Breaking Bad Opioid Sorting*

Han Ng† Stephan Sagl‡ Jia Xiang§

February 15, 2026

PRELIMINARY AND INCOMPLETE - DO NOT CITE

Abstract

In many markets, consumers rely on experts for access to goods and services. Consumer and expert preferences over outcomes are linked through consumers' choice of experts, or sorting. We investigate how consumer sorting contributes to observed differences in outcomes and its implications for expert-targeted policies. Using rich employer-sponsored health insurance claims data on opioid prescriptions for chronic pain, we first show that prescription intensity is highly dispersed across physicians. We decompose the variance of physicians' prescription decisions and find that patient sorting is more than three times as important as physicians' inherent prescription propensity for prescription intensity dispersion. Most of this sorting cannot be justified on medical grounds. We develop and estimate an equilibrium model of patient choice of physicians and physicians' prescription decisions. Patients optimally choose their physicians based on both their opioid preferences and their expectations of physicians' prescription decisions. Our counterfactual analysis shows that expert-targeted policies to curb non-medically grounded opioid prescribing will be severely attenuated by patient resorting. We propose an alternative policy that eliminates sorting based on non-medically grounded preferences, while preserving appropriate care for patients with medical needs for prescription opioids.

*We thank participants of the IU Kelley Brownbag for helpful comments. We thank Richard Meraz for his support in accessing the data used in this paper. This research was supported in part by Lilly Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute.

†Institute of Economics, Academia Sinica. Email: hlng@econ.sinica.edu.tw.

‡Kelley School of Business, Indiana University. Email: ssagl@iu.edu.

§Kelley School of Business, Indiana University. Email: xiangjia@iu.edu.

1 Introduction

In many markets, consumers rely on experts for access to goods and services. For instance, investors turn to asset managers to access complex investment opportunities, and patients consult physicians for medical advice and treatment. The stakes of choosing the right expert can be substantial: selecting the wrong asset manager may expose investors to levels of financial risk that greatly differ from their preferences, and choosing the wrong physician can significantly impact long-term health. Given the high stakes, consumers have strong incentives to sort into experts, choosing an expert whose preferences, style, or philosophy closely align with their own. As a result, market outcomes reflect both expert preferences and consumer preferences, linked together through consumer sorting. Despite the importance of disentangling these forces for policy design, empirical evidence quantifying the contribution of each channel to observed outcomes remains extremely limited.

In this paper, we study the importance of consumer sorting versus expert preferences for observed outcome differences and their implications for expert-targeted policies. Our paper's focus concerns opioid prescribing for chronic pain patients in the United States. In this market, patients must obtain prescription opioids through physicians, but in most cases can freely choose their physician. Using employer-sponsored health insurance claims data, we show that patients' choice of physicians, or patient sorting, is more than three times as important in explaining the observed dispersion in opioid prescribing rates across physicians as physicians' inherent prescribing propensity. Moreover, patient sorting is almost exclusively driven by factors unrelated to patients' medical needs. This finding suggests that policies targeting top-prescribing physicians aiming to reduce clinically unwarranted opioid prescribing may be severely attenuated by patient resorting. To understand how patient sorting interacts with top physician-targeted policies, we develop and estimate a model of equilibrium patient sorting and physician prescribing. We use this model to simulate several counterfactuals. First, mimicking top expert-targeted policies, we consider a counterfactual in which the top 10% of physicians by risk-adjusted prescription rate are excluded from the market. Despite being extreme, this policy reduces aggregate prescribing by only 12%. We show that this result is driven by patients resorting to non-excluded physicians with similar prescription propensities as the excluded physicians. In other words, resorting severely attenuates the exclusion policy. If the affected patients were instead randomly reassigned to physicians, aggregate prescribing would fall by approximately 22%. However, randomization breaks sorting indiscriminately, disregarding patients' medical needs. To address this issue, we consider a counterfactual in which we force physicians to prescribe solely based on medical needs but disregard patients' opioid preferences, mirroring the CDC guideline for

prescribing opioids (Dowell, Ragan, et al. 2022a). This eliminates patients' incentives to sort based on non-medical preferences. Breaking only this bad sorting, we find that this policy would reduce aggregate opioid prescriptions by 21% while preserving treatment for patients with medical needs.

We establish these findings using health insurance claims data from the Merative MarketScan Research Databases for the New York City Metropolitan Statistical Area over the period 2016-2022. These visit-level data allow us to link patients, their chosen physicians, and any resulting opioid prescriptions. We begin by documenting that physicians' prescription rates are highly dispersed, even after controlling for a rich set of patient severity measures. Only a relatively small fraction of physicians account for most opioid prescriptions. This pattern aligns with the well-documented heterogeneity in physicians' prescription intensity reported in the literature (Finkelstein, Gentzkow, and Williams 2016; Molitor 2018; Cutler, Skinner, Stern, and Wennberg 2019; Badinski, Finkelstein, Gentzkow, and Hull 2023; Clemens, Léger, Nandan, and Town 2024).

Next, we assess the role of patient sorting in the dispersion of prescription rates across physicians. To do so, we decompose the variance of physicians' prescription decisions into patient fixed effects, physician fixed effects, time-varying patient controls, and a patient sorting component, following the well-known two-way fixed effects approach developed by Abowd, Kramarz, and Margolis (1999). There are two identification challenges. First, physician fixed effects are identified from patients who switch physicians. When switching is limited, fixed effects estimates can be severely biased, a problem known as limited mobility bias (Andrews, Gill, Schank, and Upward 2012). We follow recent advances in the labor literature to address the limited-mobility bias by grouping physicians and patients via k-means clustering and estimating group-level fixed effects (Bonhomme, Lamadon, and Manresa 2019; Mourot 2025). Second, fixed effects estimates may be biased if patients' switching decisions are due to time-varying characteristics unobserved by the econometrician. We provide evidence to mitigate this concern following Card, Cardoso, Heining, and Kline (2018).

Using the two-way fixed effects estimator, we find that patient sorting is more than three times as important as physicians' inherent prescribing propensity in explaining the observed dispersion in opioid prescribing rates across physicians. This suggests that patients' choice rather than physicians' inherent prescribing propensity is primarily responsible for observed prescribing heterogeneity across physicians. We further investigate the determinants of patient sorting. In particular, we separate sorting into sorting due to medical needs and sorting due to non-medical opioid preferences. We find that medically grounded preferences explain little of the variation in opioid prescribing rates. Instead, non-medical preferences account for the majority of the observed dispersion. Accordingly, sorting on non-medical preferences

constitutes most of patient sorting in this market. Our findings are consistent with anecdotal evidence from online forums, where patients with strong preferences for opioids actively seek and share information about physicians' propensity to prescribe.[1] Moreover, our findings suggest that policies that punish or exclude top prescribers to curb unnecessary opioid prescriptions may be severely attenuated by patient resorting.

To understand how patient sorting interacts with top physician-targeted policies, we develop an equilibrium model of patient choice of physicians and physicians' opioid prescription decisions. We model the patients' choice stage as a static, discrete choice demand model in the spirit of Berry, Levinsohn, and Pakes (1995).We model patient types as time-persistent tuples that summarize both opioid preferences and patient information. We allow opioid preference heterogeneity to arise from both medical needs and non-medical preferences. Patients' information types determine how accurately patients perceive physicians' prescription propensities. When choosing a physician, patients consider the expected prescription probability based on their own opioid preferences and their perceptions about each physician's prescribing propensity. In the prescription stage, partially altruistic physicians optimally prescribe opioids based on their own prescribing propensities and patients' opioid preferences.

Estimating our model is challenging because it features three dimensions of unobserved heterogeneity: physicians' prescription propensity, patients' opioid preferences, and patients' information types. To reduce the estimation burden, we estimate our model in three steps using a limited-information maximum-likelihood approach. First, we use our panel on realized opioid prescriptions to estimate both physicians' prescription propensities and patients' opioid preferences with a two-way fixed-effects logit model that includes time-varying controls. This estimation closely mirrors our earlier decomposition of prescription rate variation and follows the grouped fixed-effects approach in Bonhomme, Lamadon, and Manresa (2019) and Mourot (2025). This approach addresses both limited mobility and incidental parameter biases. Second, conditional on the first-stage estimates, we estimate the patient choice stage by constrained maximum likelihood, imposing a share constraint commonly relied on in discrete choice demand models (Berry, Levinsohn, and Pakes 1995). This step recovers mean utilities for each physician, parameters governing idiosyncratic preferences for office visits, and the distribution of time-invariant patient information types by leveraging the panel structure of the data. Finally, we decompose the estimates for patient opioid preference from the first stage into medically grounded preferences and non-medical preferences

[1]For example, https://www.reddit.com/r/hudsonvalley/comments/1ccgimp/pain_management_doctor_that_will_actually_manage/, and https://www.reddit.com/r/grassvalley/comments/1kn7h0i/doctor_that_prescribes_opiates_in_town_or_nc/

using a rich vector of patient-level controls, including patient severity measures and patients' diagnosis histories.

We use our estimates to simulate several counterfactuals. We first illustrate how patient sorting interacts with top physician-targeted policies, such as prescription drug monitoring programs. Prescription drug monitoring programs identify outlier prescribers and aim to induce changes in prescribing behavior through feedback, reporting, or enhanced monitoring (Amin-Esmaeili et al. 2023). These top physician-targeted policies implicitly posit that the variation in prescription rates must reflect differences in physicians' underlying propensity to prescribe. Our results on preference-based patient sorting demonstrate that the underlying assumptions of these policies are misguided, and hence their effectiveness will be severely attenuated. To illustrate this point, we consider an extreme version of such policies by excluding the top 10% of prescribers from the market. We find that even though our policy is extreme, aggregate opioid prescriptions drop by only 12%. This muted response arises because patients do not choose physicians randomly. Instead, our results imply that patients sort toward physicians whose prescribing propensity aligns with their own preferences for opioids. When top physician-targeted policies eliminate these first-best matches, patients optimally resort to the remaining physicians. Because high-prescribing physicians are disproportionately chosen by patients with strong preference for opioids, resorting disproportionately shifts demand toward the highest-prescribing remaining physicians, attenuating the aggregate reduction in opioid prescribing.

To further illustrate how patient resorting attenuates the effect of our exclusion policy, we simulate an alternative counterfactual that eliminates patient resorting among patients who originally chose an excluded physician. In particular, patients who visited an excluded physician in the baseline choose only whether to seek care. Conditional on seeking care, they are randomly assigned to a physician. Under this counterfactual, eliminating the top 10% of physicians by prescription rate reduces aggregate opioid prescriptions by 22%. These results demonstrate the constraints faced by supply-side-focused policies that are agnostic to demand-side responses: ignoring consumer choice attenuates the effectiveness of expert-targeted policies.

An important drawback of the randomization policy is that it prevents patients from choosing their physician. The medical literature shows that the ability to choose a physician is crucial for patients remaining in care (Hsu et al. 2003). Consistent with these findings, our results predict that when patients are randomized to physicians, most patients would rather forgo care than accept random assignment. Although randomization aggressively reduces aggregate opioid prescriptions, it achieves so by breaking optimal patient-physician matches regardless of whether sorting is based on patient medical needs ("good sorting")

or on medically unjustified preference for opioids ("bad sorting"). This makes randomization an undesirable policy. To disentangle good from bad sorting, we consider a policy that restricts opioid prescribing to medically grounded patient needs, excluding non-medical preferences. By eliminating incentives for patients to choose providers based on non-medical opioid preferences, this policy reduces aggregate prescribing rates by 21%, a result close to randomization, while preserving appropriate care for patients with medical needs for opioid treatment. In particular, we show that the number of patient visits under this alternative policy closely mirrors the ones when we allowed for resorting on both medical needs and non-medical opioid preferences.

Literature Review. This paper brings tools from the labor literature to the healthcare setting (Abowd, Kramarz, and Margolis 1999; Card, Heining, and Kline 2013; Card, Cardoso, Heining, and Kline 2018; Bonhomme, Lamadon, and Manresa 2019; Bonhomme, Holzheu, et al. 2023; Kline 2024). Our main contribution is to provide the first evidence of strong positive sorting between patients and physicians in the U.S. prescription opioid market. A related paper is Mourot (2025), which studies surgeon–hospital sorting and complementarities in the context of coronary artery bypass graft surgery. We go beyond by developing an equilibrium model of endogeneous patient sorting and physician prescribing decision, which we use to study counterfactual policy interventions. Our approach is similar to "mover design" in healthcare. The most closely related study is Finkelstein, Gentzkow, and D. Li (2025), which separately identifies individual-specific and region-specific drivers of risky opioid use. We differ by focusing on patient sorting toward providers and its interaction with supply-side-targeted opioid policies. Furthermore, we emphasize a distinct counterfactual policy that targets bad sorting based on non-medical preferences, reducing opioid prescribing while preserving appropriate care for patients with legitimate medical needs.

Our analysis contributes to a large and growing literature evaluating the role of physicians in the opioid crisis (Schnell and Currie 2018; Currie and Schwandt 2020; Schnell 2025; Currie, A. Li, and Schnell 2023). While recent work shows that exposure to high-prescribing physicians worsens health and labor market outcomes (Barnett, Olenski, and Jena 2017; Staiger, Baker, and Hernandez-Boussard 2022; Eichmeyer and Zhang 2022; Eichmeyer and Zhang 2023; Alpert, Schwab, and Ukert 2025), surprisingly little is known about why patients continue to seek care from these physicians.[2] By emphasizing patient sorting, we link the demand and supply sides of the opioid market and show that patient preferences play a central role in shaping the effectiveness of supply-side interventions. This perspective is

[2]The studies above rely on quasi-random assignment of patients to physicians—for example, through emergency room encounters or physician exit from Medicaid—to identify causal effects, abstracting from endogenous patient decisions.

particularly important given the widespread adoption of policies that restrict prescribers (Meara et al. 2016; Buchmueller and Carey 2018; Sacks, Hollingsworth, Nguyen, and Simon 2021; Alpert, Dykstra, and Jacobson 2024). By endogenizing patient sorting, we provide novel evidence that patients with stronger opioid preferences disproportionately sort toward high-prescribing physicians, which attenuates the effectiveness of policies that target top prescribers.

2 Prescription Opioids for Chronic Pain

In this paper, we focus on the use of prescription opioids for the treatment of chronic pain in outpatient office visits. We first present the setting and afterwards discuss the data we use in our research.

2.1 Setting

We begin by discussing chronic pain in the United States. According to the Centers for Disease Control and Prevention (CDC), chronic pain is defined as an episode of pain that lasts for at least 3 months (Dowell, Ragan, et al. 2022b). Chronic pain affects a significant fraction of the general population. In 2023, around 24% of the United States' adult population suffered from chronic pain (Lucas and Sohi 2024). Due to its nature, chronic pain often leaves affected individuals incapacitated and unable to function properly without treatment. Treatment for chronic pain in the United States commonly involves the use of prescription opioids. Around 22% of US adults suffering from chronic pain rely on prescription opioids for pain management (Dahlhamer et al. 2021). Chronic pain is not a condition of old age and affects adults of all age groups. However, the use of prescription opioids for pain management is most common among working-age adults. According to the National Center for Health Statistics, adults aged 45-64 years exhibit the highest rates of prescription opioid use (Dahlhamer et al. 2021). Because most of the affected patients are of working age, almost 2 in 3 payments for opioid prescription collections are covered by private insurance (Schnell 2025).

In the United States, patients must obtain prescription opioids by prescription from a licensed healthcare provider. About 4 in 5 prescription opioids in the United States are prescribed by physicians (Schnell 2025). In outpatient settings, patients can generally choose the physician they want to see, with limitations on their choice set depending on their respective insurance network. Once a patient has chosen a physician, the physician examines the patient and decides whether to prescribe opioids or not. If prescribing opioids, the

physician also has to decide on the number of days of supply. Contrary to acute pain, where initial supply in 33 states is limited to 5 to 7 days or less, no such regulation exists for opioids used for treating chronic pain (Department of Health and Human Services 2025). Large studies have found that the average number of days of supply is around 28 days. After the days of supply are used, patients can get a refill if their prescription allows it, or they have to see a physician to obtain a new prescription.

Prescription opioids are a common part of pain management for both acute and chronic pain in the United States. The use of prescription opioids for *acute pain* for short periods of time is widely accepted as medically justified (Dowell, Ragan, et al. 2022b). However, the use of prescription opioids for *chronic pain* management is posing substantial health risks while lacking clear evidence of any health benefits for patients (Volkow and McLellan 2016). Metastudies have found an association of adverse health outcomes, such as increased risk of opioid abuse and dependence, all-cause mortality, and myocardial infarction, with long-term opioid treatment for chronic pain (Chou et al. 2020). There is also growing evidence that long-term prescription opioids may make it harder to effectively manage pain in chronic pain patients. Long-term opioid use may result in opioid induced hyperalgesia in patients, a condition that increases pain sensitivity that may extend beyond areas initially affected by chronic pain for which the opioids were prescribed (Lee et al. 2011). As a result, the CDC Clinical Practice Guideline for Prescribing Opioids for Pain only mentions four conditions that unequivocally warrant prescribing opioids for chronic pain: cancer-related pain treatment, sickle cell disease, palliative care, and end-of-life care (Dowell, Ragan, et al. 2022b). Outside of these conditions, prescription opioids should not be used as a first line of treatment, but the decision on whether to use prescription opioids for pain management is generally left to the physician and patient (Dowell, Ragan, et al. 2022b). This ambiguity leaves the eventual decision to use prescription opioids as a course of treatment up to the physicians' discretion and patient preference.

