The Impact of Manipulated CDS Algorithm on Opioid Prescribing Decision *

Xuelin Li† **Meizi Zhou**‡

This version: November 2025

Abstract

We document that interactions with manipulated Clinical Decision Support (CDS) systems can induce not only short-term but also long-term changes in physicians' opioid prescribing behavior. Physicians in our sample adopted electronic health record software from a list of federally certified vendors in 2011. Between 2016 and Spring 2019, one vendor secretly embedded a biased CDS function designed to promote extended-release opioid sales. Affected physicians not only increased opioid claims relative to the control group during the treatment window but also maintained a higher propensity to prescribe opioids even after the biased function was removed. This long-term behavioral change persisted even after affected physicians moved to new locations, changed their affiliations, or faced stricter state-level opioid regulations. Increasing physician awareness helped mitigate this impact. Using machine learning algorithms, we estimate that decision-making distortion accounts for approximately 54% of the treatment effects in a physician decision model with dynamic learning.

Keywords: manipulated algorithm, clinical decision support systems, long-term impact, opioid, multi-arm causal forest

*We are grateful to Tania Babina, Nick Bloom, Carri Chan, Josh Lerner, Hong Luo, Alex Moehring, Abraham Seidmann, Laura Veldkamp, conference participants at 2024 NBER Productivity, Innovation, and Entrepreneurship Program Spring Meeting, UT "Smokey" Mountain Finance Conference, Conference on Health IT and Analytics (CHITA), Emory "AI and Platform Evolution" Conference, and seminar participants at Boston University, Fudan University, NYU Shanghai, and Shanghai University of Finance and Economics for helpful comments and suggestions. An earlier version of the paper is circulated as "Can AI Distort Human Capital?"

†Columbia University, Graduate School of Business - `xuelin.li@columbia.edu`, corresponding author.

‡Boston University, Questrom School of Business - `mzzhou@bu.edu`

1 Introduction

Clinical decision support (CDS) systems, integrated with electronic health record (EHR) platforms, employ algorithms that analyze patient data to assist physicians in making personalized medical decisions, ranging from prescribing medications to ordering diagnostic tests and recommending treatments. CDS systems play a central role in modern clinical practice and have been shown to substantially improve patient safety, reduce medication errors, and enhance adherence to evidence-based guidelines. They are now pervasive across healthcare delivery: for instance, Jing et al. (2019) document utilization rates ranging from 71.8% to 100% in primary care settings. However, the increasing algorithmic complexity and opacity of CDS tools make it difficult for both regulators and clinicians to assess their validity and implications for care. In the short term, physicians differ in how they adjust their decisions in response to CDS recommendations, due to factors such as trust, time pressure, or institutional policies. More importantly, it remains unclear whether such reliance merely shapes short-run choices or instead induces persistent shifts in physicians' clinical judgment and reasoning over time.

In this paper, we highlight a unique risk factor, arguing that professional decisions can be distorted through interactions with decision support systems. We hypothesize that professional decisions adapt through collaboration with decision support systems, influencing not only current decisions but also shaping workers' habits for similar future tasks. However, this "learning-from-algorithm" phenomenon poses risks, as algorithmic predictions may contain errors, biases, or even intentional manipulation. In a worst-case scenario, our empirical setting demonstrates situations where developers deliberately manipulate these predictions. We show that not only are professionals with expertise vulnerable to such manipulations, but that algorithm-generated mistakes can also be reinforced through user habit formation. As a result, merely removing or correcting the biased algorithms cannot fully eliminate the biased practices of the users.

Empirically documenting this effect presents several challenges. First, pinning down the specific influences of a decision support function on human behaviors is difficult due to substantial heterogeneity in human-system interactions across tasks, agents, and algorithms (e.g Kleinberg et al., 2018; Angelova et al., 2023). Even in similar tasks, a single worker may face various recommended decisions across scenarios, making it challenging to identify alterations in habits consistently. Second, an essential step in this study is to obtain the long-term counterfactual of worker behaviors without further algorithm influence after their initial interaction. However, workers endogenously adopt and abandon decision support systems based on their recent experiences. As decision support systems are prevalent in many working scenarios, the system adoption can be an irreversible decision and workers rarely suspend usage, particularly due to exogenous reasons.

The background of our study is that the Health Information Technology for Economic and Clinical Health (HITECH) Act created financial incentives for physicians to use EHR software from a list of certified companies in 2011. However, from July 2016 to the Spring of 2019, one major EHR company, Practice Fusion, secretly embedded a pain clinical decision support ("Pain CDS") functionality in its EHR to promote the sales of Purdue Pharma's extended-release, also known as long-acting (LA), opioids in an "unbranded effort." This setting provides the following empirical advantages. First, the decision assistance (i.e., Pain CDS) was manipulated to favor opioid prescriptions regardless of medical necessity and potential abuse risks, but it appeared to healthcare providers as unbiased medical information. This homogeneity allows us to measure an intuitive and single-direction impact on human behaviors. Second, our identification strategy follows a simple difference-in-differences (DID) design, leveraging physician-level annual opioid prescription data from 2013 to 2021. Treated physicians adopted Practice Fusion in 2011, and we can compare their behaviors to other incentive program participants after the shock. Most importantly, our sample spans the period when the bias alert was removed (i.e., 2019 – 2021) so that we can evaluate the long-term distortions on prescribing decision.

We first confirm that in the treatment window from 2016 to 2018, physicians having adopted Practice Fusion exhibit a significantly higher propensity of prescribing LA opioids, compared to the control group. We characterize this propensity through various measures, such as the number of claims, drug costs, and days of supply. The coefficients are not only highly statistically significant (t-statistics ranging from 3.9 to 4.5) but also economically meaningful. Affected physicians had 5.9% more annual LA claims compared to the control group, resulting in an aggregated cost of approximately $2 million per year in our sample. These findings suggest that even experienced professionals tend to follow biased CDS recommendations and fail to detect intentional manipulations embedded in the system.

If the findings were purely driven by the transient manipulation, we would not observe any significant differences between the treatment and control groups after the removal of the CDS manipulation. However, in the post-treatment window from 2019 to 2021, we document evidence of long-term behavioral changes across all different measures: affected physicians continue to prescribe 11.6% more LA opioid claims relative to the control group. We also estimate the coefficient dynamics across all calendar years to characterize the time series of impacts. Supporting the parallel trend assumption, the two groups exhibit no significant differences before 2015. When the pain alert was active, affected physicians gradually increased LA opioid usage over the years. After 2019, this trend flattens, but the treatment group still maintains a significantly higher propensity for opioid prescriptions. In addition, we confirm that our treatment effects remain significant after adjusting for potential violations of the parallel trend assumption following Rambachan and Roth (2023). Besides, we also validate that our results are not confounded by log transformations using Poisson regressions, non-transformed count outcomes, and ratio-based measures.

One alternative explanation for this long-term persistence is a demand-based story: patients with addictions may continue visiting the affected physicians and requesting prescriptions after the treatment window. To address this concern, we first document that the number of unique patients receiving LA opioids increases significantly for affected physicians, while the average supply per patient remains unchanged. This pattern suggests that the higher total volume of LA opioids prescriptions is not driven by higher doses among potentially addicted patients, but rather by prescribing to a broader group of new patients. Second, we focus on the subsample of physicians who moved to new locations or changed their affiliations from 2019 to 2021, and therefore interacted with new patients. In this subsample, our estimated coefficients remain significant and are similar in magnitude to the baseline estimation.

To establish the generality of our results, we also utilize a claim-level database covering 43 million privately-insured individuals. Consistent with the findings from Medicare beneficiaries, a 1% increase in the number of Practice Fusion adopters in an MSA is associated with 0.15% more expenditure on Purdue Pharma's products in the short-term, and 0.53% more over the longer term. These results remain significant if we just focus on first-time opioid users, further alleviating the demand-based concern. Consistent with Practice Fusion's manipulation strategy, we also document significant impacts on patients both with and without recent chronic pain diagnoses. Given that most Purdue Pharma's products are approved for prolonged pains, there exist abuse and addiction concerns for patients without these conditions. Indeed, in each 3-digit zip code area, a 1% increase in the number of Practice fusion adopters is associated with 0.48% more Medicare beneficiaries requiring treatment for opioid abuse from 2020 to 2021.

How can the negative impacts of manipulated CDS be mitigated? We examine two potential channels: state regulations and physician awareness. State-level opioid control policies, including supply limits and Prescription Drug Monitoring Programs, did not meaningfully reduce the biased CDS impacts, as our estimates remain similar or even slightly larger in regulated states. These results suggest that compliance-based policies are insufficient when algorithmic biases directly shape physicians' beliefs at the point of care, as physicians can override guidelines based on their own judgment of medical necessity. In contrast, greater awareness appears more effective: using variation in public attention to the 2020 Practice Fusion settlement, we find that physicians in high-awareness states exhibit significantly smaller long-term effects, indicating that transparency and dissemination of algorithmic risks are key to offsetting habit-driven persistence.

As robustness checks, we first document that affected physicians do not have financial incentives to prescribe more opioids in the long term through detailing. In addition, we show that our results remain consistent across alternative control groups, clustering methods, and placebo tests using regulated medications not promoted by Practice Fusion.

Lastly, we use a machine-learning algorithm built upon Nie and Wager (2021) to estimate the economic magnitude of long-term behavioral changes within our conceptual framework. This algorithm uses the prescription probability of LA opioids over all claims as the outcome variable. During 2016 to 2018, it can decompose the aggregate treatment effects for each affected physician into components attributed to direct CDS manipulation and long-term belief distortions. In the data, a typical treated physician has a higher prescription probability of 2.3 bps, or 9.5% of the unconditional average. The predicted prescription probability change by our algorithm closely matches this magnitude, with an average treatment effect of 2.4 bps. We generate counterfactual prescription probabilities assuming no long-term change or that change does not accumulate after more interactions with biased CDS. The treatment effect will decrease by 54% (to 1.1 bps) in the first case and by 17% (to 2.0 bps) in the second case. These magnitudes underscore the importance of habit formation in explaining the observed prescription changes. Our algorithm also allows us to study the heterogeneity of treatment effects by estimating them as a function of observed characteristics. We hypothesize and confirm two channels through which affected physicians are more robust to CDS biases. First, physicians with more previous experience in LA opioid usage are more likely to make decisions based on their expertise, and less likely to be influenced by CDS recommendations. Second, our algorithm shows that average patient age of each physician has the highest feature importance in determining treatment effect heterogeneity. Intuitively, physicians who interact with senior patients are more cautious about the potential opioid side-effects, and therefore less likely to demonstrate negative impacts.

2 Related Work and Conceptual Framework

2.1 Related Work

Decision Making with Algorithmic Recommendations Our work contributes to the literature on the role of algorithmic decision support systems in healthcare, particularly in shaping physicians' real-time decisions and treatment outcomes (Agarwal et al., 2011). A substantial body of research has demonstrated that healthcare IT systems, particularly EHRs, can lower costs, reduce medical errors, improve care quality, and facilitate better information sharing across providers (Atasoy et al., 2018; Bardhan and Thouin, 2013; McCullough et al., 2010; Bardhan et al., 2023). Within EHR systems, CDS tools are important modules and have shown considerable benefits in improving clinical decision-making. Empirical studies have found that CDS adoption can reduce adverse medical events (Hydari et al., 2019), mitigate racial disparities in care delivery (Ganju et al., 2020), and steer physicians away from unnecessary, high-cost orders (Doyle et al., 2019). These benefits highlight the potential of algorithm-augmented systems to improve both efficiency and equity in healthcare settings.

However, regulators and healthcare organizations have remained cautious about broader algorithmic adoption because of safety and accountability concerns. For example, it was not until 2025 that CMS began reimbursing AI-powered ECG analysis under Medicare, despite the technology having been available for years.[1] Moreover, algorithms can amplify existing biases embedded in their training data. In healthcare, for instance, algorithms that use insurance costs as proxies for patient health needs have been shown to systematically underestimate the needs of Black patients, reflecting structural inequities in access to care (Obermeyer et al., 2019). Relatedly, Samorani et al. (2022) document that state-of-the-art scheduling systems cause Black patients to wait about 30% longer than non-Black patients. More broadly, Ahsen et al. (2019) show that when training clinical decision systems with human-generated signals, such as radiologists' assessments, the model inherits systematic human biases, which can ultimately degrade diagnostic accuracy. We complement this literature by documenting that firms can intentionally manipulate algorithmic outputs to exploit users' trust and achieve strategic objectives. Even trained professionals may find it difficult to detect such manipulation in real time, highlighting an additional risk aspect of algorithmic decision support.

Post-Algorithmic Effects on Human Habits Most empirical work focuses on when algorithmic

[1] For details, see https://cardiologyinnovation.com/artificial-intelligence/ai-based-diagnosis/cms-includes-ai-ecg-technology-2025-reimbursement-schedule/.

recommendations are presented. Much less is known about the long-term effects of algorithms on human decision making, particularly when individuals later make similar decisions without algorithmic input. Acemoglu (2021) theorize that prolonged reliance on decision support can erode workers' understanding of partially automated tasks, ultimately reducing productivity in areas that still require human judgment. Macnamara et al. (2024) similarly argue that AI assistants may accelerate skill decay among experts and make it harder for them to recognize these deleterious effects.

Empirical evidence remains scarce and is mostly based on small lab or short-run field experiments. For example, in a six-month study with 19 endoscopists, Budzyń et al. (2025) find that cancer detection worsened once AI assistance was removed. In education settings, Bastani et al. (2025) show that AI-based tutoring improves performance during practice, but students who rely on the technology underperform when access to AI is subsequently removed in a one-semester experiment at a Turkish high school. Kosmyna et al. (2025) find that LLM-assisted essay writing leads AI users to recall less of their work and exhibit lingering underactivation in key brainwave bands in a 54-participant lab experiment. Our work contributes to this emerging perspective by providing large-scale evidence on how repeated interaction with decision support systems can reshape professional habits over time.

2.2 Conceptual Framework

To develop our main hypotheses, we begin with a conceptual framework in which physicians make opioid prescribing decisions by forming beliefs about the necessity of treatment. A detailed version of this framework is provided in Appendix Section C. These beliefs are shaped by two components: a prior, based on observable patient characteristics, and an update, triggered by pain-related alerts from the CDS. In addition, there is a dynamic learning process whereby repeated interactions with the manipulated CDS gradually shift physicians' habits toward prescribing more opioids, thereby affecting their benchmark prior across different periods.

This framework yields two testable hypotheses, aligned with the two strands of literature discussed in Section 2.1. First, during the treatment window, affected physicians will prescribe more LA opioids. For instance, by 2018 (the third treatment period), a treated physician experiences both direct manipulation and habit reinforcement resulting from two years of exposure to biased pain alerts. Second, in the post-treatment window, even after the removal of the pain CDS, affected physicians will continue to prescribe opioids at elevated rates due to habits formed during prior exposure.

It is worth noting that, empirically, it remains an open question whether we will find null results for the first hypothesis. On the one hand, individuals may exhibit automation bias and blindly follow algorithmic recommendations (Skitka et al., 1999). Logg et al. (2019) also find that individuals display algorithm appreciation, particularly when making numeric estimates or forecasting subjective outcomes. On the other hand, a large body of research documents persistent algorithm aversion, especially among physicians, who often distrust or discount algorithmic recommendations in clinical settings. This aversion may arise from a general disinclination toward algorithms, discomfort with their opacity (Lebovitz et al., 2022), a sense of professional responsibility (Wang et al., 2024), or reputation concerns (Liang and Xue, 2022). Professional experts such as radiologists do not fully capitalize on the potential benefits of AI assistance because of deviations from Bayesian belief updating: they tend to underweight algorithmic information relative to their own signals, and thus AI support does not necessarily alter how they acquire or interpret prior information (Agarwal et al., 2023). To confirm this "first stage," we begin by examining short-term responses, testing whether affected physicians prescribe more long-acting opioids relative to the control group during the treatment window.

Testing the second hypothesis has direct implications for quality of care in the age of algorithmic support, as physician habits and beliefs jointly shape clinical judgment alongside algorithmic recommendations. For example, Jussupow et al. (2021) and Yin et al. (2025) both highlight the importance of physicians' beliefs to detect incorrect AI advice. Our findings contribute to this discussion by addressing one of the key challenges identified in the "human-centric" digital health perspective by Bardhan et al. (2025): how to effectively combine data- and analytics-based information with physicians' intuition. Empirically, we document the existence of long-term habit persistence by comparing the treatment and control groups after the removal of the biased CDS, and examine whether this behavioral bias can be mitigated through regulatory interventions or greater physician awareness.

3 Institutional Details and Data

3.1 Practice Fusion

Our empirical setting focuses on Practice Fusion's manipulation in its CDS. Practice Fusion was an EHR company founded in 2005. At one point, it was named the top EHR for customer satisfaction among primary care providers and ranked No. 1 for value among ambulatory professionals. Investors including Kleiner Perkins Caufield & Byers made a $70 million Series D investment in

September 2013. This new round valued the company at around $700 million, making it one of the largest digital health startups at the time.[2] Practice Fusion's EHR provides clinical decision support from various aspects, including generating differential diagnoses based on patient symptoms, sending notifications for necessary tests during workflows, and issuing prescribing alerts for drug history, drug-drug interactions, allergies, compliance, and drug dosage. Practice Fusion claimed that the EHR continuously updates with regulatory changes through policy exports and provider feedback, and clinicians do not need to take any action as the EHR automatically receives updates.

