Two Essays on the Impacts of Healthcare Policies

by

Hayoung Cheon

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in the University of Michigan
2024

Doctoral Committee:

Professor S.Sriram, Co-Chair
Assistant Professor Justin Huang, Co-Chair
Professor Puneet Manchanda
Associate Professor Sarah Miller
Assistant Professor Thuy Nguyen

Hayoung Cheon

cheonha@umich.edu

ORCID iD: 0009-0008-9253-6078

DEDICATION

Dedicated to my parents who constantly believed in me.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to two individuals without whom this thesis would not have been possible: Sriram and Justin. Throughout the tumultuous years of my PhD journey, the meetings and conversations with them were instrumental in shaping my perspective and guiding me towards becoming an independent thinker. Their unwavering support and contributions went far beyond the thesis itself; they played a significant role in my personal and career growth.

I also would like to express my gratitude to Puneet. Your insights, feedback and encouragement had a profound impact on both my intellectual development and personal growth.

Thank you all three of you, for being mentors, friends and pillars of strengths during the challenging endeavor.

PREFACE

This dissertation represents the culmination of several years of dedicated research, exploration, and intellectual growth. It has been a journey filled with challenges, learning, and self-discovery, and I am humbled to present the results of this endeavor.

Throughout this process, I have had the privilege of working with an exceptional group of mentors, colleagues, and friends who have provided invaluable guidance, support, and encouragement. Their collective influence has played a pivotal role in shaping the direction and quality of this dissertation.

This dissertation is a testament to the collaborative nature of academic research, and I extend my gratitude to all those who have contributed their insights, time, and expertise to this project.

In closing, I offer this dissertation as a reflection of the dedication, passion, and perseverance that have driven me to this point. It is with humility and gratitude that I present this work to the academic community and to all those who have been a part of my academic journey.

Hayoung Cheon

University of Michigan, Ann Arbor

January 2024

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Chapter 1. Prescription Opioid and Medical Marijuana Legalization

Since the late 1990s, opioids have been increasingly prescribed for pain treatment in the U.S. as a result of aggressive marketing by pharmaceutical companies, resulting in more than 450,000 opioid overdose deaths. In the same time period, several U.S. states have legalized medical marijuana, a drug that can also be used for pain relief. As a result, medical marijuana can be used as a substitute for opioids, potentially leading to a reduction in opioid prescriptions. On the other hand, marijuana use can lead to increased substance abuse, leading to a potential increase in opioid prescriptions. The lack of scientific and medical knowledge along with the uncertain regulatory environment vis-a-vis medical marijuana use also makes it possible that its legalization has no impact on opioid prescriptions. Using eleven years of claims data from a large health insurance company in the U.S., we study the effect of medical marijuana legalization on opioid prescriptions, leveraging the temporal variation in state-wise legalization. We find that, on average, opioid prescriptions decreased after medical marijuana legalization. We also find that the role of providers in reducing opioid prescriptions after legalization is more prominent than their corresponding role in increasing opioid prescriptions.

Chapter 2. Removal of Originating Site Restrictions and Telehealth Utilization: Behavioral Health Patients in Medicaid

Despite the potential and promises of telehealth in improving access to care, particularly for those in areas suffering from lack of providers, the actual utilization of telehealth had been as low as below 1 percent of total healthcare before the Covid pandemic. Among the commonly identified challenges for this dim telehealth use, originating site restriction -

where patients are required to be at the pre-approved locations for getting telehealth care - is often argued as one of major barriers for low telehealth utilization on patient sides.

In the late 2010s, states started to lift this longstanding patient setting restriction within its Medicaid program. After the removal of originating site restrictions, Medicaid patients are allowed to be home or wherever they feel appropriate when getting telehealth cares. We exploit this unique state-level Medicaid policy change to understand how removing this restriction affect telehealth utilization.

In this paper, we aim to answer two research questions. First, what is the effect of removing originating site restrictions on telehealth utilization in Medicaid? Second, how does removing originating site restrictions affect the geographical reach of telehealth providers? Using the newly released Medicaid analytic files (T-MSIS), we build a county-month panel of telehealth utilization and estimate policy effects employing a difference-in-differences model. We find that after removing the originating site restrictions, telehealth utilization increased across the number of claims, billing amount and Medicaid paid amount. Additionally, we find a larger increase in expenditure than in the number of telehealth claims, implying that cost of telehealth per claim became more expensive than before.

Interestingly, our data reveals that most telehealth consultations continued to occur within the same county. Few extended beyond a 100-mile radius, transcended varying levels of urbanity, or crossed state borders. Even with removal of originating site restrictions, the increase in telehealth utilization mostly come from patients and providers in the same county. This suggests that geographical proximity to providers remains a significant factor in telehealth visits, at least for behavioral health patients in Medicaid. Our study thus sheds light on the persistent relevance of physical distance in telehealth, even in a context where regulatory barriers are minimized.

Chapter 1 Prescription Opioid and Medical Marijuana Legalization

1.1 Introduction

Understanding healthcare ecosystem and its many distinguishing features relative to other markets has be- come a rich canvas of novel research opportunities for marketing scholars (Ailawadi et al., 2020). Beginning in the early 1990s, pharmaceutical companies in the U.S. embarked on a marketing campaign to promote opioids as a treatment for chronic pain relief, leading to a significant increase in opioid prescriptions (Department of Health and Human Services 2021). Around the same time, the nation saw a significant increase in opioid addiction and overdose deaths, with the death toll surpassing 450,000 since the late 1990s. In 2018, 125 people die from opioid overdose per day (Hedegaard et al., 2020). Meanwhile, marijuana has been gradually rolling out as a legal option for medical or recreational use across several states (McMichael et al., 2020). The efficacy of marijuana as a painkiller has been documented in medical research (Hill 2015; Cooper et al. 2018; Boehnke et al. 2019). The potential benefits of marijuana as an alternative for pain treatment include lower potential for addiction compared to opioids and significantly lower risk of hospitalization and overdose deaths (National Academies of Sciences, Engineering, and Medicine 2017). As a result, researchers conjecture that marijuana legalization might allow providers and patients to use it as a substitute for opioids, thereby reducing opioid induced harm (Bachhuber et al., 2014; Shi, 2017; Powell et al., 2018). On the other hand, opponents of marijuana legalization worry that marijuana can increase opioid misuse as some studies suggest that it is a “gateway drug” (McCabe et al., 2012; Fiellin et al., 2013). Finally, the lack of scientific and medical knowledge along with the uncertain regulatory environment vis-a-vis medical marijuana prescription and use also makes it possible that it has no meaningful impact on opioid prescriptions. Overall, the jury is still out on whether medical marijuana has an impact on opioid prescriptions, and if so, whether it is positive or negative, both at the aggregate and disaggregate (patient-provider) levels.

In this paper, we attempt to resolve this issue by quantifying the effect of medical marijuana legalization (MML) on opioid prescriptions. In particular, we consider three measures of opioid prescription: number of prescriptions, number of days of supply, and the dosage in morphine milligram equivalent (MME). We do this using a novel data set that contains details on prescription opioid claims over a ten year period. The temporal variation in MML across US states over the 2006 - 2016 period helps us to identify this causal effect using generalized random forests (Athey et al., 2019).

Given that the data are at the provider-patient level, we also investigate the extent to which providers play a role in determining the relationship between MML and opioid prescriptions. This is motivated by the fact that there are conflicting findings about the role of providers in their response to MML. For example, Nussbaum et al. (2011) and Sideris et al. (2018) show that providers are reluctant to substitute away from opioids in favor of marijuana and Kondrad et al. (2018) notes that there is poor communication about medical marijuana use between providers and their patients. On the other hand, providers do recognize that medical marijuana can act as a pain reliever and should be considered a potential therapy that can lead to beneficial outcomes (Cooper et al., 2018; Boehnke et al., 2019; Yuan et al., 2017; Bachhuber et al., 2014; Smart, 2015).

We determine the extent to which providers play a role in the change in post MML opioid prescriptions by examining the uniformity in the change in prescriptions for a given provider's patients. Specifically, if there is high within-provider uniformity (low variation across her patients) in post MML opioid prescriptions, we conclude that the provider is the main driver of the change. In contrast, if there isn't much uniformity, then it is likely that other agents, e.g., patients, are the main drivers. This approach (described in detail later) allows us to obtain the lower bound of the extent to which providers drive the change in opioid prescriptions post MML.

Our results suggest that, on average, opioid prescriptions decreased after MML. This is true for all the three opioid prescription metrics (described above) that we consider. When we consider provider- patient pair level estimates, we find that there is considerable heterogeneity in the treatment effect. While a vast majority of the treatment effects are negative (approximately 70%), a sizable fraction (28% to 32% of provider-patient pairs) exhibited an increase in opioid prescriptions after MML. This showcases one of the benefits of our disaggregate level analysis as the more prevalent negative treatment effects would have swamped the smaller proportion of positive effects if the analysis had been carried out only at the aggregate level. Our results also reveal some important differences between pairs exhibiting positive vs. negative changes in opioid prescription after MML. In particular, we find that patients with negative treatment effects in these pairs tend to be heavier users who received more prescriptions, longer days of supply, and higher potency of medication relative to those who had a positive treatment effect. Similarly, when we consider differences at the provider level, we find that negative treatment effects were for providers with larger number of patients and those writing larger number of opioid prescriptions before legalization. Overall, our results suggest that MML mostly reduced opioid prescriptions with a bigger change coming from providers and patients who prescribed/consumed a higher level of opioids.

In terms of the extent to which providers played a role in the change in opioid prescriptions after MML, we find interesting and different results for provider-patient combination where we see a decrease in opioid prescriptions versus an increase. Specifically, we find that at about 40% of the pair-level variation in the estimated treatment effects can be explained by provider influence when there is a decrease in opioid prescriptions. In contrast, only about 20% of the pair-level variation in the estimated treatment effects can be explained by provider influence when there is an increase. This suggests that providers play a more prominent role in reducing opioid prescriptions after MML relative to an increase. We discuss the implications of these findings for policy makers.

The rest of the paper is organized as follows. We first explain the institutional background vis-a-vis the opioid crisis and medical marijuana legalization in §1.2. We then describe the data and our research methodology in §1.3. §1.4 details the results. We conclude in §1.5.

1.2 Institutional Background

1.2.1 Opioid Crisis

The opioid crisis began as a part of the national effort to address the "under-diagnosis" and "under- treatment" of pain. In 1990, the American Pain Society decried the lack of improvement in pain treatment (Max 1990). In 2001, the Joint Commission[1] introduced the standards for health care organizations in terms of improving pain management. These standards incorporated quantitative measures to manage pain and encouraged the use of opioids (Baker 2017; Hirsch

[1]The Joint Commission is a non- profit organization accrediting US healthcare organizations. Majority of US state governments recognize Joint Commission accreditation as a condition for the receipt of Medicare and Medicaid reimbursement.

2017): “Some clinicians have inaccurate and exaggerated concerns about addiction, tolerance, and risk of death. This attitude prevails despite the fact that there is no evidence that addiction is a significant issue when persons are given opioids for pain control.” (Catan and Perez 2012). The emphasis on the use of opioids was further intensified with the introduction of the new value-based purchasing program by the Centers for Medicare and Medicaid Services (CMS) where a significant part of reimbursements was attached to patient satisfaction including satisfaction with pain control (Rummans et al. 2018). This, along with aggressive marketing from the pharmaceutical companies claiming that opioids are safe pain medications, could be why the medical community started to favor treatment of pain using opioids (Van Zee 2009; Højsted and Sjøgren 2007; Rischitelli and Karbowicz 2002). Opioid dispensing rate per 100 persons increased steadily to the peak of 81.3 in 2012, which is more than enough to give each American adult their own course of treatment (Centers for Disease Control and Prevention 2019). This subsequently led to widespread diversion and misuse of opioids (Chou et al. 2015). In fact, opioids are very addictive with the odds of long-term use increasing markedly after the first five days (Shah et al. 2017). Chronic pain treatment with opioids is particularly risky. Roughly 21-29% patients receiving opioids for chronic pain reported misuse, and around 10% developed an opioid use disorder (National Institute of Drug Abuse 2021). The risks are not restricted to patients; euphoria from opioid intake attracted non- medical use by non-patients as well. 55% of prescription opioid misusers reported that they accessed to the left-over pills of their family members or friends for free (National Survey on Drug Use and Health 2014).[2] Deaths from opioid overdose quickly built up, reaching 500,000 in total from 1999 to 2019. This is close to all deaths - 498,000 - during the American Civil War. In 2017, U.S. Department of Health and

[2] Among the remaining prescription opioid misusers, 25% obtained from doctors, 16% from unlawful purchase and the rest from others sources including theft.

Human Services officially declared the opioid crisis a "public health emergency."

To curb the opioid crisis, policy-makers have implemented several policies. First, federal and state governments added flexibility in prescribing antidotes as an effort to enhance the access to opioid overdose treatment. Naloxone, a drug promptly reversing opioid overdose in emergency, are now dispensable by pharmacists to individuals without prescription in 42 states. In addition, Medication Assisted Treatment (MAT) has been introduced to deal with long-term opioid use disorder, which combines medications with counseling and behavioral therapies. Several states including Maryland and Virginia have increased reimbursement rates for MAT to incentivize counseling in the treatment. In 2016, a federal law was passed to allow nurse practitioners and provider assistants to prescribe buprenorphine, one of MAT drugs, without provider oversight. Second, regulators introduced several campaigns to facilitate safe disposal of left-over pills and to further regulate opioid prescription volume. For example, the Drug Enforcement Administration (DEA), along with its law enforcement partners, have collected nearly 13.7 million pounds of prescription medications, including opioids, since the inception of the National Prescription Drug Take Back Initiative in 2010 (McWilliams 2020). CDC released new guidelines for opioid prescribing for chronic pain patients in 2016, focusing on lowering the dosage (Dowell et al., 2016). Some states implemented a mandatory Prescription Drug Monitoring Program (PDMP) for providers to check patients' substance use history before prescribing. To further prevent "rogue" clinics from dispensing opioids without medical indications, several states placed strict regulations on pain management clinics called "pill mill laws" (Rutkow et al. 2017). The combined impact of all these efforts resulted in a 28% drop in the overall prescription rate between 2012 and 2017. However, the amount of opioids in morphine milligram equivalents prescribed per person is still around 3 times higher than it was in 1999. Overall, the number of opioid overdose deaths still remains high, with over 46,000 in 2018, or 125 deaths per day (Hedegaard et al. 2020).

1.2.2 Marijuana Legalization

Around the same time of the opioid crisis, several states legalized marijuana, initially for medical purposes and subsequently for recreational use (Hollenbeck and Uetake, 2021). As of 2019, 33 states and the District of Columbia have enacted similar laws legalizing medical usage of marijuana comprehensively. Comprehensive medical marijuana programs protect patients from criminal penalties for using medical marijuana and vaporization through home cultivation, dispensaries or some other systems (National Conference of State Legislatures 2021).[3] To be eligible, patients need a licensed doctor's evaluation for medical marijuana use under qualifying medical conditions, of which 65% are associated with chronic pain in 2016 (Boehnke et al., 2019).[4]

Proponents of marijuana legalization advocate that it is an alternative pain medication to prescription opioids. Several medical studies and surveys support this argument. For example, Cooper et al. (2018) report that cannabis, in combination with a lower dosage of opioids, can be as effective as the standard dose of opioids when administered in isolation. Boehnke et al. (2019) showed that medical marijuana use was associated with 64% decrease in opioid use among chronic pain patients. In states with marijuana legalized for medical uses, studies have found lower rate of hospitalization (Yuan et al. 2017) and deaths from opioid overdose (Bachhuber et al. 2014; Smart 2015), suggesting reduced adverse outcome of opioids. Early studies have shown

[3] Comprehensive medical marijuana programs are different from the limited trial programs where only a few restricted uses are allowed. The limited trial programs, commonly called as 'Low-THC and High-CBD' laws, are running in 12 states. These programs allow the use of marijuana mostly for only seizure treatment with restricted variety of products with lower that 5% THC. Only a few medical centers are allowed to produce marijuana.
[4] This is followed by chemotherapy-induced nausea and vomiting, multiple sclerosis and cancer.

that medical marijuana laws decreased state-level annual opioid prescriptions among Medicare and Medicaid enrollees (Bradford and Bradford, 2016, 2017). McMichael et al. (2020) expanded to general population, finding that provider-level annual opioid prescribing reduced by up to 4.2% after medical marijuana laws. On the cost side, Ozluk (2017) found that patient spending on prescription opioids (for current users) decreased by $2.40 a year after MML, mainly driven by young adults.

Opponents, on the other hand, have expressed concerns over the potential harm of marijuana legalization including an increase in opioid abuse. Studies have documented a positive association between prescription opioids and marijuana among chronic pain patients and teenagers (Reisfield et al. 2009; McCabe et al. 2012). Recreational marijuana legalization has also been found to increase web search volume and advertising effectiveness for other addictive substances such as tobacco (Wang et al., 2019).[5] Furthermore, opponents note several hurdles to substituting opioids with medical marijuana for treating pain. First, there is still very limited scientific research on the long-term impact of medical marijuana, so providers might be reluctant to prescribe it (e.g. Nussbaum et al. 2011; Sideris et al. 2018; Rogers et al. 2019). Second, the dosage of medical marijuana is not standardized (Vandrey et al. 2015). Currently, patients need to go through a process of trial and error to find the right strains and dosage for their pain treatment. This is compounded by the fact that medical marijuana products are inconsistent in quality. Third, there is still no clear guidance on the extent to which the insurance plans of patients cover the cost of medical marijuana. Fourth, given that marijuana use is still illegal at the federal level,[6] providers

---

[5] In a similar vein, Bhave and Murthi (2019) find that cigarette sales in states with recreational marijuana legalization increased comparing to sales in states without such laws.

[6] Marijuana is still classified as a Schedule 1 substance under the Controlled Substances Act. Substances classified as such are seen as having no currently accepted medical use and a high potential for abuse by the Drug Enforcement Administration (DEA). Other Schedule 1 substances include heroin, lysergic acid diethylamide (LSD) and methylenedioxymethamphetamine (ecstasy).

and patients may prefer to avoid the social stigma from its use (Satterlund 2016).

In light of these arguments, it is possible that MML could have led to a decrease, increase, or no change in opioid prescriptions, both at the aggregate level and among individual provider-patient pairs. The inconsistent findings vis-a-vis the impact of MML so far (e.g., Reisfield et al. 2009; McCabe et al. 2012; Boehnke et al. 2019) suggest that it is important to go beyond an aggregate level analysis. We do so in this paper via leveraging a decade long, dis-aggregate, comprehensive prescription claims data from a major national health insurance company.

1.2.2.1 Provider Reluctance to Prescribe Marijuana

Multiple provider surveys document a high level of provider reluctance to switch away from opioids in favor of marijuana, even when providers are aware of the potential medicinal effects of medical marijuana (Charuvastra et al., 2005; Kondrad and Reid, 2013). There are at least three reasons for this.

First, providers may be skeptical about the therapeutic benefit of medical marijuana for pain relief, either due to the paucity of scientific evidence (Jensen et al., 2015) and/or the potential stigma from marijuana prescription (Hathaway et al., 2011). For example, some providers feel that "pot-docs" are cheapening the profession by acting as quasi-medical drug dealers who make make money by providing their patient with an easy, accessible high, rather than treating a serious ailment (Thompson Jr and Koenen, 2011). Additionally, this skepticism could also be driven by the lack of formal training in medical schools on this topic (Evanoff et al., 2017).

Second, providers might have not been willing to recommend medical marijuana over opioids, even if they believed in its efficacy. A recent survey shows that while 70% of providers agreed that medical marijuana should be an option for pain relief, they did not plan to register to certify patients under state medical marijuana program (though they were willing to refer their patients to registered providers) (Sideris et al., 2018). The main reason behind this reluctance was that marijuana is illegal at the federal level. Other reasons include lack of standardization on dosage and/or product quality (Vandrey et al., 2015).

Third, providers might lack the decision-making power to substitute medical marijuana for opioids, even when they were willing to recommend it. This is because their employers (medical institutions and/or big hospitals) advise them to be cautious in terms of recommending medical marijuana to their patients. Moreover, as medical marijuana is not covered by insurance while opioids are, providers might be reluctant to recommend a medication that was more expensive for patients.[7] Besides the costs for daily use, medical marijuana patients need to pay an yearly state registration fee (up to $200 depending on state) along with doctor visit fees and spend money on equipment like vape devices (if needed).

Together, these arguments suggest that providers may either be reluctant or lack agency to substitute patients away from opioids to marijuana for treating pain. If this is true, we should expect that any substitution away from opioids was predominantly a result of patient requests rather than initiated by providers.

1.3 Empirical Approach

1.3.1 Data

We obtain outpatient prescription opioid claims data between 2006 and 2016 from a leading private health insurance company in the United States.[8] For each claim, we observe four

[7] Comparing costs for medical marijuana and opioids is difficult as they do not share the common measure of dosage. The cost of medical marijuana also varies across different strains and and state tax rates. Using a back-of-the-envelope calculation, daily use of medical marijuana costs around $4-8 from legal dispensary (40-50mg use per day $*$ price per gram), while daily use of the insured opioids costs around $3-6 (90-200MME use per day $*$ price per MME). Note that "street" prices are higher than legal prices for both prescription opioids and medical marijuana.
[8] Due to the presence of a non-disclosure agreement, we are unable to reveal the name or the exact market share of the insurer.

pieces of information: a) system- encrypted provider unique ID with location at the state level; b) drug name and the unique National Drug Code; c) prescription information, including fill date, quantity dispensed and days of supply; d) patient characteristics, including age, gender, zipcode of residence and their insurance enrollment information. We restrict our sample opiate agonists.[9] These include popular opioid medications for pain treatment such as Codeine, Hydrocodone, Morphine, Oxycodone and Tramadol. We calculate the dosage strength for each opioid prescription in morphine milligram equivalent (MME), and label the drug as a long- or short-acting opioid, using CDC Oral MME Conversion guidelines (2018).[10]

We use data on patients from 19 states (7 with MML and 12 without) who were enrolled in one of the insurer's plans during the entire period between 2006 and 2016. We focus on patients who did not change their state of residence during the period and who only received prescriptions from providers in the same state. This results in 1.4 million opioid prescriptions across 116,116 patients and 70,486 providers with a total of 325,277 unique provider-patient pairs (Table 1-1). In our data, treated states (7 states) accounted for 56.1% of patients, 51.2% of providers, and 53.3% of prescriptions among the 19 states we use during 2006-2016.

Table 1-1 Summary of States during 2006-2016.

Group	No	State	Year of MML	#Pairs[1]	#Pat	#Phy	#Rx
Treated	1	AZ	Nov 2010	89,687	28,156	11,145	5,441,260
	2	DE	Jul 2011	588	281	377	16,324
	3	IL	Jan 2014	14,113	6585	5,931	846,959
	4	MD	Jun 2014	39,610	16521	7,512	1,473,559
	5	MI	Dec 2008	1,721	786	1,395	107,069
	6	MN	May 2014	38,269	12649	9,337	1,145,582
	7	NH	Jul 2013	553	276	410	17,697
Control	1	IA		6,190	2,389	2,484	407,748

[9] These bind tightly to the opioid receptors to produce maximal effects. The technical specification for these can be foundin the American Hospital Formulary Service (classification numbers 280808 and 28080800).

[10] Daily MME = Strength per Unit*(Number of Units/ Days Supply)*MME conversion factor. Long-acting opioids include Methadone and extended-release formulations such as Oxycontin. https://www.cdc.gov/drugoverdose/resources/data.html, accessed July 7, 2020.

2	ID	2,227	791	1,159	120,474
3	IN	4,723	1,918	2,773	244,631
4	KS	6,264	2,572	1,706	264,833
5	MO	37,832	13,720	5,798	2,693,088
6	NC	47,005	14,530	8,976	2,699,871
7	NE	13,086	5,166	2,316	623,735
8	SC	2,540	1,011	1,640	118,012
9	SD	572	249	405	25,419
10	UT	8,298	3,022	2,731	306,709
11	VA	11,528	5,275	4,101	421,918
12	WY	471	219	290	15,824

[1] A unique provider-patient pair.

In Table 1-2, we report the summary statistics of states with and without medical marijuana legalization during 2006-16. Through the 11 years, an average provider treated four patients and a typical patient visited two providers for opioid prescriptions (middle and lower panel in Table 1-2). The composition of patients in the treated and control states are similar in terms of demographic characteristics such as age and gender. However, treated states tend to have more patients living in urban areas. On average, the quarterly number of prescriptions and quarterly days of supply per provider-patient pair are quite comparable across states with and without MML, but the strength of the quarterly dosage (in total MME) appears 30% higher in states with MML than states without (1008.9 vs. 766.3). Conditional on prescribing, prescriptions in treated states are higher for long-acting opioids as well as for oxycodone, codeine and morphine compared to those in control states.

