STAFF REPORTS

NO. 1158
JULY 2025

A Danger to Self and Others: Health and Criminal Consequences of Involuntary Hospitalization

Natalia Emanuel | Pim Welle | Valentin Bolotnyy

FEDERAL RESERVE BANK *of* NEW YORK

A Danger to Self and Others: Health and Criminal Consequences of Involuntary Hospitalization
Natalia Emanuel, Pim Welle, and Valentin Bolotnyy
Federal Reserve Bank of New York Staff Reports, no. 1158
July 2025
https://doi.org/10.59576/sr.1158

Abstract

Involuntary hospitalization of people experiencing a mental health crisis is a widespread practice, 2.4 times as common as death from cancer and as common in the U.S. as incarceration in state and federal prisons. The intent of involuntary hospitalization is to prevent individuals from harming themselves or others through incapacitation, stabilization, and medical treatment over a short period of time. Does involuntary hospitalization achieve its goals? We leverage quasi-random assignment of the evaluating physician and administrative data from Allegheny County, Pennsylvania, to estimate the causal effects of involuntary hospitalization on harm to self (proxied by death by suicide or overdose) and harm to others (proxied by violent crime charges). For individuals whose cases are judgment calls, where some physicians would hospitalize but others would not, we find that hospitalization nearly doubles both the probability of dying by suicide or overdose and also nearly doubles the probability of being charged with a violent crime in the three months after evaluation. We provide evidence of earnings and housing disruptions as potential mechanisms. Our results suggest that, on the margin, the system we study is not achieving the intended effects of the policy.

JEL classification: I18, I12, K14
Key words: involuntary commitment, psychiatric detention, mental health treatment

---

Emanuel: Federal Reserve Bank of New York (email: natalia@nataliaemanuel.com). Welle: Allegheny County Department of Human Services (email: paul.welle@alleghenycounty.us). Bolotnyy: Hoover Institution, Stanford University (email: vbolotnyy@stanford.edu). This work was made possible by Alex Jutca. The authors thank E. Jason Baron, Mary Burke, Derek Christopher, Kaveh Danesh, Richard Frank, Ashvin Gandhi, Pat Kline, Dev Patel, Stephen Ross, Jesse Welch and participants at Brookings Institution's seminar, UC Berkeley's Applied Micro Seminar, Hoover Institution's Conference on Collaborative State & Local Policy Research, the New York Fed Lunch Seminar, UCL's CReAM Seminar, the University of Sussex's Business School Seminar, LSE's Wellbeing Seminar, Carnegie Mellon University's Heinz College seminar, and George Washington University's Seminar for helpful comments. They are grateful to the many experts, physicians, and individuals who have experienced the involuntary hospitalization process firsthand who gave their time and expertise to help them understand the practical functioning of the involuntary hospitalization system, including Drs. Awais Aftab, David Cohen, Brian Holoyda, Keith Humphreys, Nev Jones, Anna Lembke, Julia Nath, Brian Stackhouse, Morgan Shields, James Tew, David Waters, and several anonymous readers. They also thank the many professionals at the Allegheny County Department of Human Services for their support, including Director Erin Dalton, Mike Rocco, Christina Matsook, Ping Ko, Carly Ferrone, Ilinca Metes, Jewel Denne, Stuart Fisk, Erika Montana, Felipe Diaz, and Nicholas Marlton.

1 Introduction

In the US, there are 1.2 million involuntary hospitalizations every year, based on the best available estimates (Lee and Cohen, 2021). Involuntary hospitalization occurs when an emergency department physician determines that an individual poses a clear and present danger to themselves or others due to mental illness. It is intended to prevent these perceived dangers posed to oneself or others from being realized.

Involuntary hospitalizations are common and increasing in number. The present annual rate of 357 hospitalizations per 100,000 residents is 2.4 times greater than the rate of death from cancer of any type (Institute, 2024), 57% greater than the rate of homelessness (de Sousa and Henry, 2024), and is nearly identical to the 350 individuals per 100,000 residents incarcerated in state and federal prisons every year (Carson, 2022).[1] Rates of involuntary hospitalization have grown three times faster than the population since 2011 (Lee and Cohen, 2021) and emergency department visits for acute psychiatric distress have likewise risen (Larkin et al., 2005; Twenge et al., 2019; Goodwin et al., 2020; Gold, 2020; Theriault et al., 2020).

Ex ante, the impact of an involuntary hospitalization is unclear. During episodes of mental health crisis, involuntary hospitalization may create distance from dangerous settings and give a person access to care that can stabilize them. This care might facilitate connections to medications, outpatient treatment, or other continuing services. Incapacitation may also reduce risk to others, since individuals are held in inpatient care under supervision and the majority of violent offenses are impulsive (Brouwers et al., 2010). Incapacitation may likewise reduce risk to oneself, since most suicides are premeditated very briefly (Deisenhammer et al., 2009). In these ways, involuntary hospitalization may be salubrious.

On the other hand, involuntary hospitalization may disrupt beneficial social supports such as therapeutic relationships, housing, and employment. It may introduce (new) treatments that are not sustainable or appealing to the patient, or last for a duration that is not sufficient to stabilize the patient's mental illness beyond the hospitalization. Resource constraints, such as under-resourced

[1] While we benchmark against individuals who are incarcerated or are in homeless shelters, those who are involuntarily hospitalized are generally a different population. In our sample, only 4.7% were incarcerated last year and 3% were in homeless shelters. Indeed, those who are evaluated for involuntary hospitalization include many individuals who are actively engaged in society, including through formal employment: 56% were employed in the year before with average earnings of $23,111.

and porous follow-up services, can make for detrimental experiences. Moreover, if individuals find involuntary hospitalization unwelcome, it may degrade their trust in the healthcare system.[2] In these cases, an involuntary hospitalization may be neutral or deleterious.

Understanding the impact of involuntary hospitalization is important not only because the practice is widespread, but also because the population brought in for evaluation is particularly vulnerable, experiencing high levels of pain and suffering. In surveys that are used to compare different health conditions by assigning each condition a weight on a scale of 0 (perfect health) to 1 (death), mental health disorders receive among the highest weights of any ailments. The weight for acute schizophrenia, for example, is 0.778 and is 0.658 for severe major depressive disorder (Global Burden of Disease Collaborative Network, 2020).[3] For context, severe heart failure's weight is 0.51 and asthma's weight is 0.019. Indeed, individuals with behavioral health conditions experience vastly shorter lifespans, with individuals with schizophrenia-spectrum disorders, for example, estimated to live 15.4 fewer years than people without those disorders (Chan et al., 2023).

Moreover, individuals evaluated for involuntary hospitalization tend to both have high needs and require high costs. Of those who have been evaluated for involuntary hospitalization in our setting, over 60% use an emergency room within one year after the evaluation. In our setting, individuals evaluated in the previous three years account for nearly $1 of every $4 of Medicaid behavioral health spending, despite making up only 1.5% of Medicaid enrollees (Welle et al., 2023). This amounts to about $14,000 per person per year in Medicaid behavioral health spending alone. Despite these expenditures, 20% of evaluated individuals die within five years after the evaluation—a rate that is higher than that for individuals exiting jail, enrolling in homeless shelters, or living with severe mental illness in general. Further, 24% are charged with a criminal offense within a year of evaluation (Welle et al., 2023).

Several jurisdictions have begun to expand their legal ability to treat individuals with mental illness without consent. Public cases in which people who have *not* been involuntarily hospitalized

[2] The United Nation's Special Rapporteur on the right to health, Dainius Pūras, denounced the practice of involuntary treatment in 2017, noting that it "causes mistrust, exacerbates stigma and discrimination and has made many turn away, fearful of seeking help within mainstream mental health services" (Puras, 2017). He argued that "dangerousness is often based on inappropriate prejudice rather than evidence" and urged the "radical reduction and eventual elimination" of non-consensual measures.

[3] An alternative interpretation of schizophrenia's weight is that living 1.28 years with acute schizophrenia is equivalent to dying one (healthy) year earlier than expected (1.28=1/0.778). As a result, health experts estimate that the US loses the equivalent of 4,707,170 years of full health to schizophrenia and depressive and bipolar disorders each year (Institute for Health Metrics and Evaluation, 2024).

and have subsequently engaged in violent behavior have prompted calls for expansions of involuntary hospitalization (Rosen, 2023; Hirschauer, 2025). Washington DC, Maryland, New York, New Mexico, California, and Oregon are among areas where officials are actively discussing expanding such policies or have already enacted expansions (Riley, 2024; Brown, 2024; Oreskes and Newman, 2025; Newman and Fitzsimmons, 2023; Fisher, 2024; Dembosky et al., 2023). In light of the magnitude of the problem and such policy initiatives, understanding the potential impact of involuntary hospitalization on the individuals it is trying to serve is crucial.

In keeping with involuntary hospitalization's statutory mandate, we investigate how involuntary hospitalization changes the probability that a person is a danger to themselves — measured using death by suicide or overdose — or a danger to others — measured using violent crime charges. For the latter, we use the Federal Bureau of Investigation (FBI) Uniform Crime Reporting Program's definition of violent crime: "those offenses that involve force or threat of force" (FBI, 2019).

Such analysis has been challenging for several reasons: first, a raw comparison of those who are hospitalized to those who are not would conflate the sorting that physicians do when selecting whom to hospitalize with the causal impact of hospitalization. Second, data privacy laws and disconnected data systems in the US have made it challenging to track an individual across databases for hospitalization, legal services, and autopsy. We tackle these challenges by leveraging data from Allegheny County, Pennsylvania, where the city of Pittsburgh is located and approximately 1.2 million people reside. In Allegheny County, there is quasi-random assignment to the evaluating physician, allowing us to use a physician's tendency to hospitalize patients as an instrument for actual hospitalization. Further, Allegheny County has pioneered integrated data systems, which facilitate our ability to observe a variety of patient outcomes.

We find that in Allegheny County involuntary hospitalization does *not* appear to serve its statutory purpose for patients who would be hospitalized by some physicians but not by others. Specifically, for these judgment call patients, involuntary hospitalization increases the probability of being charged with a violent crime by 2.6 percentage points (off of a base of 3.3%) and increases the probability of death by suicide or overdose by 1.0 percentage points (off of a base of 1.1%) in the three months following an evaluation. These results are based on an instrumental variable measurement strategy and are consequently producing a local average treatment effect of involuntary hospitalization on the compliers.

We consider several hypotheses that might explain why involuntary hospitalization increases adverse outcomes in judgment call cases. We find evidence that involuntary hospitalization causes destabilization, as evidenced by a decrease in employment and earnings, as well as an increase in shelter usage among those who had not used a shelter in the prior year. We do not find evidence to suggest that a decrease in drug tolerance is causing deleterious effects. We also do not find evidence that hospitalization meaningfully alters utilization of continuing care, such as outpatient mental health services or adherence to prescribed medications.

This paper contributes to several strands of literature exploring mental healthcare's impacts on crime and death, as well as on employment and homelessness, two of the mechanisms we consider. Much of the current literature has focused on voluntary, outpatient therapies and their impacts, in keeping the US's shift away from longer-term institutionalization.[4] Consistent with the Penrose (1939) hypothesis that untreated mental illness increases crime, Deza et al. (2023) find that having more centers offering (outpatient, voluntary) mental health and substance use programs decreases assaults of on-duty police officers, the individuals who are often first to respond to mental crises. Jácome (2020) and Deza et al. (2024) document that losing Medicaid access increases crime, particularly among those with mental illness. Nesbit (2023) notes that when a judge mandates (usually outpatient) mental health treatment as part of a defendant's probation, recidivism declines. This paper contributes to the literature by considering how involuntary, short-term hospitalization affects the criminal trajectories of those experiencing a mental health crisis.

The paper also contributes to a literature on how mental healthcare relates to mortality, again by considering involuntary, short-term, inpatient treatment. Many papers document that individuals with serious mental illness (SMI) have much higher mortality rates (e.g., Liu et al., 2017). Yet use of medications such as lithium, neuroleptics, and antidepressants reduces mortality significantly (Baldessarini et al., 1999; Simpson and Jamison, 1999; Angst et al., 2005). Costantini (2024) finds that receiving (voluntary) healthcare decreases mortality, in part due to better handling of co-occurring chronic illnesses. Yet whether treatment is voluntary or not appears to matter: Jordan and McNiel (2020) finds that patients who perceived a hospitalization as being coercive are more likely to

[4]There is, however, a literature on longer-term institutional facilities. Raphael and Stoll (2013) find that closures of psychiatric institutions in the 1950s and 1960s did not cause movement of individuals with serious mental illness (SMI) into the prison system, but closures of the 1970s and 1980s account for four to seven percent of the growth in incarceration in the following two decades. Davis et al. (2012) suggest that those with SMI were made worse off by the closure of intensive, long-term psychiatric facilities, while those with less serious issues who could access community-based treatment became better off.

subsequently attempt suicide, consistent with other research finding that coercive perceptions are associated with less patient satisfaction with care (Katsakou et al., 2010). Kallert et al. (2008) also finds that involuntarily-admitted patients were over-represented in suicide measures, and were less likely to comply with medication than voluntarily-hospitalized individuals. Studying those who are admitted to the emergency department for self-harm in California, Goldman-Mellor et al. (2022) condition on observables and find "no clear evidence that hospitalization reduces suicide risk." Hallett et al. (2024) further document that there can be adverse consequences of mental health inpatient experiences. Numerous studies also find that there is an increased risk of death by suicide just after release from the hospital (e.g., Simpson and Jamison, 1999; Qin and Nordentoft, 2005).

The mechanisms we explore relate to a literature on mental healthcare's impact on employment and homelessness. Those with mental illness tend to have lower rates of employment and earnings (Banerjee et al., 2017; Biasi et al., 2021). But Biasi et al. (2021) find that the introduction of (voluntary) lithium medication reduced the gap in earnings by one third for people with bipolar disorder in Denmark, a result consistent with other work finding that medication can enhance employment and earnings (e.g., Bütikofer et al., 2020). This paper contributes by considering the impact of involuntary, short-term treatment on employment and homelessness, also corroborating existing work on the adverse consequences of hospitalization and job loss (e.g., Dobkin et al., 2018; Morris and Kleinman, 2020; Sullivan and Von Wachter, 2009).

Finally, this work connects to the medical literature on defensive medicine and over-treatment. Research has found that physicians' treatment decisions respond to medical liability or perceptions of liability in ways that can induce over-treatment, a phenomenon that might be at play in our setting as well given the high-risk nature of emergency psychiatric care (Kessler and McClellan, 1996; Passmore and Leung, 2002; Studdert et al., 2005).

While there is a robust and active debate among policy-makers, advocates and academics about the ethics of involuntary hospitalization, we are largely silent on this point, except to note that the aim of any policy is to achieve its explicitly stated goals.

The rest of the paper is organized as follows: Section 2 provides background context on the involuntary hospitalization process. Section 3 notes our data sources and sample construction. Section 4 explains our empirical approach. Section 5 discusses local average treatment effects of involuntary hospitalization on individuals whom some physicians would hospitalize and others would not. Section

6 explores possible mechanisms. Section 7 explores policy responses including marginal treatment effects and heterogeneity analyses based on patient risk and physician traits. Section 8 performs a costing exercise. Section 9 concludes with a discussion of our findings, avenues for future research, and implications for policy.

2 Involuntary Hospitalization Process

2.1 Overview

In every US state and in many countries across the world, involuntary hospitalization is a legal tool meant to be used in emergency mental illness situations, when other safeguards for preventing adverse outcomes—family, friends, outpatient care, and the like—are insufficient. In Pennsylvania, Section 302 of the 1976 Mental Health Procedures Act gives physicians the power to involuntarily hospitalize an individual deemed a danger to themselves or to others for psychiatric reasons (MHPA, 1976).

Evaluation proceeds as follows. If a person such as a family member or neighbor sees behavior or hears an individual indicate an intention to harm themselves or others, they call a centralized, public hotline to request that the individual be evaluated for involuntary hospitalization (see Figure A.1A). A government delegate listens to their account of the individual's behavior and language. If the reporter's information reaches a given threshold,[5] the delegate will certify that the individual may be brought to a hospital emergency department for evaluation by a physician. Overall, 84% of calls are certified. The process of calling a government delegate is in place for individuals in our sample who are referred by family, friends, and other concerned parties. Physicians and police officers may self-certify and have an individual evaluated without consulting with a government delegate. Of the referrals in our sample, 17.5% originate with medical professionals and 22.5% with police officers. The remainder come largely from friends and family (48.8% of the full sample) and other assorted individuals such as case workers or administrative personnel in healthcare settings. If the referral is certified, the individual is transferred to a hospital emergency department by police.[6]

[5] A given case may not be certified if, for example, the reporter did not themselves witness the behavior or hear the language or if the individual's language/behavior does not actually indicate that they pose a danger to themselves or others.

[6] If a person's case has been certified by a delegate, they may be brought in for evaluation any time in the next 30 days. However, the vast majority of individuals are evaluated much more rapidly (Figure B.1).

In the emergency department, a physician is tasked with determining if the individual currently poses a danger to themselves or others due to mental illness. They do this by speaking with the individual and evaluating their behavior, as well as by speaking with the referring party. Physicians doing the evaluation can have psychiatric training or other specialties; those who do not have a psychiatrist as their evaluators can be committed to inpatient without receiving a formal psychiatric consult.[7] If the physician determines that the individual poses such a danger, the individual may be involuntarily hospitalized and transferred from the emergency department to an inpatient psychiatric facility.

In Pennsylvania, an individual may be hospitalized for up to 5 days based only on the judgment of the emergency department physician.[8] If the physicians in the inpatient mental health unit believe a stay longer than 5 days is necessary before a person can safely be released, they may petition the Court of Common Pleas (CCP) for a longer stay. A CCP hearing usually occurs 3.3 days after the evaluation determination. A CCP magistrate may choose to extend the hospitalization for 20 additional days. After this initial extension, further extensions of 90 or 180 days are possible.

When a person is released, they are almost universally given a referral to voluntary outpatient services. Thus our estimates should be interpreted as the effect of being involuntarily hospitalized instead of being released with the option of attending outpatient services.

Figure A.1B shows what share of the initially evaluated individuals are involuntarily hospitalized, and the maximum number of days that they may be held. Figure A.1C describes the flow of decisions for individuals in our setting. Out of 16,630 evaluations, 78% result in a decision to hold an individual for up to 5 days; the others are released. Physicians petition for 51% of the emergency holds to be extended (39.4% of the initially evaluated cases). Of those extension petitions, 73% are upheld by the magistrate, keeping the patient hospitalized for up to 20 days total (28.6% of the initially evaluated cases). Put differently, 8,150 or 63.1% of the initial involuntary hospitalizations do not extend beyond 5 days. Of those evaluated, 6.3% are petitioned to be held up to 90 days and 0.8% are petitioned to be held up to 180 days.

Our setting has a relatively high hospitalization rate. Two years before the start of our data, in

[7] In our sample, 42% of cases are brought to the main psychiatric hospital, where a psychiatrist is the evaluating physician, and where therefore individuals receive a full psychiatric evaluation before a hospitalization determination is made.

[8] This duration varies widely across US states and across countries around the world, with 72 hours also being common (Hedman et al., 2016; Saya et al., 2019).

March 2012, a mentally ill man who had previously been involuntarily hospitalized and for whom medical personnel were considering another evaluation, opened fire at Western Psychiatric Hospital in Pittsburgh, killing two individuals and injuring seven more (CNN, 2012). Undoubtedly, this affected decision-making about whether to approve hospitalization in judgment call cases. Hospital administrators we spoke with mentioned that physicians often experience community pressure to hospitalize individuals. Sometimes this arises from exhausted care-givers who have run out of alternative paths, and sometimes from people who feel that their own safety is at risk. Academic work has also documented the practice of "defensive medicine" and over-treatment, especially by physicians who are junior and in high-risk specialties, and raised concerns about the quality and costs of care associated with such practice (Kessler and McClellan, 1996; Passmore and Leung, 2002; Studdert et al., 2005).

2.2 Features Allowing Empirical Analysis

Our identification strategy leverages variation in physicians' tendencies for hospitalizing individuals involuntarily. Several features of our setting facilitate our empirical approach. First, there are multiple physicians in any given emergency department at any given time. This means that who will serve as the evaluating physician is unclear and we can compare outcomes across physicians.

Second, there is quasi-random assignment to physician. Cases are added to the triage list, usually with a priority level of 2 on a scale of 1 (very urgent) to 5 (not so urgent).[9] This triage list determines the order in which cases are handled. When physicians are done with their case, they then take the highest urgency case on the triage list (tie-breaking by who has waited longer), creating quasi-random assignment in which physician handles a particular case. There are exceptions to this rule, however. For example, if a patient is being involuntarily hospitalized regularly and is known to the examiners, there may be an attempt to ensure that the same physician sees a patient they have seen before if that physician is on shift and available. If the individual is a child or is elderly, an attempt may be made to assign the patient to someone with pediatric or geriatric experience, and caregivers may have a greater say. As such, we limit our sample to those individuals who are being evaluated for involuntary hospitalization for the first time and who range in age between 18 and 65 years old.

It is possible that there may be endogeneity in when a patient is brought into the emergency department and which physicians serve during those times. Consequently, we residualize our instru-

[9] For example, an evaluation for involuntary hospitalization is likely to be handled before a broken rib, but after surgery for acute trauma.

ment, effectively comparing *within* a given hospital's shift so that we only compare outcomes among patients who could have been seen by another physician. This decision follows many prior papers that use similar instruments where decision-makers are conditionally randomly assigned (e.g., Dobbie et al., 2018; Frandsen et al., 2023; Norris et al., 2021; Gross and Baron, 2022).

Third, there is limited possibility for influencing which physician will conduct an involuntary hospitalization evaluation. Evaluations in Pennsylvania must occur within two hours of an individual's arrival at the emergency department, making physician availability at the time of arrival and placement on the triage list a key determinant of who performs the evaluation (Welle et al., 2023).[10] Since quasi-random assignment of cases to physicians is critical to our empirical strategy, we test the relationship between individual patients' observable traits and physician tendency to hospitalize involuntarily in Section 4.3.

Finally, what happens after the evaluation in the emergency department is handled by a different set of physicians: the physicians in the inpatient mental health ward. After a patient is transferred out of the emergency department, decisions about what therapies to provide and what medications to try are made by the inpatient care team. Decisions to apply for an extension and grant a hold beyond the 5 days allowed by an emergency department's evaluation are determined by the inpatient care team and CCP magistrate, respectively. This helps with the key exclusion restriction, which we test below as well.

3 Data Sources

We study the impact of involuntary hospitalization in Allegheny County, Pennsylvania, an area that encompasses the city of Pittsburgh and its suburban and rural surroundings. It is home to about 1.2 million residents.

The County's administrative data available to us include information on individuals who interact with government services or legal processes, of which involuntary hospitalization is one (Allegheny County Department of Human Services, 2024).