Beginning in Fall 2013, Practice Fusion solicited remuneration and negotiated with a pharmaceutical company "Pharma Co.X," later identified as Purdue Pharma, to create and embed a pain clinical decision support ("Pain CDS") functionality in its EHR.[3] The purpose is to increase prescriptions for Purdue Pharma's three extended-release opioid products. During their communication, Purdue Pharma expressed concern that "providers are hesitant about using high dosages to combat pain for a variety of reasons, mostly political pressure." In response, Practice Fusion proposed developing a Pain CDS to initiate these products and embedding Purdue Pharma's drugs as treatment options in an "unbranded effort."

The two companies signed the official contracts on March 1, 2016, in which Purdue Pharma agreed to pay Practice Fusion approximately $1 million. The Pain CDS functionality went live on July 6, 2016, providing a series of alerts and recommendations for both diagnosis and treatment. Blindly recommending opioids to every physician for all patients would likely trigger suspicions. Instead, Practice Fusion designed a two-step system to rationalize the treatment and induce the prescription. First, it recommended providers to document a pain score for the patient. Second, it displayed a "Brief Pain Inventory" with the patient's previous pain scores within the previous three months. Then the provider would go through a list of questions on the severity and impact of the patient's pain, summarizing the patient's current pain as "worst," "on average," and "least" in the previous 24 hours. If the patient reported a pain score of four or higher twice within four months, or the patient was diagnosed with chronic pain after completing the Brief Pain Inventory, the CDS utilized a drop-down menu of options for pain management. Practice Fusion added an "Opioid Therapy" treatment option without considering the patient's condition. Against the boxed warnings on the drug labels, LA opioid products are listed for patients with less than severe or non-chronic pain, and even in cases where the pain could be treated with non-opioid options.

[2] "Practice Fusion Lands A Whopping $70M To Bring A Big Data Cure To The Healthcare Crisis," TechCrunch, September 2013.

[3] All details and quotes in this section, unless otherwise specified, are from Exhibit C "Statement of Facts" of Case 2:20-cr-00011-wks, United States Attorney's Office for the District of Vermont.

This treatment option was combined with other options in the list such as non-opioid analgesics and pain-management, which were sourced from a 2016 New England Journal of Medicine article titled "Opioid Abuse in Chronic Pain — Misconceptions and Mitigation Strategies" (Volkow and McLellan, 2016), although the intention of the paper is not to provide a clinical instruction in chronic pain management.

On December 14, 2016, Practice Fusion conducted a presentation at the headquarter of Purdue Pharma. The meeting revealed that through November 30, 2016, the CDS had alerted during 21 million visits involving 7.5 million patients, generating a general shift toward LA opioids particularly from immediate-release opioids. The CDS alerted more than 230,000,000 times from July 6, 2016 to Spring 2019, in which it was suspended right before Purdue Pharma filed for bankruptcy.

3.2 EHR Adoption

Our data construction starts with a granular record of eligible physicians adopting different EHR products since April 2011, downloaded from the HealthIT.gov website. In 2009, Congress passed and President Obama signed the Health Information Technology for Economic and Clinical Health (HITECH) Act, encouraging the adoption of health IT to improve the quality and efficiency of care. The Act created incentive programs that offer physicians financial benefits for the meaningful use of EHRs. To start, the CMS and the Office of the National Coordinator for Health Information Technology (ONC) have established standards for EHRs to become certified. Practice Fusion's EHRs were certified in 2011. Eligible professionals (EPs), including most U.S. physicians, can adopt certified EHR technology and attest to fulfilling meaningful use criteria to receive incentive payments. Each physician must achieve three stages with different criteria to receive full payments.[4]

We focus on the Medicare EHR Incentive Program administered by the CMS from 2011 to 2016, providing a maximum of $44,000 in total payments across years.[5] A participating EP must demonstrate meaningful use every year to receive payments, and the CMS publicly releases a complete record of attestation information, enabling researchers to track each EP's EHR products over time. The HealthIT.gov streamlines and disambiguates the raw data into a single "EHR Products Used for Meaningful Use Attestation" file. In each attestation year, this file lists all the certified EHR products attested by an EP (identified by the NPI). For each record, it supplements the provider zip code, specialty, vendor name, product version and product setting. The initial

[4]These stages include data capturing and sharing (stage 1), advanced clinical process (stage 2), and improved outcomes (stage 3).

[5]Alternatively, EPs can participate in the Medicaid EHR Incentive Program run by every state. However, there is no centralized disclosure platform for Medicaid programs. But they can only participate in one program.

sample contains over 1.8 million records for almost 360,000 NPIs using 1,232 EHR products owned by 730 vendors.

We impose the following major restrictions on the initial sample. First, certified EHR products span various application categories, including ambulatory, cardiology imaging, financial decision support, human resources and IS security, etc. All Practice Fusion products are ambulatory EHRs. To maintain comparability, we drop all the EHR products without any ambulatory applications. There are over 1 million adoption records of the remaining 190 products by roughly 210,000 physicians. Second, we focus on the regions (3-digit zip code areas) with non-zero physicians adopting Practice Fusion products to alleviate the confounding effects due to heterogeneous healthcare conditions across regions. In these regions, we include only physicians with the same specialty as the Practice Fusion adopters in the control group to account for practice differences across various medical settings. Lastly, we drop physicians using four vendors with other misbehaviors in the Incentive Program, including Athenahealth, Nextgen Healthcare, Modernizing Medicine (ModMed), and Greenway Health. Their violations include missing the required functionality to become certified and paying physicians to falsely attest.[6] Although none of these vendors were implicated in opioid promotion or kickbacks from opioid manufacturers, their documented noncompliance raises concerns about their suitability as valid controls, as their products may have distorted clinical or reporting behaviors. The final sample has roughly 27% (57,138) of the remaining physicians following the first step. To sum up, these physicians participate in the Medicare EHR Incentive Program and use EHRs with ambulatory functions. They belong to either the treatment group, adopting Practice Fusion products, or the control group, comprising comparable neighbors of Practice Fusion users. Practice Fusion EHRs are primarily used in ambulatory (outpatient) care settings, such as physician offices, ambulatory surgery centers, and hospital outpatient departments. Because we match physicians using EHRs with comparable functionalities, Appendix Table B.1 confirms that over 60% of physicians in our sample specialize in family medicine or internal medicine, which are specialties central in opioid prescribing. For example, Levy et al. (2015) show that primary care clinicians were responsible for nearly 50% of all dispensed opioid prescriptions between 2007 and 2012, and that opioid prescribing grew more rapidly among these clinicians than among other specialties.

[6]The only exception is that ModMed solicited and received kickbacks from Miraca Life Sciences Inc. (Miraca) in exchange for recommending Miraca's pathology lab services. However, we cannot perform the analysis based on this case since we do not have utilization data of Miraca's services.

3.3 Opioid Prescription

We then link the above sample to the "Medicare Part D Prescribers by Provider" dataset by the CMS, containing annual prescription information by individual physicians under the Medicare Part D Prescription Drug Program from 2013 to 2021. Relevant to our study, it provides the total number of claims, costs, and days of supply of opioid drugs each year. Besides, it provides the same information for long-acting opioids. These are drugs formulated to release more gradually into the bloodstream with a longer duration of analgesic action. Purdue Pharma's dominant ERO products, e.g., OxyContin, are all long-acting (LA). We then back out the short-acting (SA) opioid claim amount and also utilize antibiotic and antipsychotic drugs for placebo tests. As these drugs are not promoted by Practice Fusion, the treated physicians should not exhibit significant increases in SA opioid prescriptions throughout the sample.

We include a few patient characteristics as control variables. These include the average age of beneficiaries, the percentage of male, African American, and Hispanic beneficiaries, the percentage of beneficiaries qualified for both Medicare and Medicaid benefits, and the average risk score of beneficiaries. The last variable is a hierarchical condition category developed by the CMS based on health-influencing factors.[7]

Apart from manipulating the CDS, Purdue Pharma's primary marketing strategy involves directly detailing to physicians. Indeed, since Purdue Pharma introduced OxyContin in 1996, it aggressively engaged in detailing to promote the product, leading to the decades-long opioid crisis (Van Zee, 2009; Alpert et al., 2022). We obtain Purdue Pharma's detailing information from the CMS Open Payments database. This database collects and reports financial relationships between drug and medical device companies and providers since 2013, involving detailing payments such as research, meals, travel, gifts or speaking fees. We record whether physicians in our sample receive payments from Purdue Pharma in a particular year.

Table 1 Panel A summarizes the main variables in the whole sample (all three phases). 5,401 unique physicians, i.e. more than 9% of the sample, belong to the treatment group. On average, each physician prescribes 13.55 annual LA claims, supplying 382 days of usage and costing around \$2,600 for all the beneficiaries. LA claims account for 4.26% of all the opioid claims. Since our sample focuses on the Medicare Part D prescription, the average beneficiary in our sample is above 70 years old and riskier (140%) than the national average (normalized as 1). Unconditionally, each

[7] These factors include the beneficiary's age, sex, eligibility for Medicaid, initial reason for Medicare qualification, residence in an institution such as a long-term care facility, and the diagnoses assigned to the beneficiary in inpatient, outpatient and office-based settings during a base year.

Table 1: Summary Statistics

This table reports summary statistics for the main variables in the paper. Panel A summarizes annual observations of physicians from 2013 to 2021. PF_i equals one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ equals one if year t is greater than or equal to 2016, and zero otherwise. $LAClaims_{i,t}$ is the number of long-acting opioid claims by physician i in year t. $LACost_{i,t}$ is the dollar amount of long-acting opioid prescription costs. $LASupply_{i,t}$ is the days of supply of long-acting opioids. $TotClaims_{i,t}$ is the total number of claims by physician i in year t. $LAProb_{i,t}$ is the probability of prescribing long-acting opioids (in percentage points). $LARate_{i,t}$ is the percentage of long-acting opioid claims among all opioid claims. $AvgAge_{i,t}$ is the average age of beneficiaries visiting physician i in year t. $MalePct_{i,t}$, $BlackPct_{i,t}$, and $HispanicPct_{i,t}$ denote, respectively, the fraction of male, African American, and Hispanic beneficiaries visiting physician i. $AvgRisk_{i,t}$ is the average risk score of beneficiaries visiting physician i. $DualPct_{i,t}$ is the fraction of beneficiaries eligible for both Medicare and Medicaid. $Purdue_{i,t}$ equals one if physician i receives in-kind payments from Purdue Pharma in year t, and zero otherwise. Panel B compares the mean values of opioid prescription variables for treatment and control groups across three periods: 2013–2015, 2016–2018, and 2019–2021.

Panel A: Full Sample

Variable	N	Mean	Std	p25	Median	p75
$PF \times Post$	443,286	0.060	0.237	0.000	0.000	0.000
$LAClaims$	443,286	13.546	33.311	0.000	0.000	12.000
$LACost$	443,286	2623.096	7691.018	0.000	0.000	788.490
$LASupply$	443,286	382.032	928.782	0.000	0.000	360.000
$TotClaims$	443,286	4144.498	4876.197	754.000	2616.000	5884.000
$LAProb$	443,286	0.225	0.510	0.000	0.000	0.171
$LARate$	443,286	4.255	8.898	0.000	0.000	4.628
$AvgAge$	443,286	71.925	3.930	70.090	72.415	74.381
$MalePct$	443,286	0.398	0.125	0.321	0.409	0.476
$BlackPct$	443,286	0.107	0.178	0.000	0.029	0.140
$HispanicPct$	443,286	0.072	0.154	0.000	0.000	0.077
$AvgRisk$	443,286	1.399	0.527	1.076	1.261	1.553
$DualPct$	443,286	0.223	0.184	0.085	0.181	0.321
$Purdue$	443,286	0.039	0.194	0.000	0.000	0.000

Panel B: Mean Values of Opioid Prescription in Subsample

	Pre-Treatment		**Treatment**		**Post-Treatment**	
	Control	*Treated*	*Control*	*Treated*	*Control*	*Treated*
$LAClaims$	17.490	16.204	13.859	13.717	8.733	9.505
$LACost$	3321.442	3067.516	2660.552	2689.185	1781.764	1987.570
$LASupply$	495.984	437.808	393.150	369.704	246.039	256.725
$LAProb$	0.295	0.253	0.225	0.213	0.148	0.159
$LARate$	5.315	4.440	4.367	3.823	3.089	2.751
$Purdue$	0.067	0.059	0.042	0.047	0.004	0.005

physician has a 4% chance of receiving promotional payments from Purdue Pharma.

3.4 Empirical Strategy

Our main sample includes annual observations of physician EHR usage and opioid prescriptions from 2013 to 2021. This sample spans three phases: pre-treatment (2013 – 2015), treatment (2016 – 2018), and post-treatment (2019 – 2021). Our primary specification is essentially a simple DID design:

$$Y_{i,k,t} = \alpha + \beta PF_i \times Post_t + \gamma' X_{i,t} + \delta_i + \mu_{k,t} + \varepsilon_{i,k,t}. \tag{1}$$

In the above equation, $Y_{i,k,t}$ denotes the opioid prescription made by physician i in area k during year t. PF_i indicates whether physician i adopted Practice Fusion products in the Medicare EHR Incentive Program. $Post_t$ is one if year t is greater than or equal to 2016, and zero otherwise. $X_{i,t}$ indicates the set of control variables capturing patient characteristics, detailed in Section 3.3. We include two-way fixed effects (FEs), the physician FE, δ_i, and the area-year FE, $\mu_{k,t}$, which will absorb the standalone effects of PF_i and $Post_t$. This specification effectively compares the opioid prescription of two otherwise similar physicians in the same area and year, where one is exposed to the Practice Fusion manipulation, and the other is not.

To estimate both the short-term and long-term effects, we estimate Equation (1) separately in two samples. The first sample is truncated at 2018, including only the pre-treatment and treatment phases, and captures the direct effects of CDS manipulation. The second sample excludes observations from 2016 to 2018, retaining the pre-treatment and post-treatment phases. Because the biased recommendation was removed in the post-treatment period, the estimated coefficient in this specification reflects the long-term persistence in physician prescribing habits.

Our identification strategy relies on the parallel trends assumption, which requires that treatment and control groups would have exhibited similar trends in LA opioid prescriptions absent Practice Fusion's intervention. This assumption is plausible for several reasons. First, concerns about self-selection are mitigated by the timing of events. We defined treated physicians as those having adopted Practice Fusion's EHR platform between 2011 and 2015 as part of the federal incentive program, whereas the manipulation secretly occurred only in 2016. It is highly unlikely that physicians could have anticipated the manipulation or intentionally adopted a product that would later generate biased recommendations. Second, one might alternatively argue that treated physicians had a stronger intrinsic preference for LA opioids. For instance, they may continue to prescribe OxyContin even after Purdue Pharma faced litigation and reputational damage, while control physicians were more responsive to negative publicity. Contrary to this hypothesis, Panel B of Table 1 shows that, if anything, treated physicians were more conservative in prescribing LA opioids prior to the manipulation: they prescribed fewer total days of LA opioids and were less likely to receive financial payments from Purdue Pharma. Consistently, columns (1) and (2) of Appendix Table B.2 show that the pre-treatment number of LA claims are negatively correlated with Practice Fusion adoption. Third, changing EHR vendors under the incentive program is extremely rare: in a given year, there is only a 0.08% chance that a physician switches to Practice Fusion and a 0.3% chance that they switch to any other EHR system. Thus, it is unlikely that affected physicians detected the manipulation or that physicians with a preference for opioids intentionally switched to the manipulated software. In fact, as shown in Appendix Table B.2, pre-treatment opioid claims have no significant effect on physicians' decisions to change EHR vendors. Lastly, we complement these arguments with visual evidence showing no differential pre-trends before the shock. We also perform placebo tests demonstrating null effects on SA opioids and other regulated drugs not featured in Practice Fusion's CDS recommendations.

3.5 Regulatory Background

Panel B of Table 1 also foreshadows that the focal coefficient in Equation (1) is likely to be positive. During the treatment period, the gap in LA opioid prescribing tendencies between the two groups narrowed substantially. In the post-treatment period, the pre-treatment pattern reversed, with the treatment group exhibiting a stronger tendency to prescribe LA opioids across nearly all measures.

Equation (1) estimates whether the treatment group prescribed more long-acting (LA) opioids after the shock relative to the control group. Panel B of Table 1 shows that, in absolute terms, both groups exhibited declining trends, although the treatment group maintained higher persistence following the shock. This overall decline reflects a broader tightening of opioid regulations during the sample period. In 2014, the U.S. Food and Drug Administration (FDA) required new safety labeling for extended-release and long-acting opioids, including a boxed warning on the "risks of addiction, abuse, and misuse, which can lead to overdose and death." In 2016, the CDC issued "Guideline for Prescribing Opioids for Chronic Pain", which outlined 12 evidence-based recommendations for primary care clinicians in outpatient settings (excluding cancer, palliative, and end-of-life care) to promote safer and more effective pain management, improve patient outcomes, and reduce opioid misuse and overdose.

