Release, Detain, or Surveil?

The Effect of Electronic Monitoring on Defendant Outcomes

Roman Rivera*

Columbia University

April 4, 2023

Click here for most recent public version

Abstract

This paper studies the effect of pretrial electronic monitoring (EM) as an alternative to both pretrial release and pretrial detention (jail) in Cook County, Illinois. EM often involves a defendant wearing an electronic ankle bracelet that tracks their movement and aims to deter pretrial misconduct. Using the quasi-random assignment of bond court judges, I estimate the effect of EM versus release and EM versus detention on pretrial misconduct, case outcomes, and future recidivism. I develop a novel method for the semiparametric estimation of marginal treatment effects in ordered choice environments, with which I construct relevant treatment effects. Relative to release, EM increases new cases pretrial due to bond violations while reducing new cases for low-level crimes and failures to appear in court. Relative to detention, EM increases low-level pretrial misconduct but improves defendant case outcomes and reduces cost-weighted future recidivism. Finally, I bound EM's pretrial crime reduction effect. I find that EM is likely an adequate substitute for pretrial detention. However, it is not clear that EM prevents enough high-cost crime to justify its use relative to release, particularly for defendants who are more likely to be released.

*I am grateful to Sandra Black, Bentley MacLeod, and Simon Lee for guidance and advice. For feedback, I thank Nour Abdul-Razzak, Amani Abou Harb, Livia Alfonsi, Francisca Antman, Bocar Ba, Jason Baron, Pat Bayer, Tarikua Erda, Felipe Goncalves, Bhargav Gopal, Michelle Jiang, Emily Leslie, Annie McGowan, Steve Mello, Ismael Mourifié, Brendan O'Flaherty, José Luis Montiel Olea, Aurélie Ouss, Nayoung Rim, Bernard Salanié, Rajiv Sethi, and Ashley Swanson. For detailed explanation of the Cook County Court system, I thank Ali Ammoura. I thank the Invisible Institute and Chicago Data Collaborative for their contributions to this data set. I thank the Ford Foundation Fellowship Program and the National Academies of Sciences, Engineering, and Medicine for their generous financial support. This work was sponsored by the Program for Economic Research. All errors are my own. Email: r.g.rivera@columbia.edu

1 Introduction

The social and economic costs of pretrial detention are massive: around half a million individuals are detained while presumed innocent on any given day in the United States, resulting in a direct cost to local governments of around $14 billion annually.[1] However, releasing defendants pretrial risks increased crime. Over the last two decades, an alternative to both pretrial detention and release, electronic monitoring (EM) — technology that surveils and limits the movement of defendants using an ankle bracelet — has expanded across the United States, and its adoption has accelerated due to the COVID-19 pandemic.[2] Proponents of EM believe it promises better outcomes for defendants and lower costs to taxpayers than pretrial detention (APPA (2020)) while avoiding the potential increase in crime if defendants were released from jail without EM. Critics of EM argue that its expansion leads to unnecessary surveillance and additional criminal charges against defendants for non-criminal violations while increasing their contact with the criminal justice system (Weisburd et al. (2021)).[3]

Two questions vital to understanding the value of EM remain unanswered in the literature. First, relative to detention, is the cost of pretrial crime under EM outweighed by EM's lower direct costs to taxpayers and better outcomes for defendants? Second, relative to release, does EM prevent enough pretrial misconduct to outweigh its adverse effects on defendants, such as its increase in rearrests due to EM violations? Both of these questions hinge upon how sensitive defendant decision-making is to surveillance and how many more rearrests occur on EM due to non-criminal violations.

The primary reason for the lack of evidence on the effect of pretrial EM relative to release and detention is a lack of data. Furthermore, because EM is a substitute for both release and detention — two very different alternatives — understanding its effects requires

[1]See Sawyer and Wagner (2022) and Wagner and Rabuy (2017).

[2]In 2015, more than 125,000 people in the US were on EM, more than double the number in 2005 (Pew (2016)). Weisburd et al. (2021) reviews EM's popularity across the US. Guevara (2014), Hager (2020), Barajas (2020), Federation (2020), Weisburd and Virani (2022), and Virani (2022) discuss the expansion of the use of EM in specific US municipalities.

[3]For additional discussion of the debate around EM and criticisms, see Kilgore (2013), DeMichele (2014), PRI (2016), Alexander (2018), Benjamin (2019), Lockhart (2019), and Weisburd (2022).

differentiating between both margins. In this paper, I use administrative court and jail data from Cook County, IL, one of the largest bond courts in the US and an early mass-adopter of EM. I document the effects of pretrial EM relative to both release and detention on pretrial misconduct, case outcomes, and future recidivism. I leverage the quasi-random assignment of bond court judges to recover the effects of EM versus release and EM versus detention using two-stage least squares (2SLS). While 2SLS can recover the average treatment effects on compliers, the average effect if EM were expanded to a larger population is more relevant for policy. I explore heterogeneity in treatment effects on both margins using a marginal treatment effects (MTE) framework and a novel semiparametric estimation method. Then, I construct treatment effects that are relevant for policy, such as the average effect of expanding EM to the defendants who were released or to those who were detained.

Relative to detention, I find that EM allows for increased pretrial misconduct but overall improves defendant outcomes. Defendants placed on EM are more likely to fail to appear in court and to have a new case opened against them pretrial; low-level criminal charges and charges for bond violations drive the increase in new cases. However, I find suggestive evidence that EM is less coercive and criminogenic than detention, as it lowers the likelihood of incarceration and either decreases or has a null effect on cost-weighted post-trial recidivism.[4] EM does not increase pretrial crime relative to detention when new charges are weighted by the dollar cost of crime, making it a reasonable alternative to detention for most defendants. However, selection patterns indicate that the benefits of EM are smaller for defendants who are more likely to be detained.

For the effect of EM relative to release, EM reduces failures to appear in court, has no effect on new cases pretrial with serious charges, and weakly decreases new cases pretrial with low-level charges for most defendants. The reduction in pretrial new cases with crime-related charges may understate EM's crime-reducing effect because EM increases the probability of detection through surveillance. To account for the change in detection rates, I compute

[4] I follow Miller et al. (2021) in my construction of a cost-weighted crime index.

the implied dollar cost of crime prevented by EM relative to release. Under very generous assumptions, the amount is less than \$10,000 for the average defendant and is smaller for released defendants, though I cannot reject a small or null effect.[5] Furthermore, I bound defendants' elasticities of pretrial crime with respect to the probability of detention. I find evidence consistent with defendants being elastic for low-level crimes but relatively inelastic for serious crimes.

EM imposes costs on defendants relative to release that may outweigh its potential benefits. EM increases the likelihood of a new case pretrial with charges related to bond violations and "escapes" because of the increased restrictions of movement while on EM. Naturally, these criminal charges for violations are socially costly because their punishment (e.g., re-incarceration and conviction) outweighs the harm of these activities. This highlights the trade-off with surveillance systems used to prevent undesirable behavior: additional conditions placed on individuals to enforce compliance with the system can lead to negative effects (e.g., charges for bond violation) that may undermine the benefits of reduced misconduct (e.g., crimes).

Furthermore, while the effects of EM relative to release on case outcomes and post-trial recidivism are mixed, the main results suggest that EM leads to weakly worse case outcomes and more socially costly future recidivism for most defendants. In general, selection patterns indicate EM's costs and benefits relative to release are both smaller for released defendants than for the average defendant. In contrast with EM's sizable benefits over detention, it is not clear that EM's benefits over release outweigh its costs.

This paper makes the following contributions to our understanding of the economics of pretrial detention and surveillance. First, existing work on EM in economics studies non-US contexts and has focused on the effect of EM relative to detention, pre- or post-trial, with recidivism being the main outcome of interest (Di Tella and Schargrodsky (2013), Henneguelle,

[5] Specifically, assuming EM detects 100% of crimes, all guilty charges are the crimes actually committed, and the crime-cost weighted probability of detection on release is 50%.

Monnery, and Kensey (2016), Williams and Weatherburn (2022)).[6] This paper advances our understanding of how EM influences defendant outcomes by applying a new methodology to recover heterogeneous treatment effects of EM relative to both release and detention on pretrial misconduct, case outcomes, and recidivism in a large US municipality. Beyond the location and multitude of outcomes, these results better inform policy by allowing us to explore the effects of EM on defendants who are more or less likely to be released or detained — i.e., the effect of expanding EM on either margin. Relatedly, this paper builds on the pretrial detention and judge-design literatures by incorporating multiple pretrial treatments, while prior work generally focuses on a binary set of pretrial treatments (e.g., release versus detain), as well as the literatures on probation and alternatives to detention.[7]

This paper builds on the large literature on agency theory, in which principals use costly monitoring mechanisms to deter agents from 'shirking' and to improve productivity (Alchian and Demsetz (1972), Holmstrom and Milgrom (1991)).[8] I find that EM imposes a cost by increasing new cases pretrial due to non-criminal violations (relative to both release and detention). This cost is a byproduct of EM's inability to perfectly detect criminal activity, thereby punishing a wider set of (non-criminal) activities. This highlights the importance of considering not only the direct costs of the monitoring technology (in this case, maintaining the EM system) but also the cost of punishing violations of the monitoring system itself.[9]

[6]Additional work on EM includes Marie (2008), Ouss (2013), Andersen and Andersen (2014), and Grenet, Grönqvist, and Niknami (2022). See Belur et al. (2020) for a review of the interdisciplinary literature on EM.

[7]See Gupta, Hansman, and Frenchman (2016), Leslie and Pope (2017), Dobbie, Goldin, and Yang (2018), Stevenson (2018), Arnold, Dobbie, and Yang (2018), Agan, Doleac, and Harvey (2021), and Arnold, Dobbie, and Hull (2022) for 'judge' designs in pretrial context. Exploiting 'judge' random assignment is a common identification strategy: Kling (2006), Doyle Jr. (2008), Aizer and Doyle (2015), Mueller-Smith (2015), Dobbie and Song (2015), Frandsen, Lefgren, and Leslie (2019), Bhuller et al. (2020), Norris, Pecenco, and Weaver (2021), Arteaga (2021), Jordan, Karger, and Neal (2022), Goncalves and Mello (2022), and Gross and Baron (2022). See Dobbie and Yang (2021) for a discussion of the US pretrial system. See Myers (1981) for an early analysis of the economics of pretrial detention. For work on the effect of probation, parole, and alternatives to detention, see Hawken and Kleiman (2009), Kilmer et al. (2013), Kuziemko (2013), Rose (2021), LaForest (2021), and Arbour and Marchand (2022).

[8]See also Becker and Stigler (1974), Mirrlees (1976), Holmström (1979), Grossman and Hart (1983), Shapiro and Stiglitz (1984), and Prendergast (1999).

[9]This is consistent with criticisms of EM and surveillance in the criminal justice system more broadly (Weisburd (2022)) and highlights the importance of designing policy to minimize crime rather than maximize arrests (Lum and Nagin (2017)). This also connects to the literature on elasticities of crime with respect to sentencing (Kleiman (2009), Durlauf and Nagin (2011)) and detection rates (Persico (2002), Manski (2005),

Consistent with models of agency theory and specific deterrence (Becker (1968)), I also find evidence consistent with EM reducing pretrial crime by increasing the probability of detection.[10] Furthermore, this paper connects agency theory to the literature on modern surveillance technology by studying the rise of individualized surveillance used to deter violations, such as body-worn cameras on police used to deter misconduct (Lum et al. (2020)) or remote-work monitoring technologies (Jensen et al. (2020)), whereas much prior work focuses on the economics of mass surveillance technologies (Tirole (2021)).[11]

Methodologically, this paper contributes to the literature on marginal treatment effects and identification with multiple treatments (Kirkeboen, Leuven, and Mogstad (2016), Kline and Walters (2016), Mountjoy (2022)).[12] Much like Rose and Shem-Tov (2021), I build on Heckman, Urzua, and Vytlacil (2006) by showing that identification of marginal treatment response (MTR) functions in ordered treatment environments can be achieved relying solely on variation in the probability of adjacent treatments.[13] I develop a straightforward method for the semiparametric estimation of MTRs in ordered treatment environments, which can be applied in a range of ordered treatment environments and allows for non-monotonic MTEs in contrast with the existing fully parametric method for estimating MTEs with ordered treatments (Heckman, Urzua, and Vytlacil (2006), Cornelissen et al. (2018)).

This article proceeds as follows. Section 2.2 discusses the institutional background, data, and the potential costs and benefits of each treatment. Section 3 presents the empirical strategy and results for the 2SLS analysis. Section 4 presents the empirical strategy and results for the MTE analysis. Section 5 discusses policy implications, and Section 6 presents

Dominitz and Knowles (2006), Bjerk (2007), Feigenberg and Miller (2022)).

[10]The probability of detection plays a major role in crime reduction through deterrence. See Cook (1979), Ehrlich (1996), Donohue (2007), Nagin (2013), Chalfin and McCrary (2017), and Mungan (2017).

[11]See also Beraja, Yang, and Yuchtman (2020), Beraja et al. (2021), and Acemoglu (2021) on AI and mass surveillance. Barbaro et al. (2022) discusses the post-COVID rise of surveillance and remote-work.

[12]See also Björklund and Moffitt (1987), Heckman and Vytlacil (1999), Dahl (2002), Heckman and Vytlacil (2005), Heckman, Urzua, and Vytlacil (2008), Moffitt (2008), Carneiro and Lee (2009), Brinch, Mogstad, and Wiswall (2017), Lee and Salanié (2018), Mogstad and Torgovitsky (2018), and Bhuller and Sigstad (2022). See also Cornelissen et al. (2016) and Andresen (2018).

[13]Rose and Shem-Tov (2021) develops a similar method, but their focus is on bounding MTRs (building on Mogstad, Santos, and Torgovitsky (2018)) rather than point-identification and estimation of MTRs over common support using continuous instruments.

robustness checks. Section 7 concludes.

2 Background and Data

2.1 Cook County Bond Court, Bond Types, and Treatments

Bond Court Following an arrest in Cook County, a defendant is taken into custody and "booked" based on their arrest charges, generally at the Cook County Jail (CCJ). Then they are arraigned at bond court, usually within 1 day of the arrest. For the entire "pretrial" period, the defendant is presumed innocent.

This paper focuses on the central bond court in Cook County, known as "Branch 1". Branch 1 handles almost all felony cases in Cook County and operates every day. On weekdays and non-holidays, non-felony cases (e.g., misdemeanor, traffic, and municipal code violations that require the setting of bail) are generally handled in the bond courts determined by the location of arrest (Branch 1 is the bond court for Chicago arrests) or specifically designated courts — for example, murder and violent sex offenses are handled in their own courts. On holidays and weekends, however, all such cases are handled by Branch 1. At bond court, the sitting bond court judge determines the bond conditions for a defendant, namely bond type and amount. [14]

During the period of this study, from July 2013 through 2015, a single bond court judge handled the cases that passed through Branch 1 on any given day. The judges have an irregular working schedule (Tardy et al. (2014)), which depends on their off days, vacations, and work-day preferences — Figure 1 displays a sample of the calendar from the data. There are relatively few active bond court judges within a given month (≈ 4), and only 7 were active during the period of study.[15] Defendants do not have discretion over their assigned

[14] See here for a schedule of the bond courts in Cook County. Defendants are generally processed at CCJ after they have their bond hearing at the court.

[15] Active is defined as having at least 500 cases within a year, excluding days where a judge saw fewer than 40 cases. In the full data, this filter removes 3% of observations but 90% of unique judges. Between 2010 and 2016, two active judges are recorded working in bond court on the same date on less than 10 days out of over

judge, and judges cannot choose the cases they see on a given day. Importantly, because Branch 1 is always active (including holidays and weekends) and the schedule of each bond court judge is sporadic, there is no scheduling relationship between bond court judges and prosecutors, public defenders, or trial judges.

Bond Types In the period of study, there were three main bond types used by Branch 1 judges: D-bonds, I-bonds, and IEM bonds. A D-bond is the most common bond (55% of sample bonds) and conforms to the popular understanding of bonds: it requires a defendant to post 10% of the bond amount in order to be released from jail. This 10% amount can range widely, from below \$50 to over \$200,000. The defendant can pay the amount at any point during their pretrial period and be released from jail, or they are detained in jail until the pay or their case concludes.

I-bonds (15%), or "release on recognizance" bonds, do not require the defendant to post any money in order to be released from jail. However, as with all bond types discussed, the defendant is liable to pay the full bond amount if they violate their bond conditions (e.g., they fail to appear in court).[16]

IEM bonds (29%) allow defendants to leave jail at no cost (like an I-bond), but they are placed on electronic monitoring (EM). The defendant can be released off of EM if they pay 10% of the bond amount (like a D-bond). Prior to IEM bonds, EM could be coupled with a D-bond (D-EM) such that defendants were in jail until they paid 10% of the bond amount and were released onto EM. See Appendix A.1 for more details on bond types and background on IEM's introduction.[17]

Treatment Definitions I map these bond types and release statuses into three ordered pretrial treatments ($S \in \{1, 2, 3\}$). At the lowest level, defendants can be "released" ($S = 1$)

2,500, about 0.3% of observations.

[16]While defendants are liable for the bond amount, most defendants cannot pay the large sums, and cash bail has been shown to be ineffective at ensuring court appearances (Ouss and Stevenson (2022)), and the threat of court fines has been shown to be ineffective (Albright (2021)) as the collection of such fines is rare (Pager et al. (2021)).

[17]Judges also can but rarely do deny bond (1.4%) if they determine the defendant is a flight risk or a threat to the public. There are two additional bond types in the data A and C bonds, but they account for a minute share of bonds.

if they fully exit the sheriff's office's custody through an I-bond, or because they were given an IEM or D-bond but paid the bond amount and were released within 7 days. The middle level is being on "EM" ($S = 2$), meaning the defendant was assigned an IEM bond but was not recorded as being released from it (e.g., paid the 10% amount) within 7 days. At the highest level, the defendant can be "detained" ($S = 3$) if they are given a D-bond and are not released from jail within 7 days. For the main results, the cutoff is 7 days, which is the 39th percentile for the duration in jail/EM, but 75% of releases happen within 3 days for D- and IEM bonds. This cutoff is largely arbitrary, and I construct robustness checks using alternative cutoffs. In particular, using the common 3-day cutoff yields similar results.[18]

Pretrial Period and Case Outcomes After the bond hearing, all defendants have the opportunity to plead guilty or proceed with the case (which can involve a future plea).[19] Prior to the case ending, through a trial, dismissal, plea, or being dropped, defendants can be rearrested or charged with new crimes and fail to appear in court.

Lengthy cases are common, and trials are rare, with 93% of cases with a guilty outcome involving a guilty plea. If guilty, the defendant can be sentenced to pay a fine, time served, or incarceration in prison (Illinois Department of Corrections) or in Cook County Jail. Following the case outcome and subsequent incarceration period, defendants can be rearrested or have a new case against them (a new case post-trial).

2.2 Background on Electronic Monitoring

2.2.1 EM Conditions and Details

While electronic monitoring can refer to a variety of technologies, all operate as electronic individualized surveillance systems with a similar mechanism and purpose. In this paper, I

[18]Notably, 30% of defendants classified as 'detained' are eventually released before their case ends, though this number is comparable to that of other work, such as Dobbie, Goldin, and Yang (2018).

[19]The next steps depend on if the case is a misdemeanor or felony case. Felony cases proceed to a hearing which determines if the case can proceed with felony charges (usually a preliminary hearing, grand jury, or an information) and be transferred to the criminal division; otherwise, it is dropped or proceeds with misdemeanor charges. The full evolution of a case involves many events, and a flowchart for felony and misdemeanor cases can be seen in Figures B.1 and B.2, respectively.

refer to the system operated by the Cook County Sheriff's Office between 2013 and 2015, which generally involved defendants wearing a radio-frequency ankle bracelet at all times. EM was used to ensure defendants did not leave their homes except at preapproved dates and times (e.g., a specific work or education schedule) or for a small set of "one-time movement" conditions.[20] This system is referred to as EM coupled with house arrest, and it is common across the United States (Weisburd et al. (2021)). The ankle bracelet communicates with a monitoring unit installed in the defendant's home, which informs the sheriff' office of out-of-bounds movement and tampering.[21]

Defendants on EM agree to warrantless searches of their residence while on EM, and the possession of firearms, drugs, or contraband is a violation of EM bond conditions (Rogers (2022)). As stated by the sheriff's office's EM information sheet, violations or noncompliance with the terms carry the "risk of criminal prosecution and re-incarceration," and damaged or missing equipment can be charged as felony theft (Office (2020)). Similar to other bail conditions of release (e.g., a condition of all bond releases is that the defendant appears in court), violations of the EM bond requirements can lead to re-incarceration in jail as well as additional charges.