Across the many services offered and many service-specific databases, the county connects records

[10] If an evaluation does not occur within two hours, the person is supposed to be released. We cannot tell from our data how often this happens, but from our conversations with psychiatrists, we understand this is very rare. Conversations with practitioners in other jurisdictions suggest it may not be as rare everywhere.

and generates a single unique identifier for each person. The matching algorithm is a rules based system that is an integral part of the data systems at the Allegheny County Department of Human Services (ACDHS). When a new record is loaded from a particular source, a match attempt is made against the universe of existing unique identifiers. This consists of a sixteen-step process that uses first name, last name, date of birth, and social security number to connect the record to the existing set of records. The process starts with the most strict match set possible (exact match on name, date of birth, social security number) and proceeds through less stringent match requirements (accounting for misspellings, shortened names, mis-entries of social security or date of birth). If no match is made against the existing set of identifiers, a new global unique identifier is created after de-duplicating records within that source.

Beyond the automated matching script, there are criteria that flag manual review, which is done at a weekly cadence. Examples of conditions that would flag a manual review are situations where a key piece of information is shared between two clients (if a parent, for example, writes their social security number down in place of their child's at the point of service). These collisions are handled on a case-by-case basis to ensure the high quality of the information available in the system.

While these data are available to ACDHS personnel, they are not seen by service-specific decision-makers. The County's data include demographic information such as age, legal sex, and educational level for some individuals.

Involuntary Hospitalization Data. Data on involuntary hospitalizations are aggregated by the Allegheny County Department of Human Services' Office of Behavioral Health. For each call to a government delegate, the date, the identifying information of the individual in question, the relationship of the reporter to the individual, and the identity of the delegate who decides whether or not to certify the case are all documented, as well as which of the four legal criteria — danger to others, attempted suicide, unable to care for self, or mutilated self — is being reported in the petition.

Data on involuntary hospitalization evaluations are likewise aggregated from hospitals where the evaluation occurred. For each evaluation, the reporter, criteria noted by the reporter, the physician deciding the case, and the conclusion of their evaluation are recorded. All of the information from initial call through the evaluation is recorded in the county's custom software created to manage and record data on involuntary hospitalizations.

Since any further holds beyond the initial five day hold must go through the Court of Common Pleas, this extension data is not included in the county's involuntary hospitalization system, but rather in the court data. The court records are matched back through Allegheny County's systems described above to complete the picture of each involuntary hospitalization.

Violent Crime Charge Data. Data on criminal charges come from local Magisterial District Courts and show the date on which an individual was charged and with what offenses they were charged. We focus on violent offenses, as defined by the FBI's Uniform Crime Reporting Program: "those offenses that involve force or threat of force" (FBI, 2019). This includes murder, negligent and non-negligent manslaughter, rape, simple and aggravated assault, robbery, and intimidation.

We focus on being *charged* with a violent offense as we believe that it strikes a reasonable balance of competing approaches to measuring whether someone is a danger to others. On the one hand, it is a more stringent measure of being a danger to others than arrest, as being charged with a crime generally requires two decision-makers—the arresting officer and the charging prosecutor—to believe there is sufficient evidence to regard this person as having endangered others. On the other hand, we cannot rely on conviction for a violent offense because, for instance, mental illness can be grounds for finding a defendant unfit to stand trial. So an outcome measure like conviction, while perhaps more precisely identifying guilt, could get distorted by the legal process, including whether or not someone has been involuntarily hospitalized.

Suicide and Overdose Death Data. We determine whether a death is due to either suicide or drug overdose based on autopsy reports from the Allegheny County Medical Examiner's Office. Most deaths are the result of disease and are not sent to the Medical Examiner for autopsy. However, deaths that are suspected to be due to suicide, drug overdose, homicide, or vehicular accidents are referred to the Medical Examiner, who provides an official determination of the cause of death.[11] Both deaths inside the hospital and outside of it may be sent to the Medical Examiner.

The examiner determines the immediate cause of death, which is the underlying disease or injury that explains why someone died (e.g., asphyxiation, sharp force/incised, blunt force trauma, firearm), as well as the manner of the death, which can be categorized as natural cause, (drug and non-drug)

[11] We focus on determinations made by the Medical Examiner rather than data from death certificates because a prior literature has found that there are often errors in cause and manner of death on death certificates completed by non-Medical Examiners, even when the certificate is filled out by a physician (McGivern et al., 2017).

accident, homicide, suicide, and undetermined.[12] If the cause of death is related to drugs and the manner of death is an accident or suicide, or if the manner of death is a suicide (and the cause of death is anything), we consider the death to fall into our category of focus.

Other Data Sources. We merge in additional data to obtain auxiliary outcomes and controls.

The primary additional source of data are Medicaid claims. For any individual on Medicaid, we observe every individual billing episode for services, the diagnosis the service was billed under, and the provider who billed the service. Services included could be things like inpatient or outpatient treatment, Assertive Community Treatment, service coordination, or prescription filling.

We have information on shelter usage through the county's Homeless Management Information System (HMIS). The HMIS data contains data on individual shelter stays. We use this information to analyze who has been housed in shelter in the past year, as well as who uses it in the months after an evaluation.

Data on employment and earnings come from state unemployment insurance records, maintained by Pennsylvania's Department of Labor and Industry. Data include earnings by employer at the quarterly level and so only capture formal employment and income.

Lastly, we match evaluating physicians based on name and practice location in the National Plan & Provider Enumeration System (NPPES) to retrieve a National Provider Identifier (NPI) record. Once matched, we extract publicly available information on the physicians, such as how long they have been in practice and their specialty.

4 Empirical Approach

4.1 Overview

We seek to understand the impact of the involuntary *Hospitalization* of individual i on their outcomes Y_i such as death by suicide or overdose and having a subsequent violent crime charge, when they have

[12] Many immediate causes of death could fall into various possible categorizations. For instance, some asphyxiation deaths are categorized as homicides, some as non-drug accidents, and some as suicides.

been seen by physician j on rotation r. This goal can be modeled as:

$$Y_{ijr} = \gamma_0 + \gamma_1 Hospitalization_{ijr} + X_{ij}\Gamma + \epsilon_{ijr} \tag{1}$$

where X_{ijr} is a vector of individual- and case-specific characteristics and ϵ_{ijr} is an error term.

A key problem in an ordinary least squares (OLS) regression of this relationship is that there may be a correlation between involuntary hospitalization and the unobserved traits of the individual that are also correlated with our outcomes of interest: that is, the OLS regression is likely to suffer from omitted variable bias. Indeed, given that those who are hospitalized look broadly similar to those who are not (Table 1), there may be reason to believe that physicians' discernment in deciding who is hospitalized is based on traits unobservable to the researcher.

Assuming physicians have some skill at differentiating which patients would benefit from hospitalization beyond flipping a coin, the individuals who are hospitalized are meaningfully different from those who are not hospitalized. Thus, any OLS-driven comparison would capture both the treatment effect of hospitalization and the inherent differences in the two groups that affect our outcomes of interest. For example, the evaluating physician may notice that one patient has more suicidal intent than another who has suicidal ideation (but no intent) and choose to hospitalize the former and not the latter. Such unobserved features would produce biased OLS estimates.

Moreover, OLS assumes constant treatment effects for all groups, which may obfuscate heterogeneous treatment effects. If some (unobservably different) individuals respond positively to hospitalization and others negatively, we might find no effect. Or, if many respond positively, but some respond negatively, we would likely miss the latter in an OLS approach.[13]

To overcome the challenges posed by OLS, we use an instrumental variables approach that leverages quasi-random assignment of each evaluation to an emergency department physician. Specifically, we use a physician's tendency to hospitalize an individual as an instrument for actual hospitalization. This strategy, often called "judge IV" has been used widely.[14]

In our approach, we construct a residualized measure of hospitalization tendency that removes the rotation fixed effects. For each case, we then construct a leave-on-out mean of the physician's

[13]Our empirical strategy does not necessarily capture heterogeneous treatment effects, unless they align with the complier group.

[14]For example, judge IV designs have been used to study the impacts of incarceration (Kling, 2006; Aizer and Doyle, 2015; Norris et al., 2021), foster care (Doyle, 2007; Gross and Baron, 2022), pretrial detention (Dobbie et al., 2018), and disability insurance (Dahl et al., 2014), among many others.

tendency to hospitalize, a jackknife instrument. We then use this instrument in a two-stage least squares design, which estimates the following two equations:

$$Hospitalized_{ijr} = \alpha PhysicianTendency_{ij} + X_{ijr}\phi + \mu_{ijr} \tag{2}$$

for individual i, seen by physician j on rotation r, where X_{ijr} includes hospital-by-rotation fixed effects as well as individual level demographic controls and *Physician Tendency*$_{ij}$ is our residualized measure. The second stage in our two stage least squares regression is given by:

$$Y_{ijr} = \beta \widehat{Hospitalized}_{ijr} + X_{ijr}\lambda + v_i \tag{3}$$

Our estimate for β captures the effect of involuntary hospitalization on our outcomes of interest.[15]

This approach estimates the effect of involuntary hospitalization on individuals in a gray area, who would be hospitalized by some physicians but not by others. This is called the local average treatment effect (LATE) (Imbens and Angrist, 1994) and gives us a measure of the effect of involuntary hospitalization on those individuals whose hospitalizations are "judgment calls."

The fact that our estimates pertain only to "judgment call" cases is both a weakness and a strength of the paper. On the one hand, it suggests considerable caution in generalizing our results to all involuntary hospitalizations. For example, patients' likelihood of violent behavior may be quite different for those who would be hospitalized by *all* physicians in our sample, and whose cases are therefore not judgment call evaluations. On the other hand, our estimates are the policy-relevant measures. The individuals whose effects we estimate are those who would be affected if slightly more or slightly fewer individuals were hospitalized through the mechanism of physician discretion.

Instrument Construction. We use physician tendencies to hospitalize an individual as our instrument. Specifically, we use a residualized, leave-one-out mean construction of our instrument.

This residualized measure is preferable in settings with quasi-random assignment for two main reasons, as noted in the literature (e.g., Dahl et al., 2014; Dobbie et al., 2018). First, our case assignment is only random conditional on the rotation in which a patient was brought to the emergency department, not truly random. Rotation, here, is the time-of-day and month in each hospital, and

[15] A literature (Cellini et al., 2010; Gelber et al., 2016) underscores that treatment assignment might affect outcomes both directly and indirectly by influencing future hospitalizations. β captures both effects, which is precisely the policy-relevant element that we would want to understand. We discuss this in Section 5.3 to help interpret the findings.

aims to capture which physicians are on-staff at a given time. As such, selection could bias our results if we did not residualize. For example, if more stringent physicians tend to work on days of general carousing, and individuals with substance use disorders are more likely to be brought in on these days, we may see some selection effects. Residualizing helps address any such selection issues.

Second, this approach controls for differences across hospitals. Since patients may be more likely to go to a hospital closer to their home, patient characteristics may well differ across hospitals. Moreover, hospitals may feature different norms about hospitalizing individuals.

To account for any systematic differences, we include fixed effects for the hospital-rotation in which an individual is seen. We create a residualized hospitalization measure by regressing hospitalization on hospital-rotation fixed effects and subtracting the covariates multiplied by their estimated coefficients from the $Hospitalized_{ijr}$ indicator. This approach means our analysis compares patients who encountered the same set of physicians, leveraging within-rotation variation only, and is consistent with prior work (e.g., Dobbie et al., 2018; Leslie and Pope, 2017; Frandsen et al., 2023).

We construct the instrument for each individual by computing the leave-one-out mean of the residualized measure of the physician's other cases — both past and future. We use the leave-one-out mean, which excludes the physician's decision in the present case from the instrument, so that we do not introduce estimation errors on both sides of the equation and bias into our estimates.[16]

There are 1834 unique rotations featuring first-time evaluations. The typical rotation has 8.26 physicians performing evaluations (including those who are evaluating non-first-time cases).

Variation in Physician Tendency. Figure 1 shows the distribution of our residualized instrument for physician hospitalization in the grey histogram. Our sample contains 424 physicians at 14 different hospitals. On average, each physician performs an average of 39.22 first-time involuntary hospitalization evaluations in our sample. Our residualized measure of physician tendency to hospitalize ranges from -0.427 to 0.271, with a standard deviation of 0.098. The lowest rate of hospitalization that one physician has is 11%, while 46 physicians hospitalize all of the cases they evaluate.

[16] The residualized leave-one-out mean is equivalent to the reduced form of the jackknife instrument that is recommended when the number of instruments (physician fixed effects) is likely to increase with the sample size (Stock et al., 2002; Kolesár et al., 2015).

4.2 Sample

Our sample consists of involuntary hospitalization evaluations that occurred from June 2014 through the end of 2023 in Allegheny County. We limit our sample in several ways. Table A.1 shows the cases, patients, physicians, and hospitals eliminated due to each restriction.

First, we focus on the initial time a specific patient is evaluated. Some hospitals attempt to have the same emergency department physician evaluate a patient who has been through the involuntary hospitalization process before if the physician happens to be on duty. For these patients, physician assignment is unlikely to be quasi-randomly assigned.

Second, we limit to cases where the patient is between the ages of 18 and 64. For those under 18, parents and physicians tend to be co-deciders and so the monotonicity of a leave-one-out measure of physician tendency to hospitalize may not hold. Likewise, for those over age 65, there are often caregivers who are involved, giving rise to similar concerns. Further, in these cases, hospitals often have protocols to involve physicians with pediatric or geriatric experience. As such, we focus only on patients who are likely to see whichever physician is on staff.

Third, we limit to cases handled by physicians who have encountered at least 15 involuntary hospitalization evaluations, in order to minimize noise in our measure of hospitalization tendency. This restriction ensures that we accurately measure physicians' tendencies to hospitalize as physicians' tendencies to hospitalize an individual stabilize after evaluating about 15 patients (Figure A.2).

Finally, we limit to evaluations that happen during rotations where more than one physician is available and where more than one evaluation takes place, a necessary restriction when we include rotation fixed effects.

Table 1 describes the analysis sample. Our sample consists of 16,630 evaluations of which 12,909 result in an involuntary hospitalization.

To whom do our estimates apply? Our instrumental variable analysis generates estimates about the impact of an involuntary hospitalization *only* for those individuals who would be hospitalized by one physician, but not by another. This group of people whose hospitalization is a judgment call are traditionally called "compliers" (Imbens and Angrist, 1994). An individual cannot be identified as a complier. However, by comparing physicians who hospitalize a greater or smaller share of the individuals they evaluate, we can both calculate the share of individuals who are compliers and describe their observable characteristics (Frandsen et al., 2023). Formally, if we order the physicians according to the share of people they hospitalize, $j = 1, 2, \dots J$, the proportion of compliers is given by $E[Hospitalized_{iJ}] - E[Hospitalized_{i0}]$, namely the difference in hospitalization rates between the physicians most and least inclined to hospitalize. With linearity of the first stage, this can be approximated by $\alpha(PhysicianTendency_J - PhysicianTendency_0)$. Compliers represent 43% of our sample, indicating that our estimates are relevant for a substantial portion of the population in question.[17]

These estimates cannot speak to the effects of hospitalization on those individuals whom all doctors would elect to hospitalize. To better understand to whom our estimates apply, we characterize this group of people, using the methods developed by Frandsen et al. (2023).[18] We describe them in the fourth column of Table 1.

Panel A shows demographic information. Men are slightly more likely than women to be evaluated for involuntary hospitalization. About half of the patients in our sample are between the ages of 18 and 35. Most of the cases involve a White patient. However, relative to the racial breakdown in Allegheny County, Black patients are over-represented, making up 28% of the evaluations, whereas only 13.5% of the county population is Black. About half have less education than a high school degree. For a quarter of evaluations, we do not know educational attainment. About 45% of the evaluated population is enrolled in Medicaid at the time of exam and 30% for all six months after the evaluation. Compliers are more likely to be male, more likely to be white and more likely to be enrolled in Medicaid than the sample of evaluated individuals.

Panel B shows past diagnoses for the individuals enrolled in Medicaid or with a county-paid medical claim. Within this subpopulation, about half have been diagnosed with a severe mental illness (SMI). This definition of SMIs includes schizophrenia, schizotypal disorder, schizoaffective disorder, unspecified psychosis, persistent mood disorders, bipolar, and borderline personality disorder. Over 50% have some sort of mental health diagnosis. And 18.3% have a substance use diagnosis. An individual may have multiple diagnoses and may be double-counted in these rows. Compliers are less

[17] Under the assumption of average monotonicity, $\alpha(PhysicianTendency_J - PhysicianTendency_0)$ represents a bound rather than a point-identified number. This is because the average monotonicity assumption defines compliers as all individuals whose treatment status varies across physicians. For any two physicians $j' > j$, $\alpha(PhysicianTendency_{j'} - PhysicianTendency_j)$ bounds the share of compliers, and by construction this bound is most extreme for physicians J and 0.

[18] Frandsen et al. (2023) demonstrate that how prevalent a given trait is among compliers can be recovered by regressing the treatment interacted with that trait on the treatment instrumented with the physician tendency instrument. This approach generalizes the method for binary instruments (Abadie, 2003).

likely to have mental illnesses, severe or otherwise, or substance use disorder than the general sample of evaluated individuals.

Panel C shows prior encounters in our study period with local government. About 22.3% have a prior criminal charge and about 9.6% have a prior noncriminal charge. About 30.2% have had a visit to the emergency department. Compliers are more likely to have prior charges and prior emergency department visits.

Panel D shows the traits associated with the involuntary hospitalization evaluation. Just under half of the cases are referred by a relative. Another 22.5% are referred by a police officer. Compliers are more likely to be referred by a family member. The concerns brought to the evaluating physician's attention during the process include having attempted suicide in half the cases, being unable to care for one's self in 44.1% of cases, and posing a danger to others in 47.2% of cases. These categories are not mutually exclusive. Compliers look like the general population of evaluated individuals in terms of the concerns noted.

Note also that our estimates apply, conditional on already having an experience that brought an individual to the emergency department for an evaluation. This helps reassure us about internal validity of our estimates since even the counterfactual group has had such an experience. This means that there is limited scope for selection.

4.3 Internal Validity

Several assumptions are necessary for our instrumental variable analysis to be valid, which we seek to test. These are: (1) *relevance:* physicians' tendencies to hospitalize other patients are associated with hospitalization of a given individual, (2) *random assignment:* physicians are (conditionally) randomly assigned to involuntary hospitalization cases, (3) *monotonicity:* patients who would be hospitalized by physicians with a low tendency to hospitalize would also be hospitalized by physicians with a high tendency to hospitalize and vice versa, (4) *exclusion restriction:* physician assignment impacts outcomes only through hospitalization.

Relevance. A natural test of relevance is a strong and significant first stage, as measured in α in equation 2. In both Table 2 and Figure 1, we find just such a robust first stage; the residualized measure of a physician's tendency to hospitalize in an evaluation is highly predictive of whether a patient is hospitalized.

The top panel of Figure 1 gives a graphical intuition of the first stage. The grey histogram shows the residualized physician tendency to hospitalize a patient. The upward sloping curve flexibly represents the first stage, showing that those assigned to physicians with a higher residualized tendency are more likely to be hospitalized. The rate of hospitalization increases monotonically and nearly linearly. At the 5th percentile, physicians hospitalize 64% of the patients they evaluate; at the 95th, 93%. A 10 percentage point increase in the physician tendency is associated with approximately an 8.3 percentage point increase in the probability of being hospitalized in our preferred specification.

Table 2 presents the results of estimating equation 2. Each column adds additional controls. Column 2 includes hospital fixed effects. While there is a slight decrease in the first stage when we control for hospital, this is not surprising as different populations may go to different hospitals. Column 3 shows hospital-rotation fixed effects, which capture the set of physicians who are staffing the specific emergency department at the time of the involuntary hospitalization evaluation. Column 4 includes demographic controls such as sex, age, racial category, education, and medicaid enrollment as shown in Table 1. When we add rotation- and individual-level controls, the magnitude of the first stage does not meaningfully change, consistent with quasi-random assignment of physicians. Column 5 includes controls for whether the individual has a prior criminal or noncriminal arrest or emergency department visit in our sample period. Column 6 includes the reporter's concern. Finally, column 7 includes medical controls for the subsample of those enrolled in Medicaid, including whether the individual has a severe mental illness, a mental illness, or a substance use disorder diagnosis. The penultimate specification is our preferred specification used throughout the rest of the paper.

Additionally, the effective F-statistic is 104.8, suggesting we do not have a weak instruments problem (Olea and Pflueger, 2013). Nevertheless we show results for Anderson-Rubin confidence intervals, which are impervious to weak instruments.

Our instrument is not perfectly predictive of hospitalization, possibly because physicians may respond to traits we do not observe, creating measurement error, or because their tendencies may vary over the sample period. Nevertheless, we find a very strong first-stage relationship.

Random Assignment. A key assumption is that our instrument is uncorrelated with the error term. This would hold if there were quasi-random assignment. To test that physicians are quasi-randomly assigned to involuntary hospitalization evaluations, we confirm that while demographic traits may be related to hospitalization, they are unrelated to the physician assigned to the case. Table 3 shows these tests. In column 1, we relate a set of observable traits to an individual's likelihood of being hospitalized. We find a significant F-statistic of joint significance of 10.9 (p-value ¡ 0.0001).

Column 2 shows that these traits are unrelated to the assigned physician's tendency to hospitalize. Our F-statistic of joint significance is 1.192, which is not statistically significant (p-value = 0.25).

Monotonicity. One condition needed to interpret our two stage least squares estimate as the LATE is that those who are hospitalized by a physician who hospitalizes a small share of patients would also be hospitalized by a physician who hospitalizes a large share of patients; and conversely that those who are released by a physician who releases a small share of patients would also be released by a physician who releases a large share of patients. This is referred to as average monotonicity.[19]

One test of monotonicity is that first-stage estimates should be non-negative in all subsamples. Table 4 displays the first-stage estimates separately by demographic group and Table 5 does so by the concern that precipitated the involuntary hospitalization evaluation. In all subsamples, the estimates are non-negative, stable, and significant, consistent with monotonicity.[20]

A closely related test is proposed by Frandsen et al. (2023), who show that one can recover a weighted average of individual treatment effects with a relaxed average monotonicity assumption. This assumption implies that the covariance between physician's hospitalization rate and the rate of the outcomes observed is weakly positive. This is akin to showing that the first stage should be positive in all subsampes of the data. We pass their proposed test.

We explore how physicians treat patients who are observably different (Appendix Figure A.3). We calculate a residualized hospitalization measure for each physician separately for either side of each trait (e.g., a tendency to hospitalize among female patients and another tendency to hospitalize among male patients). We plot the measures and note the correlation between the two measures. All of these correlations are positive, suggesting that physicians' tendencies toward hospitalization are similar across observably different patients. This is consistent with our monotonicity assumption. In robustness checks, we also relax this monotonicity assumption by letting our leave-out measure of judge leniency differ across case characteristics following Mueller-Smith (2015).

[19]For the marginal treatment effects discussed below, we require strict montonicity, which is a much higher threshold.

[20]These analyses are also helpful in understanding who the compliers are because heuristically, compliers are more heavily concentrated in subgroups where the instrument has a stronger relationship with hospitalization.