However, these federal-level policies, particularly the CDC guidelines, were voluntary and intended to be flexible, designed to support rather than replace individualized, patient-centered care. Consequently, they did not fully eliminate opioid misuse, as prescribing ultimately depends on physicians' clinical judgment and risk assessment. Two pieces of evidence support this point.

Table 2: Impacts of Practice Fusion's CDS on Opioid Prescription

This table reports the effects of Practice Fusion's CDS on opioid prescribing. Columns (1) to (4) show the short-term impacts using annual physician-level observations from 2013 to 2018 (pre-treatment and treatment phases). Columns (5) to (7) show the long-term impacts using physician-level observations from 2013–2015 and 2019–2021 (pre-treatment and post-treatment phases). PF_i equals one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ equals one if year $t \geq 2016$, and zero otherwise. $Log(LAClaims)_{i,t}$, $Log(LACost)_{i,t}$, and $Log(LASupply)_{i,t}$ are the logarithms of one plus $LAClaims_{i,t}$, $LACost_{i,t}$, and $LASupply_{i,t}$, respectively. $Log(SAClaims)_{i,t}$ is the logarithm of one plus the number of short-acting opioid claims. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. All regressions include physician fixed effects and area-year fixed effects. Standard errors are clustered at the three-digit zip code level, and t-statistics are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	Short-term Impacts				Long-term Impacts		
	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*	(4) *Log(SAClaims)*	(5) *Log(LAClaims)*	(6) *Log(LACost)*	(7) *Log(LASupply)*
$PF \times Post$	0.059***	0.154***	0.114***	-0.025	0.116***	0.308***	0.231***
	(3.943)	(4.474)	(4.117)	(-1.447)	(4.826)	(5.694)	(5.199)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	305,082	305,082	305,082	305,082	292,094	292,094	292,094
Adj. R^2	0.79	0.76	0.76	0.87	0.72	0.69	0.69

First, despite tighter federal guidance, prescription opioids remained the most commonly misused prescription drug in the United States in 2020, and 64.6% of individuals reporting misuse cited pain relief as their primary motive.[8] Second, although the 2016 CDC Guideline did not intend to impose legal limits, many states subsequently enacted laws restricting initial opioid prescriptions for acute pain to a seven-day supply or less. Despite these policy shifts, Appendix Figure A.1 shows that, based on raw Marketscan Outpatient Drug Claim data, most OxyContin prescriptions continued to have supplies exceeding 14 days, with little observable decline after 2016. Together, these patterns indicate that while federal and state policies introduced formal constraints, they did not eliminate opportunities for misuse, leaving room for CDS manipulation to influence prescribing behavior.

4 Regression Results

4.1 Opioid Prescription

We start by estimating Equation (1) in the initial sample from 2013 to 2018 in Table 2 to investigate whether the treatment group significantly prescribed more LA opioids when Practice Fusion manipulated the CDS. This analysis provides a first-stage result to confirm the existence of direct impacts in the short term. One might argue that as experienced professionals, affected physicians

[8]For details, see "Key Substance Use and Mental Health Indicators in the United States: Results from the 2020 National Survey on Drug Use and Health," Substance Abuse and Mental Health Services Administration.

will use their discretion to reject the manipulated recommendation and avoid unnecessary opioid prescriptions. Instead, the first three columns of Table 2 indicate that physicians utilizing Practice Fusion EHRs experienced a significant 5.9% increase in annual LA claims compared to the control group, equivalent to approximately 0.8 additional claims each year, from 2016 to 2018. Meanwhile, their total LA opioid costs surged by 15.4%, accompanied by a 11.4% increase in days of supply. A ballpark estimate of the total additional costs for Medicare Part D in our sample is approximately \$2 million annually.[9] This magnitude closely aligns with Practice Fusion's own prediction, ranging between \$8,458,232 and \$11,277,643, in opioid revenue driven by the CDS.

Column (4) in Table 2 also includes a placebo test to validate our identification assumption. As we explained earlier, Purdue Pharma's leading products are extended-release opioids used for chronic pain management, suggesting the manipulation should have minimal effects on SA opioids. Column (4) shows that affected physicians do not exhibit any significant differences in SA claims after the shock, confirming that the changes were concentrated in LA opioids. The negative coefficient, though insignificant, is in line with Practice Fusion's claim that the CDS alert generated a shift from SA opioids to extended-release products.

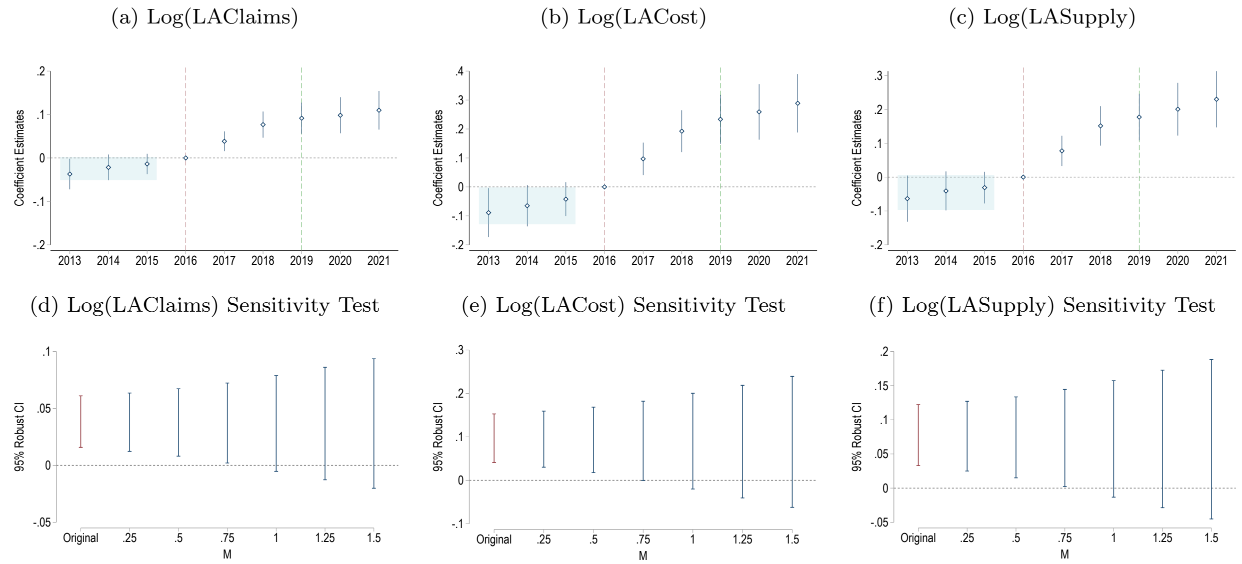
Next, we estimate Equation (1) in the sample consisting of both pre-treatment and post-treatment phases. If the distortion in opioid prescription was only transient, then affected physicians would not exhibit any significant changes in LA claims from 2019 to 2021. However, the last three columns of Table 2 demonstrate that affected physicians continue to prescribe more LA opioid claims even after the CDS stopped making recommendations after the Spring of 2019. The economic magnitudes are even larger. For example, column (1) implies a 11.6% increase in LA claims, nearly double the corresponding coefficient in the short term. Rises in LA opioid costs for affected physicians become 30.8%, implying \$4.0 million aggregated additional payments by Medicare Part D every year in our sample. The larger magnitude is consistent with our hypothesis, as long-term bias repeatedly accumulated during the treatment window. This result implies that, through learning with CDS, the short-term manipulation will generate persistent long-term impacts on professional decision making. Affected physicians develop a habit of opioid usage and ultimately maintain the tendency to prescribe opioids, regardless of the CDS recommendation.

Table 2 estimates the average treatment effects in different phases. To visualize the coefficient dynamics in the whole sample, we estimate and plot the coefficients β_c from the following equation

[9] The calculation is based on the sample average annual cost \$2,623, multiplied by the coefficient and the number of physicians in the treatment group (5,401).

Figure 1: Treatment Dynamics of Practice Fusion's Manipulation

This figure plots the treatment dynamics of LA opioid prescribing associated with Practice Fusion's manipulation. Panels (a)–(c) display the β_c coefficients from Equation (2) with 95% confidence intervals (solid lines). The x-axis indicates calendar years. The base year is 2016 (red dashed line), the first year of the manipulation. The green dashed line marks the suspension of the manipulation in 2019. Shaded regions indicate the 95% confidence interval of the average pre-treatment coefficients from 2013 to 2015. Panels (d)–(f) present results from the HonestDiD procedure of Rambachan and Roth (2023), plotting 95% confidence intervals of the average treatment effects from 2017 to 2021 under alternative bounds on relative magnitudes M. Each M imposes that the post-treatment violation of parallel trends is no more than M times the maximum violation of parallel trends in the pre-treatment period (between consecutive periods).



in Figure 1:

$$Y_{i,k,t} = \alpha + \sum_{\substack{c\in[2013,2021] \\ c\neq 2016}} \beta_c PF_i \times \mathbf{I}(t=c) + \gamma' X_{i,t} + \delta_i + \mu_{k,t} + \varepsilon_{i,k,t}. \tag{2}$$

$\mathbf{I}(t = c)$ indicates whether year t corresponds to calendar year c, and β_c measures the difference between the treatment and control groups in year c. Note that the CMS Medicare Part D provider-level prescription data are only available starting in 2013. As a result, we have only three pre-treatment years. To better visualize the pre-treatment trend, we deviate from the standard convention by using 2016, the first treatment year, as the reference year in Equation (2). Because potential treatment effects may already arise in 2016, the estimated pre-shock coefficients are not necessarily expected to be insignificantly different from zero. However, panels (a)–(c) of Figure 1 show that, in general, these coefficients are insignificantly different from zero except for some in 2013. The parallel trends assumption essentially requires that these pre-shock coefficients exhibit no systematic trend before the shock, rather than all being individually insignificant. We validate this assumption in different ways. First, we follow the standard convention by using 2015 as the reference year and demonstrate that the 2013 and 2014 coefficients are not statistically different from zero in Appendix Figure A.2. Second, we mark the 95% confidence interval of the average pre-treatment coefficients from 2013 to 2015 in shaded regions of Figure 1, suggesting that none of the coefficients are significantly different from their average values. Lastly, we apply the Honest-DiD procedure of Rambachan and Roth (2023) to assess the sensitivity of our estimated treatment effects to potential violations of the parallel trends assumption. Specifically, this procedure first computes the maximum violation of parallel trends in the pre-treatment period (i.e., the largest difference in coefficients between consecutive pre-treatment years). It then produces the 95% confidence intervals for the average treatment effects from 2017 to 2021, under the assumption that the post-treatment violation between two years is no larger than M times this maximum pre-treatment violation. Figure 1 panels (d)–(f) show that the estimated treatment effects remain statistically significant when $M \leq 0.75$. In Appendix Figure A.3, we further show that the treatment effects remain significant at the 10% level even when $M = 1$, corresponding to a post-treatment violation equal in magnitude to the maximum pre-treatment deviation.

Figure 1 panels (a)–(c) also show that while Practice Fusion manipulated the CDS (between the two vertical lines), affected physicians gradually increased LA opioid usage over the years due to increased belief distortions. While this momentum stopped after the suspension of biased recommendation in 2019, the treatment group maintained a constantly high frequency of prescription. These dynamics explain the magnitude differences in short-term and long-term effects. This pattern also supports the hypothesis: each periodic interaction with the CDS will lead to more long-term biases.

4.2 Mechanism and Robustness

Log-Transformation Chen and Roth (2024) recently show that the average treatment effects of log-transformed variables may not accurately approximate percentage changes when the variables contain zeros. In Table 3, we conduct several tests to address this concern. First, in Table 2, both *LAClaims* and *LASupply* are count variables, making them suitable for Poisson regressions. Columns (1) and (2) of Table 3 confirm that the estimated coefficients remain statistically significant for both outcome variables. For *LAClaims*, the coefficients are consistent with the percentage changes implied by the log-transformed results in Table 2. Moreover, when we winsorize the raw *LAClaims* values at the 5% level to mitigate the influence of outliers, Table 3 column (3) shows that the treatment effects estimated from the non-transformed counts are very close to the estimates

Table 3: Robustness: Alternative Specifications for Count Variables

This table reports robustness results addressing concerns about the log-transformations for count outcomes. Panel A presents short-term impacts using physician-level observations from 2013 to 2018 (pre-treatment and treatment phases). Panel B presents long-term impacts using physician-level observations from 2013–2015 and 2019–2021 (pre-treatment and post-treatment phases). PF_i equals one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ equals one if year $t \geq 2016$, and zero otherwise. Columns (1) and (2) report Poisson estimates, while Columns (3) and (4) report OLS estimates with dependent variables winsorized at the 5% level. Columns (5) and (6) use $LAProb_{i,t}$, the probability of prescribing a long-acting opioid, and $LARate_{i,t}$, the share of long-acting opioid claims among all opioid claims, as alternative outcome variables. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. All regressions include physician fixed effects and area-year fixed effects. Standard errors are clustered at the three-digit zip code level, and t-statistics are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Panel A: Short-term Impacts

	Poisson		**OLS**			
	(1)	(2)	(3)	(4)	(5)	(6)
	LAClaims	*LASupply*	*LAClaims*	*LASupply*	*LAProb*	*LARate*
$PF \times Post$	0.051**	0.052**	0.738***	20.989***	0.020***	0.236**
	(2.299)	(2.405)	(3.871)	(4.049)	(4.157)	(2.568)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	133,239	133,239	305,082	305,082	305,082	305,082
Adj. R^2			0.83	0.83	0.77	0.70

Panel B: Long-term Impacts

	Poisson		**OLS**			
	(1)	(2)	(3)	(4)	(5)	(6)
	LAClaims	*LASupply*	*LAClaims*	*LASupply*	*LAProb*	*LARate*
$PF \times Post$	0.110***	0.120***	1.429***	43.230***	0.038***	0.343**
	(2.710)	(3.012)	(4.685)	(5.093)	(5.693)	(2.476)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	125,794	125,794	292,094	292,094	292,094	292,094
Adj. R^2			0.75	0.75	0.70	0.62

implied by Table 2. In contrast, for *LASupply*, the percentage magnitudes appear inflated by the log-transformation, suggesting that the estimates in Table 2 should not be interpreted directly in percentage terms. Second, Table 3 columns (5) and (6) estimate the treatment effects on ratio outcome variables, which capture the extensive margin of opioid prescribing. *LAProb* measures the probability of prescribing a long-acting opioid (in percentage points), calculated as the number of LA claims divided by all claims in a given year for each physician. *LARate* represents the share of long-acting opioid claims among all opioid claims, a measure reported in the CMS Medicare Part D data that reflects the physician's relative preference for long-acting versus short-acting opioids in pain management. Both variables are positive and statistically significant, suggesting that physicians in the treatment group are more likely to prescribe long-acting opioids relative to the control group.

Patient Demand Given that LA opioids are highly addictive medications, it is plausible that patients with substance use disorders may repeatedly visit affected physicians in the post-treatment period, leading these physicians to maintain higher prescription frequencies. This concern is closely related to the mechanism underlying the results in Table 2: affected physicians may prescribe larger quantities of opioids to existing patients with addiction risks, or extend opioid supplies to new patients with more lenient pain conditions, or both. Unfortunately, the CMS Medicare Part D data are aggregated at the provider level due to privacy concerns, preventing us from retrieving individual patient characteristics for each claim. Nevertheless, we address this concern in two ways and show that the observed effects are more likely driven by extended coverage of opioid prescriptions rather than repeated demand from addicted patients. First, Appendix Table B.3 uses the unique number of Medicare beneficiaries who receive LA opioids as the outcome variable. In the treatment window, affected physicians prescribed LA opioids to 2.8% more patients relative to the control group. This coefficient magnitude increases to 4.3% in the post-treatment window, suggesting a larger opioid recipient group on the margin. To construct a proxy for prescribing the drug to new patients, we create indicators of whether the number of beneficiaries receiving LA opioids strictly increases from the previous year of the same physician. We find that affected physicians are 0.5% to 0.6% more likely to increase the number of patients receiving LA opioids, equivalent to roughly 20% of the unconditional average chance. We then estimate the average LA supply per beneficiary by dividing $LASupply_{i,t}$ by $LABene_{i,t}$. The last two columns of Appendix Table B.3 show that these effects are statistically insignificant, suggesting that the observed increase in total LA supply is primarily driven by a larger patient base rather than higher per-patient dosage.

