x)$ and averages the predictions from those trees to compute an estimated CATE: $\hat{\tau}(x)=B^{-1}\sum_{b=1}^{B}\hat{\tau}_b(x)$.

This aggregation scheme also helps reduce variance and smooths sharp decision boundaries (Bühlmann and Yu 2002). The variance estimate of causal forests is defined as follows (Efron 2014; Wager, Hastie, and Efron 2014; Wager and Athey 2018):

$$\widehat{V}(x)=\frac{n-1}{n}\left(\frac{n}{n-s}\right)^2 B^{-1}\sum_{i=1}^{n}\mathrm{COV}[\hat{\tau}_b(x),N_{ib}]^2, \tag{4}$$

where $\hat{\tau}_b(x)$ is the treatment effect estimate from the b^{th} tree. $N_{ib} \in \{0, 1\}$ indicates whether the bootstrap training sample i is used for the tree b, $n(n-1)/(n-s)^2$ is a finite-sample correction for forests grown by subsampling without replacement, and the covariance is taken with respect to all B trees in the forest. Equations (3) and (4) produce a treatment effect estimate and a confidence interval for each patient.

In marketing, Ascarza (2018) uses causal forests to identify customers who are particularly responsive to retention efforts, compares them against those at high risk of churning, and designs retention programs with the right targeting rules. Guo, Sriram, and Manchanda (2019) use causal forests to estimate heterogenous treatment effects of increased transparency of information disclosure on subsequent payments between firms and physicians at the level of physician-product pair. Sun et al. (2019) use instrumental forests to uncover the heterogenous treatment effects of the adoption of voice-activated shopping devices on consumers' purchase quantity, spending, and search activities.

3.3.1.3. Application to my context

Like Singal et al. (2019), I have two different treatment conditions (outreach alone and outreach with patient navigation) and three different periods. I use the following procedure to perform six causal forest estimations. Additional aspects of the estimation are summarized in Appendix B.3.

Step 1. Using patient characteristics as covariates[33], I applied causal forests to obtain each patient's treatment effect estimate in the sample that includes patients in the baseline (condition 1, sample size=600) and those in the outreach-alone condition (condition 2, sample size=600) in Period 1. For each patient *i* in condition 1 in Period 1, the patient-level treatment effect estimate is $\hat{\tau}^{1}_{i,P1,1\rightarrow 2}$ (i.e., the difference between the outcome I observe for the patient i in condition 1 and the outcome that would be realized if this patient were in condition 2); for each patient in condition 2 in Period 1, the patient-level treatment effect estimate is $\hat{\tau}^{2}_{i,P1,2\rightarrow 1}$ (i.e., the difference between the outcome I observe for the patient i in condition 2 and the outcome that would be realized if this patient were in condition 1). I term this first causal forest estimation Forest^{12}_{P1}, where P1 refers to Period 1, and the superscript 12 refers to the comparison of the baseline condition (1) and the outreach-alone condition (2).

Step 2. After excluding the patients who were deceased or diagnosed with HCC in the previous period(s), I repeated *Step 1* to obtain $\hat{\tau}^{1}_{i,P2,1\rightarrow 2}$ and $\hat{\tau}^{2}_{i,P2,2\rightarrow 1}$ (condition 1, sample size= 591; condition 2, sample size= 592) in Period 2 and $\hat{\tau}^{1}_{i,P3,1\rightarrow 2}$ and $\hat{\tau}^{2}_{i,P3,2\rightarrow 1}$ (condition 1, sample size= 577; condition 2, sample size= 584) in Period 3. As

---

[33] I scale continuous variables to zero mean and unit variance and expand categorical variables via one-hot encoding.

discussed, I included one (two) additional covariate(s), indicating whether a patient has completed the screening test in the prior period(s) in the causal forest estimation of Period 2 (3). I term these second and third causal forests $\mathrm{Forest}_{P2}^{12}$ and $\mathrm{Forest}_{P3}^{12}$, where P2 and P3 refer to Period 2 and Period 3, respectively, and the superscript 12 refers to the comparison of the baseline condition (1) and the outreach-alone condition (2).

Step 3. I repeated *Step 1* to obtain each patient's treatment effect estimate in the sample that includes patients in the baseline (condition 1, sample size=600) and those in the outreach-with-patient-navigation condition (condition 3, sample size=600) in Period 1. For each patient in condition 1, the patient-level treatment effect estimate is $\hat{\tau}_{i,P1,1\to 3}^{1}$ (i.e., the difference between the outcome I observe for the patient i in condition 1 and the outcome that would be realized if this patient were in condition 3); for each patient in condition 3, the patient-level treatment effect estimate is $\hat{\tau}_{i,P1,3\to 1}^{3}$ (i.e., the difference between the outcome I observe for the patient i in condition 3 and the outcome that would be realized if this patient were in condition 1). I term this fourth causal forest $\mathrm{Forest}_{P1}^{13}$, where P1 refers to Period 1, and the superscript 13 refers to the comparison of the baseline condition (1) and outreach-with-patient-navigation condition (3).

Step 4. I repeated *Step 2* to obtain $\hat{\tau}_{i,P2,1\to 3}^{1}$ and $\hat{\tau}_{i,P2,3\to 1}^{3}$ in Period 2 (condition 1, sample size= 591; condition 3, sample size= 589), and $\hat{\tau}_{i,P3,1\to 3}^{1}$ and $\hat{\tau}_{i,P3,3\to 1}^{3}$ (condition 1, sample size= 577; condition 3, sample size= 582) in Period 3. I term these fifth and sixth causal forests $\mathrm{Forest}_{P2}^{13}$ and $\mathrm{Forest}_{P3}^{13}$ where P2 and P3 refer to Period 2 and Period 3, respectively, and the superscript 13 refers to the comparison of the baseline condition (1) and the outreach-with-patient-navigation condition (3).

Figure 3.4 shows the distribution of the patient-level treatment effect estimates based on the causal forest estimation. I also summarize these estimates in Table 3.2.

Relative to the baseline condition, outreach alone (outreach with patient navigation) increases screening completion rate by between 10 and 20 (13 and 24) percentage points (Panel A of Table 3.2). Causal forests enable me to construct confidence intervals for patient-level treatment effect estimates. As reported in Panel B of Table 3.2, outreach-alone intervention induces positive and statistically significant treatment effects among 100%, 74%, and 66% of the patients in Period 1, 2, and 3, respectively ($p < .05$), while the outreach-with-patient-navigation intervention does so among 100%, 83%, and 89% of the patients in Period 1, 2, and 3, respectively ($p < .05$).

Germane to the focus of this paper, there is substantial heterogeneity in those significant patient-level treatment effects: (1) compared to outreach alone, outreach-with-patient-navigation intervention induces a higher proportion of patients with significant positive treatment effect estimates in Periods 2 (83%-74%=9%) and 3 (89%-66%=23%); (2) Patient-level treatment effect estimates of outreach-alone (outreach-with-patient-navigation) intervention range from 5-31 (5-37) percentage points. Next, I investigate the sources of heterogeneity.

3.3.2. Incorporating Heterogeneity in Patient-Level Treatment Effects

I examine the treatment effect heterogeneity by correlating treatment effect estimates with patient characteristics. Accordingly, I estimated the following equations:

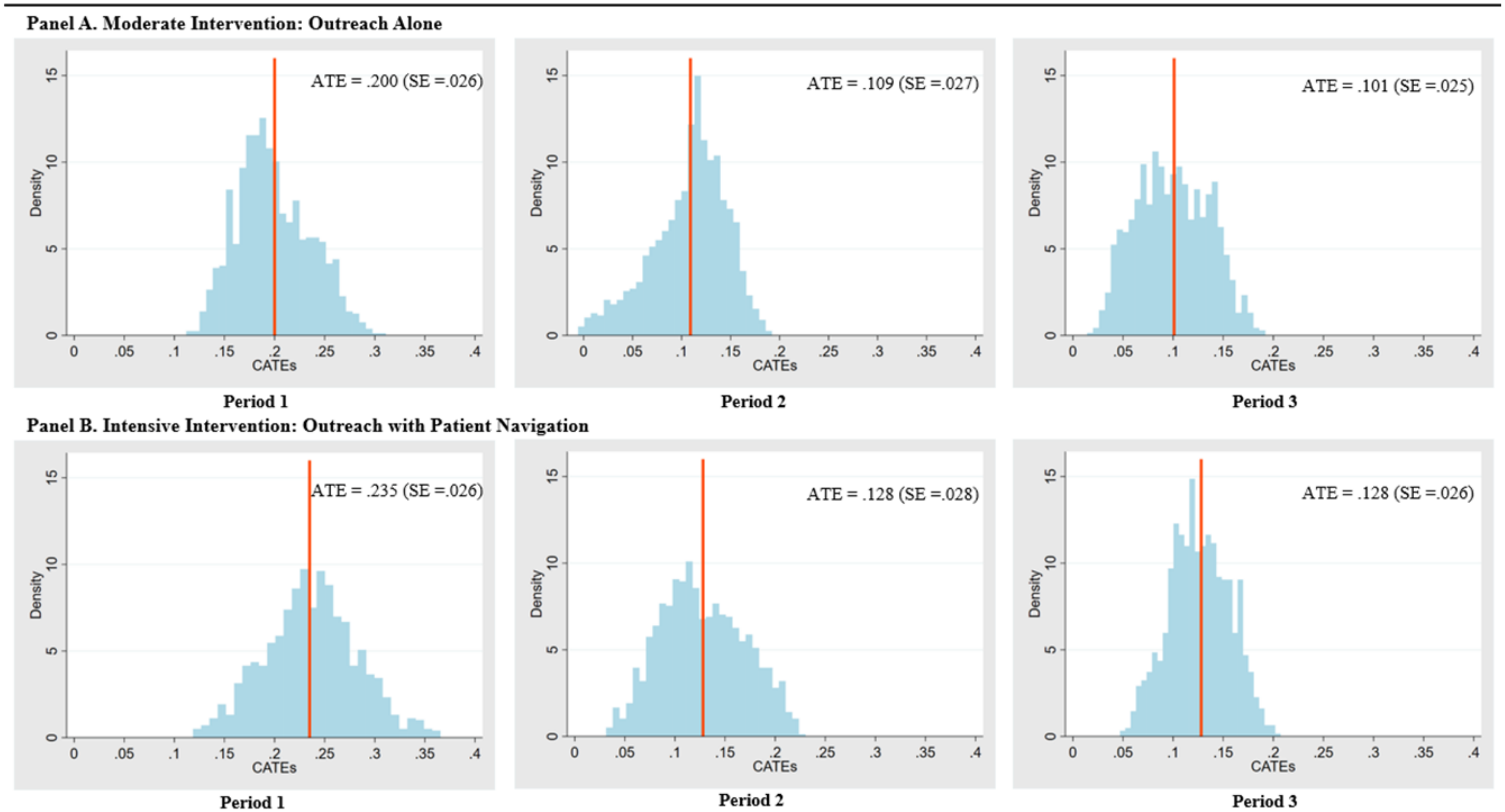


Figure 3.4 Distribution of the Patient-Level Treatment Effect Estimates

Notes. ATE refers to the average treatment effect; SE refers to the standard error; CATEs refer to conditional average treatment effects.

Table 3.2 Summary of the Average Treatment Effects and Patient-Level Conditional Average Treatment Effects by Outreach Type

Panel A	Outreach Alone					Outreach with Patient Navigation				
	N	ATE	SE	95% CI		N	ATE	SE	95% CI	
Period 1 ATE	1,200	.200	.026	.148	.251	1,200	.235	.026	.184	.287
Period 2 ATE	1,183	.109	.027	.057	.162	1,180	.128	.028	.074	.182
Period 3 ATE	1,161	.101	.025	.052	.150	1,159	.128	.026	.077	.178

Notes. ATE refers to the average treatment effect; SE refers to the standard error; CI refers to the confidence interval.

Panel B	Outreach Alone					Outreach with Patient Navigation				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Period 1										
Patient-level CATEs	1,200	.199	.036	.112	.311	1,200	.236	.046	.119	.366
Significant Patient-level CATEs*	1,200	.199	.036	.112	.311	1,200	.236	.046	.119	.366
Proportion of Significant Patient-level CATEs	100%					100%				
Period 2										
Patient-level CATEs	1,183	.108	.038	-.006	.192	1,180	.127	.042	.031	.230
Significant Patient-level CATEs*	875	.125	.024	.057	.192	975	.139	.036	.052	.230
Proportion of Significant Patient-level CATEs	74%					83%				
Period 3										
Patient-level CATEs	1,161	.099	.036	.014	.192	1,159	.126	.030	.047	.207
Significant Patient-level CATEs*	767	.118	.028	.047	.192	1,030	.131	.027	.062	.207
Proportion of Significant Patient-level CATEs	66%					89%				

Notes. CATEs refer to the conditional average treatment effects. *Statistical significance is at the 95% level.

(5a) $\hat{\tau}^2_{ijt=\{P1,P2,P3\}} = \alpha_{20} + \alpha_{21}Age_{ij} + \alpha_{22}Gender_{ij} + \alpha_{23}Ethnicity_{ij} + \alpha_{24}Language_{ij}$

$+\alpha_{25}Child\text{-}PughB_{ij} + \alpha_{26}Charlson_{ij} + \alpha_{27}Cirrhosis_{ij} + \alpha_{28}Etiology_{ij}$

$+\alpha_{29}Prior\_visit_{ij} + \alpha_{210}Hepatology\_care_{ij} + \alpha_{211}Insurance\ Coverage_{ij} + \alpha_{212}Proximity_{ij}$

$+\alpha_{213}Education_j + \alpha_{214}Income_j + \alpha_{215}Commute_j + \alpha_{216}Private_j + \alpha_{217}Public_j + \alpha_{218}Unemploy_j$

$+\alpha_{219}Population_j + \eta_t + \epsilon_{2ijt}$

(5b) $\hat{\tau}^3_{ijt=\{P1,P2,P3\}} = \alpha_{30} + \alpha_{31}Age_{ij} + \alpha_{32}Gender_{ij} + \alpha_{33}Ethnicity_{ij} + \alpha_{34}Language_{ij}$

$+\alpha_{35}Child\text{-}PughB_{ij} + \alpha_{36}Charlson_{ij} + \alpha_{37}Cirrhosis_{ij} + \alpha_{38}Etiology_{ij}$

$+\alpha_{39}Prior\_visit_{ij} + \alpha_{310}Hepatology\_care_{ij} + \alpha_{311}Insurance\ Coverage_{ij} + \alpha_{312}Proximity_{ij}$

$+\alpha_{313}Education_j + \alpha_{314}Income_j + \alpha_{315}Commute_j + \alpha_{316}Private_j + \alpha_{317}Public_j + \alpha_{318}Unemploy_j$

$+\alpha_{319}Population_j + \eta_t + \epsilon_{3ijt}$,

where $\hat{\tau}^2_{ijt=\{P1,P2,P3\}}$ and $\hat{\tau}^3_{ijt=\{P1,P2,P3\}}$ refer to patient-level treatment effect estimates of outreach alone and those of outreach with patient navigation, j denotes the three-digit zip code, and t denotes the period. I pooled the estimates across periods and included period-fixed effects (η_t) to capture common time-varying observables that may affect them and clustered standard errors at the patient level to allow for heteroskedasticity and correlated errors within patients over time.

Columns 1 and 2 of Table 3.3 report the results showing the sources of heterogeneity in patient-level treatment effects (see Appendix B.4 for a summary of these findings). Next, I discuss the patient characteristics associated with the treatment effect heterogeneity.