Table 1-2 Pair-, Provider-, Patient-level Characteristics during 2006-2016

	control		treated		
variable	mean	sd	mean	sd	t-stat
pair-quarter:					
Total Rx	0.58	1.35	0.53	1.37	81.78
Total Days	13.70	34.02	13.42	36.01	21.78
Total MME	766.37	3616.32	1008.97	7606.26	-108.41
pair-quarter $\mid Rx > 0$:					
%long acting (vs. short)	0.09	0.25	0.12	0.28	142.69

%hydrocodone	0.43	0.47	0.37	0.46	-163.82
%oxycodone	0.15	0.34	0.25	0.41	316.06
%tramadol	0.25	0.42	0.17	0.36	-264.15
%codeine	0.03	0.17	0.06	0.23	179.67
%morphine	0.02	0.13	0.05	0.20	224.95
%others	0.08	0.26	0.09	0.25	-50.315
patient:					
Nprovider	2.76	2.58	2.82	2.82	-3.79
Age (in 2020)	64.57	19.70	65.31	19.60	-6.30
Female	0.54	0.49	0.56	0.49	-7.17
Rurality[1]	1.69	1.80	1.42	1.42	28.79
provider:					
Npatient	4.09	7.92	5.11	9.06	-15.89

Notes: Comparisons of 7 treated states and 12 control states. Rurality ranges from 1 to 10, with 1 representing the highest score of urbanity, following 2010 Rural-Urban Commuting Area Codes (RUCA) codes.

1.3.2 Research Design

Recall that our objective is three-fold: (i) to understand whether MML led to substitution away from opioids, (ii) to document heterogeneity in treatment effects as a function of provider and patient characteristics, and (iii) if MML indeed led to substitution away from opioids, evaluate the extent to which providers were the driving force behind this change.

We answer the first research question using a difference-in-differences (DiD) approach where we compare the change in opioid prescriptions in states that legalized marijuana (i.e., prescriptions before legalization versus after) with the corresponding changes in control states that did not legalize marijuana. However, in order to address (ii) and (iii), we need more granular estimates of the effect of marijuana legalization on opioid prescriptions, ideally at the provider-patient level. To this end, we use generalized random forest (Athey et al., 2019). With the pair-wise treatment effect estimates, we can quantify the uniformity in the change in prescriptions across patients within a provider as a proxy for provider influence (more details in §4.3).

To estimate the pair-wise treatment effects, we use a DiD research design that exploits the temporal variation in medical marijuana legalization (MML) across states. Specifically, we compare the change in opioid prescriptions for provider-patient pairs before and after their states legalized medical marijuana, relative to the change for similar provider-patient pairs from states without MML. Out of 17 states and DC which legalized medical marijuana[11] during 2006-2016, 6 states (RI, NM, AR, OH, PA, ND) are excluded because of too short pre- or post-time periods, and 3 states (CT, MA, NJ) and DC are excluded due to too few observations. We further exclude NY to ensure that provider (and patient) behavior in both treated and control states is not impacted by the effects of two policies that went into effect around the same time as MML.[12] Our final sample consists of 7 states with MML[12] and 12 states without during 2006-2016 (Table 1-12 provides details at the state level).

We estimate the treatment effects separately for each treated state as they underwent MML at different time. Specifically, we compare each of the seven treated states with a group of 12 control states, resulting in seven DiD comparisons. For consistency, we keep the length of pre- and post-treatment time periods the same: two years before and two years after MML. We further restrict the sample to those with non- zero prescriptions pre-treatment, following the finding in Ozluk (2017). Table 1-3 compares the change in three outcome variables (number of prescriptions including refills, total days of supply, and daily strength of the prescriptions (MME)) for the treated and the control states two years before and after treatment.

---

[11] Note that 10 states (HI, ME, AK, CA, CO, OR, WA, NV, MT, VT) had already adopted medical marijuana legalization before 2006 and are excluded from our analysis.

[12] These are the comprehensive Prescription Drug Monitoring Program (PDMP) and Pain Clinic laws (the "Pill Mill" law). PDMPs attempt to prevent risky prescribing of opioids by providing prescribers timely information about prescribing opioids and patients' behavior e.g., focusing attention on providers prescribing a controlled substance for the first time. The Pain Clinic law requires that pain clinics register with the state and restricts their ownership to providers. It also explicitly controls the prescription of opioids by setting (upper) limits on quantities that can be prescribed and/or dispensed. There is some evidence that both proved somewhat effective in controlling the over-prescription of opioids (Rutkow et al. 2015; Rutkow et al. 2017)

These comparisons suggest that both the treated and the control states experienced a reduction in opioid prescriptions post MML. However, the decline is generally greater among the treated states than the control states, suggesting that MML played an effective role.

Admittedly, a simple comparison across the average prescription outcomes in the control and the treated states, as is reported in Table 1-3, is less than ideal. For example, there is sign of violation of the parallel trend assumption for DiD design (Table 1-4). Without adjusting for the comparability across the treated and the control, we cannot confidently use the outcomes observed in the control states as the reliable counterfactuals for the treated. We address the comparability issues via Generalized Random Forest in the next section.

Table 1-3 Before-After Changes in Pair-Level Prescription Behavior.

	Rx			Days			MME		
Treated State	C	T	T.stat	C	T	T.stat	C	T	T.stat
AZ	-0.15	-0.15	0.96	-5.25	-5.60	2.97	-294.26	-412.50	6.01
DE	-0.14	-0.15	2.74	-5.47	-5.82	2.95	-293.68	-393.16	5.78
IL	-0.17	-0.17	0.50	-6.39	-6.50	0.93	-306.59	-395.28	5.80
MD	-0.17	-0.17	1.68	-6.43	-6.57	1.18	-307.67	-394.22	5.76
MI	-0.16	-0.17	2.54	-4.51	-4.88	3.34	-265.78	-414.32	4.78
MN	-0.17	-0.17	1.68	-6.43	-6.57	1.18	-307.67	-394.22	5.76
NH	-0.17	-0.16	-1.08	-6.28	-6.43	1.21	-303.40	-401.40	5.99

Notes: This table reports the first differences (post-pre) in pair-level quarterly prescription behavior for treated and control states. T-stat is reported for comparisons across the first differences among treated and control states. Comparisons are restricted to provider-patient pairs with non-zero prescriptions pre-legalization, and to a time frame of +/-2 year around legalization.

Table 1-4 Pre-treatment trends

	Rx		Days		MME	
state	T	C	T	C	T	C
MI	27.3%	-7.0%	30.4%	2.7%	18.4%	0.0%
AZ	14.1%	17.3%	20.2%	20.3%	-13.7%	16.2%
DE	-24.3%	19.0%	-48.9%	23.7%	-90.9%	7.8%
NH	40.0%	7.3%	53.1%	16.2%	11.6%	12.1%
IL	3.2%	14.5%	-2.4%	19.7%	-13.1%	16.0%

MD	4.4%	6.7%	12.8%	11.6%	2.6%	10.8%
MN	12.3%	6.7%	15.9%	11.6%	19.2%	10.8%

Notes: The percentage changes in outcomes between the last and the first quarters in the 2-year window are reported across treated and control groups. Each comparison restricts to 2 years before and after focal state's legalization, and to pairs with non-zero opioid prescriptions pre-legalization.

1.3.3 Generalized Random Forest

Recall that our goal is to estimate the effect of MML on opioid prescriptions at the provider-patient pair level, reliably adjusting for the pre-treatment differences between the treated and control samples. We illustrate how we do this by looking at the case of one treatment state and the quarterly number of prescriptions as the outcome. Extending this to the case of multiple treatment states and other outcomes is relatively straightforward.

For a set of *i.i.d* provider-patient pairs $i = 1, ..., n.$, we observe a vector of $d = 166$ covariates (features) $X_i \in R^d$, a response metric $Y_i \in R$, and a treatment assignment $W_i \in \{0, 1\}$. In our case, the treatment is the state level MML and our time horizon is eight quarters before and after the MML. We take Y_i as the difference between the post-legalization and pre-legalization average quarterly prescriptions for a provider- patient pair. W_i is the indicator for whether the provider-patient pair is from the states with MML. Our X_i contain features at four levels: (a) the provider-patient pair level, (b) the provider-patient pair-quarter level, (c) the patient level and (d) the provider level. Table 1-5 lists the features at each level. As can be seen from the table, we have 166 matching features for each forest.

We infer the causal effect of MML on opioid prescriptions across provider-patient pairs with varying features by estimating the conditional average treatment effect (CATE):

$$\tau(x) = E[Y_i^{W=1} - Y_i^{W=0} \,|X_i = x]$$

under the unconfoundedness assumption:

$$W_i \perp (Y_i^0, Y_i^1) \mid X_i$$

We can estimate the CATE by considering the nearby observations in the characteristics space as if these observations come from a randomized experiment. That is, we group provider-patient pairs, in a data-driven way, by how similar they are to each other in terms of their observable characteristics X_i, with the only difference being that a subset of them are from the treated states. We then compare the outcomes across the control and treated pairs in the same group now that the differences across their pre-treatment behavior are adjusted. This gives us the treatment effect estimate conditional on the vector of features describing the matched group. In practice, we estimate CATE using the Generalized Random Forest algorithm (Athey et al., 2019). This algorithm generates robust, consistent and asymptotically normal individual estimates for heterogeneous treatment effects when the covariate space is fairly large. Its non-parametric nature avoids ex-ante specification heterogeneous subgroups, reducing the risk of finding spurious heterogeneity (e.g., Cook et al. (2004)).

We provide the intuition for the algorithm here (for application details, we refer the reader to Guo et al., 2021). In a nutshell, the algorithm bootstraps B subsamples of provider-patient pairs containing both the control and the treated units. Each subsample is then divided into two halves. One half of the subsample is used to grow a causal tree that recursively partitions the provider-patient pairs into heterogeneous subgroups (leaves) using the covariates that provide the most discrimination across the treated and the control, subject to the constraint that at least one treated and one control unit is included in the subgroup. Units in the same leaf are considered to be homogeneous. Given the partitioning structure, the algorithm then uses the remaining half of the subsample to obtain the average treatment effect estimate within each leaf.[13] Finally, the leaf-wise estimates across all *B* trees are weighted to produce the final CATE estimate as below:

$$\hat{\tau}(x) = \frac{\sum_{i=1}^{n} \alpha_i(x)(Y_i - \hat{m}^{(-i)}(X_i))(W_i - \hat{e}^{(-i)}(X_i))}{\sum_i^n \alpha_i(x)\left(W_i - \hat{e}^{(-i)}(X_i)\right)^2}$$

where $\alpha_i(x)$ is a data-adaptive weight that captures how similar the *i*th provider-patient pair is to another pair with characteristics x[14]. $\hat{m}^{(-i)}(X_i)$ is the leave-one-out estimator for the conditional expected outcome variable $m(x) = E[Y_i \mid X_i = x]$. $\hat{e}^{(-i)}(X_i)$is the leave-one-out estimator for the treatment propensity $e(x) = E[W_i \mid X_i = x]$. Recentering Y_i and W_i using mˆ (−i)(Xi) and eˆ(−i)(Xi) improves the consistency of τˆ(x) estimates in observational data (Robinson, 1988; Chernozhukov et al., 2018; Athey and Wager, 2019).

We further obtain the uncertainty measure around the CATE as

$$\hat{V}(x) = \frac{n-1}{n}\left(\frac{n}{n-s}\right)^2 \sum_{i=1}^{j} Cov[\,\hat{\tau}_b(x), N_{ib}]^2$$

where $N_{ib} \in \{0, 1\}$ indicates whether pair *i* is used for the b-th tree, $\hat{\tau}_b(x)$ is the treatment effect estimate from the b-th tree, and $\frac{n-1}{n}\left(\frac{n}{n-s}\right)^2$ is a finite-sample correction for forests grown by subsampling without replacement, and the covariance is taken with respect to all B trees in this forest.

We restrict our analysis to pairs with more than one opioid prescription during the pre-legalization period. This allows matching on the pre-legalization trend across the treated and the control units. We log-transform each outcome variable and take the first difference (i.e., between

[13] Wager and Athey (2018) showed that such an honest procedure is essential for producing unbiased CATE estimation.
[14] The similarity score captures how often the i-th provider-patient pair falls in the same leaf as other pairs with the characteristic x across B trees, adjusted by the size of the leaf.

the post and the pre period) before supplying it to the forest.[15] This differences out any time-invariant provider-patient-specific factors that may contribute to the changes in prescriptions. We grow 2,000 trees for each forest (and one forest for each treatment state and each outcome variable) with all nuisance parameters tuned with cross validation.[16]

Table 1-5 Features used in the Generalized Random Forests

Level	Count	Pre-treatment variables
Pair (23)	1	quarterly mean Rx
	1	quarterly mean days per Rx (total Days/total Rx)
	1	quarterly mean MME per Rx (total MME/total Rx)
	1	quarterly mean daily MME per Rx (total MME/total Days)
	1	quarterly mean %long-acting[1]
	6	quarterly mean %opioid types[2]
	10	quarterly mean %diagnosis types[3]
	2	the first and the last quarter of prescribing
Pair-Quarter[4](80)	3	quarterly total (Rx, MME, Days)
	1	quarterly %long-acting
	6	quarterly %opioid types
Patient (27)	1	number of providers visited
	1	quarterly mean Rx
	1	quarterly mean days per Rx
	1	quarterly mean MME per Rx
	1	quarterly mean daily MME per Rx
	1	quarterly mean %long-acting
	6	quarterly mean %opioid types
	10	quarterly mean %diagnosis types[3]
	3	socio-demographics (gender, age at the time of legalization, rurality)
	2	the first and the last quarter of prescribing
Provider (36)	1	number of patients prescribed opioids to
	1	quarterly mean Rx
	1	quarterly mean days per Rx
	1	quarterly mean MME per Rx
	1	quarterly mean daily MME per Rx
	1	quarterly mean %long-acting
	6	quarterly mean %opioid types
	10	quarterly mean %diagnosis types[3]

[15] In practice, we log transform the outcome variable in the form of log(y+1) to keep all zero observations while also allowing us to evaluate the changes in relative terms.

[16] Following the common practice in the computer science literature, we pick the hyper-parameters through cross-validation whenever the computation allows, and follow the practice in Davis and Heller (2017) and Guo et al. (2017) for the rest. These hyper-parameters include the number of trees for each forest (B), the number of covariates and the sample size considered to build a tree (S_b), as well as the minimal number of treated and the control units, k, required in a leaf. While there is no formal guideline for hyper-parameters, we follow the practice in the literature and set B = 2,000 trees per forest, k = [1,5], randomly draw [18,32] of the covariates per tree and randomly draw 50% of the data as S_b per tree in our estimation. Note that the honest estimation will further split the S_b into halves, one for tree building and the other for estimation.

	2	the first and the last quarter of prescribing
	1	specialty related to surgery
	11	specialty[5]
Total	166	

[1] Opioids are either in extended release form (i.e., long-acting) or short release form.
[2] Hydrocodone, oxycodone, tramadol, codeine, morphine and others.
[3] Top 8 categories for opioid prescriptions in ICD 9, the rest diagnoses as *others*, and *unknowns*.
[4] There are 8 quarterly pre-treatment times.
[5] Top 9 types of providers for opioid prescription, the rest as *others*, and *unknown*

1.4 Results

1.4.1 Effects of MML on Opioid Prescriptions

In Table 1-6, we report the average treatment effects on the treated (ATTs) for the quarterly number of prescriptions, days of supply per prescription, and dosage strength in MME all in log term). On average, we find that MML led to a 0.51% reduction in the quarterly opioid prescriptions, 1.9% reduction in the quarterly days supplied, and 5.2% reduction in the quarterly dosage across provider-patient pairs in the treated states.[17] Overall, our results are consistent with the previous findings at the aggregate level (Bradford and Bradford, 2016, 2017; McMichael et al., 2020) that opioid prescriptions declined after marijuana became available for medical use.

Table 1-6 Pair-level Treatment Effect Estimates from Causal Forests.

	logRx	logDays	logMME
Mean ATT	-0.0051	-0.0192	-0.0520
	(0.00008)	(0.0002)	(0.0006)

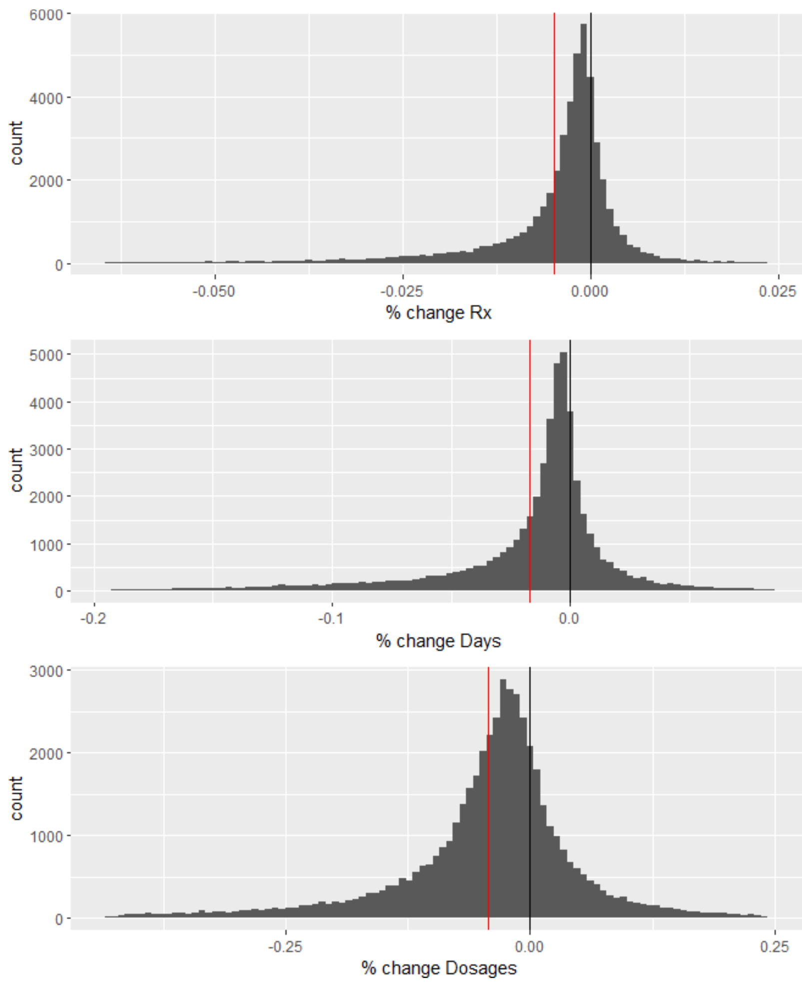
Note. Standard errors, calculated as sd(ATT)/$\sqrt{N}$, are reported in the parentheses.
All three means are significant at 99% level. Npairs (treated) = 39,161.

Next, we investigate the heterogeneity in the treatment effects across provider-patient pairs. We first plot the distribution of the estimated percentage change in the number of quarterly prescriptions, number of days of supply, and dosage strength across provider-patient pairs from

[17] All percentages are computed as 100% * ($e^{\hat{\tau}} - 1$).

the treated group in Figure 1-1. We find considerable heterogeneity in how MML influenced opioid prescriptions across all three outcome metrics. In particular, 74% to 78% of the provider-patient pairs were estimated to reduce prescriptions after MML. Among the sample of pairs with a statistically significant estimate, 75% of them were negative (Table 1-7). In the next section, we take a systematic approach to evaluate the heterogeneity in the estimated pair-level treatment effects across characteristics of providers and patients.

Figure 1-1. Distribution of Percentage Change



Notes. Red vertical line indicates the average (0.1 = 10%). Black vertical line indicates zero. Extreme values are excluded (<1%).

Table 1-7. Percentage of Negative ATTs and Significant ATTs.

	Rx	Days	MME
%pairs with ATT< 0	74.3%	73.8%	78.0%
%pairs with significant ATT	66.6%	53.7%	38.2%
%pairs with ATT< 0, conditional on significance	75.6%	75.1%	75.1%

Notes: Nobs = 39,161. Significance at 95% level.

1.4.2 Treatment Heterogeneity by Provider and Patient Characteristics

We now evaluate provider and patient characteristics that are correlated with the sign of the estimated treatment effects. We start by comparing the average pre-treatment prescription behavior of provider- patient pairs who experienced a negative (as opposed to positive) treatment effect (top panel of Table 1-8). These results suggest that the provider-patient pairs for whom opioid prescriptions went down after MML had, on average, fewer but stronger prescriptions and longer days of supply before MML. This pattern is robust when we compare the differences at the level of the patient or the provider (middle and bottom panel in Table 1-8).

Interestingly, patients who saw a reduction in their opioid prescriptions after MML visited slightly more providers for opioids than patients who saw an increase. Likewise, providers with a decrease in prescribed opioids served more patients than providers with an increase. This suggests that more popular providers seem to lean towards reducing opioid after MML, especially among patients shopping around for opioids. We do not find any significant difference in demographic characteristics such as age, gender, or rurality between patients who decreased and increased opioid prescription after MML.

Another unintended positive consequence of MML is that population groups that reduced their opioid dosage had higher average opioid dosage than their peers that saw their dosage increase. In particular, the results in Table 1-8 suggest that the average daily MME in cases with a negative treatment effect is around 10 MME higher than the corresponding value when among units with a positive treatment effect.[18] Therefore, on the daily MME dimension, riskier population groups (higher average daily MME) saw larger reduction than their less risky (lower average daily MME) counterparts. However, since positive ATT is associated with less risky (lower daily MME) population, the benefit derived among the riskier population groups in terms of reducing opioids after MML could have been partly reversed by the increase in opioid use among the less risky population groups.

Table 1-8 Provider and Patient Characteristics.

	Negative ATT	Positive ATT	t.stat
pair level:			
Nobs	28,013	11,148	
quarterly Rx	4.45	7.65	9.97
days per Rx	11.02	10.43	-4.91
daily MME	51.08	41.26	-15.12
patient level:			
Nobs	17,842	6,709	
quarterly Rx	8.40	8.96	1.18
days per Rx	10.03	9.35	-5.07
daily MME	48.17	40.04	-10.37
age	58.96	59.17	0.82
female	0.58	0.58	-0.30
rurality	1.44	1.45	0.57
Nprovider	1.86	1.62	-13.23
provider level:			
Nobs	11,858	3,573	
quarterly Rx	13.46	14.09	0.56
days per Rx	9.86	9.95	0.53
daily MME	47.50	38.80	-11.52
Npatient	2.64	2.17	-9.22

Notes: This table reports the average pre-legalization characteristics across negative vs. positive mean ATTs on dosage at pair-, patient-, and provider-level. Group mean t-tests are reported in the last column.

[18] The 2016 CDC guideline for prescribing opioids for chronic pain ask clinicians to be very cautious when prescribing a daily dosage higher than 50 MME and to try to avoid daily dosages above 90 MME.

1.4.3 Role of Providers in Driving the Change in Opioid Prescriptions

Prescription of any medication, including that of opioids, is a joint decision made by the patient and (pre- scribing) provider. We examine potential variation in the ATT across all patients of a given provider to assess the extent of the role the provider plays in this decision in our setting. There are two major patterns that we could see in terms of the ATT across patients of a given provider - the ATTs could change uniformly (higher or lower for all or most patients) or non-uniformly. We propose three possible ways that could lead to these two patterns (in ATT across patients for a given provider).

1. If providers believe that marijuana is a safer, but equally effective alternative to treating pain, legalization of marijuana in the state where they operate should have led them to act as champions for substituting away from opioids. On the other hand, if providers have been influenced by heavier marketing activities from opioid companies in response to the increased competition after MML, they may have prescribed more opioids than before.[19] In either case, the outcome will be uniform ATTs across all patients served by a provider.
2. Providers could have selectively substituted away from opioids for some of their patients while in- creasing opioid prescriptions for others based on the match value (based on the patient's condition and their preference for marijuana). In this case, the ATTs would vary a lot across patients served by the same provider.
3. Upon hearing about the potential benefits of marijuana over opioids, patients might have urged their providers to switch them away from opioids. However, if marijuana use leads to opioid misuse (as the gateway drug theory suggests), some patients might have

[19] Large opioid pharmaceutical manufacturers lobbied aggressively against MML (Angell 2018; Frances 2021).

requested their providers for higher dosages of opioids. Similar to (2) above, we should observe non-uniform changes in ATTs within the same provider due to the idiosyncratic requests from patients.

Of these three ways, the first and the second are ones where providers play the primary role. However, the first way is the strongest proxy for provider-initiated (prescription) changes. In other words, the extent to which the change in the ATT for a provider's patient pool is *uniform* allows us to quantify the extent to which providers play a role in the change. Note that the provider could also play a role in the second way (above). However, by restricting ourselves to only the first way, our quantification can be viewed as a lower bound of the role played by providers in inducing the changes in opioid prescriptions in response to MML.

