Can a Single Opioid Prescription Make a Difference? Evidence from Physician Prescribing Variation in Emergency Departments

Sarah Eichmeyer and Jonathan Zhang*

June 1, 2020

Abstract

In the past two decades, death rates from opioids have seen a fivefold increase and opioid prescribing has emerged as a leading public health problem in the United States. Clinical guidelines leave many opioid prescribing decisions to physician judgement; we study the extent to which a single opioid prescription in an emergency department, for these marginal cases, can induce long-term dependence and impact health and economic outcomes of a patient. We tackle these questions by leveraging quasi-random assignment of patients to physicians, who vary in their propensity to prescribe opioids. We analyze the universe of electronic health record data for a particularly vulnerable population—veterans—and find that a single opioid prescription can have strong adverse effects on a veteran's long-term outcomes. A single opioid prescription induces a 1.2 percentage point (pp) increase in the probability of long-term prescription opioid use, a 0.34pp increase in development of an opioid use disorder, and a 0.075pp increase in opioid overdose mortality. We find suggestive evidence of both use of and death by heroin and synthetic opioids. Moreover, in settings where the supply of legal prescription opioids is restricted, veterans are more likely to resort to illicit opioids, highlighting the complex interdependencies between legal and illicit sources of opioid supply.

Keywords: opioids, prescription drugs, physician variation, emergency department, patient outcomes, illicit drugs
JEL Classification Codes: I12, I18, H12, K42

*Department of Economics, Stanford University. Eichmeyer: saraeich@stanford.edu; Zhang: jzhang7@stanford.edu

The views expressed herein are those of the authors and do not necessarily reflect those of the United States Department of Veterans Affairs or the Veterans Health Administration. We are extremely grateful to our advisors Liran Einav, Matthew Gentzkow, and David C. Chan, and our VHA supervisor Jodie Trafton for their guidance and support on this project. We would like to thank Abby Alpert, Michael L. Barnett, Luca Braghieri, Tim Bresnahan, Kate Bundorf, Rebecca Diamond, Mark Duggan, David Green, Atul Gupta, Caroline Hoxby, Anupam Jena, Petra Persson, Joseph S. Shapiro, Molly Schnell, Todd Wagner, Heidi Williams, participants at various Stanford and VHA workshops, and the VHA Ci2i, HERC, and PERC teams for their comments and suggestions. This project is supported by VHA Enterprise Opiate Safety Team's Comprehensive Addiction and Recovery Act funds and the Leonard W. Ely and Shirley R. Ely Graduate Student Fellowship via funds given to the Stanford Institute for Economic Policy Research.

1. Introduction

In the past two decades, the opioid epidemic has emerged as a major public health issue in the United States. Every day, more than 130 Americans die from overdosing on opioids (CDC, 2018).[1] In 2016 and 2017 alone, there were 91,000 deaths involving opioids, eclipsing the total number of American casualties in the Vietnam War (NCHS, 2017).

In an attempt to combat the epidemic, by the end of 2018, 32 states had passed laws that impose limits on opioid prescribing for acute pain (Chua et al., 2019; Davis et al., 2019). The optimal design of such opioid prescribing policies needs to take into account both the short-term benefits and long-term costs physicians face when deciding to prescribe opioids to a patient. For instance, opioids are effective at managing acute pain (Furlan et al., 2011; Wu and Raja, 2011); however, continued use increases the risk of developing opioid dependence (Krashin et al., 2016). The evidence in this area lags behind legislation and the medical community is calling for further research to strike a balance between reducing opioid risk in the population and promoting compassionate patient care (Lowenstein et al., 2018; Ross et al., 2011).

Towards that end, in this paper, we quantify the long-run costs associated with prescribing opioids.[2] Specifically, we explore whether a single opioid prescription can induce a patient to become a long-term user of prescription opioids. We investigate whether this long-term opioid use can lead to opioid dependence, abuse, and illicit opioid use (e.g., heroin, fentanyl). Furthermore, we study the chain of downstream effects on long-term health and economic outcomes caused by a single opioid prescription.

We study these questions by concentrating on opioid prescriptions that originate in emergency departments (ED). Patients are quasi-randomly assigned to physicians when they

[1] Opioid analgesics are substances that bind to mu-opioid receptors. These receptors are primarily located in the brain regions that regulate pain and pleasure, and in the brain stem, which regulates respiration. The former explains the widespread use of opioids for pain relief (and its addicting potential), and the latter explains the high rates of deaths from opioid overdoses, caused by respiratory failure. See Volkow and McLellan (2016) for an overview of the pharmacologic properties and abuse-related risks related to opioids.

[2] Quantifying the short-run benefits of prescription opioids is beyond the scope of this paper; however, recent studies have found that non-opioid therapy is just as effective at treating both chronic and acute pain (Dodwell et al., 2010; Friedman et al., 2015; Holdgate and Pollock, 2004; Krebs et al., 2018; Teichman, 2004).

arrive at the ED, alleviating many of the selection issues that plague other health care settings. Furthermore, physicians exhibit wide variation in practice behavior in prescribing opioids, even within the same hospital, while following the same guidelines (wide practice variation in a variety of health care settings: Barnett et al., 2017, 2019; Chan, 2016; Epstein and Nicholson, 2009; Fadlon and Parys, 2019; Finkelstein et al., 2016; Gowrisankaran et al., 2017; Molitor, 2018; Silver, 2019; Tu, 2017; Tsugawa et al., 2017; Van Parys, 2016).[3,4] Together, quasi-random assignment in EDs and variation in physician opioid prescribing allow us to estimate the causal impact of being prescribed an opioid on a patient's long-term outcomes. Econometrically, our empirical strategy utilizes physician leniency in prescribing opioids—within the same hospital and controlling for the time of visit and patient observables—as an instrumental variable for the initial opioid prescription in the ED, much like the judges instrument in applications related to judges and incarceration.[5]

We investigate our research question with the universe of veterans receiving health benefits from the Veterans Health Administration (VHA), the largest integrated health care system in the United States. The VHA has some unique features that make it an attractive setting for our research design. First, due to its integrative nature, we observe a relatively complete history of a veteran's care along with rich data that are rarely measured in other sources, such as indicators for homelessness and lab test results. Second, our within-hospital physician leniency design requires knowledge of the prescribing history of all physicians in a few hospitals (rather than a few physicians in many hospitals). This is only possible when hospitals and providers operate in a closed and integrated health care system that serves a specific population. Lastly, veterans are a particularly vulnerable population who are more

[3]Specifically, this variation in clinical prescribing practice is due to differences in clinicians' interpretation of guidelines asking them to balance risk due to unmanaged pain and risk due to an opoid prescription.

[4]Sowicz et al. (2018) find that 66.4% of the variation in opioid prescribing in VHA EDs (our setting) in 2017 was due to within-facility, across-physician variation. Physician-level variation is even greater in earlier years, consistent with the decline in prescribing rates at the VHA (Sasson et al., 2019).

[5]This judge stringency instrument is commonly used in the economics of crime literature or any setting where an agent has discretion while making repeated decisions. Specifically, it is a residualized leave-out "jackknife" instrument. See, for example, (Doyle, 2007, entry into foster care) ; (Kling, 2006; Bhuller et al., 2019, incarceration); (Dobbie et al., 2017, bankruptcy cases); (Dobbie et al., 2018, pretrial detention); (Duggan, 2005, psychiatrists and antipsychotic drugs); (Doyle et al., 2015, ambulance companies); and (Farre-Mensa et al., 2019, patent examiners), among others.

likely to be suffering from mental health disorders, chronic pain, disability, and may have previously acquired a "taste" for opioids during service (Robin et al., 1974; Cesur et al., 2019). These characteristics have been found to be predictive of opioid dependence (Boscarino et al., 2010; Edlund et al., 2007, 2014; Seal et al., 2012).

Beyond identification, focusing on physician prescribing in emergency departments has broader practical implications. Emergency departments are an important part of the health care system in the United States. In 2016, 19.4% of Americans and 12.7% of veterans, visited an ED (CDC, 2018; Huang et al., 2018). Many emergency visits are pain related; 14.3% (12.6% in VHA EDs) of patients seen at emergency departments in 2011 were prescribed opioids (CDC, 2018), making them non-negligible sources of prescription opioids. Moreover, we observe over 100,000 veterans per year receiving their first opioid (after not taking opioids for over a year) from the ED.

This paper is structured in three stages. First, we evaluate the extent to which short-term exposure to opioids via a more lenient emergency physician can contribute to continued long-term use of legal prescription opioids. Next, we study the entire chain of long-term downstream health consequences. We estimate causal effects on three groups of outcomes: measures of opioid dependence and misuse; measures of mental and physical health, including mortality; and transition into illicit drug use. Finally, we account for unobserved (private coverage, black market, or illicit) opioid use and misuse with a simple accounting exercise.

We establish four main results. First, exposure to a single opioid prescription via being quasi-randomly assigned a more lenient physician in an emergency department can induce long-term prescription opioid use. A single opioid prescription increases the probability of long-term use of legal prescription opioids—defined as at least 180 days supply of new opioids filled in the first year following the initial ED visit—by 1.2 percentage points, on a base of 5.8%. This increased probability of opioid use following the initial ED prescription persists for over 24 months and never declines back to pre-ED opioid use levels.

Second, we find evidence that this long-term use can lead to development of opioid use disorder and is consistent with misuse and dependence rather than appropriate medical

care. An opioid prescription causes a 0.34pp increase in the probability of developing an opioid use disorder within three years of the ED visit. We adopt standard proxies from the medical literature for opioid-seeking behavior, such as having overlapping opioid prescriptions, pharmacy shopping, and repeated visits for back pain and headaches (these conditions are frequently prescribed opioids and are difficult to verify) and find that a single prescription in the ED increases the probability of all three proxies by 1.9pp, 0.30pp, and 0.55pp respectively. Veterans' self-reported pain score also increases in the first year, despite the fact that opioids should be providing pain relief.

Third, we find that beyond increasing the likelihood of opioid dependence and misuse, the ED prescription has severe adverse effects on downstream health outcomes. Recorded opioid overdoses increase by 0.075pp in the first year, accidental falls (an adverse outcome associated with opioid intoxication) increase by 0.5pp, and deaths from opioid overdose increase by 0.075pp. The magnitude on opioid overdose deaths are especially large and troubling. A single legitimate opioid prescription in the ED not only causes patients to become long-term users, but can lead to death. These findings on mortality make up a large portion of the cost of opioid prescribing when considering its trade-off with the benefits of pain management. We also study the transition into illicit opioid use (e.g., heroin and fentanyl) using heroin/fentanyl drug screens and hepatitis C diagnoses as proxies, along with heroin overdose mortality as an outcome; we find suggestive evidence, albeit not statistically significant, that the prescription opioid in the ED triggered some veterans to begin using illicit injection drugs. In patient assessment questionnaires conducted by clinicians, veterans also begin using illicit cocaine, while other substance use (e.g., marijuana, alcohol, sedatives, etc.) remains unchanged. Heroin and synthetic opioids (e.g., fentanyl and tramadol) account for 30-40% of the overall opioid mortality effect caused by a single legitimate prescription opioid.

Last, we investigate the channels of opioid supply through which veterans are satisfying their excess demand. We find suggestive evidence that veterans with a lenient opioid-prescribing primary care provider (PCP) are more likely to become long-term users of legal prescription opioids from a single ED prescription. In contrast, veterans with a strict PCP are

more likely to contract hepatitis C, overdose from opioids, and die from an opioid overdose, consistent with these patients being more likely to turn to illicit drugs. This heterogeneity may suggest that veterans satisfy their demand for opioids from the legal market if a supply of prescription opioids is available from their PCP; otherwise, they may be more likely to resort to more lethal black market or illicit opioids. This substitution pattern between the legal and illicit market can also been seen in ED visits after 2012, when the VA introduced opioid-prescribing guidelines and policies that greatly reduced the supply of prescription opioids during this period.

The findings regarding illict use, along with the observation that we do not observe a recorded overdose for the majority of opioid overdose deaths, suggest that a sizable fraction of total opioid use/abuse (including prescription opioids, black market opioids, and illicit heroin/fentanyl) is unobserved. At the end of this paper, we perform a simple accounting exercise to estimate the amount of unobserved (and total) opioid use and opioid overdose caused by an ED prescription. We conduct back-of-the-envelope calculations suggesting that for every 100,000 veterans prescribed opioids in the ED because of a more lenient physician, 2,600-6,300 will become long-term users, 1,220-3,500 will develop an opioid use disorder within three years, and eventually 75 will die from an opioid overdose (of which 20-30 are due to illicit opioids). Phrased differently, each year roughly 52 veterans die from an opioid overdose caused by an initial ED opioid prescription via a lenient prescriber.

This paper relates to a large body of medical literature about the health consequences of prescription opioid use and a growing economic literature about its economic consequences.[6] The most closely related paper is Barnett, Olenski, and Jena (2017)—henceforth BOJ—who

[6]Other related papers on the effect of opioids on various outcomes include: (Fernandez and Zejcirovi, 2017, pharmaceutical opioid promotion on opioid overdose mortality); (Savych et al., 2018, prescription length on temporary disability receipt); (Quast et al., 2018, child removals); (Currie et al., 2019, state prescription rates and employment). There is also a large growing literature on the impact of various opioid policies on opioid-related outcomes: (Mueller et al., 2015; Rees et al., 2019; Doleac and Mukherjee, 2019, access to opioid antagonists); (Rees et al., 2019, "Good Samaritan" laws); (Buchmueller and Carey, 2018; Borgschulte et al., 2018; Dave et al., 2018; Kilby, 2015; Grecu et al., 2019; Rutkow et al., 2011; Sacks et al., 2019, prescription drug monitoring programs on opioid outcomes); (Nguyen et al., 2019, prescription drug monitoring programs on pharmaceutical promotion); (Kaestner and Ziedan, 2019; Rutkow et al., 2011, "pill-mill" laws); (Packham, 2019, syringe exchange programs); (Baker et al., 2018, enrollment in Medicare advantage plan); among others.

use data on a 20% random sample of Medicare Part D beneficiaries from 2008-2011 and find that patients treated by "high-intensity" prescribers are 30% more likely to be long-term users of legal prescription opioids compared to patients treated by "low-intensity" prescribers. In a concurrent work, Barnett et al. (2019) replicate their 2017 study with VHA data and find similar but attenuated differences in long-term use. Our paper differs from these two papers in three ways. First, we research beyond long-term use by studying detailed measures of opioid-seeking behavior, health outcomes, and mortality. Second, although similar in spirit, our identification strategy differs both in the construction of the instrument and in the main econometric model. Last, we account for unobserved opioid use (non-VHA prescriptions, black market opioids, and illicit heroin/fentanyl) with a accounting exercise. Appendix A provides a detailed discussion of how this paper differs from BOJ and where our findings depart from theirs.

Utilizing an instrumental variable strategy similar to ours, de Vaan and Stuart (2019) study within-household diffusion of opioids in the context of ED patients in Massachusetts and find that spouses of patients who are prescribed opioids have a 19% increase in the probability of filling an opioid prescription within one year. By studying the family household, the authors focus on how the patient demand channel accelerated the opioid epidemic. Other closely related papers studying the extent and consequences of supply-side variation in opioid availability to patients are Laird and Nielsen (2016) and Finkelstein, Gentzkow, and Williams (2018). The former is set in Denmark and exploits variation across primary care providers' prescription rates in conjunction with patient changes in primary care providers due to moves across municipalities. It finds that a 10pp increase in a primary care provider's opioid prescription rate causes a 4.5pp increase in opioid use among patients, and a significant decrease in labor income and employment. The latter studies Medicare Part D patients in the US and exploits geographical (county-level) variation in opioid prescription rates in conjunction with patient moves across counties. It finds that moving to a county with a 20% higher opioid abuse rate increases the mover's propensity to abuse opioids by 6%. In ongoing work (Eichmeyer and Zhang, 2020), we study veterans assigned to primary care providers

in the VHA and study how potential long-term exposure to opioid prescribing in primary care settings can influence health outcomes. We find that a 3pp increase in a primary care provider's opioid prescription rate (equivalent to the difference between the 90th and 10th percentile) is associated with a 0.72pp increase in long-term opioid use. Furthermore, we find that being assigned a high primary care opioid prescriber is associated with adverse effects on opioid use disorders, mental health, attempted suicide and self-harm, visits to the ED, and vehicle accidents. These papers highlight the importance of supply-side effects to developing opioid dependence. We build on this work by directly identifying the effects of individual opioid prescriptions on dependence, health, and economic outcomes while shutting off the myriad of other supply-side factors that could be at play.

It is important to note that this paper studies the impact of an opioid prescription through a prescribing decision requiring clinical judgment rather than through specific VHA policies or differences in adherence to clinical practice guidelines. Moreover, our period of study spans eleven years (2006-2016), and the national narrative regarding the opioid epidemic, clinical guideline recommendations, and VHA practice has changed dramatically over this time. Opioid prescribing and care delivered at the VA, and studied here, were within clinical guidelines during this period. Furthermore, it is not substandard care but rather practice variance *within practice norms* that provides us with the variation for our research design and results in the outcomes presented in this paper. Notably, this design strengthens the utility of the findings, providing data on outcomes within the range of clinical cases in which prescribing decisions are not straight-forward and clinician judgment varies. This may help to fill gaps in clinical practice recommendations.

The remainder of this paper is structured as follows. The next section describes the data source and outlines our baseline sample. The empirical strategy and its accompanying identifying assumptions are laid out in Section 3. Section 4 presents the main results of an opioid prescription on health and economic outcomes. The last section concludes.

2. Data and Definitions

2.1 Data Source

Our empirical analysis uses four data sources from the Veterans Health Administration (VHA): i) VHA electronic health records, ii) Medicare and Medicaid claims records iii) community care (care provided at non-VHA facilities that get billed to the VHA), and iv) date and cause of death records. This section provides an overview of Veterans Health Benefits and summarizes the relevant information for each dataset.

Of the approximately 20 million veterans in the United States, roughly 9 million have Veterans Health Benefits, administered by the VHA. Veterans are eligible if they satisfy one of the following: i) served prior to September 7, 1980; ii) discharged for service-connected disability; or iii) served on active military duty for at least 24 consecutive months. Once eligible, veterans are enrolled in priority groups based on honorable decorations (e.g., Medal of Honor, Purple Heart), disability, exposure to adverse war events (e.g., Agent Orange, ionization radiation, Prisoners of War), and income thresholds. These priority groups determine rates of copayment for health care and medication; there is no annual premium. Eligible veterans can be treated at any of the 152 VA medical centers and approximately 1,400 community-based outpatient clinics across the country. Coverage does not extend to a veteran's spouse or dependents.

Our primary data source is the VHA Corporate Data Warehouse (CDW), which includes electronic health records, patient enrollment and benefits, and other health system records (including referrals, clinician notes, lab test orders, etc.) dating as far back as 1999. This includes standard inpatient, outpatient, and pharmacy data with variables for patient demographics, condition, physician, hospital, etc. Unlike claims data, electronic health record data also include referrals, patient assessment questionnaires, physician notes, lab results, etc. Because of the vast programs covered by the VHA, we also have data regarding non-medical outpatient clinics, for example, the Department of Housing and Urban Development-VA Supportive Housing Program, and the Residential Rehabilitation Treatment Program for

substance use disorders. Emergency department records begin in 2006.

Through the VHA health record data, we observe only care that occur at VHA facilities (this includes the 1,074 outpatient clinics contracted with the VHA). According to a VHA survey (Huang et al., 2018), 80% of VHA benefit enrollees have additional coverage through public or private insurance: 51% have Medicare, 7% have Medicaid, 20% have TRICARE (Department of Defense health care for active duty service members and retirees), and 28% have private insurance coverage. We supplement our CDW data with VA/CMS data, providing us with a more complete view of veterans' health events. This includes Medicare claims from 2011-2016 and Medicaid claims from 2011-2014 for all veterans. We observe medical claims for both Medicare (Part A and B) and Medicaid, along with prescription claims for patients enrolled in any Medicare Part D plan, and Medicaid prescription claims. Approximately 53% of our sample of veterans with ED encounters are enrolled in Medicare or Medicaid (50% have Medicare and 7% have Medicaid coverage), however, only 22% of those have Medicare claims during the year of their emergency. Conditional on using Medicare, the average veteran has 40 unique VHA outpatient encounter days compared to 4.4 unique Medicare outpatient encounter days; this ratio is not surprising given that we are selecting only patients who visit a VHA emergency department. In terms of opioid prescriptions, including CMS prescriptions increases the total amount of opioids prescribed (measured in milligrams of morphine) in a given year by 27%.

The VA often pays vendors (providers and facilities) in the community to provide care for its veterans when they cannot themselves. Examples of eligible care include emergency care, nursing homes, inpatient hospice care, and childbirth. While the VHA pays or reimburses for the aforementioned care, veterans themselves file a reimbursement claim and we observe all medical and pharmaceutical claims that veterans file, regardless of final VHA approval or payment status. More recently, under the Veterans Access, Choice, and Accountability Act ("Choice Act") in 2014, non-VHA access expanded greatly for veterans who are unable to schedule an appointment or live too far from a VA medical clinic. For medical care covered

under the Choice Act, vendor claims are automatically updated in the VHA system nightly.[7]

Our last data sources provide us with date of death and cause of death for all veterans from VHA Vital Status files and Centers for Disease Control and Prevention (CDC) National Death Index (NDI) Plus files. The former are sourced from the Social Security Administration Death Master File, Medicare Vital Status File, and interval VA records, which include hospital, family, and National Cemetery Administration records and provide dates of death. The CDC NDI Plus files include detailed information about cause of death, allowing us to differentiate between opioid drug overdose deaths from deaths by heart failure, among others, along with death by specific type of opioid (e.g., natural and semi-synthetic versus heroin, versus synthetic opioids). We observe all deaths regardless of place of death, place of treatment, or cause of death.

In sum, we observe all veteran medical care and pharmaceutical prescriptions that occur at VHA-affiliated facilities, Medicare and Medicaid claims for a subset of the years, and any care received outside the VHA that veterans choose to file a reimbursement claim for. We also observe lab results, clinician notes, patient assessment questionnaires conducted at VHA-affiliated facilities. We will miss care covered policies other than VHA, Medicare, and Medicaid (e.g., private coverage), care that veterans choose to not file for reimbursement, uncompensated care, and of course, undiagnosed conditions and illicit use of drugs.

2.2 Sample Construction

Our research design focuses on adult veterans who visit a VHA emergency department. We observe approximately 20 million such visits. Emergency department encounters include both visits leading to inpatient hospital admissions and same-day discharge visits.

We make several restrictions to the sample of all emergency visits to improve power. First, we drop encounters with diagnoses that are prescribed less than 10% of the time because these visits are not impacted by the instrument (common diagnoses dropped include heart attacks and mental health episodes). This cuts our sample to approximately 7.1 million visits.

[7]Specifically, the data comes from Fee Basis Claims System (FBCS) and Program Integrity Tool (PIT).

We also exclude visits involving patients who are already very heavy prescription opioid users prior to the ED visit. Specifically, visits by patients who had over 3,150 milligrams of morphine equivalent in the prior year are excluded (this is equivalent to 5 weeks of 90 daily MME, approximately the 84th percentile of prior year opioid usage). These visits reduce power because an additional opioid prescription will have little to no effect on the heaviest prior users and some may even be seeking opioids by coming to an ED. After removing visits with non-prescribed diagnoses and high prior opioid users, we are left with approximately 5.9 million visits.

It is a consensus in the medical community that opioids are generally appropriate for end-of-life patients. Therefore, we exclude emergency visits in which the patient has terminal cancer or is on end-of-life hospice care on the date of ED arrival. Next, we further restrict our sample to encounters involving physicians who treat over 200 ED cases per year. Finally, we take the first visit per veteran that satisfies the above criteria. Our baseline sample consists of 1,958,209 emergency visits (and veterans) treated by 5,313 physicians from 2006-2016.

2.3 Variable Definitions

Our independent variable is $Prescribed_i$, an indicator for whether patient i is prescribed *any* prescription opioid for their emergency encounter. These are opioid prescriptions that the patient fills at a VHA pharmacy, and not opioids administered during the emergency department visit. We are able to locate and assign prescriptions to the emergency provider using a combination of the prescriber identifier, location of origin, and date the prescription was written (see Appendix B for exact details). Note that the patient could potentially not fill the prescription[8] or not consume the drugs. Our main outcomes can be categorized into opioid use, opioid-seeking behavior, adverse opioid-related outcomes, health and economic outcomes, and proxies for illicit opioid use. With the exception of a reduced form event study analysis in which we construct opioid use dummy variables at the month level (for 12 months prior and 24 months after the ED visit), all outcome variables are aggregated at the year

[8]We observe that approximately 3% of all VHA opioid prescriptions written by clinicians go unfilled.

level, relative to the date of the ED visit (i.e., one year post-ED, two years post-ED, etc.).

Opioid use and opioid drug screens

We construct measures of long-term prescription opioid use using filled pharmacy prescriptions, excluding the initial emergency prescription and any prescriptions within the first seven days after the ED visit. Specifically, we follow the literature and define long-term use as 180 days supply of opioids filled in the first 12 months after the ED encounter, excluding the first seven days (Barnett et al. (2017); Barnett et al. (2019); Jena et al. (2016); Dunn et al. (2010); and Braden et al. (2010) use 90 days supply instead of 180). This implies that typically a long-term user of prescription opioids needs to fill 180 days supply of opioids from another physician who is not their initial ED physician. We also study positive opioid drug screens (urine or blood lab tests), conditional on being screened, as a proxy for opioid use not captured by pharmacy prescriptions. These opioid drug screens test for any opioid use (i.e., metabolite of morphine) and generally cannot differentiate between prescription and illicit opioids (Moeller et al., 2008).

Opioid-seeking behavior and pain scores

To distinguish between medically appropriate long-term use versus inappropriate opioid misuse, we construct three proxies for opioid-seeking behavior: i) another prescription is filled when more than 25% of a previous prescription's days remain ("overlapping prescriptions"),[9] ii) prescriptions are filled at three or more pharmacies over a 90-day period ("pharmacy shopping"), and iii) patient visits total five or more encounter days with back problems or headaches and migraines in one year ("repeated back pain and headaches"). The first two proxies are common in the opioid literature (for example, in Yang et al. (2015) and Finkelstein, Gentzkow, and Williams (2018) and have been found to be strong predictors of opioid overdose, while the last measure is based on conversations with VHA physicians. We

[9]For example, if an opioid prescription with 10 days supply is filled on January 1 and a new prescription is filled on January 7, this is considered an overlapping prescription since three days remain (30%) on the previous prescription.

also construct an average annual self-reported pain score across all clinical settings. Pain scores are a frequently recorded "fifth vital sign" at the VHA (Stephenson, 1999) that fall on a discrete scale of 0-10. Since they are self-reported and unobserved by the physician, the score can be exaggerated by the patient to easier obtain opioid prescriptions.

Opioid use disorder, overdose, and mortality

We study two clinical diagnoses directly related to opioid misuse and abuse: opioid use disorder (OUD) and opioid overdose. OUD, often referred to as opioid dependence and abuse, is defined in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) as a "problematic pattern of opioid use leading to clinically significant impairment or distress." Disjoint from OUDs, opioid overdoses are episodes when central nervous system opioid receptors are blocked, causing breathing to slow down and the brain and heart to be starved of oxygen. This can lead to unconsciousness, coma, and then death. The vast majority of opioid overdoses are treated in time and are hence nonfatal. However, some can lead to death, which we construct from the NPI Plus files using cause of death data. We follow CDC definitions for opioid overdose mortality,[10] and also investigate *specific* type of opioid mortality: heroin; synthetic (excluding methadone; e.g., fentanyl and tramadol); and natural and semi-synthetic opioids (e.g., oxycodone and hydrocodone).[11] For OUD and overdose outcomes, we construct one- and three-year measures, and mortality is a single indicator of eventually dying of an opioid overdose. Virtually *all* opioid overdose mortality events are observed, regardless of whether the patient died in a VHA hospital or overdosed on prescription opioids or illicit heroin. Both OUD and opioid overdoses are constructed from VHA, CMS, and some non-VHA medical diagnosis codes,[12] meaning veterans who are not diagnosed by a VHA clinician for their OUD, or are treated for overdoses at non-VHA

[10]Opioid overdose cause of death is constructed using ICD-10 mortality codes in which the entity axis is one of X40-44, X60-64, X85, or Y10-Y14, and at least one of the record axis conditions lists: T40.0, T40.1, T40.2, T40.3, T40.4, T40.6

[11]ICD-10 mortality codes for heroin (T40.1), synthetic opioids (T40.4), and natural and semisynthetic opioids (only T40.2 and no other opioids detected).

[12]ICD-9 codes for OUD: 304.0x, 304.5x, and 304.7x. ICD-9 codes for opioid overdoses: 965.x, E850.0-E850.2, E935.0-E935.2, and E980.0.

hospitals or at home, and do not file for VHA reimbursement are not captured in our data. There are reasons to believe that a large fraction of OUDs and overdoses are not captured in our data. Some patients may be treated with naloxone and never see a medical provider. Furthermore, only 5.3% of all veterans we observe dying from an opioid overdose had a recorded opioid overdose in the three years leading up to their death.