2.2 MarketScan Data

As we discussed above, the majority of patients who receive prescription opioid treatment for chronic pain are privately insured working-age adults. For this reason, we need to assemble a data set that covers the relevant, privately insured working-age population comprehensively. Consequently, we cannot rely on the often-used data from the Centers for Medicare & Medicaid Services. Instead, we rely on data from the Merative MarketScan Research Databases. The MarketScan Commercial Database consists of individual-level, de-identified healthcare claims for millions of enrollees and their dependents in employer-sponsored healthcare plans

in the United States. To construct our sample, we combine data from the Commercial Database and the Merative Micromedex RED BOOK for the years 2016 to 2022. We link the Outpatient Services Table, the Inpatient Services Table, and the Outpatient Pharmaceutical Claims Table from the Commercial Database with National Drug Code (NDC)-level information on opioids from the Merative Micromedex RED BOOK.

Using our data sources, we build two data sets for the New York City Metropolitan Statistical Area. We restrict our sample to patients who had an episode of chronic pain and received at least one opioid prescription between 2016 and 2022. The first data set consists of claims data information on all office visits of covered patients linked with data on their prescription collections. We refer to this data set as the *visit data set*. The second data set supplements these data with information on time periods in which a given patient chose not to see any physician. We refer to this data set as the *choice data set*. We use the former data set to establish patient sorting for our motivating evidence and in our structural estimation, and the second data set solely for our structural estimation.

Step I: Constructing the visit data set.

We construct our data sample in multiple steps. First, we develop a strategy to identify whether a particular office visit for a patient resulted in the prescription of a prescription opioid or not. Following the literature, we define an opioid prescription as any prescription of substances that fall within both, the opioid analgesics category under the American Hospital Formulary Service Classification Compilation Therapeutic Class scheme, and the Drug Enforcement Administration (DEA) drug schedule fall into schedule II or III, indicating high or moderate potential for abuse, respectively (Alpert, Dykstra, and Jacobson 2024; Staiger, Baker, and Hernandez-Boussard 2022).[3] As is common with commercial claims data, we lack national provider identification (NPI) information for the prescribing physician for prescription opioid collections from the Outpatient Pharmaceutical Claims Table. Instead, we use a well-established method to link prescription drug collections with prescribing providers. In particular, we follow Ding and Liu (2021) and assign a prescription to an observed office visit if the prescription opioid collection was within 14 days following the office visit , was the first collection within that time frame among all prescription opioid collections for that patient, and could not be associated with any prior office visit of that patient. Our final variable is a binary indicator for whether a prescription opioid was prescribed during a particular office visit. Note that we are focusing on the extensive rather than the intensive margin.

The MarketScan Commercial Database includes rich demographic and diagnosis information on possible patient risk factors. In particular, for each patient, we observe age, sex,

[3]For the American Hospital Formulary Service Classificationn Compilation (AHFSCC) Therapeutic Class, the opioid analgesics category is "28:08.08".

and up to 4 primary diagnoses per outpatient visit. These diagnoses adopt the 10th revision of the International Classification of Diseases, ICD-10 codes. We leverage a patient's medical history within 12 months preceding an office visit to create measures of a patient's health status, as well as possible controls for opioid preference. The first health measure we construct to capture a wide range of comorbidities is the Charlson Comorbidity Index (Charlson, Pompei, Ales, and MacKenzie 1987).[4] Additionally, we follow the CDC Clinical Practice Guideline for Prescribing Opioids for Pain (Dowell, Ragan, et al. 2022b) to identify whether a patient has a medical condition unequivocally qualifying for prescription opioid treatment. These conditions are cancer-related pain, sickle cell disease, palliative care, and end-of-life care. Next, we construct possible controls for opioid preference by identifying diagnoses for which the CDC guidelines recommend against prescribing or urge caution when prescribing opioids. These conditions are sleep-disordered breathing, renal or hepatic insufficiency, mental health conditions (i.e., anxiety, depression, and post-traumatic stress disorder), substance use disorders, opioid abuse, and non-fatal overdose (Dowell, Ragan, et al. 2022b). We also construct counts of emergency room visits related to non-fatal overdoses, and count inpatient and outpatient visits during the 12 months preceding each office visit. We provide detailed descriptions, including the ICD-10 diagnosis codes used to construct our health status and opioid preference measures, in appendices D and E.

Lastly, we construct measures of out-of-pocket costs for prescription opioids and total outpatient and inpatient spending during the 12 months preceding an office visit, using the financial variables from the MarketScan Commercial Database. Since we do not observe out-of-pocket costs when no opioid was prescribed, we construct out-of-pocket prescription opioid costs per outpatient visit as the mean out-of-pocket cost across insurance-type-month combinations. We provide more details on the construction of this variable in Appendix G. We summarize our first data set in table 1.

Step II: Constructing the choice data set.

For our choice data set we exclusively use to estimate our structural model, we supplement the above data with additional information on months when a patient could have seen a physician but chose not to see one.[5] For each month, we define the population of active patients (the set of potential patients) as the set of patients who had at least one office visit

[4] In particular, we use the 17 categories considered in the popular Charlson/Deyo variant (Deyo 1992). We adopt the more recent weights suggested in Schneeweiss, Wang, Avorn, and Glynn (2003) and provide additional details in the appendix.

[5] We use a month in accordance with our evidence on average days of supply.

Table 1: Summary Statistics of Chronic Pain Patients

Variable	Mean	SD	25 Perc	50 Perc	75 Perc
Panel A: Patient Severity and Constraints					
Age (years)	48.20	11.61	40.00	50.49	58.00
Sex (Male=1)	0.47	0.50	0.00	0.00	1.00
Charlson Commorbidity Index	1.23	1.59	0.00	1.00	2.00
Counts of Prior Inpatient Visits	0.13	0.40	0.00	0.00	0.00
Counts of Prior Outpatient Visits	65.26	70.60	23.50	45.50	83.40
Counts of Prior ED Visits	0.48	1.31	0.00	0.00	0.50
Prior OOP Spending	1696.09	2019.36	394.57	1065.30	2315.74
Out-of-pocket Cost	1.04	0.37	0.79	0.93	1.25
Panel B: Conditions Justifying Opioid Prescription					
Sickle Cell Disease	0.00	0.07	0.00	0.00	0.00
Cancer-Related Pain	0.05	0.23	0.00	0.00	0.00
Panel C: Conditions Discouraging Opioid Prescription					
Renal/Hepatic Deficiency	0.02	0.15	0.00	0.00	0.00
Mental Health Conditions	0.33	0.47	0.00	0.00	1.00
Opioid Misuse	0.04	0.21	0.00	0.00	0.00
Substance Use Disorder	0.12	0.33	0.00	0.00	0.00
Prior Nonfatal Overdose	0.03	0.18	0.00	0.00	0.00
Counts of ED Visits Related to Overdose	0.00	0.05	0.00	0.00	0.00
Work in Construction or Transportation	0.08	0.27	0.00	0.00	0.00
Observations					
Number of unique patients			11,909		
Total observations			59,227		

in the corresponding quarter of that year. If we observe a patient visit a physician in a given month, we record their choice.[6] If we do not observe a patient visiting a physician in a given month for which this patient is part of the active population, we record that the patient chose the outside option of not seeing a physician. If a patient chooses the outside option for a given month, we assume that their characteristics, which we computed for the first data set, are the same as the closest month in that quarter.

For our second data set, we also compute additional variables for physicians. Based on the choices of physicians or the outside option by patients, we compute the market shares for physicians for each month. The market share is defined as the number of patients we observe choosing a given physician in a given month, divided by the total number of patients active in that particular month. A common problem in defining market shares based on observable choices is that some options (physicians) may not be chosen in each month. If this happens for some of our physicians, we follow the standard approach in the discrete choice literature and assume that this physician was not active in a given month. We interpret this as the possibility that physicians might close their offices periodically for vacation. We summarize our data set for structural estimation in table 2.

Table 2: Physician Market Share Summary Statistics

Variable	Mean	SD	25 Perc	50 Perc	75 Perc
Physician Market Share (%)	0.09	0.09	0.04	0.06	0.09
Inside Share in a Market (%)	46.97	2.82	45.40	47.06	48.81
Observations					
Number of markets			84		
Number of physicians			8059		

3 Establishing Patient Sorting

In this section, we first document variation in opioid prescription rates for chronic pain patients across physicians. We then decompose the sources of this variation. A physician's prescribing rate reflects not only their underlying prescribing propensity, but also differences in patient severity, patient preferences, and patient sorting into physicians based on both medical needs and non-medical preferences. We are particularly interested in assessing the

[6] 13% of patient-months involve more than one visit. For these cases, we collapse multiple visits into a single patient-month observation if they are all with the same physician (which occurs in about two thirds of these cases), and we record an opioid prescription if any of the visits resulted in one. We exclude patient-months in which a patient saw more than one physician during that month.

relative importance of patient sorting versus physician prescribing propensity in explaining cross-physician differences in prescription rates. Finally, we extend the analysis beyond prescribing to opioid misuse, examining the extent to which patient sorting may contribute to risky prescription opioid use.

Cross-Physician Variation in Prescription Rates. Our primary outcome is an indicator equal to one if an office visit results in an opioid prescription. We estimate physician-level raw prescription rates by regressing this indicator on physician fixed effects:

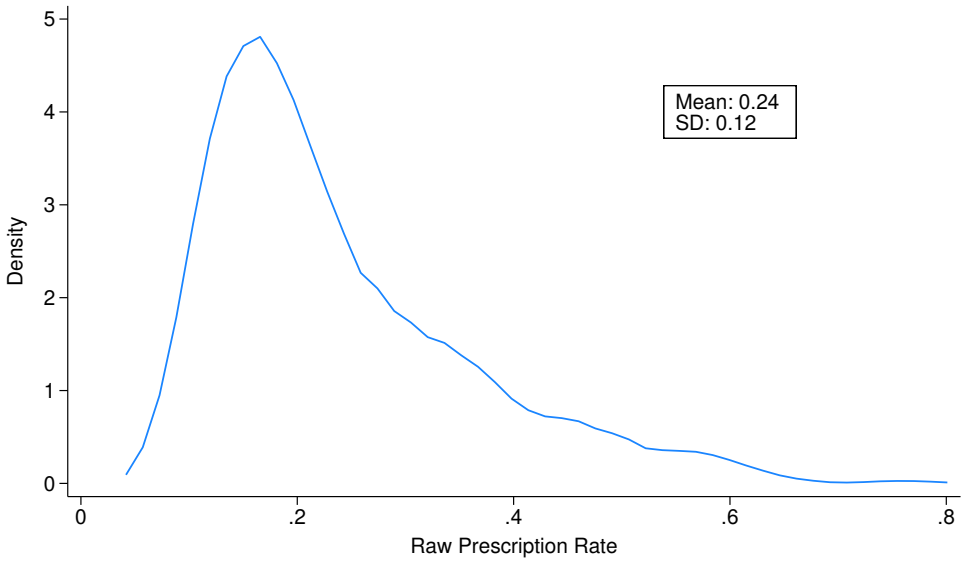
$$\text{prescribe}_{it} = a_{j(i,t)} + \varepsilon_{it}, \tag{1}$$

where prescribe_{it} equals one if patient i receives an opioid prescription at visit t, $a_{j(i,t)}$ denotes the fixed effect for physician j, capturing physician j's average prescription rate in the data, and ε_{it} is the error term. Because $a_{j(i,t)}$ may be noisily estimated for physicians with relatively few patient visits, we apply an empirical Bayes shrinkage estimator that shrinks imprecisely estimated fixed effectstowards the overall mean, following Chandra, Finkelstein, Sacarny, and Syverson (2016).

Figure 1 plots the density of the empirical Bayes–adjusted physician fixed effects, $\hat{a}_{j(i,t)}$. The distribution reveals substantial heterogeneity in raw prescription rates across physicians: the standard deviation is one-half of the mean prescription rate. This finding is consistent with prior evidence documenting large differences in drug prescribing intensity across providers (Finkelstein, Gentzkow, and Williams 2016; Molitor 2018; Cutler, Skinner, Stern, and Wennberg 2019; Badinski, Finkelstein, Gentzkow, and Hull 2023; Clemens, Léger, Nandan, and Town 2024). The distribution also features a thick upper tail, indicating that a nontrivial share of physicians prescribe opioids at rates far above the average.

Our results demonstrate that opioid prescription rates across physicians are highly dispersed. We next investigate the determinants of the observed dispersion in opioid prescription rates. A significant source of opioid prescription rate dispersion may be differences in the patient severity composition across physicians. To assess whether opioid prescription rates mainly reflect differences in patient severity, we risk-adjust physicians' prescribing decisions by controlling for a rich set of patient observables. Our data features three categories of patient observables we leverage for risk-adjustment: (1) patient health status and patient-specific constraints, such as patient age, sex, the Charlson comorbidity index, counts of historical inpatient, outpatient, and emergency-department visits, historical out-of-pocket spending, and patient cost sharing; (2) patient medical conditions for which opioid use may be clinically warranted under the CDC's opioid prescribing guidelines (Dowell, Ragan, et al. 2022a). These include sickle-cell disease and cancer-related pain; and (3) diagnoses and

Figure 1: Raw Physician Prescription Rate



Notes: This figure plots the distribution of the empirical Bayes–adjusted physician fixed effects from equation (1). We restrict the sample to patients with at least 10 visits and physicians with at least 10 patient visits over the sample period.

risk indicators for which the CDC recommends caution or avoidance in the management of chronic pain. This group includes diagnoses of renal or hepatic insufficiency, mental health conditions, opioid misuse or substance-use disorder, a prior nonfatal overdose, counts of emergency-department visits related to overdoses, and an indicator for employment in construction or transportation industries.[7]

We construct risk-adjusted prescription rates for each physician by regressing our prescription indicator on physician fixed effects, $a_{j(i,t)}$, the full set of patient observables described above, denoted by X_{it}, as well as year fixed effects, τ_t, to absorb any potential common time-varying factors such as policy changes and guideline updates using the following regression equation:

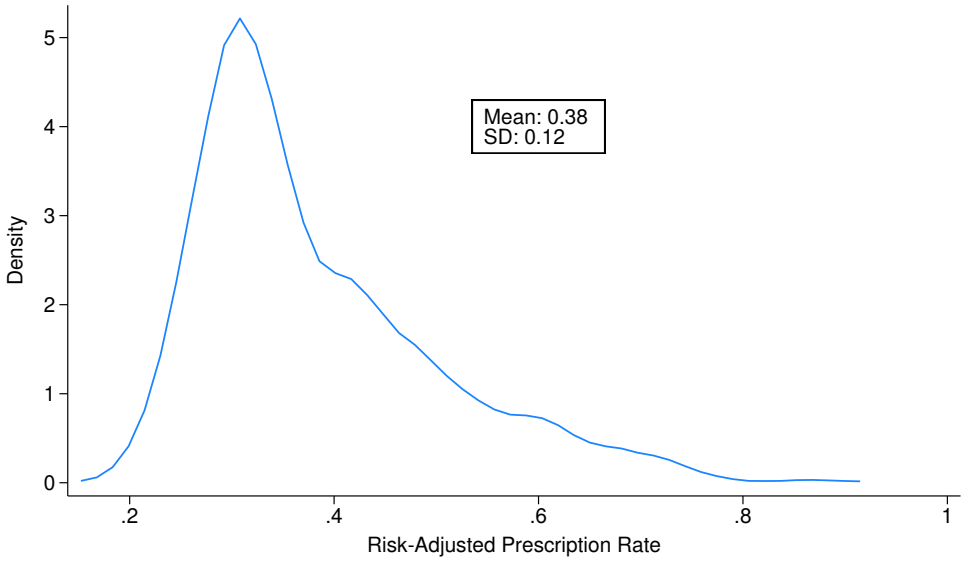
$$\text{prescribe}_{it} = a_{j(i,t)} + X_{it}b + \tau_t + \epsilon_{it}. \tag{2}$$

Here, $a_{j(i,t)}$ captures the physician-specific risk-adjusted mean prescription rate net of observable patient characteristics and common time effects. After applying the empirical Bayes shrinkage procedure to correct the estimated $\hat{a}_{j(i,t)}$ for potential measurement error, we plot

[7]Mental health conditions include anxiety, posttraumatic stress disorder, and depression (Dowell, Ragan, et al. 2022b). Occupations warranting caution include driving, using heavy equipment, climbing ladders, working at heights or around moving machinery, or working with high-voltage equipment (Dowell, Ragan, et al. 2022b).

the density for the empirical Bayes–adjusted $\hat{a}_{j(i,t)}$ in Figure 2.

Figure 2: Risk-adjusted Prescription Rate



Notes: This figure plots the distribution of the empirical Bayes–adjusted, risk-adjusted physician fixed effects from equation (2). We restrict the sample to patients with at least 10 visits and physicians with at least 10 patient visits over the sample period.

Figure 2 shows that relative to Figure 1, controlling for patient observables and year fixed effects leaves the dispersion in prescription rates largely intact. In other words, the prescription rate heterogeneity across physicians is not primarily driven by patient observables or common time trends. Specifically, the standard deviation of the distribution of risk-adjusted prescription rates equals the standard deviation of the raw prescription rates.

At face value, the patterns we documented appear to suggest a natural empirical rationale justifying policies that target high-prescribing physicians. Any remaining variation in risk-adjusted prescription rates reflects inherent differences in physicians' prescribing propensities. However, such arguments and recommendations may be severely misguided. Observed prescription rates may reflect unobserved patient preferences and positive sorting of patients with strong preference for prescription opioids into high-prescribing physicians, even after risk-adjustment for patient severity. In fact, positive sorting may amplify even modest differences in physicians' underlying prescribing propensities and lead to vast dispersion in observed prescribing rates. Conversely, if patient sorting were negative, the true dispersion in physician prescribing propensities may be larger than the one implied by observed risk-adjusted prescription rates.

Patient Sorting. To understand the importance of patient sorting versus physicians' prescribing propensities for the observed dispersion in prescription rates, we exploit the panel structure of our data. In particular, we decompose the variation in raw prescription decisions into components attributable to patient type, physician type, and sorting of patients into physicians whose prescribing styles they prefer. We further distinguish sorting based on medical needs versus non-medical preferences.