We quantify the influence of the provider by regressing the pair-level treatment effects on provider fixed effects, α_i:

$$ATT_{ij} = \alpha_i + \epsilon_{ij},$$

where the provider fixed effects capture the average change in opioid prescriptions after MML for each provider as described in (1) above and the residuals, ϵ_{ij}, capture the deviation from this average across patients treated by the same provider as in (2) and (3) above. By examining the percentage of variation in ATT_{ij} explained by provider fixed effects (α_i), we can assess the level of influence providers have exerted.[20]

We find that at least 40% of the pair-level variation in the estimated treatment effects can be explained by provider influence (Table 1-9, column 1). Specifically, provider influence can explain for at least 36.5% of the dosage reduction, 44.1% of the reduction in length of supply, and 46.6% of the reduction in number of prescriptions. Note that the extent of provider role

[20] To estimate Eq. 5, we only consider providers who have at least two patients (and hence, two provider-patient pair level treatment effects). These providers account for 47% of all treated providers, 87% of all prescriptions, and 83% of all patients.

when prescriptions increase is less than half of the extent when prescriptions decrease (Table 1-9, column 3 versus column 2). These results provide us a lower bound on the amount of provider influence. Given that the provider role is twice as large for decreases than increases, we conclude that provider influence is primarily exerted towards switching patients away from opioids.

Table 1-9. %Variation in Pair-level ATTs Explained by Provider Fixed Effects

	All Pairs	Negative Providers[a]	Positive Providers
ATT Rx	44.2%	46.6%	21.0%
ATT Days	44.5%	44.1%	16.7%
ATT MME	40.7%	36.5%	17.7%

Sum of squares from provider fixed effects over total variation of pair-level ATTs. [a] Providers whose average treatment effects across her patients is negative. They account for 75% of all providers in the sample.

One concern is that within-provider uniformity in ATT could have been a result of multiple patients requesting similar adjustments to their opioid prescriptions after MML, as opposed to a provider-initiated change. We therefore carry out a few supplementary analyses showing that this is unlikely. First, we focus on the size of a provider's patient pool. As the number of patients treated by the same provider increases, it will become progressively less likely that all these patients make the same request (in terms of the ATT). Therefore, we can be more confident in our assertion (and subsequent analysis) that uniformity in the ATT within a provider's patients reflects the role of the provider as the provider's patient pool gets larger. Table 1-10 shows that, even for providers with nine or more patients, about 30% of the variation is explained by provider fixed effects. In other words, although the conventional wisdom suggests little to no willingness from providers to prescribe marijuana as a substitute for opioid, we find significant evidence that providers do play a considerable role in initiating the drug change.

Second, we examine the percentage variation in ATT explained by provider fixed effects for providers who show a larger versus a smaller ATT (using a median split). If the percentage variation in ATT explained by the provider fixed effects is higher for the former group compared to the latter, it shows that providers play a larger role in effecting bigger reductions in opioid prescription post MML. We find that the percentage of variation explained is 72% for the providers with a bigger reduction compared to 7% for providers with a smaller reduction on Rx, 67% versus 6% on Days and 72% versus 7% on MME.

Table 1-10. % Variation in ATTs Explained by Provider Fixed Effects among Providers with Negative Mean ATTs.

		Average %phyFE		
Npat per phy	Nphy	ATT Rx	ATT Days	ATT MME
2-3	3,405	63.3%	64.9%	58.9%
4-5	1,176	51.7%	45.4%	37.8%
6-8	645	39.3%	36.7%	29.5%
≥9	556	34.4%	34.0%	27.2%

Notes: %Variation explained by provider fixed effects among providers with reduction. Nphy is number of providers in each bin with negative ATT Rx. Nphy is similar across negative ATT Days and negative ATT MME.

Table 1-11. % Variation Explained by Provider FE between Providers with Larger Reduction and Smaller Reduction.

	ATT Rx		ATT Days		ATT MME	
Npat per phy	large	small	large	small	large	small
2-3	57.7%	11.0%	57.5%	7.9%	51.8%	9.9%
4-5	45.2%	6.9%	38.7%	3.3%	33.8%	4.3%
6-8	33.7%	2.5%	28.0%	3.5%	23.5%	3.0%
≥9	30.2%	1.5%	26.7%	1.2%	22.0%	2.4%

Notes: Providers are median-split into large vs. small on mean ATT on the corresponding outcome. Large group has bigger reduction than small group. Each group's % variation of pair level changes explained by provider FE are reported.

Once again, the above analysis may be confounded with the patient pool size for each provider. In Table 1-11, we carry out the same comparison but within the same bin (of number of patients). This also has the additional benefit of controlling for the number of the observations that drive each fixed effect. As can be seen from the table, there is a consistent pattern across the three outcomes where providers with larger reductions saw a higher explained variance in ATT via provider fixed effects. In conclusion, the above analyses show that providers play a considerable role in reducing opioid prescription after MML.

Based on both analyses, we conjecture that concerns about substituting opioid with marijuana or the lack of agency to do so did not deter providers from taking the initiative. We note that there is still scope for encouraging providers to take the lead if reducing opioid prescription is the policy goal.

1.5 Conclusion

Over the last two decades, the misuse of opioids as general pain relief drugs has caused almost half a million deaths in the United States. In this paper, we investigate whether the legalization and availability of another pain relief agent - medical marijuana - has had any impact on opioid prescriptions. Leveraging a large and unique database and the phased legalization of medical marijuana across states, we find evidence of a reduction in opioid prescriptions after MML. This reduction is consistent across three prescription outcomes: the quantity of drugs prescribed, the days of supply and dosage strength. In other words, it appears that medical marijuana is being used as a substitute for opioids. Furthermore, we build upon this main finding by carrying out further analysis at the disaggregate (patient-provider pair) level. This allows us to provide a conservative estimate of the extent to which this reduction in opioid prescriptions is driven by providers. In particular, our results show that providers play a much larger role in terms of influencing a decrease in opioid prescriptions relative to an increase.

For policy makers, these findings suggest that MML has an impact in non-marijuana domains such as opioid prescriptions. If the resulting impact of this spillover i.e., reduction in opioid prescriptions, is desirable, then our results suggest that policy makers could explore the design and deliver of targeted interventions towards providers to accelerate the reduction. These interventions center around addressing the factors that block providers from appropriately adopting/recommending medical marijuana for pain relief. Some specific interventions could be the development of dosage guidelines, offering systematic training in medical schools, and clarifying the legal situation with an effort to minimize liability at the federal level.

Our study has a few limitations. First, based on findings from previous research, we focus our effort only on providers who were already prescribing opioids. Second, our claims data is from a private insurance company, which makes it difficult to generalize the findings to all insurance types. Third, given our data limitations, we can only estimate the lower bound (instead of the total value) of provider influence in driving the change in opioid prescriptions. Fourth, we do not take a stance on whether the reduction of opioids due to MML is desirable as we do not have access to long-term detailed health outcome data. We hope that future research can address these limitations.

1.6 Reference

Ailawadi, K., Chan, T., Manchanda, P., and Sudhir, K. (2020). Introduction to the special issue on marketing science and health. Marketing Science, 39(3):459–464.

Angell, T. (2018). Senator calls out big pharma for opposing legal marijuana. Forbes. (February 23), Ac- cessed July 29, 2021: https://www.forbes.com/sites/tomangell/2018/02/23/senator-calls-out-big-pharma- for-opposing-legal-marijuana/?sh=58a927151bac.

Athey, S., Tibshirani, J., Wager, S., et al. (2019). Generalized random forests. The Annals of Statistics, 47(2):1148–1178.

Athey, S. and Wager, S. (2019). Estimating treatment effects with causal forests: An application. arXiv preprint arXiv:1902.07409.

Bachhuber, M. A., Saloner, B., Cunningham, C. O., and Barry, C. L. (2014). Medical cannabis laws and opi- oid analgesic overdose mortality in the United States, 1999-2010. JAMA Internal Medicine, 174(10):1668– 1673.

Baker, D. W. (2017). History of the joint commission's pain standards: lessons for today's prescription opioid epidemic. JAMA, 317(11):1117–1118.

Bhave, A. and Murthi, B. (2019). A study of the effects of legalization of recreational marijuana on con- sumption of cigarettes. Available at SSRN 3508422.

Boehnke, K. F., Gangopadhyay, S., Clauw, D. J., and Haffajee, R. L. (2019). Qualifying conditions of medical cannabis license holders in the United States. Health Affairs, 38(2):295–302.

Bradford, A. C. and Bradford, W. D. (2016). Medical marijuana laws reduce prescription medication use in medicare part d. Health Affairs, 35(7):1230–1236.

Bradford, A. C. and Bradford, W. D. (2017). Medical marijuana laws may be associated with a decline in the number of prescriptions for medicaid enrollees. Health Affairs, 36(5):945–951.

Catan, T. and Perez, E. (2012). A pain-drug champion has second thoughts. The Wall Street Journal. December 17, 2012. https://www.wsj.com/articles/SB10001424127887324478304578173342657044604.

Centers for Disease Control and Prevention (2019). Opioid prescribing practices. August 13, 2019. Retrieved April 23, 2021, from https://www.cdc.gov/drugoverdose/data/prescribing/prescribing-practices.html.

Charuvastra, A., Friedmann, P. D., and Stein, M. D. (2005). Provider attitudes regarding the prescription of medical marijuana. Journal of Addictive Diseases, 24(3):87–93.

Chernozhukov, V., Chetverikov, D., Demirer, M., et al. (2018). Double/debiased machine learning for treatment and structural parameters. The Econometrics Journal, 21(1):C1–C68.

Chou, R., Turner, J. A., Devine, E. B., Hansen, R. N., Sullivan, S. D., Blazina, I., Dana, T., Bougatsos, C., and Deyo, R. A. (2015). The effectiveness and risks of long-term opioid therapy for chronic pain: a systematic review for a national institutes of health pathways to prevention workshop. Annals of Internal Medicine, 162(4):276–286.

Cook, D. I., Gebski, V. J., and Keech, A. C. (2004). Subgroup analysis in clinical trials. Medical Journal of Australia, 180(6):289.

Cooper, Z. D., Bedi, G., Ramesh, D., Balter, R., Comer, S. D., and Haney, M. (2018). Impact of co- administration of oxycodone and smoked cannabis on analgesia and abuse liability. Neuropsychopharma- cology, 43(10):2046–2055.

Davis, J. and Heller, S. B. (2017). Using causal forests to predict treatment heterogeneity: An application to summer jobs. American Economic Review, 107(5):546–50.

Department of Health and Human Services (2021). What is the U.S. opioid epidemic? February 19, 2021. Accessed April 23, 2021. https://www.hhs.gov/opioids/about-the-epidemic/index.html.

Dowell, D., Haegerich, T. M., and Chou, R. (2016). CDC guideline for prescribing opioids for chronic pain—united states, 2016. JAMA, 315(15):1624–1645.

Evanoff, A. B., Quan, T., Dufault, C., Awad, M., and Bierut, L. J. (2017). Providers-in-training are not prepared to prescribe medical marijuana. Drug and Alcohol Dependence, 180:151–155.

Fiellin, L. E., Tetrault, J. M., Becker, W. C., Fiellin, D. A., and Hoff, R. A. (2013). Previous use of alcohol, cigarettes, and marijuana and subsequent abuse of prescription opioids in young adults. Journal of Adolescent Health, 52(2):158–163.

Frances, A. (2021). Opioid companies lobby against medical marijuana. July 20, 2021. Accessed July 29, 2021: https://rehabs.com/pro-talk/opioid-companies-lobby-against-medical-marijuana/.

Guo, T., Sriram, S., and Manchanda, P. (2021). The effect of information disclosure on industry payments to providers. Journal of Marketing Research, 58(1):115–140.

Hathaway, A. D., Comeau, N. C., and Erickson, P. G. (2011). Cannabis normalization and stigma: Contem- porary practices of moral regulation. Criminology & Criminal Justice, 11(5):451–469.

Hedegaard, H., Minin˜o, A. M., Warner, M., et al. (2020). Drug overdose deaths in the United States, 1999-2018.

Hill, K. P. (2015). Medical marijuana for treatment of chronic pain and other medical and psychiatric problems: a clinical review. JAMA, 313(24):2474–2483.

Hirsch, R. (2017). The opioid epidemic: It's time to place blame where it belongs. Missouri Medicine, 114(2):82.

Højsted, J. and Sjøgren, P. (2007). Addiction to opioids in chronic pain patients: a literature review. European Journal of Pain, 11(5):490–518.

Hollenbeck, B. and Uetake, K. (2021). Taxation and market power in the legal marijuana industry. RAND Journal of Economics.

Jensen, B., Chen, J., Furnish, T., and Wallace, M. (2015). Medical marijuana and chronic pain: a review of basic science and clinical evidence. Current pain and headache reports, 19(10):50.

Kondrad, E. and Reid, A. (2013). Colorado family providers' attitudes toward medical marijuana. The Journal of the American Board of Family Medicine, 26(1):52–60.

Kondrad, E. C., Reed, A. J., Simpson, M. J., and Nease, D. E. (2018). Lack of communication about medical marijuana use between doctors and their patients. The Journal of the American Board of Family Medicine, 31(5):805–808.

Max, M. B. (1990). Improving outcomes of analgesic treatment: is education enough? Annals of Internal Medicine, 113(11):885–889.

McCabe, S. E., West, B. T., Teter, C. J., and Boyd, C. J. (2012). Co-ingestion of prescription opioids and other drugs among high school seniors: Results from a national study. Drug and Alcohol Dependence, 126(1-2):65–70.

McMichael, B. J., Van Horn, R. L., and Viscusi, W. K. (2020). The impact of cannabis access laws on opioid prescribing. Journal of Health Economics, 69:102273.

McWilliams, K. (2020). DEA and partners collect a record amount of unwanted medications during National Prescription Drug Take Back Day. United States Drug Enforcement Administration. October 30, 2020. Retrieved April 23, 2021, from https://www.dea.gov/press-releases/2020/10/30/dea-and-partners-collect- record-amount-unwanted-medications-during.

National Academies of Sciences, Engineering, and Medicine (2017). The health effects of cannabis and cannabinoids: the current state of evidence and recommendations for research.

National Conference of State Legislatures (2021). State Medical Marijuana Laws. April 5, 2021. Retrieved May 12, 2021, from https://www.ncsl.org/research/health/state-medical-marijuana-laws.aspx.

National Institute of Drug Abuse (2021). Opioid Overdose Crisis. March 11, 2021. Retrieved April 23, 2021, from https://www.drugabuse.gov/drug-topics/opioids/opioid-overdose-crisis.

Nussbaum, A. M., Boyer, J. A., and Kondrad, E. C. (2011). "but my doctor recommended pot": Medical marijuana and the patient–provider relationship. Journal of General Internal Medicine, 26(11):1364.

Ozluk, P. (2017). The effects of medical marijuana laws on utilization of prescribed opioids and other prescription drugs. Available at SSRN 3056791.

Powell, D., Pacula, R. L., and Jacobson, M. (2018). Do medical marijuana laws reduce addictions and deaths related to pain killers? Journal of Health Economics, 58:29–42.

Reisfield, G. M., Wasan, A. D., and Jamison, R. N. (2009). The prevalence and significance of cannabis use in patients prescribed chronic opioid therapy: a review of the extant literature. Pain Medicine, 10(8):1434–1441.

Rischitelli, D. G. and Karbowicz, S. H. (2002). Safety and efficacy of controlled-release oxycodone: a systematic literature review. Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, 22(7):898–904.

Robinson, P. M. (1988). Root-n-consistent semiparametric regression. Econometrica: Journal of the Econo- metric Society, pages 931–954.

Rogers, A. H., Bakhshaie, J., Buckner, J. D., Orr, M. F., Paulus, D. J., Ditre, J. W., and Zvolensky, M. J. (2019). Opioid and cannabis co-use among adults with chronic pain: Relations to substance misuse, mental health, and pain experience. Journal of Addiction Medicine, 13(4):287–294.

Rummans, T. A., Burton, M. C., and Dawson, N. L. (2018). How good intentions contributed to bad outcomes: the opioid crisis. In Mayo Clinic Proceedings, volume 93, pages 344–350. Elsevier.

Rutkow, L., Chang, H.-Y., Daubresse, M., Webster, D. W., Stuart, E. A., and Alexander, G. C. (2015). Effect of florida's prescription drug monitoring program and pill mill laws on opioid prescribing and use. JAMA internal medicine, 175(10):1642–1649.

Rutkow, L., Vernick, J. S., and Alexander, G. C. (2017). More states should regulate pain management clinics to promote public health.

Satterlund, K. (2016). Approaches to grammar intervention: A look at current practice. Shah, A., Hayes, C., and Martin, B. (2017). Characteristics of Initial Prescription Episodes and Likelihood of Long-Term Opioid Use — United States, 2006–2015. MMWR Morb Mortal Wkly Rep, 66:265–269.

Shi, Y. (2017). Medical marijuana policies and hospitalizations related to marijuana and opioid pain reliever. Drug and Alcohol Dependence, 173:144–150.

Sideris, A., Khan, F., Boltunova, A., Cuff, G., Gharibo, C., and Doan, L. V. (2018). New york providers' perspectives and knowledge of the state medical marijuana program. Cannabis and Cannabinoid Research, 3(1):74–84.

Smart, R. (2015). The kids aren't alright but older adults are just fine: Effects of medical marijuana market growth on substance use and abuse. Available at SSRN 2574915.

Thompson Jr, J. W. and Koenen, M. A. (2011). Providers as gatekeepers in the use of medical marijuana. The Journal of the American Academy of Psychiatry and the Law, 39(4):460.

Van Zee, A. (2009). The promotion and marketing of oxycontin: commercial triumph, public health tragedy. American Journal of Public Health, 99(2):221–227.

Vandrey, R., Raber, J. C., Raber, M. E., Douglass, B., Miller, C., and Bonn-Miller, M. O. (2015). Cannabi- noid dose and label accuracy in edible medical cannabis products. JAMA, 313(24):2491–2493.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association, 113(523):1228–1242.

Wang, P., Xiong, G., and Yang, J. (2019). Frontiers: Asymmetric effects of recreational cannabis legalization. Marketing Science, 38(6):927–936.

Yuan, J. T., Tello, T. L., Hultman, C., Barker, C. A., Arron, S. T., and Yom, S. S. (2017). Medical marijuana for the treatment of vismodegib-related muscle spasm. JAAD Case Reports, 3(5):438–440.

1.7 Appendix. State Policies around Marijuana Legalization and Opioid Prescribing

Table 1-12. Marijuana legalization and laws related to opioid prescribing in 50 states and D.C.

		Opioid Prescribing		Marijuana Legalization		
	State	PDMP[1]	Pain Clinic	Medical[2]	Recreational	Study Group[3]
1	Rhode Island	2015	x	2006	x	
2	New Mexico	2012	x	2007	x	
3	Michigan	x	x	2008	x	treated
4	D.C.	x	x	2010	2014	
5	New Jersey	2015	x	2010	x	
6	Arizona	x	x	2010	x	treated
7	Delaware	x	x	2011	x	treated
8	Connecticut	2015	x	2012	x	
9	Massachusetts	2016	x	2013	2016	
10	New Hampshire	x	x	2013	x	treated
11	Illinois	x	x	2014	x	treated
12	Minnesota	x	x	2014	x	treated
13	Maryland	x	x	2014	x	treated
14	New York	2013	x	2014	x	
15	Arkansas	x	x	2016	x	

16	Ohio	2015	2011	2016	x	
17	Pennsylvania	2015	x	2016	x	
18	North Dakota	x	x	2016	x	
1	Alabama	x	2013	x	x	
2	Florida	x	2011	x	x	
3	Georgia	x	2013	x	x	
4	Idaho	x	x	x	x	control
5	Indiana	x	x	x	x	control
6	Iowa	x	x	x	x	control
7	Kansas	x	x	x	x	control
8	Kentucky	2012	2012	x	x	
9	Louisiana	x	2005	x	x	
10	Mississippi	x	2011	x	x	
11	Missouri	x	x	x	x	control
12	Nebraska	x	x	x	x	control
13	North Carolina	x	x	x	x	control
14	Oklahoma	2015	x	x	x	
15	South Carolina	x	x	x	x	control
16	South Dakota	x	x	x	x	control
17	Tennessee	2013	2012	x	x	
18	Texas	x	2010	x	x	
19	Utah	x	x	x	x	control
20	Virginia		x	x	x	
21	West Virginia	2013	2012	x	x	
22	Wisconsin	x	2016	x	x	
23	Wyoming	x	x	x	x	control

As of Dec 2016. 10 states with legal medical marijuana before 2006 are excluded.
[1] Comprehensive PDMP where all prescribers are mandated to use at least to all initial pioid prescriptions issued to patients.
[2] Effective date of medical marijuana legalization.
[3] Final comparison group. NJ, CT, MA, MA are excluded due to small observations.
Sources: PDMP (pewtrusts.org), Pain clinic (Rutkow et al., 2017), Marijuana legalization (procon.org)

Chapter 2 Removal of Originating Site Restriction and Telehealth Utilization: Behavioral Healthcare Patients in Medicaid

2.1 Introduction

Telehealth has been touted for its potential in improving access to care, particularly in underserved areas. [21] Past surveys and randomized experiments have reported that patients can benefit from reduced travel time and improved access to care without compromise in quality.[22]

Despite its potential, the overall utilization of telehealth has been marginal before the Covid Pandemic. Research indicates that while telehealth usage had been growing over time, it remained below 1% across various insurance types and states before the pandemic (Mehrotra et al., 2017; Yu et al., 2018). In Medicaid, 0.1% of beneficiaries used telehealth in 2008 in states with telehealth reimbursement (Douglas et al. 2017). In Medicare, 0.25% of beneficiaries were reported to have used telehealth services in 2016 (CMS 2018).

One key obstacle to broader telehealth adoption had been originating site restriction imposed by the Centers for Medicare and Medicaid Services (CMS). This mandate requires patients to be at specific locations while getting telehealth. The type of originating sites had to be approved by CMS and commonly included doctor's office, skilled nursing facility, critical access hospital, rural health clinic or federally qualified health center. This only limited flexibility of

[21] In 1996, the Institute of Medicine defined telehealth as "the use of electronic information and communications technologies to provide and support healthcare when distance separates participants". This concept was echoed in 2012 by Mary Wakefield, then Health Resources and Service Administration (HRSA) Administrator. In the telehealth workshop with the Institute of Medicine, she said, "Telehealth is a key component in ensuring access to healthcare services in isolated geographic areas across the United States. More effective deployment of telehealth technologies will enhance our ability to better meet the healthcare needs in rural and frontier parts of the country. However, telehealth is important not just for rural communities, but for any underserved community." (Lustig 2012).
[22] Interested readers may refer to Butzner and Cuffee (2021), a survey on telehealth research.

telehealth by requiring patients to still travel to health care centers which are sparse already in rural areas.

Recognizing the constraints, states have been progressively lifting patient setting restrictions within their Medicaid programs, given their autonomy to define telehealth payment and policies as long as they meet federal criteria for efficiency, economy and quality of care. After removal of the originating site restriction, Medicaid patients were now allowed to receive telehealth services from their homes or any other suitable location. In essence, the lift of restriction improved access to telehealth on the patient side by reducing travel costs. This shift prompts a few critical questions: Can patient-side access to telehealth alone boost telehealth utilization, especially among underserved populations? Does this new flexibility encourage patients to seek care outside their immediate neighborhood, which has been the popular argument for telehealth benefits in underserved areas?

To address these questions, this paper focuses on the following research objectives: First, we aim to quantify the impact of removing originating site restrictions on telehealth utilization for behavioral healthcare within Medicaid. Second, we aim to investigate how these effects vary across different care types, patient demographics, and urban-rural settings. Lastly, we aim to explore whether this policy change influences the geographical reach of telehealth services. These questions are crucial for understanding the implications of policy changes on telehealth utilization in general and their potential role in enhancing healthcare access and equity.

To answer the questions, we take advantage of the state-year variation in the removal of originating site restrictions as a quasi-experimental setup. Employing a difference-in-differences approach with two-way fixed effects at the county-month level, we analyzed the impact of this policy change. We studied states that eliminated originating site restrictions in 2017, using 2016 as pre-intervention periods and 2018-2019 as post-intervention periods. This setup allows for a straightforward 2 x 2 comparison in our difference-in-differences design.

For our analysis, we utilized the Medicaid Analytic Files (T-MSIS), an upgraded version of the earlier MAX files with updates mostly on comprehensiveness and synchronization. Our process involved gathering all outpatient claims data for the study states, identifying telehealth claims using specific procedural and place of service codes, and constructing a county-month level panel for telehealth utilization. We aimed for utmost transparency in our sample construction and data documentation, hoping that our pioneering use of T-MSIS for telehealth studies can serve as a valuable reference for future research in this area.

Our empirical exercises face two primary challenges. First, there are data quality issues in T-MSIS files, particularly during the early transition year which overlaps with our pre-intervention time period. To address this, we conducted a thorough examination of our sample, ensuring a conservative interpretation of our results. Our validation process includes reliance on the Data Quality Atlas (DQ Atlas), a CMS tool for assessing T-MSIS data completeness, along with manual comparisons using aggregate numbers from public reports. We also sought to replicate a previous study that used Medicaid claims data sourced directly from the state offices, using our T-MSIS dataset. Every step of our data and sample construction process, including extensive data quality validations, is detailed in §2.10. Despite initial concerns, we found that our main policy effect estimates remain robust across various sample definitions.