An alternative specification to address this concern is through the mover approach, commonly used to separate demand and supply factors (e.g. Finkelstein et al., 2016; Keys et al., 2023). Consider the group of physicians who move to new locations in the post-treatment window. They are likely to face new patients, and their prescription behavior therefore reveals their long-term preference. Table 4 replicates the long-term impact analysis, except that we restrict the sample to physicians who move to a different city after 2019. Although the sample size reduces by more than 90% due to the large group of non-movers, the coefficients remain significant for all outcome variables with similar economic magnitudes. In the Appendix, we consider alternative ways of defining movers, requiring them to migrate to new 5-digit zip codes or states (Appendix Tables

Table 4: The Long-term Impacts in the Mover Sample

This table reports the long-term impacts of Practice Fusion's CDS on opioid prescribing in the mover sample. Columns (1) to (4) present estimates for physicians who relocated to a new city after 2019, and Columns (5) to (8) present estimates for physicians who changed primary affiliations after 2019. The analysis covers physician-level observations from 2013–2015 and 2019–2021, corresponding to the pre-treatment and post-treatment phases. PF_i equals one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ equals one if year $t \geq 2016$, and zero otherwise. $Log(LAClaims)_{i,t}$, $Log(LACost)_{i,t}$, and $Log(LASupply)_{i,t}$ are the logarithms of one plus $LAClaims_{i,t}$, $LACost_{i,t}$, and $LASupply_{i,t}$, respectively. $LARate_{i,t}$ is the share of long-acting opioid claims among all opioid claims by physician i in year t. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. All regressions include physician fixed effects and area-year fixed effects. Standard errors are clustered at the three-digit zip code level, and t-statistics are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	New City				**New Affiliation**			
	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*	(4) *LARate*	(5) *Log(LAClaims)*	(6) *Log(LACost)*	(7) *Log(LASupply)*	(8) *LARate*
$PF \times Post$	0.139*	0.407**	0.339**	0.793*	0.153***	0.407***	0.299***	0.538**
	(1.696)	(2.268)	(2.286)	(1.666)	(3.974)	(4.714)	(4.158)	(2.359)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	20,049	20,049	20,049	20,049	75,407	75,407	75,407	75,407
Adj. R^2	0.67	0.64	0.64	0.56	0.71	0.67	0.68	0.61

B.4 and B.5). The sample size will adjust correspondingly, but the results remain robust. The analysis of cross-state movers is particularly important despite a relatively small sample (above 5,000 observations). This is because physicians must hold an active license to practice in a particular state and typically hold the license only in their home state. As a result, they cannot legally prescribe for patients in the previous location once the old license expires. A related concern is that EHR adoption may occur at the facility level, and facilities can have heterogeneous policies governing opioid prescribing. To address this, we extract physicians' affiliation information from the CMS Doctors and Clinicians National Downloadable File, which lists up to five organizational affiliations per physician, and define "moving" as a change in the primary (first) affiliated facility. In the Appendix Table B.6, we expand the definition to changes in any affiliated facilities. The results remain robust under this alternative definition.

Privately Insured Patients In this paper, we focus on Medicare prescriptions because we can link physician-level identifiers to the EHR adoption in the HITECH Act. However, we acknowledge the concern that almost all Medicare beneficiaries are the elder population, and our results may not be generalized due to heterogeneous patient demographics in other insured populations. To alleviate this concern, we provide consistent evidence using the Marketscan database. The data provider sources commercial insurance claim information for employees, retirees, and dependents from over 260 medium and large employers and 40 health plans. The initial database covers over 43 million privately-insured individuals with employment-based health plans, representing roughly

14% of the whole insured population. Due to privacy concerns, the data provider de-identifies the physician information so we cannot link Practice Fusion adoption to individual providers in this database. Moreover, granular patient geographic identifiers, such as the zip code or the FIPS code, are also redacted and we can only rely on the Core-Based Statistical Area (CBSA) information. As a result, our estimation effectively concentrates on the metropolitan and micropolitan populations. For each CBSA, we define the treatment intensity as the logarithm of (one plus) the unique number of physicians who adopted Practice Fusion in the Medicare Incentive Program in that area. Note that this is still a valid measure for treatment intensity because physicians attest to meaningful use of the EHR with patients insured by all health plans, not just by Medicare. The term "Medicare" reflects federal sponsorship by CMS rather than state medical boards.

Despite its limitations, the Marketscan database complements our Medicare analysis by providing prescription-level information. This granularity allows us to directly identify prescriptions of Purdue Pharma's products based on their National Drug Codes (NDCs). We aggregate the total payments across all prescriptions related to Purdue products within each CBSA and year. We focus on payments rather than raw prescription counts, since the latter do not capture variation in quantities per order. On average, enrollees in a typical CBSA spent over $215,000 on Purdue products in the short-term sample and over $174,000 in the long-term sample. In Table 5, we show that a 1% increase in the number of Practice Fusion adopters is associated with a 0.15% increase in payments for Purdue Pharma's products in the short term. This effect persists beyond 2019, with an economic magnitude of approximately 0.53%, suggesting sustained influence on opioid purchasing behavior. In terms of the magnitude, the average number of physicians who adopted Practice Fusion in an area is 8.65, with a standard deviation of 13.88. Thus, a one-standard deviation increase from the mean is associated with a 5.8% increase in payments for Purdue Pharma's products in the short term and a 20.4% increase in the long term. Moreover, for each enrollee, we can keep track of the whole claim history in both inpatient and outpatient settings. As a result, we are able to distinguish whether or not the patient has ever used Purdue Pharma's products before. Appendix Figure A.4 shows that new users initially accounted for over 30% of total expenditures, but this share has recently declined to below 20%. In column (2), we restrict the analysis to payments associated with first-time opioid users and find similar results, suggesting that affected physicians were extending opioid prescriptions to a broader set of new patients.

We next examine the clinical conditions of patients receiving opioid prescriptions. Formally, products such as OxyContin are approved only for severe and persistent pain requiring extended treatment, where alternative therapies are inadequate. As described in Section 3, Practice Fusion's

Table 5: Robustness of the Main Results using Commercial Insurance Claims

This table reports the robustness of the main results using commercial insurance claims. Panel A presents estimates for the short-term sample (2013–2018), and Panel B presents estimates for the long-term sample (2013–2015 and 2019–2021). $Log(NumPF)_k$ is number of physicians having adopted Practice Fusion in area k. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. Columns (1)–(4) report total payments for Purdue Pharma's products across different categories in area k and year t: all claims (*All Claims*), new patients (*New Patient*), patients with a chronic pain diagnosis in the past 180 days (*Chronic*), and patients without a chronic pain diagnosis in the past 180 days (*Non-Chronic*). Column (5) reports the total number of chronic pain claims (*Chronic Claims*). In all regressions, the outcome variables are log-transformed, but the unlogged sample means of the outcome variables are reported in the "Mean" row. Control variables include local demographic characteristics: logged population, logged per-capita income, male percentage, African American percentage, Hispanic percentage, average age, unemployment rate, and health insurance coverage percentage. All regressions include CBSA and year fixed effects. Standard errors are clustered at the area level, and t-statistics are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Panel A: Short-term Impacts

	(1)	(2)	(3)	(4)	(5)
	Log(PurduePay)				Log(Chronic)
	All Claims	*New Patient*	*Chronic*	*Non-Chronic*	
$Log(NumPF)$ $\times Post$	0.151^{***} (3.088)	0.211^{***} (2.959)	0.195^{**} (2.504)	0.256^{***} (4.613)	-0.005 (-0.167)
Controls	Yes	Yes	Yes	Yes	Yes
CBSA FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Mean	215,506.12	84,449.64	78,277.13	137,228.99	2,470.19
N	1,985	1,985	1,985	1,985	1,985
Adj. R^2	0.85	0.70	0.70	0.81	0.93

Panel B: Long-term Impacts

	(1)	(2)	(3)	(4)	(5)
	Log(PurduePay)				Log(Chronic)
	All Claims	*New Patient*	*Chronic*	*Non-Chronic*	
$Log(NumPF)$ $\times Post$	0.528^{***} (3.660)	0.928^{***} (6.770)	0.537^{***} (3.261)	0.807^{***} (5.546)	0.041 (1.093)
Controls	Yes	Yes	Yes	Yes	Yes
CBSA FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Mean	174,287.34	65,768.37	59,692.91	114,594.43	1,966.91
N	1,915	1,915	1,915	1,915	1,915
Adj. R^2	0.70	0.67	0.63	0.72	0.92

biased recommendations could trigger extended-release opioid prescriptions either when patients are diagnosed with chronic pain, or when their reported pain score is sufficiently high. In Table 5 columns (3) and (4), we split the total payments for Purdue Pharma's products by whether the patient had a chronic pain diagnosis in the previous 180 days. Appendix Figure A.4 shows that in earlier years, and perhaps contrary to common expectations, most enrollees receiving Purdue products had not been recently diagnosed with chronic pain. In recent years, expenditures from chronic and non-chronic pain patients have become roughly equal. In Table 5, the coefficients for both outcomes are statistically significant, with larger magnitudes for the non-chronic pain group.

These findings suggest that affected physicians became more likely to prescribe opioids not only to patients with prolonged pain conditions but also to those without, pointing to potential overtreatment or inappropriate prescribing. Consistently, in Table 5 column (5) we do not find any evidence that the number of claims associated with chronic pain conditions (regardless of opioid prescriptions) changes following the treatment, suggesting the increased prescriptions are likely driven by changes in the prescribing preferences rather than patient conditions.

State Regulations During our sample period, several federal and state policies were implemented to mitigate the risk of opioid addiction. In our specification, with granular Area×Year fixed effects, we are effectively comparing physicians within the same geographic area. As a result, our estimates cannot be driven by cross-regional policy differences. Nonetheless, it is informative to examine whether these policies moderated the impact of the CDS shock.

We focus on two sets of regulations. First, following the 2016 CDC Guideline, several states imposed restrictions on the maximum days of supply per prescription. Second, most states adopted or strengthened their Prescription Drug Monitoring Programs (PDMPs), which are statewide databases that track controlled-substance prescriptions, including opioids. Under these programs, physicians are required to check the PDMP before issuing a prescription, and pharmacists must report dispensing data promptly. Appendix Tables B.7 to B.9 confirm that the main results remain statistically significant among the subset of states with these regulations. Moreover, the estimated coefficients are very similar to, or even slightly larger than, the baseline estimates in Table 2. There are several reasons why these policies did not meaningfully attenuate the CDS effect. First, most state-level supply limits impose hard caps only on acute pain prescriptions, while allowing exceptions for chronic pain conditions. Second, the PDMP primarily serves as an informational tool that alerts physicians to a patient's prescription history. It does not restrict prescribing authority if the physician deems the prescription medically necessary. In contrast, the CDS directly shapes physicians' beliefs, potentially leading them to perceive such prescriptions as necessary even when caution is warranted. Put differently, CDS biases operate at the point of care in a way that can effectively overrule the PDMP. Third, many PDMP systems are not integrated into EHRs, making it costly for physicians to check them. In contrast, the CDS is embedded directly in the EHR and cannot be bypassed during routine care.

Awareness Although affected physicians were not aware of the manipulation during the treatment window, in January 2020 Practice Fusion announced a $145 million settlement with the U.S. Department of Justice, publicly revealing the details of the manipulation for the first time. This raises an important question: why did physicians not recognize their altered prescribing behavior and correct their habits afterward? One possible explanation is that they exhibit behavioral bias, where physicians remain unaware of the gradual changes in their own prescribing patterns, which we cannot directly test. An alternative explanation is that many physicians were simply unaware of the settlement itself. We examine this possibility using Google Trends data, which capture public search interest over time. First, at the aggregate level, search interest in "Practice Fusion" steadily declined and showed no noticeable spike around the 2020 settlement (Appendix Figure A.5). Second, we exploit state-level variation in search interest during 2020–2021 and classify physicians into two groups based on whether their state's search volume is above or below the median. Appendix Table B.10 shows that affected physicians in high-search states, who were arguably more likely to be aware of the manipulation, exhibit significantly lower opioid prescriptions in the long term, suggesting that greater awareness can partially mitigate the persistence of habit-driven prescribing behavior.

Detailing It is plausible that the treatment group had alternative incentives to continue prescribing LA opioids, such as receiving in-kind payments from Purdue Pharma. However, several facts suggest this is unlikely to explain our results. First, Appendix Figure A.6 shows that the geographic distribution of Practice Fusion users and payment recipients is largely uncorrelated ($\rho = -0.19$), indicating that adopters did not have other financial motives to prescribe LA opioids. Second, as shown in Appendix Figure A.7, Purdue Pharma's detailing activities peaked in 2014, when nearly 4,000 physicians in our sample received promotional payments, but steadily declined thereafter, falling to only one recipient by 2019. This decline coincides with increased transparency following the 2014 CMS Open Payments disclosure and heightened litigation risks after Purdue Pharma's legal settlements with multiple states, both of which likely curtailed promotional activities.

We formally address this concern using two regression-based approaches. First, because some physicians still received in-kind payments during the short-term sample period, we augment Equation (1) with two variables: $Purdue_{i,t-1}$ and $Purdue_{i,t-1} \times Post_t$. The first variable indicates whether a physician received in-kind payments from Purdue Pharma in the previous year, while the interaction term captures the additional effect of such payments during the treatment window. If our main results were primarily driven by these payments, this interaction term would absorb the variation in prescribing behavior, making the focal treatment coefficient insignificant. However, Panel A of Appendix Table B.11 shows the opposite: the estimated effects of Purdue Pharma remain positive and statistically significant, and the influence of in-kind payments appears to have weakened after 2016 in persuading physicians to prescribe opioids.

Second, Appendix Table B.11 Panel B replicates the analysis in Table 2 by comparing pre-

Table 6: Long-term Opioid Therapy Utilization

This table reports cross-sectional evidence on the long-term costs for Medicare beneficiaries associated with Practice Fusion adoption. The sample consists of annual observations at the three-digit ZIP code level in 2020 and 2021. $Log(NumPF)_k$ is the logarithm of the number of physicians who adopted Practice Fusion in area k. $OTPBenes_{k,t}$, $OTPSvcs_{k,t}$, $OTPCharge_{k,t}$, and $OTPPay_{k,t}$ denote, respectively, the number of beneficiaries, number of services, billed charges, and actual Medicare payments for Opioid Treatment Programs in area k and year t. In all regressions, the outcome variables are log-transformed, but the unlogged sample means of the outcome variables are reported in the "Mean" row. Control variables include local demographic characteristics: logged population, logged per-capita income, male percentage, African American percentage, Hispanic percentage, average age, unemployment rate, and health insurance coverage percentage. Year fixed effects are included. Standard errors are clustered at the area level, and t-statistics are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	(1) *Log(OTPBenes)*	(2) *Log(OTPSvcs)*	(3) *Log(OTPCharge)*	(4) *Log(OTPPay)*
$Log(NumPF)$	0.475^{***}	0.763^{***}	1.269^{***}	1.245^{***}
	(4.966)	(4.984)	(4.922)	(4.895)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Mean	74.97	1,478.33	274,451.28	228,875.90
N	1,777	1,777	1,777	1,777
Adj. R^2	0.28	0.28	0.27	0.27

scribing behavior before (2013–2015) and after (2019–2021) the treatment period. The variable $Detail_i$ indicates whether physician i received in-kind payments from Purdue Pharma before 2019. Note that Purdue Pharma effectively suspended its detailing activities after filing for bankruptcy in 2019. This setup allows us to test whether prior payment recipients continued to prescribe opioids at elevated rates even without ongoing financial incentives. The results show that former payment recipients substantially reduced their long-acting opioid prescriptions after 2019, with the average number of LA claims declining by approximately 43% relative to the pre-treatment period (2013–2015).

Real Costs We then supplement additional evidence on the long-term impacts from the perspective of real costs on beneficiaries. Starting from 2020, Medicare covers Opioid Treatment Programs (OTPs) that provide medication-assisted treatment for people diagnosed with an opioid use disorder (OUD) in Medicare Part B. Using the "Medicare Physician & Other Practitioners" file, we aggregate area-level OTP utilization based on HCPCS codes and use this measure as a proxy for local opioid addiction prevalence among older adults. We prefer this measure to alternative public data sources, such as drug overdose mortality, for two reasons. First, although we cannot directly observe the OUD status of beneficiaries who received long-acting opioids from treated physicians, aggregated OTP utilization provides an imperfect but consistent indicator within the same Medicare population. Second, overdose mortality is a relatively rare event with limited variation, which reduces statistical power relative to utilization-based measures.

Because the coverage began in recent years and did not span the pre-treatment phase, we acknowledge the limitation of this analysis as correlational evidence and interpret the results with caution. In Table 6, our focal regressor is the number of unique physicians who adopted Practice Fusion in the Medicare Incentive Program in each 3-digit zip code area. The outcome variables include the number of beneficiaries, number of services, bill charges, and actual Medicare payments for OTPs. Control variables include local demographic characteristics such as logged population, logged income per capital, male percentage, African American percentage, Hispanic percentage, average age, unemployment rate, and health insurance coverage percentage. We find that Practice Fusion usage among providers indeed correlates with long-term OUDs among the Medicare beneficiaries. For example, a 1% increase in the number of Practice Fusion adopters is associated with 0.48% more Medicare beneficiaries and 1.245% higher Medicare payments for OTPs.

Robustness Checks Lastly, we perform a few robustness checks on our empirical specification in the Appendix. First, we show that our results hold with a more restricted geographic boundary (5 digit zip code areas) for the control group in Appendix Table B.12. This restriction generates fewer observations in the sample, as expected. But the coefficient magnitudes and statistical significance become even larger. Second, we consider alternative clustering methods using vendors in Appendix Table B.13, and show that the main results remain robust. Lastly, we consider other regulated medicines in Medicare Part D, namely antibiotics (due to concerns of antimicrobial resistance) and antipsychotics (due to side effects such as obesity and metabolic disorder). We do not find that treated physicians significantly increase prescriptions of these drugs in Appendix Table B.14. Consistent with column (4) of Table 2, the estimated coefficients are all negative, suggesting a potential crowd-out effect of long-acting opioid prescriptions.