2.2.2 EM as Middle Treatment

IEM bonds were introduced in June 2013 in the wake of a jail-overcrowding crisis and conflict between the court and the sheriff's office, and they quickly became popular among judges, comprising 28.7% of bonds in the following two and a half years, leading to a decline in other

[20]See this information sheet for the rules and information sheet for the EM program in 2020. While the domain of EM monitoring is generally one's home, exceptions can be made in advance for work, school, or other reasons (Federation (2020)), it but requires 2 days prior approval. The Sheriff's EM program is by far the most common form of EM (Green (2016), Federation (2020)). See Appendix A.1 for a discussion of other EM programs. Time spent on EM in the Sheriff's program counts as days served in jail and thus can reduce one's time required to be served if found guilty (Federation (2020)). In many jurisdictions, EM can require defendants to pay a fee, but this was not common for pretrial EM during the period of study based on available sources. See Dizikes and Lightly (2015) for images of the 2013-2015 system.

[21]While I do not observe whether or when the defendant has the EM system set up in their home, I do observe disposition codes in a defendant's case which provide more specific information on their EM status, such as explicitly stating a defendant was admitted into the Sheriff's EM program. I test the robustness of my results to modifying the treatment definitions using these disposition codes in Section 6.

bond types. While IEM (EM) can be seen as a tool to reduce overcrowding in Cook County Jail, in practice, EM was used as a middle option between release and detention and applied to defendants who would otherwise have been released as well as those who would otherwise have been detained. Though no official guidance on EM as the 'middle' option existed at the time, the 2016 "Decision Making Framework" in Cook County explicitly places EM between release and detention (jail) (CGL and Appleseed (2022)).

Figure 2 shows that with the introduction of the IEM bond, mid-level D-bonds were replaced by IEM bonds. To demonstrate this more directly, I regress the probability of detention using defendant observables from 2009 to 2012; using these coefficients, I then construct an index for defendant severity (propensity to be detained) as the predicted probability of detention in the post-EM period for released, detained, and EM defendants. Figure 3 displays the distribution of the index for each treatment and shows that defendants assigned to EM are between those who were released and those who were detained based on their predicted likelihood of detention.

2.3 Data and Summary Statistics

2.3.1 Data

The main data for this study comes from the Circuit Court of Cook County and Cook County Jail. The court data contains information on defendants, cases, charge counts, and outcomes for almost all Cook County court cases between 1984 and 2019. This allows me to follow cases from inception to bond court, individual motions and case events, through to final case outcomes (e.g., a defendant demanding a trial, a guilty plea on a specific charge but not others, and a sentence to time served). I connect defendants across cases, allowing me to observe extensive criminal and case histories across tens of thousands of individuals over more than three decades, including past and future cases, arrests, guilty verdicts, charges, sentencing, and pretrial misconduct. I link the court data to data from the Cook County Jail, maintained by the Cook County Sheriff's Office. This jail data, spanning from 2000 to

2017, contains information on individuals' detention spells, intake, and release. I also link this data to Chicago Police Department arrest data.

While the collective data allows me to follow tens of thousands of defendants in Cook County through the criminal justice system from arrest in Chicago to sentencing over more than 15 years, this study focuses on two and a half years (July 2013 to December 2015) in which EM was one of the most common pretrial treatments for defendants. I focus on adult cases that went through Branch 1 with known bond types. The data is summarized at the booking and defendant level — which I will refer to as a "case" for simplicity.[22]

In order to quantify the intensive margin of defendants' recidivism (new charges after bond court) and avoid common issues with binary measures of recidivism (Rosenfeld and Grigg (2022)), I quantify the social costs of different crimes by supplementing my main data with the crime cost information from Miller et al. (2021).[23]

In the final data, I focus on completed cases within the Cook County court system that do not involve murder or felony sexual assault charges (in the current case) and have categorizable pretrial treatments. In the final data there are 84,332 defendant-bookings (cases) between July of 2013 and December of 2015 which passed through Branch 1, 51,348 of which comprise the main felony sample.[24] See Appendix A.2 for more details on the data construction and filtering.

[22]Linking these cases to jail spells leaves some cases with missing jail information. In the main analysis, I code missing releases for I-bonds as immediate releases, missing EM bonds as EM, and drop missing D-bonds. I test the robustness of my results to these codings in Section 6.

[23]I use Table 5 from Miller et al. (2021), which contains the total tangible and quality of life costs associated with different crimes, to guide the construction crime costs for charges against defendants in the data as well as imputing costs for some unlisted crimes. A similar method is used in Mello (2019) to compute cost-weighted crimes per capita. However, because the cost of a single murder is so high (almost $8 million dollars, which is equivalent to about 200 police-reported robberies), I also conduct analyses using a "low" murder cost, making it equivalent to a police-reported rape (about $400,000) to determine if outlier defendants accused of murder in future cases are driving all of the results, similar to an exercise in Heckman et al. (2010).

[24]Though misdemeanor arrests are an important part of the criminal justice system (Kohler-Hausmann (2013), Mayson and Stevenson (2020)), EM was primarily used on felony cases. The main felony sample also excludes D-EM cases.

2.3.2 Summary Statistics

Table 1 displays summary statistics for defendant and case characteristics, and Table 2 displays summary statistics for a subset of outcomes by pretrial treatment, with release being the lowest treatment and detention being the highest.[25] As shown in Table 1, bond amounts, days in jail or on EM, and case durations are increasing in the treatment level. Black and male defendants (the majority of defendants) are more likely to receive a higher treatment level, and treatment level is increasing in the severity of charges against the defendant and in the severity of the defendant's case history increases. For example, released defendants are least likely to have a felony charge (45%) and have the fewest previous felony cases (1.5). In the middle, EM defendants are the most likely to have a felony charge (86%), but this is driven by drug felonies (e.g., 57% are charged with possession), and they have better case histories than detained defendants. Detained defendants are most likely to have violent felony charges (11%) and have the worst case histories.

For outcomes, released defendants are two and four times as likely to fail to appear (FTA) in court (14%) relative to EM and Detain defendants. The likelihood of a guilty felony charge increases with treatment level, from 22% for released defendants to 59% for detained defendants. Interestingly, EM defendants have the highest rates of having a new case pretrial (18%) relative to lower likelihoods for released and detained defendants (15% and 11%), likely due to detained defendants being released eventually (after 7 days) and EM defendants having new charges for bond violations. Post-trial, total new cases over four years are similar across treatments at around 1.6, though only about one-third of these are felony cases, and less than 5% are for violent felonies. The similarity across treatments is likely due to post-trial incapacitation in addition to selection and potential treatment effects.

[25]Table B.1 and Table B.2 display the same for the felony sample. Table B.3 displays the means of all outcomes by treatment for the felony sample.

2.4 Framing Costs and Benefits

To understand to better frame the results, we can categorize the costs of each pretrial treatment (release, EM, detention) into four main categories: direct costs, the dollar cost of operating each treatment; crime costs, the social cost of crime under each treatment; punishment costs, the cost of enforcing each treatment; and indirect costs, all other spillovers onto the defendant or society. This paper provides estimates for changes in crime and punishment costs using new cases pretrial and indirect costs using new cases post-trial and sentencing outcomes, while direct cost estimates are taken from prior work. Furthermore, treatments differ in which mechanisms they rely on to reduce misconduct: the probability of detection (p) to deter misconduct or the level of incapacitation (d).

Direct Costs For direct costs, the comparisons are relatively straightforward and not studied directly in this paper. Jails cost about \$100-\$200 per detainee per day to maintain.[26] While pretrial EM is significantly cheaper than detention, monitoring defendants is still costly. The total cost of operation of EM is about \$15 per defendant-day in Cook County in 2021 (CGL and Appleseed (2022)), while release is almost costless and less expensive than EM. These estimates are average costs. Marginal costs, which may be more appropriate, are likely around 20% and 60% of average costs (Wilson and Lemoine (2022)).[27] As the average defendant is in jail or on EM for over 100 days, this equates to marginal costs of jail between \$3,000 and \$9,000 for the average defendant.

Crime Costs Pretrial detention should reduce pretrial crime costs by incapacitating defendants in jail. On the other hand, EM operates through deterrence, with defendants' crimes or violations being more likely to be detected while on EM relative to release ($p^R < p^{EM}$), and through partial-incapacitation through house arrest conditions ($d^R < d^{EM} < d^D$). We can estimate changes in crime cost as the effect of EM versus release and detention on

[26]For example, in 2011, CCJ cost about \$90 per defendant-day, while in 2021 it cost \$223 per defendant-day (Institute (2022)).

[27]Marginal costs are more difficult to calculate. Wilson and Lemoine (2022) find that short-run and long-run marginal costs of incarceration are often around 20% and 60% of average costs. Thus, per defendant-day in jail, the average cost is around \$150, while the short-run and long-run marginal costs are around \$30 and \$90.

new cases pretrial; however, we must account for the fact that new cases pretrial depend on the probability of detection, which is higher for EM compared with release.

Punishment Costs Punishment costs result from defendants being punished, e.g., spending time in jail or being prosecuted or fined for pretrial misconduct. Naturally, detention involves high punishment costs as defendants effectively serve sentences while awaiting their case conclusion. And EM has lower punishment costs as house arrest is a less intense punishment than jail. However, because a defendant can be rearrested and placed in jail, release can lead to higher punishment costs if pretrial misconduct is detected and penalized (e.g., pretrial rearrests). By comparison, whether EM reduces punishment costs depends on whether being placed on EM increases or decreases the likelihood of receiving a new case pretrial. While EM aims to reduce pretrial crime through deterrence and partial incapacitation, it may result in higher punishment costs through two mechanisms.

First, because EM does not allow for perfect detection of defendants' misconduct, it applies sanctions to complementary but non-criminal activity (e.g., leaving one's home without permission), largely consistent with principal-agent models of multitasking (Holmstrom and Milgrom (1991)). As such, while on EM, non-criminal activity can lead to re-incarceration, a new case, and subsequent punishments, inducing inefficiency and higher punishment costs. Any increase in new cases from violations can be seen as a cost of monitoring the agent (defendant) beyond the direct costs. If the rules of EM are too stringent, unforgiving, or unreasonable such that defendants are punished for non-criminal activity (Weisburd et al. (2021)), then EM's punishment cost may be large.[28]

Second, if defendants are relatively inelastic (elasticity of crime, c, with respect to the probability of detection $-1 < \epsilon_p^c < 0$), and the incapacitation effect is relatively small, then defendants who are placed on EM will commit fewer crimes, but EM will still lead to increased pretrial rearrests. As a result, EM will increase punishment costs through

[28]Notably, though most new cases pretrial under EM (77.76%) do not involve a violation charge, 55.17% of new cases with a violation charge include no other charges for low-level or serious crimes.

heightened detection, partially offsetting gains through reduced crime.[29] Thus, the change in punishment cost depends on the effect of EM versus release on new cases pretrial.

Indirect Costs We can measure some of the indirect costs associated with each treatment as the effect of EM versus release and detention on future recidivism (rearrests post-trial) and case outcomes (e.g., guilty findings). Beyond individual suffering, detention has high indirect costs by damaging employment, future income, and defendants' bargaining power leading to worse case outcomes.[30] Pretrial EM may be less coercive than detention and reduce long-run recidivism while causing less individual suffering. Relative to release, however, EM may be criminogenic (increasing future crime costs) or increase indirect costs by hurting defendants' bargaining power. Beyond this, EM may damage social ties, economic opportunities, and health outcomes and increase housing insecurity, as found in interviews with EM participants (CGL and Appleseed (2022)).[31]

3 Effect of EM vs. Release and Detention

3.1 Empirical Strategy

To identify the effect of EM versus release and EM versus detention, I instrument for the two endogenous treatments ($Release_{ic}$ and $Detain_{ic}$) by exploiting the quasi-random assignment of judges to a defendant-case (i, c). I interact bond court judges with defendant characteristics to instrument for pretrial treatment assignment because there are relatively few judges in the data (7) who see many cases (12,047 on average). This allows judges to have heterogeneous preferences over defendant observables (Mueller-Smith (2015), Leslie and

[29] Since the elasticity of x with respect to y is $\epsilon^x_y = \frac{y}{x}\frac{dx}{dy}$, and arrests (detected crime) $a = c \times p$, then $\epsilon^a_p = \frac{p}{a}\frac{da}{dp} = \frac{p}{c \times p}(c + p\frac{dc}{dp}) = 1 + \frac{p}{c}\frac{dc}{dp} = 1 + \epsilon^c_p$. So, $\epsilon^a_p < 0 \iff \epsilon^c_p < -1$. Furthermore, if defendants are elastic $\epsilon^c_p < -1$, then the punishment cost will decline under EM relative to release.

[30] See Leslie and Pope (2017), Stevenson (2018), and Dobbie, Goldin, and Yang (2018). Bargaining power is a significant factor in case outcomes (Silveira (2017)). Spending time in jail may reduce future economic opportunity and increase criminal capital (Bayer, Hjalmarsson, and Pozen (2009), Stevenson (2017)).

[31] Alternatively, EM may reduce future crime through an individual deterrent effect (increasing defendants' expectations of punishment) — though the research on EM versus detention (which should have a similar effect) has not found evidence consistent with such a mechanism (Williams and Weatherburn (2022)).

Pope (2017), Stevenson (2018)). I construct the judge instruments by estimating:

$$1\{s_{ic} \in k\} = \beta_j^k 1\{Judge_{ic} = j\} X_{ic} + \kappa_{t(c)}^k + \psi_{ic}^k \tag{1}$$

where s_{ic} is the treatment to which defendant i in case c was assigned, $\kappa_{t(c)}^k$ are month and day of week (time) fixed effects for case c's bond court date, and β_j^k are judge-specific propensities for assigning defendants with characteristics X_{ic} (e.g., charges, past cases, race) to a pretrial treatment level within k. For example, $k = \{2, 3\}$ means the judge assigns the defendant to EM or detention. Figure B.5 displays the heterogeneity in estimated β_j^k's across judges over defendant characteristics. I use fitted values from estimating equation (1) with OLS as the instrument for being assigned to a treatment within k, $Z_{ic}^k = \widehat{\Pr}\{s_{ic} \in k\}$.

Given the instruments, Z_{ic}^1 and Z_{ic}^3, I estimate the effect of EM versus release and EM versus detention with two-stage least squares (2SLS), with the following first and second stages:

$$[Release_{ic}, Detain_{ic}]' = \lambda_1 [Z_{ic}^1, Z_{ic}^3]' + \lambda_2 X_{ic} + \mu_{ic}. \tag{2}$$

$$Y_{ic} = \alpha_1 \widehat{Release}_{ic} + \alpha_2 \widehat{Detain}_{ic} + \theta X_{ic} + \epsilon_{ic} \tag{3}$$

where Y is the outcome of interest and $\widehat{Release}_{ic}$ and $\widehat{Detain}_{ic}$ are the instrumented treatments.

3.1.1 Instrument Validity

The validity of the instrumental variables strategy relies on multiple assumptions. First, we assume that the assigned judge is unrelated to the defendant, conditional on time fixed effects. The judge calendar (see Figure 1) alleviates concerns over violations of this assumption because judge assignment is determined solely by the sitting judge at the time of the bond hearing and not defendant characteristics. We can indirectly test this by seeing if defendant observables are predictive of judge assignment, which would suggest unobservables may be predictive as well, and thus exogeneity is violated. Table 3 displays the results of a multinomial logit with the defendant's bond court judge as the outcome and defendant observable as regressors (Column (2)) and month and day of week fixed effects as regressors (Column (3)). Consistent with exogeneity, defendant observables have virtually no predictive power (a pseudo-R^2 of almost 0). However, time fixed effects have significant predictive power (a pseudo-R^2 of almost 0.5), consistent with the calendar determining judge assignment.

An additional concern is that the judge influences defendant outcomes through means other than their pretrial treatment — for example, the judge also influences the defendant's assigned prosecutor — thus violating the exclusion restriction. Fortunately, Branch 1 operates on an entirely separate schedule from the other elements of the Cook County court system, and bond court judges do not play a role in future portions of the case — in the sample, only 0.2% of defendants saw their bond court judge in some capacity later in their case. Overall, these facts suggest violations of the exclusion restriction are not a significant concern.

Finally, we require that the judges actually influence treatment assignments. Table 4 displays the relationship between the standardized instruments, $\hat{Z}^1_{ic}$ and $\hat{Z}^3_{ic}$, and the endogenous variables $Release_{ic}$ and $Detain_{ic}$.[32] The first-stage relationship is strong: a 1 SD increase in $\hat{Z}^1_{ic}$ increases $Release_{ic}$ by about 9pp, while a 1 SD increase in $\hat{Z}^3_{ic}$ increases $Detain_{ic}$ by 10pp, and these relationships are consistent across specifications with and without controls and fixed effects. Figure 4 visualizes the support of $\hat{Z}^1_{ic}$ and $\hat{Z}^3_{ic}$ and the first stage.[33] Finally, Table B.4 contains results of the Sanderson and Windmeijer (2016) weak instruments test for multiple endogenous variables which indicate that weak instrument bias is not a significant concern.[34]

[32]Standardized instruments are first purged of the linear influence of defendant observables. For example, estimating equation (1) with $K = \{3\}$ and predicting the values gives Z^3_{ic}, then $\hat{Z}^3_{ic}$ can be recovered by taking the residuals from regressing $Z^3_{ic} = \beta X_{ic} + \kappa_{t(c)} + \gamma_{ic}$ ($\hat{\gamma}_{ic} = Z^3_{ic} - \hat{\beta} X_{ic} - \hat{\kappa}_{t(c)}$)and standardizing them ($\hat{Z}^3_{ic} = \frac{\hat{\gamma}_{ic} - \overline{\hat{\gamma}_{ic}}}{SD(\hat{\gamma}_{ic})}$). Similar results for the non-standardized instruments can be seen in Table B.6.

[33]Table B.5 displays the reduced form results. The judge instruments ($\hat{Z}^1$ and $\hat{Z}^3$) have a strong reduced form relationship with the outcomes of interest for pretrial misconduct; however, the strength of the reduced form relationship is either mixed or weak for case outcomes and post-case outcomes.

[34]Because instrumental variation comes from judges $\times$ observables = 154, we may have a many weak instruments problem. I use the specification $1\{s_{ic} \in k\} = \beta^k_j X_{ic} + \beta^k_2 X_{ic} + \kappa^k_{t(c)} + \psi^k_{ic}$ which produces one additional instrument per judge-characteristic interaction (which was previously collected into a single instrument). Table also includes the Montiel Olea and Pflueger (2013) robust weak instruments test.

3.1.2 Interpreting 2SLS Results

2SLS does not generally recover the effect of one treatment versus another if there are multiple treatments without additional assumptions (Kirkeboen, Leuven, and Mogstad (2016), Mountjoy (2022)).[35] We wish to know if $\hat{\alpha}_1$ and $\hat{\alpha_2}$ can be interpreted as the (weighted average of) causal effects of release versus EM and detention versus EM for some population of compliers — defendants whose treatment status was influenced by the judge instrument — or a local average treatment effect (LATE). Bhuller and Sigstad (2022) provide informal tests for the conditions under which 2SLS can recover a LATE. As discussed in Appendix A.3, I find small violations of these tests, so I cautiously proceed with the interpretation of the 2SLS results as a LATE.

3.2 Results: EM vs. Release

Column (1) of Table 5 displays the 2SLS results for the effect of a defendant being placed on EM relative to release for the main felony sample. Relative to release, EM reduces FTAs by -5.5pp ($p<0.05$) (-45.31% of release mean). This indicates that EM is an effective method for ensuring defendants show up for court, though this is consistent with a reminder effect rather than necessarily preventing flight, as Fishbane, Ouss, and Shah (2020) show that reminder text messages can have large effects on reducing FTAs and are not nearly as coercive as EM.