As with any policy evaluation study, establishing causality presents inherent challenges. In our 2 x 2 difference-in-differences design, it is crucial to validate the parallel trend assumption between the treated and control units. To validate the parallel trend assumption, we employed several standard practices. We first checked event study plots to detect any preliminary issues, which did not reveal any significant concerns. Next, we examined residual plots for treated units during the pre-intervention period based on counterfactual outcome models (i.e., regressions without the policy variable on data excluding policy assignment) following Liu et al. (2022). This process also yielded reassuring results. Finally, following another recent recommendation on pre-trend testing, we reported the slope of linear trends that our sample and power could detect, as suggested by Roth (2022). The minimum detectable slope in our pre-trend tests, given our sample's power at 90%, was found to be 30-50 times smaller compared to the magnitude of our estimated effect sizes, further bolstering our confidence in causal interpretation. Furthermore, we conducted a thorough research on other telehealth policies around our study times and directly controlled for potential confounders in the regression for our robustness checks. Our estimates were not impacted by controlling for other telehealth policies.

Our study reveals a significant increase in telehealth utilization following the policy change that removed originating site restrictions. This increase is observed across various metrics, including the number of claims, billed amounts, and Medicaid payments. A particularly notable finding is the disproportionate rise in expenditures relative to the number of telehealth claims. This suggests that, on average, telehealth claims became more costly post-policy change. This aspect warrants further scrutiny, especially in light of the ongoing discussions about whether to continue waiving the patient setting requirement after the Public Health Emergency, a decision that is becoming increasingly urgent as we are already in the grace year of 2024.

When examining sources of heterogeneity in our data, we found no significant differences across several key variables. In particular, there were no difference in policy effects between urban and rural counties, which was different from competitive aspects found in Zhou et al. (2021) between rural and urban areas among Medicare population. Our finding suggests that for Medicaid behavioral healthcare, improved access to telehealth did not exacerbate existing disparities between rural and urban areas, with regard to telehealth utilization.

Additionally, we investigated whether the policy change facilitated patients in rural areas get care from urban providers. However, the majority of telehealth consultations remained within the same county, seldom crossing different urbanity levels. Our study reveals that physical distance still remains a significant factor in telehealth, highlighting an area that may require further attention to fully understand patient's preference on providers at distance and to harness telehealth's potential in bridging healthcare access gaps.

Our work builds upon previous works on telehealth in Medicaid (Douglas et al. 2017; Talbot et al. 2020; Harju and Neufeld 2021). While these precursor studies provided essential descriptive statistics on telehealth usage, they were limited to documenting the frequency of total telehealth in state level before 2016. Our research extends beyond these works by incorporating more recent data and by honing in on the causal effects of removal of the originating site restrictions within Medicaid.

Mainly, our work contributes to the fast-growing literature examining the effects of telehealth access on care utilization (Dahlstrand 2023; McCullough et al. 2021; Rabideau and Eisenberg 2021; Zeltzer et al. 2023; Zhou et al. 2021). Our work aligns most closely with Zeltzer et al. (2023) and Rabideau and Eisenberg (2022) in that shocks pertain to patients. Zeltzer et al. (2023) explored the impact of providers' propensity of telehealth adoption on patient's care utilization in primary care setting for the half of Israeli population. They debunked overutilization concerns by showing that the increase in utilization were offset by lower episode intensity and overall healthcare costs even decreased slightly. Rabideau and Eisenberg (2022) found no significant effects from patient fee waivers on telehealth spending among the employer-insured in the US. We add to the literature by focusing on Medicaid patients for mental health and substance use disorder, who are most marginalized and expected to benefit the most from having access to care that were almost non-existent before.

We also contribute to the literature on the effectiveness of telehealth in mitigating inequity in healthcare by defying distance. In the context of online health platform market in China, Wang et al. (2021) showed that compared with offline healthcare, online health enables patients to have better access to nonlocal healthcare resources and its effect is greater for underserved patients than for well-served patients. They concluded that the value of online health platform depends heavily on how it is utilized differently across areas. Dahlstrand (2024), using the nationwide conditional random assignment between patients and doctors in digital primary care service in Sweden, showed that optimal reallocation of patients across providers, defying the physical distance, can improve welfare by 20%. However, for our Medicaid population, most of telehealth use was within county and even with improved access, telehealth usage for connecting to providers across counties did not increase.

We speculate that differences lie on the unique nature of our sample: Medicaid behavioral health patients. First, healthcare delivery in Medicaid is governed by stringent regulations that dictate eligibility criteria and pricing structures. For example, outpatient visits charge out-of-pocket cost of $4 dollars for patients under 100% federal poverty rates. And providers are reimbursed with a pre-set rate by federal and state government and should hold medical licenses of state where patients reside. This rigidity does not allow pricing mechanism in the zero-demand setting in Wang et al. (2021) and restricts patient-provider matches within states. Second, behavioral health care typically requires chronic management of disease. Patients might have high preference for nearby providers for chronic disease management, expecting to have in-person exams or visits along the continuum of care. Relatedly, Dahlstrand (2023) noted that her optimal reallocation regardless of distance would be differentially impact patients across rural and urban areas when moving to hybrid cares. She addressed that rural area patients should heavily rely on remote care to continue having good match with providers in distance. Our work provides novel evidence on the impacts of telehealth on Medicaid population for behavioral health care and shows that telehealth did not defy distance in our sample. Furthermore, Medicaid population rather increased telehealth for connecting with providers *not* in distance.

The rest of the chapter is organized as follows. We first explain the institutional background in §2.2. We then describe the research design in §2.3 and our data in §2.4. §2.5 details the results. We conclude in §2.6.

2.2 Setting

2.2.1 Mental Health and Substance Use Disorder in Medicaid

Mental health and substance use disorder (MHSUD), referred to as behavioral health, are the leading causes of disease burden in the U.S. According to an analysis by Kaiser Family Foundation[23], behavioral health accounted for 3,355 disability-adjusted life years (DALYs, a measure of overall disease burden as the number of years lost due to ill health)[24], per 100,000 population in 2015, followed by cancer and tumors (3,131 DALYs) and circulatory diseases (3,065 DALYs).

[23] US specific analysis of 2015 Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) data by the Kaiser Family Foundation. https://www.healthsystemtracker.org/chart-collection/know-burden-disease-u-s. Accessed April 2023.

[24] While mental health disorders have lower death rates than cancers (12 vs. 144 death rates per 100k), with the burden of disease defined broadly to include disability as well as mortality, mental health disorders are adjusted to be higher than any other diseases.

In the U.S., Medicaid is the single largest payer for mental health services and is increasingly playing a larger role in the reimbursement of substance use disorder services.[25] Compared to general population, Medicaid enrollees have a higher portion of severe mental disorders, such as schizophrenia, major depression, and bipolar disorder, which cause substantial functional impairment (Saunders and Rudowitz 2022). Population below federal poverty level (FPL), which is one eligibility condition for Medicaid enrollment, are reported to have high 12-month prevalence was found higher than those at or above 100% FPL (6.8% vs. 3.5%).

The shortage of MHSUD providers and low insurance acceptance rates have been common challenges in getting MHSUD care. However, these challenges are notably acute for Medicaid populations, largely due to two factors: the geographic dispersion of providers and Medicaid's relatively low reimbursement rates. Providers are often concentrated in areas that don't align with where most Medicaid beneficiaries live, exacerbating geographic barriers to care. Andrilla et al. (2018) documented the large variation in the supply of psychiatrists, psychologists, psychiatrics, and psychic nurse practitioners across regions, especially between urban and rural areas. About one-quarter of metropolitan counties lacked these providers, compared with 65% of non-metropolitan counties. There was a more than tenfold difference in the percentage of counties lacking mental health providers between the New England Census Division (6%) and the West North Central Census Division (69%). Additionally, Medicaid's lower reimbursement rates further worsen the access to MHSUD care for Medicaid populations. Zhu et al. (2023) showed that Medicaid reimbursement for mental health services (vs. Medicare) might help illuminate Medicaid participation among psychiatrists. On average, Medicaid paid psychiatrists at 81% of Medicare rates, and most states had a Medicaid-to-Medicare index that

[25] https://www.medicaid.gov/medicaid/benefits/behavioral-health-services/index.html. Accessed Jan 2024.

was less than 1.0 (median = 0.76). Even the rejection rate is high for Medicaid bill, which leads physicians refusing to accept Medicaid patients (Dunn et al. 2023).

With high prevalence rates and significant access barriers as described, Medicaid MHSUD patients stand to benefit considerably from telehealth, potentially enhancing healthcare equity. Furthermore, they offer an unique opportunity to evaluate if telehealth can mitigate these access barriers without altering provider financial incentives (e.g. increase reimbursement rates).

Given the general shortage of MHSUD providers, it's hypothesized that enhanced telehealth usage, by allowing telehealth at home, might result from the reallocation of geographically dispersed providers.

2.2.2 Originating Site Restrictions for Telehealth in Medicaid

It has been required by CMS that originating site restrictions for telehealth should be imposed beginning from 1996, where telehealth was first authorized to be reimbursable as Medicare fee-for-services under the Balanced Budget Act. Specifically, originating site restrictions dictate where patients must be located to receive telehealth services—typically at designated healthcare facilities like hospitals and physician's offices[26]. This restriction, which ensures that telehealth services are provided in the presence of healthcare professionals, acts as a form of quality control but offers little incentive for patients to seek care remotely unless services are unavailable locally.

For decades, originating site restrictions have remained a steadfast component of telehealth coverage. However, recognizing the constraints these restrictions placed on the

[26] Other common eligible sites include federally qualified health center, critical access hospital, rural health center, community health center, school, skilled nursing facility (American Telemedicine Association 2017).

expansion of telehealth, several states began to dismantle this restriction in the mid-2010s, most commonly for behavioral healthcare services delivered via video.[27]

The COVID-19 pandemic precipitated a temporary federal waiver of many telehealth restrictions, including those on originating sites. However, as we transition to a post-pandemic landscape, the permanency of these waivers is under active discussion. Currently, patient setting requirements are slated to remain lifted until December 2024.

During a 2023 US Senate subcommittee hearing, healthcare leaders strongly advocated for the permanent removal of originating site and other geographic restrictions, warning that reinstating these could precipitate a 'fast death of telehealth.'[28] They cautioned that a return to pre-pandemic telehealth utilization rates, where less than 1% of providers and patients engaged with these services, is a real possibility. Underpinning the hesitation to make permanent changes to patient location requirements are concerns about potential spikes in telehealth spending that may not correspond with substantial improvements in health outcomes.[29]

2.2.3 Study States

There are three groups of Medicaid programs[30] regarding home allowance for telehealth: 1) programs with explicit eligibility site restrictions where patient's home is *eligible*, 2) programs with explicit eligibility site restrictions where patient's home is *not eligible*, and 3) programs with

[27] Video telehealth predominated. Other formats are asynchronous telehealth such as store-and-forward, remote patient monitoring (RPM) and e-visits. Live video is distinct from RPM and e-visits: RPM and live video allow patients at home for care, however RPM is for chronic health services requiring steady monitoring, including diabetes, strokes or hypertension, but not on MH/SUD while live video care is mostly for MH/SUD cares. E-visits are asynchronous patient portal messages that require medical decision-making and at least 5 minutes of clinician time over a 7-day period. The types of telehealth are detailed in §2.10.4

[28] https://mhealthintelligence.com/news/stakeholders-urge-senators-to-avert-fast-and-slow-death-of-telehealth"

[29] We had a short exchange with one of the healthcare leaders who voiced out in the Senate subcommittee hearing. The conversation revealed that the blocker for the permanent removal of patient setting restrictions was government's concerns on potential overutilization.

[30] A Medicaid program is in state-level. For example, Michigan has its own program.

no written restrictions on patient's location. Regarding the last group, it could mean either the most flexible type, or a lack of languages in early stages of telehealth.

Since our goal is to understand the effect of improved telehealth access through home allowance, we compare among programs with explicit eligibility site restrictions. Specifically, we compare the first group with the second groups. During our study time between 2016-2019, 25 Medicaid programs kept eligibility site restrictions for telehealth (Alabama, Arizona, Arkansas, Colorado, D.C., Delaware, Georgia, Hawaii, Illinois, Indiana, Maryland, Minnesota, Michigan, Mississippi, Missouri, Nevada, Ohio, South Carolina, South Dakota, Texas, Virginia, Washington, Wyoming, Wisconsin, West Virginia).

Based on the Center for Connected Health Policy (CCHP) bi-annual reports, "State Telehealth Laws and Reimbursement Policies", we tracked changes in eligibility site restrictions, particularly *home allowance* across state Medicaid programs.[31] CCHP reports are restricted to policies around fee-for-service Medicaid telehealth claims and followingly our analysis focuses on FFS outpatient claims as well[32]. Among our 25 Medicaid programs, 11 lifted restrictions sometime between 2016 and 2019, while 14 kept the restrictions during our study time. From treated states, we removed Ohio and Missouri because its restriction was on and off. In 2017, 6 programs lifted restrictions (Colorado, Delaware, Minnesota, Texas, Washington, Wyoming). In 2018, 3 programs lifted restrictions (Maryland, Michigan). In 2019, 2 states further lifted restrictions (Nevada, South Carolina). 12 states and D.C. continued not to include patient's home as eligible sites (Alabama, Arizona, Arkansas, D.C., Georgia, Hawaii, Illinois, Indiana, Mississippi, South Dakota, Virginia, Wisconsin, West Virginia).

[31] We use home allowance and removal of originating site restrictions interchangeably.

[32] For managed care plans, Medicaid programs and healthcare providers contract on capitated payment per patient per month and each contract may include different requirements and obligations that could affect telehealth use. We could not access individual conditions around home allowance for telehealth with managed care providers, so we did not include these claims in our analysis.

We further remove 2 states (HI and MI) due to zero telehealth claims in the sample, 2 states (NV and SC) due to too short post-intervention periods, 6 states (MD, CO, GA, IL, TX, AR) due to significant data quality issues noted in DQ Atlas. Total of 13 states remained, with 4 states removing originating site restriction in 2017 and 9 control states keeping the restriction throughout the study periods.

2.3 Research Design

To estimate causal effects of telehealth access on telehealth utilization, we adopt a standard Difference-in-differences (DiD) design with two-way fixed effects, where treatment intervention happened once and forever for all treated units at the same time, which is the case in our context.

$$Y_{it} = \beta Post_t * Treat_i + F_{it} + \epsilon_{it}$$

Above equation defines our research design mathematically. Our outcome variable, Y_{it} is telehealth utilization in the number of claims or expenditures, for a county *i* and monthly time *t*. If county *i* is in treated states (i.e. states which lifted originating site restrictions) then $Treat_i$ = 1, while in control states (i.e. states which did not lift originating site restrictions) then $Treat_i =$ 0. $Post_t = 1$, if monthly time *t* is post-intervention periods. $Post_t = 0$, if monthly time *t* is pre-intervention periods. F_{it} refers to all unit-time varying variables. ϵ_{it} denotes nonsystematic iid error term.

Our parameter of interest is β, a coefficient for difference in changes in telehealth utilization with home allowance between treated and control units, which is the policy effects. The causal interpretation of β requires parallel trend assumption. This assumption posits that, in the absence of the treatment, the average change over time in the outcome variable would have been the same for both the treatment and control groups. Two groups are assumed to be on the same trajectory regarding the outcome of interest, in the absence of the treatment, with level difference allowed. Using the notation in the equation, there should not exist any F_{it} that affect Y_{it} differentially across time between the treated and control units.

In F_{it}, we include county- and monthly time- fixed effects. First, county level fixed effects control for time-invariant differences across counties, allowing the constant level difference between treated and control states. One state might have high level of telehealth utilization over the other, for example, due to different Medicaid reimbursement models, overall infrastructure for tele-communication or preferences. Also, within a state, one county might be more likely to use telehealth than the other, due to different demographics. Next, monthly time fixed effects control for common shocks that affect all groups over time. These include seasonality in behavioral health care visits and overall trends in telehealth uptake nationwide.

Following the standard approach, we test for parallel trend assumption with pre-trend observations. The event studies concerned us less as there was no pre-trend effects. Despite null effects from pre-trend tests, there still can remain unit-time varying confounders. For example, states might have removed originating site restrictions expecting high utilization for telehealth with the change. Or, states might not have removed originating site restrictions, expecting too high jump in telehealth utilization concerning the Medicaid budgets. Other changes in telehealth environments might have been different between treated and control states, such as state's participation in cross-border licensure compacts and telehealth parity laws on private insurance.

2.4 Data

2.4.1 Medicaid T-MSIS Analytics Files

Our research utilizes data from the Transformed Medicaid Statistical Information System (T-MSIS), a national data repository offering comprehensive insights into beneficiaries, providers, service utilization, managed care, expenditures, and third-party liability within Medicaid and the Children's Health Insurance Program (CHIP)[33]. T-MSIS represents a significant upgrade from its predecessor, MSIS, which was fragmented across states and necessitated intensive data-cleaning and extensive validation. Prior to 2014, the scalability and cross-state comparability of Medicaid analytic files were limited. However, T-MSIS, with its streamlined data-cleaning protocols, the introduction of a unique beneficiary ID for cross-state tracking, enhanced scalability through cloud-based operations, and more frequent updates (from quarterly to monthly), has markedly improved the ease of conducting multi-state studies with scale.

In conjunction with T-MSIS, the Centers for Medicare and Medicaid Services (CMS) offers the Data Quality Atlas (DQ Atlas)[34], a web-based tool designed to assist researchers in assessing the quality of the data. During the early transition years to T-MSIS, participation varied among programs, and data quality was often compromised by high levels of missing or incomplete information[35]. However, since 2016, there has been a notable improvement and stabilization in data quality. Despite these advancements, it is advised by CMS that researchers ought to remain vigilant as data quality can still vary across different sub-samples.

[33] https://www.medicaid.gov/medicaid/data-systems/macbis/transformed-medicaid-statistical-information-system-t-msis/index.html. Accessed Dec 10, 2023.
[34] https://www.medicaid.gov/dq-atlas/. Accessed Dec 10, 2023.
[35] On a different note, TAF data excludes rejected claims, which are a significant aspect of Medicaid billing practices. Recent research by Dunn et al. (2024) highlights that Medicaid physicians lose a substantial portion of revenue due to claim denials and resubmissions, more so than with Medicare or private insurers. If the rejection on telehealth differs across treated and control states around the time of policy shock, this might lead to a bias in the policy's positive impact. We assume in our study that rejection rates did not systematically differ between treated and control states post-policy implementation.

Using the DQ Atlas, we validated the quality of Medicaid enrollment and outpatient claims volume for our study states between 2016-19. All study states in all years showed less than 5% of discrepancy against other benchmarks. Also, we scrutinized the missingness of our critical variables – ICD diagnosis codes, place of service codes, and CPT/HCPCS procedure codes – which are vital for identifying Mental Health and Substance Use Disorder (MHSUD) care claims and telehealth services. Concerningly, six states (MD, CO, GA, IL, TX, AR) demonstrated significant data gaps, with over 90% missingness in total outpatient claims. In our sample, MHSUD services constituted 16% of total outpatient claims, implying that total missingness exceeding 90% would necessarily impact the quality of our MHSUD subsamples by more than 25% [36]. In our main analysis, we excluded these six concerning states.

2.4.2 Data Comparison with Previous Literature

The T-MSIS, a relatively recent addition to the field, has not been extensively explored in empirical studies, particularly in the context of telehealth for the Medicaid population. Recognizing the novelty and potential of this dataset, we aim to provide details about how T-MSIS compares with other data on Medicaid telehealth studies. First, we conducted a review of existing literature on telehealth in the Medicaid to establish a baseline. Next, we replicated the approach used by Harju and Neufeld (2021) to validate the compatibility of the T-MSIS with their Medicaid claims data obtained directly from state offices.

Previous literature utilized a variety of data sources and study samples, posing challenges for direct comparison and validation, as summarized in Table 2-1. Douglas et al. (2017) and Talbot et al. (2020) employed the Medicaid Analytic eXtract (MAX) data, which are analytics

[36] We excluded these six states in the main analysis. Further details regarding data quality in our sample are thoroughly examined in §2.10

files based on the older MSIS, preceding T-MSIS. These studies represent early efforts to identify telehealth claims in Medicaid across multiple states in all outpatient claim types, with their samples drawn from 2008 and 2011, respectively.

The most recent data snapshot is in early 2020, from the CMS report using the T-MSIS as in our case[37]. This report primarily focuses on the surge in telehealth usage following the onset of the COVID-19 pandemic, a period marked by increased demand and supply of telehealth due to social distancing measures and the relaxation of federal and state regulations on telehealth services. Caveat is that the temporal and situational differences between the pre-pandemic and early pandemic periods are significant, limiting the comparability of their report with ours.

Harju and Neufeld (2021) presents the closest parallel to our research. Investigating Medicaid claims data directly sourced from state offices for four states between 2014 and 2017, they focused on mental health care, which offers additional commonality with our study. The similarity in the time frame and thematic focus provides a more direct basis for comparison and validation of our findings.

In our best attempt to validate the use of T-MSIS data for telehealth study before the pandemic, we replicated the study conducted by Harju and Neufeld (2021) for two overlapping states, Minnesota and Wisconsin, during the years 2016-17[38]. Harju and Neufeld (2021) reported the number of telehealth claims for mental health care in Medicaid per 10K population using Medicaid claims data sourced from the state offices.

We closely followed Harju and Neufeld (2021) in identifying mental health claims and telehealth services, based on the lists of codes provided in their supplementary documents. Their approach of identifying tele-mental health services are two-steps. First, telehealth claims were

[37] https://www.medicaid.gov/sites/default/files/2021-05/covid-19-medicaid-data-snapshot.pdf
[38] These are the only overlaps in the sample we have for direct comparison.

identified based on their set of procedural codes. Subsequently, the claims were further filtered by their set of provider taxonomy codes for certain procedures. This way, they removed mental health care services by nonspecialists e.g. family medicine or pediatrics. We reported the stepwise metrics based on the T-MSIS data in Table 2-2. Column (1) presents telehealth claims identified by their set of procedural codes and column (2) details the telehealth claims filtered by their set of provider taxonomies after column (1). Column (3) displays the final telehealth claims per 10,000 population, and column (4) offers a comparison with the numbers reported by Harju and Neufeld (2021). For the replication, we included both inpatient and outpatient following Harju and Neufeld (2021). However, we were unable to detect any inpatient claims associated with a GT telehealth identifier. Since Harju and Neufeld (2021) did not report telehealth claims separately across inpatient and outpatient setting, we cannot know whether it implies concerns on our data quality particularly for inpatient claims and also concerns on validation based on single total count metrics between ours and theirs, or literally no telehealth use case identified with the GT modifier in inpatient setting. Considering the nature of the telehealth (i.e. virtual engagements between providers and patients in distance) and successful replication for later years with better data quality in the T-MSIS, we speculate that telehealth claims identifiable with the GT modifier are predominantly, if not exclusively, from outpatient setting. In our main analysis, we focus on outpatient claims only.

There are several noteworthy observations. Assuming that metrics reported in Harju and Neufeld (2021) reveal the truth, we noticed that data quality issues in the T-MSIS were more prevalent in the earlier years. For instance, data from 2016 exhibited more issues compared to 2017. Despite their data not extending beyond 2017, if we assume that the telehealth counts remained relatively stable into 2018 [39], the data quality in the T-MSIS for that year aligns closely with their findings. This assumption suggests that the reliability of telehealth identification in T-MSIS, especially when employing Harju and Neufeld (2021)'s methodology, has improved over time and stabilized by 2018 for Minnesota (MN) and Wisconsin (WI).

Furthermore, our analysis reveals state-specific variations in data quality. Minnesota encountered more pronounced data quality challenges than Wisconsin in 2016 and 2017. Specifically, the final number of telehealth claims per 10,000 population in Minnesota was only 18% and 59% of what Harju and Neufeld (2021) reported for 2016 and 2017, respectively. In contrast, Wisconsin's figures were closer to metrics reported in Harju and Neufeld (2021), standing at 70% for both years. This disparity might be attributed to differences in the missingness of provider taxonomies between the two states, as indicated by a more significant drop in claims between columns (1) and (2) for Minnesota after matching provider taxonomies.

Our observations significantly highlight the need to consider both state-specific factors and temporal variations when assessing data quality in the T-MSIS. A key challenge for our causal inference analysis arises from the observed divergent trends in data quality between Minnesota and Wisconsin. This disparity suggests that the pace of data quality improvement, both before and after our focal policy intervention, might vary across these states. Such variation introduces a potential source of bias in our difference-in-differences estimates.