Health and non-opioid outcomes

We also construct indicators for health and economic outcomes. First, we study two different mental health outcomes: attempted suicide and self-harm (excluding suicide ideation), and major depressive disorder. Both are constructed from medical diagnoses across both inpatient and outpatient settings, using VHA and CMS claims when available. Additionally, we have depression assessment results from two different questionnaire surveys (PHQ-2 and PHQ-9) and response scores above the recommended thresholds are coded as scoring positive. Second, we construct indicators for accidental falls as a proxy for impulsivity or sedation and are predictive of opioid overdose risk (Elizabeth M. Oliva and Trafton, 2017). Finally, we proxy for homelessness in a given year from a variety of sources, including diagnosis codes for lack of housing/inadequate housing and outreach to VA homeless and/or shelter programs.[13]

Illicit drug use and other substance use

In addition to heroin and synthetic opioid deaths, we investigate transition into illicit drug use by constructing two proxies. The first is an indicator for a physician's intent to screen for heroin/fentanyl. Unfortunately, it is difficult to distinguish heroin and fentanyl from prescription opioids and morphine. There is a specific metabolite unique to heroin and fentanyl called 6-monoacetylmorphine (6-MAM), which has a short half-life of 36 minutes and is detectable in urine for only up to eight hours after heroin use (Moeller et al., 2008). Furthermore, physicians often do not know the distinction between drug screens for heroin

[13]We follow VHA homelessness definitions from Peterson et al. (2015), which uses diagnosis codes, VA homeless service codes (Department of Housing and Urban Development VA Shared Housing, telephone and health care for homeless veterans, community outreach to homeless veterans), and VA inpatient treatment such as domiciliary care for homeless veterans.

and standard morphine (Starrels et al., 2012). Therefore, we code any completed drug test that mentions heroin, fentanyl or 6-MAM in the order form, regardless of test result, as an intent to screen for illicit opioids. Our second proxy for illicit drug use is a hepatitis C diagnosis. Hepatitis C is an infection that is commonly transmitted by sharing needles. The opioid epidemic has contributed to the rise in hepatitis C infections Powell et al. (2019); Zibbell et al. (2018), and the CDC (2016b) identifies injection drug use as the main risk in over half of new reported hepatitis C cases. It is estimated that 32% of injection drug users are diagnosed with hepatitis C within one year of injection and 53% within five years (Hagan et al. (2008); see Degenhardt et al. (2017) for overview). Unlike intended heroin/fentanyl drug screens, hepatitis C proxies for any injection drug use, not necessarily heroin. However, this proxy will mainly be measuring heroin use since it is the most preferred substance among injection drug users (Novak and Kral, 2011) and roughly 80% of injection drug users injected heroin in the previous year (Al-Tayyib et al., 2017). For both proxies, we observe only patients who are tested. All patients who do not take the test or are not diagnosed with hepatitis C (in the second proxy) are coded as zero. Finally, we study questionnaire results from Brief Addiction Monitor (BAM) which asks whether a veteran has used marijuana, sedatives, cocaine/crack, other stimulants, and opiates in the past 30 days, and the AUDIT-C which screens for potentially hazardous alcohol drinkers.

2.4 Summary Statistics

Table 1 describes relevant statistics of the emergency opioid prescription, emergency visit encounters, patient veteran demographics, and medical morbidities for our baseline sample. In 26.1% of emergency encounters, physicians prescibe an opioid. This number is mechanically high because we are excluding ED visits when the patient is diagnosed with conditions that are rarely prescribed; across all emergency visits, approximately 12% are prescribed opioids in the VHA, comparable to the 14.3% national average in 2011 (NCHS, 2016). There is also substantial variation across years in the opioid prescription rate, as seen in Figure F.1. The prescription rate is as high as 28.5% in 2011 and as low as 22.9% in 2016, for our

baseline sample. The average ED prescription is for 8.2 days with 21.0 milligrams of morphine equivalent per day. Figure 1 displays the empirical CDF of days supply and daily milligrams of morphine. While the prescription length has declined over time, the intensity in terms of daily milligrams of morphine has remained unchanged. Hydrocodone, Tramadol, and Oxycodone are the three most common opioids. Table F.1 lists the top 15 most common opioid prescriptions with dosages.

The emergency department is a common point of contact for patients with pain. Accordingly, musculoskeletal and connective tissue, and injury and poisoning are the two most common major diagnosis categories in our sample, making up over 40% of all ED visits. Frequent diagnoses in these two categories include arthritis, back or knee pain, fractures, sprains and strains, and dislocations. Figure 2 shows the top 15 most common major diagnosis categories and the share that are prescribed opioids in the ED, again for the baseline sample. See Figure F.3 for the most common major diagnosis categories for all emergency visits.

The average veteran is a middle-aged white male (average age 55; 90% are male and 69% are white). This demographic group has been hit the hardest by the opioid epidemic (Case and Deaton, 2015, 2017; Scholl et al., 2019). We have substantial variation in age and race in our sample, allowing us to test for heterogeneity. The average veteran is plagued with problems such as mental health and homelessness, with 24% and 13% diagnosed with medical depression and post-traumatic stress disorder in the prior year, and 6.2% homeless at some point the last year. A total of 27% used an opioid in the previous year and 10% in the previous month. Rates of depression and PTSD among these veterans are much higher than in the adult American population,[14] and prior-year opioid use is double that of the average emergency department patient in a privately insured Optum Clinformatics Data Mart sample. While one should be cautious about generalizing our veteran sample to the average American,[15] studies have shown that mental health history and prior opioid

[14]Based on the NSDUH 2017 and the 2005 National Comorbidity Survey-Replication, 7% and 3% of adult Americans, respectively, were diagnosed with depression and PTSD in a given year.

[15]Cesur et al. (2019) argue that "war injury-induced chronic pain, lax monitoring of opioid prescriptions by VHA providers, combat-related psychological trauma, and exposure to cheap opium supplies during war deployments have placed post-9/11 combat veterans at substantial risk for opioid abuse and mortality."

use are strong predictors of opioid abuse (Boscarino et al., 2010; Edlund et al., 2007, 2014; Seal et al., 2012)—suggesting that veterans are more representative of the population that could be potentially vulnerable to spiraling into dependence from a single prescription.[16] In subsection 4.2, we investigate our research question with veterans who never saw combat, a group that is plausibly more representative of the general population.

3. Empirical Strategy

Our empirical strategy closely follows the literature that exploits quasi-random assignment of agents to cases—what is commonly known as the "judges design." Papers in this literature estimate the causal effect of different types of sentencing outcomes on defendants' outcomes, exploiting variation in the leniency of the assigned judge. In the case of medical practitioners, the strategy rests on the empirical observation that there is wide systematic variation in practice behavior among physicians. Moreover, in the emergency department, assignment of patients to physicians conditional on hospital and shift is plausibly random. The remainder of this section provides an overview of physician assignment in EDs, presents the main regression model, describes the construction of the physician leniency instrument, and discusses the validity of the instrument.

3.1 Physician Assignment in Emergency Departments

The emergency department is a health care setting where patients have little to no choice regarding which provider they see, alleviating much of the physician shopping and patient selection issues in traditional outpatient settings. Leading up to ED visits, patients cannot look up reviews and schedule an appointment with their preferred physician as they can with primary care providers. Instead, when patients arrive in emergency departments—sometimes due to a health shock—they are assigned to an emergency physician based on a triage system:

[16] The 2018 National Survey on Drug Use and Health (NSDUH) supports this claim: 65.7% of surveyees with opioid substance use disorder had some mental illness in the prior year, 28.4% had major depression and 3.2% attempted suicide.

First they are seen by a triage nurse who assigns the patient an Emergency Severity Index (ESI), a five-point system. This index is based on health factors including life-threatening status, number of resources required, and vital signs and determines the patient's position in line. Next, when physicians are available, they treat the first patient in the queue and the next once they are done. In some cases, physicians may specialize in certain diagnosis conditions. To summarize, conditional on showing up to the *same ED* at the *same time* with the *same diagnosis*, physician assignment to patients is nearly as good as random.

3.2 Leniency Instrument Construction

Our physician leniency instrument is constructed as a residualized year-varying leave-out ("jackknife") instrumental variable (e.g., Doyle et al. (2015) and Dobbie et al. (2018)) based on estimation of the following regression equation:

$$Prescribed_{ikt} = \alpha_0 + \alpha_{hym} + \alpha_{hdt} + \alpha_{diagnosis} + \alpha_{agebins} + \gamma W_{ik} + \epsilon_{ikt} \quad (1)$$

Fixed effects include α_{hym}, hospital-year-month fixed effects, to control for time and seasonal variation in opioid prescribing such as hospital-specific policies and initiatives to limit prescribing or hospital specific seasonality in ED visits (e.g., trauma injuries are more likely in the winter). We also control for "shift-level" variations that include both physician scheduling and patient arrival (e.g., patients arriving to the ED on Friday nights differ from those patients coming in on Monday mornings) with hospital-day of week-time of day fixed effects, α_{hdt}.[17] Diagnosis fixed effects,[18] $\alpha_{diagnosis}$, are included to account for physician specialization with respect to diagnosis conditions. As mentioned earlier, these three set of controls are what is required for our quasi-random assignment assumption. Next, to improve statistical precision in our leniency measure, we include controls for five-year age bin, $\alpha_{agebins}$, and W_{ik}, which includes Elixhauser Comorbidity Index, pain score in the ED, and

[17] Day of week takes on seven values: Sunday, Monday, etc. and time of day are six mutually exclusive four-hour bins: 8am-12pm, 12pm-4pm, etc.

[18] We truncate ICD-9 diagnosis code to its three-digit value; all ICD-10 codes are crosswalked to ICD-9.

number of prior visits. Under the assumption that we have captured the observables under which quasi-random assignment occurs in the ED, the unexplained variation, the physician's contribution, resides in the error term, ϵ_{ikt}.

For patient i's, encounter k, treated by physician j in calendar year y, the instrumental variable equals:

$$Leniency^{phys}_{-i,jy} = \frac{1}{N_{-i,jy}} \sum_{i' \in \{\mathbb{J} \setminus i\}} \hat{e}_{i'k} \tag{2}$$

where $\hat{e}_{ik'} = \hat{Prescribed}_{ik'} - Prescribed_{ik'}$, the residual from Equation 1, $\mathbb{J}$ is the set of all ED encounters treated by physician j in year y, and $N_{-i,jy} = |\{\mathbb{J} \setminus i\}|$, the number of cases that physician has seen that year, excluding patient i.[19] This leave-out mean eliminates the mechanical bias that stems from patient i's own case entering into the instrument. This residualized measure is interpreted as the average (leave-out) prescription rate of patient i's physician, relative to other physicians in that hospital-year-month, hospital-day of week-time of day, controlling for patient age and diagnosis.[20]

Appendix C summarizes which physicians are more likely to be lenient prescribers by our measure. We find that lenient prescribers are slightly more likely to be older men. The appendix also correlates prescribing leniency with other dimensions that physicians can vary. For example, lenient prescribers are more likely to admit a patient to an inpatient hospital, perform more intensive care, and be lower in proxies for physician quality. In subsection 4.4, we examine how these physician characteristics affect our findings and interpretations. As a preview, we attempt to control for and address these physician differences and find our results are widely robust.

[19] This leave-out mean measure is algebraically equivalent to a physician fixed effect estimated in separate leave-out regressions (Dobbie et al., 2018).

[20] This version of constructing leniency is equivalent to the one of running OLS of (1, including physician fixed effects on the right-hand side, separately for every patient (and leaving out that patient's observation), using the estimated coefficient of the patient's physician's fixed effect as the leniency measure (Dobbie et al., 2017).

3.3 Empirical Specification

To study the effects of being prescribed an opioid in the ED on a particular outcome for patient i's ED visit (for example, in the first year after the ED encounter), we estimate the following model using our baseline sample of only first visits:

$$Y_i = \beta_1 Prescribed_i + \theta X_i + \epsilon_i \tag{3}$$

where individual-level baseline controls, X_i, include the Elixhauser Comorbidity Index (three-year lookback excluding the ED visit), measures of prior opioid use (dummy for prior month and log 1+ total milligrams of morphine equivalent for prior year), gender, and race. It also includes the set of fixed effects from Equation 1. If $Prescribed_i$ is randomly assigned, then the coefficient β_1 in Equation 3 represents the average causal effect of being prescribed an opioid for an ED visit on overall opioid use over the period of study. However, the $Prescribed_i$ variable suffers from endogeneity concerns. For example, injury severity may be unobserved and correlated with opioid prescription, which in turn also affects long-term prescription opioid use. To circumvent this endogeneity problem, we instrument $Prescribed_i$ with the assigned physician j's underlying propensity to prescribe opioids, $Leniency^{phys}_{-i,jy}$. We cluster heteroskedasticity-consistent standard errors at the physician level to account for the assignment process of patients to physicians.

3.4 Instrument Validity

Formally, under the assumptions of conditional independence, relevance, exclusion, and monotonicity, 2SLS, using a discrete, multi-valued z as an instrument for an endogenous, binary variable w estimates

$$\alpha^{IV} = \sum_{k=1}^{K} \lambda_k \alpha_{z_k, z_{k-1}},$$

where $\alpha_{z_k,z_{k-1}}$ is the LATE for the compliers from switching the instrument from value z_{k-1} to value z_k, and where the weights λ_k are non-negative and add up to one (Imbens and Angrist, 1994).[21] The IV estimator thus equals a weighted average of pair-wise local average treatment effects for successive values of the instrument.

We will discuss each of the four assumptions necessary for delivering the above interpretation of the estimator and provide evidence for the validity of these assumptions in our sample.

Relevance

The first assumption, the relevance criterion, requires that the instrument $Leniency_i$ has explanatory power with respect to the endogenous regressor $Prescribed_i$. Figure 3 graphs the histogram of the instrumental variable along the x-axis and the left y-axis. A local-linear regression of the fitted probability of prescribed opioids on the instrument after residualizing is overlaid and displayed on the right y-axis. The histogram displays wide variation in the prescription rate among physicians working at similar times, in the same hospital, treating identical diagnoses. The average physician in the 10th percentile of leniency prescribes opioids at a rate of 16.8%, while a physician in the 90th percentile prescribes at a rate of 41.9%. Table 2 presents the same local-linear graphical first stage in a regression table. The association between ED opioid prescription and physician leniency is linear and strong. Being assigned to a 10 percentage point more lenient physician (relative to the average in that hospital, year, month, patient diagnosis, etc.) is associated with a roughly 17 percentage point increase in the likelihood of being prescribed an opioid in the ED. The first-stage F-statistic is 25 when all controls and fixed effects are included; the instrument does not suffer from weak instrument concerns. The first-stage coefficient is greater than 1 because *all* emergency visits are used to construct the leniency instrument, however, the first stage is calculated using the baseline sample, which excludes the rarely prescribed diagnoses. The first-stage coefficient for the entire sample of emergency visits is exactly 1, as expected with

[21] The weights λ_k equal the following: $\lambda_k = \frac{(Pr(w=1|z_k - Pr(w=1|z_{k-1})\left[\sum_{l=k}^{K} Pr(z=z_l)(z_l - E[z])\right]}{\sum_{m=1}^{K}\left((Pr(w=1|z_m - Pr(w=1|z_{m-1})\left[\sum_{l=m}^{K} Pr(z=z_l)(z_l - E[z])\right]\right)}$

the large number of encounters used to construct the leniency measure.

Conditional Independence

Second, the conditional independence assumption requires that assignment of patients to ED physicians be random, conditional on seasonality, shift, and diagnosis. One can test for patient selection by checking whether patients with particular characteristics are more likely to be assigned lenient physicians. Table 3 shows evidence of balance on observables. The first column is a regression of *Prescribed* on observables including patient demographics, priority groups, income, previous medical diagnoses, and ED diagnoses. Unsurprisingly, many observables predict opioid receipt status. For example, a unit increase in the patient's self-reported pain score is associated with a 5.9pp increase in the probability of being prescribed. The second column displays the coefficients from a regression of physician leniency instrument as the dependent variable, on the same set of observables. Conditional independence implies that these covariates should not predict the leniency status of the physician the patient sees. Figure F.4 presents the same table in a graph. With so many coefficients, we fail to reject that the instrument violates conditional independence (the joint F-stat drops by two orders of magnitude, from 280 to 2.4). In fact, some of the variables that are significant in the second column go in the opposite direction of the decision to prescribe (e.g., pain score is now negatively correlated with leniency).

The residualization in Equation 1 controls for more patient observables than required to achieve quasi-random assignment; they are included for statistical precision in measuring physician leniency (ie. age, Elixhauser, pain score, prior ED visits are controls for precision on top of the required controls for quasi-random assignment: hospital-year-month, hospital-day of week-time of day, and diagnosis). We test for quasi-random assignment without the additional controls. Figure F.5 displays the outcome of the balance test when only seasonality and shift included in the residualization of Equation 1 in the left panel. The right panel residualizes for diagnosis as well as seasonality and shift. When controlling only for seasonality and shift, there is some selection along observables present; however, once controlling for

diagnosis, the coefficients are nearly balanced. This lends support to our assumption that assignment in the ED is nearly random at the seasonality, shift, and diagnosis level.[22]

An alternate test asks whether *predicted* opioid receipt is correlated with physician leniency. Formally, we project $Prescribed_i$ on the observables and regress predicted $\hat{Prescribed_i}$ on the instrument. We find that the variation spanned by observables and previous medical comorbidities that explain whether a patient is prescribed opioids, is not correlated with our instrument, again failing to reject the conditional independence assumption.[23] We are not able to test conditional independence with respect to unobservables. As an alternate quasi-random research design, we leverage variation in the team of physicians working in the ED when a patient arrives in subsection 4.4. This specification no longer relies on within-shift random assignment.

Exclusion

With the conditional independence assumption, the reduced form identifies the causal effect of being treated by a *lenient physician*. However, our estimand of interest is the causal effect of being prescribed *an opioid*. This requires that the instrument influence the outcome of interest (such as health outcomes or subsequent opioid use) only through its effect on initial opioid consumption. This assumption is at its core untestable, but there are reasons to believe any violations are likely to be small. First, unlike in primary care settings, where the patient and primary care provider have many repeat encounters, the scope of what the emergency physician can do to impact long-term outcomes is limited. Second, any violation of the exclusion restriction needs to directly affect the specific outcome of interest. The channel by which ED physicians can influence opioid-related outcomes is likely through

[22] Looking ahead, the various levels of residualization have some effects on our baseline 2SLS causal estimates. Table F.3 presents the effects of a single opioid prescription on our main outcomes with instrumental variables that are constructed with varying levels of residualization. The main estimates are attenuated when the leniency construction only residualizes for seasonality and shift; however, when diagnosis is included in the residualizing, the 2SLS estimates are very similar to the baseline, while standard errors are slightly larger.

[23] The regression coefficient from this exercise is actually negative and economically insignificant (Table F.2. The average physician in the 90th percentile of leniency on average treats patients who are 0.1pp less likely to be prescribed an opioid (predicted by its observables) than the average physician in the 10th percentile of leniency.

opioid prescribing. Nevertheless, we model and account for additional physician practices to address exclusion restriction concerns in subsection 4.4. We augment our baseline model by controlling for admission and procedures as endogenous decisions with corresponding propensities as instruments. In another specification, we include a proxy for physician quality in the control to our baseline estimating equations. As a placebo test, we study the same outcomes for patients with diagnoses that are never prescribed opioids. We should expect to see no difference in the reduced form of leniency on opioid outcomes (e.g., long-term opioid use) for never-prescribed conditions like schizophrenia.

Monotonicity

Finally, the monotonicity assumption is necessary for interpreting the coefficient estimates obtained from the IV approach as local average treatment effects if there are heterogeneous treatment effects. It requires that any patient who is (not) prescribed opioids by a strict (lenient) physician, would also (not) be prescribed opioids by a less strict (lenient) physician. The literature leveraging the judges design typically perform two informal tests for the implications of the monotonicity assumption. The first one provides that the first stage should be weakly positive for all subsamples (Dobbie et al., 2018); there cannot be a group of patients for which the decision to prescribe an opioid cannot be partly explained by their physicians' underlying average propensity to prescribe. To test this, we run the first-stage regression separately by gender, age, education, and diagnosis groups, among others. The second implication asserts that the instrument constructed by leaving out a particular subsample has predictive power over the left-out subsample (Bhuller et al., 2019). If a physician's propensity to prescribe opioids for men has no power in predicting whether women are prescribed, then this would violate this test. Frandsen et al. (2019) provide formal motivation for these tests.[24] Table 4 presents both of these tests in the two columns for various subsamples of interest. Both columns exhibit significant and positive first-stage

[24] Frandsen et al. (2019) propose a weaker "average monotonicity" assumption which prescribes that as long as each individual patient complies with monotonicity for a sufficient amount of judges, the 2SLS estimand can still be interpreted as a well-defined weighted average. This assumption lends justification for the tests performed in Dobbie et al. (2018) and Bhuller et al. (2019).

coefficients, implying that for each subgroup, both the physician's tendency to prescribe to all patients, and to patients outside of that subgroup have strong predictive power on their opioid status. The coefficient magnitudes differ across each subgroup row because rates of opioid prescription differ across various subgroups. Finally, we check whether our main results hold using differential, mutually exclusive leniency measures (i.e., by major diagnosis category) in subsection 4.4.

4. Results

In this section we present our main findings: i) a single emergency department opioid prescription can cause veterans to become long-term users; ii) some veterans develop opioid use disorders, abuse and misuse opioids, and exhibit behavior consistent with opioid-seeking; and iii) veterans are more likely to suffer adverse health outcomes including opioid overdoses and opioid overdose mortality. We then study heterogeneous effects and investigate the channel through which patients receive opioids. Several robustness checks follow, including an alternate strategy leveraging the team of physicians working on shift, and modeling of endogenous admission and procedures follow. Finally, we conduct a back-of-the-envelope exercise to estimate total opioid use and misuse, accounting for unobserved markets.

4.1 Main Results

Long-term prescription opioid use

We begin by studying continued use of prescription opioids following the emergency prescription. Figure 4 plots the reduced form event-study analysis of new monthly opioid prescriptions filled for patients who see lenient and strict physicians. Physicians are classified as lenient (strict) if they fall into the top (bottom) quintile of the physician leniency measure. We plot average monthly (relative to ED encounter date) residualized prescription opioid use indicators for the two groups of patients. The ED prescription and any prescriptions filled in the seven days after the visit are excluded; only new filled prescriptions count toward each

30-day period data point. On average, veterans are getting sicker, experiencing more pain, and hence prescription opioid use rises in the months approaching up to their ED encounter.[25] The lenient and strict groups are on parallel paths of opioid use prior to the visit. However, after the visit, there is an increase in the probability of filling an opioid prescription for patients treated in the ED by lenient physicians relative to their strict counterparts.

This level shift in opioid use probability following the ED is roughly a quarter of a percentage point (average difference between the two groups in months 2-24) and can be attributed to a 27.4% (42.3% - 14.9%) difference in opioid prescription rates between the lenient and strict prescribers. To estimate the causal effect of the initial ED opioid, we run separate monthly regressions of Equation 3 with opioid use dummy as the outcome, instrumenting for $Prescribed_i$ with $Leniency_i$, now utilizing the continuous instrument. The results are graphed in Figure 5. The causal effect of an initial ED opioid on opioid use in subsequent months is highest for the first month at 5%, and settles in at around 1.5% at around the six-month period, when the average opioid use rate is roughly 9.5%.

We summarize continued opioid use with a single measure of long-term use as defined in subsection 2.3: 180 days supply in the first year. Table 5 presents the OLS and 2SLS regression outputs with long-term use as the dependent variable. The naïve OLS shows that on average, an opioid prescription in the ED is associated with a 2.85 pp increase in the probability of long-term use, on a base of 5.8%. This number is relatively robust to controlling for baseline characteristics such as patient demographics and the Elixhauser Comorbidity Index. However, with the IV model, the effect is more than halved to 1.17 pp in our preferred specification, a 20% increase on the overall average long-term use rate. The upward bias in the OLS is expected if, for example, patients with more severe conditions (and severity is unobserved) are more likely to be prescribed opioids and if severe conditions require more long-term opioid use to treat. The remaining rows of Table 5 present the total milligrams of

[25] This rise in opioid use prior to the ED visit, and the findings on subsequent use are *not* explained by patients shopping for opioids. In Figure F.6 and Figure F.7, the same reduced form event-study figure is displayed for veterans who are opioid-naïve and veterans who never visited an ED in the prior year, but did have some other VHA outpatient encounter. These two groups of patients are unlikely to be ED shopping for opioids.

morphine equivalent filled by the patient. Reassuringly, there are no differences in pre-period total morphine, consistent with the reduced form plot. The effect of a single opioid in the emergency department increases first-year morphine equivalent filled by 202 mg and 360 mg over two years, excluding the initial ED prescription. Including the initial ED prescription, this is equivalent to roughly 35 tablets of Oxycontin 10 mg over two years—approximately a 30% increase in total *observed* milligrams of morphine filled in the two-year period after an average emergency visit (20% increase conditional on being prescribed). The analyses and research design control for diagnosis condition, implying that the increase in subsequent opioid use is because a veteran saw a more lenient prescriber at the ED, and not due to the condition itself, provided that the conditional independence assumption is satisfied. The estimate of 1.17pp is most likely an underestimate because we do not observe any opioids filled through private coverage or obtained through other means (family and friends, black market, etc.). This underestimate is likely more pronounced for opioid abusers; Lin et al. (2019) find that among all opioid overdose decedents in 2016, only 26% had a VHA opioid receipt in the three months prior to their death.

The bottom panel of Table 5 reports the regression coefficient for positive urine drug screen for opioids for the subsample of veterans who are tested. Veterans who are prescribed opioids in the ED are 2.0pp (8%) more likely to test positive for opioids, conditional on receiving a urine drug screen even in the third year (months 25-36) following their ED visit.

The effects on subsequent opioid use originating from a single opioid prescription in the ED are large and go against current medical guidelines.[26] Both the CDC (Dowell et al., 2016) and the CMS (2019) have established recommendations and guidelines to limit opioid prescriptions to three days or fewer for acute pain and to encourage or to require physicians to consider non-opioid therapy first. There are few reasons for a single (non-cancer, non-end-of-life) transitory health shock to cause a sustained increase in the probability of opioid use for 24 months, especially if the increase is because the patient was treated by a different physician.

[26]Converting Barnett et al. (2017) and Barnett et al. (2019) to 2SLS estimates, they find a single opioid prescription increases probability of long-term use by 2.0pp and 0.9pp, respectively. However, their samples are different from ours. See Appendix A for a detailed comparison of the methods and how they affect estimates of long-term use.

In a large randomized controlled trial, Krebs et al. (2018) find that opioid treatment is no more effective than acetaminophen or nonsteroidal anti-inflammatory drugs (NSAIDS) at treating chronic back pain or knee or hip osteoarthritis pain. The same is true for acute lower back pain (Friedman et al., 2015), kidney stone pain (Teichman, 2004; Holdgate and Pollock, 2004), and minor fractures (Dodwell et al., 2010).

Opioid-seeking behavior and opioid misuse

A single opioid prescription from the ED can lead to long-term prescription opioid use. However, it is unclear whether this long-term use is through appropriate medical care or active opioid seeking, suggestive of inappropriate use and patient misuse.

As mentioned earlier, we study three proxies for opioid-seeking behavior and opioid misuse along with average self-reported pain score. The regression outcomes in Table 6 indicate that an opioid prescription in the ED increases the one-year post-ED likelihood of overlapping prescriptions by 1.9pp on a base of 9.9% and pharmacy shopping by 0.3pp on a base of 0.6%. Veterans are 0.55pp more likely to be seen by a clinician for back pain, headaches, and migraines at least fives times in the first year following the ED visit, on a base of 6.2%. In this regression we exclude patients for whom the ED diagnoses were back pain and headaches to avoid the concern that increases in this proxy could be due to their condition being treated poorly in the ED. Lastly, veterans' average self-reported pain score increases by 0.08 on a base of 2.81 (0-10 point scale). There is no significant difference in any of the four measures in the year prior to the emergency encounter.

The unanimous increase across all four measures provides evidence that an ED prescription increases not only long-term opioid use, but also unnecessary prescriptions, inappropriate opioid misuse, and potential abuse. The increase in overlapping prescriptions suggests that the long-term use results cannot be explained by appropriate medical care for the emergency condition. Furthermore, veterans exhibit opioid-seeking behavior that is consistent with patients having excess demand for opioids. The findings on back pain and headaches, and pain scores especially point in this direction as they go against standard medical intuition:

Being prescribed opioids should decrease experienced pain, rather than increase it. These two proxies, along with the uptick in pharmacy shopping, imply that some veterans attempt to game the VHA system to acquire opioids.

Adverse opioid outcomes: OUD, opioid overdose, and mortality

Thus far, we've shown that some of the long-term opioid use induced by an ED prescription is inconsistent with appropriate medical care and can be explained by opioid-seeking behavior, suggesting some veterans are misusing opioids. Next, we investigate whether this misuse can lead to severe adverse outcomes for the veteran. To shed light on this question, we study three opioid-related outcomes: opioid use disorders (commonly referred to as opioid dependence and abuse; defined as a "problematic pattern of opioid use leading to clinically significant impairment or distress" (DSM-5)), opioid overdoses, and opioid overdose mortality.

Table 7 reports the OLS and 2SLS regression output for these three outcomes. For both OUD and opioid overdoses, there is no significant difference in the prior year for patients who are prescribed in the ED because of the leniency of their physician. This placebo pre-trend test further lends support to our research design. In the year following the ED prescription, patients who are prescribed an opioid because of the leniency of their physician are 0.26pp more likely to develop an OUD, on a base of 1.98%. Within three years of the ED visit, 0.335pp of veterans will develop an OUD because of an ED opioid prescription. This effect is on a base of 3.27%, a 10% increase in development of an OUD within three years.