Our variance decomposition approach mirrors the well-known two-way fixed effects estimator of Abowd, Kramarz, and Margolis (1999) from the labor literature, but applied to prescription decisions. Following their method, we specify a two-way fixed effects regression in which we regress physicians' prescription decisions on patient fixed effects, physician fixed effects, patient observables, and year dummies.

$$\text{prescribe}_{it} = \eta_{j(i,t)} + \theta_i + X_i\beta + \tau_t + \varepsilon_{it}, \tag{3}$$

Here, $\eta_{j(i,t)}$ represents the physician prescribing propensity of the chosen physician, θ_i represents patient unobserved type, and X_i is a vector of patient observables defined as above, τ_t are year dummies, and ε_{it} is an idiosyncratic error term representing any unobserved time-varying patient characteristics.[8]

Note that in equation (3), β and θ_i are not separately identified because θ_i and X_i are both patient-specific. Therefore, we define $\Theta_i \equiv \theta_i + \beta X_i$ and treat Θ_i as patient fixed effect for estimation in the spirit of Abowd, Kramarz, and Margolis (1999). We therefore rewrite equation (3) as

$$\begin{aligned}\text{prescribe}_{it} &= \eta_{j(i,t)} + \underbrace{\theta_i + X_i\beta}_{\equiv\Theta_i} + \tau_t + \varepsilon_{it} \\ &= \eta_{j(i,t)} + \Theta_i + \tau_t + \varepsilon_{it}.\end{aligned} \tag{4}$$

Obtaining unbiased fixed effects in an Abowd, Kramarz, and Margolis (1999) style model is well known to be challenging because physician fixed effects are identified only from patients who switch physicians. When switching is limited, fixed effects estimates can be severely biased, a problem known as limited mobility bias (Andrews, Gill, Schank, and Upward 2012; Bonhomme, Holzheu, et al. 2023). Recent advances in the labor literature show that this bias can be mitigated using a two-step estimator that first groups firms and then

[8]Because most of patient observables exhibit little within-patient variation over time, we use patient-level averages and treat them as time-invariant characteristics. Including time-varying versions of these variables when available yields highly similar results in terms of the relative importance of each component in explaining variation in prescription rates.

estimates group-level fixed effects (Bonhomme, Lamadon, and Manresa 2019). We adopt the approach of Bonhomme, Lamadon, and Manresa (2019) and follow Mourot (2025) to adapt this approach to our setting of opioid prescribing. Following Mourot (2025), in the first stage, we group physicians and patients using k-means clustering based on physician-level and patient-level risk-adjusted prescription rates, respectively.[9] Appendix A provides more details on the grouping algorithm. Following standard practice in the labor literature (Kline 2024), we use ten groups for physicians and ten groups for patients. We assess robustness by varying the number of groups. In the second stage, we estimate equation (4) using group-level physician and patient fixed effects.

After obtaining estimates for the group-level fixed effects, $\hat{\eta}_{j(i,t)}$ and $\hat{\Theta}_i$, we further decompose $\hat{\Theta}_i$ into preferences we can attribute to patient medical needs and non-medical preferences. To do so, we regress patient fixed effects estimates, $\hat{\Theta}_i$, on our full vector of non-time varying patient observables X_i:

$$\hat{\Theta}_i = \underbrace{X_{1i}\beta_1}_{\text{medical needs}} + \underbrace{X_{2i}\beta_2 + \theta_i}_{\text{non-medical preferences}} . \tag{5}$$

We partition X_i into two sets of time-invariant patient characteristics: medical needs, X_{1i}, and non-medical preferences, X_{2i}. X_{1i} includes variables under (1) health severity and patient-specific constraints, as defined before, and (2) diagnoses that warrant opioid use, including sickle-cell disease and cancer-related pain. While not all of these variables reflect clinical appropriateness in a narrow sense, they capture patient characteristics commonly considered in pain management decisions emphasized in CDC guidance (Dowell, Ragan, et al. 2022a). We collect the remaining non-medically justified factors in X_{2i} and θ_i. X_{2i} includes (3) characteristics for which the CDC explicitly recommends caution or avoidance of opioid use for managing chronic pain. θ_i is the residualized patient fixed effects obtained after partialing out X_i from $\hat{\Theta}_i$.

We now decompose the variance of the prescribing decision. Following the recommendation of Kline (2024), we scale each component of the decomposition by $var(\text{prescribe}_{it}^{E})$, the variation in prescribing that can be explained by our regressors. According to Kline (2024), this normalization ensures the comparability of our results with any future related studies. In particular, we have that

[9] Mourot (2025) groups surgeons and hospitals, whereas we group patients and physicians.

$$
\begin{aligned}
var(\text{prescribe}_{it}^{E}) &\equiv var(\eta_{j(i,t)} + X_{1i}\beta_1 + X_{2i}\beta_2 + \theta_i) \\
&= var(\eta_{j(i,t)}) + var(X_{1i}\beta_1) + var(X_{2i}\beta_2 + \theta_i) \\
&+ 2cov(X_{1i}\beta_1, \eta_{j(i,t)}) + 2cov(X_{2i}\beta_2 + \theta_i, \eta_{j(i,t)}) + 2cov(X_{1i}\beta_1, X_{2i}\beta_2 + \theta_i)
\end{aligned} \tag{6}
$$

where $var(X_{1i}\beta_1)$ captures variation in prescribing explained by patient medical needs relevant for opioid prescriptions, while $var(X_{2i}\beta_2 + \theta_i)$ reflects variation associated with patient non-medical preferences. The term $var(\eta_{j(i,t)})$ represents variation attributable to physician prescribing propensity. The covariance term $cov(X_{1i}\beta_1, \eta_{j(i,t)})$ captures patient sorting based on medical needs, whereas $cov(X_{2i}\beta_2 + \theta_i, \eta_{j(i,t)})$ captures sorting based on non-medical preferences. Finally, $cov(X_{1i}\beta_1, X_{2i}\beta_2 + \theta_i)$ reflects the covariance between patients' medical needs and non-medical preferences. We report the results of our variance decomposition in table 3.

For our main results in Table 3, the baseline specification, we use 10 groups for both patients and physicians, and restrict the sample to patients with at least 2 visits and physicians with at least 2 patient visits over the sample period. We report results from alternative sample restrictions and alternative number of groups to assess the robustness of our results. Additionally, we report the results of employing an empirical Bayes–adjusted estimator to estimate risk-adjusted prescription rates we use for grouping of patients and phyisicans.

Our results in Table 3 demonstrate that across specifications, and irrespective of whether we apply Bayesian shrinkage, there exists strong positive assortative matching between patients and physicians.[10] We find that assortative matching is overwhelmingly due to non-medical preferences for opioid prescriptions rather than medical needs. Patient sorting based on medical needs explains only about 5 percent of the variation in prescription rates, whereas sorting based on non-medical preferences accounts for more than 30 percent. Perhaps surprisingly, physicians' intrinsic prescribing propensity explains only around 10 percent of the variation in prescribing decisions. Our results imply that patient sorting accounts for roughly three times as much of the observed heterogeneity in prescribing rates as providers' own prescribing propensity. Consistent with this finding, the correlation coefficient between patients' non-medical preferences and their chosen physicians' prescribing propensity is as high as 0.66.

[10]Due to space constraints, we do not report the term $2cov(X_{1i}\beta_1, X_{2i}\beta_2 + \theta_i)$, which is close to zero. As a result, summing the first five rows for a given column yields nearly 100%.

Table 3: Percentage of Variance Explained in Raw Prescription Rates (%)

	Baseline	**Alternative Number of Groups**				**Alternative Sample Restrictions,** $K = 10$			**Bayesian Shrinkage**
	$K = 10$ ≥ 2 visits/patient ≥ 2 visits/provider	$K = 5$	$K = 15$	$K = 20$	$K = 50$	≥ 1 visits/patient ≥ 1 visits/provider	≥ 5 visits/patient ≥ 5 visits/provider	≥ 10 visits/patient ≥ 10 visits/provider	$K = 10$ ≥ 10 visits/patient ≥ 10 visits/provider
Patient									
$\frac{\text{Var}(X_{2i}\beta_2 + \theta_i)}{\text{Var}(\text{prescribe}^E)}$	49.58	49.08	50.70	50.79	51.68	49.15	50.89	51.16	52.70
$\frac{\text{Var}(X_{1i}\beta_1)}{\text{Var}(\text{prescribe}^E)}$	4.70	4.62	4.74	4.81	4.65	2.82	3.66	4.42	5.04
Provider									
$\frac{\text{Var}(\eta_{j(i,t)})}{\text{Var}(\text{prescribe}^E)}$	12.12	12.65	11.29	11.28	11.04	17.16	14.49	12.26	10.11
Sorting									
$2 \times \frac{\text{Cov}(X_{2i}\beta_2 + \theta_i, \eta_{j(i,t)})}{\text{Var}(\text{prescribe}^E)}$	31.70	31.49	31.38	31.33	30.99	30.13	30.10	30.73	30.22
$2 \times \frac{\text{Cov}(X_{1i}\beta_1, \eta_{j(i,t)})}{\text{Var}(\text{prescribe}^E)}$	1.90	2.16	1.89	1.79	1.63	0.74	0.85	1.42	1.93
Correlation									
$\text{corr}(X_{2i}\beta_2 + \theta_i, \eta_{j(i,t)})$	0.66	0.65	0.66	0.66	0.65	0.52	0.55	0.62	0.67
$\text{corr}(X_{1i}\beta_1, \eta_{j(i,t)})$	0.13	0.14	0.13	0.12	0.11	0.05	0.06	0.10	0.14

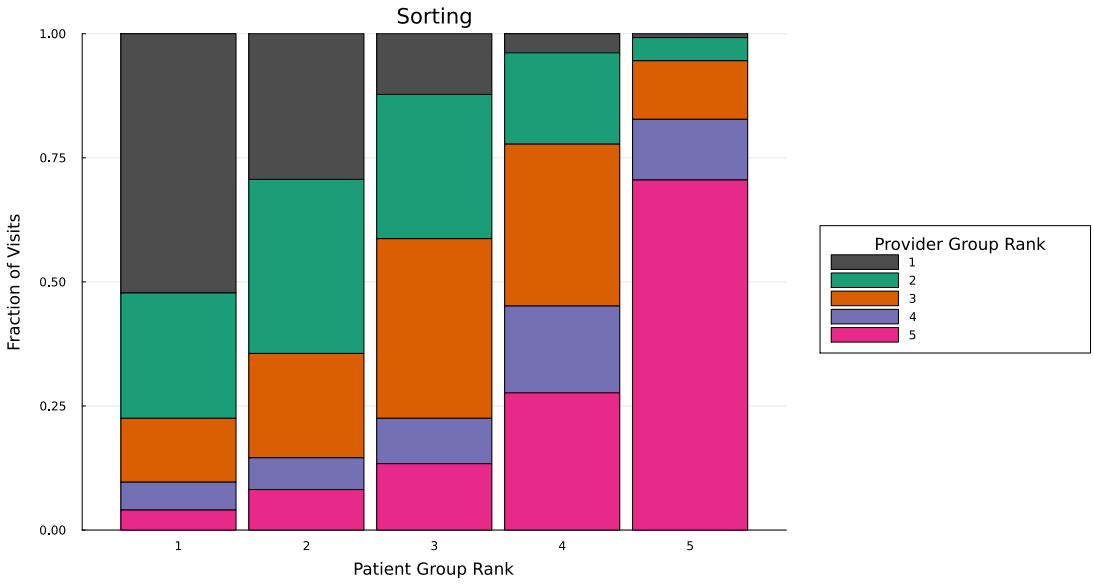


Figure 3: Patient Sorting

We further illustrate strong positive assortative matching in Figure 3. Figure 3 shows that patients with low opioid preference (group 1) make more than half of their visits to providers who prescribe the least opioids (group 1). Conversely, patients with high opioid preference (group 5) disproportionally choose providers with the highest prescription rates (group 5). Appendix Figure 13 shows the same patterns when patients and providers are divided into 10 groups instead of 5.

Positive patient sorting implies that the observed differences in prescription rates across physicians are larger than the underlying differences in physicians' true prescribing propensities. In other words, positive sorting amplifies variation in prescribing behavior, leading to a steeper gradient in the raw prescription rate. Figure 4 illustrates this mechanism: the horizontal axis represents ten physician groups ranked by their raw prescription rates. Each blue triangle shows the raw prescription rate for a physician group, while each red circle indicates the estimated physician fixed effect, i.e., the group's prescribing propensity, scaled by the base prescription rate of group 5. The fitted line for the raw prescription rate is visibly steeper, reflecting positive patient sorting. At the upper end of the distribution, physicians appear to prescribe much more liberally than they truly are: raw prescription rates overstate their true prescribing propensities. Similarly, physicians in the lower percentiles appear to prescribe far more conservatively, even though their true propensities are not that different from those of higher prescribers.

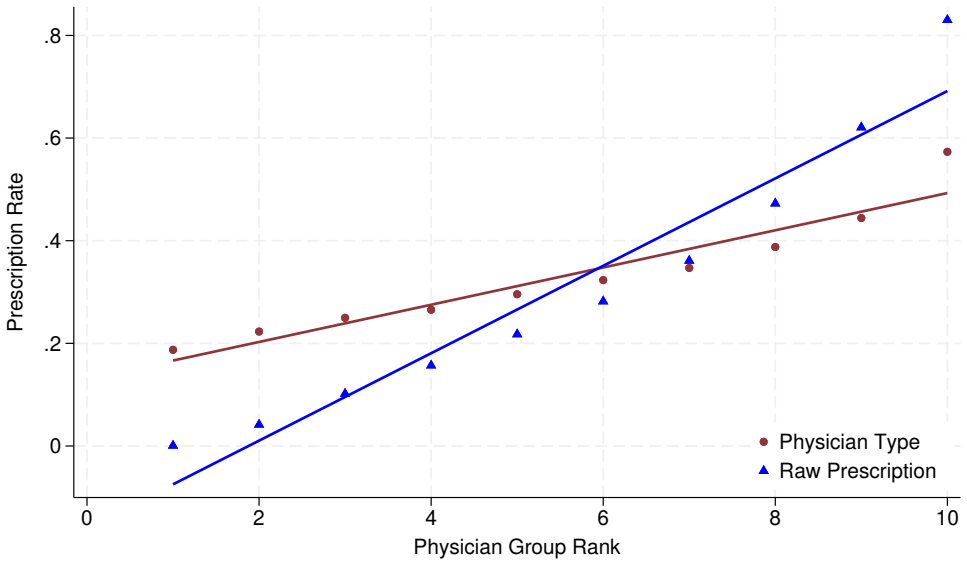


Figure 4: Patient Sorting Versus Physician Propensity Type

Patient Sorting Contributes to Risky Opioid Use. We have established that variation in observed cross-physician prescription rates is driven less by differences in physicians' prescribing propensities and more by patients sorting into physicians according to those propensities. We now examine whether positive patient sorting also plays an important role in explaining adverse health outcomes. To do so, we revisit the decomposition exercise in equation (4), replacing the prescription outcome variable with an indicator for opioid misuse at that visit. Literature has documented correlation between high prescription levels, defined as the patient's daily MED being 120 mg or greater, and opioid abuse (Meara et al. 2016; Staiger, Baker, and Hernandez-Boussard 2022; Finkelstein, Gentzkow, and D. Li 2025). Following this literature, we classify a visit as involving opioid misuse if the patient's daily MED is 120 mg or greater. We provide details on the construction of our measure of daily MED in Appendix F. We focus on this health outcome because it is well defined at the visit level. Additionally, it provides a more objective measure of opioid misuse than opioid-abuse–related diagnoses coded by the physician, which depend on both the physician's or patient's willingness to test and assessment. Moreover, opioid misuse is both a strong risk factor for, and a precursor to, opioid addiction. Based on this definition, 6% of visits in our sample involve risky opioid use.

In Table 4, we decompose the explained variation in opioid misuse into components attributable to patient medical needs ($X_{1i}\beta_1$), patient non-medical preferences ($X_{2i}\beta_2 + \theta_i$), physician prescribing propensity ($\eta_{j(i,t)}$), and patient sorting on both medical needs and non-medical preferences ($Cov(X_{1i}\beta_1, \eta_{j(i,t)})$ and $Cov(X_{2i}\beta_2 + \theta_i, \eta_{j(i,t)})$). All specifications

yield qualitatively similar results. Moreover, the patterns closely mirror those from the decomposition of raw opioid prescription rates. Patient sorting explains nearly one-third of the variation in opioid misuse, far outweighing the contribution of physician prescribing propensity, which accounts for less than 10% in the baseline specification. The correlation coefficient between patient non-medical preference type and physician propensity is as high as 0.64. Furthermore, the contribution of patient medical needs is much smaller in this context, suggesting that non-medical preferences play a central role in driving opioid misuse. These results indicate that patient sorting not only drives high opioid prescription rates but is also associated with worse health outcomes, thereby amplifying the opioid epidemic.