In the absence of a consistent, uniform benchmark applicable to all time periods and across all treated and control states, determining the exact direction and magnitude of this potential bias becomes a complex endeavor. Consequently, our findings and estimates should be

[39] In our replication process, we comprehensively examined mental health claims across all financing types. However, the primary focus of our study pertains to mental health claims financed through the fee-for-service (FFS) model. Given this concentration, it's reasonable to assume that the overall count of mental health claims would not exhibit significant fluctuations with the policy shock pertaining to FFS claims

interpreted with a heightened degree of caution. The usual concerns inherent in causal estimates are further complicated here by the varying trends in data quality across states. These discrepancies likely stem from unidentified issues within the source data, adding another layer of complexity to our analysis. In essence, we are dealing with time-varying confounders that are further exacerbated by the factor of unknown data quality. The precise impact of this on the bias of our estimates remains ambiguous and is a matter of concern both before (a priori) and after (posteriori) conducting our analysis. We delve deeper into the implications of these data quality issues and their potential effects on our findings in §2.10.

The replication exercise further showed that our data may not attain enough telehealth claims with high missingness in provider taxonomy variables. In fact, DQ Atlas also confirmed concerning missingness of provider details, including taxonomies, particularly in the earlier years. To circumvent this problem, here we defined MHSUD care visits as those either diagnosed as MHSUD *or* provided by MHSUD specialties[40]. This way, we could ensure stable and not too small sample size across years and states. Furthermore, this definition reflects the choice of Medicaid population better. Medicaid MHSUD patients typically have low access to specialists. Access to psychiatrists has been an ongoing challenge in Medicaid due to overall shortages, geographical maldistribution and also low reimbursement. For example, psychiatrists receive lower Medicaid reimbursement than primary care providers for similar services (Saunders et al., 2023). Thus, mental health care visits by Medicaid population are mostly not from specialists and we include visits from non-specialists in our study sample, which was not fully reflected in Harju and Neufeld (2021).

[40] We opted for using diagnoses rather than procedures unlike Harju and Neufeld (2021) due to easiness in data operation with similar results - the claims identified with MHSUD diagnoses and with MHSUD procedures overlapped each other heavily. Unlike Harju and Neufeld (2021), we did not filter out MHSUD care from non-specialists (e.g. primary care).

2.4.3 Sample

We extracted outpatient claims related to Mental Health and Substance Use Disorders (MHSUD) from 22 study states identified in §2.2.3. The selection criteria for Mental Health (MH) and Substance Use Disorder (SUD) diagnoses were based on specific ICD codes. MH diagnoses included ICD-9 codes 290-302 and 306-319, along with ICD-10 codes F01-F09 and F20-F99. For SUD diagnoses, we used ICD-9 codes 303-305 and ICD-10 codes F10-F19. In addition to these explicit diagnostic codes, our analysis also encompassed claims where the primary diagnosis was not specifically classified as MHSUD, provided that the billing provider was recognized under MHSUD-related taxonomies. This approach accounts for scenarios where patients, perceiving a need for mental health care, seek services from MHSUD specialists. These visits may ultimately result in a determination that the patient does not have an MHSUD diagnosis and does not require MHSUD care. Nonetheless, we included such cases in our sample to capture a broader spectrum of patient care choices influenced by their perceived mental health needs.

We classified a claim as a telehealth service based on the presence of specific procedural codes and their modifiers, as outlined in §2.10.3. This classification encompassed both synchronous services, such as audio and live-video interactions, and asynchronous services, which included patient portals, virtual check-ins, and remote monitoring. Asynchronous services represented a minimal portion of our identified telehealth claims, accounting for only about 0.1%. The overwhelming majority (99.9%) of the identified synchronous telehealth services were marked using the GT modifier in tandem with the place of service code 02. Initially, the GT modifier was predominantly used in the earlier years in our sample. This finding aligns with the results presented in Yeramosu et al. (2019), where a manual review of 300 telehealth billing codes indicated that the GT modifier had a 100% sensitivity and specificity rate for identifying live-interactive telehealth encounters in 2016. However, over time, we observed a shift in billing practices. The introduction and growing popularity of the place of service code 02 gradually overshadowed the use of the GT modifier in our data.

Recall that our policy of interest is the removal of eligibility site restrictions on FFS telehealth claims for MHSUD care. Accordingly, we restricted MHSUD claims to those financed by the fee-for-service (FFS) model. This entail types of claim payments designated as 1, A, or U, corresponding to 1) FFS Medicaid or Medicaid expansion claim, 2) FFS separate CHIP claim, and 3) other FFS claims, respectively.

For our eligible study population, we focused on nondual Medicaid enrollees diagnosed with behavioral health conditions, maintaining continuous enrollment (a minimum of 180 days each year) from 2016 to 2019. This subgroup allows us to abstract away from complexities in multiple payment sources and varied enrolled status within individuals. These sample patients may be enrolled in various types of plans capable of incurring MHSUD claims under fee-for-service models. These include: fee-for-service (FFS) plans, primary care case management (PCCM) plans, and managed care plans in states with behavioral health carve-out policies[41].

After removing claims with invalid patient's address and date of service (0.02% of any MHSUD claims, 0.06% of tele-MHSUD claims), which are necessary information for knowing state-year level treatment assignment, we identified 140,815,323 MHSUD claims financed by FFS models for eligible study population, spanning 1,290 counties in 20 states and the D.C. Among these, 322,866 (0.22%) claims were identified as telehealth services.

[41] FFS is a traditional direct payment models where providers are reimbursed for each service. In PCCM plans, primary care providers manage overall patient care receiving both the pre-set, capitated fee per beneficiary and FFS per care provided. Lastly, regarding managed care plans: despite a general trend towards managed care plans in Medicaid, some states have retained a carve-out policy for mental health services. This approach separates behavioral health services from the general managed care plans due to concerns about inadequate focus on behavioral health within the managed care systems.

Table 2-3 presents summary statistics for a sample of 332,866 tele-MHSUD claims. On average, claims are predominantly for mental health disorders (93%) compared to substance use disorders (7%), with an average billing amount of $158, of which Medicaid covered $95. Detailed breakdowns of the top six diagnoses are provided in Table 2-4. Notably, neurodevelopmental disorders represent 30% of these claims, followed by severe mental health illnesses (SMI) including depressive disorders, schizophrenia, and bipolar disorder, as defined by the Substance Abuse and Mental Health Services Administration (SAMHSA)[42]. Within the subset of substance use disorders, opioid-related disorders constitute half of the claims, accounting for 4% of the total tele-MHSUD claims.

Demographically, an average tele-MHSUD claim involves a slightly higher proportion of male patients (52%) with an average age of 27 years. Over half of these cases are for individuals with disabilities, followed by children (28%) and adults over 20 years old (18%). Among non-disabled adults, those not covered under the Affordable Care Act (ACA) exceed those who are (11% vs. 7%). Patients are typically enrolled in Medicaid for nearly an entire year, averaging 358.2 days, and it is likely due to mandatory Medicaid enrollment by states.

Geographically, most telehealth claims originate from metropolitan and adjacent to metropolitan counties, as classified by the 2013 Rural-Urban Continuum Codes[43], with 9.5% of claims coming from completely rural counties having fewer than 2,500 urban inhabitants.

In terms of service providers, claims are more frequently associated with voluntary entities (63%) compared to proprietary ones (37%), with a minimal presence from teaching facilities (0.3%). Providers generally accept new Medicaid patients of various types within a year

[42] https://www.samhsa.gov/serious-mental-illness
[43] https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/. Economic Research Service, U.S. Department of Agriculture

(0.516). However, there is a significant missingness of detailed provider information, including specialties (70% missingness in our final sample), as indicated in the replication exercise in Table 2-2 as well.

2.4.4 County-month Aggregation

For our main analysis, we aggregated the telehealth claims identified in §2.10.4 at the county and month level. Our treatment timing is definitely known only at the year level, which led us to categorize treated states into two groups based on their treatment year: 2017 and 2018. We excluded the last two states treated in 2019 (Nevada and South Carolina) due to the absence of post-treatment data. This approach yielded 59,136 county-month level observations across 7 treated states and 12 control states, encompassing 1,232 counties over 48 months (§2.10.4).

Table 2-6 presents the group mean t-tests between treated and control units using 2016 data, prior to the intervention. On a per-county, per-month basis, we observed that the number of telehealth claims, the total billing amount for telehealth, and the total payment received for telehealth services were marginally higher in treated units compared to control units. Geographically, counties in treated states were less likely to be metropolitan or adjacent to metropolitan areas. Furthermore, treated counties generally had larger populations, higher number of Medicaid enrollees, and also greater size of Medicaid enrollees with behavioral health diagnoses[44]. After adjusting for the number of Medicaid enrollees with behavioral health diagnoses - which is a key demographic for the policy in focus - we found no significant difference in the number of telehealth claims between the control and treated counties ($t = -0.92$).

[44] Correlation between county-month level population, Medicaid enrollees, Medicaid enrollees with behavioral health diagnoses are all high ($\rho > 0.9$), making adjustment by these denominators not much different to each other.

Figure 2-1 shows average trends of treated counties and control counties. First plot shows trends for 864 control counties (red) and 402 treated counties (blue) with treatment year 2017 (CO, DE, MN, TX, WA, WY), while the second plot shows trends for the same 864 control counties (red) and 24 treated counties with treatment year 2018 (MD). Light grey areas denote the treatment year for two different treatment cohorts. From the graphs only, it is difficult to find a notable change after the policy for treated counties that are different from control counties. For treated cohort 2017, it seems control counties increased more with the shock than treated counties (i.e. the slope is higher for control counties after 2017). For treated cohort in 2018, which is actually only Maryland, there seems to be an increasing trend regardless of shock in 2018. To test the policy effects, we move to next section where we employ causal inference from 2 x 2 DiD design across two different treated cohorts.

2.5 Result

2.5.1 Pre-trend Tests

To ascribe a causal interpretation to our difference-in-differences estimates, adherence to the parallel trend assumption is crucial. This assumption posits that, in the absence of treatment, the difference between the treatment and control groups would remain constant over time. Although testing this assumption directly is not feasible, it is a standard practice to examine pre-intervention trends and assume that these trends would have persisted post-intervention in the absence of treatment. Under the parallel trend assumption, the pre-intervention difference between treated and control units should show no systematic divergence away from zero. In Figure 2-2, across all three outcome variables, pre-trends between treated and control units do not show systematic difference pre-intervention in 2016 (i.e. relative treatment time $t<0$) besides a month before the treatment. The deviation in the second last month before treatment at $t = -2$, should not affect our causal estimates much given its size relative to the policy effects post-intervention (i.e. one fifth).

Recent literature has warned that pre-trend tests can be misleading in finite samples (Freyaldenhoven et al. (2019); Kahn-Lang and Lang (2020); Bilinski and Hatfield (2018)). Roth (2022) showed that pre-trend tests can suffer from low power and can potentially exacerbate the bias of point estimates and under-coverage of confidence intervals conditional on passing the low-powered tests. Following the recommendation from Roth (2022), we present the calculated slopes that are detectable 50%, 80% and 90% of the times in our finite sample in Table 2-7. This calculation helps gauge the range of bias in our estimates from linear pre-trends that might exist but not be detected from the low-powered tests. For example, with power at 90% and $\alpha\ at\ 0.01$, pre-trend tests on log (tele-claims per 10K) can fail to detect 10% of the times, when linear slopes of pre-trends are smaller than 0.013. Reassuringly, this is less than 3.3% of the estimated effect size of 0.511 from our main model reported in Table 2-8. In other words, if there is a linear pre-trend which slope is smaller than 0.013, under α at 0.01, we might fail to detect 10% of the times and still this trend we fail to detect biases less than 3.3% of the estimated effect size. For other outcome variables, the maximum slope of linear trend that we might fail to detect with higher than 10% rate are all less than 3.3% of the estimated effect sizes. Hence, across all outcome variables, we gain more confidence that our causal estimates are not affected much from the violation of parallel trend assumption particularly due to any linear trend not detectable from the low-powered pre-trend tests. Note that we restrict the violation of parallel trend assumption to one due to a *linear* divergence of trends between treated and control units.

For an additional check on the parallel trend assumption, we present the average prediction errors for treated units pre-intervention from a model trained without treated observation following Liu et al. (2022). The residuals of these predictions, plotted in Figure 2-3, show good alignment with actual outcomes for the 12 months preceding the intervention, with only minor divergence observed in November of 2016 (t = -2).

Overall, all attempts to test the parallel trend assumption indicate that the assumption is not harmed much in our sample across outcome variables. Based on these tests, we carefully interpret our estimates as causal.

2.5.2 Difference in Differences

In this chapter, we delve into the findings from our difference-in-differences (DiD) models, targeting three telehealth utilization metrics: 1) the number of telehealth claims, 2) the billed amount in dollars for telehealth, and 3) the Medicaid payment in dollars for telehealth. Each of these variables is normalized per 10,000 county population and subsequently log-transformed[45].

Our analysis reveals that the removal of originating site restrictions led to significant increases in telehealth utilization for MHSUD care under FFS financing in Medicaid (Table 2-8). Specifically, we observed increases of 67%, 447%, and 365% in the county-month volume of claims, billed amount, and Medicaid payments, respectively. However, it is important to note that these percentage increases might be misleading due to the low baseline levels of telehealth utilization pre-intervention[46]. On average, per county per month per 10,000 population, the

---

[45] Other transformation methods yielded qualitatively similar results. Normalizing by Medicaid enrollees or those with behavioral care records showed similar patterns due to their strong correlation with county population ($\rho > 0.92$). Inverse hyperbolic sine transformations paralleled results from log transformation, but lacked interpretability in terms of marginal effects. Our choice of transformation and normalization scale aligns with existing literature in Medicaid telehealth utilization (Harju and Neufeld (2021).

[46] The scaling in log transformations can cause variations in resulting percentage changes. Hence, we supplement percentage changes with actual level changes for clarity.

number of telehealth claims rose from 2.02 to 3.37, the billed amount from $337.69 to $1,509.47, and Medicaid payments from $156.79 to $572.28.

Interestingly, the expenditure variables (billed amount and Medicaid payment) exhibit a disproportionately larger increase compared to the number of claims. Translating these figures to a per-claim basis, the average billed amount per telehealth claim increased from $167 to $447, and the average Medicaid payment per claim rose from $77 to $169. Our definition of a claim corresponds to a single visit, which may include multiple procedures. A visit is classified as a telehealth visit if any of its procedures are identified as such using our predefined telehealth identifiers (§2.10.3). Therefore, the larger relative increase in expenditures could be attributed to either an increase in the number of procedures per claim, a shift towards more expensive procedures per claim post-policy change, or potential issues in data quality.

In our analysis, we initially prioritized the quality of claim volume data over the quality of expenditure data, due to concerns about losing a significant number of observations[47]. However, recognizing the importance of expenditure data accuracy, we conducted a subsequent quality check. Based on the DQ Atlas, we identified states with more than a 20% discrepancy in total outpatient expenditure compared to CMS64 reports. Four states from our control group exhibited significant expenditure discrepancies[48].

To assess the impact of potential data quality issues on our findings, we reran our analysis on a conservative sample set, excluding these four states with flagged expenditure data. The results, presented in Table 2-27, show that the Home policy's effect on billed and paid amounts for telehealth is consistent with the full sample analysis in Table 2-8, albeit slightly

[47] This decision was made based on the Data Quality Atlas (DQ Atlas), which indicated fewer quality issues in claim volume data compared to expenditure data. Consequently, our initial sample was reduced from 21 potential states to 13 states with higher-quality claim volume data (detailed in Table A1). Further details on our sampling procedures are provided in §2.10.4

[48] The discrepancies in total expenditure ranged from 20% in Arizona to as high as 46% in Virginia. The states flagged for quality issues included Arizona (20%), Washington D.C. (22%), Virginia (46%), and West Virginia (27%).

smaller. In this conservative sample, we observed an increase in telehealth utilization of 60% (1.94 to 3.37), 359% ($310.84 to $1,422.9), and 290% ($136.22 to $531.25) for the number of claims, billed amount, and paid amount, respectively. Despite these adjustments, the increase in expenditure still disproportionately outpaces the increase in the number of claims.

Further, in Table 2-9, we report DiD regression results at the county-month level, focusing on the mean number of procedures per claim and the mean billed and paid amount per procedure. The analysis indicates that the increase in expenditures per procedure (columns (2) and (3)) is significantly larger than the increase in the number of procedures per claim (column (1)). Given that most telehealth claims involve only one or two procedures [49], it appears that the substantial increase in total expenditure cannot be solely explained by an increase in the number of procedures. Instead, a more than 200% increase in the average billed or paid amount per procedure per claim likely drives this considerable rise in expenditure following the Home policy implementation.

Our comprehensive analysis demonstrates that the removal of originating site restrictions has led to a significant increase in telehealth utilization across various metrics. A critical observation from our study is that the disproportionate rise in expenditure relative to the number of telehealth claims is predominantly attributed to the increase in the average billed and paid amount per procedure per claim. This suggests that after removal of originating site restrictions, telehealth claims have become more costly, despite no substantial change in the number of procedures per claim.

2.5.3 Heterogeneous Treatment Effects

[49] 96% of telehealth claims have only one procedure, 3% of claims have two procedures.

In this chapter, we examine heterogeneous treatment effects of removing originating site restrictions across rurality, patient demographics, and care types. While our primary focus is on the number of telehealth claims, it's important to note that the effects on the other two expenditure outcomes (billed and paid amounts) are broadly similar.

As Table 2-10 illustrates, removing originating site restrictions led to an increase in telehealth utilization across all rurality types. However, our analysis did not reveal significant differences in treatment effects between completely rural, mostly rural, and mostly urban areas[50]. This uniformity suggests that the policy impact on telehealth utilization does not vary significantly across urban and rural areas at the county-month level.

In Table 2-13, we analyze the policy's impact on telehealth claims across different demographic groups[51]. Our findings show no substantial difference between male and female patients in terms of pre-intervention utilization or post-policy impact. Among age groups, children under 18, who were the predominant telehealth users pre-intervention (40%), also experienced the most significant increase in utilization (38.5%) following the policy. Adults, both young and middle-aged, each comprising about 25% of pre-intervention telehealth use, saw a 22% increase in claims post-policy. Senior adults, less represented due to our exclusion of Medicare dual enrollees, showed a smaller increase.

Table 2-14 presents the differential impacts of the policy across various diagnoses. Neuro-developmental disorders, which are closely linked to the high utilization among children, showed the highest pre-intervention usage. The most substantial percentage increases post-policy were observed in trauma-related disorders and other grouped diagnoses.

[50] We also explored differences between metro and non-metro counties, but found no significant heterogeneity in treatment effects.

[51] We also examined heterogeneity across other patient plan characteristics e.g. MCO vs. FFS and disability status, but found no significant differences.

In Table 2-15, over 95% of telehealth services were delivered via video, which also saw a positive impact from removing originating site restriction. This aligns with expectations, given that video services were predominantly eligible under the policy, in contrast to audio and asynchronous modes.

While there are observable variations in the policy impact across different subgroups, the most pronounced effects were on already popular telehealth services pre-policy (e.g., care for children, neuro-developmental disorders, and video-based services). However, given the relatively small number of telehealth claims and the further subdivision into numerous subgroups, we advise caution in interpreting these findings. With the current sample sizes and levels of telehealth utilization, a conservative approach focusing on the overall policy effects on total telehealth utilization is warranted.

2.5.4 Beyond Local Providers

The removal of originating site restrictions signifies a pivotal shift in telehealth policies, allowing patients to receive telehealth services from their homes. In theory, this shift allows patients get care from providers in distance more easily than before, as long as the providers are licensed in the state where the patients reside[52]. This aspect of telehealth is particularly important, as it offers the potential to link patients in underserved areas with healthcare resource that are geographically remote but accessible through telecommunication. In this chapter, we examine whether improved access to telehealth indeed expanded the distance between patients and providers, and particularly for rural areas. Does access to telehealth increase telehealth care

[52] Under certain conditions, such as in border counties, providers may even be located out-of-state without holding a medical license from the patient's state of residence. We explored whether this additional flexibility led to increased telehealth uptake but found no significant differences between border and inner counties, or between counties bordering states with the Federation of State Medical Boards (FSMB) policy and inner counties (see Table 2-11, 2-12).

from providers with farther distance than before? If so, does the effect vary across rurality of patient's residence? For example, do rural counties increase telehealth claims more for connecting farther providers than urban counties?

We first measured the distance between patients and providers at county level for telehealth claims.[53] Our data revealed that most telehealth interactions occurred within a 50-mile radius (85%) and nearly all within 100 miles (93%). Moreover, a significant proportion of telehealth claims involved providers from the same state (87%) and county (45%).

Table 2-16 shows the policy impact for providers within 50 miles. Post-policy change, both telehealth claims and expenditures saw an increase. However, beyond 50 miles (Table 2-17) and 100 miles (Table 2-18), the policy effects, while positive, were not statistically significant. This suggests that increased telehealth usage predominantly facilitated connections between patients and providers relatively close to each other.[54]

Our analysis also delved into whether patients in rural areas, who could greatly benefit from telehealth's potential to connect them with distant providers, demonstrated different utilization patterns compared to urban patients across distance with providers. Tables 2-19 and 2-20 present the results. For telehealth within 50 miles (Table 2-19), all areas, irrespective of rurality, showed increased utilization following the policy change, with the most significant increase in completely rural counties. Conversely, for telehealth beyond 50 miles (Table 2-20), the effects were non-significant except in more urban areas, indicating a divergence in telehealth utilization patterns between rural and urban patients post-policy change.

---

[53] We calculate county distances between patients and providers using the Haversine formula on the internal points of counties (NBER County Distance Database 2010). Due to the high missingness of provider location (50-100%), MS, SD, WY, VA are excluded from the analyses in §2.5.4. All other states had less than 12% missingness of provider location in individual telehealth claims. This exclusion leaves us with 85% of telehealth claims and 66% of county-month observations compared to our final county-month panel.

[54] We also tested the policy effect on the average distance of telehealth per county per month (operationalized with log (average miles+1)). The estimate was positive but not significant (beta = 0.579, p-value = 0.12).

These findings suggest that telehealth for Medicaid MHSUD care did not entirely transcend geographical barriers. With access to telehealth, the average telehealth utilization, regardless of rurality, improved only within a 50-mile range. There could be both demand and supply side reasons. For example, demand-side factors could include a preference among Medicaid MHSUD patients for nearby providers, potentially due to the chronic nature of their healthcare needs and expectation of office visits along the care. On the supply side, provider networks for Medicaid MHSUD patients might be geographically constrained.

The differential response between completely rural and less rural/mostly urban counties to telehealth access for providers beyond 50 miles might be also influenced by various factors. For instance, Medicaid patient acceptance rates by providers could decrease with distance, or patient demand for different types of telehealth care might vary based on urbanity due to different health care needs or preference.

2.6 Conclusion

In this study, we aimed to answer two research questions: 1) What is the impact of the removal of originating site restrictions - allowing patients being home for telehealth - on telehealth utilization for Medicaid MHSUD patients? 2) Does the removal of originating site restrictions 'defy distance', making patients to connect with providers beyond their immediate neighborhood?

Our empirical findings indicate a clear increase in telehealth utilization among Medicaid MHSUD patients following the removal of originating site restrictions. Despite potential adoption barriers such as familiarity to technology, internet or device access, and need for private space for virtual care, all of which were frequently brought up as blockers particularly for

Medicaid population (Parker et al. 2003; Ragnedda and Muschert 2013), there was a notable uptick in telehealth use.

However, it's important to note that while telehealth utilization increased, the expenditures associated with these services, including billing amounts and Medicaid payments, rose disproportionately more, compared to the number of claims. This suggests that each telehealth claim became costlier post-policy implementation warranting closer scrutiny. It could be due to improved engagement between patients and providers in telehealth. If patients could spend extra minutes that could have been spent in travel in further consultation with providers on telehealth, per visit price could increase (e.g. 10 min consultation to 1 hour consultation per visit). In this case, telehealth is likely improving quality of care and the increased cost can be warranted. On the other hand, providers might find it more difficult to evaluate and manage patients over video call, requiring longer time for the same treatment outcome compared to in-person visits. In this case, increased cost per visit actually implies overutilization and warrants further examination. One important caveat is that we study only telehealth utilization, not total utilization including in-person care. Thus, comprehensive analysis is crucial for appropriately setting and monitoring telehealth budgets to ensure positive net welfare.

To answer our second question, we examined whether telehealth services indeed connected patients and providers across wider geographical areas. Our findings reveal that, contrary to expectations, telehealth interactions largely remain confined to patients' geographical neighborhoods, even with the convenience of accessing services from home without travel. In our Medicaid MHSUD sample, telehealth use for patients and providers within 50 mile distance increased the most with the removal of originating site restrictions. For cases beyond 50-mile or 100-mile distance between patients and providers, we could not find any increase in telehealth utilization after removal of originating site restrictions. Furthermore, contrary to common belief, we did not find that patients in rural areas seek out providers beyond immediate neighborhood more than patients in urban areas. Rather, patients in urban areas increased telehealth use for connecting with farther away providers (beyond 50 miles) while patients in rural areas did not. This finding implies that commonly touted telehealth benefits for rural areas – connecting patients in underserved areas to providers in urban or resourceful areas – did not realize fully with removing originating site restrictions.