One of the most severe outcomes associated with patients with OUDs is an opioid overdose: an episode when opioid receptors affect the ability to breathe, which can lead to unconsciousness, coma, or even death. An ED prescription increases the one-year probability of overdosing on opioids by 0.88pp, on a base of 0.23%. The coefficients of three-year probability are positive and not significant, and we cannot reject the null that the three-year effect is equal to the effect on one-year overdose probability.

Fortunately, the vast majority of opioid overdoses are nonfatal. By far the worst outcome for an opioid user or abuser is an opioid overdose death. The last row in Table 7 presents the

results of a regression with opioid overdose mortality as the dependent variable. There is a positive effect effect of 0.075pp on opioid overdose mortality caused by a single ED opioid prescription, on a base of 0.167%. Conditional on a fatal opioid overdose, the median time from the ED prescription to death in our sample is is three years and five months.

Table 8 distinguishes between the type of opioid in opioid deaths, specifically, heroin, synthetic, and natural/semi-synthetic opioids. Although statistical power suffers with such mortality splits, some overall patterns arise. A single (non-heroin, non-fentanyl) opioid prescription can cause illicit opioid deaths. In fact, a substantial fraction (30-40%) of the mortality effect stemming from a single prescription is accounted by heroin and synthetic opioid related mortality.

The effects of a single opioid prescription on adverse opioid-related outcomes are staggeringly large. The emergency prescription causes a 0.335pp increase in the probability of developing an OUD within three years, more than a quarter of the effect size on long-term use. This means that on average, a quarter of recorded long-term users of prescription opioids that are induced by a more lenient ED prescriber, develop (and are *diagnosed* with) a pattern of opioid use that is characterized by "clinically significant impairment or distress." A sizable fraction of veterans will also overdose on opioids. For every 100,000 veterans who are prescribed an opioid in the emergency department due to being quasi-randomly assigned a more lenient prescriber, approximately 335 will develop an OUD, and 75 will die from an opioid overdose. These results are evidence that even a single opioid prescription in a one-time emergency setting can have strong and long-lasting effects on the patient's outcomes.

Across all three outcomes, the OLS estimates are biased downward relative to the 2SLS estimates. For instance, the average veteran who is prescribed an opioid in the ED is no more likely to develop an OUD or overdose on opioids. One potential explanation for this correlation is that on average, physicians are good at spotting patients who are at obvious risk of opioid overdose, and do not prescribe them opioids. The coefficients of prior-year OUD and overdose, along with the negative correlation between prior mental health and homelessness and being prescribed in Table 3, support this claim. Physicians are more likely to prescribe a

veteran an opioid, on average, if the patient did not have an OUD, overdose on opioids, or have mental health issues in the past. In addition to the downward OLS bias, the 2SLS effect sizes are large. For example, the causal effect on opioid overdose mortality is 0.075pp on a base of 0.167%, an effect size of approximately 45%. Of the veterans who visit an ED—with diagnosis conditions that fall into our sample restriction—and eventually die from an opioid overdose, roughly 45% of opioid overdose deaths were caused by the ED prescription. Stated differently, in general, veterans are not at risk of a fatal opioid overdose unless they receive an ED prescription due to their physician's leniency. One potential interpretation of the large effect size is that that LATE compliers affected by the instrument are veterans with particularly high overdose mortality risk that physicians take different approaches to treating. The causal effects on mortality are also large relative to the effects on recorded OUD and overdose diagnoses. For instance, the coefficient magnitudes imply that an ED prescription causes more fatal opioid overdoses than opioid overdoses in total. As mentioned earlier, we observe all mortality, while many opioid-related events go unreported because veterans seen at non-VHA hospitals or treated with naloxone by a non-medical provider are not recorded. Therefore, we perform a back-of-the-envelope calculation in subsection 4.5 to account for unobserved long-term opioid use and misuse.

Health and non-opioid outcomes

In addition to the chain of opioid-related outcomes, the emergency opioid prescription could have downstream effects on a variety of health outcomes. For instance, Scherrer et al. (2016) find that long-term use of opioids can lead to new onset of depression. Along with depression, we also study attempted suicide and self-harm, accidental falls, and homelessness.

Table 9 reports the regression output. Consistent with the other prior-year characteristics, the veterans with prior health morbidities are less likely to be prescribed opioids. In general, veterans who receive opioids were in better health prior to the ED visit. Opioids increase the probability of accidental falls—a proxy for sedation—within three years by 0.5pp. The preferred 2SLS coefficient for depression is positive but not significant. We are able to rule

out effects on depression diagnosis larger than 1.35pp, on a base of 35.5%, at the 95% level. Patient questionnaire results from two depression assessment screens (PHQ-2 and PHQ-9) are reported in Table 10. Veterans are 0.57pp more likely to screen high on depression questionnaires. From diagnosis and questionnaire results we can conclude that there may be some negative impact of the ED opioid prescription on depression. The coefficients for attempted suicide and homeless episodes are not statistically significant from zero, or the prior-year effect. The coefficient for homelessness in the prior year is positively and weakly significant, presumably due to sampling variation given the balance of other prior morbidities and patient characteristics.

Transition into other substances and illicit drugs

The majority of veterans who die from an opioid overdose do not fill a prescription for prescription opioids in the prior year (Lin et al., 2019) and around 80% of heroin users say they began by abusing prescription opioids (Muhuri et al., 2013). Alpert et al. (2019) and Evans et al. (2019) find that state-level exposure to Oxycontin sparked increases in heroin-related overdose deaths after Oxycontin's reformation to become abuse-deterrent. These facts, along with our estimates on heroin and synthetic opioid mortality suggest that some veterans are transitioning into illicit opioids or illicit drugs in general.

By analyzing the Brief Addiction Monitor (BAM) questionnaire responses of roughly 31,000 veterans in the three years after the initial ED encounter, we can paint a picture of the type and amount of drug use caused by a single prescription (Table 10). Self reported illicit/illegal or prescription drug abuse increases by 0.156pp, again roughly 20%. There is evidence of increased cocaine/crack use, suggesting that prescription opioids may be a gateway for other illicit drugs. Interestingly, use of marijuana or hazardous alcohol consumption do not increase, perhaps because these substances are more common and less stigmatized and thus unaffected by prescription opioids.

To research the effects on injection drug use, we report the output of regressions of two different proxies: i) physician intent to screen for heroin/fentanyl and ii) positive hepatitis

C diagnosis, on an indicator for any ED prescription in Table 11. The 2SLS coefficients for both proxies are positive, but statistically insignificant. This is not surprising given over 50% people living with hepatitis C are undiagnosed (CDC, 2016a), and this fraction is likely higher if they are opioid abusers or injection drug users. We fail to rule out effect sizes for intended heroin screens as large as 0.16pp, and 0.67pp for new hepatitis C diagnoses.

4.2 Heterogeneous Effects

The findings in the previous sections represent the local average treatment effects of an opioid on the complier population. However, these could mask vast heterogeneity in impacts across patients. In order to design policies that curb the number of new opioid dependents, understanding patient heterogeneity, i.e., *who* is vulnerable to opioid dependence due to a single opioid prescription, is crucial. We present 2SLS estimates of our main outcomes on subsamples constructed by splitting our baseline sample into mutually exclusive and exhaustive groups along several key margins: i) prior-year opioid use, ii) prescribing leniency of a patient's assigned primary care provider (PCP),[27] iii) year (pre/post 2011), and iv) gender. Prior opioid use is an important moderator to study because an analysis of the subsample of opioid-naïve patients informs us about the probability that a new user becomes opioid-dependent. The PCP leniency margin is motivated by the fact that patients require a consistent source of new prescriptions beyond the initial ED prescription. Knowing the prescribing leniency of the patient's assigned PCP allows us to study the heterogeneous effects of short-term prescribing shock (emergency physician) by their exposure to long-term prescribing risk and the potential substitution to black market or illicit opioids when their legitimate supply is restricted. After 2012, the VHA established various non-opioid strategies to manage pain and VHA opioid prescriptions dropped by 40% (GAO (2018)).

Table 12 presents the results of the subsample 2SLS regressions (with the preferred specification) along with p-values for the difference between the subsample coefficients.

[27]Leniency of PCPs are constructed using *all* 61 million outpatient cases for all veterans, not just our emergency cohort. It's constructed analogously to the ED prescribing leniency, residualizing for outpatient facility-year-month, outpatient facility-day of week-time of day, diagnosis, and age bin fixed effects.

Although some of the regression coefficients from the subsample regressions are not as precisely estimated due to loss in sample size, they exhibit some suggestive patterns. First, a single ED opioid prescription can induce long-term use of prescription opioids for both opioid-naïve and prior users. The effect sizes are also extremely similar. The effects on developing an OUD, opioid overdose mortality, and hepatitis C diagnosis are qualitatively smaller for opioid-naïve veterans; however, they are still large and not statistically distinguishable from the effects on prior users. For instance, the 2SLS coefficient for opioid overdose mortality is 0.057pp, (the baseline estimate is 0.075), meaning for every 100,000 veterans who are prescribed an ED opioid due to their physicians' leniency, 57 will die from an opioid overdose.

Splitting the sample along primary care provider leniency, we do not find heterogeneous effects that are statistically significant; however, the results indicate some suggestive patterns. We find that the effects on long-term use are greater for those with more *lenient* PCPs (1.3pp vs. 0.76pp). However, the effects on hepatitis C as a proxy for illicit drug use and opioid overdose mortality are larger in magnitude for veterans with *stricter* PCPs. This heterogeneity could suggest substitution between legal prescriptions and black market or illicit opioids. If a legal supply of prescription opioids via a patient's PCP is readily available, the patient chooses this source of opioids. However, if the patient's PCP does not provide ready access to opioids via prescriptions, the patient is more likely to turn to illicit opioids. These findings, although suggestive, highlight the complex interdependencies between different parts of the "supply chain" for opioids. They could also imply that reducing prescription opioid supply in one domain of medical care alone may not be sufficient, especially if patients are seeking opioids, as it will have spillover effects that extend to other sources of legitimate and illicit opioids.[28]

The causal estimates of long-term use are larger pre-2012, when the VHA had a greater supply of prescription opioids and prescriptions were easier to acquire. After 2012, the effects on hepatitis C diagnosis are larger and significant. The decreased supply of legal prescription

[28]PCP leniency is treated as an exogenous initial condition, but this may not necessarily be the case. For example, PCP leniency be correlated with ED encounters, or veterans may seek out lenient PCPs after receiving an ED prescription. We study heterogeneity based on PCP leniency restricting our sample to *opioid-naïve* veterans in Table F.4. The pattern is remarkably robust.

opioids and increased supply of illicit heroin and fentanyl after 2012 support the narrative regarding substitutions between legitimate/illicit opioids we found in the PCP heterogeneity. The smaller magnitudes of long-term use and OUD after 2012 suggest that the combined VHA efforts (including but not limited to: paradigm shift in pain care, increase in non-opioid therapy, academic detailing and educational outreach, opioid dashboards and risk monitoring, etc.) decreased observed and recorded long-term use and opioid dependence. However, they may have led to a rise in injection drug use. The coefficient for overdose mortality after 2012 are attenuated because our cause of death data ends in 2016.

The final heterogeneity margin suggests that the adverse effects of a single opioid prescription extend beyond only men. Finally, we also find that the magnitudes are greater in states with high opioid mortality. Our baseline estimates are also robust to veterans who never saw combat, lending some support for external validity (Table F.5).

4.3 Complier Analysis

We have identified a Local Average Treatment Effect for those patients for whom the assigned ED physician's leniency determined whether they received an opioid prescription. Patients with certain pre-existing conditions and demographics are at higher risk of developing opioid dependence (Ives et al., 2006; Zedler et al., 2014); therefore, it is important to characterize the compliers. We characterize the fraction and patient characteristics of the compliers based on moments in the first stage (Abadie, 2003; Dahl et al., 2014); see Appendix D for more details.

Approximately 39.4% of our baseline veteran sample—veterans who show up in the ED with conditions that are prescribed at least sometimes and are not heavy prior opioid users—are compliers. These patients would have been sent home without an opioid prescription if they saw the most strict physician. The share of compliers is large and it underscores the

wide variation and lack of consensus among physicians in prescribing opioids.[29] A potential explanation for our large share of opioid compliers is the relatively low *perceived* risk from physicians. For example, Hwang et al. (2016) find that many physicians underestimate the addictiveness of certain opioids; Doctor et al. (2018) find that physicians cut back on prescribing after learning about a patient overdose) relative to more obviously life-altering decisions that judges make. Patients can also influence physician decisions by demanding opioids, a channel that is at larger play in medical settings than in court settings. In comparison, 7.6% are always-takers, meaning their circumstance (medical condition or patient demand) would always result in an opioid from all physicians. Unsurprisingly, the majority of our baseline sample are never-takers, at 53%. These are patients for whom opioids would never be prescribed by any physician. For the universe of veteran emergency visits (including repeat visits, heavy prior users, all diagnosis conditions, etc.), the fractions of compliers, always-takers, and never-takers are 24%, 1.3%, and 75% respectively.

Table F.6 reports for each demographic subgroup, its unconditional share, its conditional probability given patients are compliers, and the relative likelihood. Compliers are roughly 4.6% more likely to be middle-aged (ages 40-60) compared to the baseline sample. Compliers are also 17% more likely to have musculoskeletal or connective tissue conditions. These conditions are commonly prescribed opioids, but can also be treated using various forms of non-opioid therapy including NSAIDs, physical therapy, yoga, etc. Patients with other conditions, including disorders of the digestive system, are significantly less likely to be compliers. Another interesting avenue to study is whether compliers are more or less likely to be at risk for severe opioid outcomes. We predict ex-ante risk of opioid overdose death using veteran characteristics and medical history.[30] Veteran compliers are 6.4% more (10.2% less) likely to be above (below) average risk for opioid overdose death prior to their ED visit.

[29] In other contexts; Dobbie et al. (2017) find 13% of consumer bankruptcy cases are compliers. The complier share for criminal defendants facing pre-trial detention is 13% in Dobbie et al. (2018). Approximately 14% of Norwegian criminals facing incarceration are compliers in Bhuller et al. (2019). For Norwegians applying for disability insurance, 25% are compliers (Dahl et al., 2014).

[30] We use opioid overdose deaths rather than overdoses because there is significant under-reporting of overdoses as mentioned in earlier sections. Risk is predicted using Least Absolute Shrinkage and Selection Operator (LASSO) for veteran observables, medical morbidities prior to ED (suicide, mental health, falls, etc.), prior opioid use, and prior opioid-seeking behavior.

In sum, compliers tend to be middle-aged veterans who show up with musculoskeletal or connective tissue conditions and are at higher risk for opioid abuse. The higher risk (complier) patients tend to be the ones for whom physicians differ in their treatment approach to pain management (i.e., opioid vs. non-opioid treatment), explaining why the 2SLS magnitudes for overdose and overdose mortality are larger than their OLS counterparts.

4.4 Alternate Specifications and Robustness Checks

In this section we discuss various threats to our identification strategy and our solutions. The concerns can be grouped into the three different identifying assumptions: conditional independence, exclusion restriction, and monotonicity. For each concern we present alternate specifications and show the sensitivity of our main outcome measures to that specification.

Addressing Threats to Conditional Independence

In Section 3.4 we showed the balance of patient observables with respect to physician prescribing leniency. However, there might be selection along unobservable margins. Non-random triaging may occur; for instance, triage nurses may assign severe patients to physicians based on physician expertise. If severity is unobserved and not accounted for, our estimates may be biased.

Our main remedy for this is to leverage variation in the composition of physicians working on a particular shift. On average, 2-3 physicians are working in a VHA emergency department at any time. Even if there is within-shift selection (e.g., experienced physicians getting more severe patients), patients are constrained by the physicians on shift at the time of arrival. We leverage this across-shift variation in the physicians working when a patient arrives in the ED and construct a team prescribing leniency instrument. The instrument is a weighted average, leave-shift-out of the individual instruments of the physicians working on shift; the details are shown in Appendix E. Observe that when there is only a single physician working, this team instrument is equivalent to the baseline physician instrument and as the number of physicians working increases, the team instrument approaches the average prescribing rate in

that ED, and loses power in predicting opioid prescription status.

The results of Equation 3, where we now instrument $Prescribed_i$ with team leniency, are displayed in column 2 of Table 13. The coefficients for long-term use, opioid overdose, overdose mortality, falls, and homelessness are very similar. The coefficient for hepatitis C is attenuated and much noisier. The causal estimates of the effects of a single opioid prescription are very similar whether the prescription receipt is induced via being assigned a lenient physician within a shift or arriving at a particular time when a team of lenient physicians are on shift. This robustness lends support to the claim that patient-physician selection is unlikely to be driving our findings.

Finally, we study patients visiting the ED for injuries and poisonings. Injuries like broken legs are plausibly unexpected random health shocks, and could be subject to fewer selection concerns. Moreover, long-term adverse outcomes caused by an ED prescription are even more alarming given the acuteness of injuries. The coefficients are reported in Table F.7 and the findings are qualitatively robust. The effects on OUD and overdose mortality death are attenuated, but the large standard errors do not rule out effect sizes as large as the baseline estimates. In addition to alleviating selection concerns, these findings reject the notion that patients only become opioid dependent from continued opioid prescriptions for chronic conditions like chronic back pain; a one-time opioid prescribed for acute injuries is not immune to the adverse effects caused by an opioid.

Addressing Threats to Exclusion Restriction

A potential threat to interpreting our findings as causal effects from opioids rather than from lenient physicians, is that physicians differ in many dimensions beyond just prescribing leniency. If these dimensions are correlated with leniency and affect our outcomes of interest, then we violate the exclusion restriction. While this is concern is slightly alleviated due to the short-term nature of emergency physician and patient relationships, emergency physicians may still make endogenous decisions that can impact patient outcomes. Two of these decisions include admission of a patient to an inpatient hospital (hospitalization) and the intensity of the

procedures performed on the patient. We explicitly model these two decisions as endogenous decisions as in Mueller-Smith (2015) and Bhuller et al. (2019). We include hospital admission as a dummy variable and proxy for intensity of procedures with work-Relative Value Units (w-RVU), which is the part of the CMS fee schedule that converts procedure codes to a payment amount.[31] We construct instruments for admission and procedure propensities analogous to prescribing leniency by replacing *Prescribed* in Equation 1 with admission dummy and total w-RVU.

We include predicted admission and total w-RVU as controls in the baseline 2SLS regression. This approach follows Mueller-Smith (2015) and is an indirect least squares version of three endogenous variables and three instruments. Column 3 of Table 13 reports the 2SLS estimates. The estimated coefficients are virtually unchanged, suggesting that an ED physician's admission and intensity of procedure decisions do not affect the patient's long-term outcomes, but rather, opioid prescriptions do.

Beyond admission and intensity of procedures, overall physician quality can bias our estimated coefficients. One proxy for physician quality is the effect on immediate patient mortality: what is the physician's effect on a veteran dying within one month of visiting the ED? With quasi-random assignment of patients to physicians, we can attribute any death that happens within a short period of one month to their practice in the ED. Such immediate mortality will not be caused by the physician's prescribing decision, but rather by other dimensions of physician care (the dimensions that may violate our exclusion restriction). We estimate this physician quality proxy analogous to our prescribing leniency instrument and the two admission and procedure propensities above. This estimated physician quality proxy is included as a control in the baseline 2SLS regressions in column 4. The coefficients are nearly identical and our main findings are robust.

Another potential violation of the exclusion restriction is the intensive margin decision of the *amount* of opioids to prescribe. We've modeled opioid prescriptions as a binary extensive

[31]The CMS fee schedule converts procedure codes to payments based on time, technical skill, and effort required. One caveat is that the VHA does not pay physicians on a fee-for-service basis, hence there is an under-reporting of procedures. To the extent that all physicians consistently under-report, this would only be a level-change without biasing our intensity of procedure estimates.

margin decision; however, physicians are also deciding on prescription length and dosage. If our extensive margin measure of prescribing leniency masks differences in the intensive margin, our 2SLS estimates will be biased. The bias will be upward (downward) if extensive margin lenient physicians are also lenient along the intensive margin. We investigate this by graphing the relationship between total milligrams of morphine equivalent (volume) on our baseline (extensive) physician leniency, conditional on being prescribed an opioid. Panel D of Figure F.9 indicates that there is a small positive relationship: the average physician in the top decile of extensive margin prescribing leniency, prescribes roughly 6 mg of morphine more than the bottom decile, conditional on being prescribed. This effect is not large since the average ED prescription has a morphine equivalent of 152 mg and the relationship is not monotonic. Nevertheless, we adopt the standard approach in accounting for the intensive margin in the judges design (Bhuller et al., 2019). We include a variable for the total milligrams of morphine prescribed (including zero if the patient is not prescribed) in Equation 3, construct an intensive margin instrument (just like the propensities above and our baseline instrument), and run a 2SLS regression with two endogenous variables and two instruments. Then we evaluate the average treatment effect conditional on being prescribed the average ED morphine equivalent dosage. We report this estimate in column 5 of Table 13, which represents the average treatment effect of being prescribed an average ED prescription, controlling for both the intensive and extensive margin decisions in opioid prescribing. The estimates on long-term use and opioid overdoses are slightly lower; for example, the effect on long-term opioid use is now 1.09pp as opposed to 1.17pp. The estimates of the other outcomes are completely unchanged; in fact, the effect on hepatitis C diagnosis is now larger. This suggests that our main findings are driven by physician variation in the decision to prescribe any opioid, and not by variation in the decision of the amount of opioids.

To further investigate the relationship between prescribing leniency and veteran outcomes, we can study patients with diagnosis conditions that are never prescribed opioids. For example, if a patient comes to an ED with a heart attack and is assigned a lenient physician, we should expect to see no reduced form effect of leniency on long-term use because heart

attacks are never prescribed opioids. If we do see a reduced form effect, then lenient physicians are doing something to the patient that results in worse care and worse outcomes. Formally, we run reduced form regressions of long-term opioid use and opioid overdose mortality on the physician leniency instrument for different conditions based on the likelihood that particular diagnoses are prescribed an opioid. The results of this exercise are displayed in Figure 6. All emergency visits are categorized into four bins based on the unconditional probability of being prescribed an opioid for their diagnosis. The bins are 0-3%, 3-10%, 10-20%, and 20%+, which roughly correspond to the four quartiles (recall that our baseline sample only includes the latter two groups). The diagnoses that are never prescribed (0-3%) serve as a "placebo" test. We find that when patients visit the ED with these diagnoses, seeing a lenient physician has no effect on their probability of long-term prescription opioid use or overdose mortality. In contrast, the effects are roughly monotonic in prescription probability, with the diagnoses that are prescribed over 20% of the time experiencing the largest effects.

Addressing Threats to Monotonicity

As mentioned in Table 4, the monotonicity assumption requires lenient physicians to be consistently lenient. We allow physicians to have differential prescribing leniency measures across different major diagnosis categories (MDC) and construct a physician-year-MDC-specific instrument. Column 6 of Table 13 reports 2SLS estimates with these differential, mutually exclusive instruments. The estimates of long-term use are qualitatively similar, and if anything, greater in magnitude (the sample is restricted to diagnoses with sufficiently numerous cases to measure leniency, and thus biased toward conditions such as injuries and musculoskeletal conditions that experience a larger causal effect).

4.5 Quantifying the Overall Effect: An Accounting Exercise

The analysis thus far does not capture the full effect of a single opioid prescription. This is because of the varying degrees of observability in our data with respect to different opioid outcomes. Consider one potential chain of downstream opioid outcomes originating from

the initial prescription: a veteran is prescribed an opioid, which may lead to a variety of opioid outcomes including long-term use, opioid dependence, and may finally result in opioid overdose death. We observe the initial ED opioid prescription status and final mortality nearly perfectly.[32] However, there is underreporting of intermediate opioid use outcomes (e.g., long-term opioid use, dependence, and overdose). For example, any prescriptions opioids filled through other insurance policies, acquired through the black market, as well as illicit heroin/fentanyl use, are unobserved. Similarly, opioid overdoses treated at non-VHA hospitals or at home with naloxone are not recorded by the VHA, nor are OUDs that go undiagnosed. There are reasons to believe the unobserved margin is large. Lin et al. (2019) find that only 23% of veterans who died from an opioid overdose in 2016 had a VHA opioid prescription in the prior three months. In our sample, only 42% of veterans who died from an opioid overdose were diagnosed with an OUD within three years of the ED visit.

The key observation that allows us to estimate unobserved and thus total opioid use is the fact that opioid overdose mortality are perfectly observed. With some empirical moments in the data, along with some assumptions on the likelihood an intermediate opioid outcome results in death by opioid overdose, we can back out the overall effect of an opioid prescription. Let us begin with estimating the total effect of a single opioid prescription on development of opioid use disorder (within three years).

Our 2SLS estimates provides us the causal (hereafter, all probabilities and effects are causal) probability that an ED opioid prescription causes a veteran to develop an *observed* OUD, $p^{O}_{OUD} = 0.335$, from Table 7. The *total* effect on OUD is the sum of the observed and unobserved parts: $p^{Total}_{OUD} = p^{O}_{OUD} + p^{U}_{OUD}$. The second term, the effect of an opioid prescription on *unobserved* OUD is unknown; however, it can be approximated in three steps.

First, we need to know how much of the causal effect on opioid overdose mortality is due to patients with an OUD we observe, which in turn inform us about the unobserved share. This probability (the probability a patient with an observed OUD—induced by an ED opioid

[32] In reality, it takes some time for the VHA to be alerted of veteran deaths. However, since our cause of death data ends at the end of 2016, all veteran deaths that take up to three years to be noted by the VHA are observed.

prescription—eventually dies from an opioid overdose), π^{O}_{OUD}, can be obtained via Bayes' rule:

$$\pi^{O}_{OUD} = P(\text{death} \mid \text{observed OUD}) = P(\text{observed OUD} \mid \text{death}) \times p_{death}/p^{O}_{OUD}. \quad (4)$$

Each term in Equation 4 can be approximated with sample moments in our data if we assume that the *causal* probability of dying from an opioid overdose, conditional on having an OUD (induced by an ED prescription), can be approximated by the unconditional population counterpart. Specifically, 42% of veterans in our baseline sample who died from an opioid overdose were diagnosed with an OUD, the opioid overdose mortality rate is 0.167%, and the three-year OUD diagnosis rate is 3.27%. These estimates yield a overdose mortality effect, conditional on having an observed OUD of $\pi^{O}_{OUD} = 0.0215$. This implies that $0.335 \times 0.0215 = 0.0072$ of the 0.075 opioid overdose mortality effect (2SLS estimate on mortality from Table 7) can be attributed to veterans with an observed OUD, who eventually die. This leaves the remaining 0.0678 of the total mortality effect to veterans with unobserved OUDs.[33]

Second, we need to know the probability of dying from an opioid overdose, conditional on having an unobserved OUD, π^{U}_{OUD}. We benchmark this probability against the probability for observed OUDs found above, and consider bounds for this ratio: $\pi^{U}_{OUD}/\pi^{O}_{OUD}$. For a lower bound, we assume that unobserved OUDs are on average, at least as lethal as observed OUDs. To obtain an upper bound, we make three assumptions. First, we assume that all unobserved OUD cases are related to illicit heroin (prescription only, and all observed cases are related to prescription opioids. Next, we assume that opioid overdose mortality among the unobserved cases is only caused by heroin, and mortality among the observed cases is only caused by prescription opioids. These two assumptions allow us to separate the observed and unobserved margins as disjoint prescription and heroin users while still preserving the upper

[33] We are using the following identity which decomposes the causal effect of an opioid prescription on opioid overdose mortality into effects attributed to patients with observed and unobserved OUDs: $p_{death} = \pi^{O}_{OUD} \times p^{O}_{OUD} + \pi^{U}_{OUD} \times p^{U}_{OUD}$

bound. Finally, we can assume that the number of opioid (or heroin) deaths is growing at the same rate as the number of Americans with opioid (or heroin) substance use disorder. These three assumptions allow us to approximate π^U_{OUD}/π^O_{OUD} using survey counts of prescription and illicit opioid users, and counts of annual deaths from prescription and illicit opioids. This calibration exercise arrives at an upper bound of 3.58: unobserved OUDs are at most 3.58 times as deadly as observed OUDs.[34]

Finally, with our lower and upper bound on the probability of dying from an opioid overdose conditional on having an unobserved OUD, we can back-out the amount of unobserved OUDs that ultimately resulted in a 0.0678 effect on deaths. This arrives at an effect of the initial ED prescription on unobserved cases of OUD of $p^U_{OUD} \in [0.88, 3.16]$. Including the observed effect, the total effect on OUDs is $p^{Total}_{OUD} \in [1.22, 3.50]$, many times larger than the effect on observed (diagnosed and recorded) OUDs caused by an ED prescription.

Analogously, we can extend the logic of this exercise and account for the overall effect of an ED opioid prescription on total long-term opioid use[35] We arrive at an estimate of 2.6-6.3pp. In summary, by using our estimated effects on opioid overdose mortality to account for the total effect of an opioid prescription, we find that for an additional 100,000 veterans who are prescribed opioids in the ED due to a more lenient physician, 2,600-6,300 will become long-term users, 1,220-3,500 will develop an opioid use disorder within three years, and 75 will eventually die from an opioid overdose.