Table 4: Percentage of Variance Explained in Opioid Misuse (%)

	Baseline	**Alternative Number of Groups**				**Alternative Sample Restrictions,** $K = 10$			**Bayesian Shrinkage**
	$K = 10$ ≥ 2 visits/patient ≥ 2 visits/provider	$K = 5$	$K = 15$	$K = 20$	$K = 50$	≥ 1 visits/patient ≥ 1 visits/provider	≥ 2 visits/patient ≥ 2 visits/provider	≥ 5 visits/patient ≥ 5 visits/provider	$K = 10$ Bayes shrinkage
Patient									
$\frac{\text{Var}(X_{2i}\beta_2 + \theta_i)}{\text{Var}(\text{opioid misuse}^E)}$	61.30	52.80	60.66	60.76	60.61	57.24	62.64	65.47	62.27
$\frac{\text{Var}(X_{1i}\beta_1)}{\text{Var}(\text{opioid misuse}^E)}$	0.70	0.90	0.72	0.76	0.87	0.53	0.67	0.72	0.92
Provider									
$\frac{\text{Var}(\eta_{j(i,t)})}{\text{Var}(\text{opioid misuse}^E)}$	8.40	13.34	8.54	8.59	8.85	13.14	9.93	7.48	8.05
Sorting									
$2 \times \frac{\text{Cov}(X_{2i}\beta_2 + \theta_i, \eta_{j(i,t)})}{\text{Var}(\text{opioid misuse}^E)}$	29.21	32.51	29.73	29.53	29.36	28.89	26.53	25.97	28.36
$2 \times \frac{\text{Cov}(X_{1i}\beta_1, \eta_{j(i,t)})}{\text{Var}(\text{opioid misuse}^E)}$	0.39	0.45	0.36	0.37	0.30	0.21	0.23	0.36	0.39
Correlation									
$\text{corr}(X_{2i}\beta_2 + \theta_i, \eta_{j(i,t)})$	0.65	0.62	0.66	0.65	0.64	0.53	0.53	0.59	0.64
$\text{corr}(X_{1i}\beta_1, \eta_{j(i,t)})$	0.08	0.07	0.07	0.07	0.05	0.04	0.05	0.08	0.07

Threats to Identification. Including patient fixed effects, which absorb all time-invariant unobserved patient characteristics, addresses many sources of possible confounders. However, fixed effects estimates may still be biased if patients' switching decisions are correlated with unobserved time-varying characteristics ε_{it}. Following the labor literature, we provide evidence that switching is not driven by omitted time-varying unobservables in our setting. In particular, Card, Cardoso, Heining, and Kline (2018) show that if worker switching is exogenous, workers exhibit symmetric wage responses with opposite signs when switching to higher- versus lower- type firms. Analogously, if patient switching is exogenous in our context, patients who switch to more aggressive prescribers should experience prescription probability changes that are symmetric, with opposite signs, to those who switch to less aggressive prescribers.

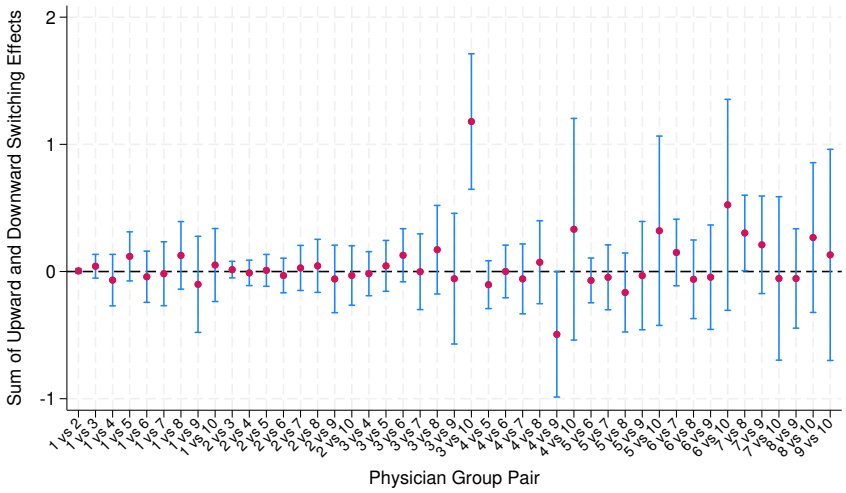
We test this implication in Figure 5 (with implementation details provided in Appendix). Figure 5 presents pairwise analyses of movers between physician groups. For each pair of physician groups, we estimate the effect of switching across groups relative to remaining within the same group. For example, for the physician group pair 1 and 2, some patients move from physician group 1 to group 2 ("move up"), others move from group 2 to group 1 ("move down"), while patients who switch physicians but remain within the same group serve as a control group. We implement a difference-in-differences–style design to estimate the effect of crossing groups relative to staying within a group.

Across all 45 group pairs, with one exception, we find that the change in prescription probability has opposite signs for move-up and move-down patients, and that the magnitudes are highly similar. As shown in Figure 5, each dot reports the sum of the estimated move-up and move-down effects, which is generally close to zero and statistically insignificant. This symmetric response pattern holds across all pairs of physician groups.

We further conduct an event-study analysis, pooling all movers into move-up and move-down categories, to examine whether switching is preceded by differential pre-trends. As shown in Figure 6, we find no evidence of pre-trends in prescribing behavior prior to the switch. This absence of pre-trends further suggests that switching is not driven by patients' gradual learning or other time-varying unobserved factors.

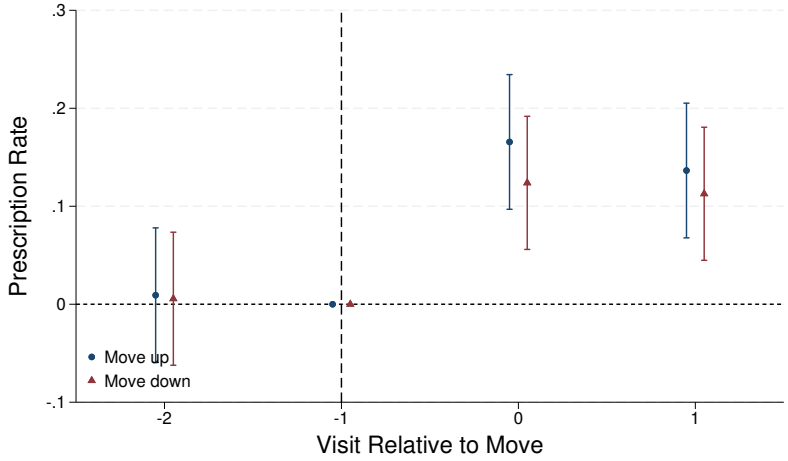
Discussion. Our finding that patient sorting dominates providers' prescribing propensity in explaining variation in both prescription rates and opioid misuse has important policy implications. First, policies that target only high prescribers are unlikely to be effective in curbing prescription rate or opioid misuse, as patients who initially sought care from these top prescribers are likely to re-sort toward other providers whose prescribing propensity is just below the targeted prescribers. Second, exploiting patient sorting might be a promising

Figure 5: Net Prescription Rate Response to Upward and Downward Patient Switching



Notes: This figure plots the sum of the changes in prescription rates for patients who switch upward and downward within each pair of physician groups. Physician groups are defined based on physician-level risk-adjusted prescription rates and are ordered from 1 (lowest-prescribing group) to 10 (highest-prescribing group).

Figure 6: Magnitude of Prescription Rate Responses to Patient Switching



Notes: This figure plots the magnitude of changes in prescription rates for patients who switch to higher-prescribing physicians ("move up") and to lower-prescribing physicians ("move down"). Time 0 corresponds to the visit at which the physician switch occurs.

avenue to address opioid over-prescription, especially if policy can discourage patient sorting based on non-medical preferences while preserving patient sorting based on medical needs.

To examine how patient sorting interacts with different physician-targeted policies and toevaluate alternative policies, we propose and estimate a structural model that explicitly accounts for patient sorting in the next section.

4 Model

Our results in section 3 suggest that patients choose physicians according to physicians' prescribing propensities. In this section, we develop a static two-stage model of physician-patient interaction that is able to capture the uncovered patient sorting. In the first stage, the *choice stage*, in each month t, patients choose either to see a physician $j \in \mathcal{J}_t$ in their market (month) or to select the outside option of not seeking care. Patients choice is based on their beliefs about each physician's prescribing behavior. In the second stage, the *prescription stage*, the chosen physician decides whether to prescribe or not. We illustrate our two-stage model for a single month t and a single patient i in figure 7.

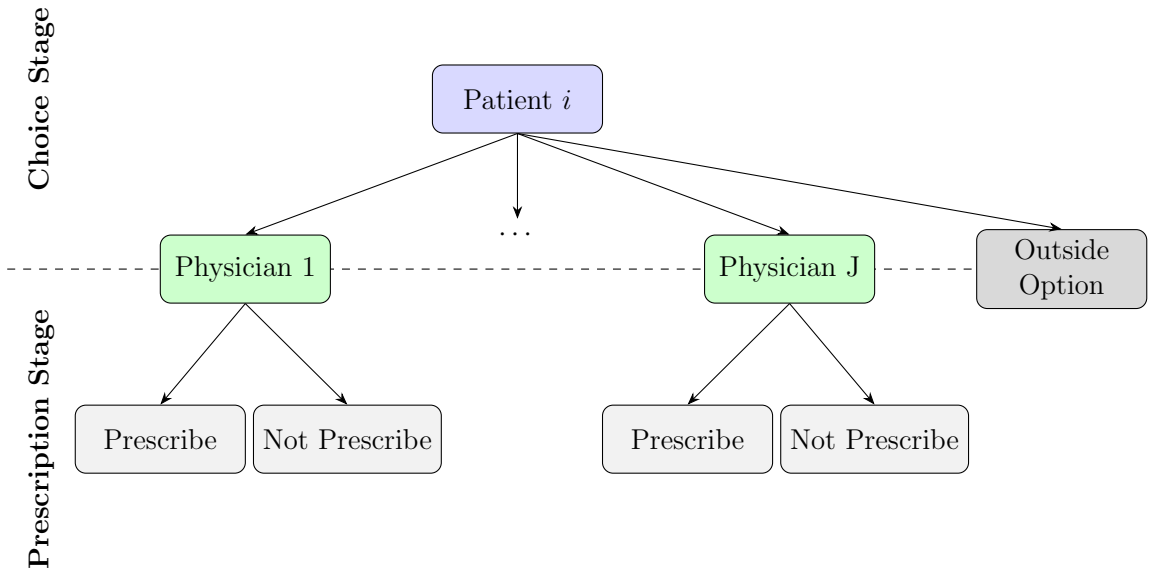


Figure 7: Illustration of two-stage model for each month t.

4.1 Indirect Utility and Information Environment

Physician's Indirect Utility Function. In each month t, conditional on being chosen by patient i, physician j makes a binary treatment decision: prescribe an opioid or not.

Physician j's indirect utility from treating patient i is given by

$$V_{ijt} = -\gamma + \pi_{jt} + \eta_j a_{ijt} + U_{ijt}(a_{ijt}) + \varepsilon_{ijt}(a_{ijt}). \tag{7}$$

Here, π_{jt} represents the baseline profit from seeing a patient, measured in utils and independent of treatment choices, such as the office visit fee, η_j is a physician-specific fixed effect capturing physician j's inherent propensity to prescribe opioids, reflecting factors such as treatment philosophy, graduate medical education, risk tolerance, and regulatory scrutiny, a_{ijt} equals 1 if physician j prescribes an opioid to patient i in month t, and 0 otherwise, γ represents Euler's constant, and ε is an i.i.d. alternative-specific Type 1 extreme value shock, independent of all other components. In this specification, everything else equal, physicians with higher η_j are more likely to prescribe, whereas those with lower η_j prescribe less often, *independent* of the patient's utility from receiving a prescription.

Physicians' Information. Physicians make prescription decisions by combining patient needs they can learn with their own treatment preferences. While physicians are medical experts, they may not fully observe a patient's preferences. We assume that the components of patient preference that are observable and unobservable to the physician are additively separable, such that

$$U_{ijt}(a_{ijt}) = u_{ijt}(a_{ijt}) + \nu_{ijt}, \tag{8}$$

where $u_{ijt}(a_{ijt})$ represents the part of patient utility observable to the physician, and ν_{ijt} the unobservable part. We assume that while physicians do not observe the realization of ν_{ijt}, they know its distribution. In addition, we assume that this informational structure is common knowledge to both patients and physicians.

Patient's Indirect Utility Function. We start by defining patient types in our model. We assume that patient i's type is a tuple (Θ_i, ζ_i), consisting of two time-invariant components. The first component, Θ_i, captures both medical needs and non-medical preferences for opioids. The second component, ζ_i, measures how informed patient i is about physicians' prescription types. We assume that Θ_i enters patients' utility, whereas the information type ζ_i affects patient choice only through patient beliefs and does not directly enter utility. We refer to Θ_i as the patient's *opioid preference type* and to ζ_i as the patient's *information type*.

Patient i's indirect utility from choosing physician j at time t is

$$U_{ijt}(a_{ijt}) = \underbrace{\delta_{jt} + a_{ijt}(\Theta_i + X_{it}\alpha) + Z_{it}\phi}_{\equiv u_{ijt}(a_{ijt})} + \nu_{ijt}, \tag{9}$$

where δ_{jt} denotes the mean utility associated with physician j in month t. The mean utility vary across physicians because they differ in amenities or are subject to physician-specific demand shocks. For example, a physician may have more friendly front desk staff, offer appointment reminders, or be perceived as having greater expertise in treating chronic pain, independent of their prescribing behavior.

Θ_i measures patient i's inherent preference for opioid prescriptions. As discussed in section 3, it includes both patient medical needs and non-medical preferences, which may arise from prior opioid misuse or addiction. As in equation (5), we assume that patients' preference types are a linearly seperable function of patients' medical needs ($X_{1i}\beta_1$) and non-medical preference ($X_{2i}\beta_2 + \theta_i$).

X_{it} denotes time-varying patient characteristics that affect the benefit from receiving an opioid prescription at time t, such as yearly trend of patient taste for prescription opioids. Z_{it} represents time-varying patient characteristics that affect the utility from an office visit at time t, and ν_{ijt} is an i.i.d. Type 1 extreme value shock. We further assume that the distribution of patients' opioid preference type Θ_i has mean zero, $\mathrm{E}[\Theta_i] = 0$, and normalize the value of the outside option (not seeing any physician) to $U_{i0t} = \nu_{i0t}$.

Patients' Information. We assume that patients do not have perfect information about physicians' preferences. First, patients cannot observe the alternative-specific preference shock $\varepsilon_{ijt}(a_{ijt})$, because this shock is realized only after a physician is chosen. We assume, however, that patients know the distribution of this shock. Second, patients do not perfectly observe physicians' inherent prescription propensity η_j when choosing a physician. Instead, we assume that patient i's perceived prescription type of physician j, denoted by ω_{ij}, equals the physician's true type plus noise:

$$\omega_{ij} = \eta_j + \zeta_i. \tag{10}$$

A patient's information type ζ_i captures their uncertainty about a physician's prescription type η_j. We allow patients to differ in their level of informedness, represented by their information type. In practice, patients may differ in their their ability to search for and interpret informal information sources, such as online forums (e.g., Bluelight, Reddit), word-of-mouth, or prior experiences, which can systematically shape how well informed they are about physicians' willingness to prescribe opioids.

We further assume that $\zeta \overset{\text{i.i.d.}}{\sim} N(0, \sigma_\zeta)$. Under this information structure, a given patient either overestimates or underestimates *all* physicians' prescription propensities by the same amount. We interpret this assumption as reflecting a uniform bias in their expectations, such as general optimism or skepticism about physicians' willingness to prescribe, rather than idiosyncratic misperceptions about individual physicians. If all patients have accurate perception on each physician, the variance of ζ_i will be zero. Finally, we assume that the information structure of the game is common knowledge to both physicians and patients.

Given the two sources of uncertainty regarding a physician's preference, patients cannot perfectly predict a physician's treatment decision at the time of choosing a physician. Instead, they form expectations about a physician's probability of prescribing based on their perceived prescription type.

4.2 Equilibrium

This two-stage model is solved by backwards induction. We start by describing a physician's optimal prescribing behavior, conditional on being chosen by the patient. We then derive patients' optimal physician choice in the first stage.

4.2.1 Optimal Prescription Choice

Because the physician can observe the patient's utility up to the realization of the shock ν_{ijt}, we can rewrite the physician's indirect utility V_{ijt} as an explicit function of the prescription decision a_{ijt}:

$$V_{ijt}(a_{ijt} = 1) = \pi_{jt} + \eta_j + \delta_{jt} + \Theta_i + X_{it}\alpha + Z_{it}\phi + \varepsilon_{ijt}(1), \quad \text{and} \tag{11}$$

$$V_{ijt}(a_{ijt} = 0) = \pi_{jt} + \delta_{jt} + Z_{it}\phi + \varepsilon_{ijt}(0). \tag{12}$$

Physician j will choose to prescribe if her indirect utility from prescribing an opioid weakly exceeds the indirect utility from not prescribing i:

$$V_{ijt}(a_{ijt} = 1) - V_{ijt}(a_{ijt} = 0) = \eta_j + \Theta_i + X_{it}\alpha + \varepsilon_{ijt}(1) - \varepsilon_{ijt}(0) \geq 0. \tag{13}$$

Note that this difference and, therefore, the physician's prescription decision, is *independent* of the demand side parameters δ_{jt}, the patient's information type ζ_i, and the base profit from treating a patient π_{jt}. Therefore, in deciding whether to prescribe, the physician does not consider her relative attractiveness to the patient or the information a patient has about

her prescription type in the first stage. Instead, the prescription decision depends solely on the patient's preference from receiving the prescription and the physician's own preference for prescribing opioids.