Further, our setting is conducive to close adherence to the physician's determination. It is nearly impossible for a person not to be hospitalized if a physician has determined that they ought to be, given that individuals may be held by force, including using police resources and/or physical restraints. It is possible, however, for an individual to be hospitalized even if a physician does not determine they must be: individuals may choose to voluntarily commit themselves to hospitalization. In the parlance of Pennsylvania's statute, this is called a 201. Some of these voluntary hospitalizations may occur after a physician has opted not to hospitalize the individual involuntarily. There are 1066voluntary commitments, 1066% of all examinations (including non-first-time examinations). All voluntary commitments are dropped from our sample accordingly (Table A.1).

Exclusion Restriction. The exclusion restriction requires that physician assignment only impacts outcomes via their decision about whether to hospitalize a patient or not. This is impossible to test empirically. However, several institutional features help give confidence that the exclusion restriction holds in our context. First, the emergency department physician who performs the involuntary hospitalization evaluation is not the physician who provides care on the mental health inpatient floor. Rather, inpatient physicians provide care while the patient is in the hospital. This means that the exact care provided is divorced from the decision of whether to hospitalize or not. This mitigates concerns about omitted treatment bias (Mueller-Smith, 2015).[21]

Second, the emergency department physician does not decide the duration of hospitalization; inpatient physicians decide whether or not to file for an extension of involuntary hospitalization with a magistrate in the CCP and it is that magistrate who decides whether a patient should remain hospitalized for longer than 5 days.

Consistent with the exclusion restriction, we find that conditional upon hospitalization, the physician's tendency to hospitalize is unrelated to further downstream outcomes (Table 6). Specifically, it is unrelated to whether an individual is prescribed antipsychotic medications, has a hearing with the Court of Common Pleas to extend their hospitalization based on the recommendation of the inpatient team or whether the magistrate extends their hospitalization.

If our exclusion restriction is violated, then the reduced form estimates can still be interpreted as

[21] In our data, 6% of hospitalizations do not show an inpatient stay. This may occur if the individual is held for a very short period of time before being released and the transfer does not occur. It can occur if there is no available spot in the mental health inpatient unit and the patient must be held in the emergency department. For this reason, we test the robustness of our findings to excluding these individuals in Section 5.3.

the causal effect of being assigned to a physician more or less inclined to hospitalize. Our reduced form estimates are reported in our tables and tend to be fairly similar to the estimates arrived at via two stage least squares. This makes sense given that we have a very strong first stage. The similarity of the reduced form and two stage least squares estimates provides confidence that we may be capturing the causal effect of hospitalization.

5 Local Average Treatment Effects of Involuntary Hospitalization

Naive ordinary least squares regressions find no statistically significant impact of hospitalization on either the probability of death or getting a violent crime charge (Figure A.9). As noted above, these regressions conflate the causal impacts of involuntary hospitalization with the sorting that physicians may be doing. A null finding could be consistent with several possible explanations. They could hypothetically belie either physicians selecting less risky individuals and then hospitalization increasing their risk or physicians hospitalizing more risky individuals and then hospitalization decreasing their risk. A null finding can also indicate that there are heterogeneous treatment effects that go in opposite directions for different groups. Under this assumption, overall null effect in the OLS case might reflect a weighted average of (a) a positive coefficient on hospitalization among compliers (increasing their risk of bad outcomes) and (b) a negative coefficient on hospitalization among non-compliers (reducing their risk of bad outcomes).

Instead we turn to our instrumental variables approach to learn the causal effects on our compliers. In the context of calculating the local average treatment effect, a relevant baseline comparison is the complier control mean, which estimates the rate of outcomes would have been among the complier population had they not been hospitalized. We follow Frandsen et al. (2023) in calculating these, finding that in the three months following an evaluation for involuntary hospitalization 3.3% of compliers are charged with a violent crime and 1.1% die by suicide or drug overdose.

5.1 A Danger to Others

Our instrumented estimates find that involuntary hospitalization increases the probability of being charged with a violent crime and the probability of death by suicide or overdose. Table 7 shows our first stage and both reduced form and instrumental variables estimates of the impacts.

For being charged with a violent crime, our reduced form estimate finds that a 10 percentage point increase in a physician's tendency to hospitalize is associated with a 2.6 percentage point increase in being charged with a violent crime within the following 3 months. Column 3 divides the reduced form estimate by the first stage, finding a slightly higher effect. Figure 2(A) shows the time trends in these effects. Each month shows the full effects through the end of that month; that is, they are cumulative estimates, not point-in-time estimates. The grey dots show the control complier mean—which can be interpreted as the cumulative share of the compliers who would have experienced the given outcome in the absence of hospitalization. The blue triangles show the level effects that would be implied by our two stage least squares estimates. They are calculated by adding the IV coefficient β_1 from equation 3 to the control complier mean. The figure shows that the statistically significant increase in the probability of a violent crime charge occurs in the first month after evaluation and is sustained nearly continuously throughout the six months following an evaluation. This is consistent both with the fact that some people are released at different time periods after an evaluation and also with the fact that it may take some time for violent crime charges to be brought.

This statistically significant increase in being charged with a violent crime is interesting insofar as it goes in the opposite direction from a possible mechanical effect. There are behaviors that are potentially considered both illegal and related to psychiatric decompensation, for instance threatening to kill a person. If a physician finds this person should be hospitalized, they may not be subject to the same strictures from the criminal justice system. If not hospitalized, it is possible that they would have been charged with a crime. So it is possible that the individuals who are not hospitalized are more likely to be incarcerated than those who are hospitalized, a mechanical effect that would push *against* finding the results that we do. We hypothesize that this dynamic is indeed occurring because among those referred for evaluation by police, we find no effect of hospitalization on the probability of violent crime charges (Figure A.6). We consider the robustness of our result by excluding cases referred by police in Section 5.3.

Our analysis focuses on being charged with a *violent* crime. To emphasize that the effects we see are concentrated in violent offenses, we consider all charges, regardless of whether they were violent or not. In Table A.2(A), we see a suggestive increase that is driven by violent crime charges. This suggests that nonviolent crime charges do not have the same connection to involuntary hospitalization.

5.2 A Danger to Oneself

Our results find that hospitalization increases death by suicide or overdose in the three months after an evaluation. A 10 percentage point increase in a physician's tendency to hospitalize yields an increase in the probability of death in the next three months of 0.98 percentage points. This represents a doubling of the control complier mean. Table 7 shows both our reduced form and instrumental variables estimates of the impacts. The effects over time are also shown in Figure 2(B). This is consistent with prior literature on the time following discharge from a hospitalization being especially high-risk for death by suicide (e.g., Qin and Nordentoft, 2005).

Our analysis focuses on death by suicide or overdose. To sharpen these results, we consider separately death by natural causes, which has been found in other contexts to be reduced by (voluntary) mental healthcare (Costantini, 2024). Table A.2(B) shows no increase and a suggestive decrease in natural deaths from involuntary hospitalization.

5.3 Robustness

Our results are robust to several different design elements in our main specification. Table A.3 shows the first stage, reduced form estimates, and instrumented LATE estimates for both violent crime charges and deaths by suicide or overdose. Row 1 shows the same estimates with no controls. Our main estimates are measured 3 months after the individual has been evaluated for hospitalization. Rows 2 and 3 show our outcomes as they are measured at 6- and 12- months after the evaluation.

We next shift which physicians are included. Our main specification requires that a physician has evaluated at least 15 cases. This is based on Figure A.2, which shows that physicians' hospitalization tendency seems to stabilize after they have evaluated approximately 15 cases. Nevertheless, we consider samples in which we include physicians who have evaluated more than 5 cases (row 4), and restrict to those who have evaluated more than 30 cases (row 5).

Our results are also robust to considering different ways of clustering our standard errors. Our main estimates cluster at the physician level in case those evaluated by the same physician are somehow correlated in ways not accounted for by the construction of the instrument. Another approach would be to cluster at the rotation level since that is the level at which treatment is assigned. Row 6 shows our results persist if we cluster at the rotation level. Yet another approach would be to cluster at the physician-rotation level, which also does not affect the significance of our effects. We have limited to only first-time evaluations to ensure that assignment to physician is random. As such, it does not make sense to cluster at the patient- or patient-physician or patient-rotation levels.

We next turn to conceptual robustness. One might worry that individuals referred by police have a different counterfactual to hospitalization than do other individuals. Specifically, some behaviors may be grounds for arrest if they were committed in sound mind. For example, verbally threatening a person may be considered either evidence of mental illness or grounds for arrest. If an individual who has exhibited such behavior is not hospitalized a police officer may arrest them. If that were the case, the control group may have elevated levels of being charged with a violent crime and the true results may be attenuated. As such, we exclude cases referred by the police in row 8 of Table A.3, finding that while our point estimates are slightly higher than our main estimates, they are not so large as to conclude that our main estimates are artifacts of mechanical incarcerations in the control group. An alternative explanation for the different rates among police officers is that the cases they refer are less likely to be extended (20 % among police referrals relative to 31 % among non-police referrals), which provides more opportunity for people to be charged with a violent crime. This further suggests that our findings are not only the product of incarceration by police immediately after an examination that does not result in hospitalization.

Another concern is that those referred by physicians may be different themselves or may be treated differently. As such, row 9 shows results excluding cases referred by physicians.

Likewise, we had noted that not everyone has an inpatient visit in their chart. We limit accordingly in row 10, finding similar results. In one specification, you could measure deaths by suicide and overdose among those who are not yet charged with a crime in case jail incapacitates suicide or overdose. This approach is also consistent with avoiding double-counting these individuals in both sets of outcomes. However, it is also reasonable to wonder about the overall effects, regardless of whether an individual has spent time in jail. To this end, Table A.3 Row 11 considers this outcome, finding slightly noisier effects that point in the same direction. In Row 12, we limit to those who individuals who have an address in the county, as our data coverage is less good among those outside the county. Specifically, we may have less good information about medical examiner data for those out of the county.

We allow our main sample to include people even if we do not observe them for the entire post-period after the evaluation (e.g., their evaluation occurred only 2 months before the end of our data) to maximize power. Row 13 shows the estimates when limiting to a balanced panel for which we can see the whole post-period. Row 14 shows estimates that exclude the first year of data, just in case a person had been evaluated before our data started and we are inadvertently failing to limit to first-time evaluations.

We also consider alternative instrument constructions. Our main estimate uses a residualized leave-one-out mean for physician tendency used in much of the recent judge IV literature. We also use a slightly more intense formulation of the jackknife instrument in Row 15, that residualizes based on all the controls. Next, we use a split sample instrument in which we estimate physician tendency separately for half of the rotations and then apply these tendencies to the other half the sample and vice versa. We find similar results in row 16. Row 17 presents results using a full set of physician fixed effects as instruments. We also construct instruments that are allowed to differ based on whether a person has been indicated to have attempted suicide or not (row 18) or be a danger to others or not (row 19). These two instruments relax the monotonicity assumption in case physicians systematically treat patients with certain concerns differently from others. The similarity of these results suggests that potential monotonicity violations are likely to introduce only minimal bias.

We consider Anderson-Rubin confidence intervals, which are valid regardless of whether an instrument is weak since their procedure uses test inversion to obtain confidence intervals (Anderson and Rubin, 1949; Andrews et al., 2019). Figure A.7 shows that the Anderson-Rubin confidence intervals are nearly identical to the conventional LATE ones. This is consistent with the fact that our effective F-tests shown in Table 2, which measure weak instruments per Olea and Pflueger (2013), are sufficiently large that we are not concerned about weak instruments.

For additional insight, we further include the results as they apply to subsets of our data. Table A.4 shows the results. These subgroups reduce the sample size considerably and estimates may lose significance as result. Nevertheless, we find them helpful for contextualizing. In Row 1, we include results where we limit to people who did not experience an extension in their stay and thus did not remain in the hospital for more than 5 days. We find that point estimates for both outcomes remain positive, but only significant for death by suicide or overdose. In rows 3 and 4 we consider those who are examined in the regional psychiatric hospital versus the other hospitals. In rows 4 and 5 we consider those who are and are not enrolled in Medicaid at the time of the evaluation, respectively. For both sets of subgroups, we find that point estimates tend to remain positive, but only one outcome is significant for each subgroup.

5.4 Interpretation

These estimates speak directly to the policy experiment of dialing up or down the share of those evaluated who are hospitalized.

First, the measured effects are those of involuntary hospitalization on "judgment call" cases, also known as compliers. As discussed above, the compliers are those who might be hospitalized by some physicians but not by others.

This is important to keep in mind because, as we noted above, our setting has a relatively high hospitalization rate. Settings with lower hospitalization rates may have a different set of judgment call cases. Settings that have a lower hospitalization rate might see judgment call cases that would benefit from involuntary hospitalization.

Second, our estimates likely encapsulate both direct and indirect effects. That is, if hospitalization changes the likelihood of future involuntary hospitalizations, as discussed in a recent literature (Cellini et al., 2010; Gelber et al., 2016), then we are capturing both effects. To more deeply understand our context, note that while we measure no causal effect of hospitalization on subsequent evaluations (Table A.5), 10.2% of treated individuals have another evaluation for involuntary hospitalization in the three-month primary outcome study period. The null effect of hospitalization on subsequent evaluations and hospitalizations is consistent with the fact that physicians' mandates are to evaluate if a person poses a danger at the *present* moment rather than being likely to do so in future moments. It is also consistent with the fact that medical records available to the evaluating physician do not always include prior evaluations, unless perhaps the individual is being evaluated within the same hospital system where they have been treated before.

Note that finding limited indirect effects via subsequent involuntary hospitalization is neither good nor bad for our empirical strategy. Capturing both direct and indirect effects in β is precisely the policy-change of interest.

6 Mechanisms

Several hypotheses could explain the increase in risks resulting from hospitalization among our compliers: disrupted employment, disrupted housing, trauma from the evaluation process itself, destabilized relationships, decreased drug tolerance, and issues with continuing care. These mechanisms are not mutually exclusive. We endeavor to consider the weight of the evidence for each hypothesis.

6.1 Employment and Earnings

Involuntary hospitalization may disrupt employment. For many jobs, an unexplained absence—and especially when it is sustained over several days—is interpreted as quitting or is grounds for termination. Individuals are rarely given an opportunity to call relevant employers before hospitalization. If the employment were a source of structure or daily rhythm, this might contribute to negative impacts. To evaluate this, we use data from Pennsylvania's Department of Labor and Industry, which includes employers and earnings in each quarter from June 2016.[22]

Among our sample, 56% reported an employer at some point during the year before the evaluation. Those who were employed had an average income of $6,065 in the quarter before the evaluation.

In the quarter following an evaluation, there is a statistically significant decline of $1,128 (Table 8A). The decrease persists throughout the quarters following the evaluation. Given the baseline earnings in the quarter before hospitalization, our estimate represents a decrease of 19%. This aligns with a decrease in employment as well (Table 8B), although the decline in employment is only significant in later quarters.

There are two mechanical effects worth discussing. First, earnings may decrease while hospitalized. This may still be costly to the extent that bills such as rent may not also mechanically decrease. Note that the extended effect over the course of several quarters is likely not due only to the mechanical incapacitation since the vast majority of individuals are held for 20-days or fewer (Figure A.1), or approximately 14 business days. That said, we find no statistically significant effect among those who

[22]These data, collated by the state's unemployment insurance system, are requested by the county from the state for reasons unrelated to this research. Not every person's data is requested by the county from the state. We are successfully able to match about 46.5% of our sample. None of the individuals evaluated in the first two years of the data are matched.

are held for 5 or fewer days, suggesting that this effect is concentrated among those who have an extended hospitalization.

Second, those who die or are incarcerated are, mechanically, going to show up as earning less income. As such, we perform the same analysis focusing on those who do not show up in jail and are not dead (for any reason). We find that in the quarter after hospitalization, the decrease in earnings is also a statistically significant decrease of $1,329.[23]

Such employment and earnings disruptions have implications for mortality and crime. An existing literature has documented that job loss can decrease life expectancy (Sullivan and Von Wachter, 2009) and increase the probability of death by suicide and alcohol use (Eliason and Storrie, 2009). Other work has found a strong connection between job loss and property crime and a more tenuous one between job loss and violent crime (Raphael and Winter-Ebmer, 2001; Lin, 2008). Moreover, while our data do not give us insight into the magnitude of out-of-pocket medical expenses individuals incur as a result of involuntary hospitalization, prior work has found that the financial shock, especially for those who are uninsured, can be substantial (Morris and Kleinman, 2020; Dobkin et al., 2018). Taken together, these studies suggest that the employment and earnings disruptions we measure in our setting, along with other possible financial shocks, could be a factor behind the adverse effects of involuntary hospitalization on death and violent crime.

6.2 Housing

We next consider that involuntary hospitalization could disrupt housing. People who have been living with others may find that their relatives or housemates are unwilling to live with someone who has been unwell enough to warrant involuntary hospitalization and has been officially deemed a "danger" to others or themselves. Conversely, if a housemate contributed to their involuntary hospitalization, individuals may not want to return to their prior housing. If individuals have lost their job, they may be unable to pay rent. If a person is living in a shelter, absence for even one night is usually sufficient to lose one's spot in the shelter.

Within our sample, shelter usage is relatively rare. In the year before evaluation, only 3% had used the shelter system. In the month before, that figure is just 0.6%. So that we do not conflate inflows to shelters of people who previously had non-shelter housing and outflows from shelters of individuals

[23]The declines in earnings measured among the whole population and among a subsample that excludes those who mechanically cannot earn money are not statistically significantly different from each other.

who had previously been living in shelters, we evaluate the impact of hospitalization on shelter usage among those who had not stayed in a shelter in the prior year, and separately for those who had been staying in a shelter. We use data from the Allegheny County's shelter system, which documents on a daily basis who is using services.

Involuntary hospitalization increases shelter usage. In the month following a hospitalization, shelter usage increases by 1.5 percentage points among those who had never used shelters before (Table 8C). Limiting to those who neither died nor were incarcerated, the results are virtually unchanged. The elevated shelter usage is statistically significant in the second month as well, and remains elevated but not statistically significant in the following months. There is no statistically significant change among prior shelter users, likely due to power issues arising from small sample size.

6.3 Trust

We explore whether the evaluation process itself may impact subsequent behavior. The hypothesis is that being brought to the hospital emergency department—possibly by police—might already destabilize a person. A related hypothesis is that the evaluation may cause a rift between the individual and people who would otherwise provide social support such as family, friends, and school members.[24]

To examine this possibility, we leverage the fact that the government delegate who chooses whether to certify that a person ought to be brought in for an evaluation is randomly assigned (see Section 2 for a description of the process). Among the 27 delegates who have heard more than 100 such calls, there is substantial heterogeneity in their tendency to certify a person for evaluation, with some certifying 73% of calls, and others certifying 95% (Figure B.2). As such, we use a nearly identical empirical approach as described in Section 4 to estimate the causal effects of experiencing an evaluation, consistent with the literature on multiphase system (Baron et al., 2024). The assumptions of relevance, independence, monotonicity, and exclusion, which are necessary for the validity of this approach, are shown in Appendix B. This instrument passes the same tests applied to our main approach in Section 4.3.[25]

We find no significant impact of being evaluated on subsequently being charged with a violent

[24]We have heard many anecdotal accounts of individuals cutting ties or trying to cut ties with people who have referred them in the wake of a bad hospitalization experience.

[25]Importantly, we find no association between the delegate's tendency to certify and the physician's tendency to hospitalize (Table B.5). Likewise, there is no association between a delegate's tendency to certify and the likelihood of treatment, conditional on being evaluated.

offense or dying from suicide or overdose. While there is a strong, positive first stage, we find no statistically significant effects on our main outcomes (Table A.6(A)). However, the results are extremely noisey and the point estimates suggest increases in negative outcomes. Several possible interpretations could make sense of these results.

Notably, these estimates apply only to people referred by family, friends, or school personnel. This is because physicians and police may self-certify, deciding to bring an individual in for evaluation without discussing first with a delegate. However, heterogeneity analyses find similarly adverse effects among referrers who are family, police, physicians or other individuals (Figure A.6). We do not have strong evidence that the effects occur only for those referred by family or friends, which would be an implication of this particular mechanism.

Second, it is possible that the (non-significant) increases arise because most cases that are certified end up being evaluated and most people who are evaluated are hospitalized. An alternative exercise seeks to eliminate the effect of the hospitalization itself and compares those whose cases are not certified against those whose cases are certified but the examination does not result in a hospitalization. We find similarly (non-significant) increases (Table A.6(B)). This suggests that the effects of evaluation are not solely picking up the effects of hospitalization.

Third, the compliers in this setting—those whose evaluation might be certified by some delegates but not by others—are likely a different group of people from the compliers in our main estimates—who may be hospitalized by some physicians but not others. It would be reasonable to assume that most of the people for whom hospitalization is a physician's judgment call are more likely to be certified for evaluation by delegates. If evaluations are meaningfully more traumatic for those who would always be certified, this analysis may miss some of the effect.

These results, while suffering from less power than we would want, suggest that evaluation itself may not be the main mechanism inducing the adverse impacts of involuntary hospitalization. We note this also because one concern is that when a physician decides *not* to involuntarily hospitalize an individual, the physician usually gives referrals to outpatient services. So one could imagine a scenario in which people who have been involuntary hospitalized and discharged (also with an outpatient referral) are less likely to take up outpatient services, whereas the people who were not hospitalized did use it and found it to be helpful. Then perhaps the adverse effects we document are simply an indication that outpatient services are stabilizing people successfully. While this dynamic would not be a problem for our interpretation, we do not find evidence for it. We find slight (and insignificant) increases in utilization of outpatient services associated with hospitalization (Table A.7). This suggests that being evaluated and then released with a referral for (voluntary) outpatient services is not so successful that it is driving the relative increases in adverse outcomes that we find.

To further plumb a potential loss of trust, we also conduct a heterogeneity analysis that considers the relationship of the referrer to the patient. While generally statistically underpowered, our estimates suggest adverse impacts both for those referred by family and by other referring groups such as social workers (Figure A.6).

6.4 Drug Tolerance

Another possible mechanism is that hospitalization decreases patients' tolerance for the illicit substances that they may be using. When they are discharged from the hospital, they may use their old dosage with especially deleterious effects.

While this hypothesis is a plausible explanation for the increase in deaths by overdose, decreased tolerance to a substance may not increase a person's proclivity toward violence. Even if cravings induce violence to obtain the drug, the connection between decreased drug tolerance and being charged with a violent crime is more tenuous.

Our main estimates in Figure 2 show that the LATE rises over the course of the first few months after an evaluation. If one thinks that, immediately upon discharge, people are returning to illicit drugs with an inaccurate sense of their own tolerance, then the profile for the treatment effect would start high upon release and then perhaps taper off. Thus, the shape that we find suggests either that diminished tolerance is not the explanation or that individuals who are hospitalized maintain sobriety for longer than just the time they are hospitalized before using drugs again.

Nevertheless, we consider whether we see more intense effects among those individuals with a diagnosis of substance use disorder than among those without such a diagnosis (Figure A.8). We find that the increase in the probability of a violent crime charge is present for both those with and without a diagnosis of substance use disorder, consistent with substance tolerance being generally unrelated to violent crime charges. We find a significant increase in deaths by suicide or overdose for those *without* diagnoses and a positive point estimate that is not statistically significant for those with such a diagnosis.