5 Magnitude of Habit Changes

5.1 Method

Table 2 indicates the presence of long-term effects on physicians' prescribing habits. We now quantify the economic magnitude of this habit change. Recall that during the treatment window (2016–2018), affected physicians experienced both direct manipulation and subsequent behavioral adjustments in their prescribing habits. The short-term estimates in Table 2 therefore capture the combined effect of these two channels. To disentangle them, we employ a machine-learning approach based on the multi-arm causal forest (Nie and Wager, 2021), which allows us to estimate the conditional average treatment effect (CATE) as a function of observed physician characteristics.

This approach enables us to compare the magnitude of prescribing habit changes to those arising from direct manipulation, and to construct counterfactual scenarios in which prescribing habits evolve differently. Formally, we will estimate the following equation using the sample from 2013 to 2018:

$$p(\bar{X}_i) = p^C(\bar{X}_i) + \mathbf{I}\{i \in \mathbf{T}, t \geq 2016\} \left(\sum_{l=0}^{t-2016} \tau_l(\bar{X}_i) \right). \tag{3}$$

$\mathbf{I}\{i \in \mathbf{T}, t \geq 2016\}$ indicates that physician i belongs to the treatment group and that year t falls within the treatment window. We also calculate the number of prior exposures to the biased CDS as $t-2016$. To illustrate, consider two otherwise similar physicians, one using Practice Fusion (treated) and one not (control), under two cases. First, in 2016 (the initial treatment period), neither physician has previously interacted with the manipulated CDS. Therefore, the last component in Equation (3) includes only one term, τ_0, which captures the immediate effect of exposure to the biased reminder. Second, in 2017 (the second treatment period), the treated physician experiences both the direct manipulation effect (τ_0) and a habit-formation effect (τ_1) arising from one year of exposure. The assumption embedded in Equation (3) is that each additional interaction with the biased reminder reinforces prescribing habits, generating cumulative effects over time. Put differently, the longer the treatment period, the stronger and more persistent the habit distortion becomes. We assume all terms in Equation (3) are characterized by physician-level variables $\bar{X}_i$, including the average patient beneficiary age, risk score, male fraction, African American fraction, Hispanic fraction, dual qualification (Medicaid and Medicare) fraction, physician seniority (years since graduation) and physician gender. This specification allows us to capture heterogeneous effects of the CDS across different physician profiles.

To estimate this relationship, for each physician i in year t, we define the prescription probability $p_{i,t}$ as the total number of long-acting (LA) opioid claims divided by the total number of all claims, multiplied by 100. For clarity, $p_{i,t}$ is expressed in percentage points, given the relatively low opioid prescribing rates in the sample. Appendix Section C presents a physician decision model that microfounds this specification by modeling the probability of prescribing an opioid as the key outcome variable. Consistent with the hypothesis that each additional interaction with the biased reminder reinforces prescribing habits, our OLS estimation of Equation (3) shows that all the estimated τ_ls are around 1.1 to 1.2 bps, being statistically significant at 1% (Appendix Table B.15).

The logic behind the multi-arm causal forest, or broadly the generalized random forest, involves constructing a collection of decision trees. Each tree is built from a random subsample of the data as the initial node, and splits it into child nodes recursively to form leaves. Each node is split based on a random subset of variables $\bar{X}_i$ using threshold strategies. This threshold strategy aims to maximize the difference in estimated treatment effects post-split. Once the trees are constructed, the final treatment effect for a given data point is estimated by comparing outcomes between treated and control units, weighted by how many times they belong to the same bottom leaf.

We estimate $\tau_l(\bar{X}_i)$ by adjusting the multi-arm causal forest by Nie and Wager (2021) as follows. In the original algorithm, each unit i belongs to one of the mutually exclusive treatment arms, or the control arm. They then estimate the CATE of each treatment arm, without any boundary conditions. In our setting, there is only one treatment arm, but treated units gradually receive heterogeneous CATE $\tau_l(\bar{X}_j)$ depending on the number of previous treatment periods. We also impose additional restrictions on regulating the values of $\hat{\tau}_l(\bar{X}_i)$ to be non-negative, guided by Appendix Table B.15. An important caveat is that $\tau_1(\cdot)$ and $\tau_2(\cdot)$ are estimated conditional on $\tau_0(\cdot)$. Therefore, we cannot extrapolate the estimated $\hat{\tau}_1(\cdot)$ and $\hat{\tau}_2(\cdot)$ to the 2019 – 2021 sample without an independence assumption.

For each observation in the data (both treated and control), this algorithm generates the predicted treatment effects $\hat{\tau}_l(\bar{X}_i)$ respectively, along with the baseline counterfactual prescription probability $\hat{p}^C(\bar{X}_i)$ in the control group. We derive the predicted probability using Equation (3)

$$\hat{p}(\bar{X}_i) = \hat{p}^C(\bar{X}_i) + \mathbf{I}\{j \in \mathbf{T}, t \geq 2016\} \left(\sum_{l=0}^{t-2016} \hat{\tau}_l(\bar{X}_i) \right). \tag{4}$$

As a benchmark, we first compare the raw prescription probability in data to the estimated $\hat{p}$. Column (1) of Table 7 implies that a typical treated physician has a higher chance of LA opioid prescription by 2.3 bps, equivalent to 9.5% of the unconditional average. To assess the fitness of our predicted probability, using $\hat{p}$ as the outcome yields a 2.4 bps coefficient, matching the ground truth closely.

We next assume no long-term distortion by setting $\hat{\tau}_1 = \hat{\tau}_2 = 0$ in Equation (4). Under this assumption, Column (3) shows a 54% reduction in the average treatment effect to 1.1 bps. Column (4) studies the possibility that there exists only a one-time habit distortion and this bias does not reinforce with additional interactions. We consider this possibility by only setting $\hat{\tau}_2 = 0$ in Equation (4). The coefficient becomes 2.0 bps, which is 17% smaller than column (2).

Appendix Table B.16 examines the cross-sectional and time-series correlations among $\hat{\tau}_l$. The first three columns imply that the three treatment effects are contemporaneously correlated. When an affected physician has a stronger response to direct manipulation, he also suffers larger habit

Table 7: Economic Magnitudes of Long-term Behavioral Change

This table reports the economic magnitudes of long-term behavioral change. The outcome variables are *LAProb* in the prediction sample in column (1), the predicted probability in column (2), the counterfactual probability without any long-term belief distortion (excluding both τ_1 and τ_2) in column (3), and the counterfactual probability with only one-time belief distortion (excluding τ_2) in column (4). No control variables are included. All regressions include physician fixed effects and area-year fixed effects. Standard errors are clustered at the three-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
		$\hat{LAProb}$		
	LAProb	*Predicted*	$\hat{\tau}_1 = \hat{\tau}_2 = 0$	$\hat{\tau}_2 = 0$
$PF \times Post$	0.023***	0.024***	0.011***	0.020***
	(4.821)	(14.173)	(7.029)	(11.857)
Controls	No	No	No	No
Physician FE	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes
N	314,145	314,145	314,145	314,145
Adj. R^2	0.76	0.69	0.69	0.69

distortion at the same time. For example, a one basis point increase in $\hat{\tau}_0$ is associated with 0.17 – 0.20 bps higher $\hat{\tau}_1$ and $\hat{\tau}_2$ in the same period. Physicians also exhibit persistent patterns in their behavior changes, documented in the remaining columns. For example, column (4) implies that a one basis point increase in $\hat{\tau}_0$ in the previous quarter correlates with a 0.79 bps increase in the current quarter as well. This simple regression has an R^2 of 0.69. A higher reaction to the direct manipulation will also imply larger long-term habit distortions in the next period. This large correlation can be due to persistent patient characteristics or inherent physician preferences.

5.2 Heterogeneity

Which types of affected physicians are less sensitive to the implicit impacts from the Pain CDS? In this section, we explore several heterogeneity tests based on ex-ante characteristics to better understand the interaction between humans and CDS. We first guide our analysis with anecdotal facts about Practice Fusion's strategy. During its communication with Purdue Pharma, it believed the CDS would target "opioid naive" users and utilize their limited knowledge on potential addiction risks. Besides, if physicians had previously prescribed opioids, they might be more attentive to the coverage of opioid risks and litigation of Purdue Pharma, thereby reluctant to accept opioids as a treatment option. So we first hypothesize that the experience of previous opioid usage for affected physicians will mitigate the treatment effects.

To confirm this hypothesis in our empirical study, we interact the treatment variable with the

Table 8: Heterogeneous Effects of Impacts: Long-acting Opioid Experience

This table reports the heterogeneous impacts of Practice Fusion's CDS by physicians' long-acting opioid prescription experience prior to the shock. Columns (1)–(4) present short-term effects using physician-level observations from 2013–2018, while columns (5)–(8) present long-term effects using physician-level observations from 2013–2015 and 2019–2021. PF_i equals one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ equals one if year $t \geq 2016$, and zero otherwise. $Log(LAClaims)_i^{2015}$ is the logarithm of one plus the number of long-acting opioid claims by physician i in 2015, i.e., the pre-shock level of LA opioid usage. $LAProb_{i,t}$ is the prescription probability. $Log(LAClaims)_{i,t}$, $Log(LACost)_{i,t}$, and $Log(LASupply)_{i,t}$ are the logarithms of one plus $LAClaims_{i,t}$, $LACost_{i,t}$, and $LASupply_{i,t}$, respectively. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. All regressions include physician fixed effects and area-year fixed effects. Standard errors are clustered at the three-digit zip code level, and t-statistics are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	Short-term Impacts				Long-term Impacts			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	LAProb	*Log(LAClaims)*	*Log(LACost)*	*Log(LASupply)*	*LAProb*	*Log(LAClaims)*	*Log(LACost)*	*Log(LASupply)*
$PF \times Post$	0.057***	0.181***	0.403***	0.319***	0.127***	0.445***	0.978***	0.788***
	(12.997)	(14.967)	(13.878)	(13.743)	(19.868)	(21.536)	(21.125)	(21.006)
$PF \times Post$	-0.039***	-0.130***	-0.265***	-0.218***	-0.093***	-0.342***	-0.697***	-0.580***
$\times Log(LAClaims)^{2015}$	(-10.691)	(-12.481)	(-11.631)	(-12.001)	(-18.384)	(-21.458)	(-19.238)	(-19.939)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	304,420	304,420	304,420	304,420	291,282	291,282	291,282	291,282
Adj. R^2	0.77	0.80	0.76	0.76	0.70	0.73	0.69	0.70

pre-shock (2015) level of LA opioid claims for each physician in Table 8. The coefficients of all interaction terms are negative with substantial statistical significance. Column (2) implies that a typical treated physician with a pre-existing median-level LA prescription amount (0) will have a treatment effect of 18.1% in the short-term. This impact will reduce to zero if that physician had around three LA claims before the shock. The last four columns document consistent heterogeneity in this alternative sample.

Next, we guide the heterogeneity test based on the feature importance of variables. Recall that leaves are split by cut-off thresholds based on a subset of $\bar{X}_j$. For each variable, feature importance measures the fraction of leaves split using thresholds related to that particular variable. Logically, this measure tells how important each variable is in determining the treatment effect heterogeneity. Appendix Table B.17 shows that $AvgAge$ is the dominant splitting variable, accounting for 48% of splits. The following two are dramatically less important: $DualPct$ (17.5%) and $AvgRisk$ (15.5%). We also believe there exists an intuitive logic for patient age to generate substantial heterogeneity. Physicians may exercise additional caution and invest more effort in considering the side effects of extended-release opioids when facing older patients. Note that feature importance does not necessarily suggest a monotonic heterogeneity pattern. Indeed, we find that higher treatment effects tend to concentrate in middle-level $DualPct$ and $AvgRisk$ groups.

The regression results in both panels of Table 9 support this hypothesis as all interaction terms

Table 9: Heterogeneous Effects of Impacts: Patient Age

This table reports the heterogeneous impacts of Practice Fusion's CDS by patient age. Columns (1)–(4) present short-term effects using physician-level observations from 2013–2018, and columns (5)–(8) present long-term effects using observations from 2013–2015 and 2019–2021. PF_i equals one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ equals one if year $t \geq 2016$, and zero otherwise. $HighAge_i^{2015}$ equals one if the average beneficiary age for physician i in 2015 is above the sample median, and zero otherwise. $LAProb_{i,t}$ is the prescription probability. $Log(LAClaims)_{i,t}$, $Log(LACost)_{i,t}$, and $Log(LASupply)_{i,t}$ are the logarithms of one plus $LAClaims_{i,t}$, $LACost_{i,t}$, and $LASupply_{i,t}$, respectively. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects and area-year fixed effects are included. Standard errors are clustered at the three-digit zip code level, and t-statistics are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	Short-term Impacts				**Long-term Impacts**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	LAProb	*Log(LAClaims)*	*Log(LACost)*	*Log(LASupply)*	*LAProb*	*Log(LAClaims)*	*Log(LACost)*	*Log(LASupply)*
$PF \times Post$	0.032***	0.099***	0.241***	0.190***	0.053***	0.198***	0.481***	0.378***
	(4.383)	(4.785)	(5.033)	(4.914)	(5.433)	(6.346)	(6.851)	(6.743)
$PF \times Post$	-0.023***	-0.076***	-0.167***	-0.143***	-0.030**	-0.156***	-0.331***	-0.280***
$\times HighAge^{2015}$	(-2.696)	(-3.019)	(-2.755)	(-2.967)	(-2.384)	(-4.124)	(-3.849)	(-4.091)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	304,420	304,420	304,420	304,420	291,282	291,282	291,282	291,282
Adj. R^2	0.77	0.79	0.76	0.76	0.70	0.72	0.69	0.69

are significantly negative. In terms of magnitudes, the high-age group has a net treatment effect of 0.9 bps in column (1), which is roughly 30% of the low-age group's effect (3.2 bps). Similarly, the high-age treated group will only increase the number of LA claims by 2.3% in the short-term, which is 39% of the baseline effect and 23% of the low-age group's effect.

6 Conclusion

This paper sheds light on the lasting impacts of CDS interactions on physician prescribing habits in the context of opioid treatment. A biased pain alert designed to promote the use of extended-release opioids led to a significant increase in opioid claims by affected physicians during the treatment window (2016–2018). Notably, this effect persisted even after the removal of the biased alert, indicating a long-term distortion in prescribing behavior.

We believe our results have important implications for both managerial decisions and the regulatory framework governing the use of CDS in supporting treatment decisions. Clinical decision support systems are commonly classified as knowledge-based or non-knowledge-based. In knowledge-based systems, decision rules are developed from literature, practice, or guidelines. In contrast, non-knowledge-based systems rely on claims or observational data inputs, where recommendations are generated using artificial intelligence, machine learning, or statistical pattern recognition, rather than being explicitly coded to follow medical guidelines. The CDS examined in this study was introduced nearly a decade ago and relied on simple rule-based reminders rather than complex algorithms. Nevertheless, understanding how algorithmic or data-driven biases can shape professionals' long-term habits remains broadly relevant in the current era of AI-assisted healthcare for several reasons.

First, with the advancement of retrieval-based models, observational data are increasingly combined with expert knowledge, blurring the distinction between the two CDS categories. Second, even knowledge-based systems often feature algorithmic complexity that obscures their recommendation logic, exposing physicians to errors unconsciously. For this reason, Sutton et al. (2020) show that such systems have not yet achieved widespread implementation. Lastly, our results highlight the need for targeted mitigation procedures. Generic regulatory frameworks that simply flag risks in certain clinical treatments may not sufficiently counteract algorithmic bias. Our findings argue for education and transparency as first-order policy levers: clinicians should receive training on algorithmic risk and have access to concise "model facts" that evaluate the recommendation performance. These steps align with emerging regulatory approaches, such as FDA authorization of autonomous AI diagnostics paired with transparency principles and the EU AI Act's "high-risk" obligations.

That said, our work has limitations. In particular, we cannot quantitatively assess how increasingly complex algorithms in modern CDS will affect physician habits. On the one hand, greater model complexity may make manipulation and bias more difficult to detect, raising the bar for effective monitoring and auditing. On the other hand, increased opacity may lower physicians' trust, encouraging them to rely more heavily on their own judgmental heuristics. Future research should explore how these opposing forces interact.

References

Acemoglu, Daron, 2021, Harms of ai, Technical report, National Bureau of Economic Research.

Agarwal, Nikhil, Alex Moehring, Pranav Rajpurkar, and Tobias Salz, 2023, Combining human expertise with artificial intelligence: Experimental evidence from radiology, Technical report, National Bureau of Economic Research.

Agarwal, R, G Gao, C DesRoches, and AK Jha, 2011, The role of information systems in healthcare: Current research and road ahead, *Information systems research* 22, 419–428.

Ahsen, Mehmet Eren, Mehmet Ulvi Saygi Ayvaci, and Srinivasan Raghunathan, 2019, When algorithmic predictions use human-generated data: A bias-aware classification algorithm for breast cancer diagnosis, *Information Systems Research* 30, 97–116.

Alpert, Abby, William N Evans, Ethan MJ Lieber, and David Powell, 2022, Origins of the opioid crisis and its enduring impacts, *The Quarterly Journal of Economics* 137, 1139–1179.

Angelova, Victoria, Will S Dobbie, and Crystal Yang, 2023, Algorithmic recommendations and human discretion, Technical report, National Bureau of Economic Research.