Since we assume that the alternative-specific error terms ε follow type 1 extreme value distribution, the probability that physician j prescribes an opioid to patient i at time t can be expressed as:

$$\rho_{ijt} = \frac{\exp(\eta_j + \Theta_i + X_{it}\alpha)}{1 + \exp(\eta_j + \Theta_i + X_{it}\alpha)}. \tag{14}$$

4.2.2 Optimal Physician Choice

Patients face a discrete choice among J_t physicians in a market, given their information about each physician. As described in section 4.1, patients do not observe the realization of the type 1 extreme value shocks ε and instead hold potentially inaccurate perception ω_{ij} about each physician's prescription propensity η_j. Based on this information, patients form beliefs over the probability of receiving a prescription from physician j:

$$\tilde{\rho}_{ijt} = \frac{\exp(\omega_{ij} + \Theta_i + X_{it}\alpha)}{1 + \exp(\omega_{ij} + \Theta_i + X_{it}\alpha)}. \tag{15}$$

Given the patient's beliefs about physician j's prescribing probability, $\tilde{\rho}_{ijt}$, conditional on being chosen, and the assumption that ν_{ijt} follows a Type 1 extreme value distribution, the probability that patient i chooses physician j at time t can be expressed as:

$$s_{ijt} = \frac{\exp(\delta_{jt} + \tilde{\rho}_{ijt} \times (\Theta_i + X_{it}\alpha) + Z_{it}\phi)}{1 + \sum_k \exp(\delta_{kt} + \tilde{\rho}_{ikt} \times (\Theta_i + X_{it}\alpha) + Z_{it}\phi)}. \tag{16}$$

Lastly, we obtain aggregate choice probabilities s_{jt} by integrating over individual choice probabilities s_{ijt}

$$s_{jt} = \int s_{ijt} \mathrm{d}i. \tag{17}$$

5 Estimation and Results

In this section, we describe how we estimate our structural model and present the estimation results.

5.1 Estimation

Since our model implies that in the second stage of prescribing, the physician does not consider her relative attractiveness to the patient or the information a patient has about her prescription type in the first stage, it is natural to estimate the choice stage and the prescription stage separately using a limited information maximum likelihood estimation procedure.

In particular, we estimate the model in three steps. In the first step, we focus on the prescription stage and estimate patients' opioid preference types Θ_i, physicians' prescribing types η_j, and parameters $\boldsymbol{\alpha}$ that govern other time-varying factors that affect prescription decisions, using realized prescription decisions. In the second step, we take our prescription stage estimates as given and estimate the demand parameters in the choice stage. These include the mean utility associated with a physician δ_{jt}, the standard deviation of patients' information types σ_ζ, and the parameter vector $\boldsymbol{\phi}$, which governs time-varying patient characteristics that affect the utility of an office visit. Identification comes from variation in patients' physician choices. In the last stage, we decompose estimated patients' types Θ_i into patients' medical needs (governed by β_1), patients' opioid preference reflected in observables (governed by β_2), and the residual unobservable patient opioid preference θ_i. We elaborate on each estimation step below.

Step I: Estimation of $\Xi_1 = \{\Theta, \eta, \alpha\}$ from the prescription stage

We first note that the prescription stage depends only on $\Xi_1 = \{\Theta, \eta, \alpha\}$. As shown in equation (14) in section 4, a physician's prescription probability is given by $\rho_{ijt} = \frac{\exp(\eta_j + \Theta_i + \alpha X_{it})}{1+\exp(\eta_j + \Theta_i + \alpha X_{it})}$. We can reformulate this equation to the following two-way fixed effects logistic regression model:

$$\log(\frac{\rho_{ijt}}{1-\rho_{ijt}}) = \eta_j + \Theta_i + \alpha X_{it}. \tag{18}$$

Note that this reduced-form equation derived from the prescription stage closely resembles our Abowd, Kramarz, and Margolis (1999)-style equation in (4) for establishing patient sorting. It differs from equation (4) only due to the parametric type 1 extreme value assumption on the error term ε in physician's indirect utility.

Given its similarity to equation (4), equation (18) is subject to the same identification concerns. In particular, limited mobility bias might cause non-trivial bias when estimating the fixed effects (Andrews, Gill, Schank, and Upward 2012; Bonhomme, Lamadon, and Manresa 2019). Moreover, logit models are known to suffer from incidental parameter bias

when including fixed effects for each patient and physician. To address these issues, we follow Bonhomme, Lamadon, and Manresa (2019) and Mourot (2025) and group physicians and patients in exactly the same manner as in section 3. We then estimate the grouped parameters using theprescription stage panel data via maximum likelihood. This step yields estimates of $\Xi_1 = \{\Theta, \eta, \alpha\}$, which we denote by $\widehat{\Xi}_1$ and take as given in step II.

Step II: Estimation of $\Xi_2 = \{\boldsymbol{\delta}, \sigma_\zeta, \boldsymbol{\phi}\}$ from the choice stage

Conditional on our estimates $\widehat{\Xi}_1$ from the prescription stage, we estimate the demand side parameters $\boldsymbol{\delta}$, $\boldsymbol{\phi}$, and σ_ζ in a second step using maximum likelihood. Our approach is closely related to commonly used estimators in the health economics and industrial organization literature (Goolsbee and Petrin 2004; Ho 2006), but additionally leverages the panel dimension of our choice stage data.

Although in principle the choice stage of our model could be estimated via unconstrained maximum likelihood, doing so would require estimating thousands of δ parameters, which poses a substantial computational burden. Our data contain thousands of physicians observed over 84 months, making it impractical to estimate a separate parameter per each physician-month combination. We therefore follow the industrial organization literature and concentrate out the demand side mean utilities $\boldsymbol{\delta}$ by imposing a market share constraint (Berry 1994; Berry, Levinsohn, and Pakes 1995). We then maximize the constrained log-likelihood of the choice stage over $\Xi_2 = \{\sigma_\zeta, \boldsymbol{\phi}\}$ only.

Conditional on our estimates from the prescription stage $\widehat{\Xi}_1$, the constrained log-likelihood of the choice stage can be written as

$$\mathcal{L}_2 = \sum_{i=1}^{N} \log \int \prod_{t \in T_i} \prod_{j \in J_t} (s_{ijt}(\zeta; \Xi_2, \widehat{\Xi}_1))^{y_{ijt}} f(\zeta; \sigma_\zeta) d\zeta \tag{19}$$

$$s.t. \quad s_{jt}(\Xi_2, \widehat{\Xi}_1) = \mathcal{S}_{jt}, \tag{20}$$

where y_{ijt} is an indicator equal to 1 if patient i chooses physician j in month t, and 0 otherwise. $\mathcal{S}_{jt}$ represents the observed market share of physician j in month t in our data.

Because we use a two-step estimator, we need to adjust the standard errors of $\widehat{\Xi}_2$ in the second stage to account for the first-step estimation. Since we use maximum likelihood for both the first and the second step, it follows that under standard regularity conditions, the variance-covariance matrix for the second-step, $\widehat{\boldsymbol{V}}_{\hat{\Xi}_2}$, is given by (Murphy and Topel 2002):

$$\widehat{\boldsymbol{V}}_{\hat{\Xi}_2} = \widehat{\boldsymbol{V}}_2 + \widehat{\boldsymbol{V}}_2 (\widehat{\boldsymbol{A}} \widehat{\boldsymbol{V}}_1 \widehat{\boldsymbol{A}}') \widehat{\boldsymbol{V}}_2, \tag{21}$$

where $\widehat{\boldsymbol{V}}_1$ is the variance-covariance matrix from the first-step logit estimation, $\widehat{\boldsymbol{V}}_2$ denotes the variance-covariance matrix from the second-step likelihood estimation, and $\widehat{\boldsymbol{A}}$ is defined as

$$\widehat{\boldsymbol{A}} = \sum_i (\frac{\partial \mathcal{L}_{i2}}{\partial \widehat{\Xi}_2})(\frac{\partial \mathcal{L}_{i2}}{\partial \widehat{\Xi}_1'}), \tag{22}$$

where $\mathcal{L}_{i2}$ denotes patient i's individual contribution to the second-step likelihood.

Step III: Separating Patient's Opioid Preference from Patient's Medical Needs

An important question for policy is whether prescription opioids are prescribed for medical reasons or instead reflect non-medical preferences. To answer this question, we decompose our estimates of patients' preference type $\hat{\Theta}_i$, as in equation (5), into three components: patients' medical needs (governed by β_1), observed opioid preference (governed by β_2), and a residual unobservable patient opioid preference term θ_i. We further adjust the corresponding standard errors of step III to account for the multi-step estimation procedure. Specifically, we adjust the standard errors analogous to equation (21) following Murphy and Topel (2002).

5.2 Results

In this work-in-progress version, we abstract from X_{it} and Z_{it}. Estimates of patients' preference and information types, as well as physicians' prescribing types, are reported in table 5. We further decompose patients' preference types in table 6.

Table 5 shows that patient preference types are highly dispersed. In particular, the estimated distribution of patient preference types Θ have long tails. The estimates further show that some patients exhibit a strong preference for prescription opioids, while others exhibit a strong distaste for them.

Table 6 further shows that our estimates for patient preference types are strongly and intuitively correlated with factors that the literature associates with opioid demand. For instance, we estimate that patients with sickle cell disease, a condition explicitly mentioned in CDC prescribing guidelines as warranting opioid prescriptions, exhibit a much higher opioid preference type than patients without the condition. Similarly, patients with higher comorbidity scores, as measured by the Charlson Comorbidity Index, and older patients tend to exhibit higher opioid preference types. At the same time, our estimates are also strongly correlated with factors for which the CDC recommends caution in opioid prescribing. For example, previous opioid misuse or emergency room visits related to overdoses strongly and positively correlate with patients' opioid preference.

Table 5: Estimated Model Parameters

	Estimate	**(Std. Error)**
Patient Preference Type (Θ_i)		
Θ_1	-4.470	(0.588)
Θ_2	-1.972	(0.168)
Θ_3	0.372	(0.136)
Θ_4	1.561	(0.133)
Θ_5	2.013	(0.132)
Θ_6	2.316	(0.133)
Θ_7	2.766	(0.133)
Θ_8	3.047	(0.135)
Θ_9	3.658	(0.140)
Θ_{10}	5.908	(0.197)
Patient Information Type		
σ_ζ	2.711	(0.094)
Physician Prescribing Type (η_j)		
η_2	-8.695	(0.608)
η_3	-6.031	(0.226)
η_4	-3.877	(0.196)
η_5	-3.562	(0.190)
η_6	-3.346	(0.188)
η_7	-3.219	(0.188)
η_8	-3.087	(0.187)
η_9	-2.903	(0.188)
η_{10}	-2.599	(0.190)
Constant	0.427	(0.227)
Year FE	✓	
Data		
Patients (N)	11,909	
Physicians (J)	8,059	
Months (T)	84	
Total Observations	126,357	

Notes: This table reports estimated model parameters from the structural model. Standard errors adjusted for two-step estimation are shown in parentheses.

The estimated variance of the patient information type, σ_ζ, indicates substantial dispersion in patients' perceptions about physicians' prescribing propensities. This suggests considerable heterogeneity in information across patients. While some patients appear to be well informed, others substantially overestimate or underestimate their likelihood of obtaining prescription opioids during office visits.

We further estimate strong heterogeneity in prescribing propensities across physicians. In particular, physicians of the lowest prescribing type are much less likely to prescribe opioids, even for patients with the highest opioid preference types, than physicians of the highest prescribing type. These results suggest that the choice of physician meaningfully affects prescribing outcomes, regardless of patient opioid preference type.

Table 6: Projecting estimated patient preference type onto medical needs and non-medical opioid preference

Panel A: Type 1 (X_1)		**Panel B: Type 2 (X_2)**	
Variable	Estimate (SE)	Variable	Estimate (SE)
Constant	-0.436 (0.225)	ER overdose	0.777 (0.475)
Cancer pain	0.074 (0.106)	Renal/hepatic	-0.236 (0.150)
Sickle cell	1.169 (0.346)	Mental health	0.063 (0.049)
Sex	0.182 (0.046)	Opioid misuse	1.582 (0.133)
Total outpatient (avg.)	-0.004 (0.000)	Substance use	0.306 (0.071)
Total inpatient (avg.)	0.195 (0.066)	Nonfatal overdose	0.124 (0.126)
Charlson CCI (avg.)	0.113 (0.021)	Safety	0.107 (0.082)
OOP visits (avg.)	0.000 (0.000)		
ER visits (avg.)	-0.021 (0.019)		
OOP days (avg.)	0.266 (0.062)		
Age group 2	-1.365 (0.268)		
Age group 3	-1.303 (0.249)		
Age group 4	-1.022 (0.222)		
Age group 5	-0.977 (0.213)		
Age group 6	-0.793 (0.193)		
Age group 7	0.144 (0.282)		
Age group 8	0.143 (0.300)		
Age group 9	0.302 (0.297)		
Age group 10	1.097 (0.288)		
Data			
Patients (N)	11,909		

Notes: OLS estimates from regressions of Θ_i on covariate sets X_1 (Panel A) and X_2 (Panel B). Standard errors adjusted for two-step estimation.

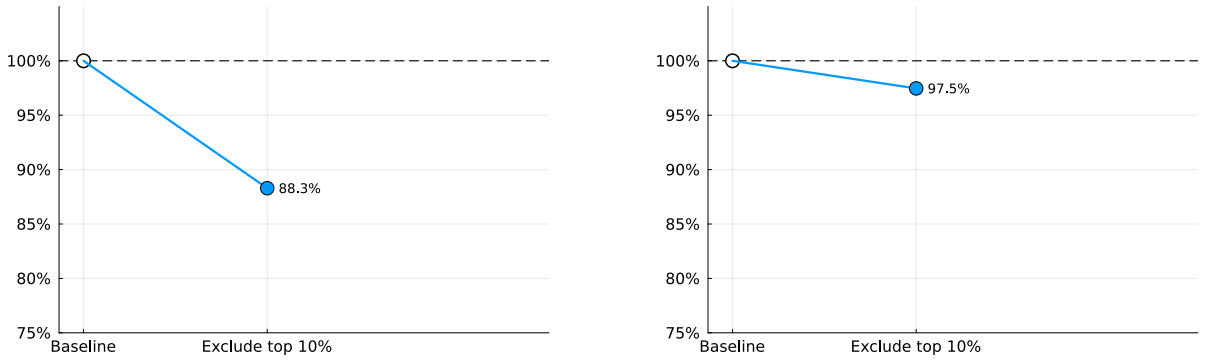
6 Counterfactual Simulations

In this section, we demonstrate how patient sorting can influence the effectiveness of top provider-targeted policies. Moreover, we show how an alternative policy eliminating prescriptions based on non-medically justified preferences would lead to a larger and more targeted reduction in aggregate prescriptions. We highlight the underlying economic mechanisms below and provide technical details on our implementation in Appendix H.

6.1 Excluding Top Prescribers

We start by simulating the *exclusion* of the top 10% of physicians by risk-adjusted prescription rate across all markets.[11] This policy prevents patients from choosing any of the excluded physicians. Compared to real-world policies monitoring or warning the top percentiles of high-prescribing physicians, our counterfactual exclusion policy is extreme but designed to illustrate the underlying economic mechanisms of these policies. We simulate the counterfactual choices for all patients who see an excluded physician in the baseline. Patients who chose an excluded physician in the baseline may either choose to see any of the remaining, non-excluded physicians or the outside option of not seeing any physician in that month. We hold fixed the choices of patients who did not see an excluded physician.

Figure 8: Excluding the top 10% of physicians by risk-adjusted prescription rate



(a) Aggregate prescriptions

(b) Aggregate visits

Notes: This figure plots the aggregate prescriptions and visits after excluding the top 10% of physicians by risk-adjusted prescription rate. The numbers are expressed as percentage points relative to the baseline.

The left panel of Figure 8 compares the baseline and counterfactual aggregate prescriptions in percentage points of the baseline. Aggregate opioid prescriptions are a key metric for policymakers because opioid prescribing in the primary market is closely associated with

[11] We define the risk-adjusted prescription rate in appendix A.

drug overdoses and mortality, with the majority of secondary market opioid supply originating from the primary market (Schnell 2025). Figure 8 shows that our policy reduces aggregate prescriptions by only around 12%, while aggregate visits drop by around 3%. In light of the evidence we presented in section 3 that the prescription intensity distribution of physicians is highly skewed, with only a couple of physicians responsible for a significant fraction of aggregate prescriptions, this muted response is surprising. Given the skewed distribution, everything else equal, excluding high-prescribing physicians should reduce opioid prescription rates more dramatically.

Figure 9: Average choice probabilities for each patient group (Θ group) by physician group (η group).

Notes: Average choice probability of patients conditional on choosing a physician in the market. Patient and physician groups are in ascending order.

However, this argument ignores the demand side response to top physician-targeted policies. To illustrate how the exclusion policy leads to patient choice reoptimization, we plot the average choice probability for each physician prescribing propensity type group η across each patient benefit type group Θ. Figure 9 shows that patients with high opioid benefit type Θ are disproportionally more likely to choose physicians who prescribe opioids more frequently, regardless of patients' benefit type. In other words, when we limit patients' choice sets to a subset of physicians, affected patients will reoptimize by choosing the physician who, ceteris paribus, most closely aligns with their preferences for prescription opioids. However, Figure 9 also demonstrates that resorting is not perfectly aligned with patients automatically choosing the highest prescribers. Our demand system flexibly captures differences in the

attractiveness of physicians, such as quality of care, friendliness of front desk staff, etc., via the mean utility terms δ. Our estimates imply that these differences matter. In particular, we show that although patients with positive benefit type Θ, groups Θ_3 to Θ_{10}, exhibit a strong preference for prescription opioids, patients are slightly more likely to trade off a higher probability for obtaining a prescription for better quality of care. Similarly, patients who dislike prescription opioids, groups Θ_1 and Θ_2, accept higher prescription probabilities for better quality of care. Nevertheless, Figure 9 clearly demonstrates that patients do not randomly choose physicians in response to exclusion policies. Patients resort into physicians.

Next, notice that aggregate visits decline under the exclusion policy as illustrated in the right panel of Figure 8. This implies that a fraction of the observed reduction in the aggregate prescription rate is driven by patients forgoing care. Our results imply that 63% of the predicted reduction in aggregate prescriptions is attributable to patients being forced to switch to physicians with lower prescription propensity, while the remaining 37% is caused by patients forgoing care.