Despite the significant decrease in FTAs, EM *increases* the likelihood of a new case pretrial 7.1pp ($p<0.05$), which is a 41.5% increase relative to the release mean. We can decompose these effects into charge types: serious charges (violent crimes and non-drug felonies), low-level charges (drug crimes and non-felonies), and bond violation/escape charges (a single case can have multiple charge types). EM increases the likelihood of each charge

[35] If treatment effects are constant across individuals, then 2SLS will return the average treatment effect even in the case of unordered treatments. However, rejection of the null in the over-identification test in Table B.4 is consistent with different judge instruments identifying different treatment effects, indicative of heterogeneous treatment effects. Furthermore, we cannot exploit the ordered structure to recover a meaningful average causal response of increased treatment (Imbens and Angrist (1995)) due to the treatments not being quantifiable in a meaningful sense.

type: 1.6pp ($p>0.1$) for serious, 3.5pp ($p<0.05$) for low-level, 4pp ($p<0.05$) for violations. As discussed in Section 2.4, the increase in arrests for crimes pretrial consistent with EM increasing the probability of detection, such that more crimes may be detected under EM while total crimes may actually decrease due to deterrence and incapacitation. Despite the increase in new cases (the extensive margin), on the intensive margin (measured by the dollar cost of new cases), EM has a noisy but negative effect. So, while EM increases arrests, this is largely driven by low-cost crimes. On the other hand, the increase in violation charges is consistent with EM as a "punitive surveillance" device with many inflexible rules by which defendants cannot reasonably abide (Abid (2014), Hager (2020), Weisburd et al. (2021), Weisburd (2022)), leading to higher punishment costs.

I find that EM has a small or null effect on case outcomes, with a small and not statistically significant increase in the likelihood of a guilty felony charge and a decrease in the likelihood of being sentenced to incarceration. Similarly, EM has small positive effects but not statistically significant effects on total new cases over 4 years for all and felony cases, and on the intensive margin, EM has a noisy effect on post-trial cost crime over 3 years and over 4 years post-bond court. These results are consistent with EM being no more coercive or criminogenic than release, though they may be null due to the weakness reduced form relationship for these outcomes. Overall, EM does not appear to have higher indirect costs relative to release.

3.3 Results: EM vs. Detention

Column (2) in Table 5 displays the 2SLS results for the effect of a defendant being placed on EM relative to detention for the main felony sample. Relative to being detained pretrial, EM results in a significant increase in FTAs, 7pp ($p<0.05$) (241.7% increase relative to the detain mean). Similarly, EM also increases the likelihood of having a new case pretrial, by 16pp ($p<0.05$) (150.26% increase relative to the detain mean). Furthermore, there is a large and uniform increase across all charge types, meaning the increase in new cases pretrial is not solely driven by violations. EM decreases the intensive margin (cost of pretrial crime) when using the full murder cost (-\$8,300 ($p>0.1$)). However, this is due to outlier murder charges because when using the low-murder cost, EM leads to a statistically significant increase in pretrial crime costs (\$2,600 ($p<0.05$)). These results suggest that EM's deterrence effect and partial-incapacitation effect are not complete substitutes for detention's incapacitation effect and allow for defendants to FTA, but the increase in crime is relatively low-cost and likely less than even the short-run marginal cost of detaining a defendant.

EM has a small negative effect on the likelihood of being found guilty on a felony charge (-1.6pp ($p>0.1$)). However, EM does result in better case outcomes, with the likelihood of being sentenced to incarceration declining by -4.9pp ($p<0.05$), corresponding to -17.23% of the EM mean. These results suggest EM improves some case outcomes for defendants relative to detention and has lower indirect costs.

Despite a reduction in post-trial incarceration, EM has small and noisy effects on total cases and felony cases post-trial. EM has a negative but noisy effect on post-trial crime-cost when using both full (-\$17,000 ($p>0.1$)) and low (-\$1,800 ($p>0.1$)) murder cost, as well as crime-cost over 4 years following bond court for full (-\$40,000 ($p<0.05$) and low murder (-\$71 ($p>0.1$)). This suggests that any reduction in post-trial incapacitation or deterrence due to being placed on EM relative to detention is not sufficient to outweigh EM's lower criminogenic effect, and thus EM appears to be less criminogenic than detention resulting in either no change in or cost savings in post-trial criminal activity, further supporting EM's lower indirect costs relative to detention.

4 Heterogenous Effects

The 2SLS results do not tell us how the expansion of EM would affect different types of defendants, which is particularly important considering how the introduction or expansion of EM will lead to different populations being assigned to EM rather than release or detention.

For this, we turn to a marginal treatment effects (MTE) framework. In this section, I build on Heckman, Urzua, and Vytlacil (2006)'s generalized ordered choice Roy (1951)-style model to identify MTEs under weaker assumptions. Then, I provide a tractable method for semiparametrically estimating said MTEs. With MTEs, we are able to construct relevant treatment parameters, such as the average effect of moving a released defendant to EM. This section focuses on the case of 3 ordered treatments while Appendix C provides the general identification result and estimation method.

4.1 Empirical Strategy

4.1.1 Model

Selection into Treatment A bond court judge is randomly assigned (conditional on time fixed effects) to a defendant and assigns the defendant to one of three mutually exclusive treatment levels ($S \in \{1, 2, 3\}$) ordered by their intensity: release ($S = 1$), EM ($S = 2$), and detention ($S = 3$). The judge's decision is based partially on the defendant's observable characteristics (X). Different judges can view observable characteristics differently (e.g., one judge is harsh on drug offenders, another lenient, another prefers EM for them), so we interact the judge with observables to construct an instrument $Z = \text{Judge} \times X$. We can capture the judge's benefit from or desire to assign a defendant to higher treatment with some function $\tau(X, Z)$.

We assume all unobservable characteristics (unobservable to the econometrician but observed by the judge) that influence the judge's decision can be summarized in a single "latent" index V, which has a continuous but unspecified distribution. V captures the cost to the judge of assigning the defendant to a higher treatment and can be seen as the defendant's resistance to higher treatment, according to judges, or just resistance to treatment. We make no assumptions on the factors which determine V; V could be a function of the likelihood of committing a future crime, wealth, or hair length.

Collectively, the "net" benefit from assigning a defendant to a higher treatment level to the judge is $\tau(X, Z) - V$, and this index moves across a series of thresholds which are a function of judge preferences over observables ($C^s(Z)$). Higher treatments have weakly higher threshold values ($C^{s-1}(Z) \leq C^s(Z)\ \forall s$), and no one can be assigned below $S = 1$ nor above $S = 3$, so $C^0(Z) = -\infty$ and $C^3 = +\infty$. A defendant is assigned to treatment level s if and only if the net benefit is between the relevant cutoffs:

$$D_s = 1(S = s) = 1[C^{s-1}(Z) < \tau(X, Z) - V \leq C^s(Z)] \tag{4}$$

where $D_s \in \{0, 1\}$ indicates whether the defendant was assigned to treatment level s. This selection model assumes strict monotonicity of unobserved defendant's index (V), in that all judges agree on the ordering of V, and that higher values of V, all else equal, should receive weakly higher treatments.[36] However, judges can set different cutoffs for which values of V (conditional on X) receive higher treatments.[37]

Assuming V is independent of Z and X and using a standard transformation from Heckman, Urzua, and Vytlacil (2006), we can reorganize the selection model into a function predicted probabilities, and we can transform V, whose distribution we do not know, into a uniform random variable. Let $F_V(V) = U \sim Unif[0, 1]$ ($F_V(\cdot)$ is the CDF of V), such that U now captures resistance to higher treatment. Let $\pi_s(Z, X)$ as the probability a defendant receives a treatment *higher than* s given Z and X. Whereas we did not know $\tau(X, Z)$ or $C^s(Z) \forall\ s$, we can estimate $\pi_s(Z, X)$ as predicted probability of higher treatment based on Z and X. $\pi_1(Z, X) = \Pr(S > 1)$ is the probability of not being released, and $\pi_2(Z, X) = \Pr(S > 2) = \Pr(S = 3)$ is the probability of being detained, given values of X and Z. Equation (4) can now be expressed as:[38]

[36] All judges agree that any defendant a with $V_a = v'$ should receive a *weakly* higher treatment than any defendant b who is identical to defendant a except that $V_b = v$ if $v' < v$. This is analogous to the monotonicity assumption discussed in Imbens and Angrist (1994) and Vytlacil (2002), but in an ordered environment (Vytlacil (2006)).

[37] For example, judge A is more lenient with felony drug charges than judge B on the EM (2) versus detention (3) margin, but more strict than judge B on the EM (2) versus release (1) margin: $C^2(A \times \text{Drug}) > C^2(B \times \text{Drug})$ and $C^1(A \times \text{Drug}) < C^1(B \times \text{Drug})$, so a larger range of values $\tau(X, Z) - V$ for defendants with drug charges will be assigned to EM under judge A relative to judge B.

[38] Because $\pi_s(Z, X) = \Pr(S > s|Z, X) = F_V(\tau(X, Z) - C^s(Z)) = \Pr(\tau(X, Z) - V \geq C^s(Z))$, then

$$S = \begin{cases} 1 \text{ (Release)}, & \text{if } \; U \geq \pi_1(Z, X) \\ 2 \text{ (EM)}, & \text{if } \; \pi_1(Z, X) > U \geq \pi_2(Z, X) \\ 3 \text{ (Detain)}, & \text{if } \; \pi_2(Z, X) > U \end{cases}$$

Essentially, the defendant is assigned to treatment higher than s if the observable factors pushing them higher, as measured $\pi_s(Z, X)$, are larger than their unobserved resistance to treatment, as measured by U.

Potential Outcomes Let Y be the outcome of interest for a specific defendant, such as whether the defendant fails to appear in court ($Y \in \{0, 1\}$) or the number of cases against them post-trial ($Y \in \{0, 1, 2, \ldots\}$). Y_s is the defendant's treatment (s)-specific potential outcome: their outcome *if* they were placed in treatment level s (Holland (1986)). Then, the observed outcome of a defendant is their potential outcome for the treatment they actually received: $Y = \sum_{s=1}^{3} Y_s \times D_s$, where Y_s is a function of observables (X) and unobservable factors ω_s $\forall s$.

For example, if we could observe both Y_1 and Y_2 for a defendant, then $Y_2 - Y_1$ would give us that defendant's treatment effect of being placed on EM ($S = 2$) relative to being released ($S = 1$). However, we only ever observe a single potential outcome for each defendant. Furthermore, $\mathbb{E}[Y_s] \neq \mathbb{E}[Y | D_s = 1]$ because treatment is not randomly assigned and is possibly assigned based on potential outcomes.

For estimation and identification in this section, I assume full independence of observables and unobservables (Carneiro, Heckman, and Vytlacil (2011)), $(X, Z) \perp (\omega_1, \omega_2, \omega_3, U)$, and that Y_s is linear in observables. This implies additive separability: $Y_s = \beta_s X + \omega_s$, meaning the way observables influence potential outcomes is unrelated to the way unobservables do.[39]

$1(S = s) = 1[F_V(\tau(X, Z) - C^{s-1}(Z)) > F_V(V) \geq F_V(\tau(X, Z) - C^s(Z))]$, so $1(S = s) = 1[\pi_{s-1}(Z, X) > U \geq \pi_s(Z, X)]$.

[39] While additive separability can be obtained with weaker assumptions than full independence (Brinch, Mogstad, and Wiswall (2017)), assuming full independence is appropriate in this case because of the interaction

4.1.2 Treatment Effects

Because all judges agree on the ranking of defendants in terms of U but disagree on where to set cutoffs and how to weigh observable factors, two nearly identical defendants in terms of U can be assigned to different treatments solely due to their quasi-randomly assigned judge. The judge only influences outcomes by influencing treatment, not by influencing potential outcomes. Given we observe X, which influences both treatment and potential outcomes, we must overcome the fact that the unobserved component, ω_s, may also influence treatment assignment in order to measure the effect of one treatment versus another. As we have assumed V (and thus U) captures all unobserved factors determining treatment, we have that: $\mathbb{E}[Y_s|X = x, U = u] = \beta_s x + \mathbb{E}[\omega_s|u]$.

So, we will use the variation in judges to study how defendants across the distribution of U respond to different treatments. Because U ranks defendants by how likely they are to be assigned to a higher or lower treatment, different treatment effects across different values of U inform how outcomes will change if, for example, EM is expanded to defendants who are likely to be detained (low U, all else equal) or released (high U, all else equal). Note that we can only make such comparisons where there is common support, meaning that judges must have sufficiently different cutoffs such that we observe individuals with $U = u$ in different treatments.

MTEs and MTRs We can define the marginal treatment effect (MTE) as the average treatment effect of a defendant with observables $X = x$ and unobservable resistance $U = u$ from one treatment s to another s' as: $MTE_{s',s}(X = x, U = u) = \mathbb{E}[Y_{s'} - Y_s|X = x, U = u]$.[40] In particular, we are interested in the effects of moving from release to EM ($MTE_{2,1}(x, u)$) and moving from detention to EM ($MTE_{2,3}(x, u)$). We can decompose any MTE into the difference between two marginal treatment response (MTR) functions, $\mathbb{E}[Y_{s'}|X, U]$ and $\mathbb{E}[Y_s|X, U]$ (Mogstad, Santos, and Torgovitsky (2018)). Given the prior assumptions, we

between observables and judges in constructing the instruments.

[40] When the treatment levels are adjacent (e.g., $s' = s + 1$), the MTE is referred to as the transition-specific marginal treatment effects (TSMTE) (Heckman and Vytlacil (2007)).

have:

$$
\begin{aligned}
MTE_{s',s}(X,U) = \mathbb{E}[Y_{s'} - Y_s|X,U] &= \underbrace{\mathbb{E}[Y_{s'}|X,U]}_{\text{MTR of } s'} - \underbrace{\mathbb{E}[Y_s|X,U]}_{\text{MTR of } s} \\
&= (\beta_{s'} - \beta_s)X + \mathbb{E}[\omega_{s'}|U] - \mathbb{E}[\omega_s|U]
\end{aligned}
\tag{5}
$$

Then, we can identify MTEs by first identifying MTRs separately. In Appendix C, I show that each MTR can be identified relying solely on variation between π_s and π_{s-1} for $\mathbb{E}[Y_s|X,U]$ $\forall s$. For example, by taking the derivative of $\mathbb{E}[Y \times D_s|\pi_s, \pi_{s-1}, X]$ (which can be estimated) with respect to π_{s-1}, we recover the MTR for $S = s$, $\mathbb{E}[Y_s|U = u, X = x]$, at $u = \pi_{s-1}$. Rose and Shem-Tov (2021) developed a similar approach but with the aim to bound MTRs.

Treatment Parameters of Interest As shown in Heckman and Vytlacil (2005), with marginal treatment effects and full support, we can construct a variety of treatment parameters of interest: in particular, the average treatment effects of moving from release or detention to EM ($ATE_{2,1}$ and $ATE_{2,3}$), the average treatment effect of moving from release to EM for released defendants (ATR), and the average of moving from detention to EM for detained defendants (ATD).

The ATE tells us the expected difference in an outcome if we took a random defendant in the population and assigned them to one treatment (EM) versus another (release or detention), fixing x at $\bar{X}$ (the mean of observables across all observations). Because u is uniformly distributed in the population, each point on the MTE is weighted equally.

$$
\begin{aligned}
ATE_{2,1} &= \int_0^1 MTE_{2,1}(u, \bar{X})du, \\
ATE_{2,3} &= \int_0^1 MTE_{2,3}(u, \bar{X})du
\end{aligned}
$$

The ATR tells us the expected difference in outcomes if we took a random defendant who was released and assigned them to EM — that is, what would happen if we removed release as an option for defendants and assigned them to EM instead? Similarly, the ATD tells us the expected difference in outcomes if we took a random defendant who was detained and assigned them to EM — that is, what would happen if we removed detention as an option for defendants and assigned them to EM instead? These effects are relevant when discussing policies under which EM is expanded to replace release or detention.[41]

$$ATR = \int_0^1 MTE_{EM,R}(u,x)W_R(u,x)du,$$
$$ATD = \int_0^1 MTE_{EM,D}(u,x)W_D(u,x)du$$

In contrast with the ATE, which evenly weights all points on the MTE curve because u is uniform in the population, the ATR(D) applies heterogeneous weights. The ATR(D) weighs the effect of the observables of individuals more likely to be released (detained) more heavily, meaning those with low (high) values of π_1 (π_2), and they weigh the effects of unobservables more heavily for individuals who are unobservably more likely to be released (detained), equating to higher weights on high (low) values of $U = \pi_1$ ($U = \pi_2$). These weights are both contained in W_R and W_D, which both integrate to 1, though their exact composition depends on functional form assumptions in practice.

4.1.3 Estimation

Estimation of MTRs requires first recovering π_1 and π_2 and then recovering MTEs as the difference between MTRs. Steps and assumptions for estimation are discussed in more detail in Appendix C.

Recovering π π_1 and π_2 can be recovered separately by estimating equation (1) with the dependent variables being $1[S_i > 1] = D_{i,2} + D_{i,3}$ and $1[S_i > 2] = D_{i,3}$, respectively, using a probit model and predicting treatment probabilities.[42]

[41] However, they differ from the policy relevant treatment effect focusing on shifting propensity scores and treatment-uptake decisions (Heckman and Vytlacil (2005), Carneiro, Heckman, and Vytlacil (2011), Mogstad and Torgovitsky (2018)).

[42] This allows for all judges' preferences over observables to be heterogeneous across treatment margins (approximating instrument-dependent $C^1(Z)$ and $C^2(Z)$), but also allows observables to uniformly affect the likelihood of crossing a treatment threshold ($\tau(X,Z)$). Note that this is similar to, but not equivalent to, estimating an ordered probit with judge-specific thresholds, as it involves additivity across judge $\times$

Recovering MTRs and MTEs As discussed above, we take the derivative of $\mathbb{E}[Y \times D_s|\pi_s, \pi_{s-1}, X]$ with respect to either π_s or π_{s-1} to recover an MTR ($\mathbb{E}[Y_s|X, U]$) at $\pi_s, \pi_{s-1} = u$. Given additive separability of Y_s, $\mathbb{E}[Y_s|X = x, U = u] = \beta_s x + \mathbb{E}[\omega_s|U = u]$.

Using the fact that we observe Y and $D_s = 1[\pi_s \leq U < \pi_{s-1}]$, we can construct $Y \times D_s$: $Y_i \times D_{i,s}$ is 0 if $S_i \neq s$ and Y_i if $S_i = s$. I define $\Lambda_s(\pi) = \int_0^\pi \mathbb{E}[\omega_s|U = u]du$, such that:

$$\begin{aligned}\mathbb{E}[Y \times D_s|\pi_s, \pi_{s-1}, X] &= \beta_s X \times (\pi_{s-1} - \pi_s) + \mathbb{E}[\omega_s \times D_s|\pi_s, \pi_{s-1}] \\ &= \beta_s X \times (\pi_{s-1} - \pi_s) + \Lambda_s(\pi_{s-1}) - \Lambda_s(\pi_s)\end{aligned}$$

where the change in the second line is due to $\int_{\pi_s}^{\pi_{s-1}} \mathbb{E}[\omega_s|U = u]du = \int_0^{\pi_{s-1}} \mathbb{E}[\omega_s|U = u]du - \int_0^{\pi_s} \mathbb{E}[\omega_s|U = u]du$. Note that the separable nature of $\Lambda_s(\pi_{s-1}) - \Lambda_s(\pi_s)$ is a result of the single index model.

We can approximate the difference $\Lambda_s(\pi_{s-1}) - \Lambda_s(\pi_s)$ using sieves which can be, for example, a B-spline or polynomial. Specifically, let $\Phi_s(\pi)$ denote a vector of basis functions used to approximate $\Lambda_s(\pi_{s-1}) - \Lambda_s(\pi_s)$, such that $\Lambda_s(\pi_{s-1}) - \Lambda_s(\pi_s) \approx \phi_s[\Phi_s(\pi_{s-1}) - \Phi_s(\pi_s)]$, where ϕ_s is a vector of coefficients for each s. For example, with a 3rd polynomial sieve, $\Phi_s(\pi) = [\pi, \pi^2, \pi^3]$ for each s. Using this sieve approximation and including ϵ as an error term, we can estimate with OLS:

$$Y \times D_s = \beta_s X(\pi_{s-1} - \pi_s) + \phi_s[\Phi_s(\pi_{s-1}) - \Phi_s(\pi_s)] + \epsilon \tag{6}$$

Then, we can approximate the MTR by taking the derivative with respect to π_s or π_{s-1} to recover $\mathbb{E}[Y_s|x, u] \approx \hat{\beta}_s x + \hat{\phi}_s \Phi'_s(u)$. Finally, we can take the difference between MTRs evaluated at the same values of X and U to recover $MTE_{s+1,s}(x, u) \approx (\hat{\beta}_{s+1} - \hat{\beta}_s)x + \hat{\phi}_{s+1}\Phi'_{s+1}(u) - \hat{\phi}_s \Phi'_s(u)$.

Main results use a 3rd degree polynomial for Φ, meaning MTRs and MTEs effectively use a 2nd degree polynomial. 95% confidence intervals are based on bootstrapped estimates with

characteristics by threshold and is less sensitive to misspecification of the distribution of V.