Another way to contextualize our causal estimates is to ask, "How many of the veteran deaths from opioid overdose during the years of our study were caused by an opioid prescription prescribed in the emergency department?" Of the approximately 9,200 veterans who died

[34] According to CDC NCHS multiple cause of death, there were 28,466 deaths from heroin and fentanyl, and 17,762 from prescription opioids in 2017. National Survey on Drug Use and Health estimate 2,110,000 adult Americans with any opioid use disorder, of which 652,000 have a heroin use disorder. Dividing number of annual deaths per number of annual users for heroin versus prescription opioids yields our upper bound of 3.58.

[35] *Long-term prescription opioid use* is defined as 180 days of prescription opioids in the first year following the ED visit. For illicit opioid use, the definition is not as straight-forward. One way is to view it as a measure of total milligrams of morphine and to convert the amount of heroin use into an equivalent days supply of Oxycontin for the average long-term prescription opioid user. Alternatively, one may view long-term prescription opioid use as an intermediate outcome that can lead to opioid dependence, in which case all regular users of illicit heroin are considered "long-term" users.

from an opioid overdose between 2006 and 2016, 5.85% were caused by the ED prescription for the diagnoses and criteria to which we restrict our baseline sample.[36] Each year, around 178,000 new veterans visit the ED who satisfy our sample restrictions. Approximately 300 of them will eventually die from an opioid overdose, of which 52 were because they were prescribed in the ED due to the leniency of their prescriber.

5. Conclusion

In this paper, we find that even a single opioid prescription, among cases where clinicians differ in their prescribing choices, can have lasting effects on a patient's long-term outcomes. A single opioid prescription—prescribed because a veteran is quasi-randomly assigned a more lenient opioid-prescribing physician in the emergency department—increases the probability of observed long-term prescription opioid use by 1.2pp. This long-term opioid use can lead to opioid misuse, abuse, and dependence. There is a 0.34pp increase in the probability of being diagnosed with an opioid use disorder within three years of the initial emergency visit. Further along the downstream chain of adverse outcomes, the single prescription can also lead to an increased probability of opioid overdose, accidental falls, and even potential illicit drug use (cocaine/crack, heroin/fentanyl, and injection drugs). Finally, opioid overdose mortality increases by 0.075pp due to the initial prescription, with 30-40% of the effect coming from illicit heroin or synthetic opioid (i.e., fentanyl and tramadol) related mortality. Because of the amount of unobserved opioid use and misuse, intermediate (non-mortality) opioid outcomes are very likely many times larger. Note, however, that these findings apply to the patient population for which clinicians differed in their choices to prescribe opioids or not. This study design cannot provide insight into the impact of opioid prescribing for the types of cases where clinicians universally chose to prescribe or not prescribe under guidelines of the time.

In sum, our findings provide strong evidence that even a single opioid prescription can

[36] The back-of-envelope calculation is conducted as follows: of the 9,200 veterans, 3,077 visited the ED and enter our baseline sample. Our complier share is estimated to be 39%. Our 2SLS causal estimate of opioid overdose mortality is 0.075 on a base of 0.1666.

have severe consequences, even for acute conditions like injuries and even for patients who are not prior opioid users. This paper studies short-term exposure to opioid risk in the form of emergency department encounters. In related work, Eichmeyer and Zhang (2020) study veterans assigned to primary care providers and find exposure to lenient opioid-prescribers in settings with more frequent and repeat interactions, is associated with adverse long-term outcomes that are qualitatively consistent with the findings in this paper.

Our research design quantifies the long-run costs associated with opioid prescribing relative to the implicitly embedded counterfactual of handing out over-the-counter nonsteroidal anti-inflammatory drugs (NSAIDs). Studies have found that prescription opioids are no more effective than NSAIDs at treating chronic pain, in terms of pain-related function and experienced pain intensity (Dodwell et al., 2010; Friedman et al., 2015; Holdgate and Pollock, 2004; Krebs et al., 2018; Teichman, 2004). This implies that for new patients, prescribing non-opioid therapy (e.g., NSAIDS, physical therapy, acupuncture, meditation, etc.) can treat pain just as effectively, without the adverse effects of opioids. At the VHA, completely cutting back on prescribing opioids for non-heavy prior users in the ED would result in roughly 52 fewer veteran deaths per year.

Many in the medical community have cautioned that opioid restriction policies should not be a one-size-fits-all solution (Lowenstein et al., 2018; HP3, 2019). For instance, patients who are already opioid dependent or are on high doses may substitute to more dangerous black market or illicit opioids if their legal supply is cut. Alpert et al. (2018) and Evans et al. (2019) find that heroin deaths increased after the introduction of abuse-deterrent Oxycontin in 2010; similarly, Meinhofer (2018) finds that required access Prescription Drug Monitoring Programs increased illicit heroin deaths. Investigating responses to opioid tapering, Mark and Parish (2019) find that faster rates of opioid tapering were associated with a great immediate probability of opioid poisoning or substance use disorder for patients already on high doses. We find suggestive evidence supporting the substitution of legal prescription opioids for black market and illicit opioids. Veterans with a lenient opioid-prescribing primary care provider are more likely to satisfy their excess demand for opioids through legal prescriptions. In contrast,

veterans with strict primary care providers may resort to illicit heroin or black market opioids, resulting in a higher probability of both opioid overdose mortality and hepatitis C infection. For opioid-dependent patients or patients already on high doses, policies should consider both the secondary (illicit) market and appropriate responses to policies targeting the primary (legitimate) market (Schnell, 2017). For patients who are opioid-dependent or have turned to illicit sources, medication-assisted treatment (MAT) such as methadone or buprenorphine has been found effective for reducing overdoses (Schwartz et al., 2013; Connery, 2015) and hepatitis C infections (Tsui et al., 2014). However, the efficacy of MAT in mitigating the opioid epidemic is unclear, largely due to lack of access to MAT programs nationwide. Further research into the design of policies targeting the opioid-dependent and solutions to the opioid crisis are needed.

References

Abadie, Alberto (2003), "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics*, 113, 231–263.

Al-Tayyib, Alia, Stephen Koester, Sig Langegger, and Lisa Ravilled (2017), "Heroin and Methamphetamine Injection: An Emerging Drug Use Pattern." *Substance Use & Misuse*, 52, 1051–1058.

Alpert, Abby, William N. Evans, Michael M.J Lieber, and David Powell (2019), "Origins of the Opioid Crisis and Its Enduring Impacts." Working paper.

Alpert, Abby, David Powell, and Rosalie Liccardo Pacula (2018), "Supply-Side Drug Policy in the Presence of Substitutes: Evidence from the Introduction of Abuse-Deterrent Opioids." *American Economic Journal: Economic Policy*, 10, 1–35.

Baker, Laurence C., Kate Bundorf, and Daniel Kessler (2018), "The Effects of Medicare Advantage on Opioid Use." Nber working paper 25327.

Barnett, Michael L., Andrew Olenski, and Anupam B. Jena (2017), "Opioid-Prescribing Patterns of Emergency Physicians and Risk of Long-Term Use." *New England Journal of Medicine*, 376, 663–673.

Barnett, Michael L., Xinhua Zhao, Michael J. Fine, Carolyn T. Thorpe, Florentina E. Sileanu, John P. Cashy, Maria K. Mor, Thomas R. Radomski, Leslie R. M. Hausmann, Chester B. Good, and Walid F. Gellad (2019), "Emergency Physician Opioid Prescribing and Risk of Long-term Use in the Veterans Health Administration: an Observational Analysis." *Journal of General Internal Medicine*, 34, 1522–1529.

Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Løken, and Magne Mogstad (2019), "Incarceration, Recidivism, and Employment." *Journal of Political Economy*, Forthcoming.

Borgschulte, Mark, Adriana Corredor-Waldron, and Guillermo Marshall (2018), "A path out: Prescription drug abuse, treatment, and suicide." *Journal of Economic Behavior & Organization*, 149, 169–184.

Boscarino, Joseph A., Margaret Rukstalis, Stuart N. Hoffman, and John J. Han Porat M. Erlich Glenn S. Gerhard Walter F. Stewart (2010), "Risk factors for drug dependence among out-patients on opioid therapy in a large US health-care system." *Addiction*, 105, 1776–1782.

Braden, Jennifer Brennan, Joan Russo, Ming-Yu Fan, Mark J. Edlund, Bradley C. Martin, Andrea DeVries, and Mark D. Sullivan (2010), "Emergency department visits among recipients of chronic opioid therapy." *Archives of Internal Medicine*, 170, 1425–1432.

Buchmueller, Thomas C. and Colleen Carey (2018), "The effect of prescription drug monitoring programs on opioid utilization in Medicare." *American Economic Journal: Economic Policy*, 10, 77–112.

Case, Anne and Angus Deaton (2015), "Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century." *PNAS*, 112, 15078–15083.

Case, Anne and Angus Deaton (2017), "Mortality and Morbidity in the 21st Century." Brookings Papers on Economic Activity.

CDC (2016a), "Hepatitis C a Silent Epidemic."

CDC (2016b), "Surveillance for Viral Hepatitis – United States, 2016. Figure 4.6a."

CDC (2018), "FastStats - Emergency Department Visits."

Cesur, Resul, Joseph J. Sabia, and W. David Bradford (2019), "Did the War on Terror Ignite an Opioid Epidemic?"

Chan, David C. (2016), "Informational Frictions and Practice Variation: Evidence from Physicians in Training." Nber working paper.

Chua, Kao-Ping, Chad M. Brummett, , and Jennifer F. Waljee (2019), "Opioid Prescribing Limits for Acute Pain Potential Problems With Design and Implementation." *The Journal of American Medical Association*, 321, 643–644.

CMS (2019), "A Prescriber's Guide to the New Medicare Part D Opioid Overutilization Policies for 2019 ."

Connery, Hilary Smith (2015), "Medication-Assisted Treatment of Opioid Use Disorder Review of the Evidence and Future Directions." *Harvard Review of Psychiatry*, 23, 63–75.

Currie, Janet, Jonas Jin, and Molly Schnell (2019), "U.S. Employment and Opioids: Is There a Connection?" *Health and Labor Markets (Research in Labor Economics, Volume 47)*, 253–280.

Dahl, Gordon B., Andreas Ravndal Kostøl, and Magne Mogstad (2014), "Family Welfare Cultures." *The Quarterly Journal of Economics*, 129, 1711–1752.

Dave, Dhaval, Monica Deza, and Brady P. Horn (2018), "Prescription Drug Monitoring Programs, Opioid Abuse, and Crime."

Davis, Corey S., Amy Judd Liebermana, Hector Hernandez-Delgadoa, and Carli Suba (2019), "Laws limiting the prescribing or dispensing of opioids for acute pain in the United States: A national systematic legal review." *Drug and Alcohol Dependence*, 194, 166–172.

de Vaan, Mathijs and Toby Stuart (2019), "Does Intra-household Contagion Cause an Increase in Prescription Opioid Use?" *American Sociological Review*, 18.

Degenhardt, Louisa, Amy Peacock, Samantha Colledge, Janni Leung, Jason Grebely, Peter Vickerman, Jack Stone, Evan B Cunningham, Adam Trickey, Kostyantyn Dumchev, Michael Lynskey, Paul Griffiths, Richard P Mattick, Matthew Hickman, and Sarah Larney (2017), "Global prevalence of injecting drug use and sociodemographic characteristics and prevalence of HIV, HBV, and HCV in people who inject drugs: a multistage systematic review." *The Lancet Global Health*, 5, 1192–1207.

Dobbie, Will, Jacob Goldin, and Crystal S. Yang (2018), "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges." *Amercian Economic Review*, 108, 201–240.

Dobbie, Will, Paul Goldsmith-Pinkham, and Crystal S. Yang (2017), "Consumer Bankruptcy and Financial Health." *Review of Economics and Statistics*, 99, 853–869.

Doctor, Jason N., Andy Nguyen, Roneet Lev, Jonathan Lucas, Tara Knight, Henu Zhao, and Michael Menchine (2018), "Opioid prescribing decreases after learning of a patient's fatal overdose." *Science*, 361, 588–590.

Dodwell, Emily R., Julius Gene Latorre, Emilio Parisini, Elisabeth Zwettler, Divay Chandra, Kishore Mulpuri, and Brian Snyder (2010), "NSAID exposure and risk of nonunion: a meta-analysis of case-control and cohort studies." *Calcified Tissue International*, 87, 193–202.

Doleac, Jennifer and Anita Mukherjee (2019), "The Moral Hazard of Lifesaving Innovations: Naloxone Access, Opioid Abuse, and Crime." Working paper.

Dowell, Deborah, Tamara M. Haegerich, and Roger Chou (2016), "CDC Guideline for Prescribing Opioids for Chronic Pain — United States."

Doyle, Joseph J., John A. Graves, Jonathan Gruber, and Samuel A. Kleiner (2015), "Measuring Returns to Hospital Care: Evidence from Ambulance Referral Patterns." *Journal of Political Economy*, 123, 170–214.

Doyle, Jospeph J. (2007), "Child Protection and Child Outcomes: Measuring the Effects of Foster Care." *American Economic Review*, 97, 1583–1610.

Duggan, Mark (2005), "Do new prescription drugs pay for themselves? The case of second-generation antipsychotics." *Journal of Health Economics*, 24, 1–31.

Dunn, Kate M., Kathleen W. Saunders, Carolyn M. Rutter, Caleb J. Banta-Green, Joseph O. Merrill, Mark D. Sullivan and; Constance M. Weisner, Michael J. Silverberg, Cynthia I. Campbell, Bruce M. Psaty, and Michael Von Korff (2010), "Opioid prescriptions for chronic pain and overdose: a cohort study." *Annals of Internal Medicine*, 152, 85–92.

Edlund, Mark, Bradley C. Martin, Joan E. Russo, Andrea Devries, Jennifer Brennan Braden, and Mark D. Sullivan (2014), "The Role of Opioid Prescription in Incident Opioid Abuse and Dependence Among Individuals with Chronic Non-Cancer Pain: The Role of Opioid Prescription." *The Clinical Journal of Pain*, 30, 557–564.

Edlund, Mark, Diane Steffick, Teresa Hudson, Katherine Harris, and Mark Sullivan (2007), "Risk factors for clinically recognized opioid abuse and dependence among veterans using opioids for chronic non-cancer pain." *Pain*, 129, 355–362.

Eichmeyer, Sarah and Jonathan Zhang (2020), "Prescription Opioid Use, Physical and Mental Health: Evidence from Primary Care." Working paper.

Elizabeth M. Oliva, Sara Tavakoli Eleanor T. Lewis Ilse Weichers Michael Harvey Amanuel Medhanie Susana martins Meenah Paik Patricia Henderson Tigran Avoundjian, Thomas Bowe and Jodie A. Trafton (2017), "Development and Applications of the Veterans Health Administration's Stratification Tool for Opioid Risk Mitigation (STORM) to Improve Opioid Safety and Prevent Overdose and Suicide." *Psychological Services*, 14, 34–49.

Epstein, Andrew and Sean Nicholson (2009), "The Formation and Evolution of Physician Treatment Styles: An Application to Cesarean Sections." *Journal of Health Economics*, 28, 1126–1140.

Evans, William N., Ethan M. J. Lieber, and Patrick Power (2019), "How the Reformulation of OxyContin Ignited the Heroin Epidemic." *The Review of Economics and Statistics*, 101, 1–15.

Fadlon, Itzik and Jessica N. Van Parys (2019), "Primary Care Physician Practice Styles and Patient Care: Evidence From Physician Exits in Medicare."

Farre-Mensa, Joan, Deepak Hegde, and Alexander Ljungqvist (2019), "What Is a Patent Worth? Evidence from the U.S. Patent 'Lottery'." *Journal of Finance*, Forthcoming.

Fernandez, Fernando and Dijana Zejcirovi (2017), "The Role of Pharmaceutical Promotion to Physicians in the Opioid Epidemic."

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams (2016), "Sources of Geographic Variation in Health Care: Evidence from Patient Migration." *Quarterly Journal of Economics*, 131, 1681–1726.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams (2018), "What Drives Prescription Opioid Abuse? Evidence from Migration." Working paper.

Frandsen, Brigham R., Lars J. Lefgren, and Emily C. Leslie (2019), "Judging Judge Fixed Effects." Working paper.

Friedman, Benjamin W., Andrew A. Dym, Michelle Davitt, Lynne Holden, Clemencia Solorzano, David Esses, Polly E. Bijur, and E. John Gallagher (2015), "Naproxen With Cyclobenzaprine, Oxycodone/Acetaminophen, or Placebo for Treating Acute Low Back Pain A Randomized Clinical Trial." *The Journal of American Medical Association*, 315, 1572–1580.

Furlan, Andrea D, Luis E Chaparro, Emma Irvin, and Angela Mailis-Gagno (2011), "A Comparison between Enriched and Nonenriched Enrollment Randomized Withdrawal Trials of Opioids for Chronic Noncancer Pain." *Pain Research and Management*, 16, 337–351.

GAO (2018), "VA HEALTH CARE: Progress Made Towards Improving Opioid Safety, but Further Efforts to Assess Progress and Reduce Risk Are Needed."

Gowrisankaran, Gautam, Keith Joiner, and Pierre-Thomas Leger (2017), "Physician Practice Style and Healthcare Costs: Evidence from Emergency Departments."

Grecu, Anca M., Dhaval M. Dave, and Henry Saffer (2019), "Mandatory Access Prescription Drug Monitoring Programs and Prescription Drug Abuse." *Journal of Policy Analysis and Management*, 38, 181–209.

Hagan, Holly, Enrique R. Pouget, Don C. Des Jarlais, and Corina Lelutiu-Weinberger (2008), "Meta-Regression of Hepatitis C Virus Infection in Relation to Time Since Onset of Illicit Drug Injection: The Influence of Time and Place." *American Journal of Epidemiology*, 168, 1099–1109.

Holdgate, Anna and Tamara Pollock (2004), "Systematic review of the relative efficacy of non-steroidal anti-inflammatory drugs and opioids in the treatment of acute renal colic." *The British Medical Journal*, 328, 1401–8.

HP3 (2019), "Professionals Call on the CDC to Address Misapplication of its Guideline on Opioids for Chronic Pain through Public Clarification and Impact Evaluation." https://docs.google.com/document/d/1RzQDSppUKhjiAsEmhW2WbTXlP5V8vJ4M_vBPQLKhK_8/edit.

Huang, Grace, Sharon Kim, Benjamin Muz, and Joseph Gasper (2018), "2017 Survey of Veteran Enrollees' Health and Use of Health Care." Westat.

Hwang, Catherine, Lydia Turner, Stefan Kruszewski, Andrew Kolodny, and G. Alexander (2016), "Family Welfare Cultures." *The Clinical Journal of Pain*, 32, 279–284.

Ives, Timothy J, Paul R Chelminski, Catherine A Hammett-Stabler, Robert M Malone, J Stephen Perhac, Nicholas M Potisek, Betsy Bryant Shilliday, Darren A DeWalt, and Michael P Pignone (2006), "Predictors of opioid misuse in patients with chronic pain: a prospective cohort study." *BMC Health Services Research*, 6.

Jena, Anupam B., Dana Goldman, and Pinar Karaca-Mandic (2016), "Hospital Prescribing of Opioids to Medicare Beneficiaries." *The Journal of American Medical Association Internal Medicine*, 176, 990–997.

Kaestner, Robert and Engy Ziedan (2019), "Mortality and Socioeconomic Consequences of Prescription Opioids: Evidence from State Policies."

Kilby, Angela (2015), "Opioids for the masses: Welfare tradeoffs in the regulation of narcotic pain medications."

Kling, Jeffrey R. (2006), "Incarceration Length, Employment, and Earnings." *American Economic Review*, 96, 863–876.

Krashin, Daniel, Natalia Murinova, and Mark Sullivan (2016), "Challenges to Treatment of Chronic Pain and Addiction During the "Opioid Crisis"." *Current Pain and Headache Reports*, 20.

Krebs, Erin E., Amy Gravely, Sean Nugent, Agnes C. Jensen, Beth DeRonne, Elizabeth S. Goldsmith, MD4 Kurt Kroenke, Matthew J. Bair, and Siamak Noorbaloochi (2018), "Effect of Opioid vs Nonopioid Medications on Pain-Related Function in Patients With Chronic Back Pain or Hip or Knee Osteoarthritis Pain: The SPACE Randomized Clinical Trial." *The Journal of American Medical Association*, 319, 872–882.

Laird, Jessica and Torben Nielsen (2016), "The Effects of Physician Prescribing Behaviors on Prescription Drug Use and Labor Supply: Evidence from Movers in Denmark." Working paper.

Lin, Lewei (Allison), Talya Peltzman, John F. McCarthy, Elizabeth M. Oliva, Jodie A. Trafton, and Amy S.B. Bohnert (2019), "Changing Trends in Opioid Overdose Deaths and Prescription Opioid Receipt Among Veterans." *American Journal of Preventive Medicine*, 57, 106–110.

Lowenstein, Margaret, David Grande, and M. Kit Delgado (2018), "Opioid Prescribing Limits for Acute Pain — Striking the Right Balance." *New England Journal of Medicine*, 379, 504–506.

Mark, Tami L. and William Parish (2019), "Opioid medication discontinuation and risk of adverse opioid-related health care events." *Journal of Substance Abuse Treatment*, 103, 58–632.

Meinhofer, Angélica (2018), "Prescription Drug Monitoring Programs: The Role of Asymmetric Information on Drug Availability and Abuse." *American Journal of Health Economics*, 4, 504–526.

Moeller, Karen E., Kelly C. Lee, and Julie C. Kissack (2008), "Urine Drug Screening: Practical Guide for Clinicians." *Mayo Clinic Proceedings*, 83, 66–76.

Molitor, David (2018), "The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration." *American Economic Journal: Economic Policy*, 10, 326–356.

Mueller, Shane R., Alexander Y. Walley, Susan L. Calcaterra, Jason M. Glanz, and Ingrid A. Binswanger (2015), "A review of opioid overdose prevention and naloxone prescribing:

Implications for translating community programming into clinical practice." *Substance Abuse*, 36, 240–253.

Mueller-Smith, Michael (2015), "The criminal and labor market impacts of incarceration." Working paper.

Muhuri, Pradip K., Joseph C. Gfroerer, and M. Christine Davies (2013), "Associations of nonmedical pain reliever use and initiation of heroin use in the United States." *CBHSQ Data Review*.

NCHS (2016), "National Hospital Ambulatory Medical Care Survey: 2016 emergency department summary tables."

NCHS (2017), "Opioid Abuse in Chronic Pain — Misconceptions and Mitigation Strategies."

Neprash, Hannah T. and Michael L. Barnett (2019), "Association of Primary Care Clinic Appointment Time With Opioid Prescribing." *Journal of American Medical Association Network Open*, 2.

Nguyen, Thuy D., W. David Bradford, and Kosali I. Simon (2019), "How do Opioid Prescribing Restrictions Affect Pharmaceutical Promotion? Lessons from the Mandatory Access Prescription Drug Monitoring Programs." Nber working paper 26356.

Van Parys, Jessica (2016), "Variation in Physician Practice Styles Within and Across Emergency Departments." *Plos One*, 11, 1–19.

Novak, Scott P. and Alex H. Kral (2011), "Comparing Injection and Non-Injection Routes of Administration for Heroin, Methamphetamine, and Cocaine Uses in the United States." *Journal of Addictive Diseases*, 30, 248–257.

Packham, Analisa (2019), "Are Syringe Exchange Programs Helpful or Harmful? New Evidence in the Wake of the Opioid Epidemic."

Peterson, Rachel, Adi V. Gundlapalli, Stephen Metraux, Marjorie E. Carter, Miland Palmer, Andrew Red, Matthew H. Samore, and Jamison D. Fargo (2015), "Identifying Homelessness among Veterans Using VA Administrative Data: Opportunities to Expand Detection Criteria." *PLoS One*, 10.

Powell, David, Abby Alpert, and Rosalie L. Pacula (2019), "A Transitioning Epidemic: How The Opioid Crisis Is Driving The Rise In Hepatitis C." *Health Affairs*, 38, 287–294.

Quast, Troy, Eric A. Storch, and Svetlana Yampolskaya (2018), "Opioid Prescription Rates And Child Removals: Evidence From Florida." *Health Affairs*, 37, 134–139.

Rees, Daniel I., Joseph J. Sabia, Laura M. Argys, Dhaval Dave, and Joshua Latshaw (2019), "With a Little Help from My Friends: The Effects of Good Samaritan and Naloxone Access Laws on Opioid-Related Deaths." *The Journal of Law and Economics*, 62, 1–27.

Robin, Lee N., Darlene H. Davis, and David N. Nurco (1974), "How Permanent Was Vietnam Drug Addiction?" *American Journal of Public Health*, 64, 38–43.

Ross, Edgar, Robert Jamison, and Robert Edwards (2011), "The Opioid Debate - Missing the Point." *The Journal of Pain*, 12, 508.

Rutkow, Lainie, Hsien-Yen Chang, Matthew Daubresse, Daniel W. Webster, Elizabeth A. Stuart, and G. Caleb Alexander (2011), "Effect of Florida's Prescription Drug Monitoring Program and Pill Mill Laws on Opioid Prescribing and Use." *The Journal of American Medical Association Internal Medicine*, 175, 1642–1649.

Sacks, Daniel W., Alex Hollingsworth, Thuy D. Nguyen, and Kosali I. Simon (2019), "Can Policy Affect Initiation of Addictive Substance Use? Evidence from Opioid Prescribing."

Sasson, Comilla, Jason Smith, Chad Kessler, Jason Haukoos, Julianne Himstreet, Melissa Christopher, and Thomas Emmendorfer (2019), "Variability in opioid prescribing in veterans affairs emergency departments and urgent cares." *The American Journal of Emergency Medicine*, 37, 2044–1047.

Savych, Bogdan, David Neumark, and Randall Lea (2018), "Do Opioids Help Injured Workers Recover and Get Back to Work? The Impact of Opioid Prescriptions on Duration of Temporary Disability."

Scherrer, Jeffrey F., Joanne Salas, Laurel A. Copeland, Eileen M. Stock, Brian K. Ahmedani, Mark D. Sullivan, Thomas Burroughs, David Schneider, Kathleen K. Bucholz, and Patrick J. Lustman (2016), "Prescription Opioid Duration, Dose, and Increased Risk of Depression in 3 Large Patient Populations." *Annals of Family Medicine*, 14, 54–62.

Schnell, Molly (2017), "Physician Behavior in the Presence of a Secondary Market: The Case of Prescription Opioids."

Scholl, Lawrence, Puja Seth, Mbabazi Kariisa, Nana Wilson, and Grant Baldwin (2019), "Drug and Opioid-Involved Overdose Deaths — United States, 2013–2017." Working paper.

Schwartz, Robert P., Jan Gryczynski, Kevin E. O'Grady, Joshua M. Sharfstein, Gregory Warren, Yngvild Olsen, Shannon G. Mitchell, and Jerome H. Jaffe (2013), "Opioid Agonist Treatments and Heroin Overdose Deaths in Baltimore, Maryland, 1995–2009." *American Journal of Public Health*, 103, 917–922.

Seal, Karen H., Ying Shi, Gregory Cohen, Beth E. Cohen, Shira Maguen, Erin E. Krebs, and Thomas C. Neylan (2012), "Association of mental health disorders with prescription opioids and high-risk opioid use in US veterans of Iraq and Afghanistan." *The Journal of American Medical Association*, 307, 940–947.

Silver, David (2019), "Haste or Waste? Peer Pressure and Productivity in the Emergency Department."

Sowicz, Timothy Joseph, Adam J. Gordon, Walid F. Gellad, Xinhua Zhao, Hongwei Zhang, Thomas Emmendorfer, and Chester B. Good (2018), "High Variability of Opioid Prescribing Within and Across Emergency Departments in the US Veterans Health Administration." *Journal of General Internal Medicine*, 33, 1831–1832.

Starrels, Joanna L., Aaron D. Fox, Hillary V. Kunins, and Chinazo O. Cunningham (2012), "They Don't Know What They Don't Know: Internal Medicine Residents' Knowledge and Confidence in Urine Drug Test Interpretation for Patients with Chronic Pain." *Journal of General Internal Medicine*, 27, 1521–1527.

Stephenson, Joan (1999), "Veterans' Pain a Vital Sign." *The Journal of American Medical Association Internal Medicine*, 281, 978.

Teichman, Joel M.H. (2004), "Acute renal colic from ureteral calculus." *The New England Journal of Medicine*, 350, 684–693.

Tsugawa, Yusuke, Ashish K. Jha, Joseph P. Newhouse, Alan M. Zaslavsky, and Anupam B. Jena (2017), "Variation in Physician Spending and Association With Patient Outcomes." *JAMA Intern Med*, 177, 675–682.

Tsui, Judith I., Jennifer L. Evans, Paula J. Lum, Judith A. Hahn, and Kimberly Page (2014), "Association of Opioid Agonist Therapy With Lower Incidence of Hepatitis C Virus Infection in Young Adult Injection Drug Users." *The Journal of American Medical Association Internal Medicine*, 174, 1974–1984.

Tu, Peter (2017), "Quasi-experimental Evidence of Physician Effects." Working paper.

Volkow, Nora D. and Thomas A. McLellan (2016), "Opioid Abuse in Chronic Pain — Misconceptions and Mitigation Strategies." *The New England Journal of Medicine*, 374, 1253–1263.

Wu, Christopher L. and Srinivasa N. Raja (2011), "Treatment of Acute Postoperative Pain." *The Lancet*, 377, 2215–2225.