Table 3.3 Sources of Heterogeneity in Patient-Level Treatment Effects

	(1) Outreach Alone			(2) Outreach with Patient Navigation			(3) Outreach Alone			(4) Outreach with Patient Navigation		
	Est		SE	Est		SE	Est		SE	Est		SE
Demographics												
Age	-.003	***	(.000)	.006	***	(.000)	-.003	***	(.000)	.005	***	(.000)
Gender (Female=1)	.014	***	(.001)	.007	***	(.001)	.013	***	(.001)	.007	***	(.001)
Hispanic	.011	***	(.001)	.004	***	(.001)	.012	***	(.001)	.005	***	(.001)
Non-Hispanic African American	.010	***	(.001)	.011	***	(.001)	.010	***	(.001)	.011	***	(.001)
Other/unknown	.009	***	(.002)	.007	*	(.003)	.008	***	(.002)	.007	*	(.003)
Spanish	.014	***	(.001)	.015	***	(.001)	.013	***	(.001)	.014	***	(.001)
Other	-.002		(.002)	.003		(.005)	-.001		(.004)	.004		(.005)
Health Status												
Child-Pugh B	-.012	***	(.001)	-.003	***	(.001)	-.011	***	(.001)	-.003	***	(.001)
Charlson Comorbidity Index	-.003	***	(.000)	-.002	***	(.000)	-.004	***	(.000)	-.003	***	(.000)
Documented cirrhosis	-.002	*	(.001)	-.002	*	(.001)	-.002	**	(.001)	-.003	**	(.001)
Hepatitis B	-.002	*	(.001)	.001		(.001)	-.001		(.001)	.002		(.001)
Alcohol-induced	-.004	*	(.002)	-.001		(.003)	-.005	**	(.001)	.000		(.002)
Nonalcoholic steatohepatitis	-.003	***	(.001)	-.000		(.001)	-.003	***	(.001)	.000		(.001)
Other	-.000		(.001)	.005	***	(.001)	-.001		(.001)	.004	***	(.001)
Visit History												
Number of prior primary care visits	.010	***	(.001)	.011	***	(.001)	.020	***	(.001)	.022	***	(.002)
Receipt of hepatology care	.003	***	(.001)	.003	*	(.001)	.002	*	(.001)	.001		(.001)
Health System Accessibility												
Commercial	-.009	***	(.002)	-.005	*	(.002)	-.009	***	(.001)	-.002		(.002)
Medicaid	-.010	***	(.001)	-.001		(.001)	-.009	***	(.001)	.001		(.001)
Medicare	-.016	***	(.001)	-.002	*	(.001)	-.015	***	(.001)	-.002	*	(.001)
Self-pay	-.012	***	(.002)	-.003		(.003)	-.012	***	(.002)	-.002		(.002)
Unknown	-.013	***	(.001)	-.008	***	(.001)	-.011	***	(.001)	-.005	**	(.001)
Proximity to clinics	.008	***	(.001)	.005	**	(.002)	.008	***	(.001)	.007	***	(.001)
Neighborhood Socioeconomic Status												
Educational attainment (%)	.002		(.004)	-.011		(.006)	.003		(.004)	-.011		(.006)
Income ($)	-.000		(.003)	.012	**	(.004)	.000		(.003)	.011	**	(.004)
Average commute time (minutes)	-.004	***	(.001)	-.000		(.002)	-.003	**	(.001)	-.002		(.001)
Private health insurance (%)	-.000		(.004)	.002		(.005)	-.002		(.003)	.008		(.005)
Public coverage (%)	.011	**	(.004)	.006		(.005)	.009	**	(.003)	.009	*	(.004)
Unemployment rate (%)	.002		(.001)	.003		(.002)	.002		(.001)	.005		(.002)
Population	.005	*	(.002)	.005	*	(.002)	.003		(.002)	.004	*	(.002)
Period Fixed Effects												
Period 2 dummy	-.091	***	(.001)	-.109	***	(.001)	-.091	***	(.001)	-.109	***	(.001)
Period 3 dummy	-.100	***	(.001)	-.110	***	(.002)	-.100	***	(.001)	-.110	***	(.002)
Exploratory Interactions												
Primary care visit^2							-.003	***	(.000)	-.004	***	(.000)
Primary care visit × Age							.000		(.000)	.000		(.001)
Primary care visit × Gender							.000		(.001)	-.003	**	(.001)
Primary care visit × Hispanic							-.003	**	(.001)	-.002		(.002)
Primary care visit × African American							.001		(.001)	-.001		(.001)
Primary care visit × Spanish							-.005	***	(.001)	-.005	***	(.002)
Primary care visit × Child-Pugh B							-.003	***	(.001)	-.003	**	(.001)
Primary care visit × Charlson Index							-.000		(.000)	-.001		(.000)
Primary care visit × Medicaid							-.000		(.001)	.003		(.002)
Primary care visit × Medical Assistance							-.002	*	(.001)	-.001		(.002)
Primary care visit × Medicare							-.001		(.001)	-.001		(.002)
Primary care visit × Proximity to clinics							-.000		(.001)	.002		(.001)
Intercept	.190	***	(.001)	.225	***	(.002)	.193	***	(.001)	.227	***	(.002)
Clustered standard error	Yes			Yes			Yes			Yes		
R^2	.770			.727			.783			.741		
N	3,544			3,539			3,544			3,539		

* $p < .05$, ** $p < .01$, *** $p < .001$.

Notes. Est (SE) refers to the estimated coefficient (standard error). The baseline categories of main effects are male, non-Hispanic Caucasian, Hepatitis C, English, medical assistance/charity, and Period 1. I scaled continuous variables to zero mean and unit variance. As I pooled the estimates of three periods, sample sizes are 3,544 (1,200+1,183+1,161) and 3,539 (1,200+1,180+1,159).

3.3.3. Results and Discussion

Age. Older patients are less responsive to the *outreach-alone intervention* than younger patients ($\hat{\alpha}_{21}$ = -.003, p < .001) while they are more responsive to *outreach-with-patient-navigation intervention* than younger patients ($\hat{\alpha}_{31}$ = .006, p < .001). A possible explanation is that older adults prefer to use information that is customized to their needs rather than generic information that can be overwhelming (Cole et al. 2008), which makes them less responsive to direct mails than younger adults (Kaldenberg, Koenig, and Becker 1994). The interactive and personalized nature of the navigation over the telephone provides targeted and useful information to older patients, making it more effective (King, Rejeski, and Buchner 1998).

Gender. Female patients are more responsive to *both outreach interventions* than male patients ($\hat{\alpha}_{22}$ = .014, p < .001 for outreach-alone; $\hat{\alpha}_{32}$ = .007, p < .001 for outreach with patient navigation). This is likely due to the higher prevention and loss-minimization focus among females (Trudeau et al. 2003). According to agency-communion theory (Carlson 1971), males focus on maximizing gains while females focus on minimizing the downside potential of their decision. Outreach messages for cancer screening, by design, approach healthcare from a prevention and loss-minimization focus.

Ethnicity and language. Hispanic patients are more responsive to *both outreach interventions* than Caucasian patients ($\hat{\alpha}_{23}^{H}$ = .011, p < .001 for outreach alone; $\hat{\alpha}_{33}^{H}$ = .004, p < .001 for outreach with patient navigation). Likewise, non-Hispanic African-American patients are more responsive to *both outreach interventions* than Caucasian patients ($\widehat{\alpha}_{23}{}^{AA}$ = .010, $p < .001$ for outreach alone; $\widehat{\alpha}_{33}{}^{AA}$ = .011, $p < .001$ for outreach with patient navigation). Due to language and access barriers, ethnic minority patients may have relatively fewer opportunities to learn about the health screening information than ethnic majority groups (Szczepura 2005; Caplan, Wells, and Haynes 1992). Given the lower baseline access, outreach interventions that provide information on screening opportunities should be more effective among minority groups than their counterparts (Lasser et al. 2011). Similarly, patients whose primary language is Spanish are more responsive to *both outreach interventions* than those whose primary language is English ($\widehat{\alpha}_{24}$ = .014, $p < .001$ for outreach alone; $\widehat{\alpha}_{34}$ = .015, $p < .001$ for outreach with patient navigation).

Health status. Patients in a poorer health status (those with Child-Pugh B) are less responsive to *both outreach interventions* than patients with a better health status ($\widehat{\alpha}_{25}$ = -.012, $p < .001$ for outreach alone; $\widehat{\alpha}_{35}$ = -.003, $p < .001$ for outreach with patient navigation). The pattern is consistent when Charlson Comorbidity Index and the presence of documented cirrhosis are used as indicators of health status. A possible explanation is that outreach interventions might make patients fearful of finding out they have cancer (Aro et al. 2001) and experience death anxiety (Grossman et al. 2018). Those with poor health will experience higher death anxiety due to lower optimism about their health (Arndt, Routledge, and Goldenberg 2006), which reduces adaptive coping and thus decreases the utilization of healthcare services (Moorman and Matulich 1993). I also find that compared to patients with Hepatitis C, those with Hepatitis B are less responsive to outreach alone (coefficients ranging from -.004 to -.002), but it is not the case for outreach with patient navigation.

Visit history. Patients with a higher number of prior primary care visits are more responsive to *both outreach interventions* than those with fewer prior primary care visits ($\widehat{\alpha}_{29}$ = .010, $p < .001$ for outreach alone; $\widehat{\alpha}_{39}$ = .011 $p < .001$ for outreach with patient navigation). Similarly, patients who previously received hepatology care are more responsive to *both outreach interventions* than patients with no prior hepatology care ($\widehat{\alpha}_{210}$ = .003, $p < .001$ for outreach alone; $\widehat{\alpha}_{310}$ = .003, $p < .05$ for outreach with patient navigation). At its core, a patient's prior visit history signifies the extent to which a patient has a favorable attitude toward utilizing healthcare services to pursue their health goals (Klein and Cerully 2007) and familiarity with the utilization process (Goldman et al. 2015). This should motivate patients to get screened.

Health system accessibility. Patients with insurance coverage through medical assistance/charity are generally more responsive to *both outreach interventions* than patients with other types of insurance ($\widehat{\alpha}_{211}$ = ranging from -.016 to -.009, $p < .001$ for outreach alone; $\widehat{\alpha}_{211}$ = ranging from -.008 to -.001, $p < .001$–*n.s.* for outreach with patient navigation). Patients who receive healthcare at a low cost due to medical assistance/charity, with access to the corresponding insurance, are more likely to respond to outreach interventions because of their ability to overcome financial hardships to utilize screening services. A patient's ease of accessing healthcare services is based not only on their ability to pay for the service but also their proximity to healthcare providers. I find that patients with closer proximity to care are more responsive to *both outreach interventions* than patients with further proximity to care ($\widehat{\alpha}_{212}$ = .008, $p < .001$

for outreach alone; $\widehat{\alpha}_{312}$ = .005, $p < .01$ for outreach with patient navigation).

Neighborhood socioeconomic status. Patients who live in more educated neighborhoods are not necessarily more or less responsive to interventions ($\widehat{\alpha}_{213}$ = .002, *n.s.* for outreach alone; $\widehat{\alpha}_{313}$ = -.011, *n.s.* for outreach with patient navigation). Yet, patients who reside in a higher-income neighborhoods are more responsive to outreach with patient navigation ($\widehat{\alpha}_{214}$ = -.000, *n.s.* for outreach alone; $\widehat{\alpha}_{314}$ = .012, $p < .01$ for outreach with patient navigation), which implies that those in a low-income neighborhood are less responsive to this intervention. As patients in a low-income neighborhood face unique challenges such as higher rates of obesity, chronic disease, environmental pollutants, and incarceration,[34] these prevalent health and environmental challenges in the communities might cause anxiety and pessimism about their health (Conger et al. 1992), showing a lower responsiveness to outreach intervention. Patients in a neighborhood with a longer average commute time are less responsive to outreach alone ($\widehat{\alpha}_{215}$ = -.004, $p < .001$ for outreach alone), but it is no longer the case for outreach with patient navigation ($\widehat{\alpha}_{315}$ = -.000, *n.s.* for outreach with patient navigation). Patient navigation alleviates perceived costs associated with a screening by providing the information on the estimated duration for the appointment, so patients who live in a highly trafficked community will no longer show resistance to a screening.

While patient-level insurance coverage should capture the impact of health system accessibility, the neighborhood-level health insurance coverage can also offer

---

[34] https://www.healthaffairs.org/do/10.1377/hpb20180817.901935/full/

additional insights. Patients' responsiveness to the outreach interventions does not vary by the degree of private health insurance coverage in their neighborhood ($\widehat{\alpha}_{216}$ = -.000, *n.s.* for outreach alone; $\widehat{\alpha}_{316}$ = .002, *n.s.* for outreach with patient navigation). However, patients in a neighborhood with a greater public health insurance coverage are more responsive to outreach-alone intervention but not to outreach-with-patient-navigation intervention ($\widehat{\alpha}_{217}$ = .011, $p < .01$ for outreach alone; $\widehat{\alpha}_{317}$ = .006, *n.s.* for outreach with patient navigation).

Neighborhood unemployment rate does not significantly affect patients' responsive to the interventions ($\widehat{\alpha}_{218}$ = .002, *n.s.* for outreach alone; $\widehat{\alpha}_{318}$ = .003, *n.s.* for outreach with patient navigation). Yet, patients from a neighborhood with more dense populations are more responsive to both interventions ($\widehat{\alpha}_{319}$ = .005, $p < .05$ for outreach alone; $\widehat{\alpha}_{219}$ =.005, $p < .05$ for outreach with patient navigation), implying that patients in rural areas are less responsive to interventions.

Additional post-hoc analysis combining patient characteristics. As described in Table 3.3, I explore possible combinations of patient characteristics with the interactions between prior primary care visits and other patient characteristics. This is akin to examining higher-order interactions in an ANOVA. I offer several interesting insights.

The marginal benefit of additional primary care visits diminishes such that a patient's primary care visit has a nonlinear effect on outreach intervention effectiveness. Referring to columns 3 and 4 of Table 3.3, there is a positive linear coefficient (b = .020, $p < .001$ for outreach alone; b = .022, $p < .001$ for outreach with patient navigation) and a negative quadratic coefficient (b^2 = -.003, $p < .001$ for outreach alone; b^2 = -.004, $p <$

.001 for outreach with patient navigation) for the effect. Jointly, the coefficients capture diminishing returns such that a patient's first few primary care visits yield large marginal returns. Given the importance of the initial visits, to further enhance the outreach effectiveness, healthcare professionals should target patients who have sought medical services only a few times rather than those who have already made numerous visits.

The interactions between prior primary care visits and patient characteristics can help practitioners further identify the responsive subgroup. For example, Spanish-speaking patients' responsiveness to outreach interventions is attenuated as primary care visits increase (coefficient = -.005 $p < .001$ for both outreach alone and outreach with patient navigation). As Spanish-speaking patients view primary care visits as substitutes for screening services or perceive that outreach interventions are less informative than primary care visits, practitioners may target Spanish-speaking patients who have no prior primary care visits in the past. Patients with Child-Pugh B are even less responsive to outreach interventions as primary care visits increase (coefficient = -.003 $p < .001$ for outreach alone; coefficient = -.003, $p < .01$ for outreach with patient navigation), suggesting that increased primary care visits compound the perception of a fear and death anxiety of having cancer triggered by outreach interventions and thus decrease the use of screening. Overall, the post-hoc analysis highlights the need to understand how outreach effectiveness varies by the combination of patient characteristics.

3.4. Return on Cancer Outreach Interventions

3.4.1. Return on Non-Targeted Cancer Outreach Interventions

I evaluate the return on outreach interventions among patients and over time:

(6)

$$\text{Return}_k = \sum_{t=1}^{3} \sum_{i=1}^{N_{t=\{1,2,3\}}} \{ \Pr(\text{Screening}_{ikt}=1) \times$$

$$\left[\text{Benefit}_{ikt} - \text{Screening Cost}_{ikt} - \Pr\left(\text{Early Tumor}_{ikt}=1 \middle| \text{Screening}_{ikt}=1\right) \times \text{Treatment Cost}_{ikt}\right]$$

$$-\Pr(\text{Screening}_{ikt}=0) \times \text{Opportunity Cost}_{ikt} - \text{Outreach Cost}_{ikt} \}$$

where $\Pr(\text{Screening}_{ikt}=1)$ refers to the probability that patient i assigned to outreach type k completes the screening in period t. Further,

- If a patient completes the screening test,
  - The healthcare organization generates Benefit_{ikt} for patient i receiving intervention type k in period t captured by the quality-adjusted life years of a patient attributable to the screening (typically expressed in the financial value in the medical literature).
  - The healthcare organization incurs $\text{Screening Cost}_{ikt}$ for patient i receiving intervention type k in period t, which includes the costs of an ultrasound/MRI/CT test or a combination of these tests (i.e., each patient can complete multiple tests).
  - Conditional on being detected with an early tumor, the health-care provider incurs $\text{Treatment Cost}_{ikt}$ for patient i intervention type k in period t, which includes the costs of tumor resection, liver transplantation, and local ablative therapies.
- If a patient does not complete the screening test,
  - The healthcare organization incurs $\text{Opportunity Cost}_{ikt}$ if patient i receiving intervention type k in period t develops advanced HCC, which creates costs.