200 runs producing non-symmetric confidence bands (see Appendix C.5 for more information).

Common Support MTEs can be identified where common support exists between MTRs. Because there are only 3 treatment levels, common support hinges upon the only middle treatment, EM, where we require overlapping support for π_1 and π_2 in the EM sample.[43] Figures B.7 - B.9 display the support and variation of π_1 and π_2. Using a 1% sample trim, the support of EM is $\pi_1, \pi_2 \in [0.25, 0.77]$ (see Figure B.9), so the main results focus on EM versus release and EM versus detention $u \in [0.25, 0.77]$, corresponding to mid-range defendants and excluding defendants at both extremes.

Constructing Treatment Parameters I follow Cornelissen et al. (2016) in the construction of these treatment effects (see Appendix A.4 for details). Importantly, the ATR, ATD, and ATEs can only be identified with full support (the support of integration is 0 to 1). Given that we lack full common support ($u \in [0.25, 0.77]$), we instead can identify common support equivalents: CATEs, CATR, CATD (i.e., common support ATEs, ATR, ATD) (Carneiro, Heckman, and Vytlacil (2011), Bhuller et al. (2020)).[44] In practice, I rescale weights to ensure they integrate to 1 over the common support. See Figure B.10 for the weights for the unobservable components under common and full support.

4.1.4 Comparison With Prior Methods

This semiparametric sieve approach has advantages over prior methods for estimating MTEs, specifically the fully parametric approach that assumes jointly normal unobservables (the "Normal" method) from Heckman, Urzua, and Vytlacil (2006), Heckman and Vytlacil (2007), and Cornelissen et al. (2018). In addition to being fast and straightforward, this approach allows for non-linear MTEs, while the Normal method produces MTEs with monotonic and largely pre-determined shapes (effectively linear for all but the more extreme values of u)

[43] Because $\pi_0 = 1$ and $\pi_3 = 0$, MTRs for release ($E[Y_1|x, u]$) and detention ($E[Y_3|x, u]$) only require variation in π_1 and π_2, respectively.

[44] Note that with Bhuller et al. (2020)'s notation, we would refer to these analogous parameters simply the ATE/ATR/ATD for the common support sample, while Carneiro, Heckman, and Vytlacil (2011) use $\widetilde{ATE}$, for example, to differentiate the common support ATE from the true ATE.

as a result of the parametric assumptions. While prior applications of the Normal method only recover MTEs, this method also allows us to recover MTRs. Furthermore, while an advantage of the Normal method is that it allows for MTE estimation across the full support, the sieve method can be easily adapted to a fully parametric polynomial model to recover MTRs beyond common support and make full support comparisons. In Section 6, I compare the main common support results to full support using polynomial and the Normal method. Appendix A.5 provides more comparisons between the method presented in this paper and the Normal method for MTEs.

4.2 Results: EM vs. Release

Columns (1) and (2) of Table 6 and the left panels of Figures 5 - 7 display the effects of EM relative to release for the average defendant (CATE) and released defendants (CATR), along with 95% confidence intervals. Figures B.11 - B.13 display the MTE results, where higher values of U (x-axis) correspond to defendants who are easier to release.[45]

Relative to release, EM decreases the likelihood of a failure to appear in court. The effect is largest for low U defendants and smallest for high U defendants, and this selection pattern is consistent with judges assigning EM to defendants who would benefit (in terms of lower FTAs) from EM relative to release, based on unobservable factors. As a result, the CATE estimate (-9.8pp ($p<0.01$)) is larger than the CATR (-5.8pp ($p<0.01$)) (close to the 2SLS estimate), meaning expanding EM to defendants more likely to be released results in smaller benefits in terms of FTA reduction.

EM increases new cases pretrial for less resistant defendants (less likely to be released), but the effect moves negative as resistance increases, and this results in a noisy and near null effect across the distribution despite a positive and large 2SLS estimate. The noisy effect is driven by a null effect on new case charges for serious crimes, a negative effect on low-level

[45] Figures B.17-B.19 display the MTRs. Recall that the MTEs are evaluated at values of $U = \pi_1$, so the x-axis is a measure of unobserved resistance to higher treatment — for example, $MTE_{2,1}$ at $U = \pi_1 = 0.25$ is for individuals who are *unobservably* harder to release (less resistant) and at $U = \pi_1 = 0.77$ are unobservably easier to release (more resistant).

charges, and a positive effect on violation charges– all sharing the same selection pattern with larger effects for low-resistance defendants and smaller or negative effects on high-resistance ones. Overall, EM reduces the likelihood of charges for low-level crimes with -3pp ($p<0.1$) and increases the likelihood of violations by 3.2pp ($p<0.01$) for the average defendant. The effects on defendants likely to be released (CATR) are more muted, with a -1.8pp ($p>0.1$) decrease for low-level charges and a 1.4pp ($p<0.05$) increase for violations. Consistent with this, the effects on pretrial case cost are negative but noisy. From this, we can conclude that EM likely has a crime-reducing effect relative to release, reducing the likelihood of a new case with crime-related charges, but there is no significant effect on the intensity or social cost associated with these charges.

For case outcomes, EM increases the likelihood of a guilty felony charge and being incarcerated for defendants least likely to be released, but this effect decreases rapidly as defendants become more likely to be released. This results in positive CATEs (6.1pp ($p<0.05$) and 15pp ($p<0.01$)) but a negative CATR for felony charges (-2.8pp ($p>0.1$)) and a barely positive effect on sentencing (1.8pp ($p>0.1$)). While these patterns could be indicative of EM being coercive for higher-severity defendants and improving outcomes for low-severity defendants — possibly by reducing FTAs and rearrests and thus strengthening their cases — these results are largely suggestive due to small 2SLS estimates and the weak reduced form relationship.

Post-trial, EM increases the total number of new cases over 4 years, with a parabolic selection pattern: small effects on high and low resistance defendants and larger effects for mid-resistance defendants. This results in similar CATE and CATR estimates (0.39 ($p<0.01$) and 0.35 ($p<0.01$)) slightly larger than the 2SLS estimate. However, the effects on felony cases are small (and null for 2SLS), and EM has a noisy but near null (or positive in the case of low-murder cost) effect on total case cost over 4 years post-bond court, suggesting no significant criminogenic effect.

Overall, relative to release, EM increases punishment costs through increasing new cases for violations, and there is weak evidence for higher indirect costs through increased future low-cost recidivism and worse sentencing. However, EM reduces crime costs through a reduction in new pretrial cases for criminal activity and lower pretrial crime costs, as well as a reduction in FTAs. However, this reduction in pretrial new cases is small and driven by low-level crime. Selection patterns indicate that both the costs and benefits of moving from release to EM are smaller for defendants who are released relative to the average defendant.

4.3 Results: EM vs. Detention

Columns (3) and (4) of Table 6 and the right panels of Figures 5 - 7 display the CATE and CATD estimates along with 95% confidence intervals, and Figures B.14 - B.16 display the MTE results.[46]

Relative to detention, EM increases the likelihood a defendant fails to appear in court across all defendants, but the effect decreases in magnitude as defendant resistance increases. So, defendants who are unobservably easier to detain experience smaller increases in FTAs, a pattern inconsistent with selection into treatment upon gains in terms of reduced FTAs. This results in a lower CATD (3.3pp ($p<0.01$)) than CATE (5.3pp ($p<0.01$)), both smaller than the 2SLS estimate.

For new cases pretrial, EM increases the likelihood of a new case for the average defendant (6.5pp ($p<0.01$)) and similarly for defendants more likely to be detained (7.3pp ($p<0.01$)). This effect is driven by low-level and violation charges with CATEs (CATDs) of 3.8pp ($p<0.05$) (3.1pp ($p<0.1$)) and 3.7pp ($p<0.01$) (5.3pp ($p<0.01$)), respectively. Importantly, defendants assigned to EM relative to detention are no more likely to have new cases with serious charges. Notably, EM's effect on violations is relatively constant across defendant resistance. These results are in contrast with the 2SLS estimates, which are uniformly positive and larger than the CATEs or CATDs, and suggest over-weighting low-resistance defendants if interpreted

[46] Figures B.17-B.19 display the MTRs. Recall that the MTEs are evaluated at values of $U = \pi_2$, so the x-axis is a measure of unobserved resistance to higher treatment — for example, $MTE_{2,3}$ at $U = \pi_2 = 0.25$ is for individuals who are *unobservably* easier to detain (less resistant) and at $U = \pi_2 = 0.77$ are unobservably harder to detain (more resistant).

as a LATE. Nevertheless, the effect on pretrial crime costs is noisy and negative using full murder cost but small and positive with low-murder costs — consistent with the small effects on serious charges. Overall, it appears EM does not result in more costly pretrial crime but does involve more rearrests for violations.

The MTE results suggest that EM improves case outcomes relative to detention, decreasing the likelihood of a guilty felony charge and of being sentenced to incarceration, with effects increasing in magnitude as defendant resistance to detention increases. This is consistent with judges having some signal over defendant guilt or case strength when determining pretrial treatment, such that defendants who are unobservably (to the econometrician) less resistant to detention experience smaller changes in outcomes. These patterns result in CATEs of -30pp (p<0.01) and -29pp (p<0.01) and CATRs of -22pp (p<0.01) and -23pp (p<0.01), respectively; however, the 2SLS estimates are significantly smaller (though still negative) than the CATD and CATE estimates.[47]

Post-trial, EM has a positive effect on total new cases over 4 years for low-resistance defendants and has a negative effect for high-resistance defendants, with a similar pattern for felony cases. While this results in null CATE estimates (similar to the 2SLS estimates), the CATDs are positive (0.28 (p<0.01) for all and 0.14 (p<0.01) for felony). However, when post-trial cases are weighted by their crime cost, EM has a noisy negative effect on the average (-$44,000 (p<0.1)) and detained (-$31,000 (p>0.1)) defendant with full murder cost, and small effects with low murder cost. On net, EM has a negative effect on total case cost over 4 years following bond court for both full and low murder cost, with effects slightly increasing in magnitude as defendant resistance increases. This results in full and low murder cost CATEs (CATDs) of -$63,000 (p<0.01) (-$52,000 (p<0.05)) and -$5,700 (p<0.05) (-$1,400 (p>0.1)), with similar 2SLS estimates.

Overall, EM allows for increased pretrial misconduct in both pretrial crime, driven by

[47] While these are large estimates, they are not inconsistent with prior work: Dobbie, Goldin, and Yang (2018)'s 2SLS estimates for the effect of release versus detention show release reduces guilty verdicts by about 24% of the detained defendant mean.

low-level charges and violations, and FTAs. However, the main estimates are consistent with EM improving defendant case outcomes (though estimates range widely) and having a small effect on future cost-weighted recidivism. Selection patterns indicate that the average defendant experiences larger benefits and smaller costs of being placed on EM relative to detention compared with defendants who are detained.

5 Policy Implications

5.1 Are Defendants Elastic?

As discussed in Section 2.4, EM aims to reduce pretrial crime through deterrence (increasing the probability of detection) and partial incapacitation (house arrest). If EM solely operated through incapacitation, and the change in detection rates was small, then EM's effect on new cases pretrial would be proportional to EM's effect on pretrial crime. However, the effect on new cases pretrial, which was relatively small relative to EM versus release, will understate the effect on pretrial crime. This is because EM likely increases the probability of detection, and thus a larger share of crimes will be detected under EM. If EM solely operated through incapacitation (house arrest), then this would suggest EM's crime-reduction effect is small. We can explore the sensitivity of pretrial crime with respect to the probability of detection and EM's crime-reduction effect.

First, we can infer a lower bound for the average defendant's elasticity with respect to the probability of detection, $\epsilon_p^c \approx \frac{\frac{\Delta c}{\frac{c^{EM}+c^R}{2}}}{\frac{\Delta p}{\frac{p^{EM}+p^R}{2}}}$, where $\Delta x = x^{EM} - x^R$, using the fact that $arrests\ (a) = crime\ (c) \times p$, to give us an expression for implied elasticity.[48] Even though we do not observe p^R or p^{EM}, we can trace the implied elasticities by assuming reasonable values for p^R across values of p^{EM}, and we observe a^R and a^{EM} as the average MTRs (weighted for

[48]This is a lower bound on $\epsilon_p^c \geq \hat{\epsilon}_p^c$ because any incapacitation effect (reducing c^{EM}) will be attributed to deterrence. $\epsilon_p^c \approx \hat{\epsilon}_p^c = \frac{\frac{\Delta c}{\frac{c'+c}{2}}}{\frac{\Delta p}{\frac{p'+p}{2}}} = \frac{\frac{\frac{a'}{p'}-\frac{a}{p}}{\frac{a'}{p'}+\frac{a}{p}}}{\frac{p'+p}{p'-p}} = \frac{\frac{p'+p}{p'-p}(\frac{a'}{p'}-\frac{a}{p})}{\frac{a'}{p'}+\frac{a}{p}} = \frac{\frac{2p+\Delta p}{\Delta p}(\frac{\Delta a+a}{\Delta p+p}-\frac{a}{p})}{\frac{\Delta a+a}{\Delta p+p}+\frac{a}{p}}$

average or released defendants) for a specific crime type, and similarly for fitted values using 2SLS estimates.[49]

The results are displayed in Figure 8 for low-level crimes and serious crimes using low, middle, and high values for p^R. The results indicate that low-level crime is likely elastic with respect to surveillance $\epsilon_p^{c,low} < -1$, though defendants more likely to be released are slightly less elastic. However, 2SLS results suggest defendants are likely relatively inelastic, consistent with the positive point estimates for new cases pretrial. We cannot reject the null that serious crime is relatively inelastic with respect to the probability of detection ($\epsilon_p^{c,serious} > -1$) across all estimate types (2SLS, CATE, CATR). The potential inelasticity of crime with respect to detection reduces the efficacy of surveillance as a social savings policy because even though crime is reduced, socially costly arrests may rise, thereby increasing punishment costs. Furthermore, these are both the lower bounds for elasticities, meaning defendants are likely more *inelastic* than $\hat{\epsilon}_p^c$ suggests.

The result that defendants may be relatively inelastic in serious crimes is in contrast with other empirical studies in which increased detection led to lower *observed* violations (Grogger (1991), Bar-Ilan and Sacerdote (2001), Hawken and Kleiman (2009)) with which the effects on low-level crimes are more consistent. However, the inelasticity for serious crimes is consistent with the discussions of arrest rates and incarceration (Kleiman (2009), Durlauf and Nagin (2011)).[50] This raises an important policy issue with respect to surveillance: if surveillance detects more crimes but punishments stay high, and defendants are not sufficiently elastic with respect to detection, then surveillance will result in less crime but more punishment.[51]

[49] Based on FBI statistics, over 40% of violent crimes are cleared, with about 30% of robberies cleared, and almost 20% of property crimes are cleared. I assume slightly higher than average clearance rates for crimes committed by defendants who are released on bail because their location is known and they were in recent contact with the system and are tracked by the court system and the Sheriff (e.g., court attendance). See https://ucr.fbi.gov/crime-in-the-u.s/2017/crime-in-the-u.s.-2017/topic-pages/clearances for 2017 clearance rates.

[50] Even if longer prison sentences deter crime, if potential criminals are relatively inelastic, then the incarcerated population will rise despite higher deterrence.

[51] Relatedly, in the theoretical literature discussing racial disparities in hit rates, it has been argued that the optimal search rate is not that which produces the highest hit rates but that which deters the most crime (i.e., officers should target the more elastic group) (Persico (2002), Manski (2005)).

5.2 How Much Crime Does Surveillance Prevent?

Second, we can use a similar method to compute the total cost of crime prevented under EM relative to release for the average defendant. Using a similar logic as above, and using $c \times g$ (cost of crime) $= \frac{a \times g}{p}$, so $\Delta cg = \frac{(a\times g)^{EM}}{p^{EM}} - \frac{(a\times g)^{R}}{p^{R}}$. This form simplifies cost-weighted crime as a single type with a single probability of detection, rather than summing over different types with different probabilities of detection. With this simplification, we can use average MTRs ($(a \times g)^R$ and $(a \times g)^{EM}$) for the crime cost of new cases (arrests) pretrial, and similarly for CATR weights.[52] Note that this calculation does not require any assumption on EM's incapacitation effect.

The top row of Figure 9 displays the implied changes in pretrial crime due to EM's crime-reduction effect for all charges. Assuming a crime-cost weighted (mainly weighted by murder) $p^{EM} = 0.5$, the point estimates suggest EM prevents nearly \$25,000 of pretrial crime for the average defendant and less than \$10,000 for the released defendants — importantly, we cannot reject a null (or even positive) effect by EM relative to release. However, this assumes that EM detects 100% of crimes and that all charges are correct. Given that a large fraction of charges does not result in a guilty finding, we can re-estimate the curve using guilty charges only, and the effects shrink significantly: the average defendant effect is less than \$10,000, and the released defendant effect is less than \$7,500 for $p^R = 0.5$ and assuming EM detects all crimes. Only the guilty CATR estimates can reject a null effect but cannot reject small gains. To avoid results being driven by outlier-murder cases, we can use the low-murder cost, which significantly reduces the gains to less than \$4,000 for average and \$3,000 for released with $p^R = 0.5$ and $p^{EM} = 1$. For comparison, the direct cost alone of EM operation for the average main sample defendant on EM (about 100 days) is about \$1,500.

[52] 2SLS estimates predict negative crime cost under EM, making the exercise more difficult to perform.

5.3 Surveillance vs. Incapacitation

The results in the prior section suggest that EM is preferable to detention (better case outcomes, less costly total new cases both pre- and post-trial), except with respect to pretrial misconduct. The central question for EM versus detention is whether or not EM's crime-reducing effect and lower indirect and direct costs are a sufficient replacement for detention's incapacitation effect and higher direct and indirect costs. Dobbie, Goldin, and Yang (2018) estimate that pretrial detention (relative to release) costs over $50,000 in lifetime costs per marginal defendant. While we cannot provide a similar calculation for EM versus detention due to a lack of individual financial data, we can compare the costs of pretrial misconduct under EM versus detention using a similar method as above.

Figure 10 displays the changes in pretrial crime cost between EM and detention (release following detention). Essentially, for EM's effect to be, on average, less effective over the pretrial period than detention, EM would need to detect *fewer* crimes than release (post-detention). This may be due to individuals released after long detention durations committing more serious crimes than those on EM due to jail's criminogenic effects.

By construction of the treatments, detained defendants may be released at some point during the pretrial period (just not within 7 days). So, we can examine the magnitude of how much crime EM allows for assuming detention is complete incapacitation, that is, defendants are held in jail until their case is complete, essentially being denied bond.[53] Figure 11 displays the results. For guilty charges, assuming $p^{EM} > 0.5$, crime costs on EM are less than $5,000 for the average and detained defendant. For comparison, Miller et al. (2021) estimates the cost of a larceny/theft at around $\$5,000$, while the cost of a single day of pretrial detention in Cook County is over $100 more expensive than a day on EM– making the direct cost savings alone at least $4,000 for the median defendant. Overall, it is highly unlikely that EM's effect is not a cost-effective substitute for detention's incapacitation effect.

[53]While useful, if implemented, detention would likely result in additionally harmful effects for defendant cases and recidivism outcomes (as well as employment loss), in addition to unmeasured costs on families, individuals, and communities, and an increase in large direct costs from holding more people in jails.

5.4 Net Costs and Savings from Expanding EM

By applying dollar amounts to the various outcomes in the framework from Section 2.4, we can approximate the net savings associated with replacing release or detention for EM. Table 7 displays the breakdown of cost differences for ATEs, ATR, and ATD estimates using approximations for various costs.[54] Figure B.20 displays the total cost MTRs for both full and low-murder costs — notably, the MTRs are generally decreasing in defendant resistance to higher treatment.

In terms of direct cost differences replacing release with EM costs a few hundred dollars, while replacing EM with detention saves more than $7,000 per defendant. Changes in failure to appear rates are neither sizable nor costly enough to make a significant difference on average. Replacing EM with release leads to increased sentencing costs as defendants are more likely to be sentenced to incarceration, at $350 and $3,000 for average and released, while EM reduces sentencing costs relative to detention at $6,000 and $5,000 for average and detained.