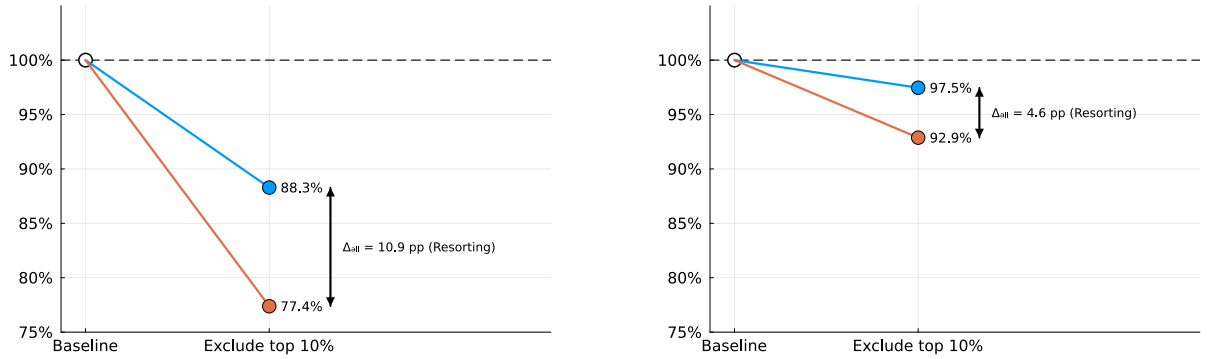
6.2 Excluding Top Prescribers & No Resorting

A surprising outcome of our first counterfactual policy is that the exclusion policy's effect is attenuated by affected patients resorting to remaining physicians who are, everything else equal, most likely to prescribe opioids. In this counterfactual, we quantify the importance of the patient resorting effect by explicitly prohibiting resorting, i.e., preventing affected patients from choosing their physicians. We simulate a counterfactual in which we allow patients to only choose between being randomly assigned to a physician and not seeking care in a given month. As before, we hold fixed all choices of unaffected patients.

Figure 10 illustrates that, without allowing patients to reoptimize their choices of physicians, the exclusion policy studied in counterfactual 1 would reduce the opioid prescription rate by an additional 11 percentage points. When policymakers naïvely ignore patient resorting, expected policy effects may be substantially overstated, aligning more closely with our results from counterfactual 1 than counterfactual 2. If a policy purely focuses on aggregate prescriptions, top provider-targeted policies combined with restrictions for patient choices appear effective.

However, while banning resorting can substantially reduce aggregate prescribing, this policy comes at a cost: patients forgoing care. Note that compared to counterfactual 1, the additional reduction in aggregate prescriptions originates from two channels. First, eliminating resorting prevents patients who remain in the market from optimally choosing their physician. This eliminates resorting to physicians whose prescribing propensities are, ceteris

Figure 10: Excluding the top 10% of physicians and banning patient resorting



(a) Aggregate prescriptions

(b) Aggregate visits

Notes: This figure plots the aggregate prescriptions and visits after excluding the top 10% of physicians by risk-adjusted prescription rate and randomly assigning patients to the remaining physicians. The numbers are expressed as percentage points relative to the baseline.

paribus, close to those of their original physicians who got excluded. Second, prohibiting resorting also induces some patients to forgo care altogether, as random assignment makes the outside option relatively more attractive. The right panel of figure 10 shows that aggregate physician visits drop to only 93% of the baseline when resorting is prohibited. Preventing affected patients from choosing their physician optimally based on their preferences leads many to forgo care. Mechanically, this happens because we estimate substantial heterogeneity in benefit types Θ and prescribing propensity types η. Random assignment makes it very likely that a patient will be matched with a physician whose prescribing type does not align with the patient's preference *irrespective* of the patient's benefit type.

We further decompose the additional 11 percentage point reduction in aggregate prescriptions into the contribution of banning resorting (conditional on market participation) and the contribution of patients forgoing care. We find that 16% of the reduction under randomization is due to the elimination of patient resorting, while the remaining 84% is driven by increased patient exit. Because a patient's benefit type includes both medical necessity and preference for opioids due to addiction, even patients who would qualify for opioid treatment under standard clinical guidelines (e.g., CDC recommendations) may be prevented from obtaining appropriate care. This constitutes a clearly undesirable policy outcome.

In our next counterfactual, we explicitly address the trade-off between limiting non-medically related preference-based sorting and the risk of denying access to medically appropriate opioid prescriptions.

6.3 Breaking Bad Opioid Sorting

We now consider policies that allow patients with objective medical needs to remain in care and receive opioid treatment while at the same time eliminating opioid prescriptions that are based on patients' non-medically justified preferences for prescription opioids. To do so, we distinguish two types of sorting as in section 3: (i) sorting based on non-medically justified preferences for prescription opioids ("bad sorting"), and (ii) sorting based on medical needs ("good sorting"). In our counterfactual simulation, we enforce that physicians prescribe only for legitimate medical reasons. Physicians cannot consider non-medically justified opioid preferences for prescribing. Essentially, this policy aims to break bad sorting based on non-medically justified preferences while preserving good sorting based on objective medical needs. Specifically, for each patient in our data set, we require physicians to prescribe based on the following rules:

1. A physician will consider the full benefit type Θ if the patient suffers from sickle cell disease, cancer-related pain, or is in end-of-life care or palliative care. In other words, we do not interfere with opioid prescribing for chronic pain associated with conditions that are explicitly identified as medically justified in the CDC guidelines.

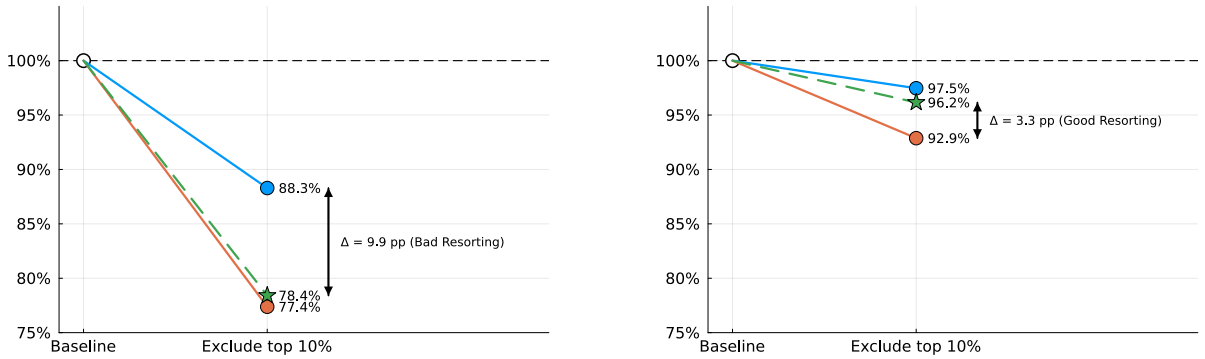
2. In all other cases, if a patient's non-medically justified preference for opioids is positive, i.e., $X_2\hat{\beta}_2+\hat{\theta}_i > 0$, the physician should ignore this preference and base the prescription decision solely on medical needs, $X_1\hat{\beta}_1$. If a patient's non-medically justified preference is negative, we do not intervene, and the physician should take this preference into account as well.[12]

Figure 12 shows that eliminating bad sorting results in almost the same reduction in opioid prescriptions as the randomization policy studied in counterfactual 2. Turning to patient visits, however, we find a stark contrast: unlike randomization, enforcing more stringent opioid prescribing guidelines leads to only a moderate decline in office visits comparable to that under counterfactual 1, because patients are allowed to resort. In other words, our counterfactual policy eliminates bad resorting almost as effectively as randomization while largely keeping patients in care. Figure 12 further shows that these results do not depend on the choice of cutoff value used to exclude physicians.

Such results further suggest that policy could eliminate discretionary prescribing for all physicians, rather than targeting only top prescribers. Our final counterfactual explores this possibiliity by applying the prescription rules to the universe of physicians. We find

[12]This rule implies that if a patient has any conditions mentioned in the CDC guideline (Dowell, Ragan, et al. 2022b) that warrant caution, physicians must not prescribe opioids for pain based on this preference.

Figure 11: Excluding the top 10% of physicians and breaking sorting based on non-medical preference for opioids

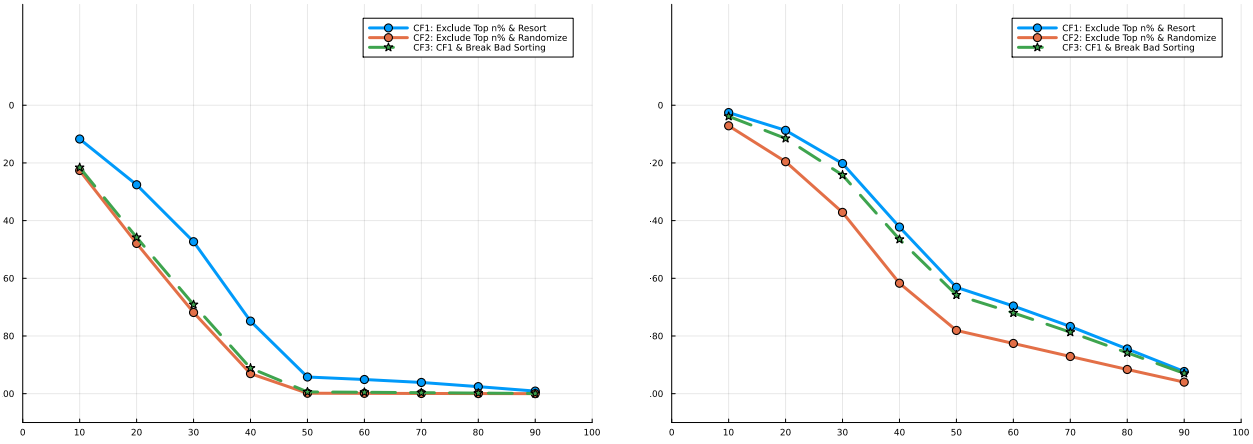


(a) Aggregate prescriptions

(b) Aggregate visits

Notes: This figure plots the aggregate prescriptions and visits after excluding the top 10% of physicians by risk-adjusted prescription rate and eleminating sorting based on non-medical preference for opioids. The numbers are expressed as percentage points relative to the baseline.

Figure 12



(a) Aggregate prescriptions

(b) Aggregate visits

Notes: Prescription rates and visits across all three counterfactuals as a function of exclusion percentile.

that a policy eliminating discretionary prescribing for all physicians, without excluding any physicians from the market, is approximately as effective as an exclusion policy that removes the top 37% of physicians by risk-adjusted prescription rate, while still maintaining access to care for patients with medically justified preference for opioids.

7 Conclusion

In this paper, we provide novel evidence on how consumers sort into experts based on aligned preferences. We demonstrate that patients strongly sort into physicians based on their preferences for prescription opioids and show how such sorting undermines policies that target high-prescribing physicians, substantially attenuating their effectiveness.

Through counterfactual analyses, we show that nearly all of this attenuation of top physician-targeted policies arises from resorting driven by non-medically justified preferences for opioids. We further show that relatively straightforward policies that eliminate physician discretion can fully eliminate undesirable patient sorting while still preserving patients' necessary medical needs.

References

Abowd, John M., Francis Kramarz, and David N. Margolis (1999). "High Wage Workers and High Wage Firms". In: *Econometrica* 67.2, pp. 251–333. (Visited on 08/13/2025).

Alpert, Abby, Sarah Dykstra, and Mireille Jacobson (Feb. 2024). "Hassle Costs versus Information: How Do Prescription Drug Monitoring Programs Reduce Opioid Prescribing?" In: *American Economic Journal: Economic Policy* 16.1, pp. 87–123. DOI: 10.1257/pol.20200579. URL: https://www.aeaweb.org/articles?id=10.1257/pol.20200579.

Alpert, Abby, Stephen D. Schwab, and Benjamin Ukert (Sept. 2025). "Opioid Use and Employment Outcomes: Evidence from the U.S. Military". In: *Review of Economics and Statistics*, pp. 1–45. ISSN: 1530-9142. DOI: 10.1162/rest.a.284. URL: http://dx.doi.org/10.1162/rest.a.284.

Amin-Esmaeili, Masoumeh, Rachel Alinsky, Renee M. Johnson, Anna Gribble, Natasha Oyedele, Taylor Parnham, Lindsey Goddard, and Ryoko Susukida (July 2023). "Impact of unsolicited reporting notifications on providers' prescribing behavior: An experimental study on Maryland PDMP data". In: *Drug and Alcohol Dependence* 248, p. 109896. ISSN: 0376-8716. DOI: 10.1016/j.drugalcdep.2023.109896. URL: http://dx.doi.org/10.1016/j.drugalcdep.2023.109896.

Andrews, M.J., L. Gill, T. Schank, and R. Upward (2012). "High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias". In: *Economics Letters* 117.3, pp. 824–827. ISSN: 0165-1765. DOI: https://doi.org/10.1016/j.econlet.2012.07.033. URL: https://www.sciencedirect.com/science/article/pii/S0165176512004272.

Badinski, Ivan, Amy Finkelstein, Matthew Gentzkow, and Peter Hull (Oct. 2023). *Geographic Variation in Healthcare Utilization: The Role of Physicians*. Working Paper 31749. National Bureau of Economic Research. DOI: 10.3386/w31749. URL: http://www.nber.org/papers/w31749.

Barnett, Michael L., Andrew R. Olenski, and Anupam B. Jena (Feb. 2017). "Opioid-Prescribing Patterns of Emergency Physicians and Risk of Long-Term Use". In: *New England Journal of Medicine* 376.7, pp. 663–673. ISSN: 1533-4406. DOI: 10.1056/nejmsa1610524. URL: http://dx.doi.org/10.1056/NEJMsa1610524.

Berry, Steven (1994). "Estimating Discrete-Choice Models of Product Differentiation". In: *The RAND Journal of Economics* 25.2, pp. 242–262. ISSN: 07416261. URL: http://www.jstor.org/stable/2555829 (visited on 10/06/2025).

Berry, Steven, James Levinsohn, and Ariel Pakes (1995). "Automobile Prices in Market Equilibrium". In: *Econometrica* 63.4, pp. 841–890. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/2171802 (visited on 10/06/2025).

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler (Apr. 2023). "How Much Should we Trust Estimates of Firm Effects and Worker Sorting?" In: *Journal of Labor Economics* 41.2, pp. 291–322.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa (2019). "A Distributional Framework for Matched Employer Employee Data". In: *Econometrica* 87.3, pp. 699–739. DOI: https://doi.org/10.3982/ECTA15722. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA15722. URL: https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA15722.

Buchmueller, Thomas C. and Colleen Carey (Feb. 2018). "The Effect of Prescription Drug Monitoring Programs on Opioid Utilization in Medicare". In: *American Economic Journal: Economic Policy* 10.1, pp. 77–112. DOI: 10.1257/pol.20160094. URL: https://www.aeaweb.org/articles?id=10.1257/pol.20160094.

Card, David, Ana Rute Cardoso, Jörg Heining, and Patrick Kline (Jan. 2018). "Firms and Labor Market Inequality: Evidence and Some Theory". In: *Journal of Labor Economics* 36.S1, S13–S70. ISSN: 1537-5307. DOI: 10.1086/694153. URL: http://dx.doi.org/10.1086/694153.

Card, David, Jörg Heining, and Patrick Kline (May 2013). "Workplace Heterogeneity and the Rise of West German Wage Inequality*". In: *The Quarterly Journal of Economics* 128.3, pp. 967–1015. ISSN: 1531-4650. DOI: 10.1093/qje/qjt006. URL: http://dx.doi.org/10.1093/qje/qjt006.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson (Aug. 2016). "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector". In: *American Economic Review* 106.8, pp. 2110–44. DOI: 10.1257/aer.20151080. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20151080.

Charlson, Mary E., Peter Pompei, Kathy L. Ales, and C.Ronald MacKenzie (Jan. 1987). "A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation". In: *Journal of Chronic Diseases* 40.5, pp. 373–383. ISSN: 0021-9681. DOI: 10.1016/0021-9681(87)90171-8. URL: http://dx.doi.org/10.1016/0021-9681(87)90171-8.

Chou, Roger, Daniel Hartung, Judith Turner, Ian Blazina, Brian Chan, Ximena Levande, Marian McDonagh, Shelley Selph, Rongwei Fu, and Miranda Pappas (2020). "Opioid Treatments for Chronic Pain". In: *Rockville (MD): Agency for Healthcare Research and*

Quality (US); 2020 Apr. (Comparative Effectiveness Review, No. 229.) URL: https://www.ncbi.nlm.nih.gov/books/NBK556253/.

Clemens, Jeffrey, Pierre-Thomas Léger, Yashna Nandan, and Robert Town (Oct. 2024). *Physician Practice Preferences and Healthcare Expenditures: Evidence from Commercial Payers*. Working Paper 33090. National Bureau of Economic Research. DOI: 10.3386/w33090. URL: http://www.nber.org/papers/w33090.

Currie, Janet, Anran Li, and Molly Schnell (Jan. 2023). *The Effects of Competition on Physician Prescribing*. Working Paper 30889. National Bureau of Economic Research. DOI: 10.3386/w30889. URL: http://www.nber.org/papers/w30889.

Currie, Janet and Hannes Schwandt (2020). *The Opioid Epidemic Was Not Primarily Caused by Economic Distress But by Other Factors that Can be More Readily Addressed*. Tech. rep.

Cutler, David, Jonathan S. Skinner, Ariel Dora Stern, and David Wennberg (Feb. 2019). "Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending". In: *American Economic Journal: Economic Policy* 11.1, pp. 192–221. DOI: 10.1257/pol.20150421. URL: https://www.aeaweb.org/articles?id=10.1257/pol.20150421.

Dahlhamer, James M., Eric M. Connor, Jonaki Bose, Jacqueline W. Lucas, and Carla E. Zelaya (Aug. 2021). "Prescription opioid use among adults with chronic pain: United States, 2019". In: *National Health Statistics Reports; no 162. Hyattsville, MD: National Center for Health Statistics. 2021.* DOI: https://dx.doi.org/10.15620/cdc:107641.