Another approach is to consider heterogeneity based on how long a person was hospitalized. Unfortunately, our data do not document the date of release. Instead, we consider whether a person who experienced a hospitalization that was extended from five days to 20 or more days by the court sees different outcomes than an individual who stayed at most 5 days in the hospital. We find similar effects among both populations.[26]

Together, these analyses do not provide compelling evidence that a decrease in drug tolerance due to abstinence while hospitalized drives our results.

6.5 Continuing Care

Another possible mechanism for the adverse effects we measure is that hospitalization is not paired with continuing care among those who are hospitalized. Possibly, hospitals are not differentially connecting those who are involuntarily hospitalized to continuing care upon discharge or hospitalized patients are unwilling to accept these services, perhaps because of a negative experience with hospitalization or general disengagement.

To test this, we explore whether involuntary hospitalization increases utilization of three types of continuing care. Outpatient care is available for anyone with mental health needs, though observable in our data only for those on Medicaid. Case management, in which a care coordinator helps navigate the medical system and ensure the patient gets connected with the care they need is available for those with a severe mental illness diagnosis and is, likewise, observable only among those who are on Medicaid. Additionally, the Community Treatment Team (CTT) is the local implementation of the Assertive Community Treatment (ACT) model that provides outreach and community-based mental health services to individuals with serious and persistent mental illnesses. CTT provides services such as service coordinators, substance use treatment, and psychiatric and social work counseling three- to five times per week per patient. They serve about 500 people in the county and there are severe capacity constraints, with only about 25 new enrollees in 2024. For CTT, we limit our observation to those with severe mental illnesses, who would be eligible for these programs.

We find no increase in any of the forms of continuing care (Table A.7) for those who are hospitalized. Notably, our data only show *utilization* of continuing care, meaning we are unable to

[26]The similarity of our results across hospitalization duration is consistent with the fact that even when a stay is *allowed* to be longer than 5 days by the CCP team, it does not automatically mean that it will be the full 20 days. It is also consistent with 5 days potentially being a sufficient amount of time to lower a person's drug tolerance, depending on their prior usage.

distinguish between services that have been recommended but not used and services that have not been recommended. As such, we cannot pinpoint precisely where (in)action occurs.

For the subset of the population that has been prescribed antipsychotic medications and is also on Medicaid (9.7% of our total sample), we can see whether involuntary hospitalization increases adherence to the prescribed medications. We compare the share of the days for which a person has filled the prescription for their medication to the total days for which they had a prescription. This is an imperfect measure of adherence since we cannot measure actual ingestion of the medications. Nevertheless, adherence is low, with only 25.9% of the prescribed patient-days having filled prescriptions in the month before evaluation. We find no increase in prescriptions for antipsychotics, nor in adhering to said prescription in the months after (Table A.8). There is, however, a slight spike in prescriptions filled among everyone evaluated because we believe both groups sometimes have prescriptions filled at the hospital.

6.6 Other Mechanisms

There are other potential mechanisms for which we do not have the data to evaluate. However, we enumerate several below to better illustrate the context and encourage future research.

First, exposure to mentally unwell patients who are dealing with unrelated issues can itself be destabilizing. For example, if a person is depressed, exposure to a person who is hallucinating or who has experienced a great deal of other trauma can make navigating one's own mental health journey harder rather than easier. This type of exposure to others' trauma can occur both through roommates and through group therapy.

Second, insurance may not cover all of the costs associated with a hospitalization. In these cases, a person may be saddled with medical debt, which may be destabilizing particularly if a person was in a better economic position beforehand (Morris and Kleinman, 2020; Dobkin et al., 2018).

Third, one of the mechanisms proffered by psychiatrists we have spoken to is that involuntary hospitalization may generate an irregular circadian rhythm that is not conducive to well-being. Hospitalization disrupts supports that impose regularity, such as employment and a familiar place to sleep. Moreover, when a person is hospitalized, their daily routines and structures are, *perforce*, set by hospital staff. When to eat or when to sleep is generally dictated. When this rhythm is removed, it may be difficult to return to one's own older, disrupted routines, especially for individuals with SMI.

7 Policy Responses

A natural question arising from our results is how to adjust policy. We try to address this question in several parts: we consider both the marginal treatment effects (Heckman and Vytlacil, 2005) and the heterogeneity of patient risk and physician traits.

7.1 Marginal Treatment Effects

First, we aim to understand the heterogeneous treatment effects across unobservable patient characteristics by exploring marginal treatment effects (MTEs) (Heckman and Vytlacil, 2005). MTEs help us with the thought experiment of dialing up or down the share of people who are hospitalized. Specifically, the MTE estimates help us consider how involuntary hospitalization might impact those individuals who would be on the cusp of hospitalization as we move from physicians more likely to hospitalize to physicians who are less likely to hospitalize. This also helps shed light on whether our local average treatment effects are likely to apply to those whom more or fewer physicians would tend to hospitalize.

MTE estimates may only be calculated where there are physicians who both hospitalize and do not hospitalize individuals, what is referred to as the area of common support. In our setting, this ranges from hospitalizing 67% to 86% of those evaluated, when trimming 5% off of each side of the distribution (Figure A.10A).

To calculate the MTE function, we first predict the probability of hospitalization with physician tendency to hospitalize as the only explanatory variable, in line with the literature (e.g., Doyle, 2007; Arnold et al., 2022). Next, we predict the relationship between our two main outcomes and the predicted probability of involuntary hospitalization using a local quadratic estimator with a bandwidth of 0.08 (Heckman et al., 2006).[27] We calculate the numerical first derivative of the local quadratic estimator. Standard errors are computed from the standard deviation of 250 bootstrapped iterations of the MTE procedure, each of which cluster resamples at the rotation level.[28]

[27] A larger bandwidth will yield flatter MTE estimates. We opt for the minimum bandwidth, minimizing the sum of squared errors between the local quadratic estimator and a fourth-degree polynomial model for each outcome. Higher and lower polynomials and both larger and smaller bandwidths yield similar results.

[28] MTEs require strong assumptions of pairwise monotonicity and separability between observed and unobserved heterogeneity in treatment effects in order to be identified over the common support of the propensity score. These assumptions have been the subject of a robust discussion (Frandsen et al., 2023; Mogstad et al., 2021; Norris et al., 2021). As such, MTE estimates should be interpreted with caution.

Figure A.10B shows the treatment effects for people just on the margin of hospitalization as we move from individuals who are unobservably more likely to be hospitalized (left side of the graph) to those who are unobservably less likely to be hospitalized (right side of the graph). The figure shows an MTE function for being charged with a violent crime that is decreasing as the physician's predicted probability of hospitalization increases (and therefore the unobservable likelihood of being hospitalized in the first place decreases) and an MTE function for death by suicide or overdose that goes in the opposite direction.

This suggests that for individuals who would otherwise be least likely to be hospitalized (that is, the ones who are on the margin of hospitalization for the physicians that hospitalize the most individuals), involuntary hospitalization has the most extreme increase in likelihood of death by suicide or overdose. But these same individuals have the largest decrease in likelihood of being charged with a violent offense. That suggests that in some sense, those who are the "least deserving" of involuntary hospitalization see the biggest increase in death by suicide or overdose yet also a decrease in being charged with a violent crime.

Given that the MTE curves are not statistically significant, we do not place much weight on the fact that the curves go in opposite directions.

We also attempt to explore whether there is heterogeneity by predicted risk. This exercise would be helpful if a program could predict which individuals are likely to experience significant treatment effects. Unfortunately, our results are too noisy to be useful (Appendix C).

7.2 Heterogeneous Effects Based on Physician Traits

Another policy response could be to require that only specific physicians can perform evaluations. Indeed, some jurisdictions require that the evaluating physician be a psychiatrist. In contrast, the setting we study has any emergency department physician perform this function. This analysis is further motivated by the fact that physicians with certain traits have different hospitalization rates. For example, psychiatrists and those with fewer than 10 years of experience tend to have lower hospitalization rates than other physicians (Figure A.11 left-most figures).

We therefore conduct heterogeneity analyses to explore whether certain types of physicians have different results among the individuals they evaluate. The data are underpowered to find statistically significant differences, so these analyses should be considered purely suggestive.

We find that individuals evaluated by psychiatrists tend to have similar outcomes as those evaluated by other emergency room physicians. Individuals who are evaluated by physicians with fewer than 10 years of experience (fewer than 10 years between NPI enumeration date and the date of the evaluation) have similar mortality outcomes to those evaluated by physicians with more than 10 years of experience. However, individuals seen by more experienced physicians have suggestively smaller increases in the probability of a violent crime charge. There are no identifiable differences in outcomes based on physician gender.

8 Costing Exercise

We conduct a costing exercise to tally up the direct and indirect costs associated with involuntary hospitalization. This is perhaps best thought of as an opportunity-cost exercise: how much is the system we study spending in direct and indirect costs to involuntarily hospitalize people that *could* be dedicated to an intervention that successfully reduces danger? The analysis is necessarily crude, but helpful for benchmarking.

Per internal Allegheny County Department of Human Services calculations, in Allegheny County it costs $120 to evaluate a person and $670 to house them in a mental health inpatient facility for one day. Medicaid billing indicates that the typical stay lasts 12.5 days and costs about $670 per day, indicating a cost of $8,375 per hospitalization.

Involuntary hospitalization causes an increase in the probability of being charged with a violent crime. We do not know exactly how closely this accords with committing a violent crime. To ensure our estimates are conservative, we assume that only half of the charges are associated with actual violent crimes having been committed. Applying this estimate only to the complier share of the approximately 1,663 annual evaluations in our analysis,[29] we estimate 11.79 more violent crimes per year. Estimates of the cost of a single violent crime range from $119,186 to $125,084 in 2025 dollars (Miller et al., 2021; Chalfin, 2015).[30] Taking the lower estimate, we would see a cost of about $1,405,000 per year in increased violent crimes as a result of the treatment effects of involuntary hospitalization.

[29]Our sample is limited by age, first-time involuntary hospitalization experience, and other criteria, further contributing to the conservative nature of our costing estimates.

[30]Miller et al. (2021) estimates that violent crime costs $18,386 in tangible costs (costs imposed on victims, costs of police investigations, costs of incarcerating offenders, etc.) and $100,800 in quality of life costs per violent crime, in 2025 dollars.

Additionally, involuntary hospitalization causes an increase in death by suicide or drug overdose. Over the course of a year, with approximately 1,663 evaluations, a complier share of 0.4296 and a treatment effect of 0.02436, implies 17.15 additional deaths per year among compliers.

The health literature often estimates the value of a statistical life at approximately $10M in the US (Kniesner and Viscusi, 2019). The typical age of the individual who dies by suicide or overdose in our sample is 40, implying a loss of 40 years. Accounting for disability weights associated with severe mental illness[31] this implies a loss of life worth $66,700,000 per year due to involuntary hospitalization.

These estimates are limited to first-time evaluations for involuntary hospitalizations among 18-64 year olds, among other limitations. Relaxing these sample restrictions would lead to different, possibly higher, cost estimates.

9 Conclusion

Involuntary hospitalization is a last-resort approach for handling individuals experiencing a mental crisis. Its use is common and current policy initiatives from local leaders in the US suggest its use is likely to expand. Given this policy context, it is important to understand the causal impact of involuntary hospitalization on its statutory goals of reducing people's likelihood of being a danger to themselves or others.

Our paper finds that for individuals whose evaluations are judgment calls insofar as some physicians would choose to hospitalize and some would not, involuntary hospitalization is not reducing danger. On the contrary, we find that in judgment call cases involuntary hospitalization nearly doubles the risk of being charged with a violent crime in the months following evaluation and likewise nearly doubles the risk of dying by suicide or overdose. For judgment call cases, we find evidence that hospitalization decreases employment, decreases earnings, and increases homeless shelter usage. We see no evidence that in judgment call cases hospitalization increases an individual's likelihood of being connected to continuing care such as outpatient visits three months after the evaluation. Taken together, while involuntary hospitalization may temporarily stabilize an individual with a judgment call case, it comes with significant disruption costs that play out over time and do not generate improved connections to

[31] As noted in our summary statistics, only about half of the evaluated individuals have been diagnosed with severe mental illness. However, we use the disability weights associated with severe mental illness rather than other (less severe) mental illnesses and apply them to *all* individuals to ensure that if our estimates are inaccurate, it is because they are *under*estimates of the true cost.

care upon release.

Our paper leaves numerous avenues open for additional research. We consider the effects of the whole involuntary hospitalization process, rather than identifying the effects of particular components of treatment. However, it is possible that specific treatments have different effects; understanding if this is the case would yield useful, actionable insight for policymakers. We do not address the general equilibrium effects of whether more or less stringent policies on involuntary hospitalization would impact trust in the medical system among those who would be hospitalized and those who would refer them. Our paper has less power to discuss heterogeneous treatment effects than we would like, which limits our ability to explore possible policy recommendations. We focus on one specific jurisdiction and may not be able to speak to the effectiveness of involuntary hospitalization in other areas, where the process is different. Further research is also needed to develop better forms of care for people facing psychiatric emergencies, both inside and outside emergency and inpatient settings. More work should be done to assist physicians in their decision-making process, especially in high-risk situations like the ones we analyze, and to reduce the variance across physicians in the tendency to hospitalize. Finding ways to more effectively target scarce healthcare resources towards this high-risk population also has the potential to improve care for people needing non-psychiatric emergency and inpatient care. Ethical, patient- and family-rights considerations are also missing from this paper and should be explored further.

Despite the areas not tackled in this paper, the results we present have important policy implications. As policymakers consider expanding involuntary hospitalization, our local average treatment effects provide insight into what would happen if physicians making the decision to hospitalize were slightly more expansive in who is hospitalized. The results we present suggest that expanding involuntary hospitalization in our setting is unlikely, without additional changes, to reduce danger to others or to the patients themselves.

References

Abadie, Alberto, "Semiparametric Instrumental Variable Estimation of Treatment Response Models," *Journal of Econometrics*, 2003, *113* (2), 231–263.

Aizer, Anna and Joseph J Jr Doyle, "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges," *Quarterly Journal of Economics*, 2015, *139* (2), 759–803.

Allegheny County Department of Human Services, "DHS Data Warehouse," `https://www.alleghenycounty.us/Services/Human-Services-DHS/News-and-Events/Accomplishments-and-Innovations/DHS-Data-Warehouse` 2024. Accessed: 2024-03-24.

Anderson, Theodore W and Herman Rubin, "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *The Annals of Mathematical Statistics*, 1949, *20* (1), 46–63.

Andrews, Isaiah, James H Stock, and Liyang Sun, "Weak Instruments in Instrumental Variables Regression: Theory and Practice," *Annual Review of Economics*, 2019, *11*, 727–753.

Angst, Jules, Felix Angst, Regina Gerber-Werder, and Alex Gamma, "Suicide in 406 Mood-Disorder Patients With and Without Long-Term Medication: A 40 to 44 Years' Follow-Up," *Archives of Suicide Research*, 2005, *9* (3), 279–300.

Arnold, David, Will Dobbie, and Peter Hull, "Measuring Racial Discrimination in Bail Decisions," *American Economic Review*, 2022, *112* (9), 2992–3038.

Baldessarini, Ross J, Leonardo Tondo, John Hennen et al., "Effects of Lithium Treatment and its Discontinuation on Suicidal Behavior in Bipolar Manic-Depressive Disorders," *Journal of Clinical Psychiatry*, 1999, *60* (2), 77–84.

Banerjee, Souvik, Pinka Chatterji, and Kajal Lahiri, "Effects of Psychiatric Disorders on Labor Market Outcomes: A Latent Variable Approach Using Multiple Clinical Indicators," *Health Economics*, 2017, *26* (2), 184–205.

Baron, E. Jason, Joseph J. Doyle Jr., Natalia Emanuel, Peter Hull, and Joseph Ryan, "Discrimination in Multi-Phase Systems: Evidence from Child Protection," *Quarterly Journal of Economics*, 2024.

Biasi, Barbara, Michael S Dahl, and Petra Moser, "Career Effects of Mental Health," *National Bureau of Economic Research*, 2021.

Brouwers, Rob C, Martin Appelo, and TI Oei, "Mental Illness and Crime: Impulsive Violent Behaviour: Factors and Forms," in "Transnational Criminology Manual, Volume 2 (Crime)," Wolf Legal Publishers (WLP), 2010, pp. 489–505.

Brown, Danielle J., "Should People with Severe Mental Health Needs Get Court-Ordered Treatment? State Senate to Decide," *Maryland Matters*, April 2024. Accessed: 2025-03-24.

Bütikofer, Aline, Christopher J Cronin, and Meghan M Skira, "Employment Effects of Healthcare Policy: Evidence from the 2007 FDA Black Box Warning on Antidepressants," *Journal of Health Economics*, 2020, *73*, 102348.

Carson, E. Ann, "Prisoners in 2021 – Statistical Tables," `https://bjs.ojp.gov/library/publications/prisoners-2021-statistical-tables` December 2022. NCJ Number 305125.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein, "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design," *The Quarterly Journal of Economics*, 2010, *125* (1), 215–261.

Chalfin, Aaron, "Economic Costs of Crime," *The Encyclopedia of Crime and Punishment*, 2015, pp. 1–12.

Chan, Joe Kwun Nam, Christoph U Correll, Corine Sau Man Wong, Ryan Sai Ting Chu, Vivian Shi Cheng Fung, Gabbie Hou Sem Wong, Janet Hiu Ching Lei, and Wing Chung Chang, "Life Expectancy and Years of Potential Life Lost in People with Mental Disorders: A Systematic Review and Meta-Analysis," *EClinicalMedicine*, 2023, *65.*

CNN, "2 Dead, 7 Hurt After Man Opens Fire at Pittsburgh Psychiatric Hospital," March 2012. Accessed 1 February 2025.

Costantini, Sydney, "How do Mental Health Treatment and Treatment Delays Impact Long Term Mortality?," Technical Report, Working Paper 2024.

Dahl, Gordon B, Andreas Ravndal Kostol, and Magne Mogstad, "Family Welfare Cultures," *Quarterly Journal of Economics*, 2014, *129* (4), 1711–1752.

Davis, Lisa, Anthony Fulginiti, Liat Kriegel, and John S Brekke, "Deinstitutionalization? Where Have All the People Gone?," *Current Psychiatry Reports*, 2012, *14*, 259–269.

de Sousa, Tanya and Meghan Henry, "The 2024 Annual Homelessness Assessment Report to Congress," Technical Report, US Department of Housing and Urban Development, Office of Community Planning and Development 2024. Accessed: 2025-03-20.

Deisenhammer, E. A., C. M. Ing, R. Strauss, G. Kemmler, H. Hinterhuber, and E. M. Weiss, "The Duration of the Suicidal Process: How Much Time Is Left for Intervention Between Consideration and Accomplishment of a Suicide Attempt?," *The Journal of Clinical Psychiatry*, 2009, *70* (1), 19–24.

Dembosky, April, Amelia Templeton, and Carrie Feibel, "When Homelessness and Mental Illness Overlap, is Forced Treatment Compassionate?," *NPR*, March 2023. Accessed: 2025-03-24.

Deza, Monica, Thanh Lu, Johanna Catherine Maclean, and Alberto Ortega, "Treatment for Mental Health and Substance Use: Spillovers to Police Safety," *National Bureau of Economic Research*, 2023.

_ , _ , _ , and _ , "Losing Medicaid and Crime," *National Bureau of Economic Research*, 2024.

Dobbie, Will, Jacob Goldin, and Crystal S Yang, "The Effects of Pre-Trial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges," *American Economic Review*, 2018, *108* (2), 201–240.

Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, and Matthew J Notowidigdo, "The Economic Consequences of Hospital Admissions," *American Economic Review*, 2018, *108* (2), 308–352.

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

Eliason, Marcus and Donald Storrie, "Does Job Loss Shorten Life?," *Journal of Human Resources*, 2009, *44* (2), 277–302.

FBI, "2019 Crime in the United States," `https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/violent-crime` 2019. Accessed: 2024-02-10.

Fisher, Austin, "NM Governor Shares Draft Proposal for Forced Mental Health Treatment," *Source New Mexico*, June 2024. Accessed: 2025-03-24.

Frandsen, Brigham, Lars Lefgren, and Emily Leslie, "Judging Judge Fixed Effects," *American Economic Review*, 2023, *113* (1), 253–277.

Gelber, Alexander, Adam Isen, and Judd B Kessler, "The Effects of Youth Employment: Evidence from New York City Lotteries," *The Quarterly Journal of Economics*, 2016, *131* (1), 423–460.

Global Burden of Disease Collaborative Network, "Global Burden of Disease Study 2019 (GBD 2019) Disability Weights," *Institute for Health Metrics and Evaluation (IHME)*, 2020. Accessed: 2025-01-01.

Gold, Mark S, "The Role of Alcohol, Drugs, and Deaths of Despair in the US's Falling Life Expectancy," *Missouri Medicine*, 2020, *117* (2), 99.

Goldman-Mellor, Sidra J, Harish S Bhat, Michael H Allen, and Michael Schoenbaum, "Suicide Risk Among Hospitalized Versus Discharged Deliberate Self-Harm Patients: Generalized Random Forest Analysis Using a Large Claims Data Set," *American Journal of Preventive Medicine*, 2022, *62* (4), 558–566.

Goodwin, Renee D, Andrea H Weinberger, June H Kim, Melody Wu, and Sandro Galea, "Trends in Anxiety Among Adults in the United States, 2008–2018: Rapid Increases Among Young Adults," *Journal of Psychiatric Research*, 2020, *130*, 441–446.

Gross, Max and E Jason Baron, "Temporary Stays and Persistent Gains: The Causal Effects of Foster Care," *American Economic Journal: Applied Economics*, 2022, *14* (2), 170–199.

Hallett, Nutmeg, Rachel Dickinson, Emachi Eneje, and Geoffrey L Dickens, "Adverse Mental Health Inpatient Experiences: Qualitative Systematic Review of International Literature," *International Journal of Nursing Studies*, 2024, p. 104923.

Heckman, James J. and Edward Vytlacil, "Structural Equations, Treatment Effects, and Econometric Policy Evaluation 1," *Econometrica*, 2005, *73* (3), 669–738.

_ , Sergio Urzua, and Edward Vytlacil, "Understanding Instrumental Variables in Models with Essential Heterogeneity," *The Review of Economics and Statistics*, 2006, *88* (3), 389–432.

Hedman, Leslie C, John Petrila, William H Fisher, Jeffrey W Swanson, Deirdre A Dingman, and Scott Burris, "State Laws on Emergency Holds for Mental Health Stabilization," *Psychiatric Services*, 2016, *67* (5), 529–535.

Hirschauer, John, "Why Pennsylvania Failed to Keep Its Governor Safe," *City Journal*, April 2025.

Imbens, G W and J D Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62*, 467–475.

Institute for Health Metrics and Evaluation, "GBD Compare Data Visualization," 2024. Accessed: 2025-01-01.

Institute, National Cancer, "Statistics at a Glance: The Burden of Cancer in the United States," *National Institute of Health*, 2024. Accessed: 2025-03-20.

Jácome, Elisa, "Mental Health and Criminal Involvement: Evidence from Losing Medicaid Eligibility," *Job Market Paper, Princeton University*, 2020.

Jordan, Joshua T and Dale E McNiel, "Perceived Coercion During Admission Into Psychiatric Hospitalization Increases Risk of Suicide Attempts After Discharge," *Suicide and Life-Threatening Behavior*, 2020, *50* (1), 180–188.