While the prior estimates are relative precise, crime cost estimates are not. Differences in pre-trial crime cost are large with full murder costs with EM saving between $19,000 and $47,000 per defendant relative to both release and detention using full murder costs for all defendant types, but confidence intervals include zero and cannot reject modest costs or even more massive savings. Using low-murder costs, EM still saves between $8,000 and $12,000 per defendant relative to either release or detention, and all the confidence intervals reject savings lower than $2,000 per defendant. Recidivism (post-trial) crime costs over 3 years add to these savings for EM relative to detention with massive but imprecise savings with

[54] I use average costs of $15 per defendant-day for EM and $150 defendant-day for jail and apply a 40% marginal cost. FTAs are valued at $1,000. Crime costs are based on aggregated arrest costs assuming a 30% cost-weighted detection rate under release, post-detention release, or post-trial, and a 60% cost-weighted detection rate under EM. Pretrial crime costs include punishment costs. Sentencing costs are assumed to be $20,000 per defendant sentenced to incarceration, based on Mai and Subramanian (2017) estimates of average prison cost per defendant in Illinois in 2015 of $33,507 and applying a 60% marginal cost discount. This also assumes defendants sentenced to incarceration are serving one year on average, which may be an underestimate (over 75% of individuals sentenced to prison or jail are sentenced to prison).

full murder cost (cannot reject losses) and smaller but precise (reject savings smaller than $4,500). However, recidivism costs reverse some of the benefits of EM relative to release. EM results in over $20,000 in post-trial crime with full murder costs (but cannot reject massive savings or costs). With low-murder costs EM costs about $10,000 per defendant in post-trial crime, with smaller confidence intervals, though only released defendants reject any savings.

On net, EM than release is more costly for released defendants ($6,000 and $2,500 using full and low-murder costs), but confidence intervals are wide. In contrast, the signs for average defendants disagree, and the confidence intervals include both massive savings and costs. Relative to detention, EM likely induces net savings as estimates are consistently negative for average and detained defendants, and confidence intervals using low-murder cost are much more precise and reject savings smaller than $4,500 per defendant. Notably, these estimates are biased in favor of higher treatment levels, as they do not account for employment, health, or personal costs, which are likely high in detention (Dobbie, Goldin, and Yang (2018)) if not EM as well. Nevertheless, these results are consistent with EM being a cost-effective alternative to detention, even for otherwise-detained defendants. At the same time, it is not clear that EM induces cost-savings relative to release, particularly for released defendants.

6 Robustness

I test the sensitivity of my results to alternative samples and specifications for five main outcomes. Figures 12-14 display the CATE, CATR, CATD, and 2SLS estimates, while Figures B.21-B.23 display the MTE curves. Additional checks are discussed in Appendix A.6.

Changing the Release Cutoff As discussed in the main results, the cutoff used for determining the treatment is 7 days. In order to test whether the results are driven by this arbitrary cutoff, I redefine the cutoff as 14 days and 3 days (as in Dobbie, Goldin, and Yang (2018) and Stevenson (2018)), and results are largely unaffected.

Felony Cases From Start In the main estimation, I include both felony and non-felony

Branch 1 cases in constructing the judge instruments as that ensures all the cases the judge sees contribute to the instrument, while the results are for the felony sample only. I re-compute the instruments using only felony cases, and these results are generally consistent with the main results.

Including D-EM Bonds and Excluding Missing EM D-EM bonds (where the defendant pays the bond amount to be released from jail to EM) comprise a small portion of observations that were dropped in the felony sample. To see if this is changing the results, I re-add them to the sample and re-compute MTEs. Also, as discussed in the data section, EM bonds with missing durations are coded as EM. To see if that is influencing the results, I drop all missing duration EM bond observations. In both cases, the results do not deviate significantly from the main results.

Judge Instrument Only The main instruments rely on variation in judge $\times$ observables, but the random assignment is between cases and judges, conditional on time fixed effects. Relying instead only on judge fixed effects ("$Z(fe)$"), which results in a weaker instrument with smaller common support, produces results that are generally consistent with the main specification except with larger effect sizes — though occasionally the estimates are highly different (e.g., for FTAs and likelihood of incarceration for EM versus detention).

Full Support Parameters The treatment parameters were constructed over the region of common support, but we can compare these estimates to those using full support (e.g., an ATE rather than a CATE) using a sieve-style polynomial parametric assumption and the fully parametric 'Normal' method. These full support ATE, ATD, and ATRs generally agree with the main results, with disagreements for outcomes with weak reduced form relationships or when the Normal method extrapolates monotonically due to its parametric assumptions (e.g., FTAs for EM versus release).

Using Disposition Codes As mentioned in Section 2.1, defendants who receive EM bonds may not actually be admitted into the EM system. To determine if this potential miscoding of treatments will influence the results, I reclassify defendants' EM status based on disposition codes observed in their case history. I construct 4 alternative codings which allow for additional conditions under which the defendant is classified as on EM based on disposition codes. The results are displayed in Figure 13, and they are broadly consistent with the main results.

Re-coding Missing Treatments As discussed in Section 2.3, jail data is unmatched to cases for a subset of the sample. In order to test the sensitivity of my results to how these cases with missing jail data were coded or dropped, I redo the analysis with 4 additional samples in which all cases with missing jail data are kept, usual filters applied, and the entire analysis is redone (e.g., π's are re-computed for each sub-sample).[55] The 4 cases are combinations of coding all missing (jail data) EM bonds as EM or release and coding all missing D-bonds as detention or release. Overall, as shown in Figure 14, these re-codings do not alter the main interpretation of the main findings.

7 Conclusion

This paper explores the effect of pretrial electronic monitoring, an individualized surveillance technology, on defendant outcomes relative to both release and detention across defendant types. In this effort, I develop a general method for semiparametrically estimating marginal treatment effects in ordered treatment environments. Compared with detention, I find that EM allows for more low-level pretrial misconduct, but EM does reduce costly pretrial crime and leads to weakly better case outcomes and fewer high-cost crimes in the future. Compared with release, I find that EM prevents low-level pretrial crime and reduces failures to appear in court, but I cannot reject a small cost-weighted crime reduction effect. Overall, the evidence suggests that EM is a viable alternative to detention for higher-level defendants, but it is not clear that its potential benefits outweigh its costs relative to release for lower-level defendants.

Crucially, I find that EM causes an increase in new cases pretrial for bond violations compared with both release and detention, which highlights an additional cost of using

[55] To reduce miscoding, I remove all D-EM cases prior to re-estimating the π's as well.

imperfect monitoring to enforce compliance. Based on testimonials and research in Cook County and other municipalities using EM, the rules and requirements of EM are often too stringent and inflexible, potentially resulting in worse economic, social, and health outcomes (CGL and Appleseed (2022)). Compliance with the threat of significant punishment involves little leniency for mistakes, legitimate emergencies, big life events (e.g., moving apartments), or daily tasks (e.g., doing laundry, picking kids up at school) (Green (2016), Hager (2020), Weisburd et al. (2021), Johnson (2022)). One improvement could be the relaxation of punishments for violations of EM and increased discretion and leniency, which would be consistent with not only the Becker (1968) model of optimal deterrence (where increasing detection can be met with a decrease in punishment) but also with successful "swift-and-certain" sanction regimes (Hawken and Kleiman (2009), Kilmer et al. (2013)).[56]

Finally, this paper calls for an increased understanding of the costs and benefits of modern surveillance technology within the criminal justice system and more broadly within economic spheres. The existing work on surveillance technology in economics focuses largely on mass surveillance and uses the lens of innovation, political economy, and social control (Tirole (2021), Beraja, Yang, and Yuchtman (2020), Beraja et al. (2021), Acemoglu (2021)). In contrast, the results of this study are consistent with the criticisms of EM and similar *individualized* surveillance technologies in use across the United States (Alexander (2018), Weisburd (2022)): defendants cannot successfully adjust to the level of monitoring, and this leads to increased criminalization and worse outcomes. Yet, this work is only a first step in understanding whether surveillance technologies are desirable in the criminal justice system.

[56]Furthermore, the results emphasize that the large costs are driven solely by outlier individuals charged with murder. This indicates that targeting resources and monitoring toward individuals who are at risk for committing murder could reduce serious crime while also reducing the scope of coercive policy on most defendants.

8 References

Abadie, Alberto. 2003. "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics* 113 (2): 231–63.

Abid, Ali. 2014. "Taskforce, Including Preckwinkle, Evans, and Alvarez, Visits Washington d.c. To Review Pretrial System; and a Warning about the Expanded Use of Electronic Monitoring." *Chicago Appleseed Center for Fair Courts*. https://www.chicagoappleseed.org/2014/07/22/bond-court-reform-update-taskforce-including-preckwinkle-evans-and-alvarez-visits-washington-d-c-to-review-pretrial-system-and-a-warning-about-the-expanded-use-of-electronic-monitoring/.

Acemoglu, Daron. 2021. "Harms of AI." Working Paper. Working Paper Series. National Bureau of Economic Research. http://www.nber.org/papers/w29247.

Afeef, Junaid, Lindsay Bostwick, Simeon Kim, and Jessica Reichert. 2012. "Policies and Procedures of the Criminal Justice System |Office of Justice Programs." 247298. Illinois Criminal Justice Information Authority.

Agan, Amanda Y., Jennifer L. Doleac, and Anna Harvey. 2021. "Misdemeanor Prosecution." w28600. National Bureau of Economic Research.

Aizer, Anna, and Joseph J. Doyle. 2015. "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges." *The Quarterly Journal of Economics* 130 (2): 759–803.

Albright, Alex. 2021. "No Money Bail, No Problems?"

Alchian, Armen A., and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization." *The American Economic Review* 62 (5): 777–95.

Alexander, Michelle. 2018. "The Newest Jim Crow." *The New York Times*. https://www.nytimes.com/2018/11/08/opinion/sunday/criminal-justice-reforms-race-technology.html.

Andersen, Lars H., and Signe H. Andersen. 2014. "Effect of Electronic Monitoring on Social Welfare Dependence." *Criminology & Public Policy* 13 (3): 349–79.

Andresen, Martin Eckhoff. 2018. "Exploring Marginal Treatment Effects: Flexible Estimation

Using Stata." *The Stata Journal: Promoting Communications on Statistics and Stata* 18 (1): 118–58.

Angrist, J. D., G. W. Imbens, and A. B. Krueger. 1999. "Jackknife Instrumental Variables Estimation." *Journal of Applied Econometrics* 14 (1): 57–67.

APPA. 2020. "APPA Response to Criticisms of Electronic Monitoring." American Probation; Parole Association. https://www.appa-net.org/eWeb/docs/APPA/announce/2020-03-18-EM-Criticisms-Response.pdf.

Arbour, William, and Steeve Marchand. 2022. "Parole, Recidivism, and the Role of Supervised Transition." Working Paper.

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

Arnold, David, Will Dobbie, and Crystal S Yang. 2018. "Racial Bias in Bail Decisions." *The Quarterly Journal of Economics* 133 (4): 1885–1932.

Arteaga, Carolina. 2021. "Parental Incarceration and Children's Educational Attainment." Working Paper.

Barajas, Michael. 2020. "As COVID-19 Cases Surge, Some Incarcerated People Remain Behind Bars After Making Bail." *The Texas Observer*. https://www.texasobserver.org/pretrial-covid-dallas-electronic-monitoring/.

Barbaro, Michael, Rikki Novetsky, Michael Simon Johnson, Mooj Zadie, Liz O. Baylen, Paige Cowett, Brad Fisher, Dan Powell, Marion Lozano, and Elisheba Ittoop. 2022. "The Rise of Workplace Surveillance." *The New York Times*, August.

Bar-Ilan, Avner, and Bruce Sacerdote. 2001. "The Response to Fines and Probability of Detection in a Series of Experiments." Working Paper 8638. National Bureau of Economic Research.

Bayer, Patrick, Randi Hjalmarsson, and David Pozen. 2009. "Building Criminal Capital Behind Bars: Peer Effects in Juvenile Corrections." *The Quarterly Journal of Economics* 124 (1): 105–47.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (2): 169–217.

Becker, Gary S., and George J. Stigler. 1974. "Law Enforcement, Malfeasance, and Compensation of Enforcers." *The Journal of Legal Studies* 3 (1): 1–18.

Belur, Jyoti, Amy Thornton, Lisa Tompson, Matthew Manning, Aiden Sidebottom, and Kate Bowers. 2020. "A Systematic Review of the Effectiveness of the Electronic Monitoring of Offenders." *Journal of Criminal Justice* 68 (May): 101686.

Benjamin, Ruha. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Cambridge, UK ; Medford, MA: Polity.

Beraja, Martin, Andrew Kao, David Y. Yang, and Noam Yuchtman. 2021. "AI-Tocracy." Working Paper 29466. National Bureau of Economic Research.

Beraja, Martin, David Y. Yang, and Noam Yuchtman. 2020. "Data-Intensive Innovation and the State: Evidence from AI Firms in China." Working Paper 27723. National Bureau of Economic Research.

Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Løken, and Magne Mogstad. 2020. "Incarceration, Recidivism, and Employment." *Journal of Political Economy* 128 (4): 1269–1324.

Bhuller, Manudeep, Gordon B Dahl, Katrine V Løken, and Magne Mogstad. 2018. "Incarceration Spillovers in Criminal and Family Networks." Working Paper 24878. National Bureau of Economic Research.

Bhuller, Manudeep, and Henrik Sigstad. 2022. "2SLS with Multiple Treatments." arXiv. http://arxiv.org/abs/2205.07836.

Bjerk, David. 2007. "Racial Profiling, Statistical Discrimination, and the Effect of a Colorblind Policy on the Crime Rate." *Journal of Public Economic Theory* 9 (3): 521–45.

Björklund, Anders, and Robert Moffitt. 1987. "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models." *The Review of Economics and Statistics* 69 (1): 42–49.

Brinch, Christian N., Magne Mogstad, and Matthew Wiswall. 2017. "Beyond LATE with a Discrete Instrument." *Journal of Political Economy* 125 (4): 985–1039.

Carneiro, Pedro, James Heckman, and Edward Vytlacil. 2011. “Estimating Marginal Returns to Education.” *American Economic Review* 101 (6): 2754–81.

Carneiro, Pedro, and Sokbae Lee. 2009. “Estimating Distributions of Potential Outcomes Using Local Instrumental Variables with an Application to Changes in College Enrollment and Wage Inequality.” *Journal of Econometrics* 149 (2): 191–208.

CGL, and Chicago Appleseed. 2022. “Electronic Monitoring Review Cook County, Illinois Final Report.” Report. Cook County, IL.

Chalfin, Aaron, and Justin McCrary. 2017. “Criminal Deterrence: A Review of the Literature.” *Journal of Economic Literature* 55 (1): 5–48.

Cook, Philip J. 1979. “The Clearance Rate as a Measure of Criminal Justice System Effectiveness.” *Journal of Public Economics* 11 (1): 135–42.

Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg. 2016. “From LATE to MTE: Alternative Methods for the Evaluation of Policy Interventions.” *Labour Economics*, SOLE/EALE conference issue 2015, 41 (August): 47–60.

———. 2018. “Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance.” *Journal of Political Economy* 126 (6): 2356–2409.

Dahl, Gordon B. 2002. “Mobility and the Return to Education: Testing a Roy Model with Multiple Markets.” *Econometrica* 70 (6): 2367–2420.

DeMichele, Matthew. 2014. “Electronic Monitoring: It Is a Tool, Not a Silver Bullet Electronic Monitoring on Social Welfare Dependence: Policy Essay.” *Criminology and Public Policy* 13 (3): 393–400.

Di Tella, Rafael, and Ernesto Schargrodsky. 2013. “Criminal Recidivism After Prison and Electronic Monitoring.” *Journal of Political Economy* 121 (1): 28–73.

Dizikes, Cynthia, and Todd Lightly. 2015. “Electronic Monitoring Spikes in Cook County.” *Chicago Tribune*, February.

Dobbie, Will, Jacob Goldin, and Crystal S. Yang. 2018. “The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges.”

American Economic Review 108 (2): 201–40.

Dobbie, Will, and Jae Song. 2015. “Debt Relief and Debtor Outcomes: Measuring the Effects of Consumer Bankruptcy Protection.” *American Economic Review* 105 (3): 1272–1311.

Dobbie, Will, and Crystal S. Yang. 2021. “The US Pretrial System: Balancing Individual Rights and Public Interests.” *Journal of Economic Perspectives* 35 (4): 49–70.

Dominitz, Jeff, and John Knowles. 2006. “Crime Minimisation and Racial Bias: What Can We Learn from Police Search Data?” *The Economic Journal* 116 (515).

Donohue, John J. 2007. “Economic Models of Crime and Punishment.” *Social Research* 74 (2): 379–412.

Doyle Jr., Joseph J. 2008. “Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care.” *Journal of Political Economy* 116 (4): 746–70.

Durlauf, Steven N., and Daniel S. Nagin. 2011. “Imprisonment and Crime.” *Criminology & Public Policy* 10 (1): 13–54.

Ehrlich, Isaac. 1996. “Crime, Punishment, and the Market for Offenses.” *Journal of Economic Perspectives* 10 (1): 43–67.

Federation, The Civic. 2017. “The Impact of Cook County Bond Court on the Jail Population: A Call for Increased Public Data and Analysis.” Chicago: The Civic Federation.

———. 2020. “Cook County Seeks Consulting Services to Review Electronic Monitoring Practices.” *The Civic Federation*. https://www.civicfed.org/civic-federation/blog/cook-county-seeks-consulting-services-review-electronic-monitoring-practices.

Feigenberg, Benjamin, and Conrad Miller. 2022. “Would Eliminating Racial Disparities in Motor Vehicle Searches Have Efficiency Costs?*.” *The Quarterly Journal of Economics* 137 (1): 49–113.

Fishbane, Alissa, Aurelie Ouss, and Anuj K. Shah. 2020. “Behavioral Nudges Reduce Failure to Appear for Court.” *Science* 370 (6517).

Frandsen, Brigham, Lars Lefgren, and Emily Leslie. 2019. “Judging Judge Fixed Effects.” w25528. Cambridge, MA: National Bureau of Economic Research.

Gaure, Simen. 2013. "Lfe: Linear Group Fixed Effects." *The R Journal* 5 (2): 104.

Goldsmith-Pinkham, Paul, Peter Hull, and Michal Kolesár. 2022. "Contamination Bias in Linear Regressions." Working Paper.

Goncalves, Felipe, and Steven Mello. 2022. "Should the Punishment Fit the Crime? Deterrence and Retribution in Law Enforcement." Working Paper.

Green, Larry. 2016. "Home Is No Castle for Some Cook County Defendants; It's Jail." *Injustice Watch*, November.

Grenet, Julien, Hans Grönqvist, and Susan Niknami. 2022. "The Effects of Electronic Monitoring on Offenders and Their Families." Working Paper.

Grogger, Jeffrey. 1991. "Certainty Vs. Severity of Punishment." *Economic Inquiry* 29 (2): 297–309.

Gross, Max, and E. Jason Baron. 2022. "Temporary Stays and Persistent Gains: The Causal Effects of Foster Care." *American Economic Journal: Applied Economics* 14 (2): 170–99.

Grossman, Sanford J., and Oliver D. Hart. 1983. "An Analysis of the Principal-Agent Problem." *Econometrica* 51 (1): 7–45.

Guevara, Marydell. 2014. "Miami-Dade Corrections and Rehabilitation Department, Biannual Report 2013 and 2014." Biannual Report. Miami-Dade Corrections; Rehabilitation Department.

Gupta, Arpit, Christopher Hansman, and Ethan Frenchman. 2016. "The Heavy Costs of High Bail: Evidence from Judge Randomization." *Journal of Legal Studies* 45 (2): 471–505.

Hager, Eli. 2020. "Where Coronavirus Is Surging - and Electronic Surveillance, Too." *The Marshall Project*. https://www.themarshallproject.org/2020/11/22/where-coronavirus-is-surging-and-electronic-surveillance-too.

Hawken, Angela, and Mark Kleiman. 2009. "Managing Drug Involved Probationers with Swift and Certain Sanctions: Evaluating Hawaii's HOPE: (513502010-001)." American Psychological Association.

Heckman, James, Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz.

2010. "The Rate of Return to the HighScope Perry Preschool Program." *Journal of Public Economics* 94 (1): 114–28.

Heckman, James, Sergio Urzua, and Edward Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." Working Paper 12574. National Bureau of Economic Research.

———. 2008. "Instrumental Variables in Models with Multiple Outcomes: The General Unordered Case." *Annales d'Économie Et de Statistique*, no. 91/92: 151–74.

Heckman, James, and Edward Vytlacil. 1999. "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." *Proceedings of the National Academy of Sciences* 96 (8): 4730–34.

———. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica* 73 (3): 669–738.

———. 2007. "Chapter 71 Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments." In *Handbook of Econometrics*, edited by James Heckman and Edward E. Leamer, 6:4875–5143. Elsevier.