- Regardless of whether the patient completes the screening test,
  - Outreach $Cost_{ikt}$ is incurred if the healthcare organization employs an outreach program. The outreach costs are higher for the outreach-with-patient-navigation than the outreach-alone condition and are zero for the baseline condition.

Research has documented that HCC screening completion with biannual ultrasound extends patients' quality-adjusted life expectancy by 1.3 months, and HCC screening utilization with MRI does so by 2 months (Goossens et al. 2017). Patients may either complete ultrasound, MRI, CT, or a mix of these tests. I assume the average quality-adjusted life expectancy to be 1.65 months for each patient who completes the screening. The medical literature posits the financial value per quality-adjusted life year is $50,000 (Andersson et al. 2008; Goossens et al. 2017). Thus, total benefits can be obtained by multiplying the number of quality-adjusted life years by the financial value per quality-adjusted life year ($50,000) (i.e., multiply total number of patients who complete the screening by average quality-adjusted life expectancy).

Table 3.4 presents the results of the benefit-cost calculation using Equation (6) for each condition in each period. I use the *observed values* in the data (e.g., actual number of patients who visit) in conjunction with parameters (e.g., early detection rate) from the medical literature to calculate the return in each condition in each period. Details on parameters from the medical literature are documented in Appendix B.5.

3.4.1.1. No-outreach condition

I observe that 150/600, 156/591, and 126/577 patients in no-outreach condition

Table 3.4 Return on Non-Targeted Outreach Interventions

	No Outreach				Outreach Alone				Outreach with Patient Navigation			
	Period 1	**Period 2**	**Period 3**	***Total***	**Period 1**	**Period 2**	**Period 3**	***Total***	**Period 1**	**Period 2**	**Period 3**	***Total***
Cost and Benefit by Intervention Type and across Period												
Sample Size	600	591	577		600	592	584		600	589	582	
Costs												
Outreach costs	$0	$0	$0	$0	$52,157	$39,099	$32,030	$123,285	$60,594	$51,338	$45,617	$157,548
Screening costs	$88,541	$90,667	$76,339	$255,547	$111,335	$118,351	$94,233	$323,919	$127,598	$105,015	$100,258	$332,871
Treatment costs	$557,975	$580,293	$468,699	$1,606,967	$1,000,634	$944,837	$859,281	$2,804,752	$1,075,031	$1,015,514	$967,156	$3,057,700
Opportunity costs	$539,226	$521,252	$540,424	$1,600,902	$396,631	$405,019	$422,993	$1,224,642	$372,665	$378,656	$385,846	$1,137,168
Total costs	$1,185,742	$1,192,212	$1,085,462	$3,463,416	$1,560,756	$1,507,305	$1,408,536	$4,476,598	$1,635,888	$1,550,523	$1,498,876	$4,685,287
Benefits												
Health benefits	$1,031,250	$1,072,500	$866,250	$2,970,000	$1,849,375	$1,746,250	$1,588,125	$5,183,750	$1,986,875	$1,876,875	$1,787,500	$5,651,250
Net Benefits (Benefits-Costs)	-$154,492	-$119,712	-$219,212	-$493,416	$288,619	$238,945	$179,589	$707,152	$350,987	$326,352	$288,624	$965,963
Net Benefits per Patient	-$257	-$203	-$380	-$840	$481	$404	$308	$1,192	$585	$554	$496	$1,635
Net Benefits among Population (N=3,217)	-$276,111	-$217,211	-$407,397	-$900,718	$515,825	$432,818	$329,759	$1,278,402	$627,292	$594,157	$531,788	$1,753,237

Return on Non-Targeted Outreach Interventions

	Period 1	**Period 2**	**Period 3**	**Total**
Total Net Benefits among Population (N=3,217)	$867,007	$809,764	$454,150	**$2,130,921**

Notes. I extrapolated the calculation to the 3,217 patients eligible for randomization. Outreach costs = number of call hours × cost per hour. Screening costs = number of ultrasound completed × cost per ultrasound + number of CT/MRI completed × Average unit cost (CT/MRI). Treatment costs = Early detection probability × Average cost of treatment (if detected early). Opportunity costs = Annual HCC probability × Annual cost of advanced HCC. Health benefits = Quality-adjusted life gain × Financial value per quality-adjusted life year.

completed the screening in Periods 1, 2, and 3, respectively. The total benefits of the no-outreach condition across three periods are estimated to be $2,970,000. The screening costs are the total costs of ultrasound, CT, and MRI tests completed. The number of ultrasound (CT and MRI) tests completed is 127 (69), 149 (68), and 113 (59) in Periods 1, 2, and 3. The cost per ultrasound is $143, while the average cost of CT and MRI is $1,020. Thus, the total screening costs are estimated to be $255,547. Focusing on treatment costs, 5% of the total number of screening tests typically result in early tumor detection. The average treatment cost per patient for early tumor detection is $74,397 (Goossens et al. 2017). Given 150/600, 156/591, and 126/577 patients in the no-outreach condition completed the screening in Periods 1, 2, and 3, 5% of them would undergo treatment costs, resulting in a total treatment cost of $1,606,967. Opportunity cost is incurred if patients who have not completed the screening develop advanced HCC. The annual cost of advanced HCC is $41,320, and the annual HCC probability is 2.9% (Goossens et al. 2017). Multiplying the number of patients who have not completed the screening in the no-outreach condition by the probability of HCC (2.9%) results in a total opportunity cost of $1,600,902. Finally, outreach costs are zero in the no-outreach condition. Subtracting the total cost from the total benefit, the total return in the no-outreach condition is -$493,416, which translates to a loss of $840 per patient to the healthcare system.

3.4.1.2. Outreach-alone condition

Compared to the no-outreach condition, there are higher benefits in the outreach-alone case ($5,183,750 vs. $2,970,000) as there are 269/600 patients, 254/592 patients, and 231/584 patients who have completed the screening in Periods 1, 2, and 3. Using the same approach as the one for no-outreach condition to calculate the costs, the screening costs, treatment costs, and opportunity costs are $323,919, $2,804,752, and $1,224,642 in the outreach-alone condition. In addition, the total number of hours devoted to outreach calls is 3,477, 2,607 and 2,135. Assuming a $15 hourly wage, the total outreach cost in the outreach-alone condition is $123,285. Thus, the total return in the outreach-alone condition is $707,152, which translates to a gain of $1,192 per patient to the healthcare system.

3.4.1.3. Outreach-with-patient-navigation condition

Compared to the no-outreach condition, there are higher benefits in the outreach-with-patient-navigation condition ($5,651,250 vs. $2,970,000) as there are 289/600 patients, 273/589 patients, and 260/582 patients who would complete the screening. Following the same calculation approach, the screening costs, treatment costs, opportunity costs, and outreach costs are $332,871, $3,057,700, $1,137,168, and $157,548 in the outreach-with-patient-navigation condition. Thus, the total return in the outreach-with-patient-navigation condition is $965,963, which is substantially greater than that in the no-outreach condition and translates to a gain of $1,635 per patient to the healthcare system.

3.4.1.4. Summary

No outreach results in a net loss of $840 per patient to the medical hospital, while outreach alone (outreach with patient navigation) generates a monetary gain of $1,192 ($1,635) per patient. When extrapolated to the 3,217 patients eligible for randomization from the hospital’s patient database, the cancer intervention results in a loss of $900,718 from no outreach or usual care, a gain of $1,278,402 from outreach alone, and a gain of $1,753,237 from outreach with patient navigation. In this scenario, the total gain is $2,130,921.

3.4.2. Return on Patient-Level Targeted Cancer Outreach Interventions: A Simulation Exercise

Thus far, the calculation of return on cancer interventions is based on the (1) random assignment of patients, and (2) patients remaining in the same condition over three periods. However, (a) given patient characteristics, there is heterogeneity in patient-level treatment effects of outreach-alone intervention and in those of outreach-with-patient-navigation intervention; (b) not all patient-level treatment effect estimates are statistically greater than 0; (c) treatment effect heterogeneity varies across periods; (d) the net return on outreach interventions varies across intervention types and over time. In other words, given each patient’s characteristics in a particular period, outreach with patient navigation is unlikely to be uniformly more effective than outreach alone. This poses two questions: (1) given each patient’s observed characteristics, which intervention type is most suitable for each patient? and (2) for the same patient, does the most suitable intervention vary across periods?

Accordingly, I conduct a simulation that assigns each patient to the most suitable condition in each period based on two types of allocation schemes: 1) predicted treatment effect, and 2) predicted net return (see detailed procedure in Appendix B.6).

3.4.2.1. Recommended allocation based on predicted treatment effect

Conceptually, in each period, given each patient's profile, I compare each patient's treatment effect estimate in their corresponding condition (e.g., condition 2) with their simulated treatment effect estimate in the counterfactual condition (e.g., condition 3). Then I assign this patient to the most suited intervention that generates a significantly higher treatment effect estimate (e.g., condition 3, $p < .05$). If none of these estimates is significantly larger than 0, I assign this patient to condition 1.

The recommended allocation for each period is:

1. Period 1: 0%, 99.9%, and .1% of patients are assigned to the no-outreach, outreach-alone and outreach-with-patient-navigation condition.
2. Period 2: 9.0%, 74.0%, and 16.9% of patients in each condition, respectively.
3. Period 3: 8.4%, 66.9%, and 24.7% of patients in each condition, respectively.

There are four noteworthy points in this recommendation. First, the recommended split deviates from the original allocation based on the randomized controlled trial (1:1:1), suggesting that targeting induces asymmetric allocation of patients to different conditions. Second, there is a fraction of patients who stay in the baseline condition in Periods 2 and 3. For these patients, neither of the interventions is more effective than the baseline. Third, I reallocate most patients to the outreach-alone condition, suggesting that healthcare organizations can achieve the same level of effectiveness by aligning only moderate outreach efforts with these patients. Fourth, over time, the outreach-with-patient-navigation condition seems to be more effective based on higher allocation to this condition (.1% in Period 1, 16.9% in Period 2, and 24.7% in Period 3).

Table 3.5 shows the return on patient-level targeted outreach interventions. When extrapolated to the 3,217 patients eligible for randomization from the hospital's patient database, patient-level targeted outreach program across conditions generates a gain of $1,547,662, $1,204,318, and $952,290 in Periods 1, 2, and 3, respectively. The total net return on patient-level targeted outreach program is $3,704,270, or 74% higher than that on non-targeted outreach program based on the random assignment ($2,130,921).

3.4.2.2. Recommended allocation based on predicted net return

What if the purpose of the healthcare system is to maximize the overall return derived from assigning each patient to the most suitable intervention? As such, I can assign each patient to the intervention that gives the *highest predicted net return* based on Equation 6 rather than only the *highest predicted treatment effect*. Specifically, in each period, I compare each patient's predicted net return in the corresponding condition (e.g., condition 2) with his/her simulated patient's estimated net return in the counterfactual condition (e.g., condition 3). Then I assign this patient to the most suited intervention that generates a significantly higher net return (e.g., condition 3, $p < .05$). If none of these estimates is significantly larger than 0, I assign this patient to condition 2 since the net return in the baseline condition is negative.

The recommended allocation for each period is:

1. Period 1: 0%, 87.2%, and 12.8% of patients are assigned to the no-outreach, outreach-alone and outreach-with-patient-navigation condition.
2. Period 2: 0%, 79.3%, and 20.7% of patients in each condition, respectively.
3. Period 3: 0%, 68.7%, and 31.3% of patients in each condition, respectively.

Table 3.5 Simulation Results Based on Patient-Level Treatment Effect Estimates

	No Outreach			Outreach Alone			Outreach with Patient Navigation		
	Period 1	**Period 2**	**Period 3**	**Period 1**	**Period 2**	**Period 3**	**Period 1**	**Period 2**	**Period 3**
Cost and Benefit by Intervention Type and across Period									
Sample Size after Reallocation	0	160	147	1,799	1,312	1,166	1	300	430
Net Benefits per Patient	-$257	-$203	-$380	$481	$404	$308	$585	$554	$496
Total Net Benefits among Population (N=3,217)	$0	-$58,838	-$103,076	$1,546,617	$961,383	$661,788	$1,045	$301,773	$393,578
Return on Patient-Level Targeted Outreach Interventions									
	Period 1	**Period 2**	**Period 3**	**Total (Improvement)**					
Total Net Benefits among Population (N=3,217)	$1,547,662	$1,204,318	$952,290	**$3,704,270 (74%)**					

Table 3.6 Simulation Results Based on Patient-Level Estimated Return

	No Outreach			Outreach Alone			Outreach with Patient Navigation		
	Period 1	**Period 2**	**Period 3**	**Period 1**	**Period 2**	**Period 3**	**Period 1**	**Period 2**	**Period 3**
Cost and Benefit by Intervention Type and across Period									
Sample Size after Reallocation	0	0	0	1,570	1,406	1,198	230	366	545
Net Benefits per Patient	-$257	-$203	-$380	$481	$404	$308	$585	$554	$496
Total Net Benefits among Population (N=3,217)	$0	$0	$0	$1,349,743	$1,030,263	$679,950	$240,462	$368,163	$498,837
Return on Patient-Level Targeted Outreach Interventions									
	Period 1	**Period 2**	**Period 3**	**Total (Improvement)**					
Total Net Benefits among Population (N=3,217)	$1,590,205	$1,398,426	$1,178,788	**$4,167,419 (96%)**					

There are two noteworthy points in this recommendation. First, nobody stays in the baseline condition under this allocation scheme, reflecting the goal of maximizing overall return. Second, while I still reallocate most patients to the outreach-alone condition, the outreach-with-patient-navigation intervention seems to be even more effective over time based on higher allocation to this condition than the previous allocation (e.g., 31.3% vs. 24.7% in Period 3).

Table 3.6 shows the return on patient-level targeted outreach interventions. When extrapolated to the 3,217 patients eligible for randomization from the hospital's patient database, the total net return on patient-level targeted outreach program is $4,167,419, or 96% higher than that on non-targeted outreach program based on the random assignment ($2,130,921). In summary, patient-level targeted outreach interventions improve the payoffs to each individual as well as to the healthcare system by 74%-96%.

Notably, the difference in allocation based on predicted treatment effect versus predicted net return shows the versatility of my approach in providing practical guidance to medical professionals and policy makers. The nature and magnitude of benefits can shift based on the goals set by a healthcare institution. Therefore, it is critical for an organization to have well-defined goals to maximize benefits from personalized medicine, as illustrated here. Moreover, these results confirm that the cumulated benefits from repeated and upgraded health education through outreach with patient navigation are enhanced through individually-tailored and dynamically-adaptive education.

3.5. Concluding Remarks

In this study, the main-effects analysis would have led scholars to conclude that the outreach with patient navigation and outreach alone are equally effective. The application of causal forests uncovers patient heterogeneity in outreach effectiveness and leads to different conclusions. Specifically, as described earlier, patients with different characteristics respond very differently to each intervention. For example, patients who are more responsive to outreach alone or outreach with patient navigation tend to be female, part of minority populations, in better health status, covered by medical assistance, have closer proximity to clinics, and reside in a populated neighborhood. Furthermore, patients who are more responsive to outreach alone tend to be younger, commute faster, and reside in neighborhoods with more public insurance coverage, while those to outreach with patient navigation tend to be older and reside in a higher-income neighborhood. It is noteworthy that over time the outreach-with-patient-navigation intervention becomes more effective for an increased proportion of patients. As such, I illustrate time-varying heterogeneity in the outreach effectiveness.