Kallert, Thomas W, Matthias Glöckner, and Matthias Schützwohl, "Involuntary vs. Voluntary Hospital Admission: A Systematic Literature Review on Outcome Diversity," *European Archives of Psychiatry and Clinical Neuroscience*, 2008, *258*, 195–209.

Katsakou, Christina, Len Bowers, Tim Amos, Richard Morriss, Diana Rose, Til Wykes, and Stefan Priebe, "Coercion and Treatment Satisfaction Among Involuntary Patients," *Psychiatric Services*, 2010, *61* (3), 286–292.

Kessler, Daniel and Mark McClellan, "Do Doctors Practice Defensive Medicine?," *The Quarterly Journal of Economics*, 1996, *111* (2), 353–390.

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

Kniesner, Thomas J. and W. Kip Viscusi, "The Value of a Statistical Life," 07 2019.

Kolesár, Michal, R. Chetty, J. Friedman, E. Glaeser, and G. Imbens, "Identification and Inference with Many Invalid Instruments," *Journal of Business and Economic Studies*, 2015, *33*, 474–484.

Larkin, Gregory Luke, Cynthia A Claassen, Jennifer A Emond, Andrea J Pelletier, and Carlos A Camargo, "Trends in US Emergency Department Visits for Mental Health Conditions, 1992 to 2001," *Psychiatric Services*, 2005, *56* (6), 671–677.

Lee, Gi and David Cohen, "Incidences of Involuntary Psychiatric Detentions in 25 US States," *Psychiatric Services*, 2021, *72* (1), 61–68.

Leslie, Emily and Nolan G Pope, "The Unintended Impact of Pretrial Detention on Case Outcomes: Evidence from New York City Arraignments," *The Journal of Law and Economics*, 2017, *60* (3), 529–557.

Lin, Ming-Jen, "Does Unemployment Increase Crime?: Evidence from US Data 1974–2000," *Journal of Human Resources*, 2008, *43* (2), 413–436.

Liu, Nancy H, Gail L Daumit, Tarun Dua, Ralph Aquila, Fiona Charlson, Pim Cuijpers, Benjamin Druss, Kenn Dudek, Melvyn Freeman, Chiyo Fujii et al., "Excess Mortality in Persons with Severe Mental Disorders: A Multilevel Intervention Framework and Priorities for Clinical Practice, Policy and Research Agendas," *World Psychiatry*, 2017, *16* (1), 30–40.

McGivern, Lauri, Leanne Shulman, Jan K Carney, Steven Shapiro, and Elizabeth Bundock, "Death Certification Errors and the Effect on Mortality Statistics," *Public Health Reports*, 2017, *132* (6), 669–675.

MHPA, "Mental Health Procedures Act," *Congress of the State of Pennsylvania*, 1976.

Miller, Ted R, Mark A Cohen, David I Swedler, Bina Ali, and Delia V Hendrie, "Incidence and Costs of Personal and Property Crimes in the USA, 2017," *Journal of Benefit-Cost Analysis*, 2021, *12* (1), 24–54.

Mogstad, Magne, Alexander Torgovitsky, and Christopher R Walters, "The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables," *American Economic Review*, 2021, *111* (11), 3663–3698.

Morris, Nathaniel P and Robert A Kleinman, "Involuntary Commitments: Billing Patients for Forced Psychiatric Care," *American Journal of Psychiatry*, 2020, *177* (12), 1115–1116.

Mueller-Smith, Michael, "The Criminal and Labor Market Impacts of Incarceration," 2015. Unpublished working paper, No. 18.

Nesbit, Rachel, "The Role of Mandated Mental Health Treatment in the Criminal Justice System," Technical Report, Working Paper 2023. `http://arxiv.org/abs/2212.06736`.

Newman, Andy and Emma G. Fitzsimmons, "New York City to Involuntarily Remove Mentally Ill People From Streets," *The New York Times*, June 2023. Accessed: 2025-03-24.

Norris, Samuel, Matthew Pecenco, and Jeffrey Weaver, "The Effects of Parental and Sibling Incarceration: Evidence from Ohio," *American Economic Review*, 2021, *111* (9), 2926–2963.

Olea, José Luis Montiel and Carolin Pflueger, "A Robust Test for Weak Instruments," *Journal of Business & Economic Statistics*, 2013, *31* (3), 358–369.

Oreskes, Benjamin and Andy Newman, "Will New York Force More Mentally Ill People Into Treatment?," *The New York Times*, February 2025. Accessed: 2025-03-24.

Passmore, Kirsty and WC Leung, "Defensive Practice Among Psychiatrists: A Questionnaire Survey," *Postgraduate Medical Journal*, 2002, *78* (925), 671–673.

Penrose, Lionel Sharples, "Mental Disease and Crime: Outline of a Comparative Study of European Statistics.," *British Journal of Medical Psychology*, 1939.

Puras, Dainius, "Report of the Special Rapporteur on the Right of Everyone to the Enjoyment of the Highest Attainable Standard of Physical and Mental Health," *Human Rights Council*, 2017, *35*, 1–21. `https://documents.un.org/doc/undoc/gen/g17/076/04/pdf/g1707604.pdf?OpenElement`. Accessed: 2025-02-15.

Qin, Ping and Merete Nordentoft, "Suicide Risk in Relation to Psychiatric Hospitalization: Evidence Based on Longitudinal Registers," *Archives of General Psychiatry*, 2005, *62* (4), 427–432.

Raphael, Steven and Michael A Stoll, "Assessing the Contribution of the Deinstitutionalization of the Mentally Ill to Growth in the US Incarceration Rate," *The Journal of Legal Studies*, 2013, *42* (1), 187–222.

_ **and Rudolf Winter-Ebmer**, "Identifying the Effect of Unemployment on Crime," *The Journal of Law and Economics*, 2001, *44* (1), 259–283.

Riley, Fiona, "D.C. Passes Bill Updating Involuntary Commitment Process for People with Severe Mental Illness," *Street Sense Media*, December 2024. Accessed: 2025-03-24.

Rosen, Jonathan, *The Best Minds: A Story of Friendship, Madness, and the Tragedy of Good Intentions*, Random House, 2023.

Saya, Anna, Chiara Brugnoli, Gioia Piazzi, Daniela Liberato, Gregorio Di Ciaccia, Cinzia Niolu, and Alberto Siracusano, "Criteria, Procedures, and Future Prospects of Involuntary Treatment in Psychiatry Around the World: A Narrative Review," *Frontiers in Psychiatry*, 2019, *10*, 271.

Simpson, Sylvia G and Kay R Jamison, "The Risk of Suicide in Patients with Bipolar Disorders," *Journal of Clinical Psychiatry*, 1999, *60* (2), 53–56.

Stock, James, Jonathan H. Wright, and Motohiro Yogo, "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, 2002, *20* (4), 518–529.

Studdert, David M, Michelle M Mello, William M Sage, Catherine M DesRoches, Jordon Peugh, Kinga Zapert, and Troyen A Brennan, "Defensive Medicine Among High-Risk Specialist Physicians in a Volatile Malpractice Environment," *JAMA*, 2005, *293* (21), 2609–2617.

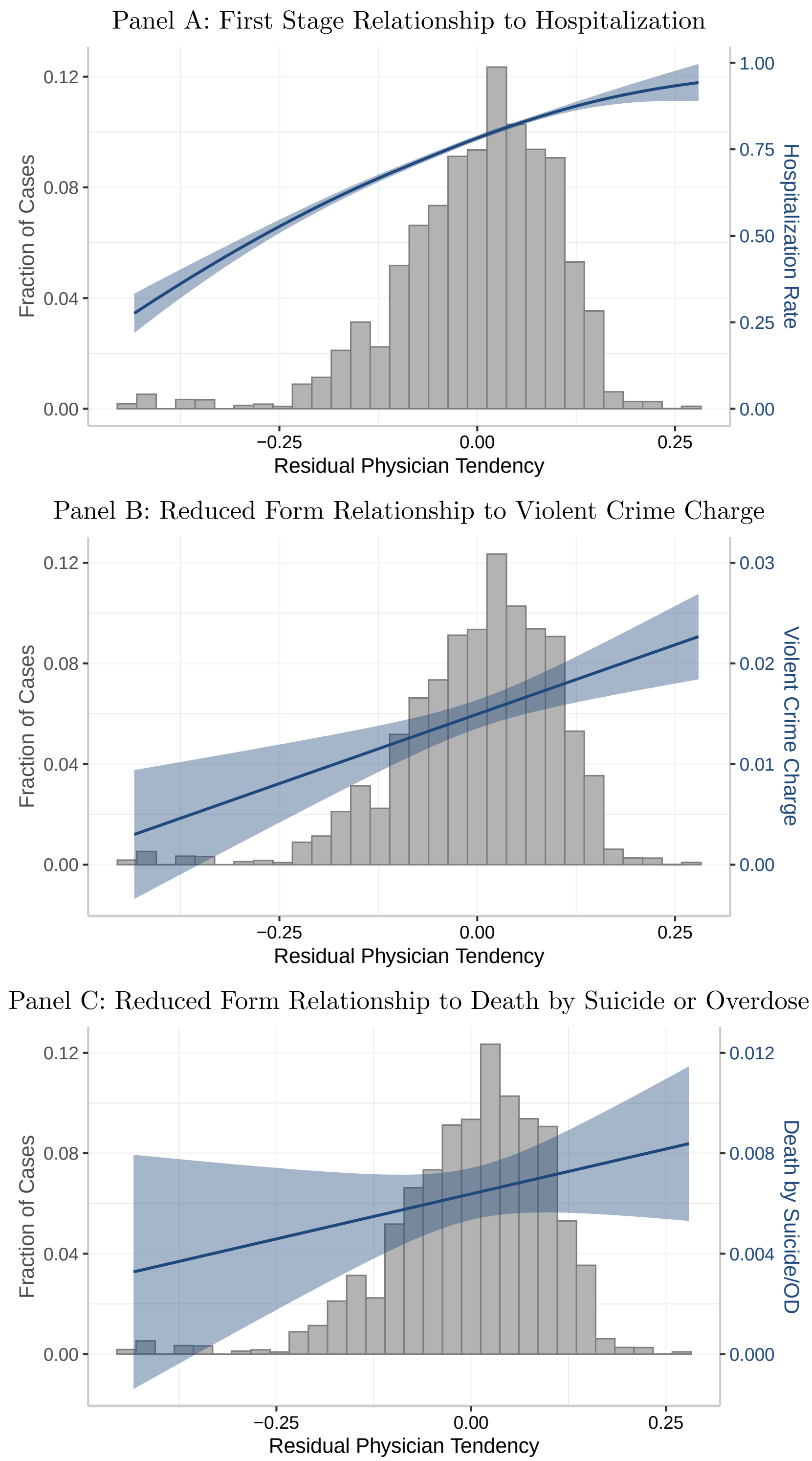
Sullivan, Daniel and Till Von Wachter, "Job Displacement and Mortality: An Analysis Using Administrative Data," *The Quarterly Journal of Economics*, 2009, *124* (3), 1265–1306.

Theriault, Kayla M, Robert A Rosenheck, and Taeho Greg Rhee, "Increasing Emergency Department Visits for Mental Health Conditions in the United States," *The Journal of Clinical Psychiatry*, 2020, *81* (5), 5456.

Twenge, Jean M, A Bell Cooper, Thomas E Joiner, Mary E Duffy, and Sarah G Binau, "Age, Period, and Cohort Trends in Mood Disorder Indicators and Suicide-Related Outcomes in a Nationally Representative Dataset, 2005–2017.," *Journal of Abnormal Psychology*, 2019, *128* (3), 185.

Welle, Pim, Erika Montana, Nic Marlton, and Shuyan Zhan, "Analysis of Allegheny County's Involuntary Hospitalization (302) Program," *Allegheny County Department of Human Services*, 2023. Accessed: 2025-06-16.

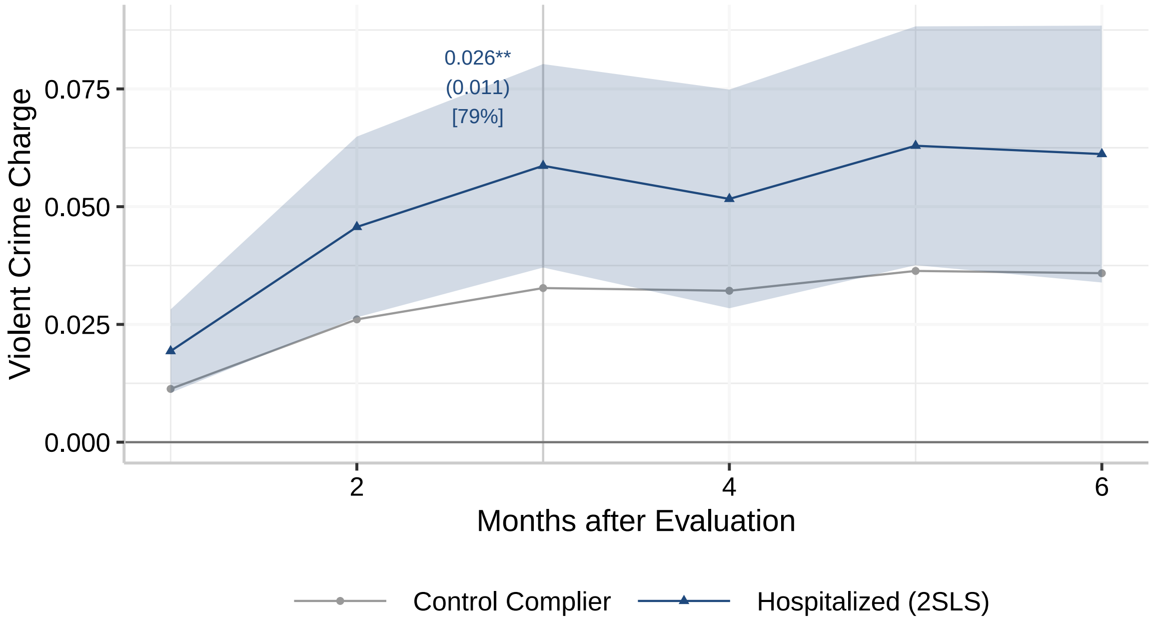
Figure 1: Distribution of Physician Tendency Instrument



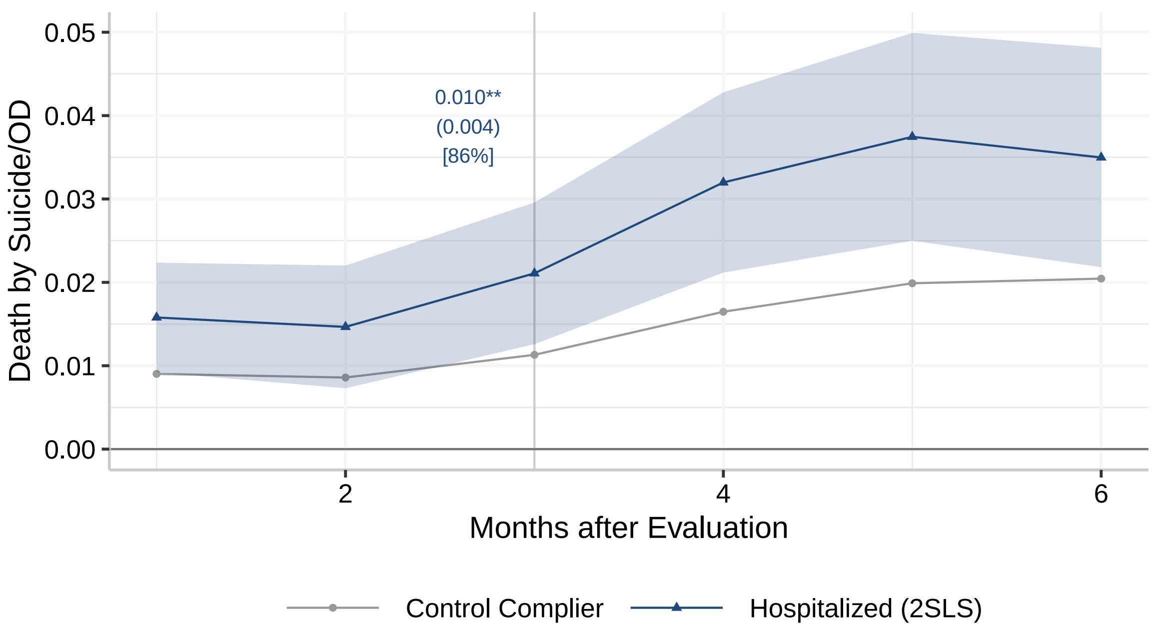
Note: This figure shows the distribution of the physician tendency to hospitalize along with the involuntary hospitalization evaluations instrument. The instrument construction is described in Section 4.1. The blue line shows the relationship between the instrument and hospitalization (top), violent crime charge (middle), and death by suicide or overdose (bottom), all in the reduced form.

Figure 2: Local Average Treatment Effects

Panel A: Violent Crime Charge



Panel B: Death by Suicide or Overdose



Note: This figure shows the local average treatment effect of hospitalization on the probability of a violent crime charge in Panel A and death by suicide or overdose in Panel B for judgment call cases. Y-axis shows the share of people who are charged with a violent crime (Panel A) and the share who have died as a result of suicide or overdose (Panel B). The ribbons indicate 95-percent confidence intervals based on standard errors clustered at the physician level. The annotations indicate the point estimate at 3 months, with standard errors in parentheses and percent change over the control complier means in square brackets. Table 7 shows the numerical version of these estimates.

Table 1: Summary Statistics

	All	Hospitalized	Not Hospitalized	Compliers
(A) Demographics				
Female	0.447	0.442	0.463	0.342*
Age 18-35	0.508	0.497	0.543	0.468
White	0.660	0.664	0.644	0.733*
Black	0.273	0.267	0.294	0.23
Education				
Education Unknown	0.232	0.234	0.224	0.171
High School or Less	0.532	0.527	0.551	0.545
High School/GED	0.167	0.169	0.159	0.23**
Bachelor's or More	0.069	0.070	0.067	0.054
On TANF/SNAP	0.222	0.207	0.273	0.266
On SSI	0.127	0.127	0.127	0.182**
Enrolled in Medicaid	0.447	0.434	0.491	0.558***
(B) Past Diagnoses (Medicaid & County-Paid Population)				
Serious Mental Illness	0.408	0.447	0.286	0.347**
Any Mental Illness	0.500	0.509	0.472	0.447**
Substance Use Disorder	0.183	0.183	0.180	0.168**
(C) Prior Encounters				
Prior Criminal Charge	0.223	0.215	0.252	0.338**
Prior Noncriminal Charge	0.096	0.092	0.109	0.146**
Prior ED Visit	0.302	0.292	0.337	0.506***
(D) 302 Evaluation Traits				
Reporter				
Relative	0.488	0.504	0.435	0.633**
Police Officer	0.225	0.196	0.324	0.259
Other Reporter	0.287	0.300	0.241	0.108*
Concern				
Attempted Suicide	0.510	0.484	0.598	0.586
Unable to Care for Self	0.441	0.465	0.356	0.38
Mutilated Self	0.063	0.059	0.075	0.066
Danger to Others	0.472	0.475	0.459	0.532
Observations	16,630	12,909	3,721	[0.43]

Note: This table describes our study sample of individuals evaluated for involuntary hospitalization in Allegheny County for 2014-2023. Information on past diagnoses is only available for the individuals who are enrolled in Medicaid or have some county-paid claims. The first column contains information on the whole sample. The middle two columns show the population based on whether they are hospitalized or not. The final column shows the observable traits for those who are compliers—those individuals for whom the evaluating physician changes whether or not they would be hospitalized. The final row shows the number of evaluations in the whole sample, evaluations among those hospitalized and not hospitalized, and in square brackets the share of compliers calculated. See the Section 4.2 for additional details on the sample and variable construction. *p< 0.1; **p< 0.05; ***p< 0.01.

Table 2: LATE First Stage: Physician Tendency and Hospitalization

	Hospitalized						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Physician Tendency	0.92***	0.88***	0.83***	0.83***	0.83***	0.83***	0.89***
	(0.03)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)
Hospital FEs		✓					
Rotation FEs			✓	✓	✓	✓	✓
Demographic Controls				✓	✓	✓	✓
History Controls					✓	✓	✓
Concern Controls						✓	✓
Medical Controls							✓
Mean Hospitalization	0.776	0.776	0.776	0.776	0.776	0.776	0.776
Effective F	158.11	120.23	109.19	107.08	104.83	106.17	89.15
Observations	16,630	16,630	16,630	16,630	16,630	16,630	7,434

Note: This table shows first stage results, documenting the relationship between our residualized physician tendency to hospitalize a patient instrument and their likelihood of hospitalization. The regressions are estimated in the our main sample, using an instrument constructed as described in Section 4.1. Each column adds successive control variables. Demographic controls include sex, age, racial category, education and medicaid enrollment. History controls include prior jail stay, prior use of a shelter, prior violent and criminal charges, prior emergency department use, prior use of community treatment teams, prior use of case management. Concern controls include indicators for whether the referring individual indicated that the individual had attempted suicide, posed a danger to others, had mutilated themselves or was unable to care for themselves. Medical controls, which are only available for those enrolled in Medicaid or who have had bills paid by the county include whether the individual has a severe mental illness, a mental illness, or a substance use diagnosis. Because the final set of controls is available only for a subset of cases, the penultimate specification is our preferred specification used throughout the rest of the paper. Standard errors are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table 3: Randomization Test

	Hospitalized	Physician Tendency
F-Statistic	10.901	1.1916
p-value	0.0000	0.25
Observations	16,630	16,630

Note: This table relates observable characteristics to an individuals' likelihood of being hospitalized (Column 1) and to the assigned physician's tendency to hospitalize on a involuntary hospitalization evaluation (Column 2). The F-statistic jointly tests the significance of the control variables and the p-value of the joint test is reported below. Standard errors are clustered at the physician level. For the sake of space, the coefficients for Column 2 are shown in Figure A.4, Panel A.

Table 4: First Stage by Demographic Characteristic

	Hospitalized									
	Female	White	Some HS	< 35 Years	Medicaid	SMI	MH	DA	SNAP	SSI
Tendency	0.81***	0.83***	0.85***	0.81***	0.95***	0.80***	0.96***	1.09***	1.01***	0.94***
	(0.10)	(0.08)	(0.10)	(0.09)	(0.12)	(0.10)	(0.10)	(0.16)	(0.12)	(0.14)
Rotation Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	7,426	10,972	8,855	8,049	4,914	3,843	5,438	1,624	3,693	2,107

Note: This table presents the first stage relating a physician's tendency to hospitalize a patient on their hospitalization. The columns that subset to individuals who have a diagnosis of severe mental illness (SMI), of a mental illness (MH) or of substance use disorder (DA), are available only for individuals who are enrolled in Medicaid or have had a County-paid medical claim. SNAP and SSI indicate that a person was enrolled in TANF/SNAP or Social Security Insurance at the time of the examination. Standard errors are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table 5: First Stage by Involuntary Hospitalization Evaluation Concern

	Hospitalized			
	Unable to Care	Danger to Others	Suicide Att	Mutilated Self
Physician Tendency	0.68***	0.79***	0.89***	0.94***
	(0.06)	(0.07)	(0.12)	(0.27)
Rotation Fixed Effects	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓
Observations	7,331	7,843	8,473	1,042

Note: This table presents the first stage relating a physician's tendency to hospitalize a patient to the petitioner's concerns about the patient. Standard errors are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table 6: Physician Tendency and Downstream Events

	Prescribed Antipsychotics	CCP Hearing	CCP Extension
Physician Tendency	−0.117	−0.011	−0.056
	(0.078)	(0.062)	(0.055)
Observations	5,608	12,899	12,902

Note: This table considers the relationship between a physician's tendency to hospitalize individuals and the downstream events among those who are hospitalized. For those enrolled in Medicaid, physician tendency to hospitalize is unrelated to whether the person has a been prescribed antipsychotics. Conditional upon hospitalization, the physician's tendency is unrelated to whether an individual has a hearing with the Court of Common Pleas (CCP) or whether the magistrate extends the individual's stay. *p< 0.1; **p< 0.05; ***p< 0.01.