Henneguelle, Anaïs, Benjamin Monnery, and Annie Kensey. 2016. "Better at Home Than in Prison? The Effects of Electronic Monitoring on Recidivism in France." *The Journal of Law and Economics* 59 (3): 629–67.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association*, 27.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, & Organization* 7: 24–52.

Holmström, Bengt. 1979. "Moral Hazard and Observability." *The Bell Journal of Economics* 10 (1): 74–91.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local

Average Treatment Effects." *Econometrica* 62 (2): 467–75.

———. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association* 90 (430): 431–42.

Imbens, Guido W., and Donald B. Rubin. 1997. "Estimating Outcome Distributions for Compliers in Instrumental Variables Models." *The Review of Economic Studies* 64 (4): 555–74.

Institute, Vera. 2022. "What Jails Cost: Cities Chicago, IL." *Vera Institute of Justice*. https://www.vera.org/publications/what-jails-cost-cities/chicago-il.

Jensen, Nathan, Elizabeth Lyons, Eddy Chebelyon, and Ronan Le Bras. 2020. "Conspicuous Monitoring and Remote Work." *Journal of Economic Behavior & Organization* 176 (August): 489–511.

Johnson, Jeremey. 2022. "June 16 2022 10:54 Am i Was Called with a False Violation in the Middle of an Interview." https://www.youtube.com/watch?v=vodHTbo2rWI.

Jordan, Andrew, Ezra Karger, and Derek A. Neal. 2022. "Heterogeneous Impacts of Sentencing Decisions." SSRN Scholarly Paper. Rochester, NY. https://papers.ssrn.com/abstract=3927995.

Kilgore, James. 2013. "Progress or More of the Same? Electronic Monitoring and Parole in the Age of Mass Incarceration." *Critical Criminology* 21 (1): 123–39.

Kilmer, Beau, Nancy Nicosia, Paul Heaton, and Greg Midgette. 2013. "Efficacy of Frequent Monitoring with Swift, Certain, and Modest Sanctions for Violations: Insights from South Dakota's 24/7 Sobriety Project." *American Journal of Public Health* 103 (1).

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. "Field of Study, Earnings, and Self-Selection." *The Quarterly Journal of Economics* 131 (3): 1057–1111.

Kleiman, Mark A. R. 2009. *When Brute Force Fails: How to Have Less Crime and Less Punishment*. Princeton University Press.

Kline, Patrick, and Christopher R. Walters. 2016. "Evaluating Public Programs with Close

Substitutes: The Case of Head Start." *The Quarterly Journal of Economics* 131 (4): 1795–1848.

Kling, Jeffrey R. 2006. "Incarceration Length, Employment, and Earnings." *American Economic Review* 96 (3): 863–76.

Kohler-Hausmann, Issa. 2013. "Misdemeanor Justice: Control Without Conviction." *American Journal of Sociology* 119 (2): 351–93.

Kuziemko, Ilyana. 2013. "How Should Inmates Be Released from Prison? An Assessment of Parole Versus Fixed-Sentence Regimes *." *The Quarterly Journal of Economics* 128 (1): 371–424.

LaForest, Michael. 2021. "Early Release and Parole Conditions at the Margin." Working Paper.

Lee, Sokbae, and Bernard Salanié. 2018. "Identifying Effects of Multivalued Treatments." *Econometrica* 86 (6): 1939–63.

Leifeld, Philip. 2013. "Texreg: Conversion of Statistical Model Output in r to LATEX and HTML Tables." *Journal of Statistical Software* 55 (November): 1–24.

Leslie, Emily, and Nolan G. Pope. 2017. "The Unintended Impact of Pretrial Detention on Case Outcomes: Evidence from New York City Arraignments." *The Journal of Law and Economics* 60 (3): 529–57.

Lockhart, P. R. 2019. "Thousands of Americans Are Jailed Before Trial. A New Report Shows the Lasting Impact." *Vox.* https://www.vox.com/2019/5/7/18527237/pretrial-detention-jail-bail-reform-vera-institute-report.

Lum, Cynthia, Christopher S. Koper, David B. Wilson, Megan Stoltz, Michael Goodier, Elizabeth Eggins, Angela Higginson, and Lorraine Mazerolle. 2020. "Body-Worn Cameras' Effects on Police Officers and Citizen Behavior: A Systematic Review." *Campbell Systematic Reviews* 16 (3).

Lum, Cynthia, and Daniel S. Nagin. 2017. "Reinventing American Policing." *Crime and Justice* 46 (1): 339–93.

Mai, Chris, and Ram Subramanian. 2017. "The Price of Prisons: Examining State Spending Trends, 2010-2015." Vera Institute for Justice.

Manski, Charles F. 2005. "Optimal Search Profiling with Linear Deterrence." *The American Economic Review* 95 (2): 122–26.

Marie, Olivier. 2008. "Early Release from Prison and Recidivism: A Regression Discontinuity Approach." Working Paper.

Mayson, Sandra G., and Megan Stevenson. 2020. "Misdemeanors by the Numbers." *Boston College Law Review* 61 (3): 971–1044.

McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, 105–42. New York: Academic Press.

Mello, Steven. 2019. "More COPS, Less Crime." *Journal of Public Economics* 172 (April): 174–200.

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

Mirrlees, James A. 1976. "The Optimal Structure of Incentives and Authority Within an Organization." *The Bell Journal of Economics* 7 (1): 105–31.

Moffitt. 2008. "Estimating Marginal Treatment Effects in Heterogeneous Populations." *Annales d'Économie Et de Statistique*, no. 91/92: 239.

Mogstad, Magne, Andres Santos, and Alexander Torgovitsky. 2018. "Using Instrumental Variables for Inference about Policy Relevant Treatment Parameters." *Econometrica* 86 (5): 1589–619.

Mogstad, Magne, and Alexander Torgovitsky. 2018. "Identification and Extrapolation of Causal Effects with Instrumental Variables." *Annual Review of Economics* 10 (1): 577–613.

Montiel Olea, José Luis, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics* 31 (3): 358–69.

Mountjoy, Jack. 2022. "Community Colleges and Upward Mobility." *American Economic*

Review 112 (8): 2580–630.

Mueller-Smith, Michael. 2015. "The Criminal and Labor Market Impacts of Incarceration." Working Paper.

Mungan, Murat C. 2017. "The Certainty Versus the Severity of Punishment, Repeat Offenders, and Stigmatization." *Economics Letters* 150 (January): 126–29.

Myers, Samuel L. 1981. "The Economics of Bail Jumping." *The Journal of Legal Studies* 10 (2): 381–96.

Nagin, Daniel S. 2013. "Deterrence: A Review of the Evidence by a Criminologist for Economists." *Annual Review of Economics* 5 (1): 83–105.

Norris, Samuel, Matthew Pecenco, and Jeffrey Weaver. 2021. "The Effects of Parental and Sibling Incarceration: Evidence from Ohio." *American Economic Review* 111 (9): 2926–63.

Office, Cook County Sheriff's. 2020. "Cook County Sheriff's Office Community Corrections - Electronic Monitoring (EM) Program (GPS) Information Sheet." Cook County Sheriff's Office.

Ouss, Aurelie. 2013. "Sensitivity Analyses in Economics of Crime: Do Monitored Suspended Sentences Reduce Recidivism?" Working Paper.

Ouss, Aurelie, and Megan Stevenson. 2022. "Does Cash Bail Deter Misconduct?" SSRN Scholarly Paper. Rochester, NY. https://papers.ssrn.com/abstract=3335138.

Pager, Devah, Rebecca Goldstein, Helen Ho, and Bruce Western. 2021. "Criminalizing Poverty: The Consequences of Court Fees in a Randomized Experiment."

Persico, Nicola. 2002. "Racial Profiling, Fairness, and Effectiveness of Policing." *American Economic Review* 92 (5): 1472–97.

Pew. 2016. "Use of Electronic Offender-Tracking Devices Expands Sharply." http://pew.org/2cpDaNx.

Prendergast, Canice. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature* 37 (1): 7–63.

PRI. 2016. "Ten-Point Plan on Reducing Pre-Trial Detention." Penal Reform International.

Rogers, Phil. 2022. "More Than 80 Guns Found in Homes of Individuals on Electronic Monitoring This Year, Cook County Sheriff's Office Says." *NBC Chicago*. https://www.nbcchicago.com/investigations/more-than-80-guns-found-in-homes-of-individuals-on-electronic-monitoring-this-year-cook-county-sheriffs-office-says/2856573/.

Rose, Evan K. 2021. "Who Gets a Second Chance? Effectiveness and Equity in Supervision of Criminal Offenders*." *The Quarterly Journal of Economics* 136 (2): 1199–1253.

Rose, Evan K, and Yotam Shem-Tov. 2021. "How Does Incarceration Affect Reoffending? Estimating the Dose-Response Function." *Journal of Political Economy*, July.

Rosenfeld, Richard, and Amanda Grigg, eds. 2022. *The Limits of Recidivism: Measuring Success After Prison*. Washington, DC: The National Academies of Sciences, Engineering,; Medicine.

Roy, A. D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (2): 135–46.

Sanderson, Eleanor, and Frank Windmeijer. 2016. "A Weak Instrument f-Test in Linear IV Models with Multiple Endogenous Variables." *Journal of Econometrics*, Endogeneity Problems in Econometrics, 190 (2): 212–21.

Sawyer, Wendy, and Peter Wagner. 2022. "Mass Incarceration: The Whole Pie 2022." Press Release. Prison Policy Initiative.

Shapiro, Carl, and Joseph E. Stiglitz. 1984. "Equilibrium Unemployment as a Worker Discipline Device." *The American Economic Review* 74 (3): 433–44.

Sheriff, Cook County. 2020. "Sheriff's Office Announces Electronic Monitoring Program Transition from Radio Frequency to GPS Bracelets." *Cook County Sheriff's Office*. https://www.cookcountysheriff.org/sheriffs-office-announces-electronic-monitoring-program-transition-from-radio-frequency-to-gps-bracelets/.

Silveira, Bernardo S. 2017. "Bargaining with Asymmetric Information: An Empirical Study of Plea Negotiations." *Econometrica* 85 (2): 419–52.

Stevenson, Megan. 2017. "Breaking Bad: Mechanisms of Social Influence and the Path to Criminality in Juvenile Jails." *The Review of Economics and Statistics* 99 (5): 824–38.

———. 2018. "Distortion of Justice: How the Inability to Pay Bail Affects Case Outcomes." *The Journal of Law, Economics, and Organization* 34 (4): 511–42.

Tardy, Michael J, Margie Groot, Rich Adkins, Monica Allen, Gregg Anderson, Amy Bowne, Tom Doyle, et al. 2014. "Administrative Office of the Illinois Courts Pretrial Operational Review Team." Illinois Supreme Court Administrative Office of the Illinois Courts.

Tirole, Jean. 2021. "Digital Dystopia." *American Economic Review* 111 (6): 2007–48.

Virani, Alicia. 2022. "Pretrial Electronic Monitoring in Los Angeles County." Report. Los Angeles: UCLA School of Law.

Vytlacil, Edward. 2002. "Independence, Monotonicity, and Latent Index Models: An Equivalence Result." *Econometrica* 70 (1): 331–41.

———. 2006. "Ordered Discrete-Choice Selection Models and Local Average Treatment Effect Assumptions: Equivalence, Nonequivalence, and Representation Results." *The Review of Economics and Statistics* 88 (3): 578–81.

Wagner, Peter, and Bernadette Rabuy. 2017. "Following the Money of Mass Incarceration." Prison Policy Initiative.

Weisburd, Kate. 2022. "Punitive Surveillance." *Virginia Law Review* 108: 75.

Weisburd, Kate, Varun Bhadha, Matthew Clauson, Jeanmarie Elican, Fatima Khan, Kendall Lawrenz, Brooke Pemberton, et al. 2021. "Electronic Prisons: The Operation of Ankle Monitoring in the Criminal Legal System." Report. The George Washington Law School.

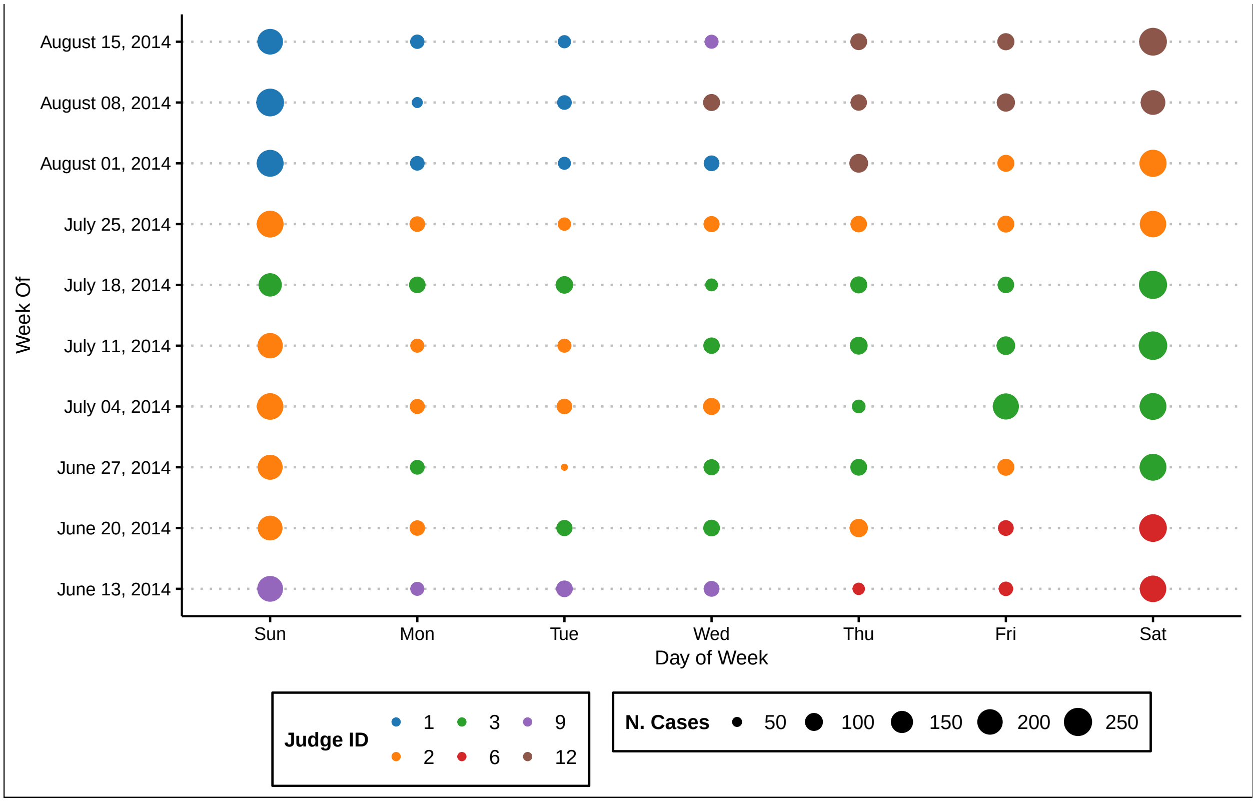
Weisburd, Kate, and Alicia Virani. 2022. "Op-Ed: The Monster of Incarceration Quietly Expands Through Ankle Monitors." *Los Angeles Times*. https://www.latimes.com/opinion/story/2022-03-15/electronic-monitoring-pretrial-bail-mass-incarceration.

Williams, Jenny, and Don Weatherburn. 2022. "Can Electronic Monitoring Reduce Reoffending?" *The Review of Economics and Statistics* 104 (2): 232–45.

Wilson, Stuart John, and Jocelyne Lemoine. 2022. "Methods of Calculating the Marginal

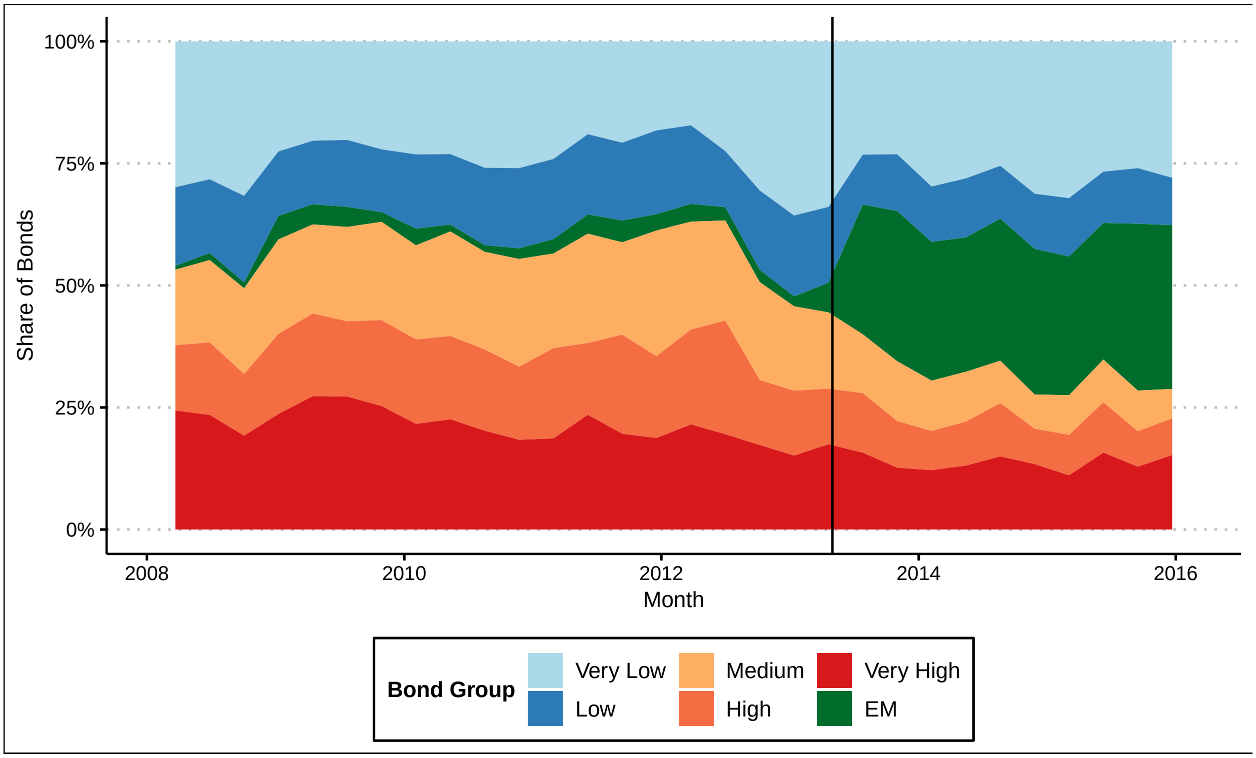
Cost of Incarceration: A Scoping Review." *Criminal Justice Policy Review* 33 (6): 639–63.