A cost-benefit analysis shows that the baseline condition results in a net loss of $840 per patient whereas outreach alone (outreach with patient navigation) generates a gain of $1,192 ($1,635) per patient. When extrapolated to the 3,217 eligible patients, the total net gain of the non-targeted cancer outreach program across conditions is $2,130,921, which implies that outreach marketing provides a substantial positive payoff to the healthcare system. The simulation shows that targeted outreach interventions can enhance such return by 74%-96%.

3.5.1. Research Implications

For the marketing discipline, this paper provides a framework for better understanding and analyzing sufficiently powered field experiments that are based on random assignment of heterogeneous customer groups. Instead of focusing only on the main effects of the interventions or a subset of individual-level covariates, causal forests flexibly predict personalized treatment effects based on high-dimensional, nonlinear functions of those covariates. Such an approach also obviates the need for choosing several one-way interactions a priori to test for heterogeneity or searching over many interactions for particularly responsive subgroups. Accordingly, this paper provides a methodological solution to the field's concern of external validity. Many empirical findings are typically much less generalizable than I imagine, because researchers lack a process and corresponding insights to identify moderators (i.e., the interaction of treatment and unmodeled/unmanipulated background factors) (Cook and Campbell 1979; Lynch 1982).

For the emerging discipline of personalized healthcare, I show how causal forests can identify particularly responsive subgroups without the need for a larger number of experimental conditions. While modern healthcare has implemented personalized medicine using genetic information, most healthcare outreach and educational programs still rely on untailored communications. Practitioners who manage these programs should recognize that the use of a large number of patient characteristics can substantially improve the outreach responsiveness through a tailored approach.

This research also responds to a recent call for boundary-breaking marketing-relevant research (MacInnis et al. forthcoming) in several ways. First, many of the covariates are driven from "real-world phenomena, rather than the constructs and theories in the marketing” (p. 24). The findings on the treatment effects that vary across covariates—which are both pragmatic and theoretical in its origin—not only engage “academics in other disciplines” (p. 2) but also offer important implications to the extant literature and theory going forward. Second, the findings from this paper has “life and death implications” (p. 20) because it can contribute to detecting liver cancer at early stages. Third, the covariates, such as ethnicity, language, insurance coverage, and neighborhood socioeconomic status, help understand how the outreach effectiveness will vary by “under-studied consumers such as minorities, privileged or impoverished classes, and marginalized consumers like special needs populations” (p. 10).

3.5.2. Healthcare Marketing Implications: Toward a Patient-Centric Healthcare

I urge hospitals and medical centers with outreach programs to fully leverage patient information to improve the effectiveness of outreach investments. The use of machine learning can power data-driven patient-centric outreach programs. Accordingly, hospitals and healthcare practitioners should realize that a “one-size-fits-all” outreach program is neither effective nor economic. Furthermore, the outreach programs should become more dynamically adaptive. This finding shows that outreach with patient navigation yields greater effectiveness and return over time, so practitioners should consider both cross-sectional and temporal adaptation of outreach programs as necessary.

I urge policy makers in the federal, state, and local health departments, American Hospital Association, American Cancer Society, and American Liver Foundation to financially support personalized outreach programs. In particular, more hospitals should reach out to the underrepresented populations (e.g., ethnic minority) as they are more responsive to outreach messages, and doing so will require additional resources and staff training. There is a need to incentivize hospitals to reach out to patients with these varying personal, clinical, structural, and socioeconomic backgrounds. In addition, they should engage a multidisciplinary group from healthcare, marketing, computer science, and other disciplines to fund an accumulation of comprehensive databases to facilitate better targeting of patients to improve outcomes.

3.5.3. Limitations and Future Research

First, as patients have different barriers to screening, future research should test the effectiveness of different barrier-reduction strategies by analyzing the nature of communication between patients and the staff with the use of call recordings. Second, this study focuses on the endpoint outcome, screening completion. Future research could apply the notion of customer journey to disentangle which parts of the intervention (e.g., barrier discussion during an outreach call vs. reminder calls) are more effective at not only increasing completion but also reducing no-show rates or time-to-response, further enhancing the return. Third, while I track individual patients, outreach designed to serve an individual may have influenced other members of the household. Future research could study possible spillover effects of outreach interventions.

3.6. References

Allenby, Greg M. and Peter E. Rossi (1998), "Marketing Models of Consumer Heterogeneity," *Journal of Econometrics*, 89 (1/2), 57–78.

American Cancer Society (2019), "Liver Cancer Survival Rates," [available at https://www.cancer.org/cancer/liver-cancer/detection-diagnosis-staging/survival-rates.html].

Andersson, Karin L., Joshua A. Salomon, Sue J. Goldie, and Raymond T. Chung (2008), "Cost Effectiveness of Alternative Surveillance Strategies for Hepatocellular Carcinoma in Patients with Cirrhosis," *Clinical Gastroenterology and Hepatology*, 6 (12), 1418–24.

Arndt, Jamie, Clay Routledge, and Jamie L. Goldenberg (2006), "Predicting Proximal Health Responses to Reminders of Death: The Influence of Coping Style and Health Optimism," *Psychology and Health*, 21 (5), 593–614.

Aro, Arja R., Harry J. de Koning, Pilvikki Absetz, and Marjut Schreck (2001), "Two Distinct Groups of Non-Attenders in an Organized Mammography Screening Program," *Breast Cancer Research and Treatment*, 70 (2), 145–53.

Ascarza, Eva (2018), "Retention Futility: Targeting High-Risk Customers Might Be Ineffective," *Journal of Marketing Research*, 55 (1), 80–98.

Athey, Susan and Guido Imbens (2016), "Recursive Partitioning for Heterogeneous Causal Effects," *Proceedings of the National Academy of Sciences*, 113 (27), 7353–60.

Basch, Charles E., Randi L. Wolf, Corey H. Brouse, Celia Shmukler, Alfred Neugut, Lawrence T. DeCarlo, and Steven Shea (2006), "Telephone Outreach to Increase Colorectal Cancer Screening in an Urban Minority Population," *American Journal of Public Health*, 96 (12), 2246–53.

Basil, Michael D. and William J. Brown (1997), "Marketing Aids Prevention: The

Differential Impact Hypothesis Versus Identification Effects," *Journal of Consumer Psychology*, 6 (4), 389–411.

Benoit, Ronald M., Henrik Grönberg, and Michael James Naslund (2001), "A Quantitative Analysis of the Costs and Benefits of Prostate Cancer Screening," *Prostate Cancer and Prostatic Diseases*, 4 (3), 138–45.

Blanchard, Simon J., Wayne S. DeSarbo, A. Selin Atalay, and Nukhet Harmancioglu (2012), "Identifying Consumer Heterogeneity in Unobserved Categories," *Marketing Letters*, 23 (1), 177–94.

Bolton, Lisa E., Americus Reed II, Kevin G. Volpp, and Katrina Armstrong (2008), "How Does Drug and Supplement Marketing Affect a Healthy Lifestyle?" *Journal of Consumer Research*, 34 (5), 713–26.

Borsky, Amanda, Chunliu Zhan, Therese Miller, Quyen Ngo-Metzger, Arlene S. Bierman, and David Meyers (2018), "Few Americans Receive All High-Priority, Appropriate Clinical Preventive Services," *Health Affairs*, 37 (6), 925–28.

Breiman, Leo (2001), "Random Forests," *Machine Learning*, 45 (1), 5–32.

Bruix, Jordi and Morris Sherman (2011), "Management of Hepatocellular Carcinoma: An Update," *Hepatology*, 53 (3), 1020–22.

Bühlmann, Peter and Bin Yu (2002), "Analyzing Bagging," *The Annals of Statistics*, 30 (4), 927–61.

Caplan, Lee S., Barbara L. Wells, and Suzanne Haynes (1992), "Breast Cancer Screening Among Older Racial/Ethnic Minorities and Whites: Barriers to Early Detection," *Journal of Gerontology* 47, 101–10.

Carlson, Rae (1971), “Sex Differences in Ego Functioning: Exploratory Studies of Agency and Communion,” *Journal of Consulting and Clinical Psychology*, 37 (2), 267–77.

Chubak, Jessica and Rebecca Hubbard (2016), “Defining and Measuring Adherence to Cancer Screening,” *Journal of Medical Screening*, 23 (4), 179–85.

Cole, Catherine A., Gilles Laurent, Aimee Drolet, Jane Ebert, Angela Gutchess, Raphaëlle Lambert-Pandraud, Etienne Mullet, Michael I. Norton, and Ellen Peters (2008), “Decision Making and Brand Choice by Older Consumers,” *Marketing Letters* 19 (3/4), 355–65.

Conger, Rand D., Katherine J. Conger, Glen H. Elder Jr, Frederick O. Lorenz, Ronald L. Simons, and Les B. Whitbeck (1992), “A Family Process Model of Economic Hardship and Adjustment of Early Adolescent Boys,” *Child Development*, 63 (3), 526–41.

Cook, Thomas D. and Donald T. Campbell (1979), *Quasi-Experimentation Design & Analysis Issues for Field Settings*. Boston, MA: Houghton Mifflin Company.

Efron, Bradley (2014), “Estimation and Accuracy after Model Selection,” *Journal of the American Statistical Association*, 109 (507), 991–1007.

Ferrante, Jeanne M., Ping-Hsin Chen, and Steve Kim (2008), “The Effect of Patient Navigation on Time to Diagnosis, Anxiety, and Satisfaction in Urban Minority Women with Abnormal Mammograms: A Randomized Controlled Trial,” *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 85 (1), 114–24.

Figueroa, Jose F. and Ashish K. Jha (2018), "Approach for Achieving Effective Care for High-Need Patients," *JAMA Internal Medicine*, 178 (6), 845–46.

Goossens, Nicolas, Amit G. Singal, Lindsay Y. King, Karin L. Andersson, Bryan C. Fuchs, Cecilia Besa, Bachir Taouli, Raymond T. Chung, and Yujin Hoshida (2017), "Cost-Effectiveness of Risk Score–Stratified Hepatocellular Carcinoma Screening in Patients with Cirrhosis," *Clinical and Translational Gastroenterology*, 8 (6), e101, 1–9.

Goldman, Shira N., David T. Liss, Tiffany Brown, Ji Young Lee, David R. Buchanan, Kate Balsley, Ana Cesan, Jordan Weil, Bridget H. Garrity, and David W. Baker (2015), "Comparative Effectiveness of Multifaceted Outreach to Initiate Colorectal Cancer Screening in Community Health Centers: A Randomized Controlled Trial," *Journal of General Internal Medicine*, 30 (8), 1178–84.

Grossman, Christopher H., Joanne Brooker, Natasha Michael, and David Kissane (2018), "Death Anxiety Interventions in Patients with Advanced Cancer: A Systematic Review," *Palliative Medicine*, 32 (1), 172–84.

Guo, Tong, S. Sriram, and Puneet Manchanda (2019), "The Effect of Information Disclosure on Industry Payments to Physicians," working paper.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (2009), *The Elements of Statistical Learning: Prediction, Inference, and Data Mining*. New York: Springer-Verlag.

Humiston, Sharon G., Nancy M. Bennett, Christine Long, Shirley Eberly, Lourdes Arvelo, Joseph Stankaitis, and Peter G. Szilagyi (2011), "Increasing Inner-City

Adult Influenza Vaccination Rates: A Randomized Controlled Trial," *Public Health Reports*, 126 (2_suppl), 39–47.

Hutchinson, J. Wesley, Wagner A. Kamakura, and John G. Lynch, Jr. (2000), "Unobserved Heterogeneity as an Alternative Explanation for "Reversal" Effects in Behavioral Research," *Journal of Consumer Research*, 27 (3), 324–44.

Johns Hopkins Medicine (2019), "The Sidney Kimmel Comprehensive Cancer Center: Outreach and Education Programs in Your Community,".

Kaldenberg, Dennis O., Harold F. Koenig, and Boris W. Becker (1994), "Mail Survey Response Rate Patterns in a Population of the Elderly: Does Response Deteriorate with Age?" *Public Opinion Quarterly*, 58 (1), 68–76.

King, Abby C., W. Jack Rejeski, and David M. Buchner (1998), "Physical Activity Interventions Targeting Older Adults: A Critical Review and Recommendations," *American Journal of Preventive Medicine*, 15 (4), 316–33.

Klein, William M.P. and Jennifer L. Cerully (2007), "Health-Related Risk Perception and Decision-Making: Lessons from the Study of Motives in Social Psychology," *Social and Personality Psychology Compass*, 1 (1), 334–58.

Lasser, Karen E., Jennifer Murillo, Sandra Lisboa, A. Naomie Casimir, Lisa Valley-Shah, Karen M. Emmons, Robert H. Fletcher, and John Z. Ayanian (2011), "Colorectal Cancer Screening Among Ethnically Diverse, Low-Income Patients: A Randomized Controlled Trial," *Archives of Internal Medicine*, 171 (10), 906–12.

Lisjak, Monika and Angela Y. Lee (2014), "The Bright Side of Impulse: Depletion

Heightens Self-Protective Behavior in the Face of Danger," *Journal of Consumer Research*, 41 (1), 55–70.

Luce, Mary Frances and Barbara E. Kahn (1999), "Avoidance or Vigilance? The Psychology of False-Positive Test Results," *Journal of Consumer Research*, 26 (3), 242–59.

Lynch, John G. (1982), "On the External Validity of Experiments in Consumer Research," *Journal of Consumer Research*, 9 (3), 225–39.

MacInnis, Deborah J., Vicki G. Morwitz, Simona Botti, Donna L. Hoffman, Robert V. Kozinets, Donald R. Lehmann, John G. Lynch Jr., and Cornelia Pechmann (forthcoming), "Creating Boundary-Breaking, Marketing-Relevant Consumer Research," *Journal of Marketing*.

Marrero, Jorge A., Laura M. Kulik, Claude B. Sirlin, Andrew X. Zhu, Richard S. Finn, Michael M. Abecassis, Lewis R. Roberts, and Julie K. Heimbach (2018), "Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases," *Hepatology*, 68 (2), 723-50.

Marketing Science Institute (2016), "2016-2018 Research Priorities: A Guide to MSI Research Programs and Procedures," Cambridge, MA: Marketing Science Institute.

McCarthy, Ona, Irrfan Ahamed, Firuza Kulaeva, Ravshan Tokhirov, Salokhiddin Saibov, Marieka Vandewiele, Sarah Standaert, Baptiste Leurent, Phil Edwards, Melissa Palmer, and Caroline Free (2018), "A Randomized Controlled Trial of an Intervention Delivered by Mobile Phone App Instant Messaging to Increase the Acceptability of Effective Contraception among Young People in Tajikistan," *Reproductive Health*, 15 (28), 1–14.

McMahon, Pamela M., Chung Yin Kong, Bruce E. Johnson, Milton C. Weinstein, Jane C. Weeks, Karen M. Kuntz, Jo-Anne O. Shepard, Stephen J. Swensen, and G. Scott Gazelle (2008), "Estimating Long-Term Effectiveness of Lung Cancer Screening in the Mayo Ct Screening Study," *Radiology*, 248 (1), 278–87.

Menon, Geeta, Lauren G. Block, and Suresh Ramanathan (2002), "We're at as Much Risk as We Are Led to Believe: Effects of Message Cues on Judgments of Health Risk," *Journal of Consumer Research*, 28 (4), 533–49.

Moorman, Christine and Erika Matulich (1993), "A Model of Consumers' Preventive Health Behaviors: The Role of Health Motivation and Health Ability," *Journal of Consumer Research*, 20 (2), 208–28.

National Cancer Institute (2018), "Cancer Statistics," [available at https://www.cancer.gov/about-cancer/understanding/statistics].

Nyström, L., S. Wall, L. E. Rutqvist, A. Lindgren, M. Lindqvist, S. Rydén, J. Andersson, N. Bjurstam, G. Fagerberg, J. Frisell, L. Tabár, and L. G. Larsson (1993), "Breast Cancer Screening with Mammography: Overview of Swedish Randomised Trials," *The Lancet*, 341 (8851), 973–78.

Rosenbaum, Paul R. and Donald B. Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70 (1), 41–55.

Rubin, Donald B. (1974), "Estimating Causal Effects of Treatments in Randomized and

Nonrandomized Studies," *Journal of Educational Psychology*, 66 (5), 688–701.