Department of Health and Human Services (2025). "Substance Use-Disorder Prevention that Promotes Opioid Recovery and Treatment for Patients and Communities (SUPPORT) Act Section 7024. Report to Congress on Opioid Prescribing Limitations". In: (). URL: https://www.fda.gov/media/147152/download (visited on 11/10/2025).

Deyo, R (June 1992). "Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases". In: *Journal of Clinical Epidemiology* 45.6, pp. 613–619. ISSN: 0895-4356. DOI: 10.1016/0895-4356(92)90133-8. URL: http://dx.doi.org/10.1016/0895-4356(92)90133-8.

Ding, Yu and Chenyuan Liu (Dec. 2021). "Alternative payment models and physician treatment decisions: Evidence from lower back pain". In: *Journal of Health Economics* 80, p. 102548. ISSN: 0167-6296. DOI: 10.1016/j.jhealeco.2021.102548. URL: http://dx.doi.org/10.1016/j.jhealeco.2021.102548.

Dowell, Deborah, Tamara M. Haegerich, and Roger Chou (Mar. 2016). "CDC Guideline for Prescribing Opioids for Chronic Pain — United States, 2016". In: *MMWR. Recommenda-*

tions and Reports 65.1, pp. 1–49. ISSN: 1545-8601. DOI: 10.15585/mmwr.rr6501e1. URL: http://dx.doi.org/10.15585/mmwr.rr6501e1.

Dowell, Deborah, Kathleen R. Ragan, Christopher M. Jones, Grant T. Baldwin, and Roger Chou (Nov. 2022a). "CDC Clinical Practice Guideline for Prescribing Opioids for Pain — United States, 2022". In: *MMWR Recomm Rep 2022; 71(No. RR-3):1-95*. DOI: http://dx.doi.org/10.15585/mmwr.rr7103a1.

— (Nov. 2022b). "CDC Clinical Practice Guideline for Prescribing Opioids for Pain—United States, 2022". In: *MMWR. Recommendations and Reports* 71.3, pp. 1–95. ISSN: 1545-8601. DOI: 10.15585/mmwr.rr7103a1. URL: http://dx.doi.org/10.15585/mmwr.rr7103a1.

Eichmeyer, Sarah and Jonathan Zhang (Oct. 2022). "Pathways into Opioid Dependence: Evidence from Practice Variation in Emergency Departments". In: *American Economic Journal: Applied Economics* 14.4, pp. 271–300. ISSN: 1945-7790. DOI: 10.1257/app.20210048. URL: http://dx.doi.org/10.1257/app.20210048.

— (Jan. 2023). "Primary care providers' influence on opioid use and its adverse consequences". In: *Journal of Public Economics* 217, p. 104784. ISSN: 0047-2727. DOI: 10.1016/j.jpubeco.2022.104784. URL: http://dx.doi.org/10.1016/j.jpubeco.2022.104784.

Finkelstein, Amy, Matthew Gentzkow, and Dean Li (Aug. 2025). "What Drives Risky Prescription Opioid Use? Evidence From Migration". In: *The Quarterly Journal of Economics* 140.4, pp. 3133–3189. ISSN: 1531-4650. DOI: 10.1093/qje/qjaf037. URL: http://dx.doi.org/10.1093/qje/qjaf037.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams (July 2016). "Sources of Geographic Variation in Health Care: Evidence From Patient Migration*". In: *The Quarterly Journal of Economics* 131.4, pp. 1681–1726. ISSN: 1531-4650. DOI: 10.1093/qje/qjw023. URL: http://dx.doi.org/10.1093/qje/qjw023.

Goolsbee, Austan and Amil Petrin (2004). "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV". In: *Econometrica* 72.2, pp. 351–381. DOI: https://doi.org/10.1111/j.1468-0262.2004.00494.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0262.2004.00494.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0262.2004.00494.x.

Ho, Katherine (2006). "The welfare effects of restricted hospital choice in the US medical care market". In: *Journal of Applied Econometrics* 21.7, pp. 1039–1079. DOI: https://doi.org/10.1002/jae.896. eprint: https://onlinelibrary.wiley.

com/doi/pdf/10.1002/jae.896. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.896.

Hsu, Jhon, Julia Schmittdiel, Edward Krupat, Terry Stein, David Thom, Bruce Fireman, and Joe Selby (Mar. 2003). "Patient choice. A randomized controlled trial of provider selection." eng. In: *Journal of General Internal Medicine* 18 (5), pp. 319–325. ISSN: 1533-4406. DOI: 10.1046/j.1525-1497.2003.20145.x.

Kline, Patrick (2024). "Chapter 2 - Firm wage effects". In: ed. by Christian Dustmann and Thomas Lemieux. Vol. 5. Handbook of Labor Economics. Elsevier, pp. 115–181. DOI: https://doi.org/10.1016/bs.heslab.2024.11.005. URL: https://www.sciencedirect.com/science/article/pii/S1573446324000051.

Lee, Marion, Sanford M Silverman, Hans Hansen, Vikram B Patel, and Laxmaiah Manchikanti (2011). "A comprehensive review of opioid-induced hyperalgesia". In: *Pain Physician*, pp. 145–61.

Lucas, Jacqueline W. and Inderbir Sohi (Nov. 2024). "Chronic Pain and High-impact Chronic Pain in U.S. Adults, 2023". In: *NCHS data brief, no 518, Hyattsville, MD: National Center for Health Statistics. 2024.* DOI: https://dx.doi.org/10.15620/cdc/169630.

Meara, Ellen, Jill R. Horwitz, Wilson Powell, Lynn McClelland, Weiping Zhou, A. James O'Malley, and Nancy E. Morden (July 2016). "State Legal Restrictions and Prescription-Opioid Use among Disabled Adults". In: *New England Journal of Medicine* 375.1, pp. 44–53. ISSN: 1533-4406. DOI: 10.1056/nejmsa1514387. URL: http://dx.doi.org/10.1056/NEJMsa1514387.

Molitor, David (Feb. 2018). "The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration". In: *American Economic Journal: Economic Policy* 10.1, pp. 326–56. DOI: 10.1257/pol.20160319. URL: https://www.aeaweb.org/articles?id=10.1257/pol.20160319.

Mourot, Pauline (2025). "Should Top Surgeons Practice at Top Hospitals? Sorting and Complementarities in Healthcare". In: *Working Paper, Boston University.* URL: https://pmourot.github.io/pmourot_main/doctor_hospital_quality_pmourot.pdf.

Murphy, Kevin M. and Robert H. Topel (2002). "Estimation and Inference in Two-Step Econometric Models". In: *Journal of Business Economic Statistics* 20.1, pp. 88–97. ISSN: 07350015. URL: http://www.jstor.org/stable/1392152 (visited on 10/06/2025).

Sacks, Daniel W., Alex Hollingsworth, Thuy Nguyen, and Kosali Simon (Mar. 2021). "Can policy affect initiation of addictive substance use? Evidence from opioid prescribing".

In: *Journal of Health Economics* 76, p. 102397. ISSN: 0167-6296. DOI: 10.1016/j.jhealeco.2020.102397. URL: http://dx.doi.org/10.1016/j.jhealeco.2020.102397.

Schneeweiss, Sebastian, Philip S. Wang, Jerry Avorn, and Robert J. Glynn (Aug. 2003). "Improved Comorbidity Adjustment for Predicting Mortality in Medicare Populations". In: *Health Services Research* 38.4, pp. 1103–1120. ISSN: 1475-6773. DOI: 10.1111/1475-6773.00165. URL: http://dx.doi.org/10.1111/1475-6773.00165.

Schnell, Molly (2025). "Physician behavior in the presence of a secondary market: The case of prescription opioids". In: *Econometrica* forthcoming.

Schnell, Molly and Janet Currie (Aug. 2018). "Addressing the Opioid Epidemic: Is There a Role for Physician Education?" In: *American Journal of Health Economics* 4.3, pp. 383–410. ISSN: 2332-3507. DOI: 10.1162/ajhe_a_00113. URL: http://dx.doi.org/10.1162/ajhe_a_00113.

Staiger, Becky, Laurence Baker, and Tina Hernandez-Boussard (2022). "Provider Opioid Prescribing Behaviors and Opioid Use in Medicaid". In: *Working Paper*.

Volkow, Nora D. and Thomas McLellan (Mar. 2016). "Opioid Abuse in Chronic Pain–Misconceptions and Mitigation Strategies." eng. In: *The New England Journal of Medicine* 374, pp. 1253–1263. ISSN: 1533-4406. DOI: 10.1056/NEJMra1507771.

A Grouping Algorithm

To group both patients and providers, we first run two regressions to construct the average risk-adjusted prescription rate across patients for providers, and the average risk-adjusted prescription rate across providers for patients. In particular, we run

$$\text{prescribe}_{ijt} - X_{it}\beta = \tilde{\eta}_j + \epsilon_{ijt}, \tag{23}$$

and

$$\text{prescribe}_{ijt} - X_{it}\beta = \tilde{\theta}_i + \epsilon_{ijt}, \tag{24}$$

where X_{it} is the patient characteristics defined as before. We then use our estimates for $\tilde{\eta}_j$ and $\tilde{\theta}_i$ to group patients and providers using the k-mean clustering algorithm (Bonhomme, Lamadon, and Manresa 2019) as in Mourot (2025). In the second stage, we can then consistently estimate the two-way fixed effects model in equation (4) using our groups from the first stage.

B Figures

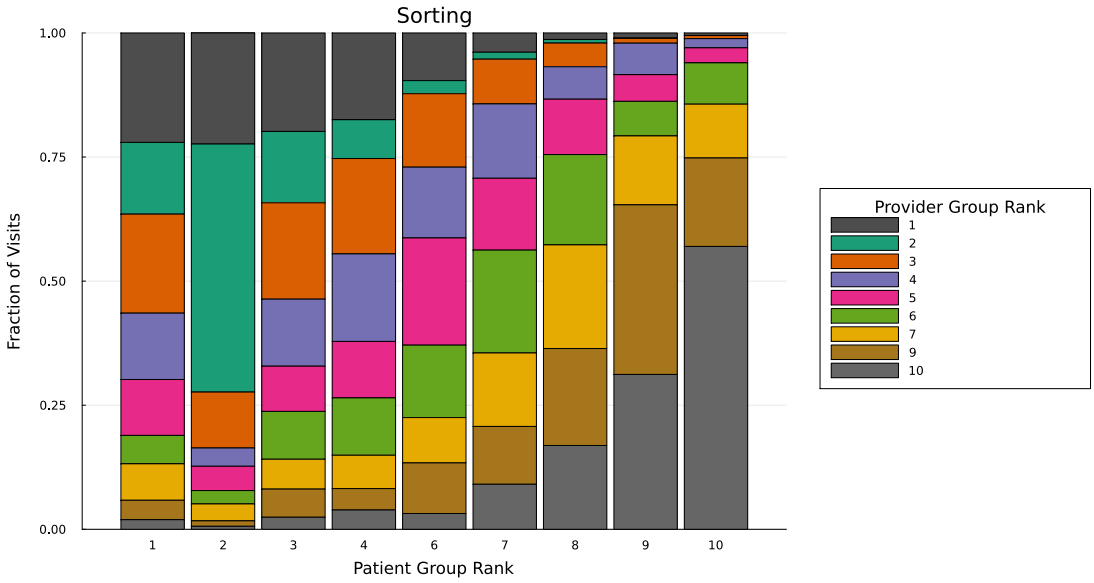


Figure 13: Sorting

C Gradients

First, we define

$$s_{ijt} = \int_{-\infty}^{+\infty} s_{ijth} di, \tag{25}$$

and

$$\tilde{\rho}_{ijt} = \int_{-\infty}^{+\infty} \tilde{\rho}_{ijth} di, \tag{26}$$

where s_{ijth} is the individual choice probability at each integration node h, and $\tilde{\rho}_{ijth}$ is the perceived prescription probability at each node. We can now derive the following partial derivatives. For each element m of the ϕ vector, we have

$$\frac{\partial s_{jt}}{\partial \phi^{\{m\}}} = \sum_i \int_{-\infty}^{+\infty} \frac{\partial s_{ijth}}{\partial \phi^{\{m\}}} di = \sum_i \int_{-\infty}^{+\infty} z_i^{\{m\}} s_{ijth} s_{i0th} di. \tag{27}$$

Let ζ_h be the value of the integration node. Then with some abuse of notation (integral vs sum)

$$\frac{\partial s_{jt}}{\partial \sigma_\zeta} = \sum_i \int_{-\infty}^{+\infty} \frac{\partial s_{ijth}}{\partial \sigma_\zeta} di = \sum_i \int_{-\infty}^{+\infty} s_{ijth}(\Theta_i + \alpha X_{it}) \zeta_h [\tilde{\rho}_{ijth}(1-\tilde{\rho}_{ijth}) - \sum_k \tilde{\rho}_{ikth}(1-\tilde{\rho}_{ikth}) s_{ikth}] di. \tag{28}$$

For each element m of the η vector we have

$$\frac{\partial s_{jt}}{\partial \eta_j^{\{m\}}} = \int_{-\infty}^{+\infty} \frac{\partial s_{jit}}{\partial \eta_j^{\{m\}}} di = \sum_i \int_{-\infty}^{+\infty} s_{ijth}(\Theta_i + \alpha X_{it}) [\tilde{\rho}_{ijth}(1-\tilde{\rho}_{ijth}) - \sum_k \tilde{\rho}_{ikth}(1-\tilde{\rho}_{ikth}) s_{ikth}] di. \tag{29}$$

Similarly, for the constant term, we have

$$\frac{\partial s_{jt}}{\partial cons} = \int_{-\infty}^{+\infty} \frac{\partial s_{jit}}{\partial cons} di = \sum_i \int_{-\infty}^{+\infty} s_{ijth}(\Theta_i + \alpha X_{it}) [\tilde{\rho}_{ijth}(1-\tilde{\rho}_{ijth}) - \sum_k \tilde{\rho}_{ikth}(1-\tilde{\rho}_{ikth}) s_{ikth}] di. \tag{30}$$

For each element m of Θ_i, we have that

$$\frac{\partial s_{jt}}{\partial \Theta_i^{\{m\}}} = \int_{-\infty}^{+\infty} \frac{\partial s_{jit}}{\partial \Theta_i^{\{m\}}} di = \tag{31}$$

$$\sum_i \int_{-\infty}^{+\infty} s_{ijth}[\tilde{\rho}_{ijth}(1-\tilde{\rho}_{ijth})(\Theta_i + \alpha X_{it}) + \tilde{\rho}_{ijth} - \sum_k (\tilde{\rho}_{ikth}(1-\tilde{\rho}_{ikth})(\Theta_i + \alpha X_{it}) + \tilde{\rho}_{ikth}) s_{ikth}] di. \tag{32}$$

For each element m of α we have

$$\frac{\partial s_{jt}}{\partial \alpha^{\{m\}}} = \int_{-\infty}^{+\infty} \frac{\partial s_{jit}}{\partial \alpha^{\{m\}}} di = \tag{33}$$

$$\sum_i \int_{-\infty}^{+\infty} s_{ijth} X_{it}^{\{m\}}[\tilde{\rho}_{ijth}(1-\tilde{\rho}_{ijth})(\Theta_i + \alpha X_{it}) + \tilde{\rho}_{ijth} - \sum_k (\tilde{\rho}_{ikth}(1-\tilde{\rho}_{ikth})(\Theta_i + \alpha X_{it}) + \tilde{\rho}_{ikth}) s_{ikth}] di. \tag{34}$$

For the partial derivatives with respect to δ, we have

$$\frac{\partial s_{jt}}{\partial \delta_{jt}} = \sum_i \int_{-\infty}^{+\infty} \frac{\partial s_{ijth}}{\partial \delta_{jt}} di = \sum_i \int_{-\infty}^{+\infty} s_{ijth}(1 - s_{ijth}) di, \tag{35}$$

and

$$\frac{\partial s_{jt}}{\partial \delta_{lt}} = \sum_i \int_{-\infty}^{+\infty} \frac{\partial s_{ijth}}{\partial \delta_{lt}} di = \sum_i \int_{-\infty}^{+\infty} -s_{ijth}(1 - s_{ilth}) di. \tag{36}$$

The fixed point imposes that the predicted shares are equal to the observed shares. That is, $0 = s_{jt}(\Xi) - \mathcal{S}_{jt}$. By the implicit function theorem, we have that

$$\left(\frac{d\boldsymbol{\delta_t}}{d\sigma_\zeta} \quad \frac{d\boldsymbol{\delta_t}}{d\boldsymbol{\phi}} \quad \frac{d\boldsymbol{\delta_t}}{d\boldsymbol{\eta}} \quad \frac{d\boldsymbol{\delta_t}}{dcons} \quad \frac{d\boldsymbol{\delta_t}}{d\boldsymbol{\Theta}} \quad \frac{d\boldsymbol{\delta_t}}{d\boldsymbol{\alpha}} \right) = -\frac{\partial \boldsymbol{s_t}}{\partial \boldsymbol{\delta_t}}^{-1} \left[\frac{\partial \boldsymbol{s_t}}{\partial \sigma_\zeta} \quad \frac{\partial \boldsymbol{s_t}}{\partial \boldsymbol{\phi}} \quad \frac{\partial \boldsymbol{s_t}}{\partial \boldsymbol{\eta}} \quad \frac{\partial \boldsymbol{s_t}}{\partial cons} \quad \frac{\partial \boldsymbol{s_t}}{\partial \boldsymbol{\Theta}} \quad \frac{\partial \boldsymbol{s_t}}{\partial \boldsymbol{\alpha}} \right].$$

Now, we can use these to form the gradients of the loglikelihood with respect to σ_ζ and ϕ.