Figure 1: Example of Bond Court Judge Rotation Calendar



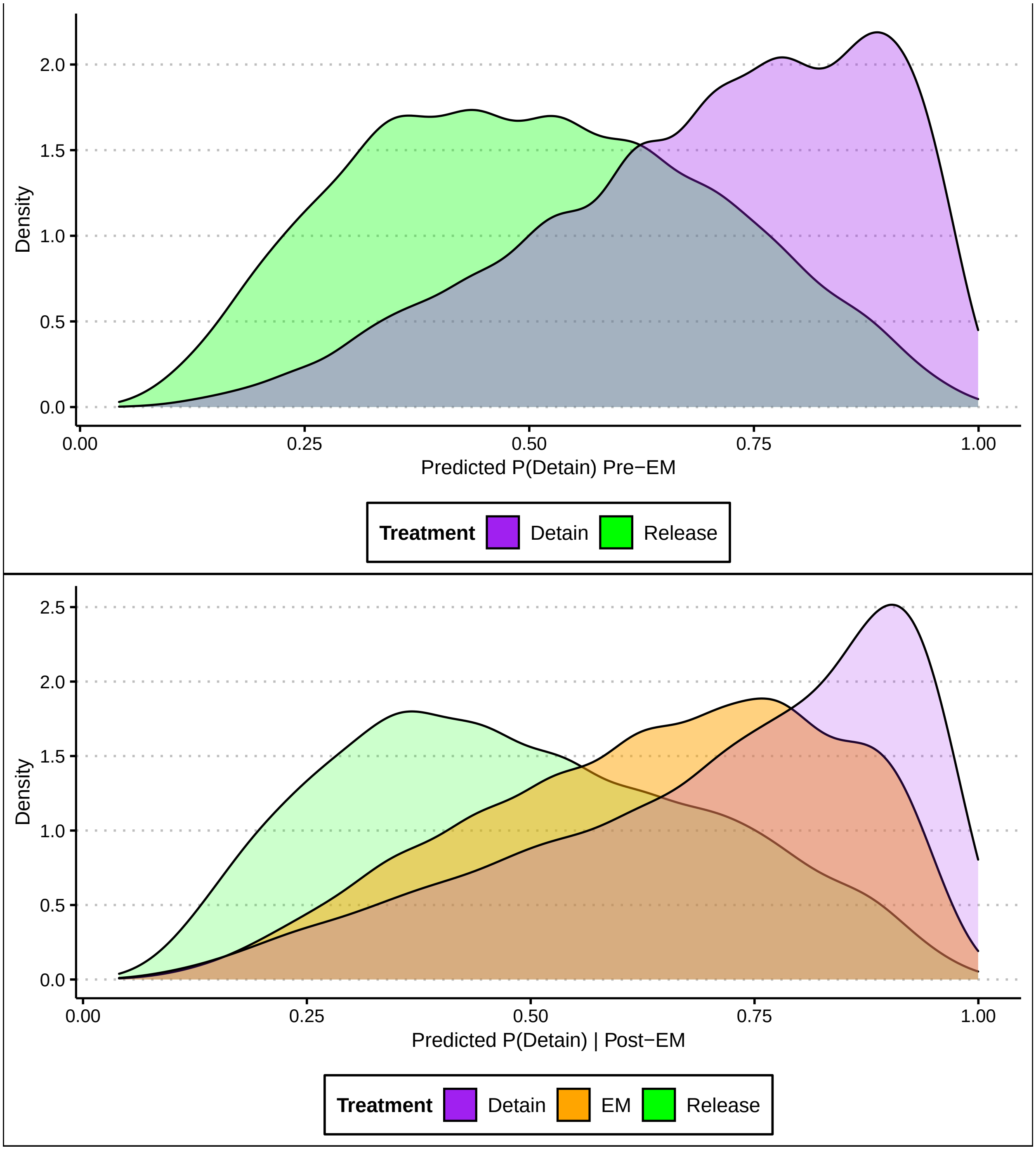
Note: Figure displays the caseloads of active judges in the Cook County Bond Court (Branch 1, Room 100) by week and day of the week between the weeks of June 13, 2014, and August 15, 2014. There were 6 active judges in this period (dot color), while the size of each dot denotes the number of cases they saw that day.

Figure 2: Bond Types Over Time



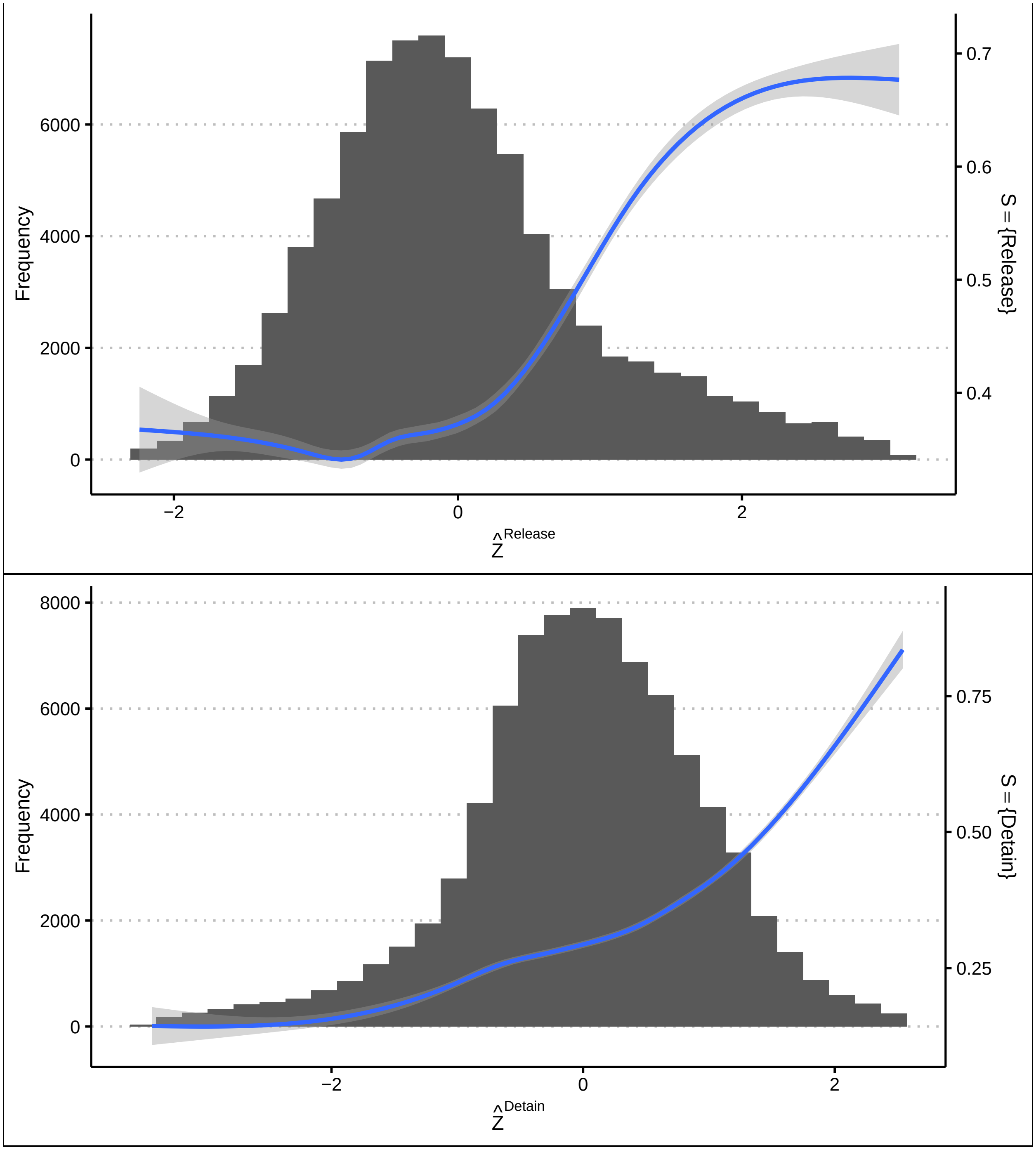
Note: Figure displays the composition of bond types within the sample between March 2008 and 2015 aggregated by the year-month of bond date. Bond group EM denotes IEM bonds, and other bond groups are determined by the bond price required for release (i.e., the bond price for any D-bond, \$0 for I-bonds, and $+\$\infty$ for bond denial). Very low contains bonds with amounts between \$0 and \$7,500; low contains bonds with amounts between \$7,500 to \$20,000; medium contains bond with amounts between \$20,000 to \$40,000; high contains bonds with amounts between \$40,000 to \$60,000; and very high contains all bonds with amounts above \$60,000.

Figure 3: Distribution of Defendants by Treatment, Pre- and Post-IEM



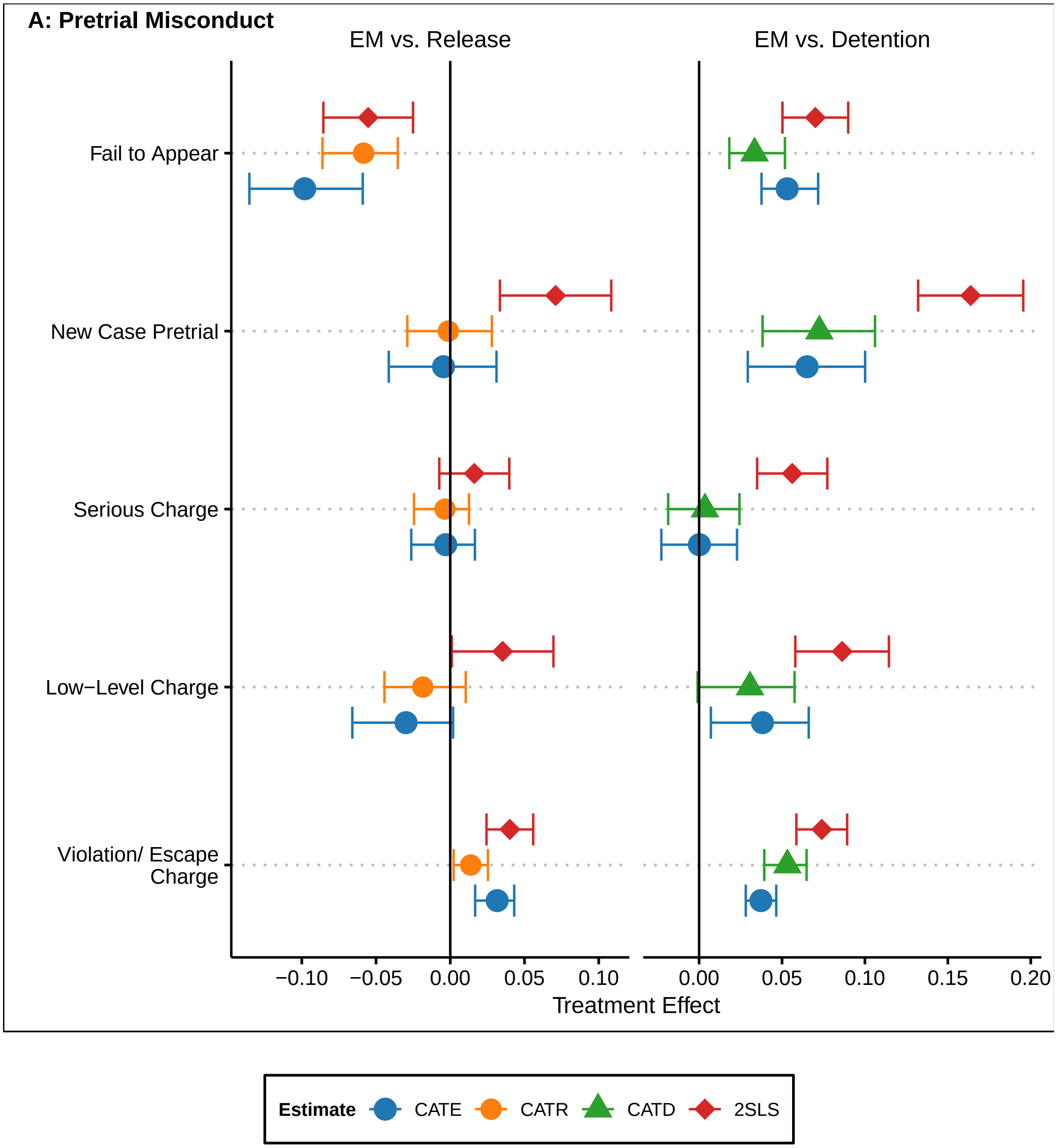
Note: Figures display the distributions of pretrial treatments during the period (Detain, Release) for pre-IEM (2009-2012) and post-IEM (July 2013 - 2015). X-axis is the defendant's predicted likelihood of being detained in the pre-IEM period based on their case observables. Coefficients for predicting likelihood of detention are recovered from regressing detention on defendant observables in the Pre-IEM period, then predicted values are computed using the coefficients on data from the Pre-IEM period (top) and Post-IEM period (bottom).

Figure 4: Distribution of Instrumented Treatment and First Stage



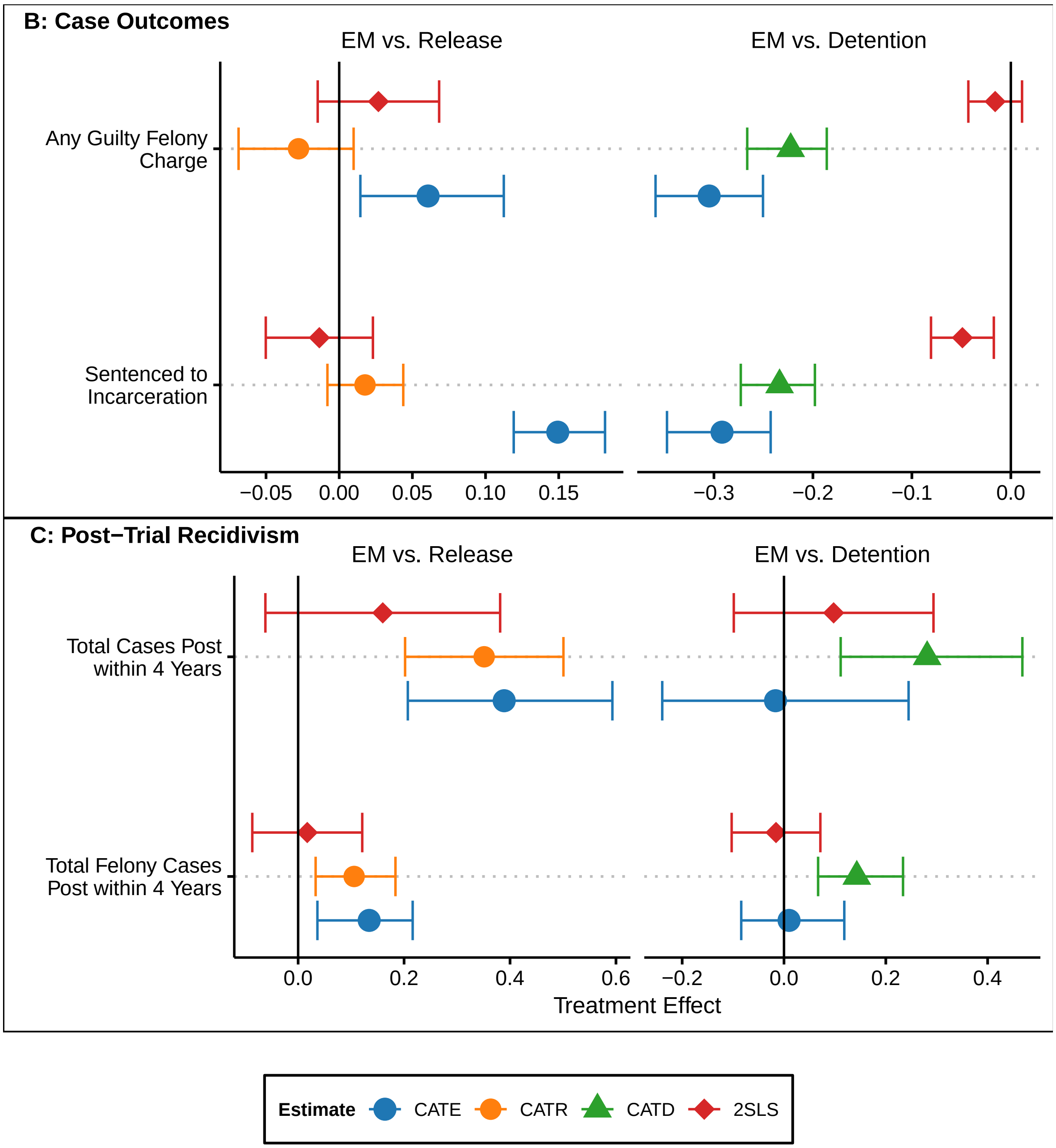
Note: Figure displays the local linear fit and support of the instrumented probabilities of a defendant being assigned to Release (top) and Detention (bottom). Instrumented probabilities are constructed using an LPM and are residualized to remove the linear influence of observables and time fixed effects, and are standardized. The x-axis is the value of the instrument (standardized); the left y-axis is the frequency of the instrument value; the right y-axis is the likelihood of the defendant being placed on the treatment of interest.

Figure 5: Treatment Effect Plots for Pretrial Misconduct



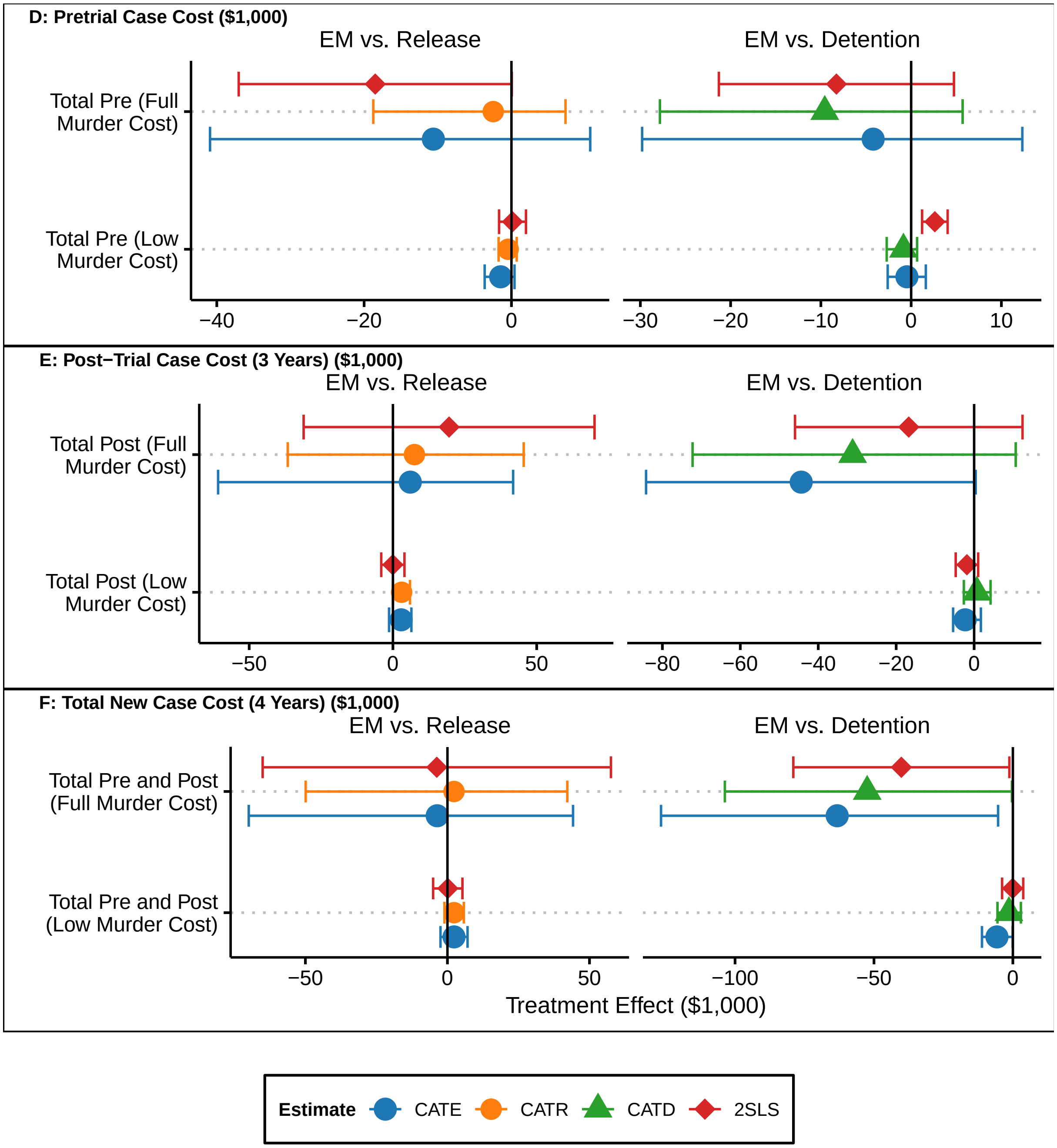
Note: Figure displays treatment effects for EM versus Release and EM versus Detention for main felony sample cases. Outcomes consist of binary variables relating to pretrial outcomes. ATE = average treatment effect, ATR = average treatment effect on the released, ATD = average treatment effect on the detained. (C) denotes parameter constructed using common support with the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs for CATE, CATR, and CATD, and standard errors clustered at the defendant level are used for the 2SLS estimates.