Santi, Valentina, Franco Trevisani, Annagiulia Gramenzi Alice Grignaschi, Federica Mirici-Cappa, Paolo Del Poggio, Maria Anna Di Nolfo, Luisa Benvegnù, Fabio Farinati, Marco Zoli, Edoardo Giovanni Giannini, Franco Borzio, Eugenio Caturelli, Maria Chiaramonte, and Mauro Bernardi (2010), "Semiannual Surveillance is Superior to Annual Surveillance for the Detection of Early Hepatocellular Carcinoma and Patient Survival," *Journal of Hepatology*, 53 (2), 291-97.

Singal, Amit G., Jasmin A. Tiro, Caitlin C. Murphy, Jorge A. Marrero, Katharine McCallister, Hannah Fullington, Caroline Mejias, Akbar K. Waljee, Wendy Pechero Bishop, and Noel O. Santini (2019), "Mailed Outreach Invitations Significantly Improve HCC Surveillance Rates in Patients with Cirrhosis: A Randomized Clinical Trial," *Hepatology*, 69 (1), 121–30.

Singal, Amit G., Jorge A. Marrero, and Adam Yopp (2014), "Screening Process Failures for Hepatocellular Carcinoma," *Journal of the National Comprehensive Cancer Network*, 12 (3), 375-82.

Singal, Amit G., Anjana Pillai, and Jasmin Tiro (2014), "Early Detection, Curative Treatment, and Survival Rates for Hepatocellular Carcinoma Surveillance in Patients with Cirrhosis: A Meta-Analysis," *PLOS Medicine*, 11 (4), e1001624, 1–20.

Singal, Amit G., Adam Yopp, Celette S. Skinner, Milton Packer, William M. Lee, and Jasmin A. Tiro (2012), "Utilization of Hepatocellular Carcinoma Surveillance among American Patients: A Systematic Review," *Journal of General Internal Medicine*, 27 (7), 861–67.

Skolnik, Neil S. (2011), *Electronic Medical Records*. Totowa, New Jersey: Humana Press.

Sun, Chenshuo, Zijun (June) Shi, Xiao Liu, Anindya Ghose, Xueying Li, and Feiyu Xiong (2019), "The Effect of Voice AI on Consumer Purchase and Search Behavior," working paper.

Szczepura, A. (2005), "Access to Health Care for Ethnic Minority Populations," *Postgraduate Medical Journal*, 81 (953), 141–47.

Trinchet, Jean-Claude, Cendrine Chaffaut, Valérie Bourcier, Françoise Degos, Jean Henrion, Hélène Fontaine, Dominique Roulot, Ariane Mallat, Sophie Hillaire, Paul Cales, Isabelle Ollivier, Jean-Pierre Vinel, Philippe Mathurin, Jean-Pierre Bronowicki, Valérie Vilgrain, Gisèle N'Kontchou, Michel Beaugrand, and Sylvie Chevret (2011), "Ultrasonographic Surveillance of Hepatocellular Carcinoma in Cirrhosis: A Randomized Trial Comparing 3- and 6-Month Periodicities," *Hepatology*, 54 (6), 1987–97.

Trudeau, Kimberlee J., Sharon Danoff-Burg, Tracey A. Revenson, and Stephen A. Paget (2003), "Agency and Communion in People with Rheumatoid Arthritis," *Sex Roles*, 49 (7/8), 303–11.

Tzartzeva Kristina, Joseph Obi, Nicole E. Rich, Neehar D. Parikh, Jorge A. Marrero, Adam Yopp, Akbar K. Waljee, and Amit G. Singal (2018), "Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients with Cirrhosis: A Meta-Analysis," *Gastroenterology*, 154 (6), 1706–18.

Wager, Stefan and Susan Athey (2018), "Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests," *Journal of the American Statistical Association*, 113 (523), 1228–42.

Wager, Stefan, Trevor Hastie, and Bradley Efron (2014), "Confidence Intervals for Random Forests: The Jackknife and the Infinitesimal Jackknife," *The Journal of Machine Learning Research*, 15 (1), 1625–51.

Wang, Wenbo, Hean Tat Keh, and Lisa E. Bolton (2010), "Lay Theories of Medicine and a Healthy Lifestyle," *Journal of Consumer Research*, 37 (1), 80–97.

Wetherell, Julie Loebach, Andrew J. Petkus, Steven R. Thorp, Murray B. Stein, Denise A. Chavira, Laura Campbell-Sills, Michelle G. Craske, Cathy Sherbourne, Alexander Bystritsky, Greer Sullivan, and Peter Roy-Byrne (2013), "Age Differences in Treatment Response to a Collaborative Care Intervention for Anxiety Disorders," *The British Journal of Psychiatry*, 203 (1), 65–72.

Zeithaml, Valarie A., Bernard J. Jaworski, Ajay K. Kohli, Kapil R. Tuli, Wolfgang Ulaga, and Gerald Zaltman (2020), "A Theories-in-Use Approach to Building Marketing Theory," *Journal of Marketing*, 84 (1), 32–51.

Zhang, Bo-Heng, Bing-Hui Yang, and Zhao-You Tang (2004), "Randomized Controlled Trial of Screening for Hepatocellular Carcinoma," *Journal of Cancer Research and Clinical Oncology*, 130 (7), 417–22.

4. CONCLUSIONS

While most research in marketing focuses on determining the financial return on marketing investment for organizations in commercial sectors, my dissertation consisting of two essays augments it with social impact of marketing investment for social-sector organizations such as K-12 public schools and healthcare institutions. The common theme of these essays is to measure the social impact of marketing interventions and provide an implementable approach that can improve the effectiveness of marketing interventions for social-sector organizations.

Essay 1 offers insights concerning whether and how internet technology investment contributes to school-level outcomes. Even though many parents and school district administrators advocate investing in internet access to improve academic outcomes, the contribution of SDIAS to school performance is ex-ante ambiguous. I quantify the effect of SDIAS on school academic performance and school disciplinary problems in Texas. Increased school district internet access spending simultaneously enhances school performance in 9 out of 10 performance indicators, and drives a 7% increase in the number of school disciplinary problems related to Part II offenses. These effects are exacerbated in regions where households have better internet access. I establish that these effects by a combination of different identification strategies and a rich set of robustness checks to rule out unobservable factors that might drive our results.

Essay 2 offers guidance on improving cancer outreach effectiveness through a combination of field experiment and personalization. Relying only on the main-effects analysis, scholars might conclude that the outreach with patient navigation and outreach alone are equally effective. However, the application of causal forests uncovers patient heterogeneity in outreach effectiveness and leads to different conclusions and important practical implications. Specifically, patients with different characteristics respond very differently to each intervention. For example, patients who are more responsive to outreach alone or outreach with patient navigation tend to be female, be part of minority populations, be in better health status, be covered by medical assistance, have closer proximity to clinics, and reside in a populated neighborhood. Patients who are more responsive to outreach alone tend to be younger, have faster commutes, and reside in neighborhoods with more public insurance coverage. Patients responsive to outreach with patient navigation tend to be older and reside in a higher-income neighborhood. Over time, the outreach-with-patient-navigation intervention becomes more effective for an increased proportion of patients. As such, I illustrate time-varying heterogeneity in the outreach effectiveness.

Furthermore, a cost–benefit analysis shows that the baseline condition results in a net loss of $840 per patient, whereas outreach alone (outreach with patient navigation) generates a gain of $1,192 ($1,635) per patient. When extrapolated to the 3,217 eligible patients, the total net gain of the nontargeted cancer outreach program across conditions is $2,130,921. The simulation shows that targeted outreach interventions can enhance this return by 74%–96%, implying that outreach marketing provides a substantial positive payoff to the healthcare system.

APPENDIX A

APPENDICES FOR ESSAY 1

A.1. Pilot Study: Customers' Perspective on School Internet Access

I examine the extent to which overall parental satisfaction, a measure of customer equity, with a school is associated with the satisfaction with digital technology and internet access in school. I obtain the data from a proprietary survey conducted during 2018–2019 on 3,924 parents whose children enroll in traditional public schools or charter schools in the United States. Respondents rated their satisfaction with digital technology and internet access in school. In addition, I obtain (1) their overall satisfaction with the school, (2) their reenrollment intention into the same public school, and (3) switching intention to private schools.

A.1.1. Measures and model specification

Satisfaction with internet access. Each participant rated a statement "Taking everything into account, please rate your overall satisfaction with digital technology and internet access at your child's school" (1 = very dissatisfied, 5 = very satisfied).

Overall satisfaction. Overall satisfaction was measured as: "Taking into account your total experience, please rate your overall satisfaction with your child's school?" (1 = very dissatisfied, 5 = very satisfied).

Reenrollment intention. Each participant rated "Taking into account your total experience, how likely will you be to re-enroll your child at this school if the opportunity arose?" (1 = very unlikely, 5 = very likely).

Switching intention. Each participant rated "Taking into account your total experience, how likely will you be to enroll your child in a private school if the opportunity arose?" (1 = very unlikely, 5 = very likely).

Control variables. In the equation of overall satisfaction, I include satisfaction with three additional attributes that capture parents' perception of school quality—i.e., academics and learning ("preparing your child for college"), teachers ("teachers provide regular updates about your child's progress"), and school environment and facilities ("the school's facilities foster a positive learning environment")—which were measured on a 5-point scale (1 = very dissatisfied, 5 = very satisfied). In addition, I control for participants' involvement with their child's school and education (1 = not at all involved, 2 = somewhat involved, 3 = extremely involved), the grade in which the child is (Kindergarten–5th, 6th–8th, 9th–12th, or other), education (high school, associates/undergraduate degree, masters/graduate degree, and Ph.D., MD, JD), ethnicity, and household income (1 = less than \$25,000, 2 = \$26,000-\$50,000, 3 = \$51,000-\$100,000, 4 = \$101,000-\$200,000, 5 = over \$200,000). The same controls are included in the equations estimating reenrollment intention and switching intention.

The system of equations is specified as follows:

(A1.1) Overall satisfaction
$=\alpha_0+\alpha_1$Satisfaction with internet access$+\alpha_2$Satisfaction with academics and learning $+\alpha_3$Satisfaction with teachers$+\alpha_4$Satisfaction with school environment and facilities $+\boldsymbol{\Theta_1'}\mathbf{Controls}+\varepsilon$,

(A1.2) Reenrollment intention
$=\beta_0+\beta_1$Overall satisfaction$+\boldsymbol{\Theta_2'}\mathbf{Controls}+\zeta$, and

(A1.3) Switching intention
$=\gamma_0+\gamma_1$Overall satisfaction$+\boldsymbol{\Theta_3'}\mathbf{Controls}+\epsilon$.

I estimated Equations A1.1–A1.3 using a seemingly unrelated regression since the Breausch-Pagan test of independence rejected the null hypothesis such that the error terms are not independent of each other ($\chi^2(3) = 323.9$, $p < .01$).

A.1.2. Results

The results are presented in the table below. Specifically, parents' satisfaction with school internet access is positively associated with overall satisfaction ($\beta = .199$, $p < .01$), which in turn is a) positively associated with reenrollment intention into the same public school ($\beta = .863$, $p < .01$) and b) negatively associated with switching intention to private schools ($\beta = -.290$, $p < .01$). These results show that school internet access is a key attribute of value used in customers' overall evaluation. Further, internet access strengthens school-parent relationship through an increased reenrollment intention as well as a decreased switching intention.

Associations among Satisfaction with School Internet Access, Overall Satisfaction, Reenrollment Intention, and Switching Intention

	Overall Satisfaction		Reenrollment Intention		Switching Intention	
	Coef.	SE	Coef.	SE	Coef.	SE
Satisfaction with internet access	.199***	.012				
Satisfaction with academics and learning	.196***	.013				
Satisfaction with teachers	.285***	.013				
Satisfaction with school environment and facilities	.302***	.014				
Overall satisfaction			.863***	.012	-.290***	.023
Involvement	-.027	.017	-.012	.023	.206***	.042
6th - 8th	-.066***	.025	-.092***	.032	.008	.060
9th - 12th	-.054**	.022	-.033	.028	-.117**	.053
Other	.040	.168	.015	.220	-.159	.409
Associates/Undergraduate degree	-.055**	.022	.051*	.029	.027	.053
Masters/Graduate Degree	.031	.030	-.084**	.039	.353***	.073
Ph.D., MD, JD	-.020	.058	-.094	.076	.648***	.141
Asian	-.034	.051	.065	.067	-.152	.124
Caucasian / White	-.112***	.031	.154***	.041	-.449***	.075
Hispanic/Latino	-.080*	.042	.070	.056	-.083	.103
Native American / American Indian	-.200**	.086	.289**	.113	-.456**	.210
Other	.053	.079	.011	.104	-.195	.192
Income	.028***	.010	.035***	.013	-.051**	.024
Intercept	.217***	.071	.466***	.089	3.832***	.166
N	3,924		3,924		3,924	

* $p < .10$, ** $p < .05$, *** $p < .01$

A.2. Procedures to Clean and Concatenate the Data

I constructed the data by concatenating information from a variety of sources. I collected administrative data from the Universal Service Administrative Company (USAC) on every E-rate subsidy application in the category of internet access between 2000 and 2014, school performance data from Academic Excellence Indicator System (AEIS) and Texas Academic Performance Reports (TAPR), school discipline data from Texas Education Agency (TEA), supplemental school reference information from TEA and National Center for Education Statistics (NCES), county-level population data from the National Cancer Institute's Surveillance Epidemiology and End Results (Cancer-SEER) program, county-level unemployment rates from the Bureau of Labor Statistics Local Area Unemployment Statistics, and county-level median household income from U.S. Census Small Area Income and Poverty Estimates (SAIPE). For analyses regarding the heterogeneity, I collected school internet download speed from USAC open platform, and the number of broadband service providers from the FCC Form 477.

A.2.1. School district internet access spending

I obtain Schools and Libraries E-rate Program data in Texas spanning from 2000 to 2014, including various information related to applications (i.e. applicant type, service category, total funds requested), funding commitments (i.e. total funds granted and subsidy rate), and disbursement (i.e. authorized amount for disbursement by USAC).[1]

As I focus on the impact of internet access spending on public school performance, I remove all applications by libraries, private schools, and multidistrict consortia, while retaining all applications in the service category of internet access. In addition, as both individual schools and school districts can apply for subsidy, I aggregate the application data to the school district level and use the total fund requested as the focal explanatory variable of interest. I end up with 12,961 district-year observations representing 1,166 Texas public and charter school districts that had applied Internet access subsidy spanning funding year 2000-2014.

A.2.2. School academic performance

Step 1. Collect school-level data from AEIS for school year 2003-04 until 2011-12, and from TAPR for school years after school year 2012-13.[2]

Step 2. Clean data by year (from school year 2000-2001 to 2014-2015)

Step 3. Focus on TAKS performance metrics

The focal period of analysis spans from 2000 to 2014, and the Texas assessment system changed from TAAS to TAKS since 2003 and then to STAAR since 2012. Thus, I focus on TAKS performance indicators as it spans a longer time horizon. Details about TAKS assessment in are documented Appendix E.

Step 4. Construct the final sample of Texas Public Schools

a. Append all annual data from AEIS (2000-01-2011-12) and TAPR (2012-13-2014-15)

The sample includes 122,393 school-year observations that represent 10,424 Texas public schools. In total, I collect the data of 10 academic performance indicators, including six TAKS commended performance indicators (mathematics, reading/ELA, writing, science, and social studies, all subjects combined), SAT/ACT meet criterion rate, AP/IB meet criterion rate, advanced course completion rate, and RHSP/DAP graduates.

b. Merge with school location data

I contact Texas Education Agency and obtain the full 1996-2016 school directory that contains zip code of all Texas public schools. In addition, I obtain the latitude & longitude data from NCES and merge with the dataset. Finally, I have an unbalanced sample of 122,377 observations that represent 10,422 schools.

c. Merge with school type information (obtained from TEA and NCES)

There are anomalies in the data (e.g., very small enrollment count, higher than 100% disciplinary rate), which come from schools that are not regular instructional campus. Thus, I separately obtain the school type information from TEA to indicate whether a school is an instructional campus, alternative instructional unit, juvenile justice alternative program, DAEP campus, or budgeted campus, and information from NCES to indicate whether a school is a regular education provider.