Atasoy, Hilal, Pei-yu Chen, and Kartik Ganju, 2018, The spillover effects of health it investments on regional healthcare costs, *Management Science* 64, 2515–2534.

Bardhan, Indranil, Rajiv Kohli, Eivor Oborn, Abhay Mishra, Chuan Hoo Tan, Monica Chiarini Tremblay, and Suprateek Sarker, 2025, Human-centric information systems research on the digital future of healthcare, *Information Systems Research* 36, 1–20.

Bardhan, Indranil R, Chenzhang Bao, and Sezgin Ayabakan, 2023, Value implications of sourcing electronic health records: the role of physician practice integration, *Information Systems Research* 34, 1169–1190.

Bardhan, Indranil R, and Mark F Thouin, 2013, Health information technology and its impact on the quality and cost of healthcare delivery, *Decision Support Systems* 55, 438–449.

Bastani, Hamsa, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakcı, and Rei Mariman, 2025, Generative ai without guardrails can harm learning: Evidence from high school mathematics, *Proceedings of the National Academy of Sciences* 122, e2422633122.

Budzyń, Krzysztof, Marcin Romańczyk, Diana Kitala, Paweł Kołodziej, Marek Bugajski, Hans Olov Adami, Johannes Blom, Marek Buszkiewicz, Natalie Grace Halvorsen, Hassan Cesare, et al., 2025, Endoscopist de-skilling after exposure to artificial intelligence in colonoscopy: A multicenter observational study, *The Lancet Gastroenterology & Hepatology* .

Chen, Jiafeng, and Jonathan Roth, 2024, Logs with zeros? some problems and solutions, *The Quarterly Journal of Economics* 139, 891–936.

Doyle, Joseph, Sarah Abraham, Laura Feeney, Sarah Reimer, and Amy Finkelstein, 2019, Clinical decision support for high-cost imaging: a randomized clinical trial, *PLoS One* 14, e0213373.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams, 2016, Sources of geographic variation in health care: Evidence from patient migration, *The quarterly journal of economics* 131, 1681–1726.

Ganju, Kartik K, Hilal Atasoy, Jeffery McCullough, and Brad Greenwood, 2020, The role of decision support systems in attenuating racial biases in healthcare delivery, *Management science* 66, 5171–5181.

Hydari, Muhammad Zia, Rahul Telang, and William M Marella, 2019, Saving patient ryan—can advanced electronic medical records make patient care safer?, *Management Science* 65, 2041–2059.

Jing, Xia, Lina Himawan, and Timothy Law, 2019, Availability and usage of clinical decision support systems (cdsss) in office-based primary care settings in the usa, *BMJ Health & Care Informatics* 26.

Jussupow, Ekaterina, Kai Spohrer, Armin Heinzl, and Joshua Gawlitza, 2021, Augmenting medical diagnosis decisions? an investigation into physicians' decision-making process with artificial intelligence, *Information Systems Research* 32, 713–735.

Keys, Benjamin J, Neale Mahoney, and Hanbin Yang, 2023, What determines consumer financial distress? place- and person-based factors, *The Review of Financial Studies* 36, 42–69.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan, 2018, Human decisions and machine predictions, *The quarterly journal of economics* 133, 237–293.

Kosmyna, Nataliya, Eugene Hauptmann, Ye Tong Yuan, Jessica Situ, Xian-Hao Liao, Ashly Vivian Beresnitzky, Iris Braunstein, and Pattie Maes, 2025, Your brain on chatgpt: Accumulation of cognitive debt when using an ai assistant for essay writing task, *arXiv preprint arXiv:2506.08872* .

Lebovitz, Sarah, Hila Lifshitz-Assaf, and Natalia Levina, 2022, To engage or not to engage with ai for critical judgments: How professionals deal with opacity when using ai for medical diagnosis, *Organization science* 33, 126–148.

Levy, Benjamin, Leonard Paulozzi, Karin A Mack, and Christopher M Jones, 2015, Trends in opioid analgesic–prescribing rates by specialty, us, 2007–2012, *American journal of preventive medicine* 49, 409–413.

Liang, Huigang, and Yajiong Xue, 2022, Save face or save life: Physicians' dilemma in using clinical decision support systems, *Information Systems Research* 33, 737–758.

Logg, Jennifer M, Julia A Minson, and Don A Moore, 2019, Algorithm appreciation: people prefer algorithmic to human judgment, *Organizational Behavior and Human Decision Processes* 151, 90–103.

Macnamara, Brooke N, Ibrahim Berber, M Cenk Çavuşoğlu, Elizabeth A Krupinski, Naren Nallapareddy, Noelle E Nelson, Philip J Smith, Amy L Wilson-Delfosse, and Soumya Ray, 2024, Does using artificial intelligence assistance accelerate skill decay and hinder skill development without performers' awareness?, *Cognitive Research: Principles and Implications* 9, 46.

McCullough, Jeffrey S, Michelle Casey, Ira Moscovice, and Shailendra Prasad, 2010, The effect of health information technology on quality in us hospitals, *Health affairs* 29, 647–654.

Nie, Xinkun, and Stefan Wager, 2021, Quasi-oracle estimation of heterogeneous treatment effects, *Biometrika* 108, 299–319.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan, 2019, Dissecting racial bias in an algorithm used to manage the health of populations, *Science* 366, 447–453.

Rambachan, Ashesh, and Jonathan Roth, 2023, A more credible approach to parallel trends, *Review of Economic Studies* 90, 2555–2591.

Samorani, Michele, Shannon L Harris, Linda Goler Blount, Haibing Lu, and Michael A Santoro, 2022, Overbooked and overlooked: machine learning and racial bias in medical appointment scheduling, *Manufacturing & Service Operations Management* 24, 2825–2842.

Skitka, Linda J, Kathleen L Mosier, and Mark Burdick, 1999, Does automation bias decision-making?, *International Journal of Human-Computer Studies* 51, 991–1006.

Sutton, Reed T, David Pincock, Daniel C Baumgart, Daniel C Sadowski, Richard N Fedorak, and Karen I Kroeker, 2020, An overview of clinical decision support systems: benefits, risks, and strategies for success, *NPJ digital medicine* 3, 17.

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

Volkow, Nora D, and A Thomas McLellan, 2016, Opioid abuse in chronic pain—misconceptions and mitigation strategies, *New England Journal of Medicine* 374, 1253–1263.

Wang, Weiguang, Guodong Gao, and Ritu Agarwal, 2024, Friend or foe? teaming between artificial intelligence and workers with variation in experience, *Management Science* 70, 5753–5775.

Yin, Jiamin, Kee Yuan Ngiam, Sharon Swee-Lin Tan, and Hock Hai Teo, 2025, Designing ai-based work processes: How the timing of ai advice affects diagnostic decision making, *Management Science* .

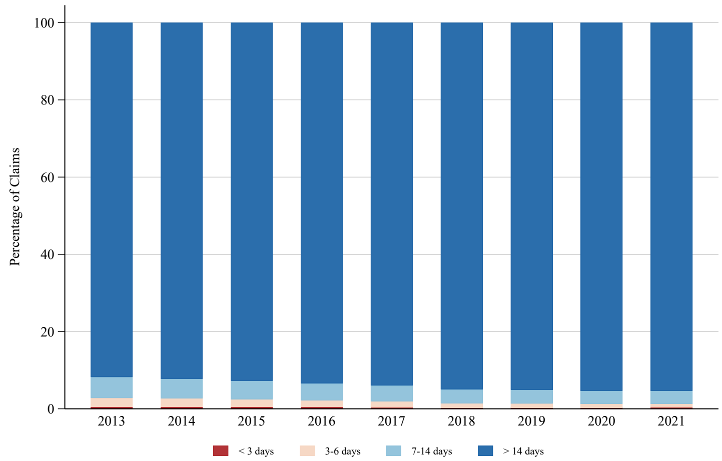
Appendix

A Appendix Figures

Figure A.1: Distribution of OxyContin Claims by Days of Supply

This figure shows the annual distribution of OxyContin prescription claims by days of supply. Panel (a) presents the distribution for initial-use claims, and Panel (b) presents the distribution for all claims, including refills. All claims are drawn from the Marketscan Outpatient Drug Claim database, and OxyContin is identified by National Drug Code (NDC) numbers beginning with 59011.

Panel A: Initial-Use OxyContin Claims by Days of Supply



Panel B: OxyContin Claims by Days of Supply

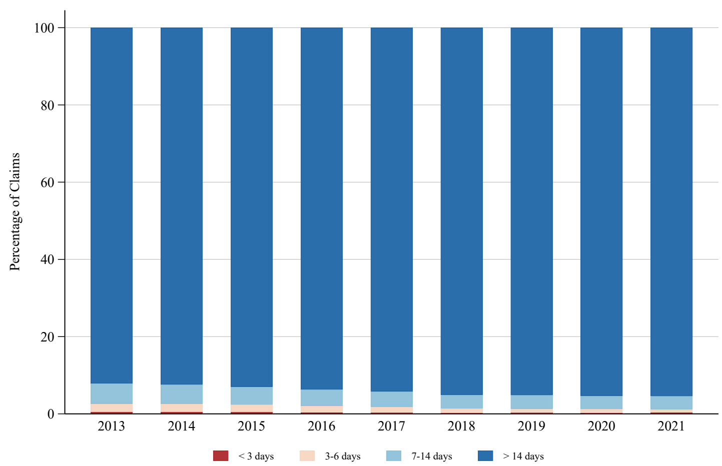


Figure A.2: Robustness of the Parallel Trend: Using 2015 as the Reference Year

This figure plots the treatment dynamics of LA opioid prescribing associated with Practice Fusion's manipulation. Panels (a)–(c) display the β_c coefficients from Equation (2) with 95% confidence intervals (solid lines). The x-axis indicates calendar years. The base year is 2015 (red dashed line), the year before manipulation. The green dashed line marks the suspension of the manipulation in 2019. Shaded regions indicate the 95% confidence interval of the average pre-treatment coefficients from 2013 to 2014.

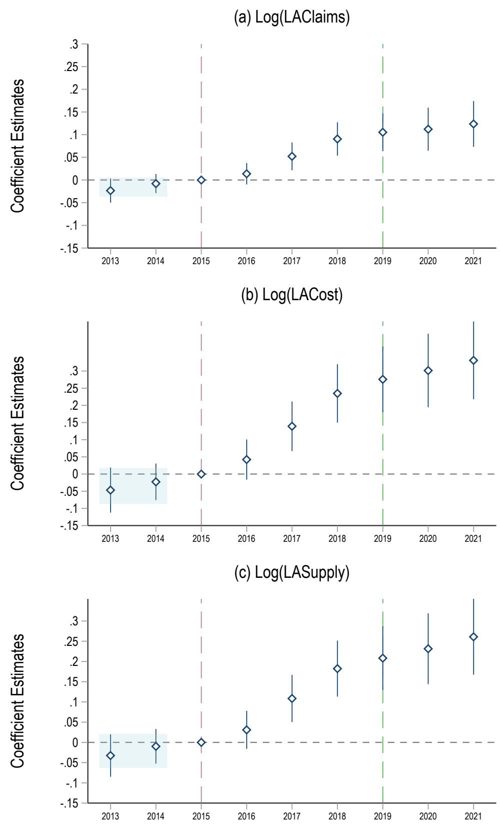


Figure A.3: Robustness of the Parallel Trend: 90% Confidence Interval with HonestDiD

Panels (a)–(c) present results from the HonestDiD procedure of Rambachan and Roth (2023), plotting 90% confidence intervals of the average treatment effects from 2017 to 2021 under alternative bounds on relative magnitudes M. Each M imposes that the post-treatment violation of parallel trends is no more than M times the maximum violation of parallel trends in the pre-treatment period (between consecutive periods).

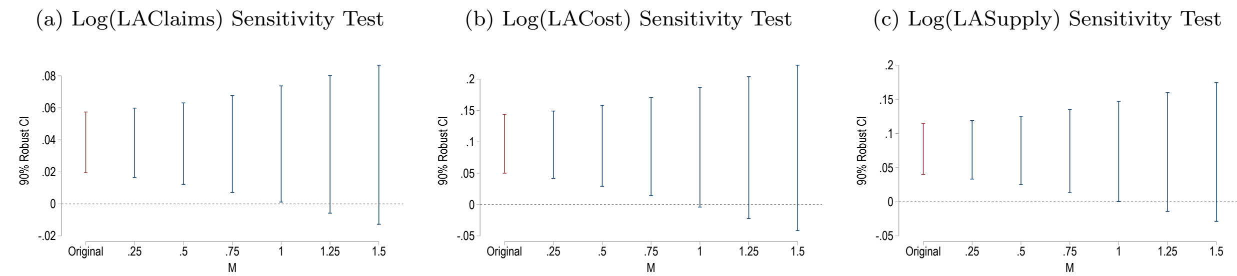
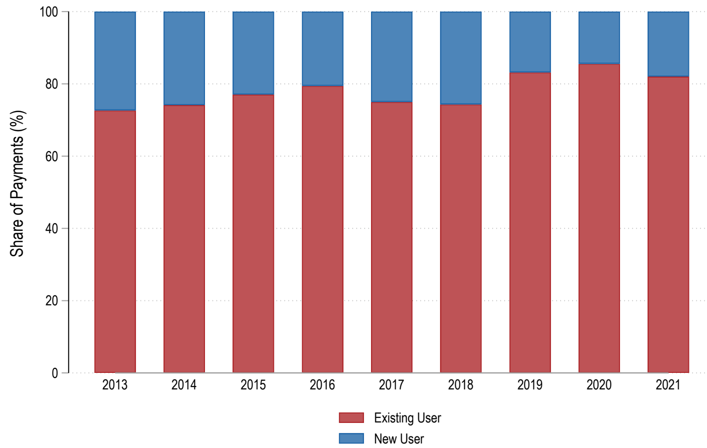


Figure A.4: Distribution of OxyContin Payments

This figure shows the annual distribution of OxyContin prescription payments by patient characteristics. Panel (a) presents the distribution by separating initial users from existing ones, and Panel (b) presents the distribution by separating patients with or without chronic pain conditions. All claims are drawn from the Marketscan Outpatient Drug Claim database, and OxyContin is identified by National Drug Code (NDC) numbers beginning with 59011.

Panel A: Share of OxyContin's Payments by New or Existing Users



Panel B: Share of OxyContin's Payments by Pain Conditions

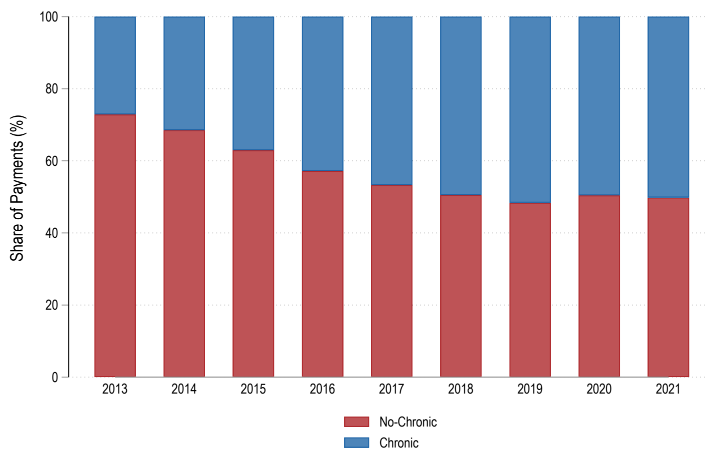


Figure A.5: Time Series of Google Search Interest of Practice Fusion

This figure shows the monthly search interest for "Practice Fusion" based on Google Trends data. The Google Trends index does not report absolute search volumes. Instead, it normalizes search intensity so that the highest observed month (January 2016–December 2022) equals 100. The dashed vertical line marks January 2020, when the U.S. Department of Justice announced its settlement with Practice Fusion.