Table 7: Impacts of Involuntary Hospitalization on Violent Crime Charge and Death by Suicide/Overdose

	Reduced Form	LATE Estimate	Control Complier Mean
Violent Crime Charge	0.021**	0.026**	[0.033]
	(0.008)	(0.011)	
Suicide/OD Death	0.008**	0.010**	[0.011]
	(0.004)	(0.004)	
Observations	16,630	16,630	

Note: This table presents the reduced form and instrumental variables (LATE) estimates for the impact of involuntary hospitalization on 3-month outcomes that proxy for being a danger to others (violent crime charge) or to oneself (death by suicide or overdose). The control complier means are in square brackets. Standard errors are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table 8: Impacts of Involuntary Hospitalization on Earnings and Employment

Panel A: Earnings

	-2Q	-1Q	0Q	1Q	2Q	3Q
Hospitalized	−45.77	−63.09	−1,679.00**	−1,128.00**	−1,036.00**	−548.40
	(502.20)	(432.30)	(784.50)	(491.90)	(506.30)	(493.20)
Control Complier Mean	2,581.1	2,460.9	2,196.7	1,986.2	2,166.9	2,134.3
Observations	7,782	7,725	7,683	7,222	6,941	6,707

Panel B: Employment

	-2Q	-1Q	0Q	1Q	2Q	3Q
Hospitalized	0.010	−0.033	−0.007	−0.056	−0.040	−0.112*
	(0.046)	(0.040)	(0.056)	(0.069)	(0.059)	(0.067)
Control Complier Mean	0.243	0.267	0.301	0.239	0.239	0.255
Observations	7,782	7,725	7,683	7,222	6,941	6,707

Note: This table shows the local average treatment effect of involuntary hospitalization on earnings and employment. Estimates are derived from estimating equation 3. The sample is limited to those individuals for whom Allegheny County has data from the Pennsylvania Department of Labor and Industry. Regressions show the effects in the indicated quarter relative to the quarter of the evaluation (0Q). Regressions include the usual controls as well as earnings or employers over the quarters prior to evaluation. Standard errors in all panels are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table 9: Impacts of Involuntary Hospitalization on Shelter Usage

Panel A: Shelter Usage Among Non-Users

	1 Mo	2 Mo	3 Mo	4 Mo	5 Mo	6 Mo
Hospitalized	0.015*	0.016*	0.015	0.014	0.014	0.009
	(0.008)	(0.009)	(0.011)	(0.014)	(0.014)	(0.014)
Control Complier Mean	0.0124	0.0211	0.0245	0.0262	0.0286	0.0253
Observations	16,234	16,116	16,006	15,891	15,769	15,695

Panel B: Shelter Usage Among Prior Users

	1 Mo	2 Mo	3 Mo	4 Mo	5 Mo	6 Mo
Hospitalized	0.318	0.063	0.087	−0.248	−0.001	−0.166
	(0.473)	(0.666)	(0.693)	(0.714)	(0.654)	(0.652)
Control Complier Mean	0.8944	0.8192	0.3129	0.1345	0.2999	0.3681
Observations	307	306	305	303	303	301

Note: This table shows the local average treatment effect of involuntary hospitalization on homeless shelter usage among those with and without a prior history of homeless shelter usage. Estimates are derived from estimating equation 3. The table divides the whole sample among people who have not used a shelter in the year prior to the evaluation and those who have. Regressions show cumulative effects over the indicated months after an evaluation. Standard errors in all panels are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Appendix to

A Danger to Self and Others: Health and Criminal Consequences of Involuntary Hospitalization

Natalia Emanuel · Pim Welle · Valentin Bolotnyy

July 16, 2025

A Appendix: Additional Tables and Figures

Table A.1: Sample Restrictions

Restriction	Evaluations	People	Physicians	Hospitals
All Cases	45,100	27,689	2,168	18
Limit to first case	27,689	27,689	1,810	18
Exclude voluntary hospitalizations	26,623	26,623	1,783	18
Limit to MDs with $\geq$ 15 cases	22,978	22,978	427	18
Limit to patients with known sex	22,978	22,978	427	18
Limit to patients over age 18	20,985	20,985	427	18
Limit to age $<$ 65	17,226	17,226	426	18
Limit to hospitals that see $>$ 40 people	17,151	17,151	424	14
Limit to rotations with variation in cases/physicians	16,630	16,630	424	14

Note: This table shows the number of evaluations, individuals, physicians and hospitals in our sample at each stage of the sample construction process.

Table A.2: Effects of Involuntary Hospitalization on Comparison Outcomes

Panel A: All (Violent and Non-Violent) Charges

	1 Mo	2 Mo	3 Mo	4 Mo	5 Mo	6 Mo
	(1)	(2)	(3)	(4)	(5)	(6)
Hospitalized	−0.005	0.014	0.015	0.024	0.023	0.015
	(0.012)	(0.015)	(0.018)	(0.021)	(0.021)	(0.023)
Control Complier Mean	0.021	0.0498	0.0665	0.0849	0.0846	0.0864
Observations	16,554	16,441	16,339	16,237	16,129	16,063

Panel B: Natural Death

	1 Mo	2 Mo	3 Mo	4 Mo	5 Mo	6 Mo
	(1)	(2)	(3)	(4)	(5)	(6)
Hospitalized	−0.003	−0.005	−0.003	−0.005	−0.012	−0.012
	(0.004)	(0.005)	(0.006)	(0.007)	(0.008)	(0.008)
Control Complier Mean	0.0018	0.0002	0.004	0.0026	-0.0026	-0.0033
Observations	16,483	16,337	16,202	16,064	15,915	15,826

Note: This table shows the instrumental variables relationship between involuntary hospitalization and comparison outcomes: being charged with any crime (even a non-violent crime), and death for a natural cause. Regressions show cumulative effects over the indicated months after an evaluation. Standard errors are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table A.3: LATE Estimate Robustness

			Violent Crime Charge		**Suicide/OD Death**		
		First St.	Red. Form	LATE	Red. Form	LATE	N
1	No Controls	0.829***	0.022**	0.027**	0.009**	0.011**	16,630
		(0.079)	(0.009)	(0.011)	(0.004)	(0.005)	
2	6 months	0.826***	0.021**	0.025*	0.012**	0.015**	16,630
		(0.080)	(0.010)	(0.014)	(0.005)	(0.007)	
3	12 months	0.826***	0.028**	0.033**	0.020**	0.024**	16,630
		(0.080)	(0.012)	(0.015)	(0.008)	(0.010)	
4	> 5 Cases	0.732***	0.015**	0.020**	0.011**	0.015**	18,153
		(0.081)	(0.007)	(0.010)	(0.005)	(0.007)	
5	> 30 Cases	0.916***	0.021**	0.023**	0.011***	0.012**	14,135
		(0.071)	(0.010)	(0.011)	(0.004)	(0.005)	
6	Clustered by Rotation	0.826***	0.021***	0.026***	0.008**	0.010**	16,630
		(0.040)	(0.007)	(0.009)	(0.004)	(0.005)	
7	Clustered by Rotation-Physician	0.826***	0.021***	0.026**	0.008**	0.010**	16,630
		(0.078)	(0.008)	(0.010)	(0.003)	(0.004)	
8	No Police	0.729***	0.033***	0.046***	0.008*	0.011*	12,890
		(0.073)	(0.010)	(0.016)	(0.004)	(0.006)	
9	No Physician	0.954***	0.022**	0.024**	0.009*	0.009*	14,294
		(0.101)	(0.010)	(0.011)	(0.004)	(0.005)	
10	Inpatient	0.819***	0.019**	0.023**	0.008**	0.010*	12,735
		(0.077)	(0.009)	(0.012)	(0.004)	(0.005)	
11	Allow Incapacitation	0.826***	0.021**	0.026**	0.006	0.007	16,630
		(0.080)	(0.008)	(0.011)	(0.004)	(0.004)	
12	Address In County	0.863***	0.025***	0.029**	0.010**	0.012**	14,298
		(0.085)	(0.009)	(0.012)	(0.004)	(0.005)	
13	Balanced Panel	0.828***	0.022***	0.026**	0.008**	0.010**	15,788
		(0.082)	(0.008)	(0.011)	(0.004)	(0.004)	
14	Sans First Year of Data	0.867***	0.021**	0.025**	0.011***	0.013***	14,721
		(0.084)	(0.009)	(0.011)	(0.004)	(0.005)	
15	UJIVE Instrument	0.649***	0.015**	0.023**	0.004	0.006	16,625
		(0.058)	(0.006)	(0.011)	(0.003)	(0.005)	
16	Split Sample Instrument	0.587***	0.012*	0.020*	0.009**	0.015**	16,614
		(0.081)	(0.007)	(0.011)	(0.004)	(0.006)	
17	Physician FE Instrument	0.826***	0.021**	0.017**	0.008**	0.005	16,630
		(0.080)	(0.008)	(0.007)	(0.004)	(0.004)	
18	Split Danger to Others Instrument	0.591***	0.020***	0.034**	0.009**	0.016**	16,615
		(0.093)	(0.007)	(0.014)	(0.004)	(0.007)	
19	Split Suicide Attempt Instrument	0.620***	0.017**	0.028**	0.010***	0.015**	16,614
		(0.088)	(0.007)	(0.012)	(0.004)	(0.006)	
	Control Mean	0.776	0.01	[0.033]	0.004	[0.011]	16,630

Note: This table presents robustness of the estimates shown in Table 7. Each row shows the estimates where one of our design decisions is changed slightly. The bottom row shows the average hospitalization rate, the outcomes in the control population under the reduced form columns and the control complier means in square brackets. Where not otherwise specified, standard errors are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table A.4: Subgroups of Interest

			Violent Crime Charge		Suicide/OD Death		
		First St.	Red. Form	LATE	Red. Form	LATE	N
1	Stay 5 Days or Fewer	0.957***	0.015	0.016	0.013*	0.014*	11,873
		(0.070)	(0.012)	(0.012)	(0.007)	(0.007)	
2	Psychiatric Hospital	0.585***	0.036**	0.062**	0.004	0.007	6,978
		(0.079)	(0.014)	(0.028)	(0.007)	(0.011)	
3	Non-psychiatric Hospitals	0.953***	0.014	0.014	0.011***	0.012**	9,652
		(0.084)	(0.011)	(0.011)	(0.004)	(0.005)	
4	No Medicaid at Time of Exam	0.754***	0.022**	0.030**	0.009*	0.012	9,195
		(0.075)	(0.009)	(0.012)	(0.005)	(0.007)	
5	Medicaid at Time of Exam	0.925***	0.026*	0.028	0.008	0.009	7,434
		(0.091)	(0.015)	(0.017)	(0.007)	(0.007)	
	Control Mean	0.776	0.01	[0.033]	0.004	[0.011]	16,630

Note: This table presents core results for several subgroups of interest. Each row shows a different subgroup. The bottom row shows the average hospitalization rate, the outcomes in the control population under the reduced form columns and the control complier means in square brackets. Where not otherwise specified, standard errors are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table A.5: Effects of Involuntary Hospitalization on Subsequent Interactions

	Subsequent Interactions	
	Evaluation	Hospitalization
	(1)	(2)
Hospitalized	−0.013	−0.001
	(0.031)	(0.031)
Control Complier Mean	0.1067	0.1023
Observations	16,613	16,613

Note: This table shows the local average treatment effect of involuntary hospitalization on subsequent evaluation and subsequent involuntary hospitalization. Standard errors are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table A.6: Impacts of Evaluation

Panel A: Main Effects

	Reduced Form	LATE Estimate	Control Complier Mean
Violent Crime Charge	0.002	0.003	[-0.002]
	(0.029)	(0.031)	
Suicide/OD Death	0.005	0.006	[0.011]
	(0.014)	(0.015)	
Observations	14,102	14,102	

Panel B: Limiting to Unhospitalized Individuals

	Reduced Form	LATE Estimate	Control Complier Mean
Violent Crime Charge	0.101	0.051	[-0.002]
	(0.062)	(0.032)	
Suicide/OD Death	0.001	0.000	[0.011]
	(0.028)	(0.015)	
Observations	4,947	4,947	

Note: This table presents the reduced form and instrumental variables (LATE) estimates for the impact of an *evaluation* for involuntary hospitalization on 3-month outcomes that proxy for being a danger to others (violent crime charge) or to oneself (death by suicide or overdose). Panel A shows all people whose cases were considered by a delegate. Panel B limits the sample to those individuals whose cases were not certified, or were certified but the individuals were not hospitalized. The estimation strategy uses random assignment of delegates' tendency to certify a case, as described in Section 2 and analyzed in Section 6. The control complier means are in square brackets. Standard errors are clustered at the delegate level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table A.7: Local Average Treatment Effects of Hospitalization on Continuing Care

	Outpatient	Case Management	Community Treatment Team
	(1)	(2)	(3)
Hospitalized	0.028	−0.020	0.011
	(0.063)	(0.143)	(0.022)
Control Complier Mean	0.4733	0.2518	0.0183
Population	Medicaid	Medicaid + SMI	Medicaid + SMI
Observations	7,346	2,993	2,972

Note: This table shows the local average treatment effect of involuntary hospitalization on outpatient treatment, case management and participation in community treatment teams. The sample is limited to people on Medicaid and, in columns (2) and (3), also to those with severe mental illness. Standard errors are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

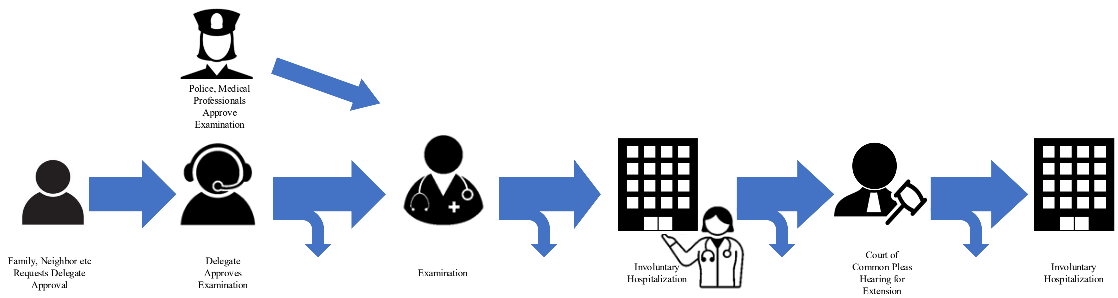
Table A.8: Effects of Hospitalization on Adherence to Antipsychotic Medication

	Days of Filled Antipsychotic Medication Prescription						
	Prescribed	1 Mo	2 Mo	3 Mo	4 Mo	5 Mo	6 Mo
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hospitalized	0.026	0.377	−8.461	−31.460	−34.750	−49.910	−62.040
	(0.028)	(10.380)	(19.760)	(29.400)	(39.190)	(47.660)	(58.230)
Control Mean	0.097	11.615	23.521	35.188	46.76	58.807	70.078
Observations	16,630	1,621	1,621	1,621	1,621	1,621	1,621

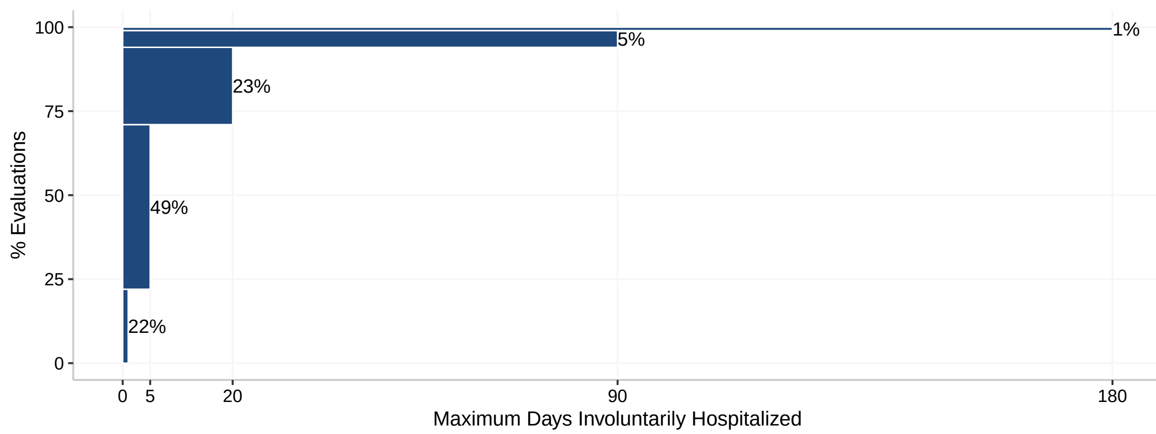
Note: This table shows the local average treatment effect of involuntary hospitalization on having a prescription for an antipsychotic medication (column 1), and the number of days of prescription in which a person has filled their prescription for the antipsychotic medication. The sample for columns 2-7 is limited to those who had a prescription for an antipsychotic. 1 Mo - 6 Mo indicate months since evaluation and the effects shown are cumulative over those months. Standard errors are clustered at the physician level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Figure A.1: Involuntary Hospitalization Process

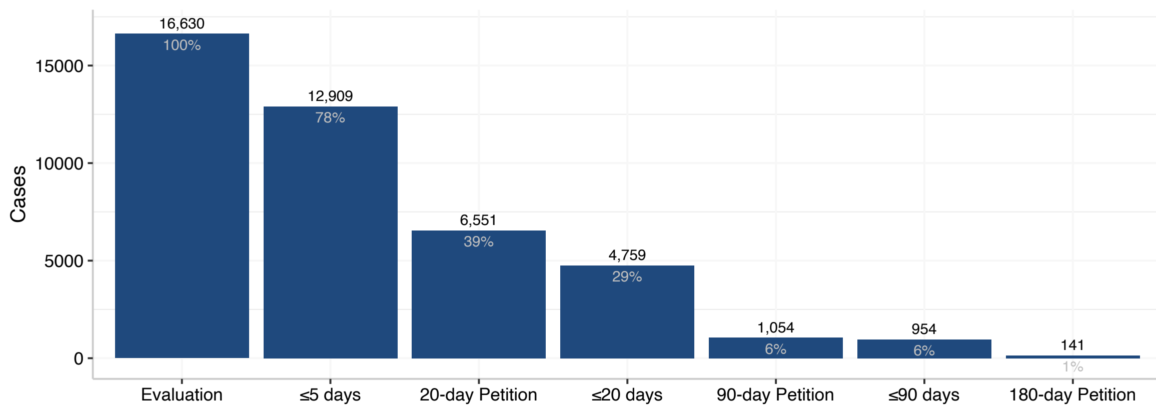
Panel A: Administrative Process



Panel B: Share of Evaluations that are Hospitalized and For How Long

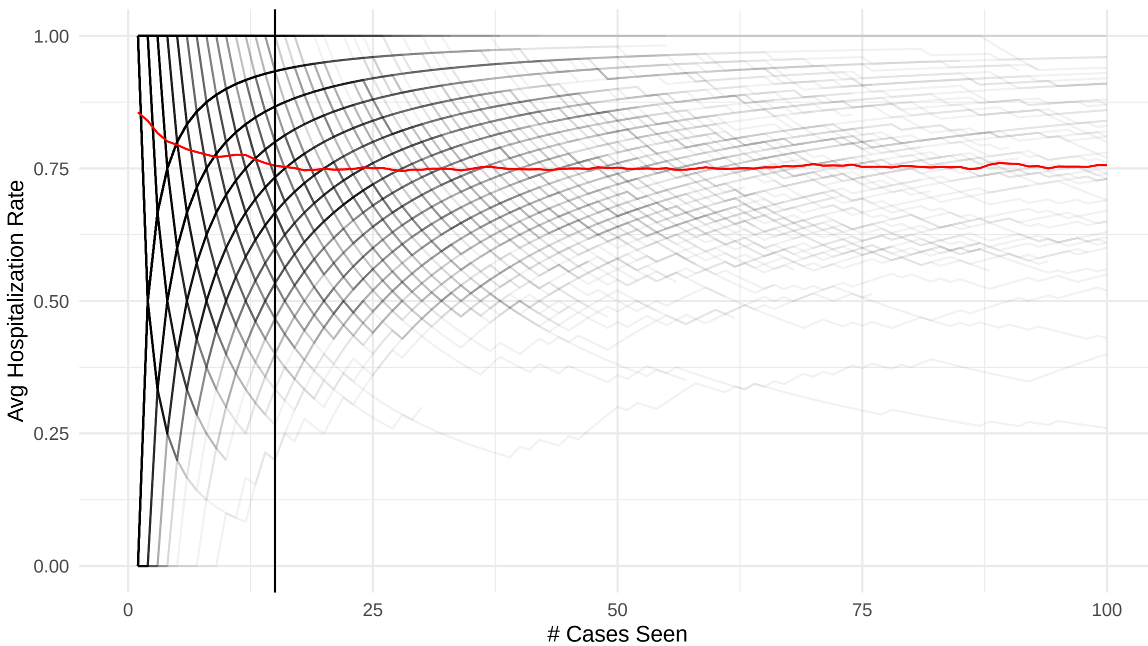


Panel C: Hearings, Petitions, and Maximum Day Held



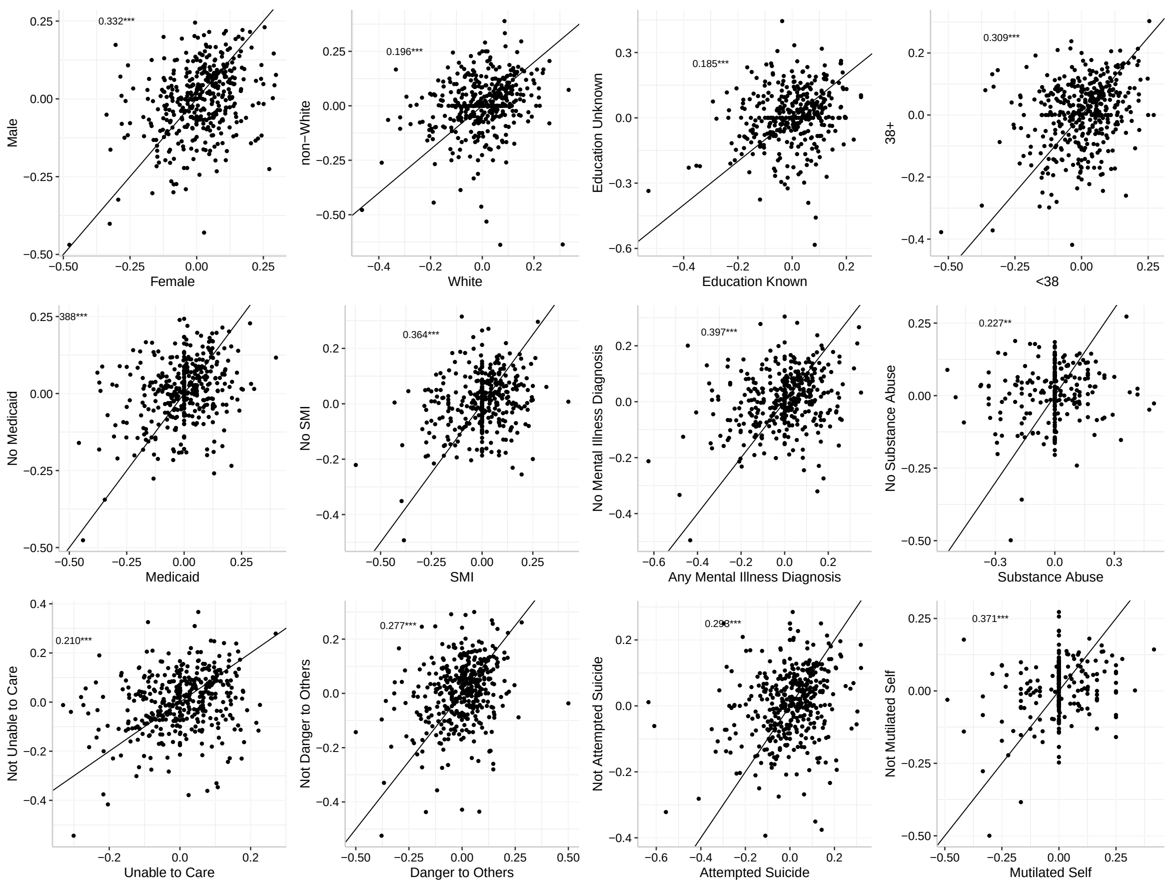
Note: Panel A shows the steps in the administrative process in our setting. First family members, neighbors or other concerned lay-people contact a government delegate that either certifies or does not certify a referral. Police and medical professionals may "self-certify." Individuals' whose cases have been certified are brought to the emergency department for an evaluation of whether they post a danger to themselves or others. The physician decides whether to hospitalize them. If so, they may be held for up to five days in an inpatient mental health unit. A Court of Common Pleas magistrate hears the case for an extension if the inpatient physicians believe more time is needed for a person to be stabilized. The magistrate may extend the hospitalization for up to 20, 90 or 180 days. Our main point of quasi-randomization occurs at the Examination phase. Panel B shows the share of first-time evaluations that result in hospitalization and the maximum days that the individuals can be held. Panel C shows the number of people evaluated for the first time and the share of initial evaluations who progress to the next step.