Until now, I have a merged sample of 122377 school-year observations representing 10422 unique Texas schools. After removing anomalies in the data (negative values of school finance and student teacher ratio), I have a final sample of 122,048 observations representing 10,418 unique Texas schools. In the sample, a school is an instructional campus (N=108,363), alternative instructional unit (N=8,736), juvenile justice alternative education program (JJAEP, N=2,423), disciplinary alternative education programs campus (DAEP, N=2,515), or budgeted campus (N=11). My main analysis focuses on instructional campus and alternative instructional units as the goals of JJAEP and DAEP are distinct. For instance, JJAEP intends to reduce delinquency, increase offender accountability and rehabilitate offenders through a comprehensive, coordinated community-based juvenile probation system.

A.2.3. School discipline data

I collect school-level discipline data from TEA spanning from 2000 to 2014, including the counts of disciplinary actions by action reason code. As TEA uses -999 to mask all counts that range from 1 to 4 due to the federal Family Educational Rights and Privacy Act (FERPA), I replaced -999 with 2.
http://ritter.tea.state.tx.us/peims/standards/1314/c165.html

a. Merge school district internet access spending, school academic performance, and school disciplinary problems

I have a matched sample (academic performance) of 9,266 schools that belong to 1,165 school districts over 2000-2014, and a matched sample (disciplinary problems) of 8,771 schools that belong to 1,139 school districts over 2000-2014.

b. Handle mask rules

AEIS and TAPR employ masking rules (the use of special symbols to conceal the performance results) in order to comply with the federal Family Educational Rights and Privacy Act (FERPA). Different rules apply for different performance indicators, and I document the details in Appendix E. Accordingly, I exclude these values in my main analyses, resulting in different sample sizes across all performance outcome variables.

A.2.4. School internet connection speed

I use E-rate Request for Discount on Services: Connectivity Information (FCC Form 471 and Related Information from (https://opendata.usac.org/), which contains school-level internet download and upload speed since 2016. When the data is available for multiple years (e.g., 2016, 2017, 2018), I retain the data in the earliest year (i.e. 2016). After matching school names with those in my main sample, I match 6986 schools with their academic outcomes, internet access spending, and internet download and upload speed.

A.2.5. Household internet access

Step 1. Obtain the number of broadband service providers at *zip code level* between 1999 and 2008[3]

I first retained all Texas zip codes (1758 zip codes, 15433 observations), and then replaced missing values with 2 (Kolko 2012; Chan, Ghose, Seamans 2016). Note that zip codes with an asterisk (*) indicate that there are one to three holding companies reporting service to at least one customer in the zip code.

Step 2. Obtain the number of broadband service providers at *census tract level* from 2009 to 2013[4]

- I replaced "1" values of variable "Providers of Fixed Connections over 200 kbps in at least one direction" with "2" as "1" denotes 1 to 3 providers in the dataset
- I matched census tract with zip code using 2000 tract-zip relationship files for year 2009 and 2010 and census tract with zip code using 2010 tract-zip relationship files for year 2011-2013
- I aggregated the number broadband service providers to zip level by averaging across census tract.
- I end up with 1864 zip codes in 2009 and 2010, and 1939 zip codes in 2011, 2012, and 2013

Step 3. Append data obtained from Step 1 and Step 2 to construct the number of broadband service providers at zip code level from 1999-2013. I end up with 2000 zip codes and 24978 observations.

Step 4. Aggregate the number broadband service providers to county level by averaging across zip codes from 1999-2013.

[1] http://www.slforms.universalservice.org/DRT/Default.aspx
[2] https://rptsvr1.tea.texas.gov/perfreport/aeis/index.html, https://tea.texas.gov/perfreport/tapr/index.html
[3] https://www.fcc.gov/general/form-477-data-zip-codes-number-high-speed-service-providers
[4] The definition of the service provider is consistent with the dataset used in Step 1 as they both capture number of broadband subscriptions (in-service connections), in contrast to broadband deployment.

A.3. Institutional Details about E-Rate Program

A.3.1. E-rate funding process

Step 1. Competitive bidding
A formal process to identify and request the products and services schools/districts need so that potential service providers can review those requests and submit bids for them. The bidding process is open and fair as 1) all bidders are treated the same; 2) no bidder can have advance knowledge of the project information; 3) all bidders have common information and know what is required of them; 4) with limited exceptions, service providers and potential service providers cannot give gifts to applicants; and 5) the value of free services (e.g., promotional offers) must be deducted from the pre-discount cost of funding.

Step 2. Selection of service provider
When an applicant examines and evaluates the bids received for eligible services, it must select the most cost-effective bid. The price of the eligible products and services must be the primary factor in the evaluation but does not have to be the sole factor. Other relevant evaluation factors may include: prior experience including past performance; personnel qualifications including technical excellence; management capability including schedule compliance; or environmental objectives.

Step 3. Funding request
Applicants file an FCC Form 471 to request discounts on eligible services and equipment for the upcoming funding year. Applicants must include information on the recipients of services and service provider(s); provide detailed descriptions of services including costs and dates of service or equipment; and certify compliance with program rules. Discount is determined by the percentage of students eligible for NSLP in the school district, the urban or rural status of the school district, and service type (category one versus two). After reviewing the FCC Form 471, USAC issues one or more Funding Commitment Decision Letters (FCDLs), setting forth its funding decision for each funding request.

Step 4. Service confirmation
Applicants the applicant certifies the FCC Form 486, Receipt of Service Confirmation Form, which lists the service start date for each separate funding request for which the service provider has begun to deliver services.

Step 5. Invoicing
After USAC has processed FCC Form 486, applicants or their service provider can begin the process of invoicing USAC for the discount share of the approved eligible services by completing FCC Form 472 or FCC Form 474.

Discount Matrix

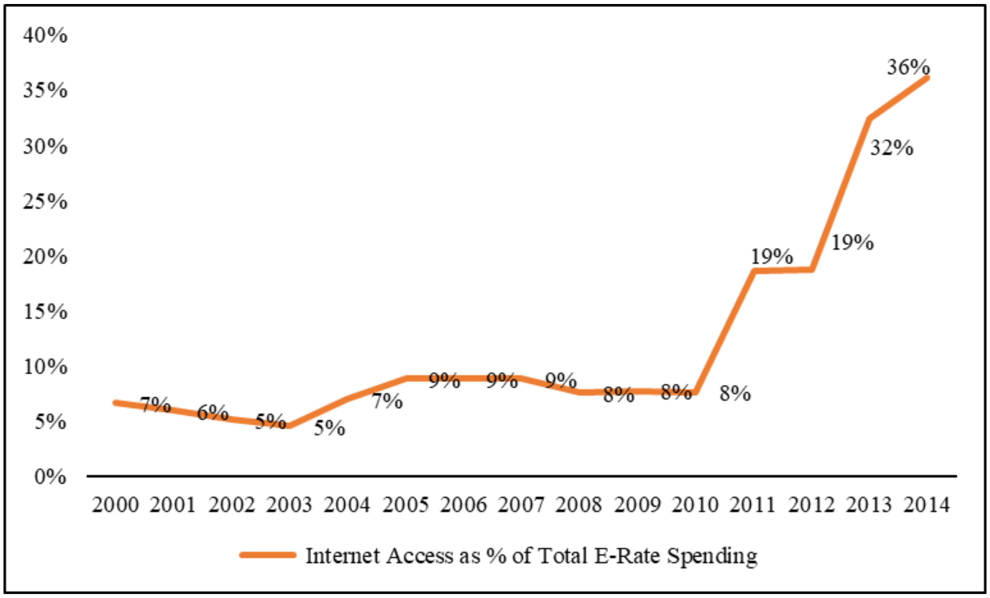
	Discount rate	
% of students eligible for the NSLP	Urban	Rural
<1%	20%	25%
1%-19%	40%	50%
20%-34%	50%	60%
35%-49%	60%	70%
50%-74%	80%	80%
75%-100%	90%	90%

Breakdown of Total Funds Requested by Service Types and over Time (in millions; Texas)

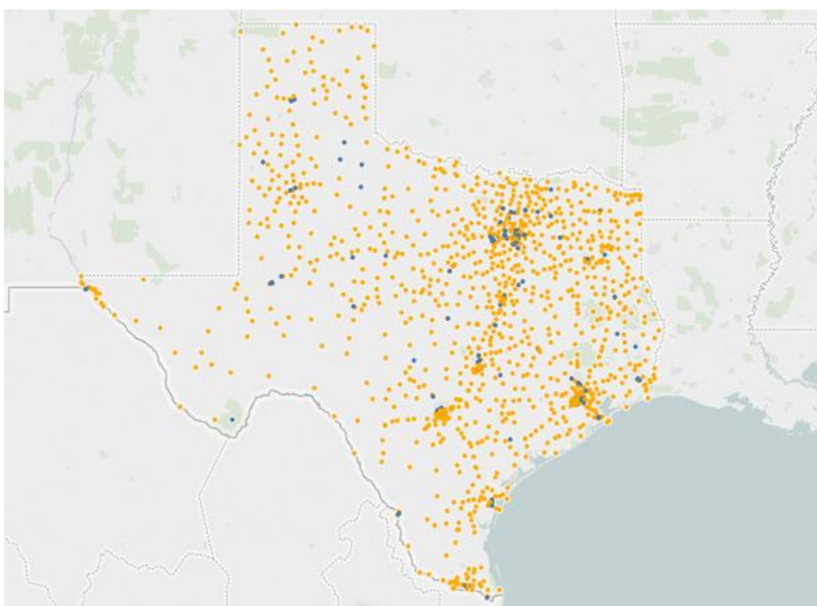
Funding Year	Internal Connections	Managed Internal Broadband Services	Basic Maintenance of Internal Connections	Internet Access	Telecommunication Services	Total
2000	105.2	0.0	0.0	15.5	77.7	198.4
2001	189.4	0.0	0.0	18.3	84.4	292.1
2002	229.1	0.0	0.0	18.4	91.2	338.7
2003	388.5	0.0	0.0	26.9	95.8	511.1
2004	259.6	0.0	0.0	27.6	98.6	385.7
2005	132.8	0.0	14.6	25.2	103.5	276.2
2006	95.9	0.0	17.2	23.5	110.2	246.8
2007	135.9	0.0	17.1	26.1	116.5	295.6
2008	190.3	0.0	21.2	30.3	123.2	365.0
2009	240.5	0.0	25.8	35.3	132.4	433.9
2010	353.0	0.0	25.8	45.3	139.7	563.8
2011	81.9	0.0	15.7	53.6	138.7	290.0
2012	113.1	0.0	14.8	65.8	149.6	343.4
2013	0.0	0.0	0.0	66.5	142.9	209.4
2014	0.0	0.0	0.0	79.5	146.0	225.6
Total	2515.3	0.0	152.3	557.8	1750.3	4975.7
Service Type Proportion	51%	0%	3%	11%	35%	100%

Note. Total funds requested represent the spending figures.

Growth in School District Internet Access Spending
(as % of Total E-Rate Spending)



Geographical Distribution of Applicants versus Non-applicants



Notes. Blue dots denote non-applicants, and orange dots denote applicants.

A.4. Tables and Figures about Texas School Performance

TAKS Assessment by Grade and by Course

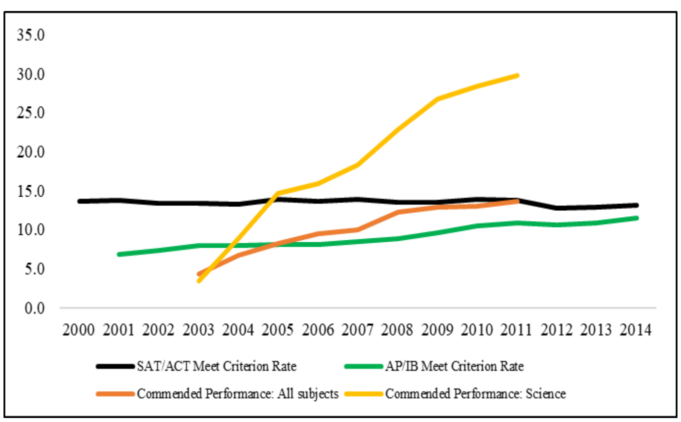
	Mathematics	Reading/ELA	Writing	Science	Social studies	All subjects
Grade 3	√	√				√
Grade 4	√	√	√			√
Grade 5	√	√		√		√
Grade 6	√	√				√
Grade 7	√	√	√			√
Grade 8	√	√		√ (added since 0506)	√	√
Grade 9	√	√				√
Grade 10	√	√		√	√	√
Grade 11	√	√		√	√	√
Sample size of commended performance metric (all grades combined)	N=62,704	N=62,777	N=44,877	N=46,802	N=24,288	N=61,566

Notes. The blanks imply that students in a given grade do not take the exam of that subject (e.g., grade 3 students do not take writing exam).

Mask Rules of Performance Indicators

Examples	Numerator	Denominator	Percent	What is shown in the data	Meaning
a	1	4	25.00%	-1	Denominator is less than five (excluding 0).
	0	2	0.00%	-1	
	3	3	100.00%	-1	
b	0	0	–	•	Denominator is 0.
c	> 0	0	–	-2	Denominator is 0.
d	–	–	–	•	Data reporting is not applicable.
e	8	6	133.00%	-2	Denominator is five or more; and, Percentages are statistically improbable or were reported outside a reasonable range.
	35	30	117.00%	-2	
f	0	5	0%	-3	Denominator is five or more; and, Percent is 0 or rounds to 0.
	1	201	0%	-3	
g	24	24	100%	-4	Denominator is five or more; and, Percent is 100 or rounds to 100.
	995	1000	100%	-4	
	199	200	100%	-4	

Temporal Variation in Academic Performance Indicators (in percentage-point)