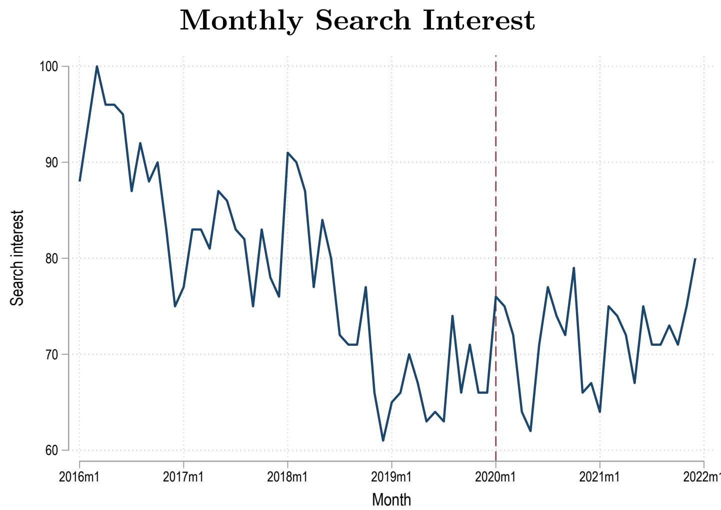
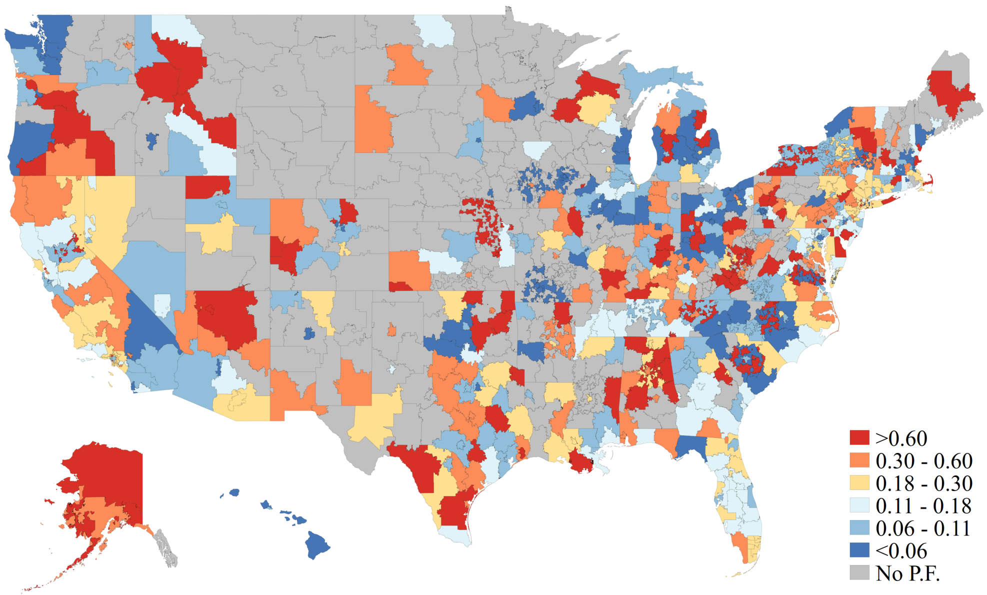


Figure A.6: Number of Recipients of Purdue Pharma Detailing

This figure plots the geographic distribution of Practice Fusion adopters and detailing recipients in our sample. Each area is a 3-digit zip code area. In each area, we calculate the fraction of Practice Fusion adopters and detailing recipients over the total number of physicians in that area from our sample. Areas without Practice Fusion adopters are marked in gray and dropped from the sample in our analysis.

Panel A: Distribution of Practice Fusion Adopters



Panel B: Distribution of Detailing Recipients

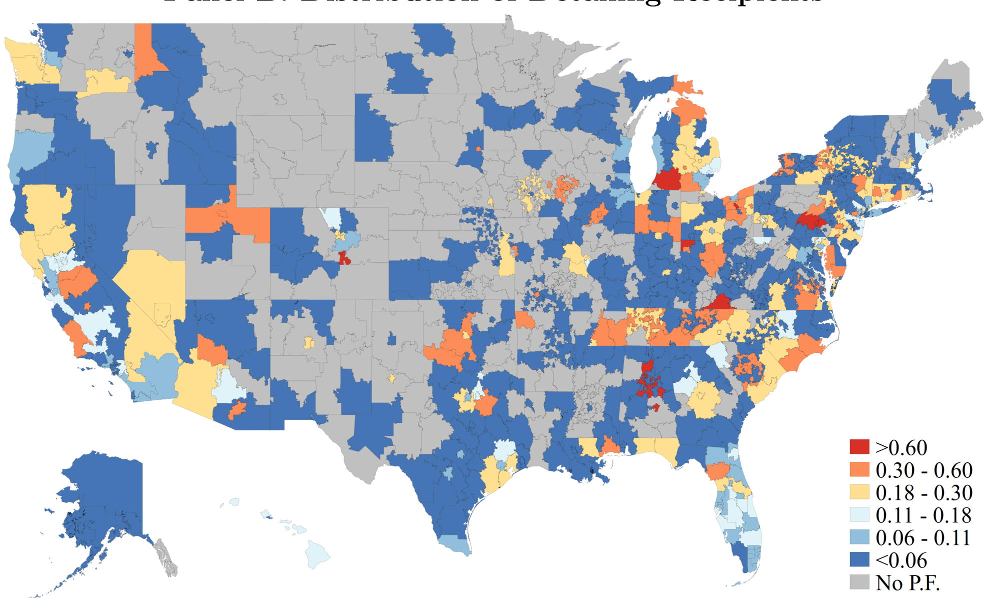
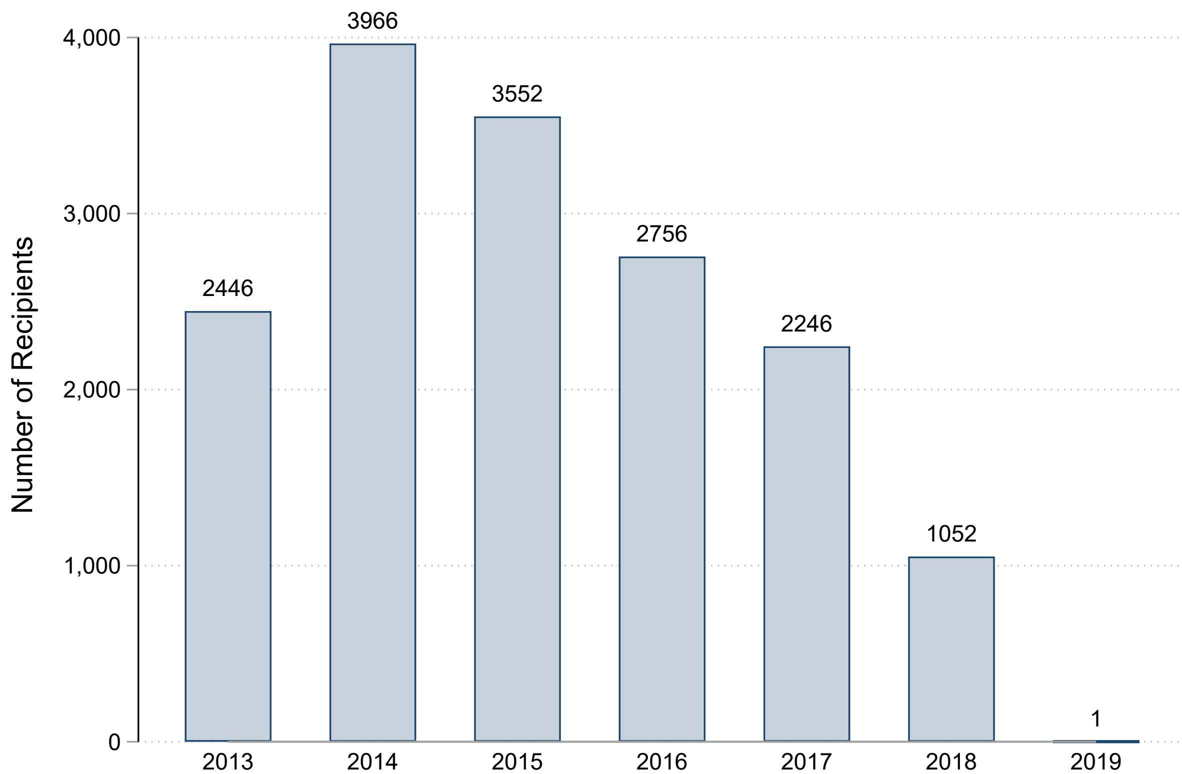


Figure A.7: Number of Recipients of Purdue Pharma Detailing

This figure plots the yearly number of recipients of Purdue Pharma detailing in our sample. We stopped at 2019 since in-kind payments to physicians were effectively suspended after Purdue Pharma filed for bankruptcy in 2019.



B Appendix Tables

Table B.1: Top 5 Specialty in the Sample

This table reports the top five physician specialties in the sample. Percent indicates the share of physicians within each specialty. Specialties are ranked by frequency, from most to least common. Cumulative Percent shows the cumulative share of physicians up to each specialty.

Specialty	**Percent**	**Cumulative Percent**
Family Medicine	31.95	31.95
Internal Medicine	28.18	60.13
Cardiovascular	4.64	64.77
Specialist	2.92	67.69
Gastroenterology	2.80	70.49

Table B.2: Adoption and Switching of Practice Fusion

This table reports the adoption and switching of Practice Fusion. Column (1) studies whether Practice Fusion adoption is associated with more LA claims in 2015. Columns (2) to (4) study whether the pre-treatment LA claims are associated with more Practice Fusion adoption, switching to Practice Fusion, or any kinds of changes in EHR in 2016. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1) *PF*	(2) *PF*	(3) *PF Swtich*	(4) *Any Swtich*
$Log(LAClaims)_t$	-0.049***			
	(-6.197)			
$Log(LAClaims)_{t-1}$		-0.050***	0.006	0.004
		(-6.279)	(0.183)	(0.376)
Sample	2015	2016	2016	2016
N	56,906	55,731	55,731	55,731

Table B.3: Short-term and Long-term Impacts on the Unique Number of Beneficiaries

This table provides the evidence that treated physicians significantly prescribe LA opioid drugs to significantly more patients. PF_i is one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. $LABene_{i,t}$ is the unique number of Medicare beneficiaries receiving LA opioids from physician i in year t. $BeneIncrease_{i,t}$ is one if the unique number of Medicare beneficiaries receiving LA opioids from physician i in year t strictly increases compared to the previous year. $Avg\ Supply_{i,t}$ is the average days of supply per beneficiary in a year, estimated by dividing $LASupply_{i,t}$ with $LABene_{i,t}$. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects, area-specialty fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Panel A: Short-term Impacts

	(1) *Log(LABene)*	(2) *BeneIncrease*	(3) *Avg Supply*	(4) *Log(Avg Supplu)*
$PF \times Post$	0.028***	0.005*	3.094	0.021
	(3.051)	(1.660)	(1.071)	(0.857)
Controls	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes
N	305,082	252,014	24,110	24,110
Adj. R^2	0.72	0.22	0.73	0.77

Panel B: Long-term Impacts

	(1) *Log(LABene)*	(2) *BeneIncrease*	(3) *Avg Supply*	(4) *Log(Avg Supplu)*
$PF \times Post$	0.043***	0.006**	-0.856	0.008
	(2.957)	(1.968)	(-0.158)	(0.186)
Controls	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes
N	292,094	238,963	17,859	17,859
Adj. R^2	0.61	0.19	0.72	0.75

Table B.4: The Long-term Impacts in the Mover Sample: 5-digit Zip Code

This table provides the results on the long-term impacts of Practice Fusion's CDS on opioid prescription in the mover sample. The sample consists of annual observations among the physicians that move to a new 5-digit zip code after 2019. The sample period is from 2013 to 2015, and 2019 to 2021, i.e. the pre-treatment and post-treatment phases. PF_i is one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. $Log(LAClaims)_{i,t}$ is the logarithm of one plus $LAClaims_{i,t}$. $Log(LACost)_{i,t}$ is the logarithm of one plus $LACost_{i,t}$. $Log(LASupply)_{i,t}$ is the logarithm of one plus $LASupply_{i,t}$. $LARate_{i,t}$ is the percentage of long-acting opioid claims out of all opioid claims by physician i in year t. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*	(4) *LARate*
$PF \times Post$	0.198***	0.510***	0.428***	1.060**
	(2.689)	(3.106)	(3.215)	(2.513)
Controls	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes
N	24,852	24,852	24,852	24,852
Adj. R^2	0.68	0.64	0.65	0.57

Table B.5: The Long-term Impacts in the Mover Sample: State

This table provides the results on the long-term impacts of Practice Fusion's CDS on opioid prescription in the mover sample. The sample consists of annual observations among the physicians that move to a new state after 2019. The sample period is from 2013 to 2015, and 2019 to 2021, i.e. the pre-treatment and post-treatment phases. PF_i is one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. $Log(LAClaims)_{i,t}$ is the logarithm of one plus $LAClaims_{i,t}$. $Log(LACost)_{i,t}$ is the logarithm of one plus $LACost_{i,t}$. $Log(LASupply)_{i,t}$ is the logarithm of one plus $LASupply_{i,t}$. $LARate_{i,t}$ is the percentage of long-acting opioid claims out of all opioid claims by physician i in year t. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*	(4) *LARate*
$PF \times Post$	0.395***	0.979***	0.835***	2.452**
	(3.403)	(3.613)	(3.841)	(2.401)
Controls	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes
N	5,156	5,156	5,156	5,156
Adj. R^2	0.59	0.56	0.56	0.51

Table B.6: The Long-term Impacts in the Mover Sample: Any Affiliation Changes

This table provides the results on the long-term impacts of Practice Fusion's CDS on opioid prescription in the mover sample. The sample consists of annual observations among the physicians that have any affiliation changes after 2019. The sample period is from 2013 to 2015, and 2019 to 2021, i.e. the pre-treatment and post-treatment phases. PF_i is one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. $Log(LAClaims)_{i,t}$ is the logarithm of one plus $LAClaims_{i,t}$. $Log(LACost)_{i,t}$ is the logarithm of one plus $LACost_{i,t}$. $Log(LASupply)_{i,t}$ is the logarithm of one plus $LASupply_{i,t}$. $LARate_{i,t}$ is the percentage of long-acting opioid claims out of all opioid claims by physician i in year t. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*	(4) *LARate*
$PF \times Post$	0.103***	0.275***	0.197***	0.369**
	(3.451)	(4.042)	(3.527)	(2.228)
Controls	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes
N	187,524	187,524	187,524	187,524
Adj. R^2	0.72	0.69	0.69	0.61

Table B.7: States with Opioid Use Limits

This table provides the results on the long-term impacts of Practice Fusion's CDS on opioid prescription in the sample of states with limits on days of supply for opioid prescription. PF_i is one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. $Log(LAClaims)_{i,t}$ is the logarithm of one plus $LAClaims_{i,t}$. $Log(LACost)_{i,t}$ is the logarithm of one plus $LACost_{i,t}$. $Log(LASupply)_{i,t}$ is the logarithm of one plus $LASupply_{i,t}$. $Log(SAClaims)_{i,t}$ is the logarithm of one plus number of short-acting opioid claims. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	Short-term Impacts				**Long-term Impacts**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log(LAClaims)	*Log(LACost)*	*Log(LASupply)*	*Log(SAClaims)*	*Log(LAClaims)*	*Log(LACost)*	*Log(LASupply)*
$PF \times Post$	0.057***	0.143***	0.110***	-0.014	0.092***	0.258***	0.189***
	(2.891)	(3.185)	(3.072)	(-0.596)	(2.676)	(3.394)	(3.028)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	145,453	145,453	145,453	145,453	139,125	139,125	139,125
Adj. R^2	0.80	0.76	0.77	0.87	0.73	0.70	0.70

Table B.8: States with Comprehensive PDMP Adoption by 2018

This table provides the results on the long-term impacts of Practice Fusion's CDS on opioid prescription in the sample of states with comprehensive PDMP adoption by 2018. PF_i is one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. $Log(LAClaims)_{i,t}$ is the logarithm of one plus $LAClaims_{i,t}$. $Log(LACost)_{i,t}$ is the logarithm of one plus $LACost_{i,t}$. $Log(LASupply)_{i,t}$ is the logarithm of one plus $LASupply_{i,t}$. $Log(SAClaims)_{i,t}$ is the logarithm of one plus number of short-acting opioid claims. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	Short-term Impacts				**Long-term Impacts**		
	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*	(4) *Log(SAClaims)*	(5) *Log(LAClaims)*	(6) *Log(LACost)*	(7) *Log(LASupply)*
$PF \times Post$	0.088***	0.199***	0.158***	0.007	0.133***	0.338***	0.266***
	(3.353)	(3.369)	(3.303)	(0.208)	(2.883)	(3.457)	(3.257)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	89,956	89,956	89,956	89,956	86,096	86,096	86,096
Adj. R^2	0.79	0.76	0.76	0.87	0.73	0.69	0.70

Table B.9: States with Comprehensive PDMP Adoption by 2021

This table provides the results on the long-term impacts of Practice Fusion's CDS on opioid prescription in the sample of states with comprehensive PDMP adoption by 2021. PF_i is one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. $Log(LAClaims)_{i,t}$ is the logarithm of one plus $LAClaims_{i,t}$. $Log(LACost)_{i,t}$ is the logarithm of one plus $LACost_{i,t}$. $Log(LASupply)_{i,t}$ is the logarithm of one plus $LASupply_{i,t}$. $Log(SAClaims)_{i,t}$ is the logarithm of one plus number of short-acting opioid claims. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	Short-term Impacts				**Long-term Impacts**		
	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*	(4) *Log(SAClaims)*	(5) *Log(LAClaims)*	(6) *Log(LACost)*	(7) *Log(LASupply)*
$PF \times Post$	0.086***	0.212***	0.164***	-0.000	0.136***	0.350***	0.267***
	(4.899)	(5.353)	(5.069)	(-0.019)	(4.748)	(5.584)	(5.159)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	201,139	201,139	201,139	201,139	192,769	192,769	192,769
Adj. R^2	0.80	0.76	0.76	0.87	0.73	0.69	0.70