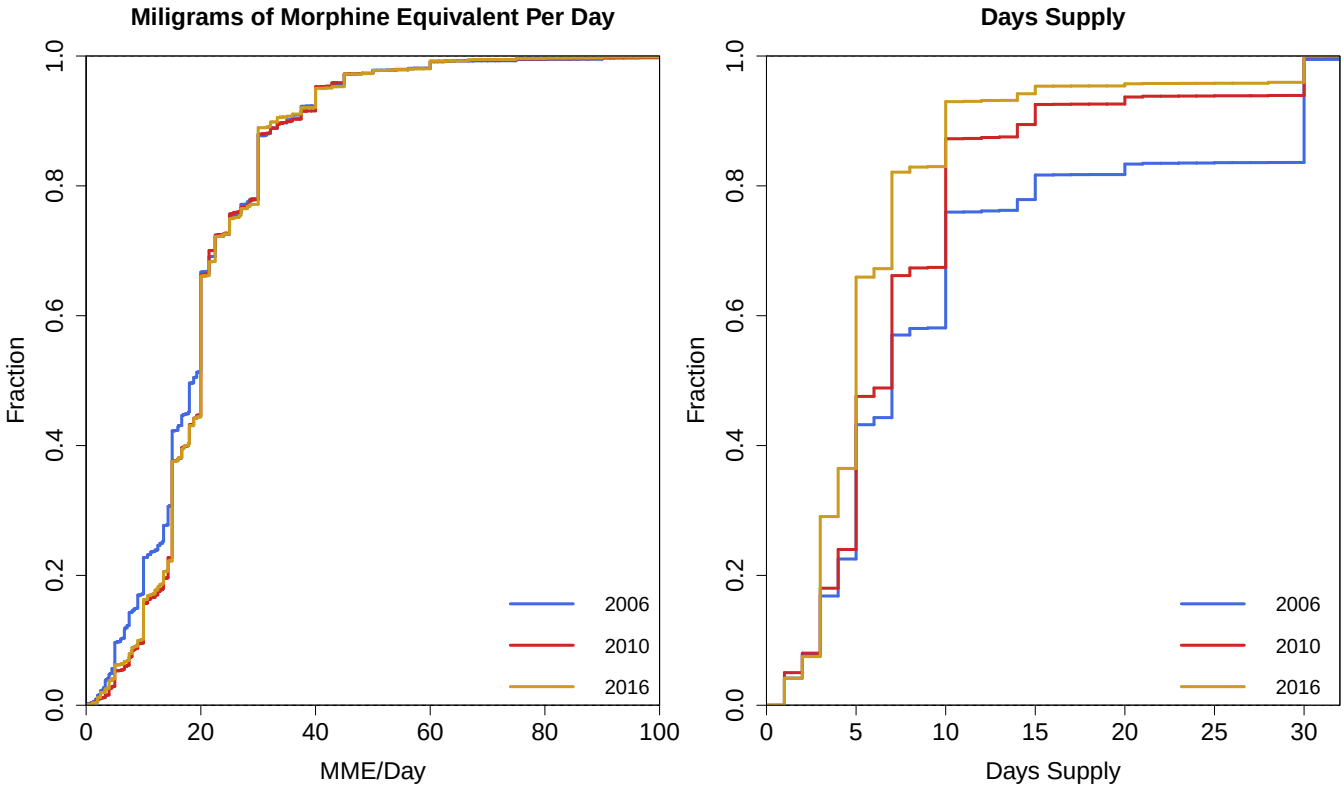
Yang, Zhuo, Barth Wilsey, Michele Bohm, Meghan Weyrich, Kakoli Roy, Dominique Ritley, Christopher Jones, and Joy Melnikowx (2015), "Defining Risk of Prescription Opioid Overdose: Pharmacy Shopping and Overlapping Prescriptions Among Long-Term Opioid Users in Medicaid." *The Journal of Pain*, 16, 445–453.

Zedler, Barbara, Lin Xie, Li Wang, Andrew Joyce, Catherine Vick, Furaha Kariburyo, Pradeep Rajan, and Onur Baserand Lenn Murrelle (2014), "Risk Factors for Serious Prescription Opioid-Related Toxicity or Overdose among Veterans Health Administration Patients." *Pain Medicine*, 15, 1911–1929.

Zibbell, Jon E., Alice K. Asher, Rajiv C. Patel, Ben Kupronis, Kashif Iqbal, John W. Ward MD, and Deborah Holtzman (2018), "Increases in Acute Hepatitis C Virus Infection Related to a Growing Opioid Epidemic and Associated Injection Drug Use, United States, 2004 to 2014." *American Journal of Public Health*, 108, 175–181.

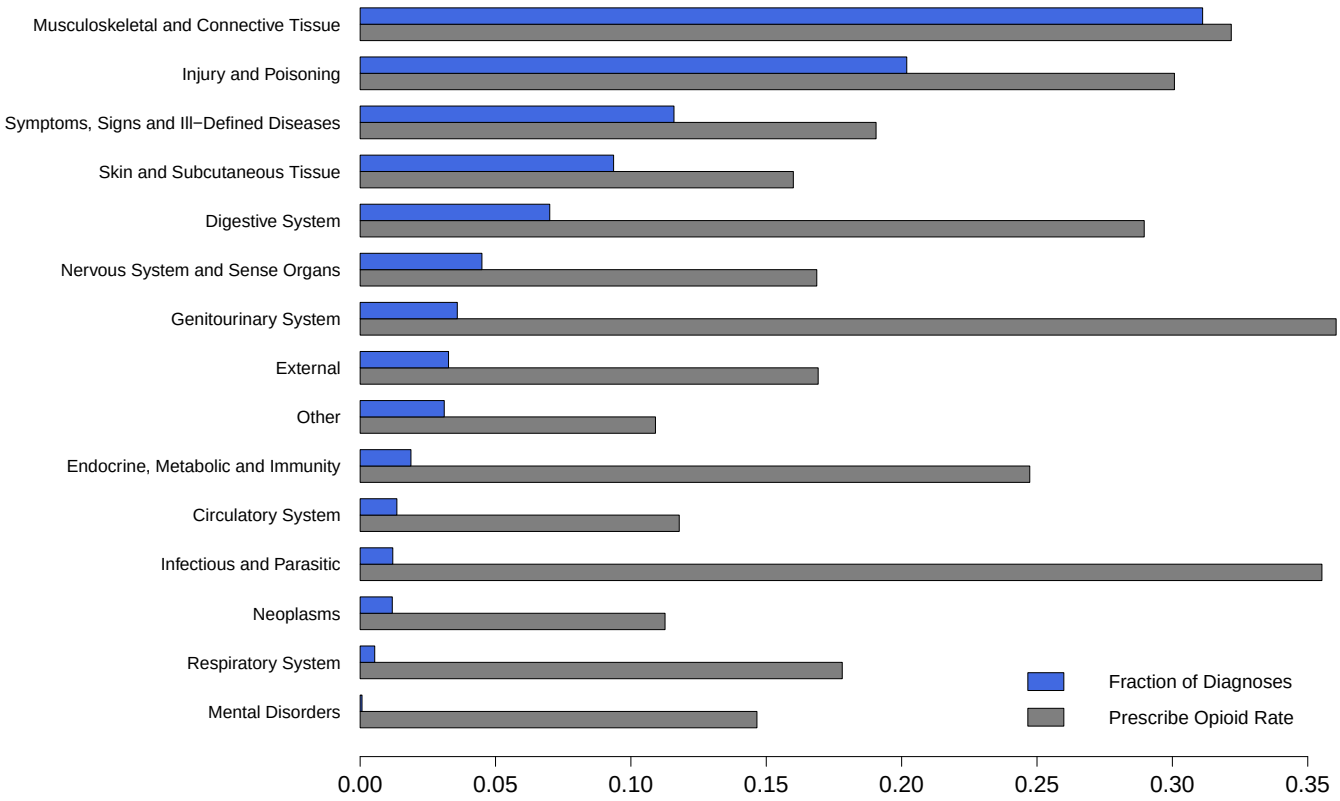
Figures

Figure 1: Empirical CDF of Opioid Prescription Characteristics



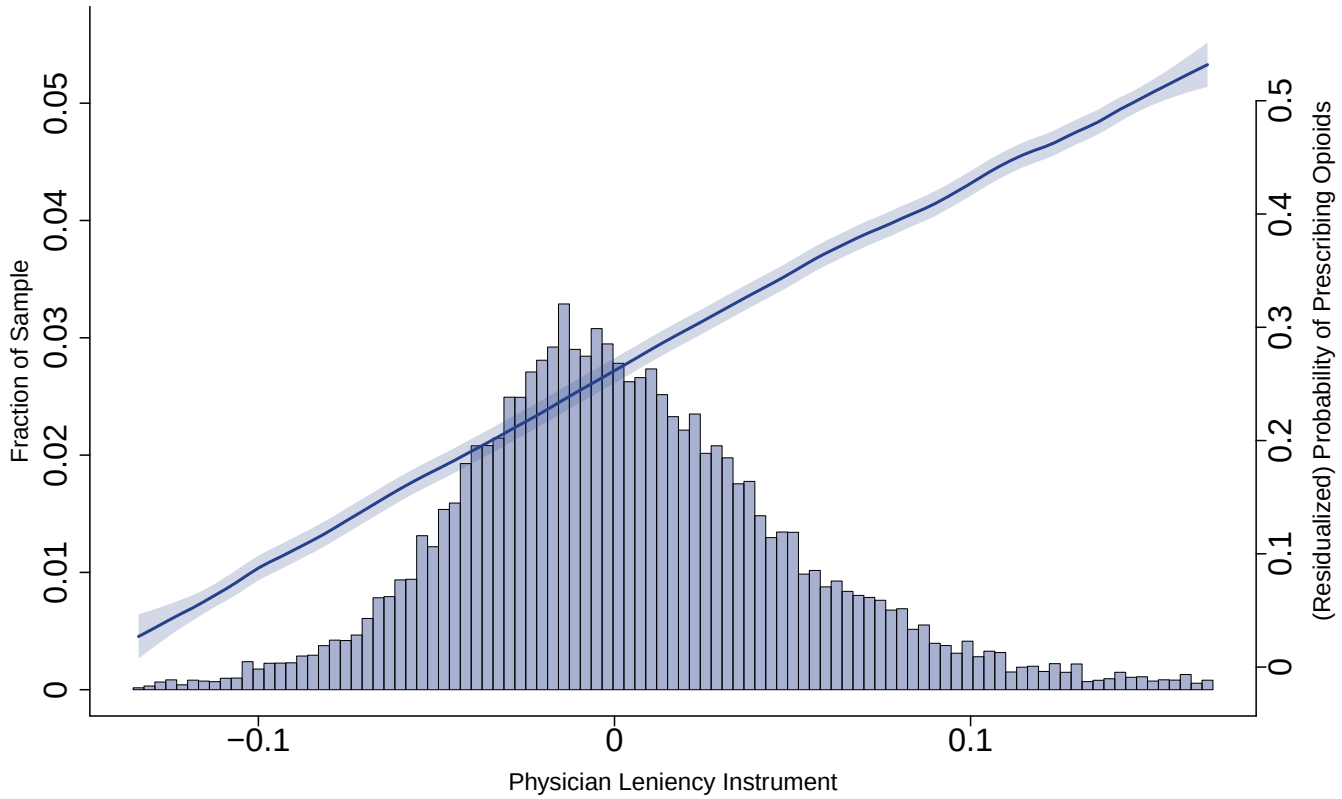
Notes: The empirical cumulative distribution function for days supply and milligrams of morphine equivalence per day, conditional on being prescribed, by year.

Figure 2: Frequent Diagnoses Occurring in Emergency Departments



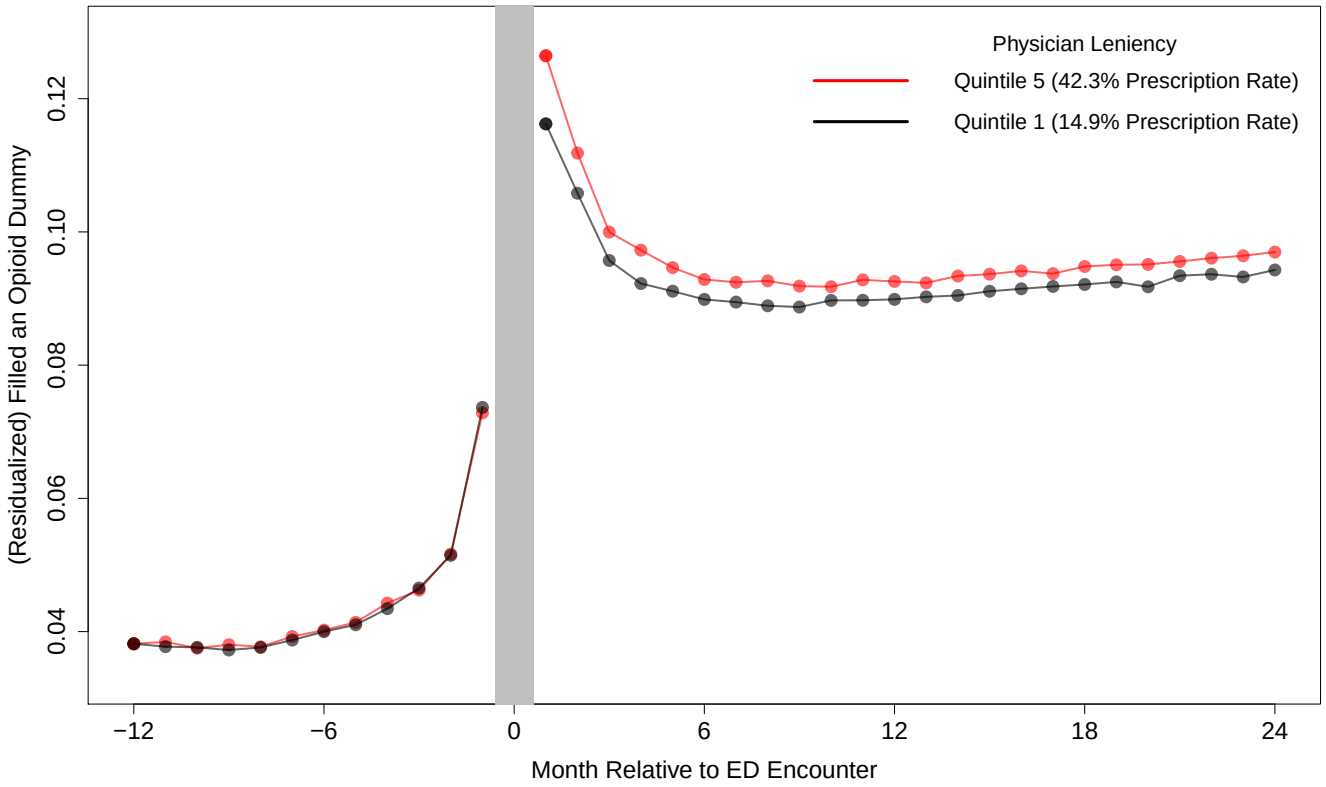
Notes: The 15 most common major diagnosis categories (ICD-9 major chapters) for ED visits and the un-adjusted rate they are prescribed opioids in our baseline sample. Rarely-prescribed categories are excluded, see the text for details.

Figure 3: Distribution and First Stage of Instrument



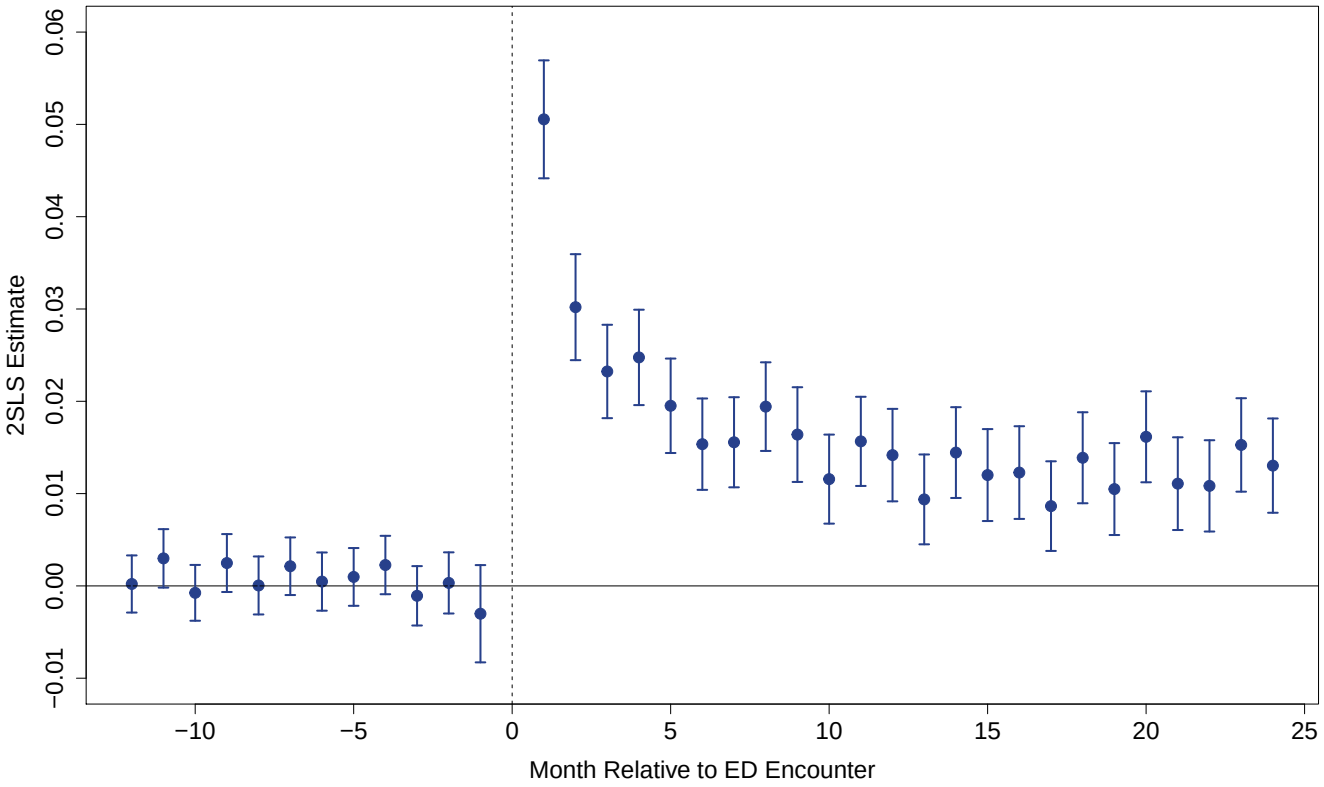
Notes: This figure plots the histogram of the baseline instrument along the x-axis and the left y-axis. A local-linear regression of the fitted probability of prescribed opioids on the instrument after residualizing (see text for baseline fixed effects and controls in residualization) is overlaid and displayed on the right y-axis. 95% confidence bands are also shown.

Figure 4: Reduced Form Event-Study for Opioid Use



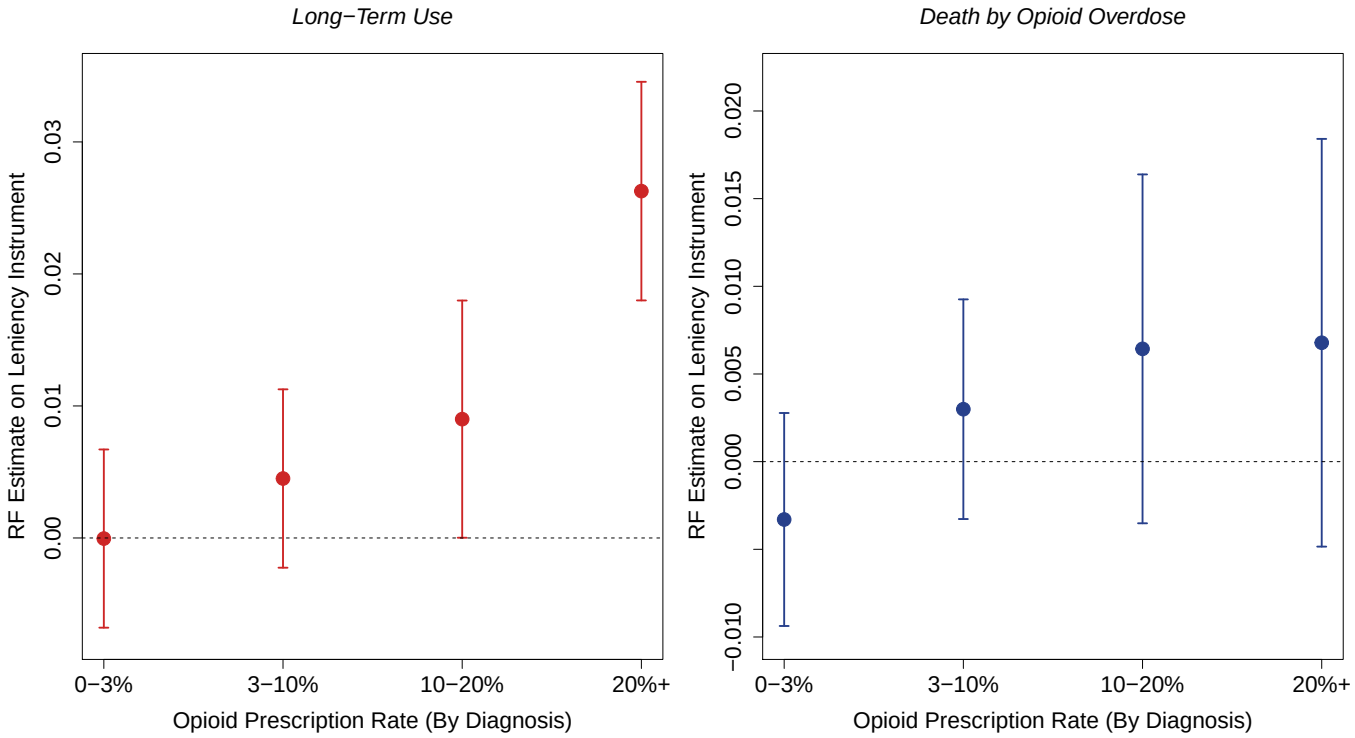
Notes: This figure plots the reduced form, event-study of opioid use on leniency of physician. Monthly residualized opioid use indicators (see text for baseline fixed effects and controls in residualization), for patients who see a strict and lenient physician are shown from one year pre-ED to 2 years post-ED. Only new filled opioid prescriptions contribute to each point and the initial ED prescription is omitted.

Figure 5: Two-Stage Least Squares Estimates for Opioid Use



Notes: The two-stage least squares version of Figure 4. Each point is an estimate from a separate regression of opioid use dummy that month on $Prescribed_i$ (the omitted opioid use dummy at time zero), instrumented with physician leniency. Controls are described in the text. 95% confidence intervals constructed using robust standard errors clustered at the physician level are also displayed. Only new filled opioid prescriptions contribute to each point and the initial ED prescription is omitted.

Figure 6: Reduced Form Coefficients by Diagnosis Bins Based on Opioid Prescription Rate



Notes: This figure plots the coefficient and standard error bars from the reduced form regression of long-term use (left panel) or opioid overdose death (right panel) on the instrument, physician leniency, by diagnosis bins. The diagnosis bins are categorized based on the unconditional probability each diagnosis is prescribed an opioid. Baseline controls as described in the text are included.

Tables

Table 1: Summary Statistics

	Mean	Q1	Median	Q3
Panel A: Prescription Characteristics				
Prescribed	0.261			
Prescribed | ED visit in 2011	0.285			
Prescribed | ED visit in 2016	0.229			
Days Supply | Prescribed	8.2	4	5	10
Daily Milligrams of Morphine | Prescribed	21.0	15	20	25
Opioid: Hydrocodone | Prescribed	0.576			
Opioid: Tramadol | Prescribed	0.165			
Opioid: Oxycodone | Prescribed	0.127			
Panel B: Emergency Department Characteristics				
Patient is admitted	0.083			
ED Diagnosis: Musculoskeletal and Connective Tissue	0.311			
ED Diagnosis: Injury and Poisoning	0.202			
ED Diagnosis: Symptoms, Signs and Ill-Defined Diseases	0.116			
ED Diagnosis: Skin and Subcutaneous Tissue	0.094			
ED Diagnosis: Digestive System	0.07			
Panel C: Veteran Characteristics				
Age	55.1	44	57	66
Income	21,009	4,068	15,000	31,128
Female	0.104			
White	0.69			
Black	0.24			
Married	0.43			
Panel D: Patient Medical History				
Prior Year Opioid Use Indicator	0.27			
Prior Month Opioid Use Indicator	0.10			
Prior Year Total MME	214.3	0	0	60
Prior Year Depression Diagnosis Indicator	0.241			
Prior Year PTSD Diagnosis Indicator	0.134			
Prior Year Homeless Indicator	0.062			
Elixhauser Comorbidity Index (3-year look-back)	1.2	-1	0	2
Observations: 1,958,209				

Notes: This table reports summary statistics for the baseline sample of emergency department visits between 2006-2016 described in the text. Panel A and B summarize characteristics related to the ED opioid prescription and ED visit. Panel C and D summarize patient veteran and patient medical history. All variables include VHA and CMS data when available.