Our study, while providing valuable insights into the effects of removing geographical restrictions on telehealth utilization, is subject to several limitations that merit some consideration. First, despite our rigorous efforts to ensure data accuracy, especially during the early years of the Transformed Medicaid Analytical Files (TAF), concerns on the data completeness remain. Second, we study telehealth utilization, not total care utilization. Future research should evaluate the impact on total care utilization, examining the interaction with in-person visits. Third, we restrict to MHSUD telehealth financed through fee-for-services due to the policy itself. However, managed plan financing became more popular in Medicaid programs over time, prompting investigation in different financing models as well. Also, note that our results may not be generalizable in other insurance types (e.g. Medicare and private) and post-pandemic.

Considering the current interest on expanding the waivers on originating site restrictions, our work is timely and uniquely important. In October of 2023, American Hospital Association voiced support for the CONNECT for Health Act (S.2016/H.R. 4189), legislation that would permanently remove geographic restrictions that limit where patients can access telehealth, add homes and other clinically appropriate sites. We believe our work can serve as one available reference point, particularly when there is no empirical research focusing on removal of originating site restrictions using causal identification.

2.7 Reference

Abbasi-Feinberg, F. (2020). Telemedicine coding and reimbursement-current and future trends. Sleep Medicine Clinics, 15(3):417–429.

Bilinski, A. and Hatfield, L. A. (2018). Nothing to see here? non-inferiority approaches to parallel trends and other model assumptions. arXiv preprint arXiv:1805.03273.

Butzner, Michael and Cuffee, Yendelela (2021). Telehealth interventions and outcomes across rural communities in the United States: Narrative review. Journal of Medicine Internet Research, 23(8): e29575. doi: 10.2196/29575. PMID: 34435965; PMCID: PMC8430850.

CCHP (2017). State telehealth laws and reimbursement policies: A comprehensive scan of the 50 states and district of columbia. https://www.cchpca.org/resources/. Center for Connected Health Policy, Accessed: 2023-12-10.

CMS (2018). Information on Medicare Telehealth. The Centers for Medicare and Medicaid. https://www.cms.gov/About-CMS/Agency-Information/OMH/Downloads/Information-on-Medicare-Telehealth-Report.pdf, Accessed Jan 2024.

Dahlstrand, A. (2023). Defying distance: the provision of services in the digital age. Working Paper.

Dong, X. (2022). The impact of telehealth parity laws on health expenses. Available at SSRN 4246602.

Douglas, M. D., Xu, J., Heggs, A., Wrenn, G., Mack, D. H., and Rust, G. (2017). Assessing telemedicine utilization by using medicaid claims data. Psychiatric Services, 68(2):173–178.

Dunn, A., Gottlieb, J. D., Shapiro, A. H., Sonnenstuhl, D. J., and Tebaldi, P. (2024). A denial a day keeps the doctor away. The Quarterly Journal of Economics, 139(1):187–233.

Freyaldenhoven, S., Hansen, C., and Shapiro, J. M. (2019). Pre-event trends in the panel event-study design. American Economic Review, 109(9):3307–3338.

Gilman, M. and Stensland, J. (2013). Telehealth and medicare: payment policy, current use, and prospects for growth. Medicare & medicaid research review, 3(4).

Harju, A. and Neufeld, J. (2021). The impact of the medicaid expansion on telemental health utilization in four midwestern states. Telemedicine and e-Health, 27(11):1260–1267.

Kahn-Lang, A. and Lang, K. (2020). The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications. Journal of Business & Economic Statistics, 38(3):613–620.

Lacktman, N. M., Acosta, J. N., and Levine, S. J. (2019). 50-state survey of telehealth commercial payer statues. Foley and Lardner LLP.

Liu, L., Wang, Y., and Xu, Y. (2022). A practical guide to counterfactual estimators for causal inference with time-series cross-sectional data. American Journal of Political Science.

Lustig, Tracey A. (2012). The role of telehealth in an evolving health care environment: workshop summary. Institute of Medicine of the National Academies, p 158.

McCullough, J.S., Ganju, K.K., and Ellimoottil C. (2021). Does Telemedicine Trascend Disparities or Create a Digital Divide? Evidence from the Covid-19 Pandemic. SSRN. April, 2021. http://dx.doi.org/10.2139/ssrn.3834445

Mehrotra, A., Jena, A. B., Busch, A. B., Souza, J., Uscher-Pines, L., and Landon, B. E. (2016). Utilization of telemedicine among rural medicare beneficiaries. Jama, 315(18):2015–2016.

National Bureau of Economic Research (2010). County Distance Database. https://www.nber.org/research/data/county-distance-database. Accessed Jan 2024.

Parker, RM., Ratzan, SC., Lurie, N (2003). Health literacy: a policy challenge for advancing high-quality health care. *Health Aff (Millwood)*. 2003;22(4):147-153. doi:10.1377/hlthaff.22.4.147

Rabideau, B. and Eisenberg, M. D. (2022). The effects of telemedicine on the treatment of mental illness: Evidence from changes in health plan benefits. SSRN. http://dx.doi.org/10.2139/ssrn.4065120.

Ragnedda M, Muschert GW. (2013). *The Digital Divide*. Routledge Florence; 2013. doi:10.4324/9780203069769

Ratcliffe, M., Burd, C., Holder, K., and Fields, A. (2016). Defining rural at the US Census Bureau. https://www2.census.gov/geo/pdfs/reference/ua/Defining Rural.pdf. Accessed Jan 2024.

Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. American Economic Review: Insights.

Saunders, H., Guth, M., and Eckart, G. (2023). A look at strategies to address behavioral health workforce shortages: Findings from a survey of state medicaid programs. Accessed: 2023-12-10.

Saunders, H. and Rudowitz, R. (2022). Demographics and Health Insurance Coverage of Nonelderly Adults with Mental Illness and Substance Use Disorders in 2020. KFF. https://www.kff.org/mental-health/issue-brief/demographics-and-health-insurance-coverage-of-nonelderly-adults-with-mental-illness-and-substance-use-disorders-in-2020/, Accessed: 2024 Jan.

Talbot, J. A., Jonk, Y. C., Burgess, A. R., Thayer, D., Ziller, E., Paluso, N., and Coburn, A. F. (2020). Telebehavioral health (tbh) use among rural medicaid beneficiaries: Relationships with telehealth policies. Journal of Rural Mental Health, 44(4):217.

Tilhou AS, Jain A, DeLeire T. Telehealth Expansion, Internet Speed, and Primary Care Access Before and During COVID-19. *JAMA Netw Open.* 2024;7(1):e2347686. doi:10.1001/jamanetworkopen.2023.47686

Thomas, L. and Capistrant, G. (2017). State telemedicine gaps analysis: Coverage and reimbursement. American Telemedicine Association, Accessed: 2023-12-16.

Yeramosu, D., Kwok, F., Kahn, J. M., and Ray, K. N. (2019). Validation of use of billing codes for identifying telemedicine encounters in administrative data. BMC health services research, 19:1–9.

Yu, J., Mink, P. J., Huckfeldt, P. J., Gildemeister, S., & Abraham, J. M. (2018). Population-level Estimates of Telemedicine Service Provision Using an All-Payer Claims Database. Health Affairs, 37(12), 1931-1939. https://doi.org/10.1377/hlthaff.2018.05116

Zeltzer, D., Einav, L., Rashba, J., and Balicer, R. D. (2023). The impact of increased access to telemedicine. Journal of the European Economic Association, page jvad035.

Zhou, M., Li, X., and Burtch, G. (2021). Healthcare across boundaries: Urban-rural differences in the finan- cial and healthcare consequences of telehealth adoption. Boston University Questrom School of Business Research Paper, (3807577).

2.8 Tables

Table 2-1.Studies on telehealth utilization in Medicaid

	(1) Douglas et al. (2017)	(2) Talbot et al. (2019)	(3) Harju and Neufeld (2021)	(4) CMS Report (2021 May)	(5) Ours
Data	MAX	MAX	state offices	TAF	TAF
Study Sample					
State	22 states	42 states	IA, NE, MN, WI	All states	13 states
Year	2008	2011	2014-17	2020 (Mar-Jun)	2016-19
Diagnosis	all	all	MH	all	MHSUD
Setting	outpatient	outpatient	in/out-patient	outpatient	outpatient
Duality	all	no dual	all	all	no dual
Financing	all	all	all	all	fee-for-services
Telehealth Identifiers	GT, Q3014	GT, GQ, G0406-G0408, G0425-G0427, G0459, 0188T, 0189T	GT	a combination of CPT, HCPCS, procedural codes, place of service[a]	§2.10.3

Notes: Previous multiple state studies on telehealth utilization for Medicaid population using Medicaid claims data before the pandemic.
a. The list of codes is similar to ours detailed in §2.10.3.

Table 2-2: Replication of Harju and Neufeld (2021)

		Ours (TAF Mediated data)			HN (2021)
		(1)	(2)	(3)	(4)
		Procedure	+Taxonomy	Claims/10K	Claims/10K
Minnesota					
	year 2016	28,781	2,022	3.6	~[a]20
	year 2017	40,582	6,561	11.8	~[a]20
	year 2018	56,923	11,222	20.3	-[b]
Wisconsin					
	year 2016	28,931	10,075	17.5	~[a]20

year 2017	31,688	9,994	17.3	~[a]20
year 2018	42,022	13,826	23.9	-[b]

Note: The number of telehealth claims for mental health services for in- and out-patient settings across all plan (including both dual and nondual enrollees) and financing types (including both capitation per patient per time and fee-for-services). Column (1) represents the number of claims remained after procedure code matching. Column (2) represents the number of claims remained after both procedure and taxonomy code matching. Column (3) and column (4) report the final metric (i.e. the number of claims per 10K population) for our replication and Harju and Neufeld (2021). The replication process carefully followed Harju and Neufeld (2021).

a. We added ~ to denote approximation. Harju and Neufeld (2021) reports rates of telehealth claim per 10K population only in graphs with sparse tick labels, without the exact numbers.

b. HN (2021)'s sample is till 2017, so there are no reports for year 2018.

Table 2-3 Summary statistics of telehealth MHSUD claims

Variables	N	Mean	St. Dev.	Min	Max
claim characteristics					
I(MH diagnosed)	322,673	0.934	0.248	0	1
I(SUD diagnosed)	322,673	0.078	0.269	0	1
dollar billed ($)	322,866	158	318	0	99,213
dollar paid ($)	322,837	95	141	0	14,570
I(place of service 02)[a]	322,866	0.362	0.481	0	1
I(modifier GT)[a]	322,866	0.851	0.356	0	1
Patient characteristics					
age	322,866	27.821	18.521	1	93
sex (1 = male)	322,866	0.524	0.499	0	1
I(adults under ACA)[b]	322,866	0.074	0.262	0	1
I(adults not under ACA)[b]	322,866	0.111	0.314	0	1
I(children)[b]	322,866	0.288	0.453	0	1
I(disabled)[b]	322,866	0.523	0.499	0	1
I(mandatory enroll)[c]	322,866	0.911	0.285	0	1
days enrolled	322,866	358.265	38.488	0	366
I(metro)	322,866	0.650	0.477	0	1
I(adjacent to metro)	322,866	0.192	0.394	0	1

I(completely rural)	322,866	0.095	0.293	0	1
%rurality	317,861	33.280	28.534	0	100
provider characteristics					
I(teaching facility)	271,135	0.003	0.053	0	1
I(proprietary)	238,942	0.373	0.483	0	1
I(accepting new Medicaid)[d]	271,135	0.516	0.500	0	1

Note: Summary statistics for 322,866 telehealth MHSUD services between 2016-2019 in 13 study states. If there are missingness in variables, summary statistics are reported based on the subsample without missingness.

a. A claim can have multiple telehealth codes and is classified as telehealth service with at least one telehealth code present.

b. Medicaid enrollees can be classified into these four buckets, exclusive to each other.

c. Some states mandate enrollment for some types of enrollees. For these types, states assign managed care plans automatically if there is not request for change on patient end.

d. Whether providers accepted any new Medicaid patients in a year, not necessarily for mental health or via telehealth.

Table 2-4: Types of care for MHSUD telehealth: Diagnosis

ICD codes	Diagnosis Description	N	cum%
F70-73, F78-82, F84, F88-90, F95	Neurodevelopmental disorder	95,432	0.296
F32-33	Depressive disorders	50,177	0.451
F06, F20-25, F28-29	Schizophrenia spectrum and other psychotic disorders	36,495	0.564
F30-31	Bipolar and related disorders	31,763	0.662
F40-41	Anxiety and fear-related disorders	21,380	0.729
F43-44, F94	Trauma- and stressor-related disorders	20,588	0.792
F11	Opioid-related disorders	14,145	0.836
Others	Others	52,886	1

Note: Top diagnoses of 322,866 telehealth MHSUD services. We followed categorization of ICD from the Clinical Classifications Software Refined (CCSR), developed as part of the Healthcare Cost and Utilization Project (HCUP), a Federal-StateIndustry partnership sponsored by the Agency for Healthcare Research and Quality (AHRQ). Corresponding CCSR codes are MBD014 (neurodevelopmental disorder), MBD002 (depressive disorders), MBD001 (schizophrenia), MBD003 (bipolar), MBD005 (anxiety), MBD007 (trauma). Diagnoses with less than 4% prevalence were collectively categorized as *Others*.

Table 2-5.Types of care for MHSUD telehealth: Provider taxonomy

Taxonomy Classification	N	cum%
Clinic/Center	35,285	0.109
Community/Behavioral Health	26,525	0.191
Counselor	15,834	0.239
Psychiatry & Neurology	11,411	0.275
Psychologist	4,645	0.289
Others	8,567	0.315
Missing	221,947	1

Note: Top taxonomy groups for 322,866 telehealth MHSUD services. Taxonomy categorization for MHSUD providers follows Harju and Neufeld (2021). Our data shows high missingness (70%) in provider taxonomy and NPI codes.

Table 2-6.Group mean T-tests between treated and control units pre-intervention

Variables	Control	Treated	Tstatistics	Pvalue
Telehealth claims	2.93	3.25	-1.66	0.10
Log (Telehealth per 10K population)	0.37	0.30	7.82	0.00
Log (Telehealth per 10K BH patients: Medicaid[a])	1.22	1.25	-0.92	0.36
Dollar billed	418.31	490.25	-2.55	0.01
Dollar paid	235.87	267.80	-1.83	0.07
I(metro)	0.40	0.38	2.56	0.01
I(adjacent to metro)	0.37	0.32	5.30	0.00
Population	74607.79	114932.42	-7.94	0.00
Medicaid enrollee	16611.61	21928.66	-5.10	0.00
BH patients: Medicaid[a]	3787.62	5253.25	-6.63	0.00
BH patients: nondual[a]	2782.92	4392.99	-8.67	0.00
BH patients: nondual FFS plans[a]	278.84	258.12	0.98	0.33

Notes. Group mean t-tests for county-month level variables in year 2016, pre-intervention.

a. BH patients are defined as Medicaid enrollees with behavioral health diagnosis for a year. The groups are ordered in sizes: Population ⊇ Medicaid enrollees ⊇ BH patients: Medicaid ⊇ BH patients: nondual ⊇ BH patients: nondual FFS.

Table 2-7.Calculated slopes that are detectable in pre-trend tests

	(1) log(tele-claims per 10K)	(2) log(tele-billed per 10K)	(3) log(tele-paid per 10K)
power 50%	0.006	0.024	0.022
power 80%	0.011	0.044	0.040
power 90%	0.013	0.054	0.049
effect size (DiD)	0.504	1.642	1.481

Note: Minimum detectable linear slopes in pre-trend tests across varying following approaches in Roth (2022).

Table 2-8: Home policy effects on county-month telehealth utilization

	Dependent variable:		
	log(claims per 10K)	log(billed per 10K)	log(paid per 10K)
	(1)	(2)	(3)
I(Home)	0.517**	1.700*	1.537*
	(0.102)	(0.439)	(0.421)
Observations	22,392	22,392	22,392
R^2	0.556	0.558	0.537
Adjusted R^2	0.542	0.545	0.523
Residual Std. Error (df = 21734)	0.449	1.787	1.621

Notes. Unit = county-month aggregation of telehealth claims. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017 and DE. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; ***p<0.005.

Table 2-9: Home policy effects on the number of procedures per claim, billed and paid amount per procedure

	Dependent variable:		
	No. prcdrs per claim	log(bill per procedure)	log(paid per procedure)
	(1)	(2)	(3)
I(Home)	0.254	1.221*	1.075
	(0.093)	(0.403)	(0.409)
Observations	22,392	22,392	22,392
R^2	0.450	0.556	0.532
Adjusted R^2	0.434	0.543	0.518
Residual Std. Error (df = 21734)	0.483	1.673	1.475

Notes. Unit = county-month aggregation of telehealth claims. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017 and DE. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; ***p<0.005.

Table 2-10: Heterogeneous treatment effects: rurality

	Dependent variable:		
	log(claims per 10K) (1)	log(billed per 10K) (2)	log(paid per 10K) (3)
I(Home)	0.720**	1.542**	1.353**
	(0.091)	(0.323)	(0.301)
I(Home):MostlyRural	−0.300	−0.104	−0.094
	(0.108)	(0.382)	(0.341)
I(Home):MostlyUrban	−0.194	0.388	0.466
	(0.078)	(0.247)	(0.239)
Observations	22,392	22,392	22,392
R^2	0.557	0.558	0.538
Adjusted R^2	0.544	0.545	0.524
Residual Std. Error (df = 21731)	0.449	1.786	1.620

Notes. Unit of analysis: county-month. Baseline: completely rural counties. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017 and Delaware. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. Counties are categorized into three: completely rural, mostly rural and mostly urban, following the USDA Economic Research Service (Ratcliffe et al. (2016)). In our sample, there are 127 completely rural counties, 252 mostly rural counties and 246 mostly urban counties. Sample means of log(claims per 10K) pre-intervention are 0.427, 0.299, 0.187 for completely rural, mostly rural and mostly urban counties, respectively. *p<0.05; **p<0.01; ***p<0.005.

Table 2-11: Policy effects across inner vs. border counties

	Dependent variable:		
	log(claims per 10K) (1)	log(billed per 10K) (2)	log(paid per 10K) (3)
I(Home)	0.488**	1.415*	1.302*
	(0.095)	(0.372)	(0.361)
I(Home):border	0.087	0.694	0.603
	(0.090)	(0.342)	(0.319)
Observations	22,392	22,392	22,392
R^2	0.556	0.559	0.538

Adjusted R^2	0.543	0.545	0.524
Residual Std. Error (df = 21732)	0.449	1.786	1.620

Notes. Unit of analysis: county-month. Baseline = inner counties. Pre mean Y for control units are 0.294, 1.569, 1.382 for log(teleclaims per 10K), log(billed per 10K), log(paid per 10K), respectively. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017 and DE. Standard errors in parentheses are not yet clustered at state-year, the policy implementation level. The reported S.Es are adjusted with clustering, which is then used in the statistical significance testing. *p<0.05; **p<0.01; ***p<0.005.

Table 2-12: Policy effects for counties bordering FSMB states

	Dependent variable:		
	log(claims per 10K)	log(billed per 10K)	log(paid per 10K)
	(1)	(2)	(3)
I(Home)	0.452**	1.353*	1.256*
	(0.094)	(0.333)	(0.341)
I(Home):border FSMB	0.226	1.081	0.928
	(0.117)	(0.426)	(0.392)
Observations	22,392	22,392	22,392
R^2	0.557	0.560	0.539
Adjusted R^2	0.544	0.546	0.525
Residual Std. Error (df = 21732)	0.449	1.784	1.618

Notes. Unit of analysis: county-month. Baseline = counties not bordering FSMB states. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017 and DE. Standard errors in parentheses are not yet clustered at state-year, the policy implementation level. The reported S.Es are adjusted with clustering, which is then used in the statistical significance testing. *p<0.05; **p<0.01; ***p<0.005.

Table 2-13: Heterogeneous treatment effects: patient demographics

	Dependent variable: log(claims per 10K)						
	Sex		Age				
	Male	Female	0-18	18-26	26-44	44-65	65+
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Frequency (pre)	50%	49%	40%	9%	23%	26%	0.3%
Mean Y (pre)	0.166	0.159	0.146	0.031	0.071	0.091	0.001
I(Home)	0.419**	0.409**	0.325*	0.102*	0.229*	0.196*	0.010**
	(0.080)	(0.069)	(0.083)	(0.028)	(0.059)	(0.057)	(0.002)
Observations	22,392	22,392	22,392	22,392	22,392	22,392	22,392
R^2	0.494	0.511	0.502	0.306	0.417	0.406	0.130
Adjusted R^2	0.479	0.497	0.487	0.285	0.399	0.388	0.104
Res. SE (df = 21837)	0.368	0.329	0.396	0.132	0.201	0.202	0.043

Notes. Unit of analysis: county-month. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017 and DE. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; ***p<0.005.

Table 2-14: Heterogeneous treatment effects: care diagnosis

	Dependent variable: log(claims per 10K)							
	Neuro	Trauma	Depress	Schizo	Opioid	Bipolar	Anxiety	Others
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Frequency (pre)	49%	12%	12%	1.3%	0.3%	0.3%	0.15%	23%
Mean Y (pre)	0.146	0.038	0.031	0.001	0.0004	0.0008	0.0009	0.083
I(Home)	0.139**	0.330*	0.031	0.039***	-0.020	-0.002	0.007	0.337*
	(0.030)	(0.102)	(0.019)	(0.004)	(0.015)	(0.001)	(0.003)	(0.112)
Observations	22,392	22,392	22,392	22,392	22,392	22,392	22,392	22,392
R^2	0.547	0.616	0.632	0.172	0.511	0.121	0.107	0.588
Adjusted R^2	0.533	0.605	0.621	0.147	0.496	0.094	0.080	0.576
Res. SE (df = 21837)	0.316	0.227	0.110	0.017	0.068	0.019	0.029	0.247

Notes. Unit of analysis: county-month. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017 and DE. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; ***p<0.005.

Table 2-15: Heterogeneous treatment effects: modality

	Dependent variable: log(claims per 10K)		
	Synchronous video	Synchronous audio	Asynchronous
	(1)	(2)	(3)
Frequency (pre)	95%	2%	0.5%
Mean Y (pre)	0.256	0.016	0.0001
I(Home)	0.473**	-0.001	-0.001
	(0.103)	(0.002)	(0.0005)
Observations	22,392	22,392	22,392
R^2	0.542	0.518	0.056
Adjusted R^2	0.528	0.504	0.027
Residual Std. Error	0.426	0.109	0.007

Notes. Unit = County-month level aggregation of telehealth utilization. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017 and DE. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; ***p<0.005

Table 2-16: Telehealth with providers within 50 miles

	Dependent variable:		
	log(claims per 10K) (1)	log(billed per 10K) (2)	log(paid per 10K) (3)
I(Home)	0.431*	0.914**	0.844**
	(0.109)	(0.192)	(0.174)
Observations	14,796	14,796	14,796
R^2	0.552	0.386	0.377
Adjusted R^2	0.538	0.364	0.355
Residual Std. Error (df = 14349)	0.430	1.949	1.721

Notes. County-month aggregation of telehealth claims with +50 miles between patients and providers. Distance is calculated using patient's and provider's location in county level based on great-circle distances. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017 and DE. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; ***p<0.005.

Table 2-17: Telehealth with providers beyond 50 miles

	Dependent variable:		
	log(claims per 10K) (1)	log(billed per 10K) (2)	log(paid per 10K) (3)
I(Home)	0.040	0.603	0.559
	(0.031)	(0.310)	(0.287)
Observations	14,796	14,796	14,796
R^2	0.417	0.443	0.443
Adjusted R^2	0.399	0.425	0.416
Residual Std. Error (df = 14349)	0.206	1.936	1.694

Notes. County-month aggregation of telehealth claims with +50 miles between patients and providers. Distance is calculated using patient's and provider's location in county level based on great-circle distances. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017 and DE. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; ***p<0.005.

Table 2-18: Telehealth with providers beyond 100 miles

	Dependent variable:		
	log(claims per 10K)	log(billed per 10K)	log(paid per 10K)
	(1)	(2)	(3)
I(Home)	0.012	0.290	0.266
	(0.023)	(0.264)	(0.237)
Observations	14,796	14,796	14,796
R^2	0.350	0.409	0.408
Adjusted R^2	0.330	0.391	0.389
Residual Std. Error (df = 14349)	0.136	1.451	1.237

Notes. County-month aggregation of telehealth claims with +100 miles between patients and providers. Distance is calculated using patient's and provider's location in county level based on great-circle distances. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017 and DE. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; *** p<0.005.