A.5. Classification of School Disciplinary Problems

Part I offense-related disciplinary problems	
2	Conduct Punishable as A Felony – TEC §37.006(a)(2)(A)
9	Based on Conduct Occurring off Campus and While The Student Is Not In Attendance At A School-Sponsored Or School-Related Activity For Felony Offenses In Title 5, Penal Code – TEC §37.006(c), TEC §37.007(b)(4), and TEC §37.0081
10	Based On Conduct Occurring Off Campus And While The Student Is Not In Attendance At A School-Sponsored Or School-Related Activity For Felony Offenses Not In Title 5, Penal Code – TEC §37.006(d) and TEC §37.007(b)(4)
16	Arson – TEC §37.007(a)(2)(B)
17	Murder, Capital Murder, Criminal Attempt To Commit Murder, Or Capital Murder – TEC §37.007(a)(2)(C)
19	Aggravated Kidnapping – TEC §37.007(a)(2)(E)
22	Criminal Mischief – TEC §37.007(f)
26	Terroristic Threat – TEC §37.006(a)(1) or §37.007(b)
29	Aggravated Assault Under Penal Code §22.02 Against a school district employee or volunteer – TEC §37.007(d)
30	Aggravated Assault Under Penal Code §22.02 Against someone other than a school district employee or volunteer TEC §37.007 (a)(2)(A)
31	Sexual Assault Under Penal Code §22.011 Or Aggravated Sexual Assault Under Penal Code §22.021
32	Sexual Assault Under Penal Code §22.011 Or Aggravated Sexual Assault Under Penal Code §22.021
36	Felony Controlled Substance Violation – TEC §37.007(a)(3)
37	Felony Alcohol Violation – TEC §37.007(a)(3)
46	Aggravated Robbery – TEC §37.007(a)(2)(F), TEC §37.006(C)-(D) (HB 9680)
47	Manslaughter – TEC §37.007(a)(2)(G)
48	Criminally Negligent Homicide – TEC §37.007(a)(2)(H)
49	Engages in Deadly Conduct – TEC §37.007(b)(3)
55	Student Is Required to Register as A Sex Offender Under Chapter 62 Of the Code of Criminal Procedure and Is Under Court Supervision - TEC §37.304.
56	Student Is Required to Register as A Sex Offender Under Chapter 62 Of the Code of Criminal Procedure and Is Not Under Court Supervision - TEC §37.305.
57	Continuous Sexual Abuse of Young Child or Children Under Penal Code §21.02
Part II offense-related disciplinary problems	
Type I	**Illegal Usage, Possession, or Exchange**
4	Possessed, Sold, Used, Or Was Under the Influence of Marijuana or Other Controlled Substance – TEC §§37.006(a)(2)(C) and 37.007(b)
5	Possessed, Sold, Used, Or Was Under the Influence of An Alcoholic Beverage – TEC §§37.006(a)(2)(D) and 37.007(b)
6	Abuse of A Volatile Chemical – TEC §37.006(a)(2)(E)
7	Public Lewdness or Indecent Exposure – TEC §37.006(a)(2)(F)
11	Used, Exhibited, Or Possessed A Firearm
12	Used, Exhibited, Or Possessed an Illegal Knife
13	Used, Exhibited, Or Possessed A Club – TEC §37.007(a)(1)(C)
14	Used, Exhibited, Or Possessed A Prohibited Weapon Under Penal Code §46.05 – TEC §37.007(a)(1)(D)
33	Possessed, Purchased, Used, or Accepted a Tobacco Product As defined in the Health and Safety Code, Section 3.01, Chapter 161.252
35	False Alarm/False Report – TEC §§37.006(a)(1) and 37.007(b)
50	Used, Exhibited, Or Possessed A Non-Illegal Knife as Defined by Student Code Of Conduct And As Allowed Under TEC 37.007.
58	Breach of Computer Security – TEC §37.007(a)(5) (HB1224)
Type II	**Physical Violence**
8	Retaliation Against School Employee – TEC §§37.006(b) and 37.007(d)
18	Indecency with A Child – TEC §37.007(a)(2)(D)
27	Assault Under Penal Code §22.01(a)(1)
28	Assault Under Penal Code §22.01(a)(1)
34	School-Related Gang Violence
41	Fighting/Mutual Combat – Excludes all offenses under Penal Code §22.01
General Expulsion	
1	Permanent Removal by A Teacher from Class
23	Emergency Placement/Expulsion – TEC §37.019
59	Serious Misbehavior, as defined by TEC §37.007(c), while expelled to/placed in a Disciplinary Alternative Education Program (DAEP)-
Unclassified reason codes	
21	Violation of Student Code of Conduct Not Included Under TEC §§37.002(b), 37.006, or 37.007 (not include violations covered in reason codes 33 and 34)
42	Truancy (failure to attend school) – Parent contributing to truancy – TEC §25.093(a)
43	Truancy (failure to attend school) – Student with at least 3 unexcused absences – TEC §25.094
44	Truancy (failure to attend school) – Student with 10 unexcused absences – TEC §25.094
45	Truancy (failure to attend school) – Student failure to enroll in school – TEC §25.085

A.6. List of Auxiliary Tables

First Stage Results

Second-stage DV	**SAT/ACT**		**AP/IB**		**Advanced course completion**		**RHSP/DAP graduates**	
First-stage DV: Internet Access Spending	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Reimbursement	1.1564***	.0066	1.2253***	.0092	1.1300***	.0070	1.0572***	.0084
N	14646		10366		19411		16556	
County-level controls	Yes		Yes		Yes		Yes	
School-level controls	Yes		Yes		Yes		Yes	
Year fixed effects	Yes		Yes		Yes		Yes	
School fixed effects	Yes		Yes		Yes		Yes	
First stage F statistics	175.8		132.9		161.8		126.1	

Second-stage DV	**CP all**		**CP science**		**CP math**		**CP reading**		**CP social studies**		**CP writing**	
First-stage DV: Internet Access Spending	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Reimbursement	1.2623***	.0020	1.2872***	.0026	1.2617***	.0019	1.2606***	.0019	1.2649***	.0032	1.2582***	.0022
County-level controls	Yes		Yes		Yes		Yes		Yes		Yes	
School-level controls	Yes		Yes		Yes		Yes		Yes		Yes	
Year fixed effects	Yes		Yes		Yes		Yes		Yes		Yes	
School fixed effects	Yes		Yes		Yes		Yes		Yes		Yes	
First stage F statistics	625.9		489.7		659.8		661.4		391.6		559.6	

Second-stage DV	**Disciplinary Problems**	
Control Function First-stage DV: Internet Access Spending	Coef.	Std. Err.
Reimbursement	1.1129***	.0057
County-level controls	Yes	
School-level controls	Yes	
Year fixed effects	Yes	
School fixed effects	Yes	

Intercepts omitted from the results * $p < .1$, ** $p < .05$, *** $p < .01$

Proportion of Funds Committed and Reimbursed by USAC in Texas

	Internet Access	
Year	**% committed**	**% reimbursed**
2000	65%	29%
2001	74%	45%
2002	69%	46%
2003	68%	52%
2004	71%	58%
2005	73%	59%
2006	73%	60%
2007	73%	60%
2008	73%	60%
2009	73%	62%
2010	71%	57%
2011	72%	60%
2012	76%	57%
2013	75%	64%
2014	76%	64%

Relationship among Internet Access Spending, Household Internet Access, and Socioeconomic Characteristics

	Coef.	Std. Err.
Household internet access	-420.66	757.31
Enrollment	13.24**	6.51
Instructional spending	0.70	0.75
Household income	-659.64	785.53
Population	-71034.89	113163.90
Unemployment rate	1191.49	1642.93
2001	3198.36	6920.77
2002	-1081.16	3979.35
2003	6862.63**	3295.51
2004	10176.00***	3777.67
2005	6375.02*	3321.16
2006	4903.27	5550.55
2007	10391.28	7531.70
2008	14703.35	9087.87
2009	16397.09**	7961.03
2010	24642.98***	8126.95
2011	32511.14***	9160.71
2012	43886.96***	13604.25
2013	45566.77***	13592.93
2014	58477.00***	17569.12
Intercept	2544.34	39869.66
School district fixed effects	Yes	
Clustered standard errors	Yes	
N	12,754	

* $p < .1$, ** $p < .05$, *** $p < .01$

APPENDIX B

APPENDICES FOR ESSAY 2

B.1. Study Design

B.1.1. Eligibility criteria

The hospital started with a pool of 199,202 patients assessed for eligibility. First, the study included patients who have documented or suspected cirrhosis with at least one outpatient clinic visit in the year before the field experiment.[35] Second, the study excluded patients with significant comorbid conditions with life expectancy less than 1 year or those with Child C cirrhosis (who are ineligible candidates for a liver transplant). Third, the study excluded patients who have known HCC or a suspicious appearing mass on imaging within six months prior to the eligibility assessment (as they require further diagnostic testing instead of routine screening). Lastly, the study excluded patients a) whose primary language is not English or Spanish[36], and b) those who have no address and/or phone number available on file. As a result, 3,217 patients were eligible.

[35] To identify patients with documented cirrhosis, I used the international classification of diseases codes for cirrhosis or cirrhosis-related complications. Patients with liver disease and an aspartate aminotransferase to platelet ratio index equal to 1.0 or above were suspected to have cirrhosis.

[36] There are six patients whose primary language is neither English nor Spanish. They were not excluded by the hospital since they interacted in English during the outreach interventions. In our analyses, I coded them as a separate category (i.e., Other).

B.1.2. Mailed outreach invitation letter

«pat_first_name» «pat_last_name»
«addr1» «addr2»
«city», TX «zip»

Dear «Eng_prefix». «pat_first_name» «pat_last_name»,

There are two things your doctors at our Health & Hospital System want you to know:

1) Your risk of getting liver cancer goes up if you have liver disease.

2) Screening tests can find many liver problems early, so problems can be treated before they get serious.

That's why we're inviting you to have a liver ultrasound and a simple blood test–two of the screening tests for liver cancer. Research has shown that these tests can have big benefits. If everyone with liver disease did these tests every 6 months, there would be many fewer deaths from liver cancer.

A liver ultrasound is a quick and painless procedure. During a liver ultrasound, you lie on a table while the doctor puts some gel on the skin above your stomach. Then the doctor touches your skin with a wand to look at images of your liver on a computer screen. If the test is normal, you repeat the test in 6 months.

Our Financial Assistance (previously Health Plus) and private insurance usually covers part, if not all, of the cost of a liver ultrasound. You may want to check with our Financial Services to see what your coverage will be.

Please call «projnum» to make an appointment for your liver ultrasound.

B.1.3. Key differences in phone scripts between outreach alone and outreach with patient navigation

The following question was asked to the patient: "People have different concerns when it comes to getting tested for cancer. Would you mind sharing the reason or reasons why you do not want to have a liver ultrasound?" In the outreach-alone intervention, the research staff said, "Thank you for speaking with me today. We will remove you from our contact list. Have a good «dayevening»." In contrast, in the outreach-with-patient-navigation condition, the research staff responded differently to address barriers to care.

Patient's Answer	Patient Navigator's Response
Prep involved	A liver ultrasound is a quick procedure. The ultrasound usually takes less than 30 minutes and the appointment should take around 1 hour from start to finish.
Pain during the test	You will feel no pain but may be asked to change positions or hold your breath during the test.
Time it takes to do the test	A liver ultrasound is a quick procedure. The ultrasound usually takes less than 30 minutes and the appointment should take around 1 hour from start to finish.
Having to take a day off of work	Liver ultrasounds do not require you to take an entire day off work. With planning, most people are able to take the time to complete a test.
Potential risks	A liver ultrasound is a safe procedure. An ultrasound takes pictures of organs inside the body. It does not use X-rays or other types of radiation.
Clothing during test	Your privacy is very important to us. You will be asked to wear comfortable, loose clothing. The doctor will have you lift up your shirt and possibly slide your pants down to expose only your stomach.
Chance that a test might find a problem	Liver cancer testing is important. It could even save your life. That's why we mailed you an invitation to have a liver ultrasound–one of the screening tests for liver cancer. This screening test can actually find many liver problems early, so problems can be treated before they get serious. If ignored, these problems typically get worse without treatment.
Cost of the test	Our financial assistance and private insurance usually cover part, if not all, of the cost of a liver ultrasound. Do you have our financial assistance or private insurance? [No] Okay. I can give you the number to our financial services so you can ask what your coverage will be. It is «finacialnum». We will make a note to contact you again in the future. If you would like to call us back, you can do so at: «projnum». Have a good «dayevening». [Yes] Okay. We will submit a request. The Radiology department will contact you by phone in about a week to schedule the appointment for your liver ultrasound. The blood test will be scheduled for the same day as your ultrasound. It will be in the main lab after your ultrasound. If you would like to call us back, you can do so at: «projnum». Have a good «dayevening».
Liver cancer screening doesn't apply to me	Your lab results suggest that you have or may have liver disease. This means you are possibly at increased risk for developing liver cancer in the future. Liver cancer often does not cause symptoms until it is in its later stages, so it is seldom found early without screening. That's why we mailed you an invitation to have a liver ultrasound–one of the screening tests for liver cancer. This screening test can find many liver problems early, so problems can be treated before they get serious.
Already had liver ultrasound (project or usual care)	You should have a liver ultrasound every 6 months. We can submit a request now and the Radiology department will contact you to schedule your next liver ultrasound 6 months after your last one.
Other reason	We're inviting you to have a liver ultrasound–one of the screening tests for liver cancer. This screening test can find many liver problems early, so problems can be treated before they get serious. Research has shown that this test can have big benefits. If everyone who may have liver disease did this test every 6 months, there would be many fewer deaths from liver cancer.
Don't want to share	Thank you for speaking with me today. We will remove you from our contact list, but please consider talking to your doctor about liver cancer testing. Don't hesitate to raise any concerns you may have. Have a good «dayevening».

B.2. Randomization Checks

Variable	Study Conditions: No Outreach (Usual Care)	Study Conditions: Outreach Alone	Study Conditions: Outreach with Patient Navigation	Total	Multi-group Mean Test (p-value)
Demographics					
Age	55.1	55.0	55.9	55.3	.26
Gender (%; female=1)	41.7	39.8	40.3	40.6	.80
Ethnicity (%)					
Non-Hispanic Caucasian	30.3	27.5	27.2	28.3	.41
Hispanic	36.2	38.3	39.0	37.8	.57
Non-Hispanic African American	31.0	32.8	32.5	32.1	.77
Other/unknown	2.5	1.3	1.3	1.7	.20
Language (%)					
English	78.8	76.0	76.0	76.9	.40
Spanish	20.7	24.0	23.5	22.7	.33
Other	.5	0	.5	.3	.22
Health Status					
Child-Pugh B (%)	28.0	27.5	29.3	28.3	.77
Charlson Comorbidity Index	2.9	2.7	2.9	2.9	.41
Documented cirrhosis (%)	78.7	79.8	80.3	79.6	.76
Etiology of liver disease (%)					
Hepatitis C	53.3	47.5	52.2	51.0	.10
Hepatitis B	3.5	4.5	2.3	3.4	.12
Alcohol	16.3	19.2	17.3	17.6	.43
Nonalcoholic steatohepatitis	17.3	16.8	15.7	16.6	.73
Other	9.5	12.0	12.5	11.3	.21
Visit History					
Number of prior primary care visits	5.0	5.2	5.3	5.2	.61
Receipt of hepatology care (%)	25.5	25.5	26.2	25.7	.95
Health System Accessibility					
Insurance coverage (%)					
Commercial	3.0	2.8	3.3	3.0	.88
Medicaid	19.3	20.8	21.5	20.6	.64
Medical assistance/charity	40.7	42.8	39.5	41.0	.49
Medicare	24.7	23.7	25.8	24.7	.68
Self-pay	2.0	1.5	2.5	2.0	.47
Unknown	10.3	8.3	7.3	8.7	.17
Proximity to clinics (%)	67.7	66.2	66.3	66.7	.83
Neighborhood Socioeconomic Status					
Educational attainment (%)	33.8	33.4	33.8	33.6	.40
Income ($)	35,375.5	35,021.2	35,274.7	35,223.8	.31
Average commute time (minutes)	27.0	27.3	27.1	27.1	.15
Private health insurance (%)	58.5	58.7	58.7	58.6	.82
Public coverage (%)	28.6	28.7	28.6	28.6	.81
Unemployment rate (%)	4.0	4.0	4.0	4.0	.90
Population (16 years and over)	1,144,378	1,141,705	1,158,066	1,148,050	.72
N	600	600	600	1,800	

B.3. Methodological Details

B.3.1. Need for a suitable methodology to enable targeted outreach interventions

There are several established approaches that accomplish the objective of investigating treatment effect heterogeneity. The first strategy is analysis of variance (ANOVA) with the treatment and manipulated moderators. Researchers could design field experiments that allow for intervention types to vary by different subgroups and infer how intervention effectiveness varies in each case. Here researchers precisely predict the nature of the subgroup effect. The advantage of this approach is that it allows for theory testing by controlling the experimental design *ex-ante*. However, this approach is less useful when the number of contingencies is large (e.g., 10-12 moderators). The number of at-risk patients available at a medical center, or even at multicenter trials, will not offer a large enough sample size to populate such designs with sufficient power. Second, given the practical restrictions, researchers can include only a limited number of subgroups (covariates) in the design, making it difficult to obtain personalized treatment effects. Third, the design of cancer outreach is an emergent field of research where moderators are not clearly *ex-ante*; therefore, such *ex-ante* designs make it difficult to discover strong but unexpected treatment effect heterogeneity.

The second strategy is regression with moderators. Researchers could obtain data from randomized control trials and introduce a set of interaction terms among the outreach interventions and patient characteristics in the form of auxiliary regressions to infer how the average treatment effect varies by contingencies. The advantage of this approach is that it allows for theory testing even with a small number of experimental conditions. In addition, this *ex-post* strategy seeks to discover systematic sources of heterogeneity in outreach interventions, and in some cases even uncover crossover interventions that further show the boundary condition of the intervention. Yet, it lacks the ability to systematically capture complex interactions with a large number of contingencies since the threat of multicollinearity would yield biased estimates because of the unbalanced distribution of observations within each subset of contingencies.