Table 2: First Stage: Effect of Physician Leniency on ED Opioid Prescription

	Dependent Variable: Prescribed in ED		
	(1)	(2)	(3)
Physician Leniency	1.691***	1.702***	1.710***
	(0.012)	(0.012)	(0.012)
Hospital, Seasonality, Shift FE?	Yes	Yes	Yes
Diagnosis and Elixhauser?	No	Yes	Yes
Patient Observables?	No	No	Yes
F-Stat	12	21	25
Observations:	1,958,209	1,958,209	1,958,209

Notes: Estimates of the first stage for the baseline sample described in the text. Hospital, seasonality, shift fixed effects include Hospital-Year-Month and Hospital-Day of week-Hour of day fixed effects. Elixhauser comorbidity is constructed with a 3-year look-back period, excluding the ED encounter. Patient observables include female, black, prior month opioid use, age bins, and log prior year total milligrams of morphine equivalent. Column 3 corresponds to the baseline controls. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table 3: Balance Test for Random Assignment (Conditional Independence)

	Dependent variable:	
	Prescribed	Physician Leniency
Age	−0.002*** (0.0001)	−0.0001 (0.0001)
Female	−0.046*** (0.003)	−0.002 (0.003)
White	0.049*** (0.002)	0.002 (0.002)
Married	0.013*** (0.002)	−0.001 (0.002)
Elixhauser (3 year look-back)	−0.002*** (0.0001)	0.0004*** (0.0002)
Pain Score	0.059*** (0.001)	−0.001** (0.0003)
Log Income	−0.0003 (0.0002)	0.0003 (0.0002)
Service Connected Disability (%)	−0.0004*** (0.0001)	0.00002 (0.0001)
Period of Service: Korean War	−0.022*** (0.005)	0.004 (0.005)
Period of Service: Other	−0.017*** (0.006)	0.002 (0.005)
Period of Service: Post-Korean	−0.005 (0.005)	0.005 (0.005)
Period of Service: Post-Vietnam	0.027*** (0.003)	−0.001 (0.003)
Period of Service: Vietnam	0.020*** (0.003)	0.002 (0.003)
Period of Service: WWII	−0.047*** (0.006)	0.008 (0.006)
Priority Group 2	−0.016*** (0.004)	−0.004 (0.004)
Priority Group 3	−0.018*** (0.006)	0.002 (0.005)
Priority Group 4	−0.070*** (0.009)	0.003 (0.009)
Priority Group 5	−0.0002 (0.007)	−0.001 (0.006)
Priority Group 6	−0.032*** (0.007)	−0.003 (0.007)
Priority Group 7	−0.018** (0.008)	−0.001 (0.007)
Priority Group 8	−0.017** (0.007)	0.001 (0.006)
Previous Year Opioid Use	0.036*** (0.002)	−0.0002 (0.002)
Previous Month Opioid Use	−0.197*** (0.006)	0.006 (0.006)
Visit Number 2-5	−0.021*** (0.002)	0.002 (0.002)
Visit Number 6-10	−0.045*** (0.003)	0.007** (0.004)
Visit Number 11-15	−0.057*** (0.007)	0.017** (0.007)
Visit Number 16-20	−0.083*** (0.012)	0.039*** (0.012)
Previous Year Opioid Overdose	−0.017 (0.023)	0.014 (0.022)
Previous Year Suicide	−0.057*** (0.009)	−0.005 (0.010)
Previous Year Homelessness	−0.059*** (0.004)	0.003 (0.003)
Previous Year Depression	−0.014*** (0.002)	−0.0001 (0.002)
Previous Year PTSD	0.001 (0.002)	0.003 (0.002)
Previous Year Fall	0.036*** (0.004)	−0.001 (0.005)
Blood and Blood-Forming Organs	−0.015 (0.034)	0.123*** (0.042)
Circulatory System	−0.085*** (0.007)	−0.039*** (0.010)
Digestive System	0.206*** (0.006)	−0.007 (0.008)
Endocrine, Metabolic, and Immunity	0.079*** (0.010)	−0.006 (0.013)
External	0.042*** (0.007)	−0.008 (0.011)
Genitourinary System	0.371*** (0.008)	0.010 (0.007)
Infectious and Parasitic	0.429*** (0.012)	−0.007 (0.012)
Injury and Poisoning	0.250*** (0.005)	0.002 (0.009)
Mental Disorders	−0.087*** (0.026)	−0.022 (0.031)
Musculoske. and Connective Tissue	0.260*** (0.006)	−0.028*** (0.009)
Neoplasms	0.031*** (0.007)	0.007 (0.014)
Nervous System and Sense Organs	−0.048*** (0.005)	−0.003 (0.010)
Other	−0.008 (0.008)	0.012 (0.022)
Respiratory System	−0.054*** (0.013)	0.019 (0.018)
Skin and Subcutaneous Tissue	−0.017*** (0.005)	0.005 (0.008)
Joint-F Stat	279.9	2.4

Notes: This table tests for random assignment of patients to physicians in EDs for our baseline sample. The first column regresses prescribed opioid indicator on all observables, jointly (patient demographics, ED visit variables, and previous medical history), and the second column is the same regression but with the physician leniency instrument as the dependent variable. Both dependent variables are standardized. Construction of instrument is described in the text. Residualization fixed effects include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins. The joint F-statistics are reported. The number of observations is $1,672,553$ for both regressions. See Figure F.4 for a graphical version of this table. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table 4: Testing the Monotonicity Assumption

First Stage:	*Dependent variable:* Prescribed Opioid	
	Baseline Leniency	Reverse-Sample Leniency
Subsample:	(1)	(2)
Male	1.696***	1.046***
	(0.012)	(0.014)
Female	1.758***	1.917***
	(0.027)	(0.035)
Black	1.836***	1.951***
	(0.02)	(0.028)
White	1.649***	1.378***
	(0.013)	(0.016)
Opioid-Naïve	1.745***	1.524***
	(0.015)	(0.024)
Prior Users	1.584***	1.628***
	(0.016)	(0.024)
No Depression or PTSD	1.688***	1.608***
	(0.014)	(0.022)
Depression or PTSD	1.74***	1.839***
	(0.015)	(0.019)
Priority Groups 1-4	1.738***	1.851***
	(0.014)	(0.019)
Priority Groups 5-8	1.695***	1.733***
	(0.014)	(0.018)
Injury and Poisoning	1.714***	1.905***
	(0.028)	(0.039)
Musculoskeletal & Connective Tissue	2.267***	2.845***
	(0.023)	(0.043)
Digestive System	1.683***	1.807***
	(0.035)	(0.039)
Circulatory System	0.711***	0.746***
	(0.088)	(0.095)

Notes: Column 1 displays the first stage coefficient of prescribed opioid on the baseline physician leniency instrument for the corresponding subsample. Column 2 constructs a new physician leniency instrument using *all* emergency visits, excluding the corresponding subsample, and displays the coefficient of the first stage regression back on that subsample. Hence the name "reverse-sample". Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table 5: Subsequent Opioid Use and Long-Term Use Outcomes

	Mean	OLS		2SLS		N=
Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)
Long-Term Use (×100)						
	5.8	2.85***	2.63***	1.32***	1.17***	1,879,150
		(0.070)	(0.061)	(0.209)	(0.202)	
Total Milligrams of Morphine						
Prior Year	214	0.39	−12.94***	12.35**	3.57	1,958,209
		(1.51)	(0.66)	(5.08)	(3.01)	
Year 0-1	759	374.3***	327.6***	216.5***	202.8***	1,879,150
		(8.2)	(7.2)	(24.0)	(23.1)	
Year 0-2	1573	704.1***	600.1***	385.6***	360.3***	1,825,450
		(16.3)	(14.4)	(47.2)	(45.6)	
Positive Opioid Urine Drug Screen (×100)						
Prior Year	16.2	0.92***	−0.38**	0.68*	0.75	239,519
		(0.20)	(0.19)	(1.03)	(1.00)	
Year 2-3	25.6	4.37***	3.52***	1.94*	1.99**	273,115
		(0.21)	(0.21)	(1.03)	(1.00)	
Residualization FE?	-	Yes	Yes	Yes	Yes	-
Baseline Controls?	-	No	Yes	No	Yes	-

Notes: This table reports the effect of an opioid prescription on long-term use (180 days supply in the first 12 months), milligrams of morphine equivalent filled, and positive urine drug screens for opioids (conditional on being tested). Column 1 reports the mean dependent variable, columns 2 and 3 report ordinary least squares regression estimates, and columns 4 and 5 report the two-stage least squares IV regression estimates. All outcomes exclude the initial ED prescription and any prescriptions filled in the first 7 days after the emergency visit. Residualization fixed effects include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins; see text for baseline controls. The samples are constrained such that the patients are alive for the entire period the outcome is measured. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table 6: Proxies for Opioid-Seeking Behavior and Self-Reported Pain Score

		OLS		2SLS		N=
Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)
Overlapping Prescriptions (×100)						
Prior Year	2.9	-0.09^{**}	0.07^{**}	0.17	0.20	1,840,595
		(0.04)	(0.03)	(0.15)	(0.14)	
Year 0-1	9.9	4.4^{***}	4.1^{***}	1.9^{***}	1.9^{***}	1,840,595
		(0.1)	(0.1)	(0.2)	(0.2)	
Pharmacy Shopping (×100)						
Prior Year	0.2	0.07^{***}	0.09^{***}	0.06	0.06	1,840,595
		(0.01)	(0.01)	(0.04)	(0.04)	
Year 0-1	0.6	0.45^{***}	0.43^{***}	0.30^{***}	0.30^{***}	1,840,595
		(0.02)	(0.02)	(0.07)	(0.07)	
Repeated Back Pain & Headaches (×100)						
Prior Year	3.8	-0.24^{***}	-0.53^{***}	0.23	0.30	1,532,610
		(0.04)	(0.04)	(0.21)	(0.20)	
Year 0-1	6.2	0.83^{***}	0.38^{***}	0.47^{*}	0.55^{***}	1,532,610
		(0.06)	(0.05)	(0.27)	(0.27)	
Self-Reported Pain Score						
Prior Year	2.39	0.17^{***}	0.18^{***}	0.02	0.03	1,501,244
		(0.01)	(0.005)	(0.02)	(0.02)	
Year 0-1	2.81	0.36^{***}	0.18^{***}	0.05^{**}	0.08^{***}	1,682,968
		(0.01)	(0.005)	(0.02)	(0.02)	
Residualization FE?	-	Yes	Yes	Yes	Yes	-
Baseline Controls?	-	No	Yes	No	Yes	-

Notes: This table reports of effects of an opioid prescription on proxies for opioid-seeking behavior, and self-reported pain scores for the prior year and the 1 year period following the emergency visit. The proxies are defined in the text. Self-reported pain score is the mean across all outpatient settings in a given year. Column 1 reports the mean dependent variable, columns 2 and 3 report ordinary least squares regression estimates, and columns 4 and 5 report the two-stage least squares IV regression estimates. Residualization fixed effects include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins; see text for baseline controls. The samples are constrained such that the patients are alive for the entire period the outcome is measured. Sample for repeated back pain and headaches exclude emergency visits that were for back pain or headaches. Regression coefficients, standard errors, and mean dependent variables are scaled as indicated. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table 7: Opioid Use Disorder, Opioid Overdose, and Opioid Overdose Mortality

	Mean	OLS		2SLS		N=
Dependent Variable (×100):	(1)	(2)	(3)	(4)	(5)	(6)
Opioid Use Disorder (OUD)						
Prior Year	1.56	−0.338***	−0.493***	0.033	0.041	1,958,209
		(0.024)	(0.025)	(0.106)	(0.105)	
Year 0-1	1.98	−0.070**	−0.262***	0.246*	0.255**	1,879,150
		(0.029)	(0.030)	(0.126)	(0.126)	
Year 0-3	3.27	0.318***	0.013	0.323*	0.335*	1,775,800
		(0.039)	(0.039)	(0.161)	(0.160)	
Opioid Overdose						
Prior Year	0.11	−0.008	−0.019***	0.015	0.013	1,958,209
		(0.006)	(0.006)	(0.030)	(0.030)	
Year 0-1	0.23	0.027***	0.005	0.091**	0.088**	1,879,150
		(0.009)	(0.009)	(0.043)	(0.043)	
Year 0-3	0.59	0.082***	0.026*	0.044	0.039	1,775,800
		(0.015)	(0.015)	(0.069)	(0.069)	
Opioid Overdose Death						
	0.167	0.065***	0.044***	0.074**	0.075**	1,846,133
		(0.008)	(0.008)	(0.034)	(0.034)	
Residualization FE?	-	Yes	Yes	Yes	Yes	-
Baseline Controls?	-	No	Yes	No	Yes	-

Notes: This table reports the effect of an opioid prescription on opioid use disorder (OUD) and opioid overdose at various points in time, and eventual death from an opioid overdose. Column 1 reports the mean dependent variable, columns 2 and 3 report ordinary least squares regression estimates, and columns 4 and 5 report the two-stage least squares IV regression estimates. Residualization fixed effects include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins; see text for baseline controls. The samples for the OUD and overdose regressions are constrained such that the patients are alive for the entire period the outcome is measured. Regression coefficients, standard errors, and mean dependent variables are scaled as indicated. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table 8: Opioid Overdose Mortality by Type of Opioid

	Mean	OLS		2SLS		N=
Dependent Variable (×100):	(1)	(2)	(3)	(4)	(5)	(6)
Type of Opioid:						
Heroin	0.055	0.016***	0.009**	0.022	0.022	1,846,133
		(0.005)	(0.005)	(0.019)	(0.019)	
Synthetic (excl. methadone)	0.038	0.010***	0.006	0.020	0.020	1,846,133
		(0.004)	(0.004)	(0.017)	(0.017)	
Heroin or Synthetic Opioids	0.084	0.026***	0.017***	0.030	0.030	1,846,133
		(0.006)	(0.006)	(0.024)	(0.023)	
Natural Opioids *Only*	0.050	0.027***	0.020***	0.034*	0.034*	1,846,133
		(0.005)	(0.005)	(0.020)	(0.020)	
Residualization FE?	-	Yes	Yes	Yes	Yes	-
Baseline Controls?	-	No	Yes	No	Yes	-

Notes: This table reports the effect of an opioid prescription on opioid overdose mortality by type of opioid. ICD-10 mortality codes: heroin (T40.1); synthetic opioids excluding methadone (T40.4); natural opioids only (T40.2 and no other opioid type). Column 1 reports the mean dependent variable, columns 2 and 3 report ordinary least squares regression estimates, and columns 4 and 5 report the two-stage least squares IV regression estimates. Residualization fixed effects include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins; see text for baseline controls. Regression coefficients, standard errors, and mean dependent variables are scaled as indicated. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table 9: Health and Non-Opioid Outcomes

	Mean	OLS		2SLS		N=
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Fall (×100)						
Prior Year	5.4	0.015	−0.042	0.137	0.120	1,958,209
		(0.041)	(0.041)	(0.221)	(0.218)	
Years 0-3	8.4	0.299***	0.052	0.552**	0.504**	1,775,800
		(0.052)	(0.051)	(0.262)	(0.257)	
Depression (×100)						
Prior Year	24.3	-0.143	−1.011***	0.268	0.433	1,958,209
		(0.087)	(0.083)	(0.422)	(0.411)	
Years 0-3	36.4	1.324***	0.195**	0.242	0.427	1,775,800
		(0.101)	(0.095)	(0.471)	(0.459)	
Suicide Attempt/Self-Harm (×100)						
Prior Year	0.65	−0.098***	−0.116***	−0.019	−0.019	1,958,209
		(0.014)	(0.015)	(0.079)	(0.079)	
Years 0-3	1.59	0.030	−0.038	−0.027	−0.026	1,775,800
		(0.024)	(0.024)	(0.117)	(0.117)	
Homeless (×100)						
Prior Year	6.2	−0.842***	−1.078***	0.374*	0.414*	1,958,209
		(0.048)	(0.050)	(0.224)	(0.221)	
Years 0-3	11.9	−0.408***	−0.798***	0.174	0.201	1,775,800
		(0.068)	(0.068)	(0.297)	(0.295)	
Residualization FE?	-	Yes	Yes	Yes	Yes	-
Baseline Controls?	-	No	Yes	No	Yes	-

Notes: This table reports the effects of an opioid prescription on fall, depression, attempted suicide/self-harm and homeless dummies for the prior year and the 3-year period following the emergency visit. Column 1 reports the mean dependent variable, columns 2 and 3 report ordinary least squares regression estimates, and columns 4 and 5 report the two-stage least squares IV regression estimates. Residualization fixed effects include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins; see text for baseline controls. The samples are constrained such that the patients are alive for the entire period the outcome is measured. Regression coefficients, standard errors, and mean dependent variables are scaled as indicated. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table 10: Patient Questionnaire Outcomes for Mental Health and Substance and Alcohol Use

	Prior Year		Year 0-3	
Dependent variable (×100)	Coef (1)	Mean Dep. Var. (2)	Coef (3)	Mean Dep. Var. (4)
Depression screen	0.121 (0.213)	5.89	0.571* (0.325)	14.7
Any drug *abuse* (illicit or prescription)†	0.063 (0.050)	0.224	0.156* (0.088)	0.776
Opiate *use*†	0.003 (0.028)	0.077	0.066 (0.054)	0.274
Cocaine/Crack *use*†	0.026 (0.034)	0.088	0.143** (0.057)	0.317
Sedatives *use*† (e.g., benzodiazepines)	−0.016 (0.023)	0.049	0.061 (0.042)	0.172
Other stimulant *use*† (e.g., amphetamines)	−0.015 (0.021)	0.047	0.047 (0.040)	0.155
Marijuana *use*†	0.023 (0.038)	0.134	0.004 (0.068)	0.471
Alcohol screen	0.484** (0.217)	9.46	0.214 (0.355)	20.5
N=	1,958,209		1,775,800	
†: in the past 30 days				

Notes: This table reports the effect of an opioid prescription on mental health questionnaire outcomes. All outcomes are binary and veterans who do not take the questionnaire are coded as zero. Depression questionnaire measured with the Patient Health Questionnaire: PHQ-2 and PHQ-9. Its goals are to assess for depression as a "first-step approach"; scores of and above 3 and 15 are coded as positive screens. Drug abuse and use outcomes are from the Brief Addiction Monitor (BAM and BAM-R) questions 6 and 7. The questions ask "In the past 30 days, how many days did you use..." and all non-zero answers are coded as positive. Finally, alcohol screen is based on the AUDIT-C which identifies "hazardous drinkers or active alcohol use disorders"; scores of 4 or greater are coded as positive screens. All regressions include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins fixed effects and standard baseline controls. The samples are constrained such that the patients are alive for the entire period the outcome is measured. Regression coefficients, standard errors, and mean dependent variables are scaled as indicated. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table 11: Proxies for Illicit Drug Use

	Mean	OLS		2SLS		N=
Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)
Intended Heroin/Fentanyl Drug Screen (×100)						
Prior Year	0.10	−0.012**	−0.023***	−0.038	−0.039	1,958,209
		(0.006)	(0.007)	(0.030)	(0.029)	
Year 0-3	0.70	0.166***	0.113***	0.023	0.015	1,775,800
		(0.020)	(0.020)	(0.073)	(0.072)	
Positive Hepatitis C Diagnosis (×100)						
Prior Year	4.3	0.182**	−0.020	0.211	0.177	1,958,209
		(0.038)	(0.038)	(0.177)	(0.175)	
Year 0-3	5.9	0.411***	0.127***	0.328	0.259	1,775,800
		(0.046)	(0.046)	(0.210)	(0.209)	
Residualization FE?	-	Yes	Yes	Yes	Yes	-
Baseline Controls?	-	No	Yes	No	Yes	-

Notes: This table reports of effects of an opioid prescription on proxies for illicit drug use for the prior year and the 3-year period following the emergency visit. The proxies are defined in the text. See text for definition and rationale for proxies. Column 1 reports the mean dependent variable, columns 2 and 3 report ordinary least squares regression estimates, and columns 4 and 5 report the two-stage least squares IV regression estimates. Residualization fixed effects include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins; see text for baseline controls. The samples are constrained such that the patients are alive for the entire period the outcome is measured. Regression coefficients, standard errors, and mean dependent variables are scaled as indicated. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table 12: Heterogeneous Effects of a Single Opioid Prescription on Outcomes

	Heterogeneity Margin:							
	Opioid-Naïve?		PCP Leniency		Year		Gender	
	Naïve	Prior User	Lenient	Strict	2006-11	2012-16	Male	Female
Outcome:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Long-Term Use	1.217***	1.328**	1.297***	0.762**	1.375***	0.979***	1.193***	1.050*
	(0.165)	(0.648)	(0.491)	(0.381)	(0.312)	(0.264)	(0.215)	(0.547)
p-value	0.964		0.445		0.414		0.791	
Opioid Use Disorder	0.228	0.736*	0.651*	0.307	0.520**	0.109	0.283	0.566
	(0.186)	(0.414)	(0.358)	(0.322)	(0.239)	(0.266)	(0.190)	(0.442)
p-value	0.302		0.534		0.306		0.665	
Opioid Overdose Death	0.057	0.118	−0.098	0.185***	0.113*	0.032	0.069*	0.102
	(0.035)	(0.088)	(0.065)	(0.059)	(0.061)	(0.034)	(0.037)	(0.092)
p-value	0.811		0.011		0.362		0.828	
Hepatitis C	0.218	0.433	0.181	0.315	−0.109	0.687***	0.263	0.558
	(0.229)	(0.493)	(0.410)	(0.394)	(0.334)	(0.249)	(0.229)	(0.374)
p-value	0.732		0.839		0.097		0.712	
Observations	1,351,472	424,328	411,791	484,931	982,671	793,129	1,577,038	198,762

Notes: This table reports the effect of an opioid prescription on the main outcomes along 4 margins: previous year opioid use, PCP prescribing leniency, pre/post 2012, and gender. The 2SLS regressions are estimated on each subsample separately, and the p-value associated with testing the difference between the coefficients are reported below each set regression coefficients. Residualization fixed effects include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins; see text for baseline controls. The samples are constrained such that the patients are alive for the entire period the outcome is measured. Regression coefficients, standard errors, and mean dependent variables are scaled as indicated. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table 13: Alternate Specifications and Robustness Checks

	2SLS Estimates					
	Main Baseline	Team Leniency	Endog. Admit & Procedures	Physician Quality	Intensive Margin	MDC IV
Dep. Var. (×100):	(1)	(2)	(3)	(4)	(5)	(6)
Long-Term Use	1.172***	1.313***	1.165***	1.104***	1.092***	1.921***
	(0.202)	(0.305)	(0.202)	(0.203)	(0.201)	(0.258)
OUD (Year 0-3)	0.335*	0.375	0.308*	0.344**	0.321**	0.232
	(0.160)	(0.237)	(0.160)	(0.161)	(0.176)	(0.182)
Overdose (Year 0-1)	0.088**	0.092	0.085**	0.091**	0.071*	0.132***
	(0.043)	(0.067)	(0.043)	(0.044)	(0.043)	(0.049)
Opioid Overdose Death	0.075**	0.104*	0.076**	0.077**	0.074***	0.066
	(0.034)	(0.055)	(0.034)	(0.034)	(0.035)	(0.042)
Fall (Year 0-3)	0.504**	0.525	0.468*	0.488*	0.521***	0.611**
	(0.257)	(0.393)	(0.255)	(0.258)	(0.256)	(0.281)
Hepatitis C (Year 0-3)	0.259	0.011	0.257	0.240	0.322	0.444*
	(0.207)	(0.321)	(0.207)	(0.209)	(0.205)	(0.244)
Observations	1,775,800	1,775,800	1,775,800	1,739,337	1,775,800	982,679

Notes: This table reports 2SLS regression coefficients of *Prescribed* on the main outcomes for the main baseline empirical model in column (1) and various alternate specifications in columns (2)-(6). Column (2) uses team leniency as the instrumental variable. Column (3) includes predicted propensity to admit (hospitalize) and intensity of procedure (measured with w-RVUs) as controls to the baseline via an indirect least squares. Column (4) constructs a proxy for physician quality analogous to propensity to prescribe, but replacing opioid prescription indicator with 1 month mortality indicator. Column (5) includes an intensive margin (total milligrams of morphine) endogenous variable and instrument, and evaluates the average treatment effect at the mean ED opioid prescription (152mg of morphine). Column (6) uses physician-diagnosis-year leniency instruments. See subsection 4.4 for more details. The samples are constrained such that the patients are alive for the entire period the outcome is measured. Observations reported are for death by year 3; they differ slightly by outcome. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

For Online Publication

Appendices

A. Comparison with Barnett, Olenski, and Jena (2017) and Barnett et al. (2019)

We would like to begin by thanking Michael L. Barnett, Walid Gellad, Anupam B. Jena, and their coauthors for their suggestions, comments, and clarifications. In this appendix, we describe the differences between our paper and Barnett, Olenski, and Jena (2017) and Barnett et al. (2019), as well as study how and where our findings depart from theirs.

The two papers listed above are the two most closely related to ours. In Barnett, Olenski, and Jena (2017), the authors study long-term use (180 days supply in 12 months) following an ED visit for opioid-naïve Medicare beneficiaries who see a high or low intensity prescriber. With a 20% random sample of Medicare claims from 2008-2011, physicians are classified as high (low) intensity if their overall prescription rate over those four years falls in the top (bottom) quartile within their hospital. The authors find that being treated by a high intensity prescriber is associated with a 0.35pp (30%) increase in the probability of long-term use. They also study a set of secondary outcomes including hospitalizations, ED visits, falls or fractures, constipation, respiratory failure, and opioid poisoning in the following year. They find higher rates of falls or fractures and opioid poisoning associated with high intensity prescribing.

Barnett et al. (2019) use 2012 VHA data to replicate their previous study (identical sample selection and research design) and find a 0.13pp (11%) increase in the probability of long-term opioid use among veterans. They study the same secondary outcomes and fail to find any statistically significant difference.

Differences Between the Papers

The key differences between these two papers and ours can be grouped into two categories: i) patient outcomes, and ii) econometric specification and sample construction. In terms of patient outcomes, both Barnett, Olenski, and Jena (2017) and Barnett et al. (2019) focus primarily on long-term prescription opioid use (180 day supply in the first year after the ED visit) as their main outcome, along with opioid-related hospitalizations such as falls, fractures, and poisonings as secondary outcomes. Our paper studies additional long-term outcomes including opioid use disorder, proxies for opioid-seeking behavior, mortality, depression, homelessness, attempted suicide, and proxies for illicit opioid use. In addition, we supplement the legal observed VHA opioid prescriptions in Barnett et al. (2019) with Medicare and Medicaid claims and attempt to account for unobserved black market opioids.

Econometrically, Barnett, Olenski, and Jena (2017) and Barnett et al. (2019) classify emergency physicians as high and low "intensity" prescribers, similar in spirit to our "leniency" instrument. They do this by first calculating each physician's raw opioid-prescribing rate as the number of emergency visits resulting in a prescription, divided by the total number of emergency visits. They construct one aggregate rate (lumping all years together in the 2017 paper) per physician. They then classify physicians as high (low) intensity prescribers if they fall in the top (bottom) quartile within their hospital.

Our paper utilizes a residualization approach, as described in subsection 3.2, leveraging detailed information about time of day, day of week, age, diagnosis, and pain score, thus eliminating some selection of patient arrival to ED or physician work schedules. Further, we leave out patient-physician pairs' own residual, eliminating the mechanical bias that stems from a patient's own case entering into the instrument. When the number of cases observed for each physician is small, this bias is large and approaches the OLS bias. Our leniency measure is also year-varying, allowing for physicians learning about the risks and benefits of prescription opioids during this time period.

The papers differ in terms of our sample selection as well. Barnett, Olenski, and Jena (2017) focus on all non-admitted emergency department conditions (diagnoses) of opioid-naïve patients between the years 2008 and 2011. Barnett et al. (2019) focus on VHA emergency

department and urgent care clinic visits in 2012. We are not as restrictive regarding prior opioid use, excluding only the top 15th percentile (3,150 mg of morphine in the prior year). However, we are more restrictive regarding conditions, excluding diagnoses that are rarely prescribed (anything less than a 10% prescription rate). Our study years also do not align; we focus on 2006-2016. This affects the interpretation of the estimates. Their estimates are for "new" opioid users following their first opioid prescription, whereas our estimates are for one (additional) prescription for veterans who come to the ED for particular conditions.

Reconciling the Differences in Long-Term Use Estimates

In this section, we investigate how the differences in the studied samples, and in methods in measuring prescriber intensity affect the estimate on long-term prescription opioid use (the only shared outcome studied in both their papers and ours). We begin by replicating Barnett et al. (2019), then we make incremental changes to the sample construction, eventually ending up at the baseline sample studied in this paper. We do this all while keeping the high/low intensity classification based on a physician's opioid prescription rate within a facility, as in their papers. Then, we move to our residualization approach as described in subsection 3.2, also incrementally including more controls, finally arriving at the estimate reported in this paper. With each incremental step, we report the mean long-term prescription opioid use associated with high and low intensity physicians, the ratio between the two (odds ratio), and the Wald estimate (an analog to the 2SLS estimate but with a binary high vs. low "instrument" to aid in comparison and interpretation with Barnett, Olenski, and Jena (2017) and Barnett et al. (2019)).

Table F.8 reports the result of this exercise. The first three columns of row 1 are taken directly from Barnett et al. (2019); the Wald estimate[37] (column 4) of 0.903. Column 2 is our best attempt at replicating their main finding. The odds ratio and Wald estimate are very similar; however, the base long-term use means are greater, presumably due to minor differences in data definitions. Next, we make incremental changes to the sample restrictions

[37]This wald estimate is called "number needed to harm" in Barnett et al. (2017). It is not reported in Barnett et al. (2019), but scaling their high vs. low long-term differences by their prescription rate, yields 0.903.

and data definitions to arrive at the baseline sample in this paper. High and low intensity physicians are classified by top and bottom quartile opioid prescribing rate, within a facility, after the corresponding sample restriction change. Some examples of such changes include: changing the definition of long-term opioid use to days supply of opioids filled[38] (row 3), excluding urgent care clinics (row 4), including admitted patients and some prior users (rows 7 and 8), excluding diagnosis conditions that are rarely prescribed (row 9), adding opioid prescriptions from Medicare and Medicaid (row 10), and including all years from 2006-2016 (row 11). Since these changes alter the relevant sample of veterans, they have varying effects on the Wald estimate. For example, including CMS opioid prescriptions increases the Wald estimate, implying that patients who see a more lenient ED physician, are also more likely to fill new opioid prescriptions through Medicare or Medicaid. With the within-facility intensity classification of Barnett, Olenski, and Jena (2017) and Barnett et al. (2019) on our baseline sample, we have a Wald estimate of 2.75 (column 4 of row 11), more than double the main effect reported in this paper. If we allow physician prescribing intensity to vary across years (i.e., top vs. bottom quartile within a facility-year; row 12), then the Wald estimate drops to 1.75, still 50% larger than our estimate of 1.17 with our residualization approach.

In the next four rows of Table F.8 (rows 13-16), with our baseline sample, we now classify physicians as high/low-intensity with our residualization approach, incrementally residualizing for additional covariates. The first level of residualization is at the hospital-year-month level. That is, we construct our physician leniency as described in subsection 3.2, but with only hospital-month fixed effects to control for hospital specific seasonality. We then select the top and bottom quartiles of prescribers per hospital based on their mean residuals. Finally, we compute the difference in (residualized) long-term use divided by (residualized) prescription rate—the Wald estimate—in column 4. By residualizing for hospital specific seasonality, the Wald estimator drops in magnitude substantially. This implies that much of the variation between physicians, even within a facility, is endogenous. The next three rows controls for "shift-level" variation in physician work schedule and patient arrival, diagnosis condition, and

[38]If a patient has two on-going opioid prescriptions with overlapping days, Barnett et al. (2019) do not count the overlapping days towards the 180 days supply needed to be classified as a long-term user, whereas we would count it overlapping days, because those opioid pills are available to be abused. Therefore, their measure of long-term use is days of opioids consumed, while ours is days of opioids available.

patient covariates including age, Elixhauser Comorbidity Index and pain score, finally arriving at a Wald estimate of 1.25. Recall that our baseline 2SLS estimate (with the continuous leniency instrument) was 1.17. This exercise implies that residualization in both the leniency construction and the second stage can yield different estimates.

Ranking Physicians by Prescribing Leniency Using Barnett et al. (2017, 2019) vs. Our Method

The comparison in the previous section teaches us that sample selection and physician leniency construction lead to differences in estimates of an ED prescription's effects on long-term use. Our long-term use probabilities are larger because they include some prior users and focus on diagnoses that are typically prescribed opioids. Moreover, even by keeping the sample fixed, the two empirical approaches used to construct prescribing leniency arrive at different estimates. The classification of lenient physicians hinges on patient diagnosis, age, risk, and time of arrival at the ED. Figure F.8 demonstrates this by graphing the reshuffling of prescribing ranking after controlling for said covariates for the Tampa VA Medical Center (the largest ED in 2012). Each physician (provided they have treated 30 cases) is sorted by his/her ranking after our residualization method on the x-axis. The y-axis represents their corresponding ranking using the Barnett et al. intensity measure. If both methods yield identical rankings, the physicians align perfectly on the dashed diagonal line. Next, we classify physicians as low and high intensity prescribers based on the top and bottom quartiles using either method. The blue squares correspond to physicians who are classified in the top or bottom quartile by both methods, and the red triangles correspond to physicians about whom the two methods disagree. The physicians at the tails of the distribution tend to be classified as top or bottom prescribers by both methods; however, there is substantial disagreement outside of the tails. There are 46 physicians whom both methods agree are either high or low intensity prescribers, and 34 who are classified by one method but not the other.