$$\frac{ds_{ijt}}{d\sigma_\zeta} = \int_{-\infty}^{+\infty} \frac{ds_{ijth}}{d\sigma_\zeta} di = \int_{-\infty}^{+\infty} s_{ijth}[\frac{d\delta_{jt}}{d\sigma_\zeta} + \frac{d\tilde{\rho}_{ijth}}{d\sigma_\zeta} - \sum_k (\frac{d\delta_{kt}}{d\sigma_\zeta} + \frac{d\tilde{\rho}_{ikth}}{d\sigma_\zeta}) s_{ikth}] di \tag{37}$$

$$\frac{ds_{ijt}}{d\phi^{\{m\}}} = \int_{-\infty}^{+\infty} \frac{ds_{ijth}}{d\phi^{\{m\}}} di = \int_{-\infty}^{+\infty} s_{ijth}[\frac{d\delta_{jt}}{d\phi^{\{m\}}} + z_i^{\{m\}} - \sum_k (\frac{d\delta_{kt}}{d\phi^{\{m\}}} + z_i^{\{m\}} s_{ikth}]di \tag{38}$$

$$\frac{ds_{ijt}}{d\eta_j} = \int_{-\infty}^{+\infty} \frac{ds_{ijth}}{d\eta_j} di = \int_{-\infty}^{+\infty} s_{ijth}[\frac{d\delta_{jt}}{d\eta_j} + (\Theta_i + \alpha X_{it}) \frac{d\tilde{\rho}_{ijth}}{d\eta_j} - \sum_k (\frac{d\delta_{kt}}{d\eta_j} + (\Theta_i - \alpha X_{it}) \frac{d\tilde{\rho}_{ijth}}{d\eta_j}) s_{ikth}]di \tag{39}$$

$$\frac{ds_{ijt}}{dcons} = \int_{-\infty}^{+\infty} \frac{ds_{ijth}}{dcons} di = \int_{-\infty}^{+\infty} s_{ijth}[\frac{d\delta_{jt}}{dcons} + (\Theta_i + \alpha X_{it}) \frac{d\tilde{\rho}_{ijth}}{dcons} - \sum_k (\frac{d\delta_{kt}}{dcons} + (\Theta_i - \alpha X_{it}) \frac{d\tilde{\rho}_{ijth}}{dcons}) s_{ikth}]di \tag{40}$$

$$\frac{ds_{ijt}}{d\Theta_i} = \int_{-\infty}^{+\infty} \frac{ds_{ijth}}{d\Theta_i} di = \int_{-\infty}^{+\infty} s_{ijth}[\frac{d\delta_{jt}}{d\Theta_i} + (\Theta_i + \alpha X_{it}) \frac{d\tilde{\rho}_{ijth}}{d\Theta_i} + \tilde{\rho}_{ijth} - \sum_k (\frac{d\delta_{kt}}{d\Theta_i} + (\Theta_i - \alpha X_{it}) \frac{d\tilde{\rho}_{ijth}}{d\Theta_i} + \tilde{\rho}_{ikth}) s_{ikth}]di \tag{41}$$

$$\frac{ds_{ijt}}{d\alpha^{\{m\}}} = \int_{-\infty}^{+\infty} \frac{ds_{ijth}}{d\alpha^{\{m\}}} di = \tag{42}$$

$$\int_{-\infty}^{+\infty} s_{ijth}[\frac{d\delta_{jt}}{d\alpha^{\{m\}}} + X_{it}^{\{m\}} \tilde{\rho}_{ijth} + (\Theta_i + \alpha X_{it}) \frac{d\tilde{\rho}_{ijth}}{d\alpha^{\{m\}}} - \sum_k (\frac{d\delta_{kt}}{d\alpha^{\{m\}}} + X_{it}^{\{m\}} \tilde{\rho}_{ikth} + (\Theta_i - \alpha X_{it}) \frac{d\tilde{\rho}_{ijth}}{d\alpha^{\{m\}}}) s_{ikth}]di \tag{43}$$

Note that we can easily calculate the gradients for the outside good choice probabilities because $s_{i0t} = 1 - \sum_k s_{ikt}$. Therefore, the respective gradients are

$$\frac{ds_{i0t}}{d\sigma_\zeta} = -\sum_j \frac{ds_{ijt}}{d\sigma_\zeta}, \tag{44}$$

and

$$\frac{ds_{i0t}}{d\phi^{\{m\}}} = -\sum_j \frac{ds_{ijt}}{d\phi^{\{m\}}}. \tag{45}$$

Lastly, we can plug the total gradients directly into the gradients of the loglikelihood.

$$\frac{d\mathcal{L}}{d\sigma_\zeta} = \sum_t \sum_j \sum_i y_{ijt} \frac{1}{s_{ijt}} \frac{ds_{ijt}}{d\sigma_\zeta} \tag{46}$$

$$\frac{d\mathcal{L}}{d\phi^{\{m\}}} = \sum_t \sum_j \sum_i y_{ijt} \frac{1}{s_{ijt}} \frac{ds_{ijt}}{d\phi^{\{m\}}} \tag{47}$$

D Charlson Commorbidity Index (CCI)

We use weights in Schneeweiss, Wang, Avorn, and Glynn 2003 for each of the 17 conditions suggested for the Charlson/Deyo variant (Deyo 1992). The 17 conditions are:

On top of these conditions, patients in the following age group receive additional points for their Charlson Commorbidity Index:

Comorbid Condition	**ICD10 Diagnosis Codes**	**Weight**
Myocardial Infarction	I21, I22, I25.2	1
Congestive Heart Failure	I09.9, I11.0, I13.0, I13.2, I25.5, I42.0, I42.5–I42.9, I43, I50, P29.0	1
Peripheral Vascular Disease	I70, I71, I73.1, I73.8, I73.9, I77.1, I79.0, I79.2, K55.1, K55.8, K55.9, Z95.8, Z95.9	1
Cerebrovascular Disease	G45, G46, H34.0, I60–I69	1
Dementia	F00–F03, F05.1, G30, G31.1	1
Chronic Pulmonary Disease	I27.8, I27.9, J40–J47, J60–J67, J68.4, J70.1, J70.3	1
Connective Tissue Disease	M05, M06, M31.5, M32–M34, M35.1, M35.3, M36.0	1
Peptic Ulcer Disease	K25–K28	1
Mild Liver Disease	B18, K70.0–K70.3, K70.9, K71.3–K71.5, K71.7, K73, K74, K76.0, K76.2–K76.4, K76.8, K76.9, Z94.4	1
Diabetes w/o Complications	E10.0–E10.1, E10.6, E10.8–E10.9, E11.0–E11.1, E11.6, E11.8–E11.9, E12.0–E12.1, E12.6, E12.8–E12.9, E13.0–E13.1, E13.6, E13.8–E13.9, E14.0–E14.1, E14.6, E14.8–E14.9	1
Diabetes w/ Chronic Complications	E10.2–E10.5, E10.7, E11.2–E11.5, E11.7, E12.2–E12.5, E12.7, E13.2–E13.5, E13.7, E14.2–E14.5, E14.7	2
Paraplegia and Hemiplegia	G04.1, G11.4, G80.1, G80.2, G81, G82, G83.0–G83.4, G83.9	2
Renal Disease	I12.0, I13.1, N03.2–N03.7, N05.2–N05.7, N18, N19, N25.0, Z49.0–Z49.2, Z94.0, Z99.2	2
Cancer	C00–C26, C30–C34, C37–C41, C43, C45–C58, C60–C76, C81–C85, C88, C90–C97	2
Moderate or Severe Liver Disease	I85.0, I85.9, I86.4, I98.2, K70.4, K71.1, K72.1, K72.9, K76.5–K76.7	3
Metastatic Carcinoma	C77–C80	6
HIV/AIDS	B20–B22, B24	6

Table 7: Charlson Comorbidity Index: ICD10 Diagnoses and Weights

Age	Score
<50 years	0
50–59 years	+1
60–69 years	+2
70–79 years	+3
≥80 years	+4

E CDC Clinical Practice Guideline for Prescribing Opioids for Pain

Dowell, Haegerich, and Chou 2016; Dowell, Ragan, et al. 2022b provides recommendations for clinicians providing pain care, including those prescribing opioids, for outpatients aged ≥18 years. For BREAKIN BAD SORTING SECTION, we follow the guidelines in determining which patients deserve an opioid prescription and which patients do not deserve an opioid prescription.

Medical Diagnosis	ICD10 Diagnosis Codes
sickle cell disease	D57
cancer-related pain	G89.3
palliative / end-of-life care	Z51.5

Table 8: Medical Diagnosis deserving of opioids

Medical Diagnosis	ICD10 Diagnosis Codes
sleep-disordered breathing	G47.3
renal or hepatic insufficiency	N17, K72
mental health condition (anxiety, depression and post-traumatic stress disorder)	F32, F41, F43
opioid abuse	F11, T40
substance use disorder	F10, F12-F19
non-fatal overdose	T36 - T39, T41 - T50

Table 9: Medical Diagnosis not deserving of opioids

F Constructing a measure of opioid misuse

Following the literature, we classify an outpatient visit as involving opioid misuse if the patient's daily Morphine Equivalent Dosage (MED) exceeds 120 mg, given the documented correlation between such prescription levels and opioid abuse (Meara et al. 2016; Finkelstein, Gentzkow, and D. Li 2025; Staiger, Baker, and Hernandez-Boussard 2022). In this section, we describe how we calculate the MED for an outpatient visit.

For a visit t, we observe all prescription drugs d, their active opioid ingredients o, and the respective dosages in milligrams (mg). [13] Following the CDC guidelines (Dowell, Ragan, et al. 2022b), we consider commonly prescribed opioid ingredients in pain management: codeine, fentanyl, hydrocodone, hydromorphone, methadone, morphine, oxycodone, oxymorphone, tapentadol, and tramadol. For each prescription drug d, we define its morphine equivalent dosage (MED) as the sum of all opioid's dosage multiplied by the conversion factor:

$$MED_d = \sum_{o \in N_d^{opioid}} dosage_o \times conversion_o$$

N_d^{opioid} denotes all the opioid ingredients found in drug d. $conversion_o$ is the conversion factor used by the CDC to standardize the comparisons between opioids relative to morphine and are found in Dowell, Ragan, et al. 2022b. Hence, patient i's daily MED in outpatient visit t is defined as the sum of all drug's total daily MED:

$$MED_{it} = \sum_{d \in N_{it}^{drug}} \frac{MED_d \times METQTY_{dit}}{DAYSUPP_{dit}}$$

N_{it}^{drug} denotes all the drugs prescribed to patient i in outpatient visit t. $METQTY_{dit}$ is the metric quantity of drug d for patient i in outpatient visit t (i.e., number of pills prescribed). $DAYSUPP_{dit}$ is drug d's days of supplement for patient i in outpatient visit t. Hence, an outpatient visit t is defined as an opioid misuse if MED_{it} exceeds 120 mg.

G Constructing expected out-of-pocket prescription opioid costs

In our data, out-of-pocket opioid costs are observed only for chronic pain patients who fill an opioid prescription. Following the demand-estimation literature (Berry, Levinsohn,

[13] Dosages in micromilligram (mcg) are divided by 1,000 to convert to milligrams (mg).

and Pakes 1995; Goolsbee and Petrin 2004), we construct a measure of expected out-of-pocket costs by computing the average opioid OOP cost at the plan–year–month level among observed users. The year-month level of aggregation lets us smooth high-frequency variation arising from insurance contract nonlinearities and seasonality—such as deductible exhaustion and end-of-year resets—that generate predictable spikes in patient cost sharing.

$$\bar{OOPD}_{it} = \frac{1}{N_{plan(i)t}} \sum_{i' \in \text{plan}(i)t} OOPD_{i't}$$

$N_{plan(i)t}$ is the number of patients in the same plan as patient i in year-month t that had an opioid prescription. $\bar{OOPD}_{it}$ is the average OOPD for all patient i' that are in the same insurance plan as patient i in year-month t.

H Technical Details on Counterfactual Simulations

We consider four counterfactual scenarios. Table 10 presents each of the scenarios. When we exclude any physicians, we focus on the counterfactual shares of the patients who chose the excluded physicians while holding the others fixed. We need (s_{ijt} and ρ_{ijt}), along with the affected patient identity, to examine the counterfactual physician choice and opioid prescriptions. Note: BBS = break bad sorting.

	Did s_{ijt} change?	Did ρ_{ijt} change?	affected patient
CF1 (exclude, resort)	Yes	No	Patients that chose excluded physicians
CF2 (exclude, randomize)	Yes	No	Patients that chose excluded physicians
CF3 (exclude, BBS)	Yes, through $\tilde{\rho}$	Yes	Category B and C patients that chose excluded physicians
CF4 (BBS)	Yes, through $\tilde{\rho}$	Yes	Category B and C patients that chose inside physicians

Table 10: Our counterfactuals

H.1 CF1: Exclude top % physician and affected patients resort

We simulate the exclusion policy by dropping physicians in the top n^{th} percentile of risk-adjusted prescription rate. We set the mean utility δ_{jt} of the excluded physicians as $-\infty$. We use Gaussian quadrature to calculate the choice probabilities s_{ijt}. Thus, the patient's perceived prescription probabilities $\tilde{\rho}_{ijth}$ at each integration node h is:

$$\tilde{\rho}_{ijth} = \frac{\exp(\eta_j + \Theta_i + X_{it}\alpha) * \exp(\sigma_h)}{1 + \exp(\eta_j + \Theta_i + X_{it}\alpha) * \exp(\sigma_h)}$$

and each $\sigma_h = \sqrt{2} * \sigma_\zeta * \zeta_h$, where ζ_h is the integration node. The patient's choice probabilities s_{ijth} are:

$$s_{ijth} = \frac{\exp(\delta_{jt} + \tilde{\rho}_{ijth} \times (\Theta_i + X_{it}\alpha) + \phi Z_{it})}{1 + \sum_k \exp(\delta_{kt} + \tilde{\rho}_{ikth} \times (\Theta_i + X_{it}\alpha) + \phi Z_{it})}$$

then s_{ijt} is:

$$s_{ijt} = \sum_h w_h s_{ijth}$$

H.2 CF2: Exclude top % physician and randomize affected patient choise

For this counterfactual, we use Monte Carlo integration to calculate choice probabilities, s_{ijt}. At each Monte Carlo draw m, we have $J+1$ T1EV errors and one standard normal draw for

each patient i, scaled by the dispersion of information type, σ_ζ. The perceived prescription probabilities are:

$$\tilde{\rho}_{ijtm} = \frac{\exp(\eta_j + \Theta_i + X_{it}\alpha) * \exp(\zeta_m)}{1 + \exp(\eta_j + \Theta_i + X_{it}\alpha) * \exp(\zeta_m)}$$

The indirect utility of a non-excluded physician is given by:

$$u_{ijtm} = \delta_{jt} + [\tilde{\rho}_{ijtm} * (\Theta_i + X_{it}\alpha)] + Z_{it}\phi + \nu_{jtm}$$

We calculate the weighted-average utility of all non-excluded physicians and compare it with the realized T1EV error of the outside option. The weights are the conditional market share of the non-excluded physicians. Patients choose the outside option if and only if:

$$\nu_{0tm} \geq \bar{u}_{itm}$$

where ν_{0tm} is the T1EV of the outside option in market t at MC draw m and $\bar{u}_{itm}$ is the weighted average utility of the non-excluded physicians and takes the form:

$$\bar{u}_{itm} = \sum_{j \notin \text{excluded}} s^*_{jt} \times u_{ijtm}$$

s^*jt is the conditional market share of the non-excluded physicians. We implement this 1000 times per patient and use the average across 1000 trials to calculate the counterfactual probability of choosing the outside option, s_{i0t}. We calculate the counterfactual choice probabilities s_{ijt} by multiplying $1 - s_{i0t}$ with the conditional market shares of the non-excluded physicians.

H.3 CF3: CF1 + lower prescription rate for Category B and C

This counterfactual exercise is similar to CF1, where we exclude physicians and allow patients to resort. Additionally, we enforce the CDC guidelines and remove elements in Θ_i related to opioid preference (i.e., $X_2 b_2$ and θ_i) for patients in Categories B and C. The prescription probabilities for patients in Category A remain unchanged. For example:

$$\text{Category A: } \rho_{ijt} = \frac{\exp(\eta_j + \overbrace{(c + X_1 b_1 + X_2 b_2 + \theta_i)}^{\Theta_i} + X_{it}\alpha)}{1 + \sum_k \exp(\eta_j + \underbrace{(c + X_1 b_1 + X_2 b_2 + \theta_i)}_{\Theta_i} + X_{it}\alpha)}$$

$$\text{Category B, C: } \rho_{ijt} = \frac{\exp(\eta_j + \overbrace{(c + X_1 b_1)}^{\hat{\Theta}_i} + X_{it}\alpha)}{1 + \exp(\eta_j + \underbrace{(c + X_1 b_1)}_{\hat{\Theta}_i} + X_{it}\alpha)}$$

As a result, the choice probabilities s_{ijt} for Category A patients remain unchanged; whereas the choice probabilities for Categories B and C patients are:

$$s_{ijt} = \frac{\exp(\delta_{jt} + \overbrace{\hat{\hat{\rho}}_{ijt}}^{f(\hat{\Theta}_i = c + X_1 b_1)} \times (\overbrace{\Theta_i}^{unchanged} + X_{it}\alpha) + \phi Z_{it})}{1 + \sum_k \exp(\delta_{kt} + \underbrace{\hat{\hat{\rho}}_{ikt}}_{f(\hat{\Theta}_i = c + X_1 b_1)} \times (\underbrace{\Theta_i}_{unchanged} + X_{it}\alpha) + \phi Z_{it})}$$

H.4 CF4: Lower prescription rate for Category B and C

Similar to CF3, we follow the CDC guidelines and lower prescription rates for Category B and C patients. We do not exclude any physicians and allow patients to resort, knowing that physicians are now strictly following the CDC guidelines for opioid prescriptions. We examine the counterfactual shares of patients who originally chose a physician, holding fixed the shares of patients who chose the outside option.