The third strategy is a latent class experiment with auxiliary regressions. The intuition behind this approach is to separately fit a linear model to each of the segments instead of fitting a single linear model to all the data. As a result, one could use this *ex-post* approach to uncover finite (e.g., 3 or 4) segments with different treatment effects by modeling unobserved heterogeneity in patient responses. Hutchinson, Kamakura, and Lynch (2000) used this approach to uncover segments that differ in their response to treatment conditions in an experiment. While this approach overcomes the multicollinearity barrier of the interaction approach, it requires repeated measures to be obtained from each individual (which is not always feasible in field experiments), requires subjective choices on the number of segments, and does not directly provide individual-level treatment effect estimates. Moreover, even when latent segments are discovered they are typically difficult to be quantified and described based on policy-relevant, observable characteristics of patients.

As a fourth strategy, the random coefficient model (RCM) can be used to estimate the treatment effect heterogeneity in the form of personalized regression coefficients. Instead of obtaining an average point-estimate of the coefficient for each explanatory variable, an RCM allows the coefficient to vary for each individual unit in the sample (patient in our case). With this feature, researchers can comment on the presence and sources of personalized treatment effects. However, in RCM, researchers assume functional forms such as linear or logit regression. In addition, the number of parameters needed to be estimated increases with the number of covariates that are allowed to vary, so the accuracy of RCM estimation is largely affected by small sample sizes, as is often the case in social experiments.

As a fifth strategy, causal forests enable nonparametric estimation of patient-level treatment effects with valid asymptotic confidence intervals, while maintaining the *ex-post* discovery advantage of the interaction and segments approaches without the need for a larger number of experimental conditions, restrictions on the number of covariates, concern about spurious treatment heterogeneity due to searching over many subgroups, or repeated measures. In the section C2, I summarize and compare the key aspects of causal forests with those of four standard approaches.

Estimation Approaches of Conditional Average Treatment Effect (CATE)

	ANOVA with manipulated moderators	Regression with moderators	Latent class experiment with auxiliary regression	Random coefficient model	Causal forests
Treatment effect heterogeneity					
Subgroup-based CATEs	Yes	Yes	Yes	Yes	Yes
Segment-based treatment effect	No	No	Yes	Yes	Yes
Personalized treatment effects	No	No	No	Yes	Yes
Able to include covariates?	No	Yes	Yes	Yes	Yes
Parametric assumptions?	Yes	Yes	Yes	Yes	No
Allow for flexible modeling of interactions in high dimensions?	No	No	No	No	Yes

B.3.2. Causal forest estimation

I use double-sample causal trees with honest estimation (see Procedure 1 in Wager and Athey (2018) and p. 548 of Davis and Heller (2017) for the implementation). For each tree, I draw a bootstrap training sample by subsampling without replacement from the full dataset. I further split the bootstrap training sample into two subsamples of equal sample size: one used to build the tree structure, and another used to estimate the treatment effects given the tree structure. To obtain the treatment effects for each patient, I use out-of-bag predictions: i.e., a prediction is made using only the trees that did not use patient i during training (Athey and Wager 2019). I follow Davis and Heller (2017; 2019) and Guo, Sriram, and Manchanda (2019) for choices of hyperparameters: i.e., the fraction of the data used to create each bootstrapped sample (s=.5), the number of covariates used for each split when building the tree with the corresponding bootstrapped sample (1/3 of the number of covariates), the number of trees for each forest (B=4,000), and the minimal number of treated and control observations in each leaf (k=10).

Key Attributes in the causal forest estimation

In the table below, I report patient characteristics that are the most important features that determine the tree split in the causal forest estimation.

Key Attributes Determining Tree Split in the Causal Forest Estimation

Causal forest	Key attributes (ordered by variable importance)
Forest^{12}_{P1}	Age, primary care visits, Charlson index, Child-Pugh B, English, Hepatitis C, Spanish, gender, medical assistance/charity, Medicare
Forest^{12}_{P2}	Age, primary care visits, gender, Charlson index, commute time, Child-Pugh B, Caucasian, medical assistance/charity, African American, proximity to clinics
Forest^{12}_{P3}	Age, Y_{P2}, primary care visits, Charlson index, Hepatitis C, $Y_{P1,}$ proximity to clinics, Medicare, commute time, private health insurance coverage
Forest^{13}_{P1}	Age, Charlson index, primary care visits, Hepatitis C, medical assistance/charity, Child-Pugh B, private health insurance coverage, unemployment rate, gender, income
Forest^{13}_{P2}	Age, primary care visits, Charlson index, gender, African American, medical assistance/charity, Hepatitis C, Medicare, receipt of hepatology care, proximity to clinics, Y_{P1}
Forest^{13}_{P3}	Age, primary care visits, Y_{P2}, Charlson index, Y_{P1}, proximity to clinics, African American, Caucasian, private health insurance coverage, income

B.4. Summary of Relative Effectiveness across Cancer Outreach Interventions

Patient Characteristics	Effectiveness of Intervention	
	Outreach Alone	Outreach with Patient Navigation
Demographics		
Older patients	–	+
Female patients	+	+
Ethnicity		
Hispanic patients	+	+
Non-Hispanic African-American patients	+	+
Other	+	+
Language		
Spanish-speaking patients	+	+
Other	n.s.	n.s.
Health Status		
Patients with Child-Pugh B	–	–
Patients with a higher Charlson Comorbidity Index	–	–
Patients with documented cirrhosis	–	–
Etiology of liver disease		
Hepatitis C (relative to Hepatitis B, alcohol-induced, nonalcoholic steatohepatitis)	+	n.s.
Hepatitis C (relative to Other)	n.s.	–
Visit History		
Patients with a higher number of prior primary care visits	+	+
Patients with a prior hepatology care	+	+
Health System Accessibility		
Patients with the medical assistance/charity coverage	+	+
Patients with a closer proximity to clinics	+	+
Neighborhood Socioeconomic Status		
Educational attainment	n.s.	n.s.
Income	n.s.	+
Average commute time	+	n.s.
Private health insurance coverage	n.s.	n.s.
Public health insurance coverage	+	n.s.
Unemployment rate	n.s.	n.s.
Population	+	+

B.5. Parameters and Observed Values Used in the Calculation

	Period 1			Period 2			Period 3		
	No Outreach	Outreach Alone	Outreach with Patient Navigation	No Outreach	Outreach Alone	Outreach with Patient Navigation	No Outreach	Outreach Alone	Outreach with Patient Navigation
Patient Category									
# of patients who do not visit	450	331	311	435	338	316	451	353	322
# of patients who visit	150	269	289	156	254	273	126	231	260
Total	600	600	600	591	592	589	577	584	582
Outreach Cost									
# of call hours	0	3,477.1	4,039.6	0	2,606.6	3,422.5	0	2,135.3	3,041.1
cost per hour ($)	$15	$15	$15	$15	$15	$15	$15	$15	$15
Screening Cost									
# of ultrasound completed	127	265	286	149	257	285	113	231	266
Unit cost of ultrasound	$143	$143	$143	$143	$143	$143	$143	$143	$143
# of CT/MRI completed	69	72	85	68	80	63	59	60	61
Average unit cost of CT/MRI	$1,020	$1,020	$1,020	$1,020	$1,020	$1,020	$1,020	$1,020	$1,020
Treatment Cost (if visit)**									
Probability of early detection	5%	5%	5%	5%	5%	5%	5%	5%	5%
Average cost of treatment (if early detection)	$74,397	$74,397	$74,397	$74,397	$74,397	$74,397	$74,397	$74,397	$74,397
Opportunity Cost (if no visit)**									
Annual HCC probability	2.9%	2.9%	2.9%	2.9%	2.9%	2.9%	2.9%	2.9%	2.9%
Annual cost of advanced HCC	$41,320	$41,320	$41,320	$41,320	$41,320	$41,320	$41,320	$41,320	$41,320
Health Benefits**									
Quality-adjusted life year (QALY) gain (months)	1.65	1.65	1.65	1.65	1.65	1.65	1.65	1.65	1.65
Financial value per QALY	$50,000	$50,000	$50,000	$50,000	$50,000	$50,000	$50,000	$50,000	$50,000

Notes. **Andersson et al. (2008); Goossens et al. (2017)

B.6. Simulation Procedure

Recall that I have three conditions: baseline condition (condition 1), outreach-alone condition (condition 2), outreach-with-patient-navigation condition (condition 3); three periods: Periods 1, 2, 3; and six causal forests: $\text{Forest}^{12}_{\text{P1}}$, $\text{Forest}^{12}_{\text{P2}}$, $\text{Forest}^{12}_{\text{P3}}$, $\text{Forest}^{13}_{\text{P1}}$, $\text{Forest}^{13}_{\text{P2}}$, $\text{Forest}^{13}_{\text{P3}}$. Note that τ_i ($\hat{\tau}_i$) denotes treatment effect (estimate) of patient i. For illustration, I describe the simulation procedure that determines each patient's most suited intervention in Period 1, and I have repeated the procedure for Periods 2 and 3:

1. I used the causal forest $\text{Forest}^{12}_{\text{P1}}$ to obtain each patient's treatment effect estimate in the sample that includes patients in the baseline (condition 1, sample size=600) and those in the outreach-alone condition (condition 2, sample size=600) in Period 1. For each patient *i* in condition 1 in Period 1, the patient-level treatment effect estimate is $\hat{\tau}^{1}_{i,P1,1\to2}$ (i.e., the difference between the outcome I observe for the patient i in condition 1 and the outcome that would be realized if this patient were in condition 2); for each patient in condition 2 in Period 1, the patient-level treatment effect estimate is $\hat{\tau}^{2}_{i,P1,2\to1}$ (i.e., the difference between the outcome I observe for the patient i in condition 2 and the outcome that would be realized if this patient were in condition 1).

2. Next, I used the causal forest $\text{Forest}^{13}_{\text{P1}}$ to obtain each patient's treatment effect estimate in the sample that includes patients in the baseline (condition 1, sample size=600) and those in the outreach-with-patient-navigation condition (condition 3, sample size=600) in Period 1. For each patient in condition 1, the patient-level treatment effect estimate is $\hat{\tau}^{1}_{i,P1,1\to3}$ (i.e., the difference between the outcome I observe for the patient i in condition 1 and the outcome that would be realized if this patient were in condition 3); for each patient in condition 3, the patient-level treatment effect estimate is $\hat{\tau}^{3}_{i,P1,3\to1}$ (i.e., the difference between the outcome I observe for the patient i in condition 3 and the outcome that would be realized if this patient were in condition 1).

3a. Allocation rules for patients in the baseline condition

Comparison. Based on 1 and 2, for each patient i in condition 1 in Period 1: I compare the patient-level treatment effect estimate of outreach-alone intervention on the untreated patient i $\hat{\tau}^{1}_{i,P1,1\to2}$ (obtained from $\text{Forest}^{12}_{\text{P1}}$) with that of outreach-with-patient-navigation intervention on the untreated patient i $\hat{\tau}^{1}_{i,P1,1\to3}$ (obtained from $\text{Forest}^{13}_{\text{P1}}$). See Figure below for the assignment rule. The statistical significance is evaluated at .05 level.

3b. Allocation rules for patients in the outreach-alone condition

Counterfactual estimate. I used the causal forest $\text{Forest}^{13}_{\text{P1}}$ to simulate what each patient's treatment effect estimate in the outreach-alone condition (condition 2) would be were they to be in the outreach-with-patient-navigation condition (condition 3) in Period 1: $\hat{\tau}^{2}_{i,P1,(2\to3)\to1}$.

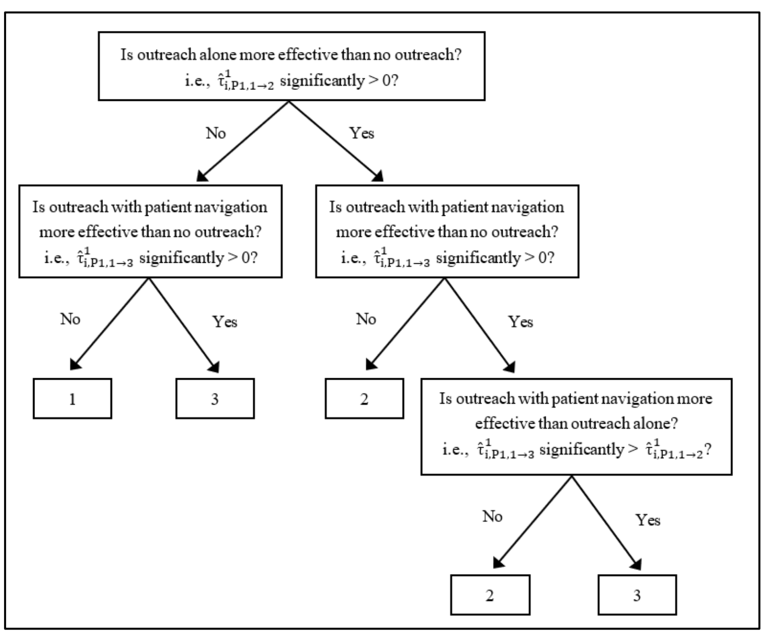
Comparison. For each patient i in condition 2 in Period 1: I compare the patient-level treatment effect estimate of outreach-alone intervention $\hat{\tau}^{2}_{i,P1,2\to1}$ (obtained from Forest^{12}_{P1}) with that of outreach-with-patient-navigation intervention $\hat{\tau}^{2}_{i,P1,(2\to3)\to1}$ (simulated from Forest^{13}_{P1}). See Figure below for the assignment rule. The statistical significance is evaluated at .05 level.

3c. Allocation for patients in the outreach-with-patient-navigation condition

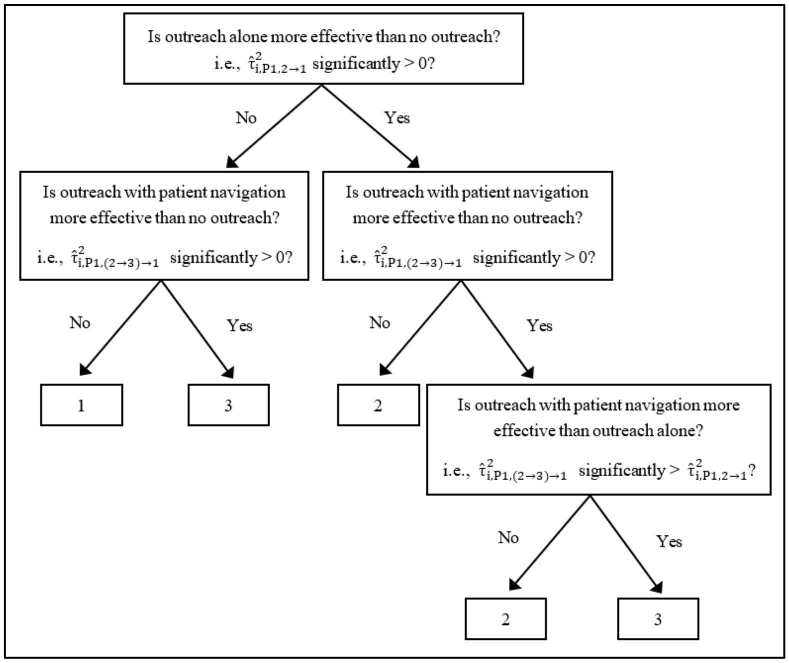
Counterfactual estimate. I used the causal forest Forest^{12}_{P1} to simulate what each patient's treatment effect estimate in the outreach-with-patient-navigation condition (condition 3) would be were they to be in the outreach-alone condition (condition 2) in Period 1: $\hat{\tau}^{3}_{i,P1,(3\to2)\to1}$.

Comparison. For each patient i in condition 3 in Period 1: I compare the patient-level treatment effect estimate of outreach-alone intervention $\hat{\tau}^{3}_{i,P1,3\to1}$ (obtained from Forest^{13}_{P1}) with that of outreach-with-patient-navigation intervention $\hat{\tau}^{3}_{i,P1,(3\to2)\to1}$ (simulated from Forest^{12}_{P1}). See Figure below for the assignment rule for the assignment rule that details which condition (condition 1, 2, or 3) where each patient is assigned. The statistical significance is evaluated at .05 level.

Assignment Rule (Condition 1, Period 1)



Assignment Rule (Condition 2, Period 1)



Assignment Rule (Condition 3, Period 1)