Table B.10: States with High Google Search Interest in 2020 and 2021

This table provides the results on the long-term impacts of Practice Fusion's CDS on opioid prescription by whether physicians are in states with high Google search interest after the settlement. This sample has the pre-treatment and post-treatment periods. PF_i is one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. $HighAware$ is one if physician i's states have high (above median) search interest of "Practice Fusion" on Google in 2020 and 2021, and zero otherwise. $HighAware$ is not absorbed by Physician FEs because physicians may move to new states. $Log(LAClaims)_{i,t}$ is the logarithm of one plus $LAClaims_{i,t}$. $Log(LACost)_{i,t}$ is the logarithm of one plus $LACost_{i,t}$. $Log(LASupply)_{i,t}$ is the logarithm of one plus $LASupply_{i,t}$. $Log(SAClaims)_{i,t}$ is the logarithm of one plus number of short-acting opioid claims. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*
$PF \times Post$	0.223***	0.505***	0.393***
	(4.392)	(4.840)	(4.519)
$PF \times Post$	-0.135**	-0.249**	-0.204**
$\times HighAware$	(-2.359)	(-2.068)	(-2.046)
$HighAware$	-0.034	-0.081	-0.060
	(-0.930)	(-0.943)	(-0.857)
Controls	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes
N	291,530	291,530	291,530
Adj. R^2	0.72	0.69	0.69

Table B.11: Explicit Detailing

This table provides the results on the impacts of Purdue Pharma's detailing on opioid prescription. $Pudue_{i,t-1}$ is one if physician i has received in-kind payments from Purdue Pharma at $t-1$, and zero otherwise. $Detail_i$ is one if physician i had received in-kind payments from Purdue Pharma by 2019, and zero otherwise. PF_i is one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. $Log(LAClaims)_{i,t}$ is the logarithm of one plus $LAClaims_{i,t}$. $Log(LACost)_{i,t}$ is the logarithm of one plus $LACost_{i,t}$. $Log(LASupply)_{i,t}$ is the logarithm of one plus $LASupply_{i,t}$. $LARate_{i,t}$ is the percentage of long-acting opioid claims out of all opioid claims by physician i in year t. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Panel A: Short-term Impacts

	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*	(4) *LARate*
$PF \times Post$	0.055***	0.147***	0.109***	0.209**
	(3.703)	(4.322)	(3.974)	(2.393)
$Purdue_{t-1} \times Post$	-0.133***	-0.221***	-0.186***	-0.617***
	(-6.625)	(-4.744)	(-5.021)	(-5.254)
$Purdue_{t-1}$	0.190***	0.399***	0.323***	0.675***
	(9.716)	(8.665)	(8.634)	(6.728)
Controls	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes
N	251,879	251,879	251,879	251,879
Adj. R^2	0.81	0.78	0.78	0.72

Panel B: Long-term Impacts

	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*	(4) *LARate*
$Detail \times Post$	-0.425***	-0.805***	-0.661***	-2.191***
	(-5.606)	(-4.866)	(-5.026)	(-4.969)
Controls	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes
N	283,358	283,358	283,358	283,358
Adj. R^2	0.71	0.68	0.68	0.60

Table B.12: Robustness Check: 5-digit Zip Code Areas

This table shows the robustness of Tables 2 by requiring the control group to be in the same 5-digit zip code areas and with the same specialty of the treatment group. Panel A estimates the effects in the short-term sample (2013 – 2018) and Panel B in the long-term sample (2013 – 2015 & 2019 – 2021). PF_i is one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. $Log(LAClaims)_{i,t}$ is the logarithm of one plus $LAClaims_{i,t}$. $Log(LACost)_{i,t}$ is the logarithm of one plus $LACost_{i,t}$. $Log(LASupply)_{i,t}$ is the logarithm of one plus $LASupply_{i,t}$. $LARate_{i,t}$ is the percentage of long-acting opioid claims out of all opioid claims by physician i in year t. $Log(SAClaims)_{i,t}$ is the logarithm of one plus number of short-acting opioid claims. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Panel A: Short-term Effects

	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*	(4) *LARate*	(5) *Log(SAClaims)*
$PF \times Post$	0.070***	0.172***	0.133***	0.403***	−0.004
	(3.895)	(4.209)	(4.010)	(3.741)	(−0.191)
Controls	Y	Y	Y	Y	Y
Area×Year FEs	Y	Y	Y	Y	Y
Physician FEs	Y	Y	Y	Y	Y
N	86,934	86,934	86,934	86,934	86,934
R^2	0.79	0.75	0.76	0.71	0.87

Panel B: Long-term Effects

	(1) *Log(LAClaims)*	(2) *Log(LACost)*	(3) *Log(LASupply)*	(4) *LARate*
$PF \times Post$	0.145***	0.376***	0.291***	0.500***
	(5.305)	(6.121)	(5.783)	(3.249)
Controls	Y	Y	Y	Y
Area×Year FEs	Y	Y	Y	Y
Physician FEs	Y	Y	Y	Y
N	83,059	83,059	83,059	83,059
R^2	0.72	0.69	0.69	0.63

Table B.13: Alternative Clustering by Vendors

This table shows the placebo test of Tables 2 using alternative clustering methods. Both panels follow from Table 2, except that standard errors are clustered at the vendor level in Panel A, and double clustered at the vendor and the 3-digit zip code level in panel B. t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Panel A: Clustered by Vendor ID

	Short-term Impacts			Long-term Impacts		
	(1)	(2)	(3)	(4)	(5)	(6)
	Log(LAClaims)	*Log(LACost)*	*Log(LASupply)*	*Log(LAClaims)*	*Log(LACost)*	*Log(LASupply)*
$PF \times Post$	0.116***	0.308***	0.231***	0.059***	0.154***	0.114***
	(5.526)	(6.958)	(6.453)	(5.574)	(6.408)	(6.004)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	292,094	292,094	292,094	305,082	305,082	305,082
Adj. R^2	0.73	0.69	0.69	0.79	0.76	0.76

Panel B: Clustered by Vendor ID and Zip Codes

	Short-term Impacts			Long-term Impacts		
	(1)	(2)	(3)	(4)	(5)	(6)
	Log(LAClaims)	*Log(LACost)*	*Log(LASupply)*	*Log(LAClaims)*	*Log(LACost)*	*Log(LASupply)*
$PF \times Post$	0.116***	0.308***	0.231***	0.059***	0.154***	0.114***
	(5.521)	(6.921)	(6.347)	(5.288)	(6.285)	(5.887)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	292,094	292,094	292,094	305,082	305,082	305,082
Adj. R^2	0.72	0.69	0.69	0.79	0.76	0.76

Table B.14: Placebo Test: Antibiotics and Antipsychotic

This table shows the placebo test of Tables 2 using Antibiotics and Antipsychotic. Panel A estimates the effects in the short-term sample (2013 – 2018) and Panel B in the long-term sample (2013 – 2015 & 2019 – 2021). PF_i is one if physician i adopts Practice Fusion's EHR, and zero otherwise. $Post_t$ is one if year t is greater or equal to 2016, and zero otherwise. Columns (1) and (2) are for Antibiotics and columns (3) and (4) are Antipsychotic. $Claims_{i,t}$ is the number of claims by by physician i in year t for the corresponding drug. $DrugCost_{i,t}$ is the dollar amount of drug costs by physician i in year t for the corresponding drug. Control variables include $AvgAge$, $MalePct$, $BlackPct$, $HispanicPct$, $DualPct$, and $AvgRisk$. Physician fixed effects, area-specialty fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Panel A: Short-term Impacts

	(1)	(2)	(3)	(4)
	Antibiotics		**Antipsychotic**	
	Log(Claims)	*Log(DrugCosts)*	*Log(Claims)*	*Log(DrugCosts)*
$PF \times Post$	-0.025*	-0.012	-0.018	-0.023
	(-1.780)	(-0.515)	(-1.466)	(-0.785)
Controls	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes
N	305,082	305,082	223,193	223,193
Adj. R^2	0.86	0.81	0.91	0.89

Panel B: Long-term Impacts

	(1)	(2)	(3)	(4)
	Antibiotics		**Antipsychotic**	
	Log(Claims)	*Log(DrugCosts)*	*Log(Claims)*	*Log(DrugCosts)*
$PF \times Post$	-0.032	-0.018	-0.034*	-0.016
	(-1.478)	(-0.484)	(-1.911)	(-0.395)
Controls	Yes	Yes	Yes	Yes
Physician FE	Yes	Yes	Yes	Yes
Area×Year FE	Yes	Yes	Yes	Yes
N	292,094	292,094	211,315	211,315
Adj. R^2	0.82	0.77	0.88	0.86

Table B.15: OLS Estimation of Equation (3)

This table estimates Equation (3) using the OLS regression:

$$p_{i,t} = \tau_0 \mathbf{I}\{i \in \mathbf{T}, t \geq 2016\} + \tau_1 \mathbf{I}\{i \in \mathbf{T}, t \geq 2017\} + \tau_2 \mathbf{I}\{i \in \mathbf{T}, t \geq 2018\} + \delta_i + \mu_{k,t} + \varepsilon_{i,t}.$$

$p_{i,t}$ is the prescription probability of physician i in year t. $\mathbf{I}\{i \in \mathbf{T}, t \geq s\}$ indicates whether physician i is in the treatment group and the calendar is greater or equal to s. No control variables are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)
	$p_{i,t}$
τ_0	0.012***
	(2.653)
τ_1	0.011***
	(2.672)
τ_2	0.012***
	(3.094)
Controls	No
Physician FE	Yes
Area×Year FE	Yes
N	314,145
Adj. R^2	0.76

Table B.16: Correlation between Direct Manipulation and Long-term Belief Distortion

This table exhibits the correlation between effects of direct manipulation and long-term belief distortion. Columns (1) to (3) are estimated in sample years 2016 to 2018. The remaining columns are estimated in sample years 2017 and 2018 due to the lagged explanatory variables. No control variables and fixed effects are included. Physician fixed effects and area-year fixed effects are included. Standard errors are clustered at the 3-digit zip code level, and t-statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_2$	$\hat{\tau}_0$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_2$
$\hat{\tau}_0$	0.201***	0.174***	0.106***				
	(35.468)	(38.163)	(19.864)				
$\hat{\tau}_1$			0.336***				
			(50.583)				
$\hat{\tau}_{0,t-1}$				0.788***	0.213***	0.169***	0.087***
				(122.570)	(35.793)	(44.368)	(18.551)
$\hat{\tau}_{1,t-1}$							0.405***
							(67.119)
Controls	155,766	155,766	155,766	155,176	155,176	155,176	155,176
Physician FE	0.10	0.07	0.17	0.70	0.12	0.08	0.23

Table B.17: Variable Importance

This table exhibits variable importance in estimating the causal forest. Importance is calculated as a weighted sum of how many times feature j was split on at each depth in the forest, defined as

$$imp(x_j) = \frac{\sum_{k=1}^{4} \left(\frac{\text{\# Tree splits on } x_j \text{ at depth } k}{\text{\# All tree splits at depth } k} \right) k^{-2}}{\sum_{k=1}^{4} k^{-2}}.$$

Seniority is the number of years since physician's graduation in medical school. *PhysicanGender* is one if physician is female and zero otherwise. Other variables are defined in Table 1.

(1) *Variable*	(2) *Importance*
AvgAge	0.481
DualPct	0.175
AvgRisk	0.150
MalePct	0.059
Seniority	0.055
BlackPct	0.046
HispanicPct	0.030
PhysicanGender	0.002

C Model

Baseline Model Our conceptual framework features a physician's decision making process by maximizing the chance of correct prescription in a setting of dynamic learning. For physicians i facing case j at year t, they take a binary decision $a_{ij} \in \{0, 1\}$, where $a_{ij} = 1$ indicates prescribing LA opioids. The underlying state $\omega_j \in \{0, 1\}$, unobserved to the physician, indicates the true fitness of opioid usage. Physicians have disutility from both unnecessary prescription (choosing $a = 1$ when $\omega = 0$) and under-treatment (choosing $a = 0$ when $\omega = 1$), with payoffs modeled as

$$u_i = -\mathbf{I}\{a = 1, w = 0\}c_1 - \mathbf{I}\{a = 0, w = 1\}c_2. \tag{C.1}$$

In each case j, physicians observe patient characteristics Z_j and possibly a Practice Fusion pain alert, indicated by $s \in \{0, 1\}$. They then form a belief $p_{it}(\omega|Z_j, s)$ and generate the likelihood ratio as

$$l_{it}(Z_j, s) = \frac{p_{it}(\omega = 1|Z_j, s)}{p_{it}(\omega = 0|Z_j, s)} = \underbrace{\frac{p_{it}(\omega = 1|Z_j)}{p_{it}(\omega = 0|Z_j)}}_{q_{it}(Z_j)} \times \underbrace{\frac{p_i(s|Z_j, \omega = 1)}{p_i(s|Z_j, \omega = 0)}}_{\delta_i(Z_j,s)}.$$

The decomposition follows from the Bayes rule and has an intuitive interpretation. $q_{it}(Z_j)$ is the prior belief of physician i observing Z_j in case j at year t. $\delta_i(Z_j, s)$ represents the periodic belief update due to CDS alerts. For simplicity, we assume δ_i is a static function (independent of t). Physicians in the control group do not suffer from possible short-term distortions so by definition, $\delta_i(Z_j, 0) = 1$. To model the long-term belief distortion, denote k by the number of years that a physician j has observed CDS alerts by t. We assume that every additional period of exposure to CDS manipulation will bias the likelihood upward by a certain degree:

$$q_{it}(Z_j) = q_{i0}(Z_j) \prod_{m=1}^{k} \gamma_i^m.$$

$q_{i0}(Z_j)$ is the initial belief without learning. γ_i^m is the belief bias due to the m^{th} interaction with manipulated CDS. Note that γ_i^m is implicitly a function of patient characteristics in that interaction. Finally, the payoff function in Equation (C.1) implies that physicians follow a simple cut-off strategy by

$$a_{it}(Z_j, s) = \mathbf{1}\left\{ q_{i0}(Z_j) \times \left(\prod_{m=1}^{k} \gamma_i^m \right) \times \delta_i(Z_j, s) \geq \frac{c_1}{c_2} \right\}. \tag{C.2}$$

The main prediction of this conceptual framework is that the treatment group will prescribe more LA opioids relative to the control group from 2016 to 2021, if $\left(\prod_{m=1}^{k} \gamma_i^m\right) \times \delta_i(Z_j, s) > 1$. To illustrate, consider two otherwise similar physicians but one uses Practice Fusion (treated) and one does not (control) in the following three cases. First, in the year of 2016 (the initial treatment period), both physicians have not interacted with manipulated CDS, and thus the second component in Equation (C.2) is one. The treated physician may have a higher prescription volume only due to the instantaneous belief update $\delta_i(Z_j, s)$. Second, in the year of 2020 (a post-treatment period), the last component equals one since the alerts are removed. But the treated physicians still have higher chances of prescription due to $\prod_{m=1}^{3} \gamma_i^m$ (three years of previous interaction). Lastly, in the year of 2017 (the second treatment period), the treatment group suffers both manipulation $\delta_i(Z_j, s)$

and belief distortion γ_i^1 after one period of learning.

Estimation via Multi-Arm Causal Forest Our data are aggregated at the physician-year level without the granular information of each visit j's decision a_{ij} and patient characteristics Z_j. Therefore, we need to estimate the expected probability of prescription based on Equation (C.2):

$$Pr(a_{it} = 1|\bar{X}_i) = \int \mathbf{1}\left\{ q_{i0}(z) \times \left(\prod_{m=1}^{k} \gamma_i^m \right) \times \delta_i(z,s) \geq \frac{c_1}{c_2} \right\} f\left(z|\bar{X}_i\right) dz. \tag{C.3}$$

In the above equation, $f\left(z|\bar{X}_i\right)$ is the conditional distribution of patient characteristics. We assume this distribution is characterized by physician-level variables $\bar{X}_i$, including the average patient beneficiary age, risk score, male fraction, African American fraction, Hispanic fraction, dual qualification (Medicaid and Medicare) fraction, physician seniority (years since graduation) and physician gender. Denote $p^T(\bar{X}_i)$ and $p^C(\bar{X}_i)$ as the equilibrium prescription probability for the treatment group and control group respectively. Following Equation (C.3), we can decompose $p^T(\bar{X}_i)$ into

$$\begin{aligned}
& p^T(\bar{X}_i) \\
= \quad & \int \mathbf{1}\left\{ q_{i0}(z) \geq \frac{c_1}{c_2} \right\} f\left(z|\bar{X}_i\right) dz \\
+ \quad & sgn\left(\delta_i - 1\right) \int \mathbf{1}\left\{ \frac{c_1}{c_2} > q_{i0}(z) \geq \frac{c_1}{c_2 \delta_i} \right\} f\left(z|\bar{X}_i\right) dz \\
+ \quad & \sum_{l=1}^{k} sgn\left(\gamma_i^l - 1\right) \int \mathbf{1}\left\{ \frac{c_1}{c_2 \delta_i \times \left(\prod_{m=1}^{l-1} \gamma_i^m\right)} > q_{i0}(z) \geq \frac{c_1}{c_2 \delta_i \times \left(\prod_{m=1}^{l} \gamma_i^m\right)} \right\} f\left(z|\bar{X}_i\right) dz.
\end{aligned}$$

Each line in the above equation has an intuitive interpretation. The second line represents the expected probability for a similar physician in the control group, i.e. $p^C(\bar{X}_i)$. The third line represents the marginal treatment effect due to direct manipulation, denoted by $\tau_0(\bar{X}_i)$. In the last line, each summation term represents the additional long-term impact due to the lth interaction, denoted by $\tau_l(\bar{X}_i)$.