Table 2-19: Telehealth with providers within 50 miles across ruralities

	Dependent variable:		
	log(claims per 10K) (1)	log(billed per 10K) (2)	log(paid per 10K) (3)
I(Home) *baseline = completely rural*	0.768** (0.125)	1.499 ** (0.306)	1.287 ** (0.271)
I(Home): mostly rural	-0.444**	-0.931	-0.789
	(0.082)	(0.372)	(0.348)
I(Home): mostly urban	-0.384***	-0.542**	-0.330
	(0.007)	(0.082)	(0.167)
Observations	14,796	14,796	14,796
R^2	0.556	0.387	0.378
Adjusted R^2	0.543	0.365	0.356
Residual Std. Error (df = 14347)	0.428	1.947	1.720

Notes. County-month aggregation of telehealth claims in the same county between patients and providers. Baseline is completely rural counties. Distance is calculated using patient's and provider's location in county level based on great-circle distances. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017. After excluding states with high missingness in provider addresses, 85% of telehealth claims remained. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; ***p<0.005.

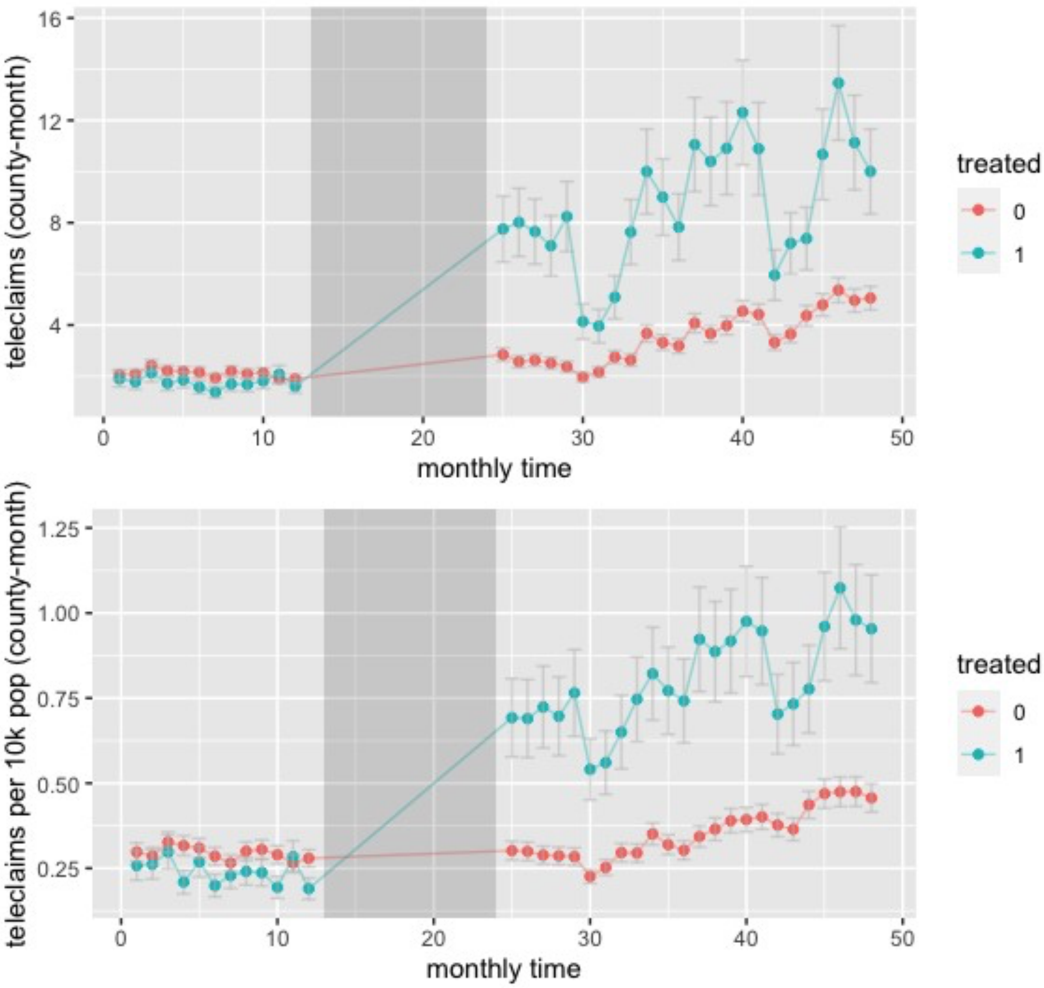
Table 2-20: Telehealth with providers beyond 50 miles across ruralities

	Dependent variable:		
	log(claims per 10K) (1)	log(billed per 10K) (2)	log(paid per 10K) (3)
I(Home)	-0.008	-0.195	-0.132
baseline = completely rural	(0.023)	(0.208)	(0.182)
I(Home): mostly rural	0.065**	0.837**	0.683**
	(0.007)	(0.138)	(0.120)
I(Home): mostly urban	0.054**	1.096**	0.984**
	(0.010)	(0.165)	(0.181)
Observations	14,796	14,796	14,796
R^2	0.417	0.444	0.435
Adjusted R^2	0.399	0.427	0.417
Residual Std. Error (df = 14347)	0.206	1.934	1.691

Notes. County-month aggregation of telehealth claims in the same county between patients and providers. Baseline is completely rural counties. Distance is calculated using patient's and provider's location in county level based on great-circle distances. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017. After excluding states with high missingness in provider addresses, 85% of telehealth claims remained. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; ***p<0.005.

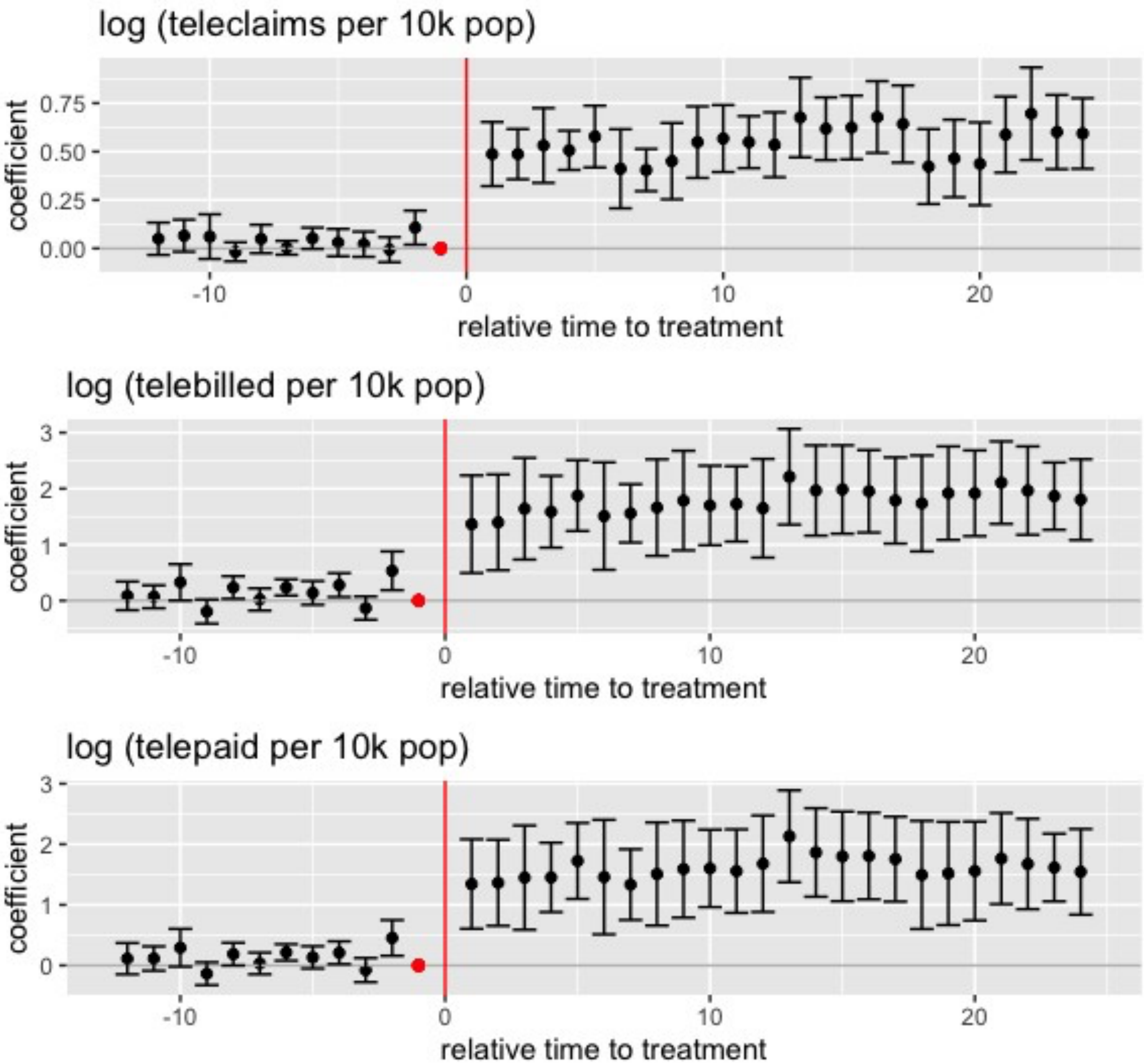
2.9 Figures

Figure 2-1: Telehealth claims per county per month



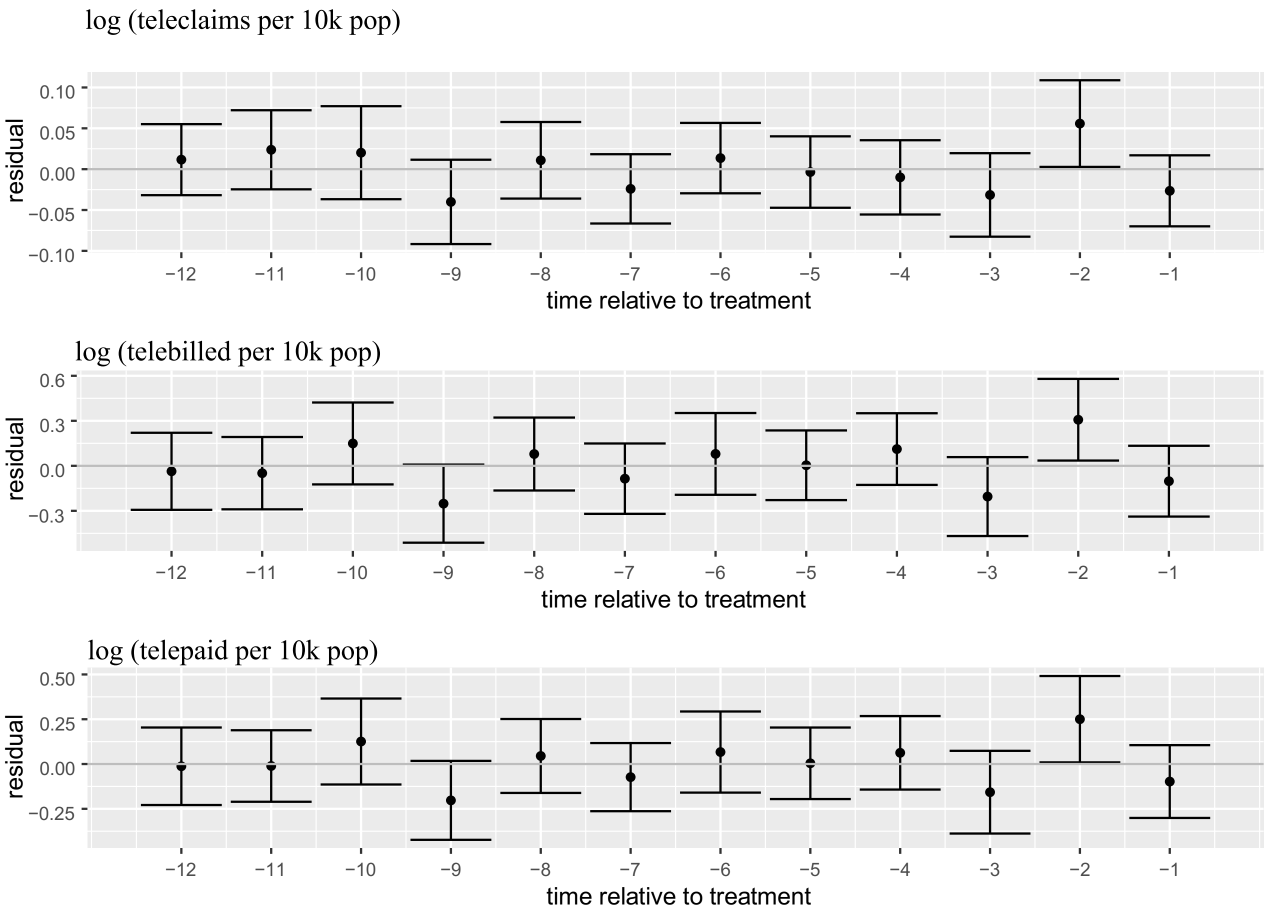
Notes: County-month trends of telehealth claims in treated and control units in year 2016, 2018, 2010. Year 2017 is excluded (grey boxed) because we do not know the exact month of the policy implementation but year. Confidence level is at 99%. Fluctuation in treated units after intervention reflects seasonality in MHSUD care - drops in summer.

Figure 2-2: Event study



Notes: County-month difference of telehealth utilization between treated and control units both pre- (left to the vertical red bar) and post-intervention (right to the vertical red bar). Year 2017 is excluded (i.e. t = -1 in Dec 2016, t = 1 in Jan 2018). Dec 2016 (t = -1) serves the reference point for coefficients. Confidence level is at 99%.

Figure 2-3: Residuals for treated units from models trained without treatment



Notes: Residuals of outcome variables for treated units pre-intervention, from models trained with units not under treatment (i.e. treated units pre-intervention and control units pre- and post- intervention). Confidence level is at 99%.

2.10 Appendices.

2.10.1 Study States

Table 2-21. Telehealth Policies across Potential Study States

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
No	State	Sample	Treatment	FSMB	Bordering states with FSMB	Private	Medicaid
1	CO	−*a*	2017	2017 (16)	AZ, KS, NE, OK (19), UT, WY	yes	B
2	DE	yes	2017		MD (19), PA	yes	B
3	MN	yes	2017	2017 (15)	IA, MI (19), ND (19), SD, UT	yes	B
4	TX	−*a*	2017		OK (19)	no	B
5	WA	yes	2017	2017	ID	yes	A
6	WY	yes	2017	2017 (15)	CO, ID, MT, NE, SD, UT	no	C
7	MD	−*a*	2018	2019	PA, WV	yes	C
8	MI	−*c*	2018		WI	no	C
9	NV	−*b*	2019	2017 (15)	AZ, ID, UT	no	A
10	SC	−*b*	2019		GA (19)	no	C
11	AL	yes		2017 (15)	GA (19), MS, OK (19), TN (19)	no	C
12	AZ	yes		2017 (16)	CO, NV, UT	no	B
13	AR	−*a*			MS, OK (19), TN (19)	yes	C
14	DC	yes			MD (19)	no	C
15	GA	−*a*		2019	AL, TN (19)	yes	C
16	HI	−*c*		2019		yes	A
17	IL	−*a*		2017 (15)	IA, MI (19), KY (19), WI,	no	C
18	IN	yes			IL, KY (19), MI (19)	no	C
19	MS	yes		2017 (16)	AL, TN (19)	yes	A
20	SD	yes		2017 (15)	IA, MN, MT, NE, ND (19), WY	no	C
21	VA	yes			KY (19), MD (19), TN, WV	no	B
22	WI	yes		2017 (15)	IL, IA, MI (19), MN	no	B

23	WV	yes	2017 (15)	KY (19), MD (19), PA	no	C

Notes: 23 states with originating site restrictions between 2016-19. Column (2) shows our study sample after careful examination on data quality (a. Removal due to data quality issues according to the DQ Atlas. b. Removal due to no post-treatment observation. c. Removal due to no telehealth claims identified). Column (3) shows years when our focal policy of interest (i.e. removal of originating site restrictions) was implemented (CCHP (2017)). States with years in column (3) are treated ones. Column (4)-(7) report other telehealth policies that can be confounding with our focal policy. Column (4) reports years when cross state licensure compacts were implemented (CCHP (2017)). FSMB was implemented with years of delay before its first day in 2017 April, in which case years of announcement of joining the compact was noted in the parentheses. Column (5) reports bordering states with cross state licensure compacts, which can incentivize providers in nearby markets to render telehealth. States without parentheses introduced licensure compacts in 2017 April. All other states which joined later are denoted with their joining years in the parentheses. Column (6) shows whether the state had private payment parity law for telehealth (Thomas and Capistrant (2017)). Column (7) denotes the letter grades for Medicaid parity laws from the American Telehealth Association (Thomas and Capistrant (2017)). During 2016-2019, states showed little change on these parity laws.

2.10.2 DQ Atlas

Below are our red flags for year 2016, the most incomplete year for TAF claims data during our study time:

1. Missing key variables (>20%)
   - diagnosis (identifies MHSUD claims): MD (100%)
   - procedure for professionals (identifies telehealth claims): all less than 10%
   - procedure for institutions (identifies telehealth claims): CO (90%), GA (99%), IL (suppressed), MD (suppressed), TX (100%)
   - place of service (identifies telehealth claims): AR (94%), MO (23%), WV (23%)
2. Difference with other data (>20%)
   - enrollment and claims (>10% diff. with PI): DC (15%), IN (12%), MS (20%)
   - dollar expenditure (>20% diff. with CMS 64): AZ (20%), AR (71%), CO (39%), DC (22%), MI (51%), OH (47%), TX (46%), VA (46%), WV (27%)

The caveat is that DQ Atlas reports missingness in all outpatient claims, while our study sample focuses on MHSUD fee-for-service claims (6% of total outpatient claims). Thus, we removed states with high missingness above 97% for key variables as a rule-of-thumb (i.e. at least half of our data would be missing). Since procedure codes are more frequently used for telehealth identifiers (CPT code ’GT’ consists 80% of telehealth claims), we do not remove further states based on high missingness of place of service which was implemented only after 2017. For other concerning differences in enrollment or dollar expenditure with other data source, we further checked once we pulled our data.

2.10.3 Telehealth Identifiers

Referring on Abbasi-Feinberg (2020) and AMA reports [55], we identified both synchronous and asynchronous telehealth services with CPT/HCPCS procedural codes and place of service codes.

1. Synchronous telehealth codes
   - audio: 98966, 98967, 98968, 99441, 99442, 99443, FQ, 93
   - video: GT, 95 (2017-), place of service 02 (2017-), facility fee Q3014, T1014
2. Asynchronous telehealth codes
   - related to conducted E/M: GQ, 99421-99423, 98970-98972, G2061-2063
   - not originating nor leading to E/M: G2012, G2010, G0071
   - remote physiological monitoring: 99457, 99458, 99453, 99454, 99091, 99451

Asynchronous telehealth services only account for less than 0.2% of total telehealth claims identified. For synchronous telehealth, modifier GT and place of service 02 consist 99.9% of identified telehealth. For asynchronous telehealth, modifier GQ consists 96% before 2017 and losses its prevalence to 47% in 2019.

2.10.4 Sample

To construct study sample, we first identified eligible population (Table 2-20) and then pulled MHSUD FFS claims from those eligible population (Table 2-21). On the pulled data, we removed some claims with invalid county information (Table 2-22) to get the final study sample.

Table 2-22: Steps for identifying eligible population

no	steps	total enrollees	% previous	% initial
1	Identifiable in 2016	86,534,354	1	1
2	Continuously enrolled till 2019	57,025,406	0.65	0.65
3	Nondual	45,305,490	0.79	0.52

[55] https://www.ama-assn.org/practice-management/digital/ama-telehealth-policy-coding-payment

4	No move across states	40,223,900	0.89	0.46

Note: Among all Medicaid enrollees beyond study sample states. % previous refers to the percentage of remaining enrollees after additional step out of previous step enrollees and % total refers to the percentage of remaining enrollees after additional step out of initial enrollees.

Table 2-23: Steps for identifying sample claims

no	steps	total claims	telehealth claims	%previous (total)	%previous (tele)
1	All outpatient MHSUD	1,862,555,753	9,402,026 (0.5%)	1	1
2	22 States and FFS	301,116,113	881,982 (0.3%)	0.16	0.09
3	Eligible population	160,313,414	374,461 (0.2%)	0.53	0.42

Note: Sequential steps for our sample construction is shown. Percentages in parentheses are the percentage of telehealth claims over total claims. %total represents the percentage of remaining total claims in each additional step and %telehealth represents the percentage of remaining telehealth claims in each additional step.

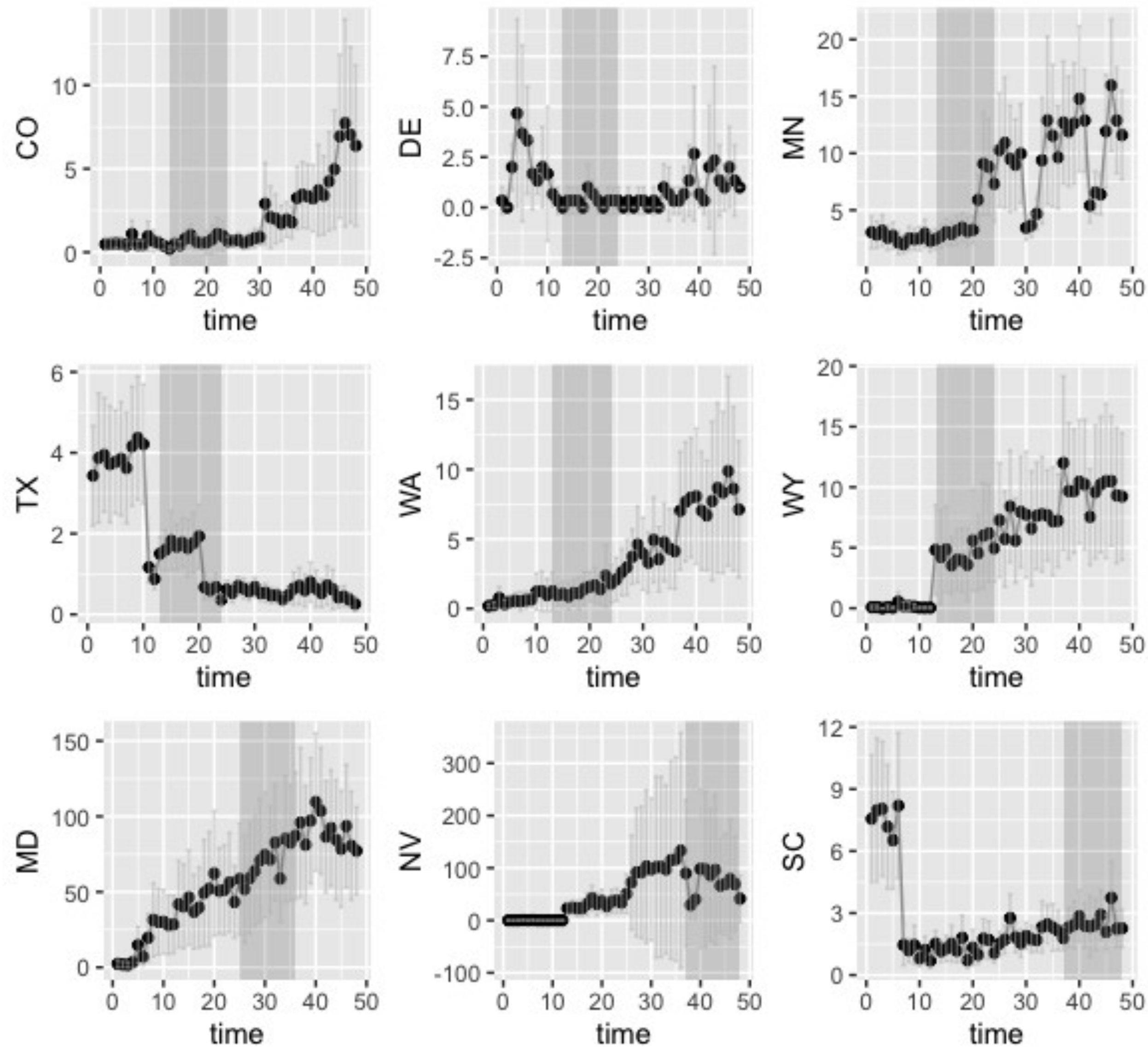
Table 2-24: Total telehealth claims in our sample between 2016-2019.

no	states	units (census)	units (sample)	telehealth claims	% remain
1	AL	67	62	2,543	0.72
2	AR	75	75	20,760	0.99
3	AZ	15	13	7,549	0.99
4	CO	64	52	4,673	0.99
5	DC	1	1	261	1
6	DE	3	3	138	0.98
7	GA	159	159	70,792	1
8	IL	102	92	8,638	1
9	IN	92	89	8,316	0.99
10	MD	24	24	64,709	0.99
11	MN	87	87	28,462	1
12	MS	82	52	1,515	0.96
13	NV	17	12	28,876	0.97
14	SC	46	46	5,613	0.98
15	SD	66	60	5,316	0.98
16	TX	254	205	14,420	0.99