B. Identifying VHA Emergency Departments, Linking Opioid Prescriptions, and Identifying Primary Care PACT Visits

In this section we describe in detail how we identify VHA emergency visits, linking opioid prescriptions to its originating emergency department (what counts as prescribed), and identifying primary care PACT visits.

Emergency Departments

Emergency departments in the VHA were standardized beginning in 2006 with VHA Directive 2006-051 "Standards for Nomenclature and Operations in VHA Facility Emergency Departments". Therefore, we start looking for ED visits in 2006.

Emergency department visits are identified off VA stop codes. We do not consider urgent care centers are emergency departments. After March 2007, we use visits with primary stopcode of 130. Prior to March 2006, we use i) primary-secondary stopcode combination 102-101 OR ii) primary stopcode of 102 with an emergency department CPT procedure code. In addition, we require the visit to originate in a station number (`DivisionSID` that is listed as an emergency department (excluding facilities that have joint emergency and urgent care) in the 2007 Survey of Emergency Departments and Urgent Care Clinics in the VHA. Lastly, we also require emergency departments to have at least 5000 annual visits and non-negligible visit share between 12-4am, following VHA Directive 2006-051, which required emergency departments to operate 24 hours a day, seven days a week.

Opioid Prescriptions

Opioid prescriptions need to be linked back to its origin (i.e., was it from an emergency department or primary care clinic?). We employ the following algorithm in coding an emergency department as prescribed an opioid:

1. We restrict attention to opioid prescriptions that are written (`IssueDate` within a day of the emergency encounter.

2. If there is a perfect provider-prescriber ID match, we code the emergency encounter as $Prescribed = 1$.

3. If the prescription was written on the same day, or on the next day (provided the emergency visit happened after 8pm) and the facility ID (`DivisionSID` match, we code the emergency encounter as $Prescribed = 1$.

4. All other emergency cases are coded as $Prescribed = 0$.

We do not require a perfect provider match because the a patient may see more than one clinician in the ED, and the (head) attending physician may not be the prescriber name on the prescription. Out of the cases we code as *Prescribed*, 88% of them have a prescriber and provider ID match, and the other 12% that match on facility ID and date/time, we code the prescription as *Prescribed* by the attending physician for the purpose of constructing leniency. Here we are assuming that the attending physician influences the decision to prescribe and has oversight what other providers (e.g., nurse practitioner) are doing. Note that if a patient is admitted and prescribed an opioid following their hospitalization, the patient will be considered prescribed provided the prescription was written within a day of the emergency visit, and the prescription will be assigned to the emergency physician.

Primary Care PACT Teams

PACT teams are identified off historical patient provider relationships where the assignment and team purpose are primary care. We exclude specialty teams: substance, ambulance, dialysis, psych, mental, geriatric or extended care, neurology, injury. Primary care outpatient visits are identified off primary or secondary stopcodes 322, 323. Lastly, the patient-provider relationship need to be current at the time of the visit for it to be considered a PACT outpatient visit. A visit is coded as prescribed an opioid (for PCP leniency construction) if there is a perfect provider-prescriber and visit date-issue date match.

C. Who Are Lenient Opioid-Prescribing Physicians and How Do They Vary Along Other Dimensions?

In this appendix we summarize the characteristics of lenient physicians based on their observables, then correlate prescribing leniency with other physician dimensions along four margins: i) decision to admit a patient to an inpatient hospital, ii) decision to perform invasive procedures, iii) likelihood of causing a patient death within one month (proxy for physican quality), and iv) intensive margin decision regarding amount of opioids to prescribe, conditional on prescribing an opioid (based on total milligrams of morphine equivalent).

Table F.9 presents characteristics of lenient and strict physicians-years. Recall that our leniency measure is defined at the year level. Physicians are classified as lenient (strict) if they are in the top (bottom) quartile of our leniency measure each year. Lenient physicians are much more likely to be male and slightly older in age and on average work more. Lenient physicians on average work nine extra days per year compared to strict physicians. They also see more patients per day, however, it is unclear whether this is due to working longer shifts or working quicker. This could be in line with findings that physicians prescribe more opioids when they are busier and more fatigued (i.e., later in the workday or when appointments are running behind schedule as seen in Neprash and Barnett, 2019).

Next, we investigate how physician opioid-prescribing leniency correlates with other physician dimensions. In particular, we study dimensions that may violate our exclusion restriction. We study the graphical first-stage of our baseline physician prescribing leniency on four different dimensions: admission, intensity of procedures, physician quality, and the intensive margin opioid-prescribing decision (conditional on prescribing an opioid). All four proxies are discussed in greater detail in subsection 4.4. Figure F.9 displays these first-stage correlations over the histogram of opioid-prescribing leniency values. Across all four dimensions, there is a positive relationship with opioid-prescribing leniency in the main mass of the histogram. The relationships are generally small and often non-monotonic at the tails. For instance, the average physician in the top decile of prescribing leniency performs on average 1.628 w-RVU compared to 1.557 in the bottom decile, a 4.6% increase in payment if paid for by CMS. In terms of one-month mortality, for physicians in the top decile of

opioid-prescribing leniency, 0.634% of their ED patients die within a month, compared to 0.614% in the bottom decile. These effects are modest and we've shown in subsection 4.4 that they do not have significant effects on our findings.

D. Calculating and Characterizing Compliers

In this section we describe the method we use to calculate the share of compliers (and always-takers and never-takers), and its characteristics. The method follows Dahl et al. (2014) and Dobbie et al. (2018).

First, compliers are defined as patients who would not have been prescribed an opioid if they had been seen by the most strict physician, but would have been prescribed an opioid if they had been seen by the most lenient physician:

$$\pi_{complier} = P(D_{\bar{z}i} > D_{\underline{z}i} = E(D_{\bar{z}i} - D_{\underline{z}i} = P(D_i|Z_i = \bar{z} - P(D_i|Z_i = \underline{z}$$

where D_i represents the prescription decision for veteran i, Z_i represents the leniency of veteran i's physician, and $\bar{z}$ and $\underline{z}$ represent the most and least lenient physicians.

Similarly, always-takers are patients who would be prescribed an opioid by every physician:

$$\pi_{always-taker} = P(D_{\bar{z}i} = D_{\underline{z}i} = 1) = P(D_{\underline{z}i} = 1)$$

where the last step follows from the monotonicity assumption. Last, the share of never-takers (patients who would never be prescribed an opioid by any physician) is found by:

$$\pi_{never-taker} = P(D_{\bar{z}i} = 0)$$

By defining the most lenient physicians ($\bar{z}$) as physicians with a leniency instrument in the top percentile and the most strict physicians ($\underline{z}$) as physicians with a leniency instrument in the bottom percentile, we can calculate the share of compliers, always-takers, and never-takers from moments in the first stage. For instance, we fit a local linear regression of $Prescribed_i$ on physician leniency, take the share of veterans who are prescribed by the top percentile of leniency, and subtract the share of veterans who are prescribed by the bottom percentile of leniency.

We can also characterize our compliers by observable characteristics. For example, we can calculate the share of veterans who are prior users, conditional on being a complier. In particular, we can compute $P(X_i = x|complier)$:

$$
\begin{aligned}
P(X_i = x|complier) &= P(X_i = x|D_{\bar{z}i} > D_{\underline{z}i} \\
&= \frac{P(X_i = x \cap D_{\bar{z}i} > D_{\underline{z}i}}{P(D_i|Z_i = \bar{z} - P(D_i|D_i = \underline{z}} \\
&= \frac{P(D_{\bar{z}i} > D_{\underline{z}i}|X_i = x)P(X_i = x)}{\pi_{complier}} \\
&= \frac{\pi_{c|x}P(X_i = x)}{\pi_c}
\end{aligned}
$$

This moment is calculated by computing the share of compliers for the subsample $X_i = x$ (i.e., checking the moments of the first stage for that subsample) and scaling it by the unconditional share of that subsample, divided by the overall share of compliers. This is the second column in Table F.6.

E. Team (Across-Shift) Leniency

As mentioned in section 3.4 and section 4.4, the identification of our *within*-"shift" quasi-random assignment strategy breaks down if there is selection in patient-physician assignment along unobserved margins. Examples of such violations include such situations as senior physicians delegating difficult, frequent ED visitors who refuse to leave to newer physicians or physicians takeing cases of severe conditions on which they are experts. In such cases, assignment to a physician, say A vs. B, is non-random at a given point in time, t. However, if only physicians A and B are working at that ED at time t, we can use the average leniency of that "team" by utilizing the fact that at some other time t', physicians A and B have been replaced with physicians C and D. Specifically, as an alternate robustness strategy, we no longer rely on random assignment to patients conditional on showing up at the ED (*within*-"shift"), but leverage variation in the timing of their visit and the available personnel working at that ED (*across*-"shift").

For individual i arriving at the emergency department at time t, we define $s = [t-1h, t+1h]$, a two hour "shift" window, and the leniency of their potential physician:

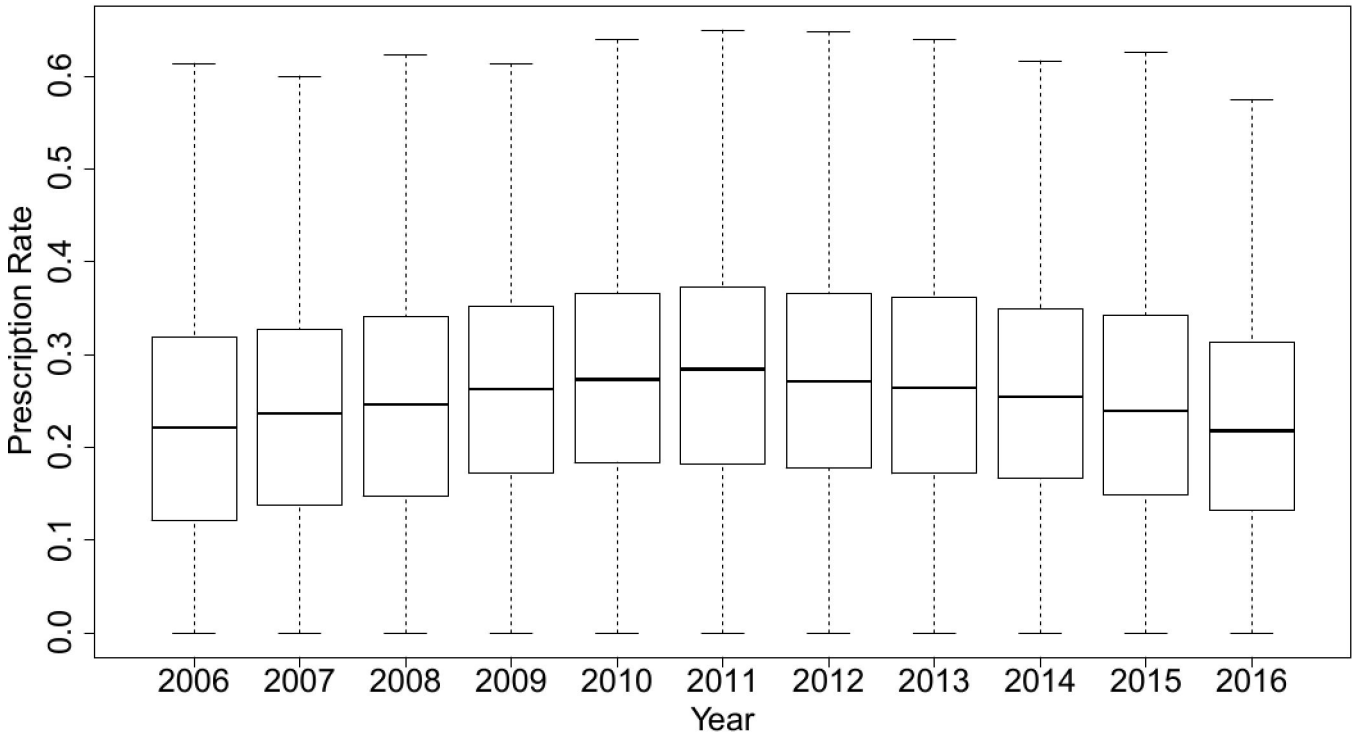
$$Leniency_s^{team} = \frac{1}{\sum_{j \in \mathbb{S}} N_{jy}} \left(\sum_{j \in \mathbb{S}} N_{jy} \times Leniency_{-s,jy}^{phys} \right) \tag{5}$$

where $Leniency_{-s,jy}^{phys}$ is as defined in Equation 2 except leaving out all patient cases occurring in shift s, $\mathbb{S}$ is the team of physicians j who are working at any point during shift s, and N_{jy} is the total number of cases seen in year y by physician j. This team-based leniency instrumental variable is a weighted average of the potential physicians a patient could have seen at the time they arrive in the ED. The weighting is based on the number of cases the physician sees that year to account for variance in our measure of individual physician leniency.

Figure F.10 graphs the histogram of the team leniency along with its first stage in comparison with the baseline physician leniency. As expected, the range of possible values shrinks, however, the first-stage slope remains unchanged.

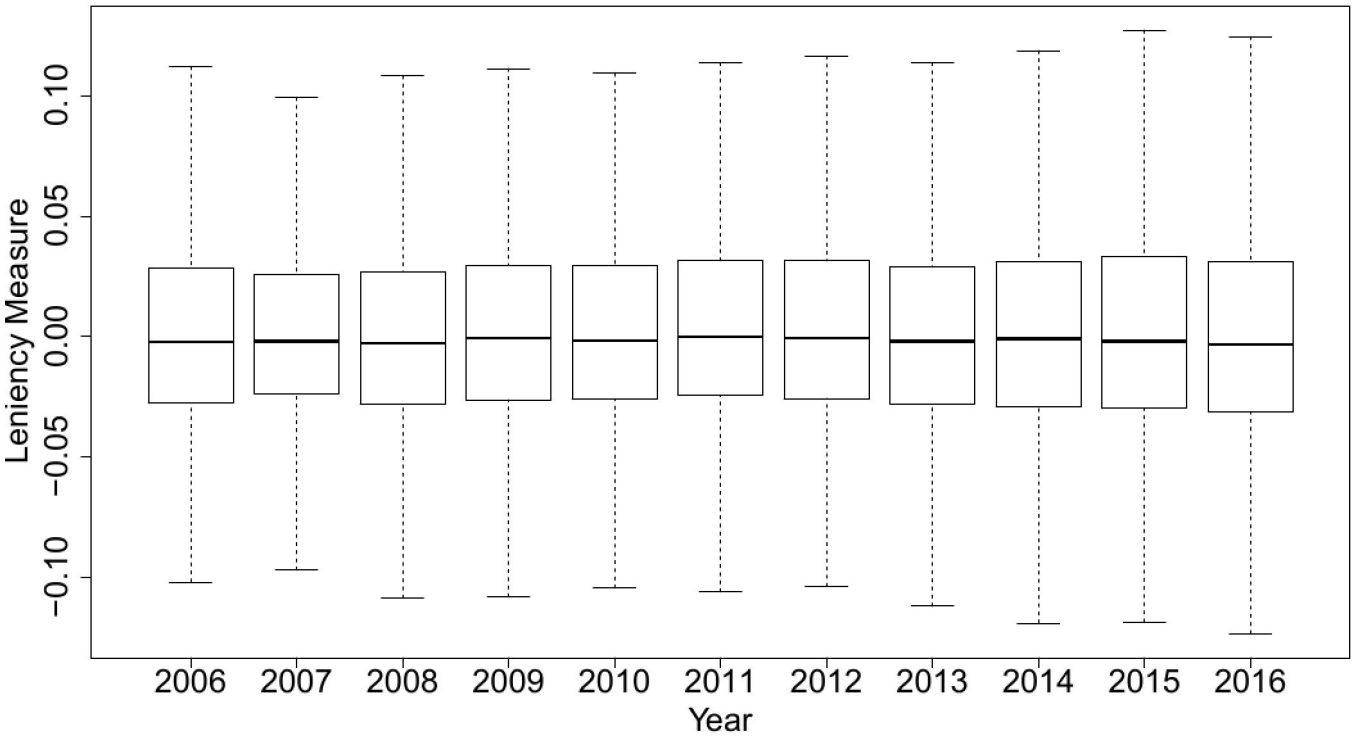
F. Additional Figures and Tables

Figure F.1: Raw Opioid Prescribing Rate among Physicians in VHA EDs for our Baseline Sample



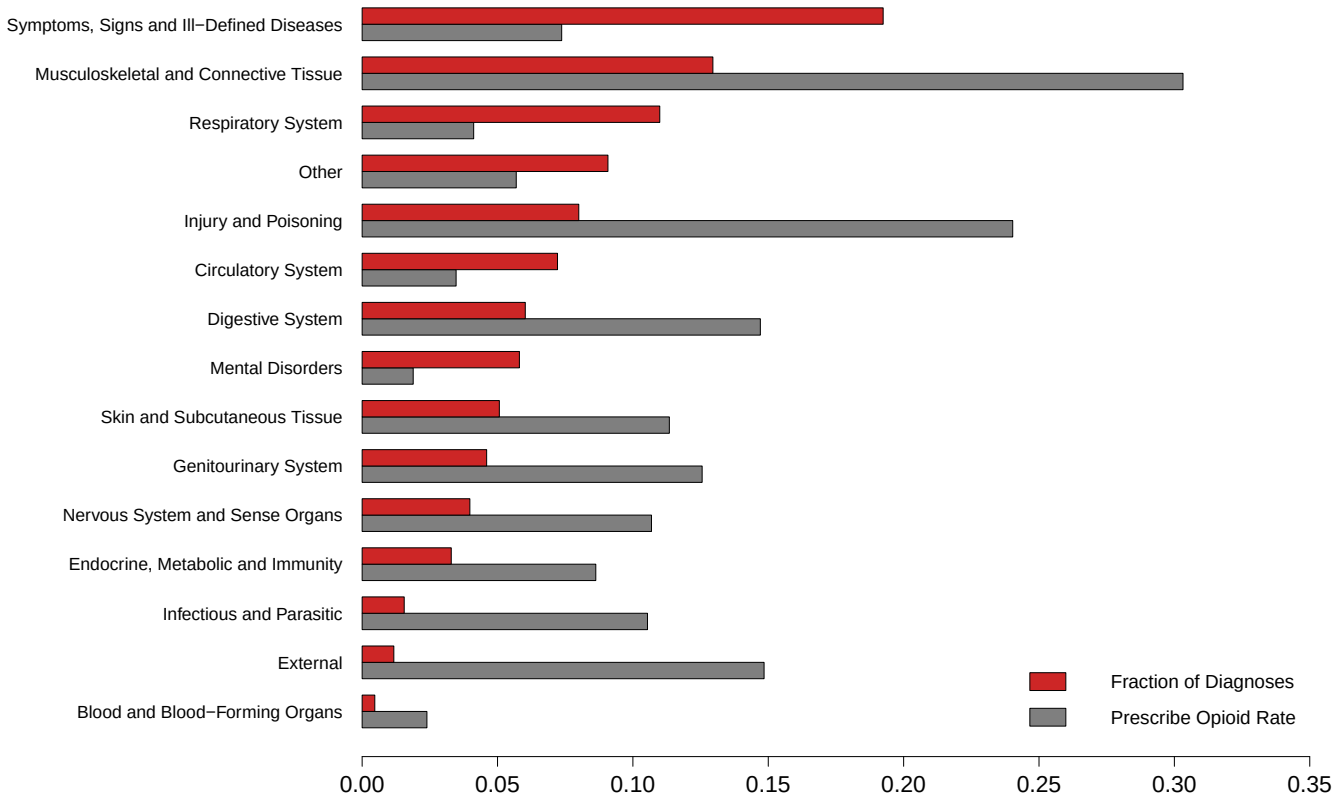
Notes: Boxplot of the raw opioid prescription rate for physicians in our baseline sample as defined in the text. The opioid prescribing rate peaks in 2011.

Figure F.2: Boxplot of Residualized Prescribing Leniency Measure



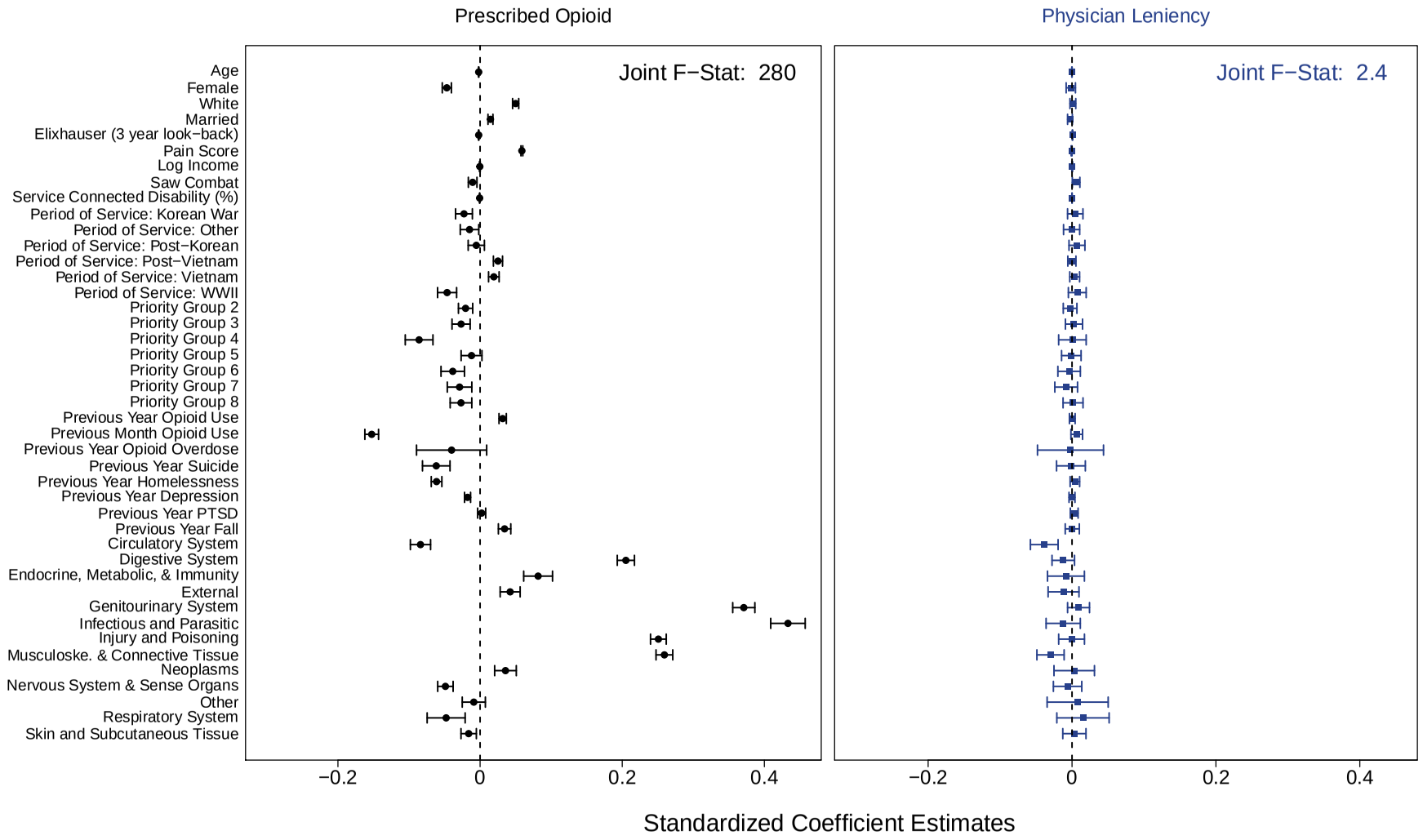
Notes: Boxplot of the residualized leniency measure for physicians in our baseline sample as defined in the text.

Figure F.3: Frequent Diagnoses Occurring in Emergency Departments



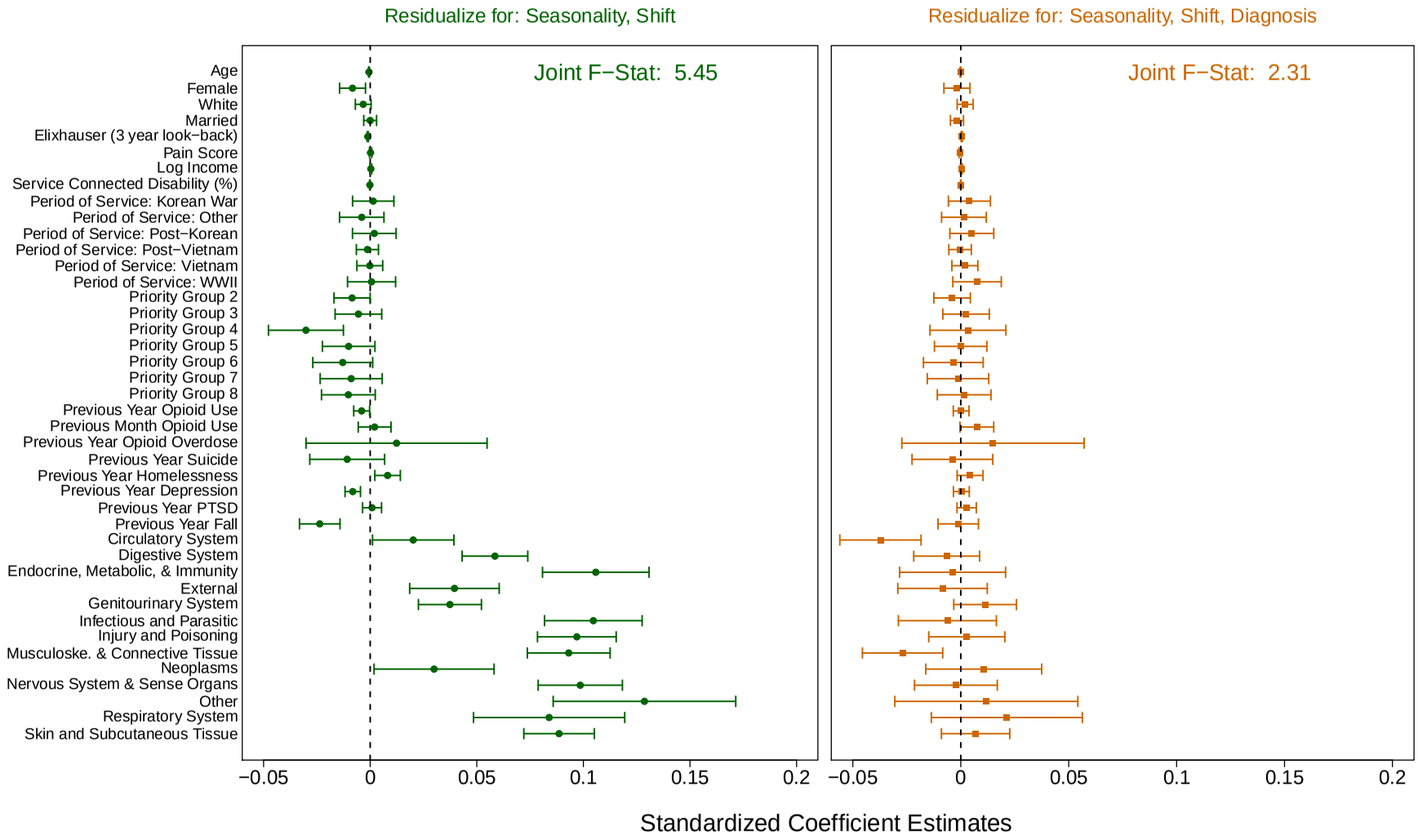
Notes: The 15 most common major diagnosis categories (ICD-9 major chapters) for *all* ED visits and the un-adjusted rate they are prescribed opioids

Figure F.4: Balance Test for Quasi-Random Assignment (Graphical Version)



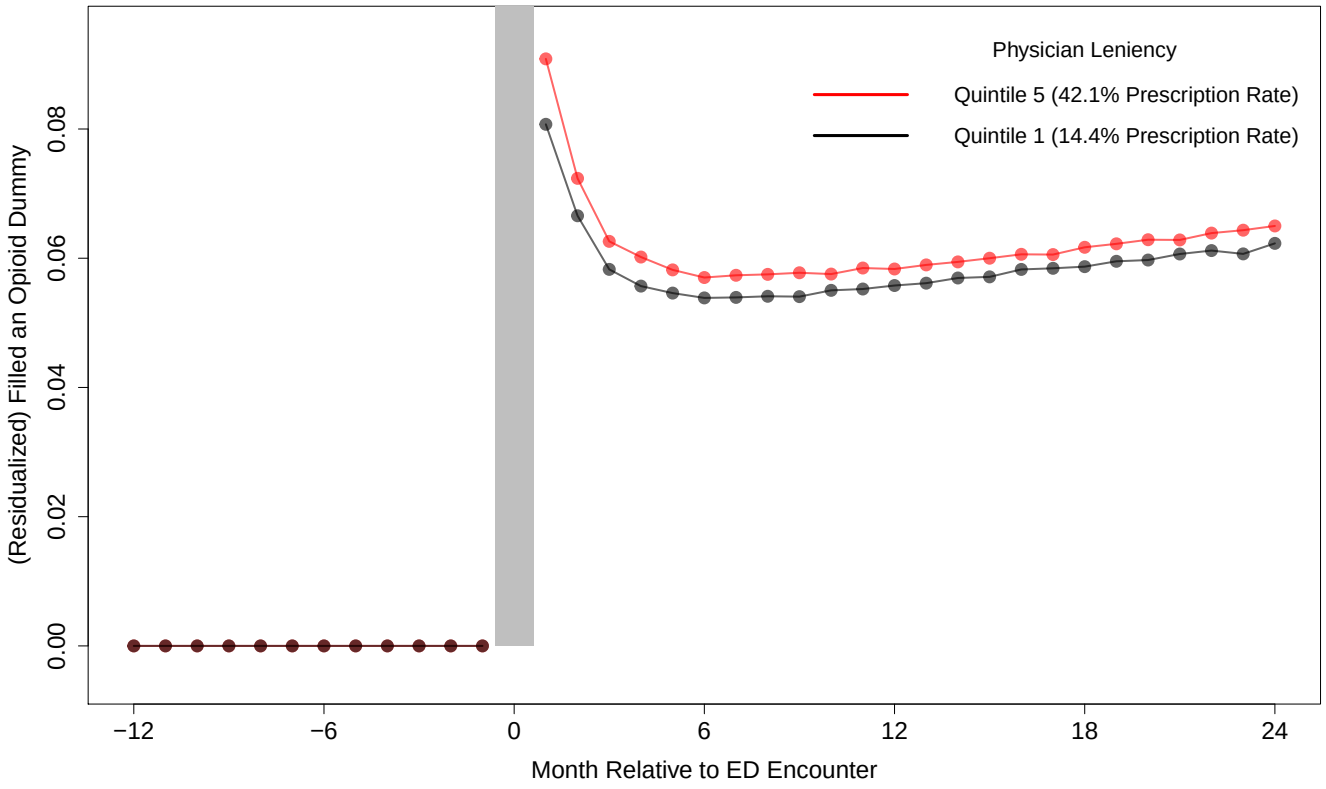
Notes: The graphical version of Table 3 where the left panel corresponds to the first column and the right panel corresponds to the second column.