Figure A.2: Hospitalization Rate Across Physicians as Case Number Increases



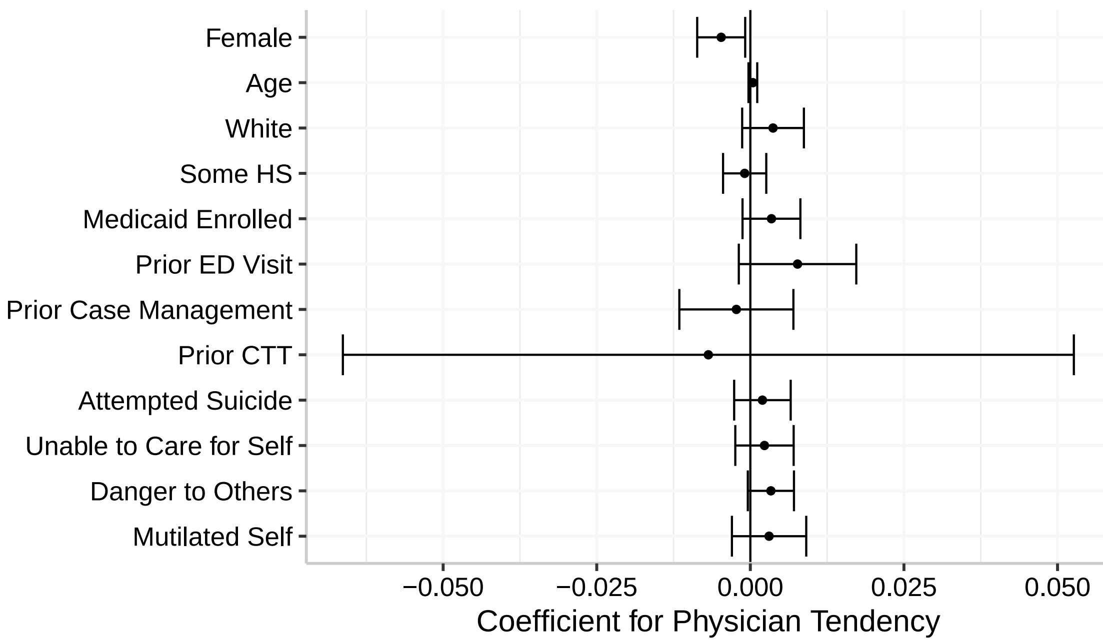
Note: This figure plots physicians' average hospitalization rate as a function of how many evaluations for involuntary hospitalization they have seen. The red line shows the average hospitalization rate. The vertical black line shows 15 cases, which is the number required in our main sample. Table A.3 shows our main estimates in samples that include physicians who have seen only 5 or more cases, and physicians who have seen 30 or more cases.

Figure A.3: Physician Tendency Instrument by Case Characteristics



Note: This figure shows the correlation between a residualized measure of physician tendency to hospitalize for different groups of patients. Each point shows the average of the residualized tendency within each subgroup for each physician. The lines show the OLS relationship between physicians' tendencies to hospitalize within one group relative to the other group. The coefficient is noted in the top left of each graph. *p< 0.1; **p< 0.05; ***p< 0.01.

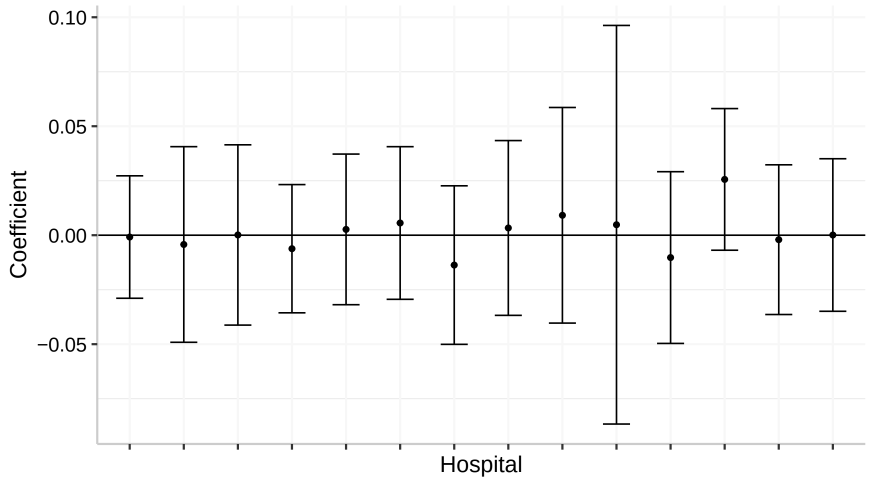
Figure A.4: Randomization Test: Relationship of Observables to Physician Tendency



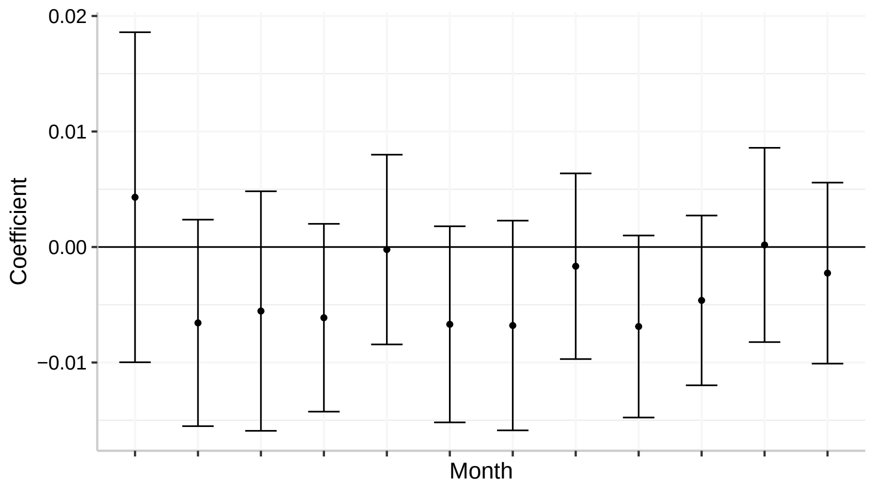
Note: This figure shows coefficients from Table 3, column 2, relating observable traits to our instrument for a physician's tendency to hospitalize a patient. The fact that nearly every predictor is not significant lends confidence that assignment to evaluating physician is indeed random. The coefficient on prior violent crime charge is mildly significant, which is not beyond the pale given the number of coefficients we include. All coefficients are from a single regression and standard errors are clustered at the physician level.

Figure A.5: Randomization Test: Hospital, Season and Shift

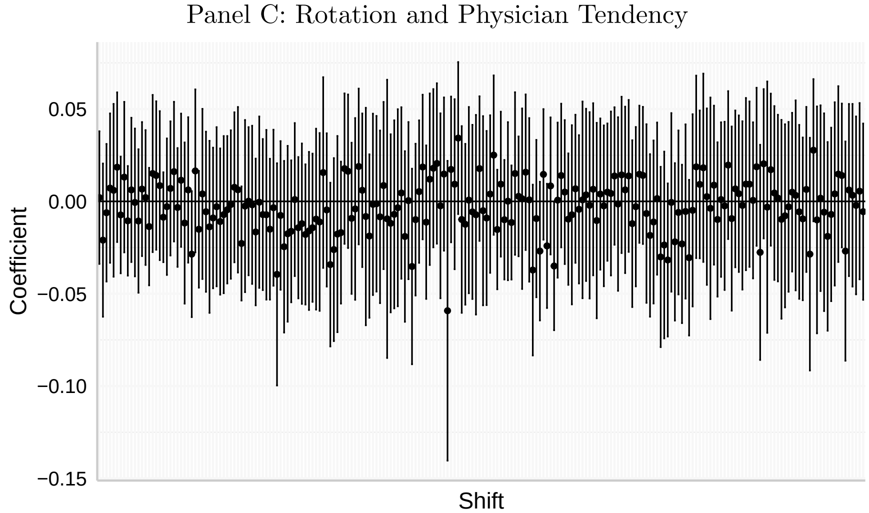
Panel A: Hospitals and Physician Tendency



Panel B: Month and Physician Tendency

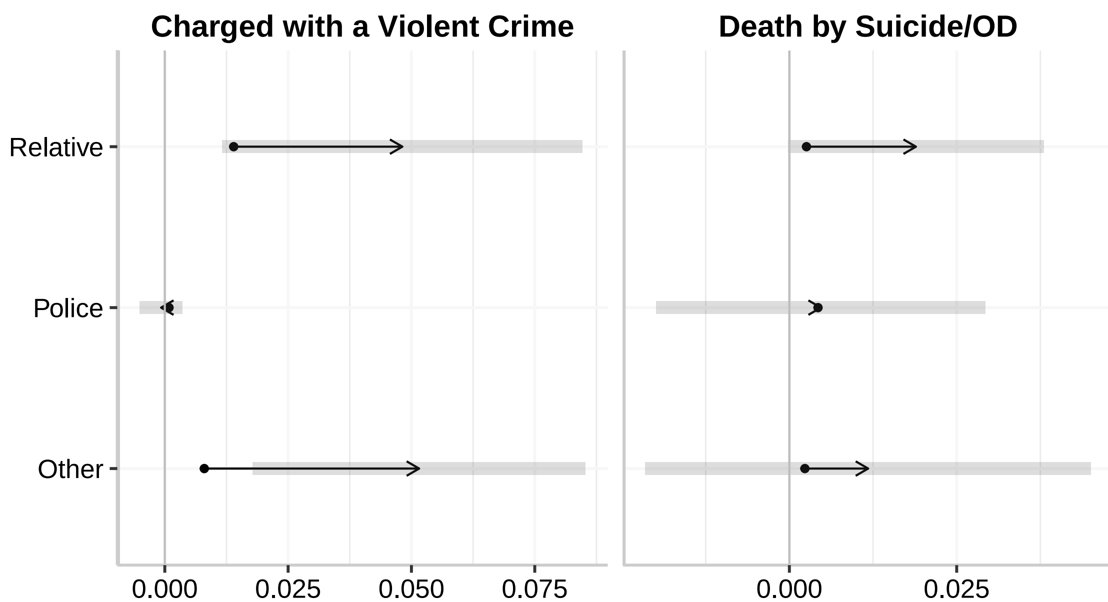


Panel C: Rotation and Physician Tendency



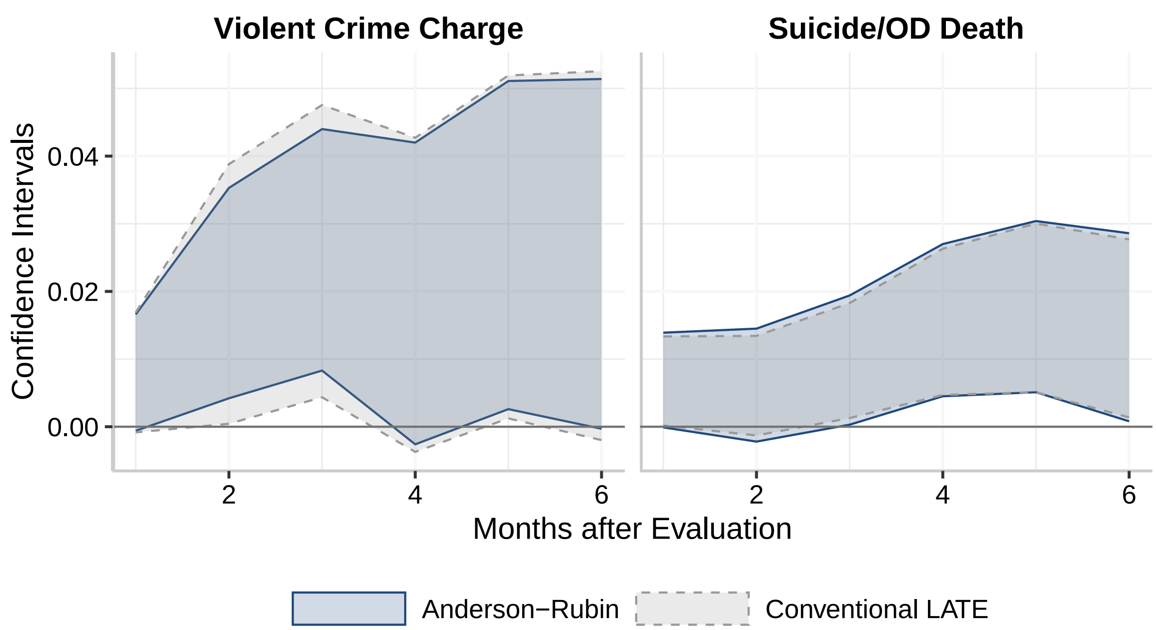
Note: This figure shows that neither the hospital, nor the month, nor the rotation is systematically related to the physician tendency. All coefficients in each panel are from a single regression and standard errors are clustered at the physician level.

Figure A.6: Heterogeneity by Referrer



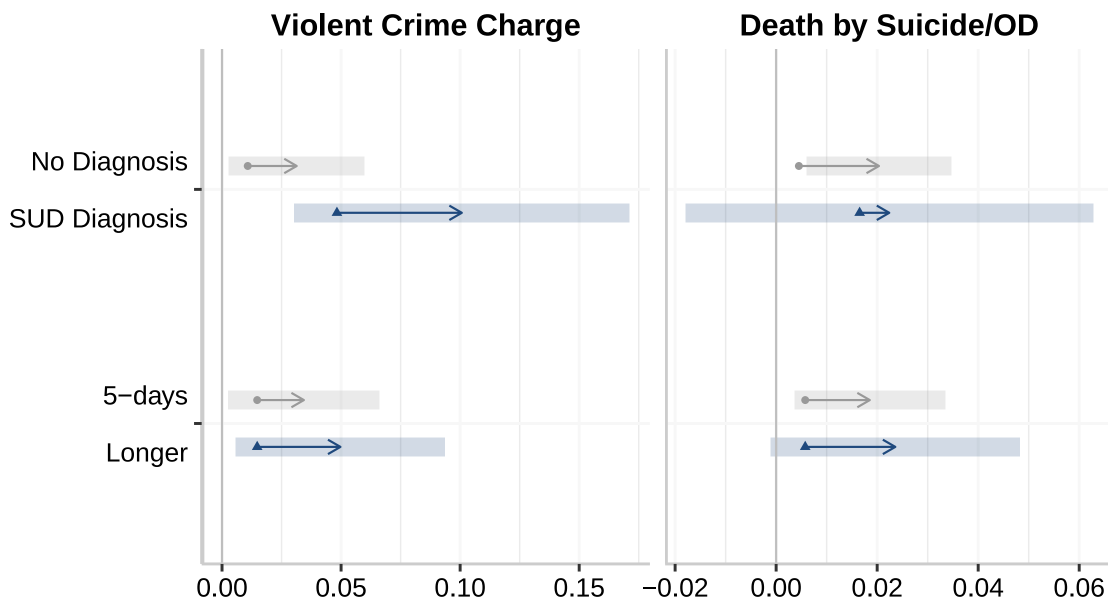
Note: This figure shows the point estimates for the effects of involuntary hospitalization based on whether the referral originated with a family member, police office or another party. Outcomes are measured 6 months after the evaluation. The dots show the control mean; the arrows the point estimate; the shaded region the confidence intervals. Each panel shows estimates from three different regressions.

Figure A.7: Anderson-Rubin Confidence Intervals



Note: This figure shows Anderson-Rubin confidence intervals in the months following an evaluation for both violent crime charges and for deaths by suicide or overdose, in blue outlined with a solid line. These intervals are impervious to weak instruments. For comparison, we include conventional LATE confidence intervals as well, outlined in a grey dashed line. See Section 5.3 for more details.

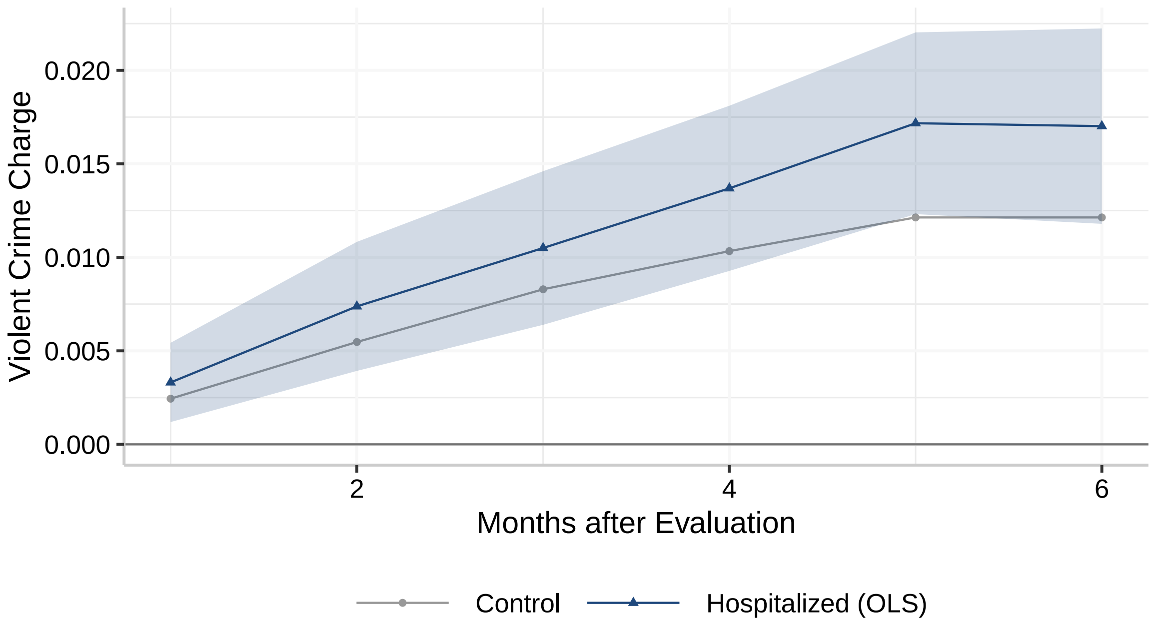
Figure A.8: Heterogeneity by Substance Use Tolerance



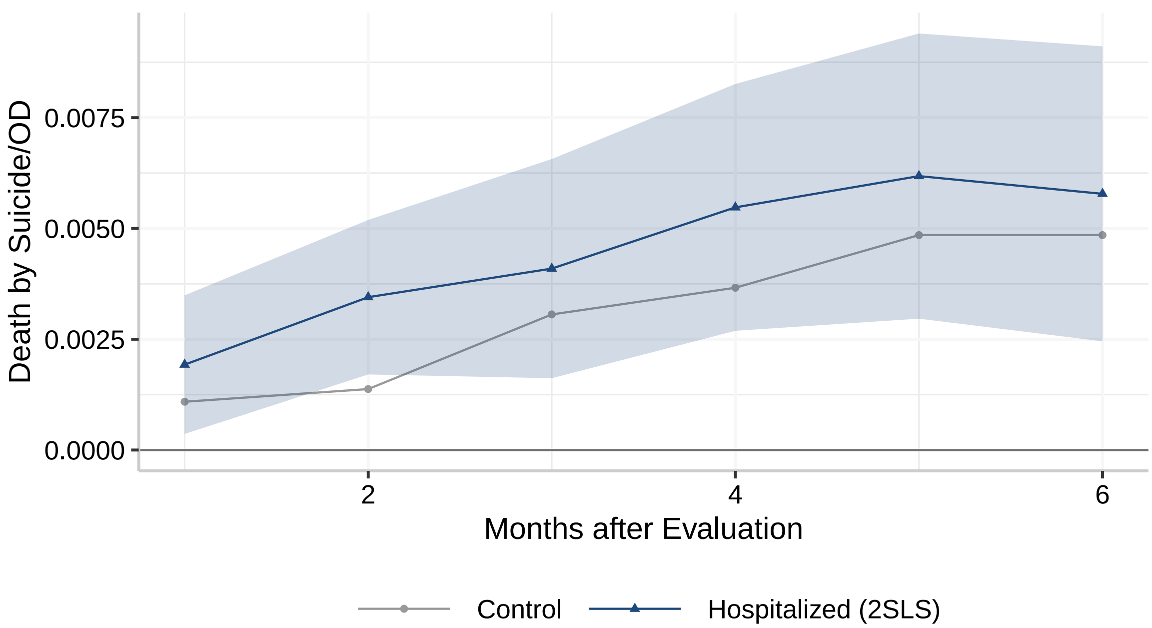
Note: This figure shows three heterogeneity analyses aimed at understanding if a decrease in substance use tolerance due to drug-abstinence while hospitalized drives our results. We plot the point estimates for the effects of involuntary hospitalization based on whether the patient has a substance use disorder (SUD) diagnosis or not, whether they received MAT treatment for opioid use during the evaluation and whether the hospitalization is at most 5-days or has been extended by the Court of Common Pleas (CCP). The dots show the control mean; the arrows the point estimate; the shaded region the confidence intervals. Each comparison is estimated from a single regression, meaning this figure shows 6 regressions. Outcomes are measured 6 months after the evaluation.