17	VA	133	77	2,747	1
18	WA	39	33	5,412	0.99
19	WI	72	71	24,977	1
20	WV	55	55	11,352	0.99
21	WY	23	22	5,797	0.97

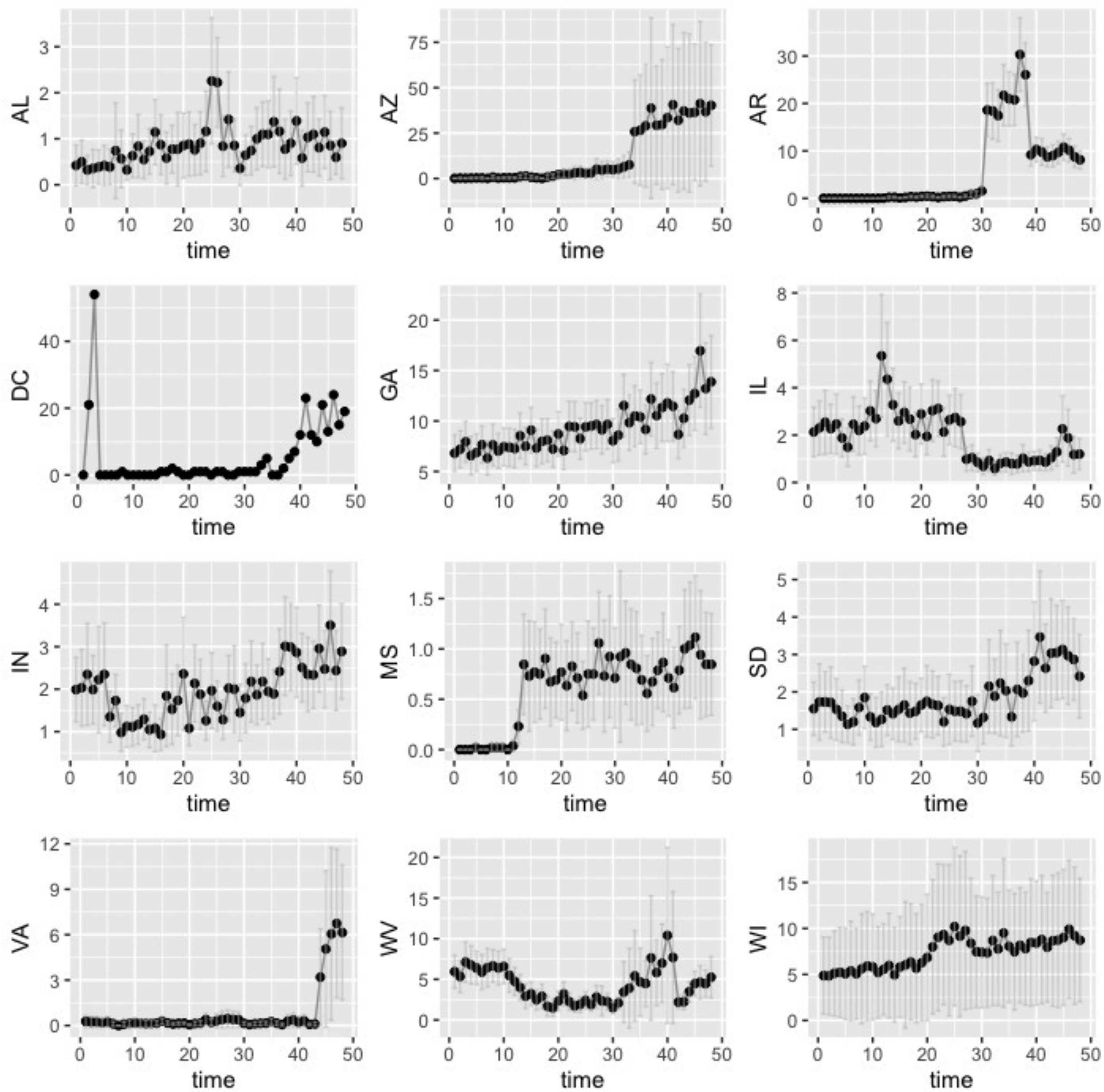
Note: Counts of units (counties) and total telehealth claims in the final sample across states. units (census) represents the number of distinct counties per state based on 2020 census report, while units (sample) are our counts of counties per state in the sample. Telehealth claims refers to total telehealth claims in our sample between 2016-2019. %remain refers to percentages of total telehealth claims after sample processing. Michigan and Hawaii are removed due to invalid county codes.

Figure 2-4: Telehealth claims per county per month across treated states.



Note. County-month average of telehealth claims (not adjusted by population) with +-2*SD in grey bar. The grey box is the year when the state lifted originating site restrictions. In our final sample, we only included MN, WA, WY.

Figure 2-5: Telehealth claims per county per month across control states.



Note. County-month average of telehealth claims (not adjusted by population) with +-2*SD in grey bar. In our final sample, we included AL, AZ, DC, IN, MS, SD,VA, WI, WV.

2.10.5 Misspecification

One potential confounder is the Federation of State Medical Boards (FSMB)'s Interstate Medical Licensure Compact (§2.10.7.3). This policy, acting as a state-level incentive to facilitate telehealth supply through easier cross-state license registration, represents a potential confounding factor. While our focal policy affects Medicaid MHSUD care financed by fee-for-service, FSMB's licensure compacts affect all medical providers with participating state licenses regardless of their specialty or whether they accept Medicaid patients. Considering this difference and particularly with relatively low reimbursement rates for Medicaid care, providers are not likely to apply for cross-state licenses in order to provide more telehealth across state borders. Despite not targeted toward care for Medicaid patients, this provider-side incentive could change the availability of telehealth providers in the region, which can affect the telehealth choice of Medicaid patients. Additionally, Zhou et al. (2021) found a positive impact on Medicare spending with the FSMB's licensure compact through actual increase in the number of state licenses per provider, which suggests controlling for this licensure compact can be important for Medicaid telehealth utilization.

For simplicity, we will henceforth use the term *Home* to denote our focal policy (i.e. removal of originating site restrictions for telehealth in medicaid) and *FSMB* to denote our potential confounding policy (i.e. the cross state licensure compact by the FSMB).

The DiD estimates adding FSMB as a control are reported in Table 2-23. There are two main observations. First, The simple DiD models (column (1), (5), (9)) show statistically significant and positive policy impacts on telehealth utilization uptake after *Home*, which is consistent with strong divergence of raw Y trends for the treated and the control following the policy (Figure 2-1). Second, *FSMB* are found not to directly affect telehealth utilization for Medicaid population (column (2), (6), (10)) even as a sole predictor. Added as a linear control, *FSMB* do not sway our *Home* coefficients at all (column (3), (7), (11)). However, when we interact *FSMB* with *Home* (column (4), (8), (12)), the results imply that most of positive *Home* effects are from interaction effects with *FSMB*. In other words, *Home* policy impacts are larger and positive with *FSMB* and smaller and even negative without *FSMB*. This pattern is found to be consistent across all outcome variables.

Before we take the second observation (i.e. Home effects switching direction with and without FSMB) seriously, we re-visit the identification of each coefficient. The main Home effects (i.e. coefficients for I(Home)), separate from the additional effects with FSMB (i.e. coefficients for I(Home):I(FSMB)) is identified from comparison of Home effects for Delaware without FSMB and Home effects for other treated states with FSMB. It means the identification depend on

Delaware, the only treated state without FSMB. However, this identification may not be reliable since Delaware has only three counties and shows a spurious spike preintervention in raw Y trend - which was interestingly the same case for all three counties. This observation led us to be cautious in interpreting the negative treatment effect for Delaware. The negative treatment effects seem to be resulting from not controlling for unobserved confounders in DE (e.g. spike in 2016), rather than a valid treatment effect. Since the interaction effects with FSMB rely on the identification of the main Home effects of Delaware, we also remain cautious in generalizing huge positive effects from Home x FSMB as well.

In order to confirm the positive effects of *Home* while controlling for FSMB, rather than relying on DE as the only state for identification of the main effects over interaction effects in the full sample, we report the main effect estimates for the simplest model (column (1), (5), (9)) using sub-sample analysis. In Table 2-24, 2-25 and 2-26, the main Home effects are estimated using (1) full sample, (2) full sample without DE, (3) subsample of states with FSMB, (4) subsample of states with FSMB controlling for years of FSMB adoption. Across all outcome variables, estimates for the main policy effects are robust across sample. Removing Delaware would increase treatment effects (i.e. (1) vs. (2)), as examined from the observed spike pre-intervetion in Delaware. FSMB does not have much effect on our Medicaid telehealth utilization, as shown in robustness in estimates between using only FSMB controls and using FSMB/no-FSMB controls (i.e. (3) vs. (2)). Adjusting for the length of gap years for FSMB implementation (MN, WY vs. WA), consistently showed that joining FSMB without gap years reduced Home effects. However, the interpretation on this interaction effects (Home x WA) should be cautious, since there is no control states with no gap years as in WA's case. Basically, it is separating out WA-specific treatment effects from MN and WY and this specific effect is not necessarily from having no gap years in FSMB.

Restricting to states with FSMB reduces our sample size without much gain in bias and precision of results. So we choose to use full sample just without Delaware. Since there are only 3 counties in DE, the treatment effect estimates do not change much by excluding DE anyways. Moreover, the treatment timing for Home policy can be all in the latter half of 2017 if we exclude Delaware.

Table 2-25: Difference-in-difference results with FSMB

	Dependent variable:											
	log(claims per 10K)				log(billed per 10K)				log(paid per 10K)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
I(Home)	0.504**		0.504**	-0.112***	1.642*		1.606*	-1.117**	1.481*		1.461*	-1.141***
	(0.099)		(0.104)	(0.010)	(0.422)		(0.447)	(0.148)	(0.402)		(0.427)	(0.071)
I(FSMB)		0.149	0.001	-0.016		0.581	0.108	0.034		0.492	0.062	-0.008
		(0.083)	(0.056)	(0.055)		(0.269)	(0.223)	(0.220)		(0.244)	(0.211)	(0.207)
I(Home):I(FSMB)				0.634**				2.806**				2.681**
				(0.111)				(0.432)				(0.439)
Observations	22,500	22,500	22,500	22,500	22,500	22,500	22,500	22,500	22,500	22,500	22,500	22,500
R^2	0.556	0.535	0.556	0.556	0.557	0.543	0.557	0.558	0.536	0.522	0.536	0.537
Adjusted R^2	0.542	0.521	0.542	0.543	0.543	0.530	0.543	0.544	0.522	0.508	0.522	0.523
Residual Std. Error	0.449	0.459	0.449	0.448	1.787	1.813	1.787	1.785	1.621	1.645	1.621	1.619

Notes. Unit = County-month aggregation of telehealth claims. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. I(Home) is the policy of interest and I(FSMB) is the confounding policy. Both Home and FSMB were implemented in 2017. Note that even though FSMB had been announced early beginning 2014-2016 for participating states in our sample, we used year of implementation as our controlling variable. In our sample states, besides WA, all FSMB states joined early (before our study time). *p<0.05; **p<0.01; ***p<0.005.

Table 2-26: Home policy effects on the number of telehealth claims

	Dependent variable:			
	log(claims per 10K)			
	(1)	(2)	(3)	(4)
I(Home)	0.504**	0.517**	0.522**	0.622**
	(0.099)	(0.102)	(0.111)	(0.126)
I(Home):Washington				−0.430*
				(0.116)
Observations	22,500	22,392	16,380	16,380
R^2	0.556	0.556	0.550	0.554
Adjusted R^2	0.542	0.542	0.536	0.541
Residual Std. Error	0.449 (df = 21839)	0.449 (df = 21734)	0.490 (df = 15889)	0.488 (df = 15888)

Notes. Regression on telehealth utilization with county and monthly time fixed effects using county-month aggregated panel data between 2016-2019 excluding the implementation year 2017. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. Model (1) refers to full model same as the column (1) in Table 2-23. Model (2) refers to the same specification but without DE in the sample. Model (3), (4) refer to the same specification but only among states with FSMB. Model (4) separates out Washington due to different gap years until FSMB implementation. Note that for all states with FSMB, besides WA, FSMB announcement occurred several years before the actual implementation, allowing enough gap years for agents to prepare telehealth adoption. *p<0.05; **p<0.01; ***p<0.005.

Table 2-27: Home policy effects on the billed amount for telehealth

	Dependent variable:			
	log(claims per 10K)			
	(1)	(2)	(3)	(4)
I(Home)	1.642*	1.700*	1.688*	1.796*
	(0.422)	(0.439)	(0.479)	(0.615)
I(Home):Washington				−0.463
				(0.560)
Observations	22,500	22,392	16,380	16,380
R^2	0.557	0.558	0.543	0.543
Adjusted R^2	0.543	0.545	0.529	0.529
Residual Std. Error	1.787 (df = 21839)	1.787 (df = 21734)	1.884 (df = 15889)	1.883 (df = 15888)

Notes. Regression on telehealth utilization with county and monthly time fixed effects using county-month aggregated panel data between 2016-2019 excluding the implementation year 2017. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. Model (1) refers to full model same as the column (5) in Table 2-23. Model (2) refers to the same specification but without DE in the sample. Model (3), (4) refer to the same specification but only among states with FSMB. Model (4) separates out Washington due to different gap years until FSMB implementation. Note that for all states with FSMB, besides WA, FSMB announcement occurred several years before the actual implementation, allowing enough gap years for agents to prepare telehealth adoption. *p<0.05; **p<0.01; ***p<0.005.

Table 2-28: Home policy effects on the paid amount for telehealth

	Dependent variable:			
	log(claims per 10K)			
	(1)	(2)	(3)	(4)
I(Home)	1.481*	1.537*	1.539*	1.572*
	(0.402)	(0.421)	(0.460)	(0.579)
I(Home):Washington				−0.139
				(0.531)
Observations	22,500	22,392	16,380	16,380
R^2	0.536	0.537	0.519	0.519
Adjusted R^2	0.522	0.523	0.504	0.504
Residual Std. Error	1.621 (df = 21839)	1.621 (df = 21734)	1.718 (df = 15889)	1.718 (df = 15888)

Notes. Regression on telehealth utilization with county and monthly time fixed effects using county-month aggregated panel data between 2016-2019 excluding the implementation year 2017. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. Model (1) refers to full model same as the column (9) in Table 2-23. Model (2) refers to the same specification but without DE in the sample. Model (3), (4) refer to the same specification but only among states with FSMB. Model (4) separates out Washington due to different gap years until FSMB implementation. Note that for all states with FSMB, besides WA, FSMB announcement occurred several years before the actual implementation, allowing enough gap years for agents to prepare telehealth adoption. *p<0.05; **p<0.01; ***p<0.005.

Table 2-29: Home policy effects from subsample with less data quality concern

	Dependent variable:		
	log(claims per 10K) (1)	log(billed per 10K) (2)	log(paid per 10K) (3)
I(Home)	0.470**	1.526*	1.363*
	(0.101)	(0.446)	(0.422)
Observations	17,136	17,136	17,136
R^2	0.559	0.543	0.517
Adjusted R^2	0.545	0.529	0.502
Residual Std. Error (df = 16624)	0.448	1.852	1.676

Notes. Unit = county-month aggregation of telehealth claims. Regression results with county and monthly time fixed effects between 2016-2019 excluding the implementation year 2017. Study states satisfy: 1) joining FSMB early before the study time (i.e. excluding WA, DE from treated group), 2) high data quality in expenditure variables according to DQ Atlas (i.e. less than 20%

divergence from CMS 64 expenditure data in all outpatient claims). These include 2 treated (MN, WY) and 4 control (AL, MS, SD, WI) states. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; ***p<0.005.

2.10.6 Measurement error

One potential source of the measurement errors is on policy timing, both for focal policy and the main confounding law. We only know the policy in year, but not in month. In our main results, we report the model trained without the intervention year of 2017 (i.e. thus 2016 as pre-intervention and 2018-19 as post-intervention periods) to abstract away from uncertain policy timing in month. In this section, we delineate the issues on lack of granular policy timing information and report the estimates across different assumptions on policy timing and across different subsets of sample with less concerns on measurement errors in policy timing.

Regarding the focal policy timing, we rely on the bi-annual reports from the CCHP on Medicaid telehealth reimbursement laws (CCHP (2017)). However, there are lack of information on the granular effective dates for the policy change in these reports. Even though we showed results without intervention year in our main analyses, we can report estimates including the intervention year and using release months of the CCHP reports, which is in April and October as our proxy for policy starting months.

Next, regarding the timing of the main confounding laws - FSMB, state licensure compacts have been devised in early 2010s, however only in 2017 April, participating states could start enrollment. This engenders two big issues in estimating causal impact of our focal policy. First, in April 2017, three out of four treated states (MN, WA, WY) started implementing FSMB and removing originating site restrictions at the same time. Thus, the separate policy impact comes from heterogeneous treatment effects which is in Delaware compared to other three treated states. This means, the estimation depends on 138 telehealth claims (0.2% of telehealth claims in 4 treated states) in Delaware. Second, since there was a delay in the policy implementation after announcement, states might have different degrees of readiness in adopting telehealth. For example, states announced early in 2015 to participate in FSMB would have incentivized its providers to be more ready for cross-border telehealth services during 2 years of delay, compared to the states that announced and joined later after 2017. Thus, controlling for implementation date without considering early announcement years before might not be enough for removing FSMB impacts which could have been underlying before our focal policy.

Table 2-30: Inclusion of intervention year 2017

	Dependent variable:		
	log(tele-claims per 10K) (1)	log(tele-billed per 10K) (2)	log(tele-paid per 10K) (3)
I(Home)	0.394**	1.206*	1.098*
	(0.115)	(0.432)	(0.400)
Observations	30,000	30,000	30,000
R^2	0.535	0.536	0.532
Adjusted R^2	0.525	0.525	0.521
Residual Std. Error (df = 16624)	0.440	0.440	1.803

Notes. Unit = county-month aggregation of telehealth claims. Regression results with county and monthly time fixed effects between 2016-2019 *including* the implementation year 2017. Reported standard errors in parentheses are not yet cluster-adjusted. Statistical testing is based on the adjusted standard errors - clustered at policy implementation level, which is state and year. *p<0.05; **p<0.01; ***p<0.005.

2.10.7 Other telehealth policies

2.10.7.1 Medicare

In Medicare, which covers 16% of the U.S. population, telehealth was introduced as early as in 1996. The initial telehealth policy was very limited in many aspects. It was restricted to rural and Health Professional Shortage Areas (HPSAs) and for consultation only. The rationale at the time to limit Medicare coverage to rural areas with provider deficits was that this population would benefit the most while limiting the risk of unnecessary use. Moreover, the originating sites (i.e., where the patient is during care) were limited to practitioner's offices, hospitals, critical access hospitals, rural health clinics, and federally qualified health centers in rural health professional shortage areas, not allowing patients getting care at home. In this case, two practitioners were required - one in a distant site (i.e., a provider connected virtually to distant patients) and the other in an originating site (i.e., provider assisting patients while on telehealth visits), who split the single payment at 75% and 25%, respectively.

In 2001, the Benefits Improvement and Protection Act (BIPA) removed provider presence requirement in originating sites and increased payment by Medicare - equal payment as in face-to-face care for distant practitioners and additional facility fees for originating sites. In net, the policy shifted from requiring two providers for one single payment to one provider with two separate payments. Small additions have followed since; however, none have changed the program significantly, particularly around patient location. Only recently, Medicare lifted geographical restrictions for certain conditions. In 2019 January, Bipartisan Budget Act (P.L. 115-123) allowed home to be eligible for end-stage renal disease and expanded the coverage to urban areas for end-stage renal disease and strokes. Focusing on opioid recovery and treatment, Support for Patients and Communities Act (H.R.6 115-271) went in effect in late 2019, to allow telehealth at home for substance use disorder and mental health only if they are co-diagnosed.

Over time, the use of telehealth among Medicare FFS beneficiaries steadily increased but the rate of adoption has been still limited. In 2009, the beneficiaries made about 38,000 telehealth visits - an increase from 26,000 in 2006 (Gilman and Stensland, 2013). In 2016, almost 90,000 Medicare FFS beneficiaries utilized just over 275,000 telehealth services. It was still a small fraction of total Medicare FFS beneficiaries (0.25%) and spending (0.01%) (CMS, 2018). Across care types, 85% of telehealth users in Medicare were mental health patients and less than 10% were stroke patients in 2016.

2.10.7.2 Private Insurance

Commercial payers cover more than half of the U.S. population, mostly under the employer-insurance. State governments regulate telehealth of commercial payers with parity laws which require commercial payers to cover services provided through telehealth to the same extent as those services are covered in person. Since 1995, many states have been enacting telehealth parity legislation, with its number starkly increasing after 2012. As of October 2019, 42 states and D.C. have passed some private health parity statutes (Lacktman et al., 2019). There are variations across states, with full coverage parity more dominant than full reimbursement parity. In 2019, 32 states and D.C. enacted coverage parity while 16 states enacted reimbursement provision[56], leaving only 15 states with full parity in both coverage and reimbursement rates (Lacktman et al., 2019).

Dong (2022) examined how telehealth parity laws on commercial payers affected healthcare expenditure at the state-year level. Deploying difference-in-differences under the staggered implementation of parity laws, he found that parity laws (either coverage or payment) decreased total healthcare expenses per capita by 3.9% in state-year. Furthermore, he showed that there were steady effects till the next four years up to 6.8% reduction. He found heterogeneous effects across parity types (i.e., reduction mostly by coverage parity, offset by cost-shifting and no effect from payment parity), though he did not explain fully. He speculated that reduction in healthcare expense in general might have come from health improvement from telehealth or different price negotiation results between insurers and providers.

[56] Limited coverage or reimbursement are excluded.

The commercial parity laws were reported to have spillover effects on Medicare telehealth utilization which has been notoriously limited in coverage. Comparing states with and without commercial telehealth parity laws, Mehrotra et al. (2016) showed that telehealth visits for Medicare FFS enrollees are higher in states with parity laws than in states without ones.

2.10.7.3 Cross-border Licensing

Widespread telehealth adoption can be hampered if physicians have no portability with their state licenses. In 2012, American Telemedicine Association (ATA) pointed out that cross-border physician licenses were the biggest barrier to telehealth[57]. A typical license application takes four to twelve weeks and sometimes seven months, with a single application fee of $500[58]. Not only that, the application process is very different across states. First, all providers must complete the United States Medical Licensing Examination before applying to a certain state's medical board. Then each state medical board may ask for idiosyncratic requirements, including citizenship requirements, educational requirements, FBI criminal background checks, in-person interviews, board certification, and assessments of mental and physical health. Even after these, state medical boards still have complete discretion on license issuance (Zhou et al., 2021).

To circumvent these hassles while encouraging telehealth, some state medical boards issued telehealth-specific licenses or certificates. These licenses could allow an out-of-state provider to render services via telehealth in a state where they are not located or allow a clinician to provide services via telehealth in a state if certain conditions are met, such as agreeing that they will not open an office in that state[59]. The number of states with telehealth licenses remained fairly consistent, though some repealed or/and adopted them. When states did not have specific licenses, sometimes they made restrictive exceptions, such as allowing cross-border care in border counties while limiting the frequency or types of cross-border interaction.

The comprehensive shift in cross-border licensing came only after telehealth licensure compact. Telehealth licensure compacts are agreements between states on how to address licensing of practitioners with out-of-state health licenses. In 2013, the Federation of State Medical Boards proposed the compact which could streamline the traditional application processes. With this compact, providers in member states are qualified to practice medicine across state lines as long as they hold a full, unrestricted medical license in at least one Compact member state[29]. This can

---

[57] The American Telemedicine Association, "Physician Licensure Barriers to 21st Centry Healthcare", 2012.

[58] Medicus Healthcare Solutions, "Physician Licensure Application Fees and Timelines by State", 2019.

[59] The Center for Connected Health Policy Report 2014-19. "State telehealth laws and Medicaid reimbursement policies". As of 2019, there are 10 states issuing telehealth special licenses - AL, LA, MT, NV, NM, OH, OK, OR, TN, TX [29]Interstate Medical Licensure Compact, https://www.imlcc.org/a-faster-pathway-to-physician-licensure/

reduce the time for application from several months to 19 days to acquire all the compact state licenses. In 2014, the compact was renamed the Interstate Medical Licensure Compact (IMLC) and states gradually announced their intention to join after 2015 with actual implementation in 2017.As of 2019, 29 states and the District of Columbia adopted IMLC. Among these, four states (Alabama, Maine, Nevada, and Tennessee) joined the Compact on top of their own telehealth licenses.

Zhou et al. (2021) studied the effect of IMLC on provider licensing and Medicare payments in physician-quarter level using staggered entries by states and difference-in-differences design. They showed that with the compact, physicians increased the number of licenses (1.5%), the number of Medicare services (1.6%), the number of Medicare patients received (1.4%) and total Medicare payment (1.1%). They found competitive effects between urban and rural providers. Urban providers financially benefit by expanding service to a wider geographic market, while rural providers experienced declining patient volume and revenue loss.