Figure F.5: Balance Test at Varying Levels of Residualization Controls



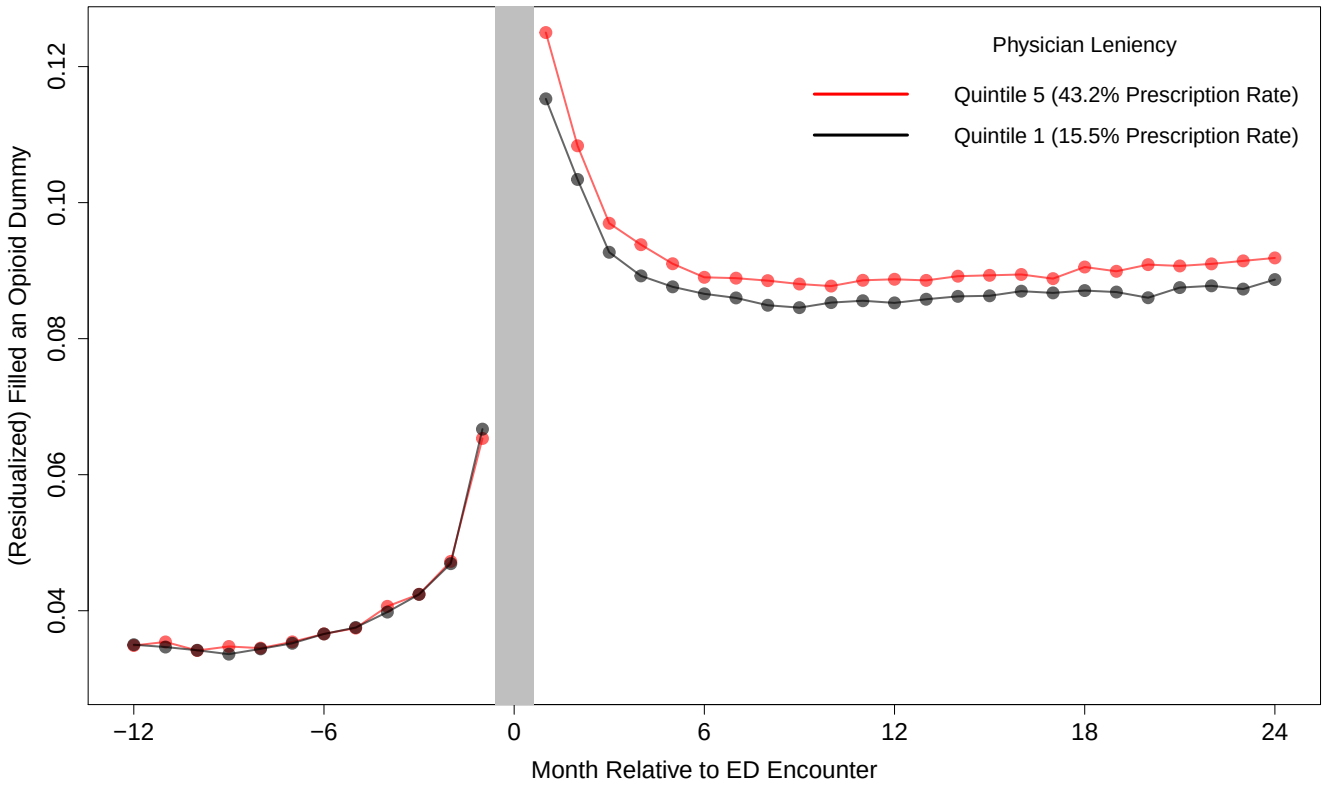
Notes: This figure displays the standard balance test (as seen previously) for the physician leniency measure constructed with different controls in the residualization of Equation 1, effectively, testing for different quasi-random assignment assumptions. In the left panel, the physician leniency instrument is constructed with only controls for seasonality and shift. The right panel, the diagnosis condition is included as an additional control in the residualization.

Figure F.6: Reduced Form: Subsequent Opioid use for Opioid-Naïve Patients



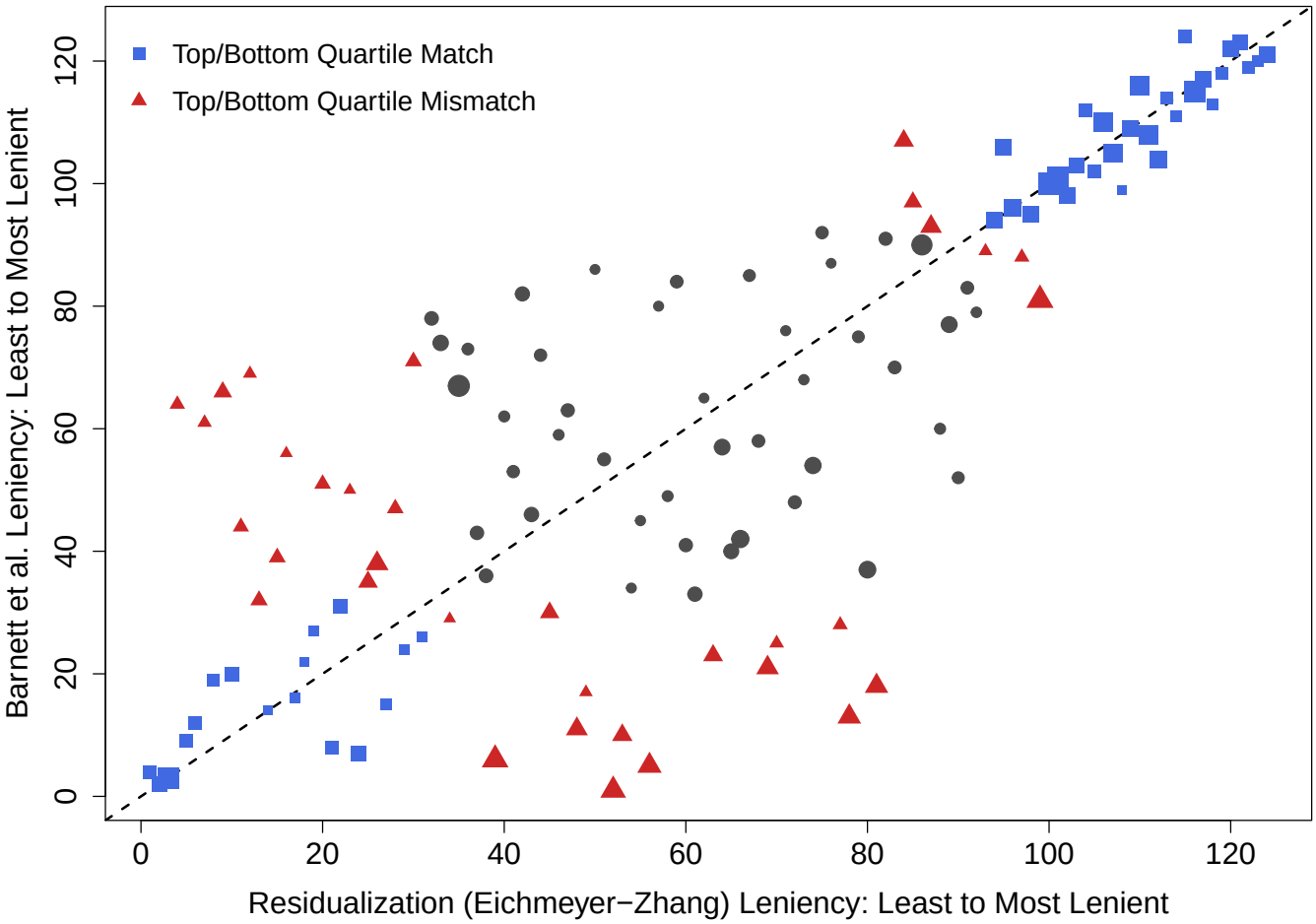
Notes: The reduced form event-study figure corresponding to Figure 4, but for opioid-naïve patients.

Figure F.7: Reduced Form: Subsequent Opioid use for Patients Without an ED Visit in the Prior Year



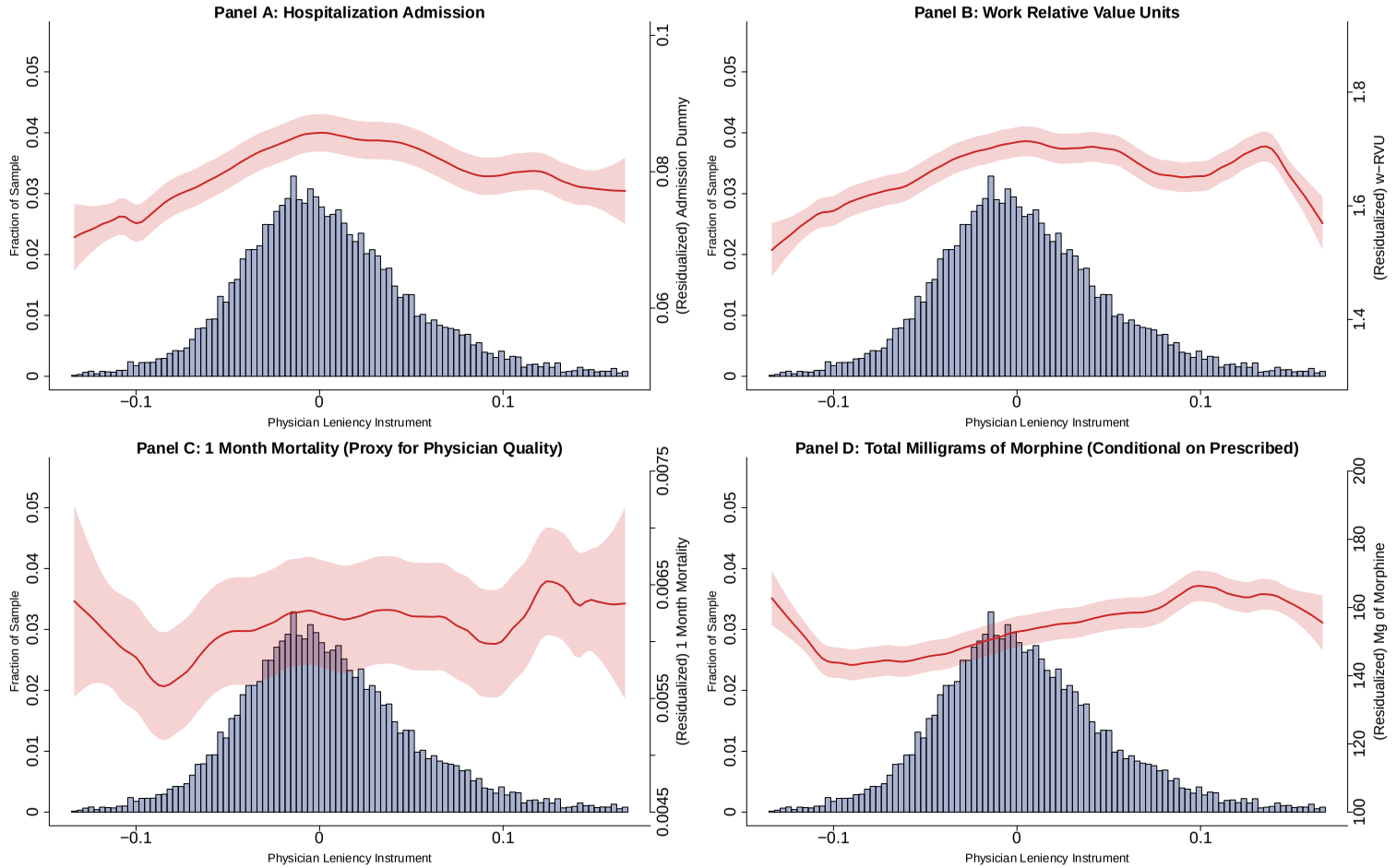
Notes: The reduced form event-study figure corresponding to Figure 4, but for patients who did not visit an ED in the prior year (for any condition), but did utilize VHA outpatient care. Presumably this group are not ED shopping for opioids.

Figure F.8: Ranking Physician Prescribing Leniency in Tampa Veteran Affairs Medical Center



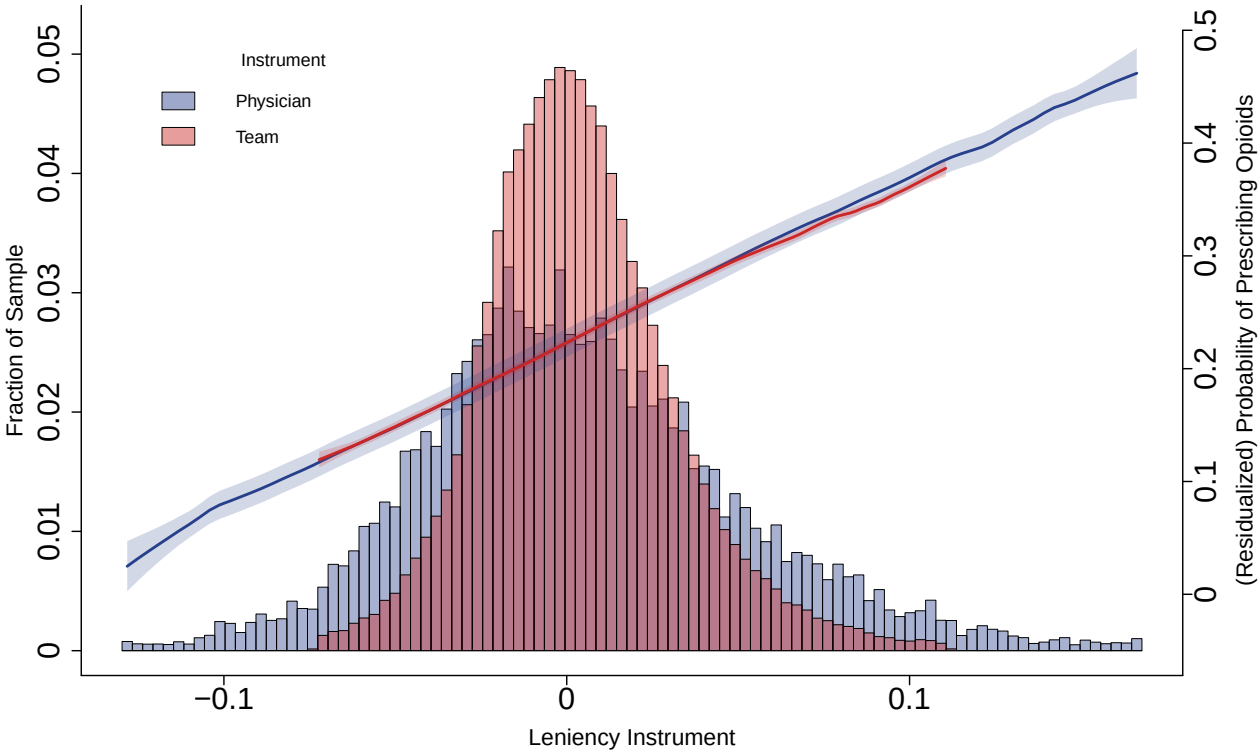
Notes: This graph shows the re-shuffling of physician ranking based on Barnett et al. method and our residualization method for physicians with at least 30 cases in Tampa Veteran Affairs Medical Center, the largest in the country by ED volume in 2012. Each point is a physician and the size of the point is proportional to the log number of cases seen. The blue boxes correspond to physicians that would be classified in the top or bottom quartile by both methods, and the red triangles correspond to physicians that the two methods disagree on.

Figure F.9: First Stage of Baseline Physician Opioid-Prescribing Leniency Instrument on Other Dimensions of Physician Characteristics



Notes: Each panel overlays a first stage local linear regression of a particular (residualized) physician dimension on a histogram of our baseline physician opioid-prescribing leniency. Panel A corresponds to the decision of admitting a patient, panel B is the total work relative value units as a proxy for intensity of procedures performed, panel C is physician quality, proxied by the one month mortality rate, and panel D is the intensive margin of total volume of milligrams of morphine equivalent, conditional on being prescribed an opioid. 95% confidence bands are also displayed in the shaded red region.

Figure F.10: Distribution and First Stage of Team Instrument



Notes: This figure plots the histogram of the alternate team leniency instrument (overlaid on top of the baseline physician leniency instrument) along the x-axis and the left y-axis. A local-linear regression of the fitted probability of prescribed opioids on the instrument after residualizing is overlayed and displayed on the right y-axis. 95% confidence bands are also shown.

Table F.1: Top Specific Opioids with Dosage Prescribed in Emergency Departments

2011		2016	
Drug Name with Dosage	Fraction	Drug Name with Dosage	Name
Hydrocodone 5mg/Acetaminophen 500mg Tab	34%	Hydrocodone 5mg/Acetaminophen 325mg Tab	30%
Tramadol HCL 50mg Tab	15%	Tramadol HCL 50mg Tab	17%
Oxycodone HCL/Acetaminophen 325mg Tab	10%	Oxycodone HCL/Acetaminophen 325mg Tab	13%
Codeine 30mg/Acetaminophen 300mg Tab	6.5%	Oxycodone HCL 5mg Tab	9.2%
Oxycodone HCL 5mg Tab	5.6%	Hydrocodone 10mg/Acetaminophen 325mg Tab	7.7%
Hydrocodone 7.5mg/Acetaminophen 500mg Tab	3.9%	Codeine 30mg/Acetaminophen 300mg Tab	6.4%
Hydrocodone 5mg/Acetaminophen 325mg Tab	3.8%	Hydrocodone 7.5mg/Acetaminophen 325mg Tab	4.1%
Hydrocodone 10mg/Acetaminophen 325mg Tab	3.0%	Morphine SO4 15mg Tab, SA	1.1%
Hydrocodone 10mg/Acetaminophen 500mg Tab	2.2%	Morphine SO4 2mg/ml Inj	0.8%
Morphine SO4 15mg Tab, SA	1.4%	Morphine SO4 4mg/ml Inj	0.8%
Hydrocodone 10mg/Acetaminophen 650mg Tab	1.2%	Hydromorphine HCL 2mg Tab	0.7%
Morphine SO4 30mg Tab, SA	1.1%	Morphine SO4 30mg Tab, SA	0.7%
Morphine SO4 15mg Tab	0.9%	Morphine SO4 15mg Tab	0.7%
Morphine SO4 4mg/ml Inj	0.8%	Oxycodone HCL 15mg Tab	0.6%
Hydrocodone HCL 2mg Tab	0.8%	Hydromorphone HCL 2mg/ml Inj	0.6%

Notes: The top 15 opioid drugs with dosage in 2011 and 2016. The quantity (e.g., number of tablets) and days supply will vary per prescription.

Table F.2: Conditional Independence Test: Does Predicted Prescription Status Correlate with Leniency?

	Dependent variable (×100): $\widehat{Prescribed}$
Leniency	-0.004^{***}
	(0.001)
N =	1,824,101

Notes: Predicted prescription status, $\widehat{Prescribed}$ is constructed by projecting *Prescribed* on the set of covariates in Table 3. The result of a regression of predicted prescription on physician leniency instrument is presented here. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table F.3: Main Outcomes with Physician Leniency IV Constructed at Varying Levels of Residualization

	IV Residualization Level:		
	Seasonality & Shift	+ Diagnosis	+ Additional Ctrl. (baseline)
Dependent variable:	(1)	(2)	(3)
Long-Term Use	1.113***	1.171***	1.172***
	(0.215)	(0.201)	(0.202)
OUD (Year 0-3)	−0.133	0.300*	0.335*
	(0.186)	(0.175)	(0.160)
Opioid Overdose (Year 0-1)	0.014	0.068	0.088*
	(0.048)	(0.044)	(0.043)
Opioid Overdose Death	0.068*	0.074**	0.075**
	(0.037)	(0.034)	(0.034)
Falls (Year 0-3)	−0.604**	0.461	0.504**
	(0.275)	(0.256)	(0.257)
Hepatitis C (Year 0-3)	0.080	0.250	0.259
	(0.218)	(0.206)	(0.209)

Notes: This table reports the 2SLS causal effect of an opioid prescription on the main outcomes with three different physician leniency instruments. The three instruments are constructed with varying levels of controls in the residualization in Equation 1: hospital-year-month and hospital-day of week-time of day (Column 1), hospital-year-month and hospital-day of week-time of day, and diagnosis (Column 2), and the above including Elixhauser Comorbidity Index, pain score, five-year age bins, and number of prior ED visits (i.e., the baseline IV; Column 3). All three regressions include the standard controls described in the text. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table F.4: Effect of Opioid on Outcomes by PCP Leniency for Opioid-Naïve Veterans

	Dependent variable:			
	Long-Term Use	Opioid Use Disorder	Opioid Overdose Mortality	Hepatitis C
	(1)	(2)	(3)	(4)
Lenient PCP	1.456***	−0.063	−0.063	−0.122
	(0.426)	(0.397)	(0.077)	(0.492)
Strict PCP	1.276***	0.620*	0.089	0.583
	(0.335)	(0.357)	(0.062)	(0.454)

Notes: This table reports the effect of an opioid prescription on the main outcomes for patients with above and below median PCP prescribing leniency for veterans with no observed opioid use in the year prior to the ED. The 2SLS regressions are estimated on each subsample separately. Residualization fixed effects include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins; see text for baseline controls. The samples are constrained such that the patients are alive for the entire period the outcome is measured. Regression coefficients, standard errors, and mean dependent variables are scaled as indicated. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table F.5: Effect of a Single ED Opioid on Outcomes by Veteran Combat Status

	Dependent variable:			
	Long-Term Use	Opioid Use Disorder	Opioid Overdose Mortality	Hepatitis C
	(1)	(2)	(3)	(4)
Combat	0.459	0.429	−0.076	0.233
	(0.666)	(0.564)	(0.122)	(0.585)
Non-Combat	1.315***	0.482**	0.123***	0.370
	(0.246)	(0.217)	(0.041)	(0.256)

Notes: This table reports the effect of an opioid prescription on the main outcomes for combat and non-combat veterans. The 2SLS regressions are estimated on each subsample separately. Residualization fixed effects include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins; see text for baseline controls. The samples are constrained such that the patients are alive for the entire period the outcome is measured. Regression coefficients, standard errors, and mean dependent variables are scaled as indicated. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table F.6: Characterization of Compliers

	$P(X = x)$	$P(X = x\|\text{complier})$	$\frac{P(X=x\|\text{complier})}{P(X=x)}$
White	0.690	0.657	0.952
Black	0.239	0.238	0.999
Age < 40	0.201	0.197	0.977
Age $\in [40, 60)$	0.363	0.379	1.046
Age ≥ 60	0.436	0.377	0.865
Opioid-Naïve	0.752	0.766	1.018
Prior Opioid User	0.248	0.222	0.896
Depression or PTSD	0.294	0.296	1.005
No Depression, No PTSD	0.706	0.696	0.986
Musculoskeletal and Connective Tissue	0.311	0.364	1.170
Injury and Poisoning	0.202	0.195	0.966
Digestive System	0.070	0.057	0.819
Other Major Diagnosis Categories	0.417	0.321	0.770
Above Average Risk for Opioid Overdose Death	0.376	0.400	1.064
Below Average Risk for Opioid Overdose Death	0.376	0.338	0.898

Notes: This table reports for each demographic subgroup: its unconditional share, its conditional probability given they are a complier, and the relative likelihood.

Table F.7: Opioid Prescription on Outcomes for ED Conditions Pertaining to Injuries and Poisonings

	Dependent variable (×100):								
	Long-Term Use	OUD Year 0-3	Overdose Year 0-1	Overdose Death	Fall Year 0-3	Depression Year 0-3	Suicide Year 0-3	Homeless Year 0-3	Hep C Year 0-3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Prescribed	1.438***	0.255	0.166*	0.031	0.340	0.624	−0.211	−1.36**	0.788*
	(0.391)	(0.360)	(0.098)	(0.070)	(0.619)	(0.969)	(0.283)	(0.634)	(0.422)
N=	382,034	382,034	382,034	373,178	361,090	361,090	361,090	361,090	361,090

Notes: This table reports 2SLS regression coefficients of *Prescribed* on the main outcomes the subset of the baseline sample with ED diagnosis of injury and poisoning (ICD-9 800-999, excluding drug poisonings). Residualization fixed effects include hospital-year-month, hospital-day of week-time of day, diagnosis, and age bins; see text for baseline controls. The samples are constrained such that the patients are alive for the entire period the outcome is measured. Regression coefficients, standard errors, and mean dependent variables are scaled as indicated. Robust standard errors are clustered at the physician level. *p<0.1; **p<0.05; ***p<0.01

Table F.8: Long-Term Use Estimate: Incrementally Moving From Barnett et al. (2019) to Eichmeyer and Zhang (2019)

Outcome: Long-Term Prescription Opioid Use	High Intensity (1)	Low Intensity (2)	High/Low Ratio (3)	Wald Estimate (4)
1. **Barnett et al. (2019)**	1.39	1.26	1.10	0.903
2. **Replicating Barnett et al. (2019)**	1.96	1.79	1.10	0.987
Incremental changes to sample restriction and data definition:				
3. +Extend long-term use defn. to opioid avail.	2.59	2.33	1.11	1.46
4. +Exclude urgent care clinics	2.53	2.30	1.10	1.39
5. +No prior enrollment/encounter restriction	2.70	2.38	1.14	2.01
6. +No post-ED cancer restriction	2.74	2.44	1.12	1.84
7. +Include admitted patients	2.97	2.68	1.11	2.00
8. +Include prior users	5.36	5.17	1.04	1.32
9. +Exclude rarely prescribed conditions	6.73	6.37	1.06	1.36
10. +Add CMS prescriptions	7.63	7.17	1.06	1.97
11. +Include all years (2006-2016)	6.05	5.36	1.13	2.75
12. **Year-varying physician intensity**	6.01	5.60	1.07	1.75
Incremental controls in leniency residualization:				
13. +Hospital-Year-Month (seasonality)	5.87	5.67	1.04	1.08
14. +Hospital-DayOfWeek-TimeOfDay (shift)	5.90	5.71	1.03	1.10
15. +Diagnosis	5.90	5.70	1.04	1.20
16. +Age, Elixhauser, pain score	5.90	5.69	1.04	1.25

Notes: This table begins with the estimate on long-term opioid use obtained in Barnett et al. (2019), and incrementally alters the sample and empirical approach (i.e., residualization in leniency construction) to arrive at the main estimates in this paper. Column 1 reports the mean long-term use associated with physicians in the top quartile of intensity (defined based on that specific sample restriction and/or leniency construction). Column 2 reports the same mean long-term use for physicians in the bottom quartile. The ratio of the two (odds ratio) is reported in column 3. The fourth column is a Wald estimate—mirroring the 2SLS estimate—for veterans treated by the top and bottom quartile physicians; BOJ 2017 calls this "number needed to harm". Row 1 reports the estimates found in Barnett et al. (2019) and row 2 is our best attempt at replication. Rows 3-11 *incrementally* alter the sample selection and data definitions, moving from Barnett et al. (2019) to our baseline sample in this paper. High/low intensity is defined with respect to the particular sample restriction, across all years. Row 12 classifies physicians-year as high/low within a facility-year for our baseline sample (row 11). Rows 13-16 employ our residualization approach described in Equation 1, *incrementally* including additional controls. In these four rows, both the outcome variable (long-term use) and endogenous variable (prescribed) are residualized with the baseline controls described in the text.

Table F.9: Average Characteristics of Physicians in the Top and Bottom Quartile of Leniency

	Lenient	Strict
Male	0.717	0.612
Age	47.4	46.1
Cases per year	929	789
Days worked per year	114	105
Patients per day	8.25	7.68

Notes: This table displays the simple mean of each variable for physician-years classified as lenient or strict. Lenient and strict are based on the top and bottom quartile of our leniency instrument measure each year. Only physician-years that treat at least 200 patients per year are included.