Figure A.9: Ordinary Least Squares

Panel A: OLS for Violent Crime Charge



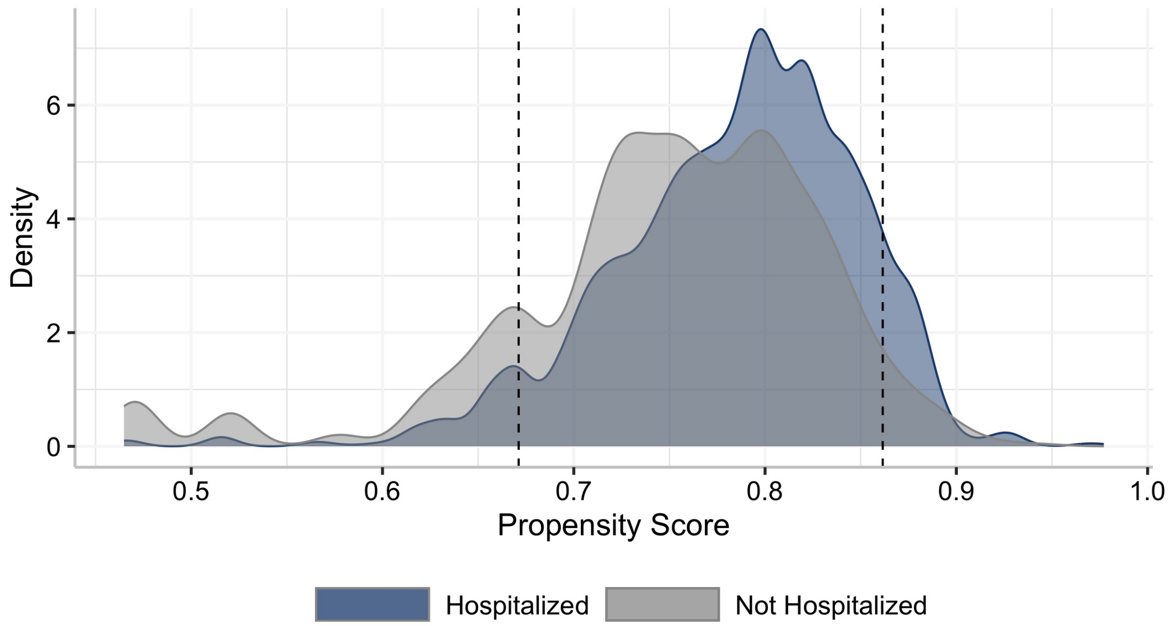
Panel B: OLS for Death by Suicide/Overdose



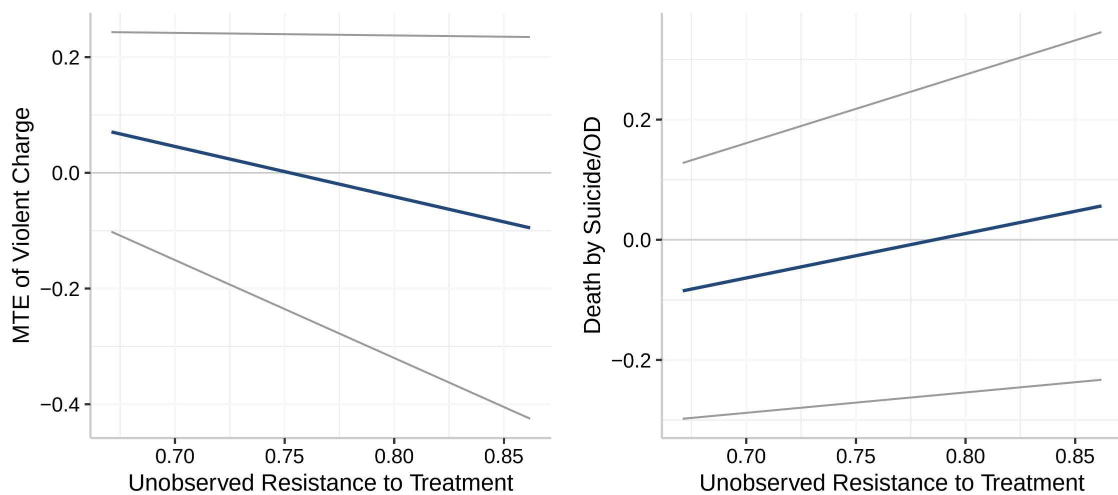
Note: This figure shows the ordinary least squares relationships between involuntary hospitalization and our two main outcomes. The grey line shows the control mean; the blue line shows the coefficient added to the control mean. Y-axis shows the share of people who are charged with a violent crime (Panel A) and the share who die as a result of suicide or overdose (Panel B). The ribbon shows the 95% confidence interval. Point estimates show cumulative effects over time.

Figure A.10: Marginal Treatment Effects

Panel A: Area of Common Support



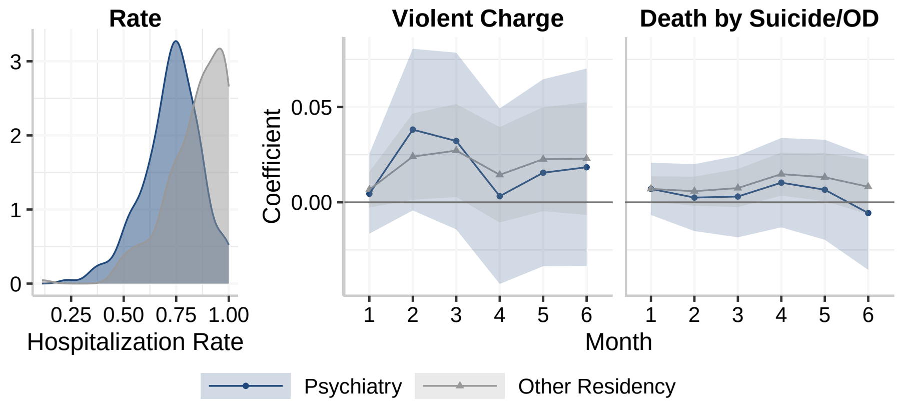
Panel B: Marginal Treatment Effects



Note: Panel A shows the area of common support between the treated and untreated individuals. Panel B shows marginal treatment effects for those who would be marginal among physicians who are most likely to hospitalize (on the left) toward those who are marginal among physicians who are least likely to hospitalize (on the right). Standard errors are based off of 250 bootstraps.

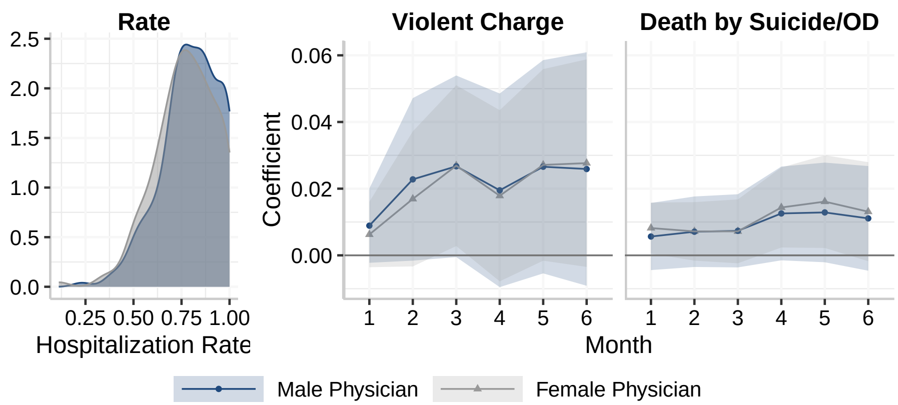
Figure A.11: Treatment Effects by Physician Trait

Panel A: Psychiatry Training



Panel B: Experience

Panel C: Gender



Note: Each panel shows three figures. The left-most figure shows the hospitalization rate for physicians with each trait. Next, the cumulative treatment effects for patients evaluated by each type of physician are shown over the six months following the evaluation for being charged with a violent crime and dying by suicide or overdose. Panel A shows those with psychiatry training as separate from those without psychiatry training. Panel B shows those who have fewer than 10 years since they graduated residency at the time of evaluation compared to those with 10 or more years. Panel C shows female versus male physicians.

B Appendix: Causal Effects of Evaluation: Evidence from Delegate Tendency to Certify

This appendix presents additional information, as well as tables and figures, testing assumptions associated with the instrument validity of using call delegates to assess the causal impacts of being evaluated for involuntary hospitalization.

We note that unlike in our main analysis, it is possible to have people who do not comply with the treatment in this context. A case may be certified, but then an individual may not be brought in for an evaluation. We find 71instances in which the case was certified but no examination took place, suggesting that a very small share of certified cases do not comply. Note that one may not be brought in for an evaluation without going through this process, so there are no always-takers.

Sample: We construct the sample to include all calls answered by delegates who have handled 100 or more cases in the past, so as to reduce noise around delegate tendency to certify cases. As with our main analysis, we limit our sample here based on age, focusing on individuals over age 18 and younger than age 65. While this is not necessary for the randomization of our cases (as it is in our main analysis), we do this to keep the sample under consideration here as similar to the sample in our main analysis as possible (Table B.1).

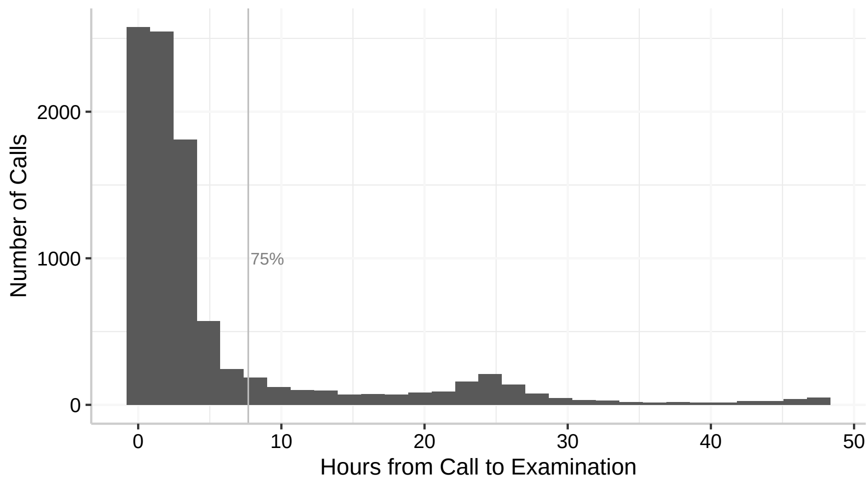
Table B.1: Delegate Sample Restrictions

Restriction	Evaluations	People	Delegates	Hospitals
All Calls	21,835	21,835	35	17
Limit to non-missing delegates	21,527	21,527	34	17
Limit to delegates with $\geq$ 100 cases	21,358	21,358	27	17
Limit to patients with known sex	21,358	21,358	27	17
Limit to patients over age 18	19,208	19,208	27	17
Limit to age $<$ 65	15,193	15,193	27	17
Limit to rotations with variation in cases/delegates	14,102	14,102	27	17

Note: This table shows the number of evaluations, individuals, delegates, and hospitals in our sample at each stage of the sample construction process.

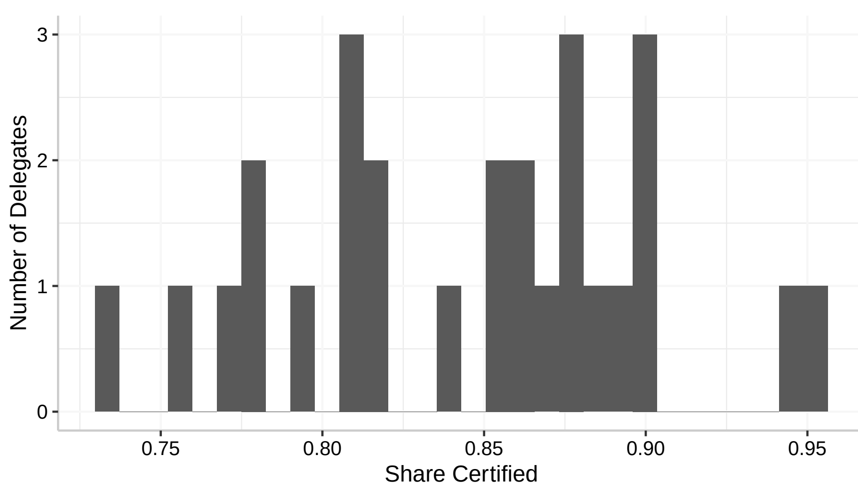
Table B.2 describes the sample as well as compliers. Compliers are less likely to be aged 18-35. They are more likely to be have a referral from a relative and less likely to have a referral from a school or other reporter (e.g., friends). They are less likely to have been reported as having attempted suicide.

Figure B.1: Time Between Delegate Certification and Hospital Evaluation



Note: This figure shows the hours between a call to a government delegate and the evaluation in the hospital. While technically, upon certification, an individual may be brought in for evaluation any time in the next 30 days, the vast majority of individuals are evaluated within hours of certification.

Figure B.2: Distribution of Tendency to Certify



Note: This figure shows the distribution in delegates' tendency to certify the cases that they hear. All delegates in this distribution have heard at least 100 cases.

Table B.2: Summary Statistics of Calls

	All	Certified	Not Certified	Compliers
(A) Demographics				
Female	0.441	0.440	0.448	0.404
Age 18-35	0.488	0.490	0.476	0.134***
White	0.664	0.669	0.640	0.759
Other Race	0.063	0.062	0.068	-0.01
Education				
Education Unknown	0.211	0.210	0.213	0.214
Some HS or Less	0.541	0.540	0.551	0.537
HS Diploma/GED	0.175	0.178	0.162	0.059**
Bachelor's or More	0.072	0.072	0.073	0.19***
(D) 302 Evaluation Traits				
Reporter				
Relative	0.717	0.702	0.797	0.364***
School	0.075	0.076	0.066	0.323***
Other Reporter	0.276	0.290	0.202	0.682***
Concern				
Attempted Suicide	0.480	0.509	0.323	0.242**
Unable to Care for Self	0.472	0.462	0.526	0.608
Mutilated Self	0.056	0.061	0.028	0.027
Danger to Others	0.500	0.501	0.494	0.442
Observations	14,112	11,869	2,243	[0.146]

Note: This table describes our study sample of individuals for whom there has been a call in which a delegate can decide whether they should be evaluated for involuntary hospitalization in Allegheny County from 2014 to 2023. The first column contains information on the whole sample. The middle two columns show the population based on whether their case was certified or not. The final column shows the observable traits for those who are compliers—those individuals for whom the evaluating delegate changes whether or not their case is certified and they are evaluated for involuntary hospitalization. The final row shows the number of evaluations in the whole sample, evaluations among those whose cases are certified and not certified, and in square brackets the share of compliers calculated. *p< 0.1; **p< 0.05; ***p< 0.01.

Table B.3: First Stage: Delegate Tendency and Case Certification

	Case Certified		
	(1)	(2)	(3)
Delegate Tendency	1.08***	0.98***	0.97***
	(0.07)	(0.08)	(0.09)
Date FEs		✓	✓
Demographic Controls			✓
History Controls			
Effective F	2.33	146.53	140.35
Observations	14,112	14,112	14,112

Note: This table shows first stage results for the delegate analysis, documenting the relationship between our residualized delegate tendency to certify a case instrument and the individual's likelihood of having their case certified for evaluation. The regressions are estimated in the our delegate sample, using an instrument constructed as described in Section 4.1, only applied to delegates. Each column adds successive control variables. Demographic controls include only those variables visible to delegates while assessing a case: the individual's sex, age, and the reporter's relationship to the individual in question. Standard errors are clustered at the delegate level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table B.4: Randomization Test

	Case Certified	Delegate Tendency
F-Statistic	47.1204	2.5038
p-value	0.0000	0.0573
Observations	14,112	14,112

Note: This table relates observable characteristics to an individual's likelihood of having their case certified (Column 1) and to the assigned delegate's tendency to certify a case on a involuntary hospitalization evaluation (Column 2). The F-statistic jointly tests the significance of the control variables and the p-value of the joint test is reported below. Standard errors are clustered at the delegate level.

Table B.5: Delegate Tendency and Downstream Events

	Delegate Tendency			
Physician Tendency to Hospitalize	0.007			
	(0.005)			
Hospitalized		−0.001		
		(0.001)		
CCP Hearing			0.0002	0.003
			(0.002)	(0.003)
CCP Extension			0.001	−0.004
			(0.002)	(0.003)
F-stat from joint test	0.666	2.5717	0.4069	1.2369
p-val from joint test	0.4145	0.1088	0.6657	0.2904
Observations	9,220	11,869	9,054	3,912

Note: This table considers the relationship between a delegate's tendency to certify a case and the downstream events among those who are evaluated. Conditional upon certification, the delegate's tendency is unrelated to the physician's tendency to hospitalize, whether the individual is hospitalized, whether the individual has a hearing with the Court of Common Pleas (CCP), or whether the magistrate extends their stay. The F-statistic and p-value from a test of joint significance is shown at the bottom. *p< 0.1; **p< 0.05; ***p< 0.01.

Table B.6: First Stage by Demographic Characteristic

	Certified			
	Female	White	Some HS	< 35 Years
Tendency	0.74***	1.09***	1.02***	0.91***
	(0.18)	(0.12)	(0.13)	(0.17)
Rotation Fixed Effects	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓
Observations	6,222	9,374	7,641	6,886

Note: This table presents the first stage relating a delegate's tendency to certify a case for evaluation to the demographic characteristics of those whose cases are brought before them. Standard errors are clustered at the delegate level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

Table B.7: First Stage by Reported Concern

	Case Certified			
	Unable to Care	Danger to Others	Suicide Att	Mutilated Self
Delegate Tendency	0.97***	1.14***	0.75***	1.23
	(0.19)	(0.14)	(0.11)	(1.24)
Rotation Fixed Effects	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓
Observations	6,659	7,054	6,767	788

Note: This table presents the first stage relating a delegate's tendency to certify involuntary hospitalization petitions to a particular certification, subsetting the sample by the reported concern. Standard errors are clustered at the delegate level and shown in parentheses. *p< 0.1; **p< 0.05; ***p< 0.01.

C Appendix: Heterogeneous Effects Based on Predicted Risk

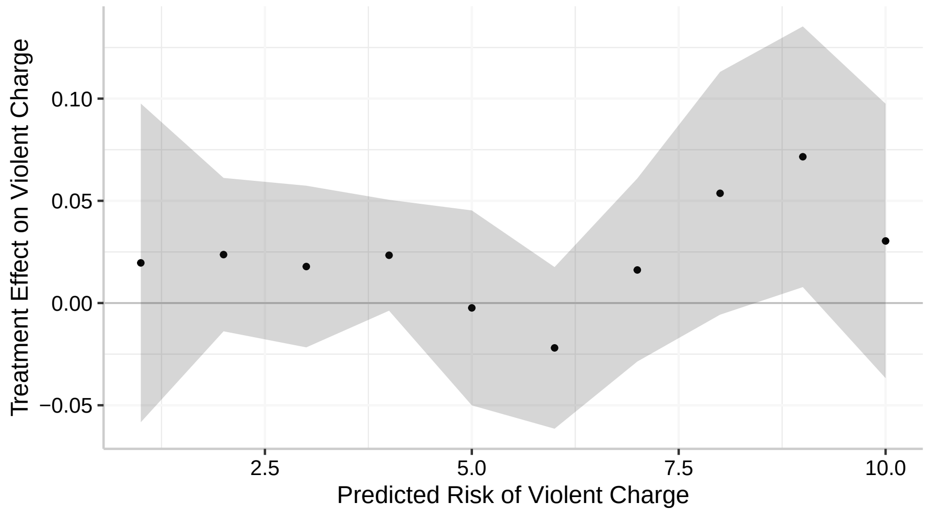
Another way to consider the estimates is to see whether hospitalization has differential effects on certain groups of people, specifically those who are most at risk of being charged with a violent crime or dying by suicide or overdose. This analysis simulates a thought experiment in which physicians' determinations of whether an individual poses a danger to themselves or others are supplemented with a prediction of risk based on available data.

We begin by using data available to us to predict an individual's likelihood of being charged with a violent crime, and separately the individual's likelihood of dying by suicide or overdose in the year following an examination. Note that we use data that are not available to the evaluating physician to construct these risk estimates. We generate these predictions based on a cross-validated generalized linear model, estimated on untreated individuals and predicted among all individuals who are evaluated.

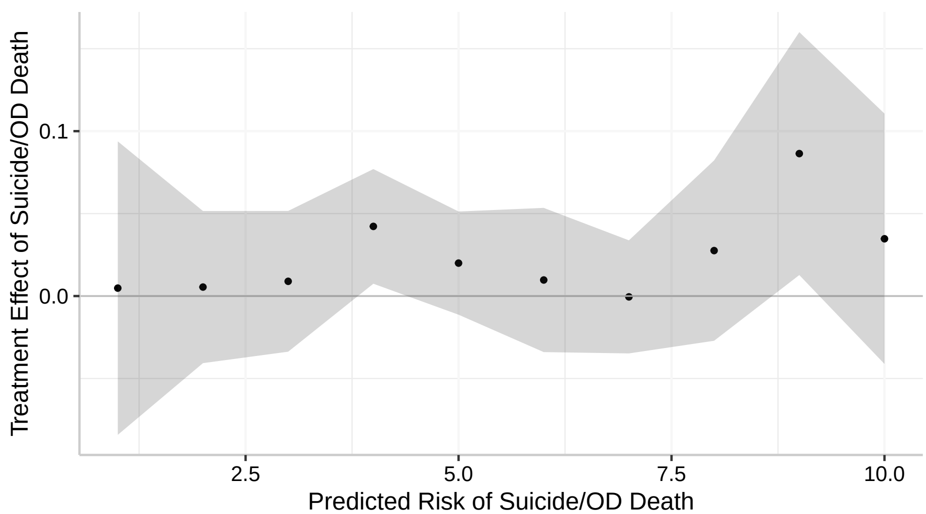
Our results are too noisy to conclude anything instructive, in line with the sparse outcomes we consider and the poor predictive power of the observable variables (See Figure C.1).

Figure C.1: Treatment Effects by Predicted Risk

Panel A: Violent Crime Charge Treatment Effects



Panel B: Death by Suicide/Overdose Treatment Effects



Note: Panel A shows the increase in the probability of a violent crime charge within 3 months of evaluation for each decile of predicted risk for having a violent crime charge. Panel B shows the increase in the probability of death by suicide or overdose within 3 months of evaluation for each decile of predicted risk for having a such a death.