Figure 6: Treatment Effect Plots for Case and Post-Trial Outcomes



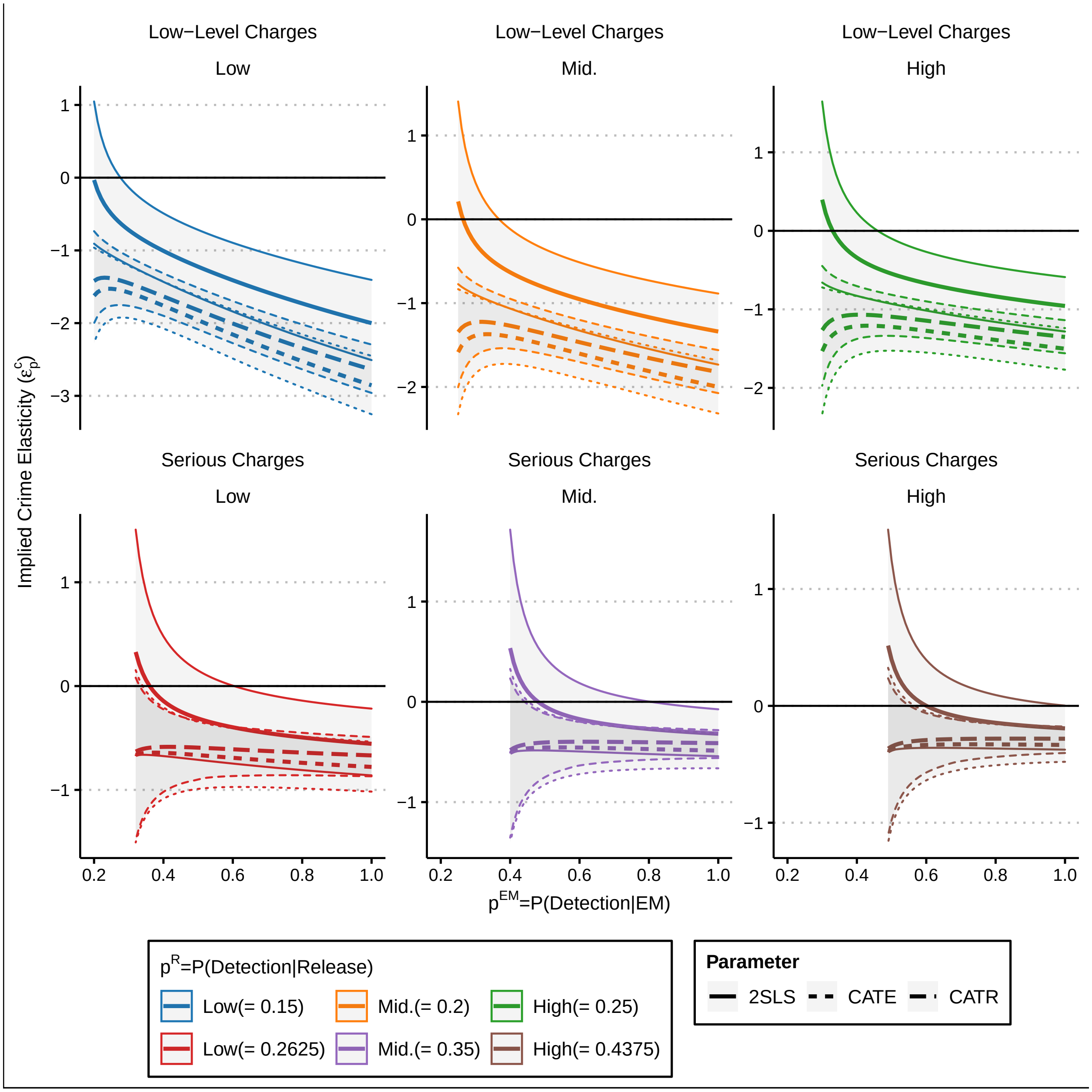
Note: Figure displays treatment effects for EM versus Release and EM versus Detention for main felony sample cases. Outcomes consist of variables relating to case and post-trial outcomes. ATE = average treatment effect, ATR = average treatment effect on the released, ATD = average treatment effect on the detained. (C) denotes parameter constructed using common support with the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs for CATE, CATR, and CATD, and standard errors clustered at the defendant level are used for the 2SLS estimates.

Figure 7: Treatment Effect Plots for New Case Costs



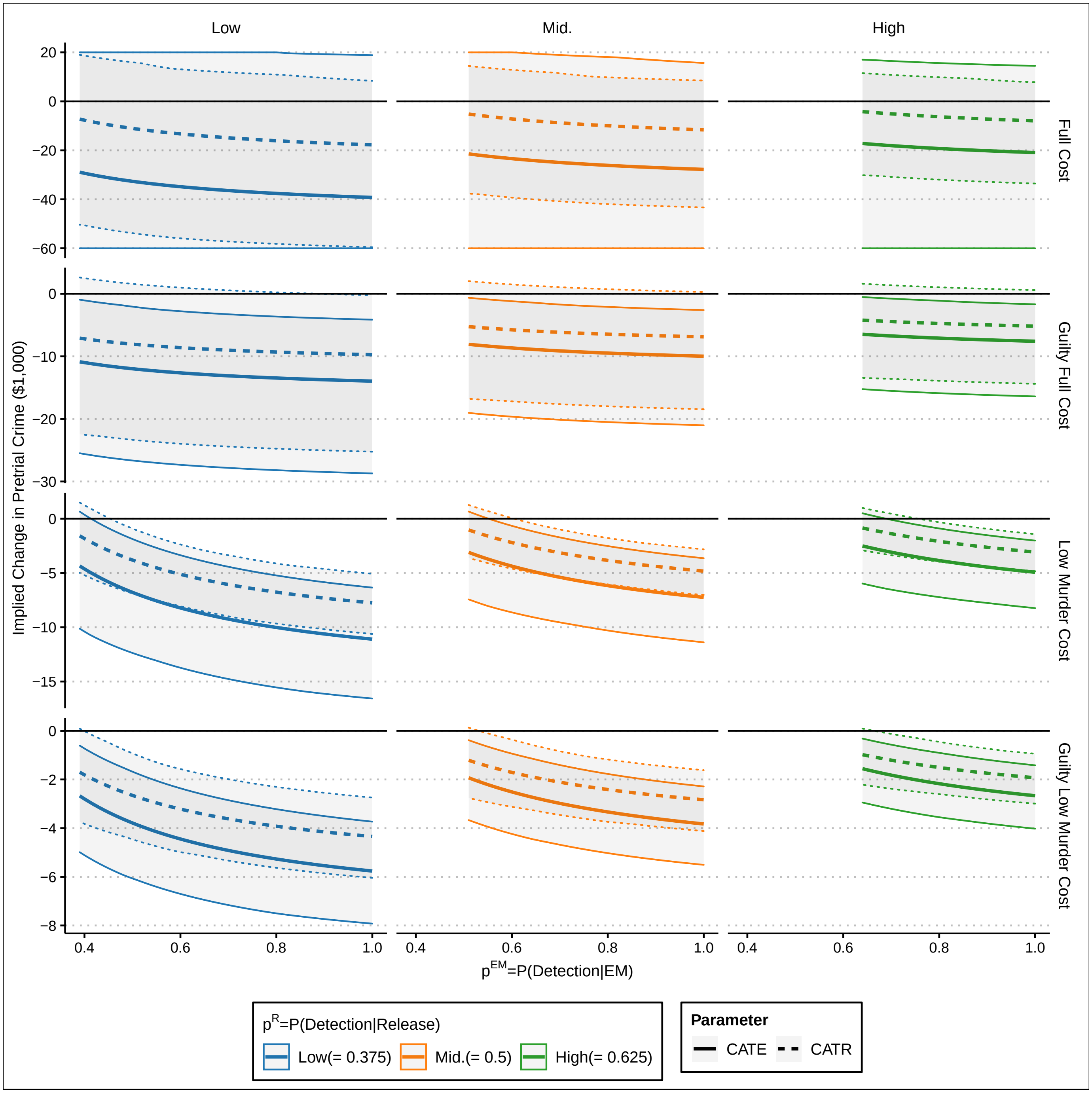
Note: Figure displays treatment effects for EM versus Release and EM versus Detention for main felony sample cases. Outcomes consist of variables pre- and post-trial new cases weighted by their incidence costs cost based on estimates from Miller et al. (2021) (in which the full murder cost is around $8,000,000). Low murder cost uses a value of around $400,000. ATE = average treatment effect, ATR = average treatment effect on the released, ATD = average treatment effect on the detained. (C) denotes parameter constructed using common support with the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs for CATE, CATR, and CATD, and standard errors clustered at the defendant level are used for the 2SLS estimates.

Figure 8: Estimated Bounds for Elasticities



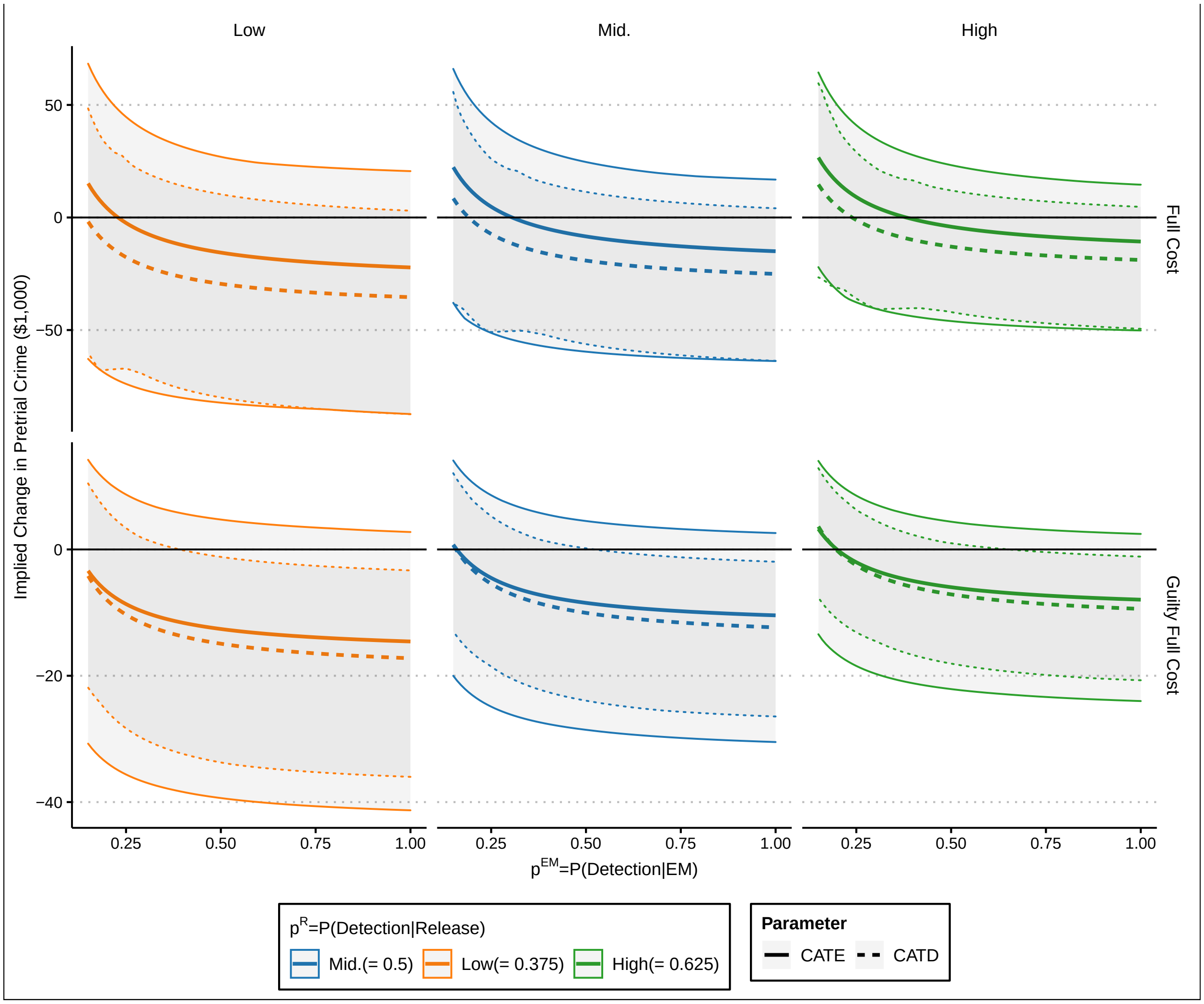
Note: Figure displays implied elasticities for different pretrial criminal charge types using a single value for the initial (release) probability of detection across other potential values for the increased probability of detection under EM. Colors correspond to crime type and probabilities of detection under release. Line types correspond to weights being used are for CATE (even weights across defendants and common support) or CATR (higher weights for released defendants across common support). 95% confidence intervals are computed using 200 bootstrap runs or CATE and CATR, and 95% confidence intervals for 2SLS are based on standard errors clustered at the defendant level.

Figure 9: Estimated Change in Pretrial Crime Cost: EM vs. Release



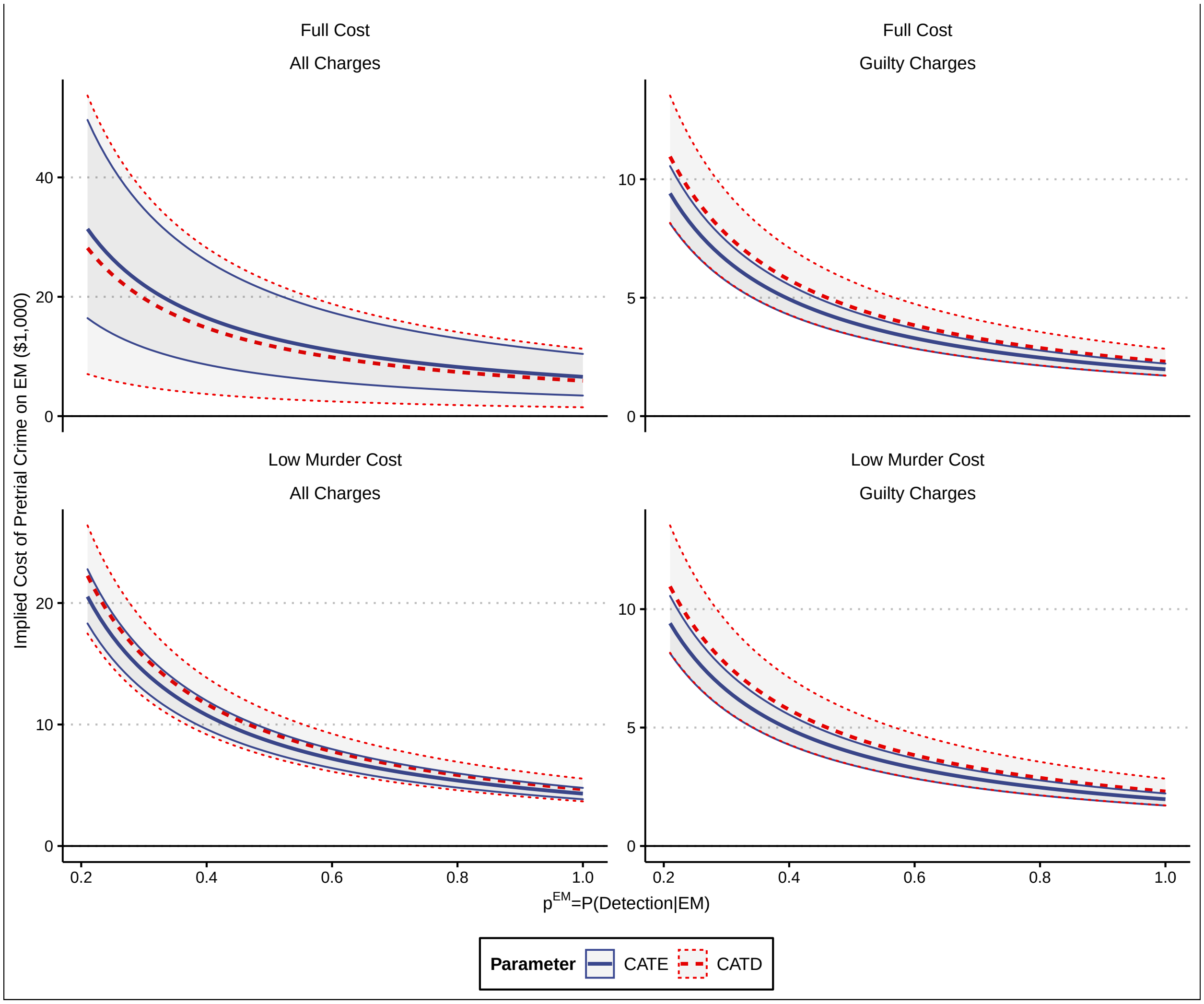
Note: Figure displays implied change in pretrial crime costs using a single value for the initial (release) probability of detection across other potential values for the increased probability of detection under EM. Colors correspond to probabilities of detection under release. Guilty charges refers to counting the crime cost only on charges which has a guilty finding. Low murder refers to using a reduced cost of murder in crime cost computations. Line types correspond to weights being used are for CATE (even weights across defendants and common support) or CATR (higher weights for released defendants across common support). 95% confidence intervals are computed using 200 bootstrap runs but are censored in extreme cases for readability.

Figure 10: Estimated Change in Pretrial Crime Cost: EM vs. Detention



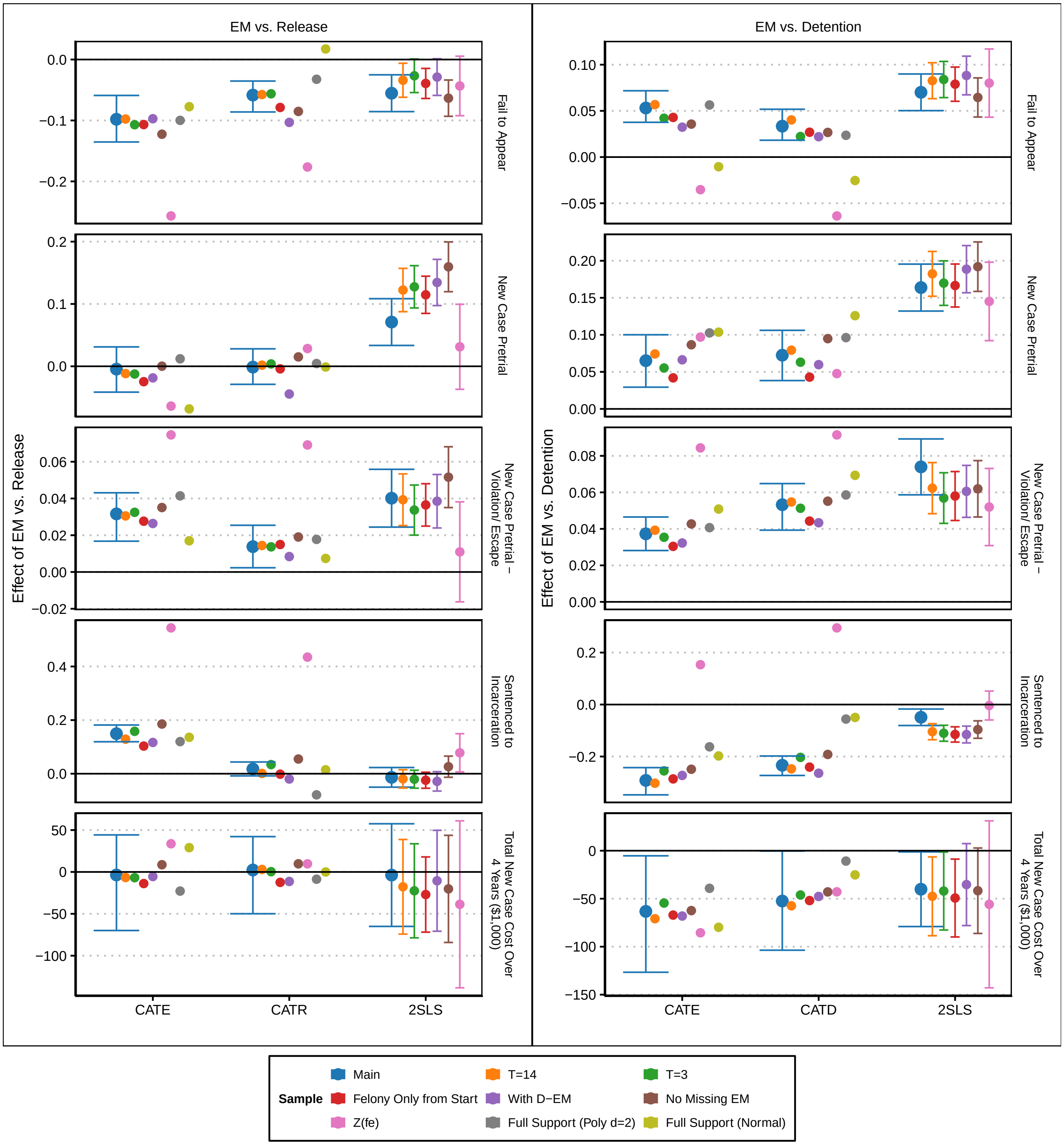
Note: Figure displays implied change in pretrial crime costs using a single value for the initial (release following detention) probability of detection across other potential values for the increased probability of detection under EM. Colors correspond to probabilities of detection under detention. Guilty charges refers to counting the crime cost only on charges which has a guilty finding. Line types correspond to weights being used are for CATE (even weights across defendants and common support) or CATD (higher weights for detained defendants across common support). 95% confidence intervals are computed using 200 bootstrap runs.

Figure 11: Estimated Amount of Pretrial Crime Cost under EM



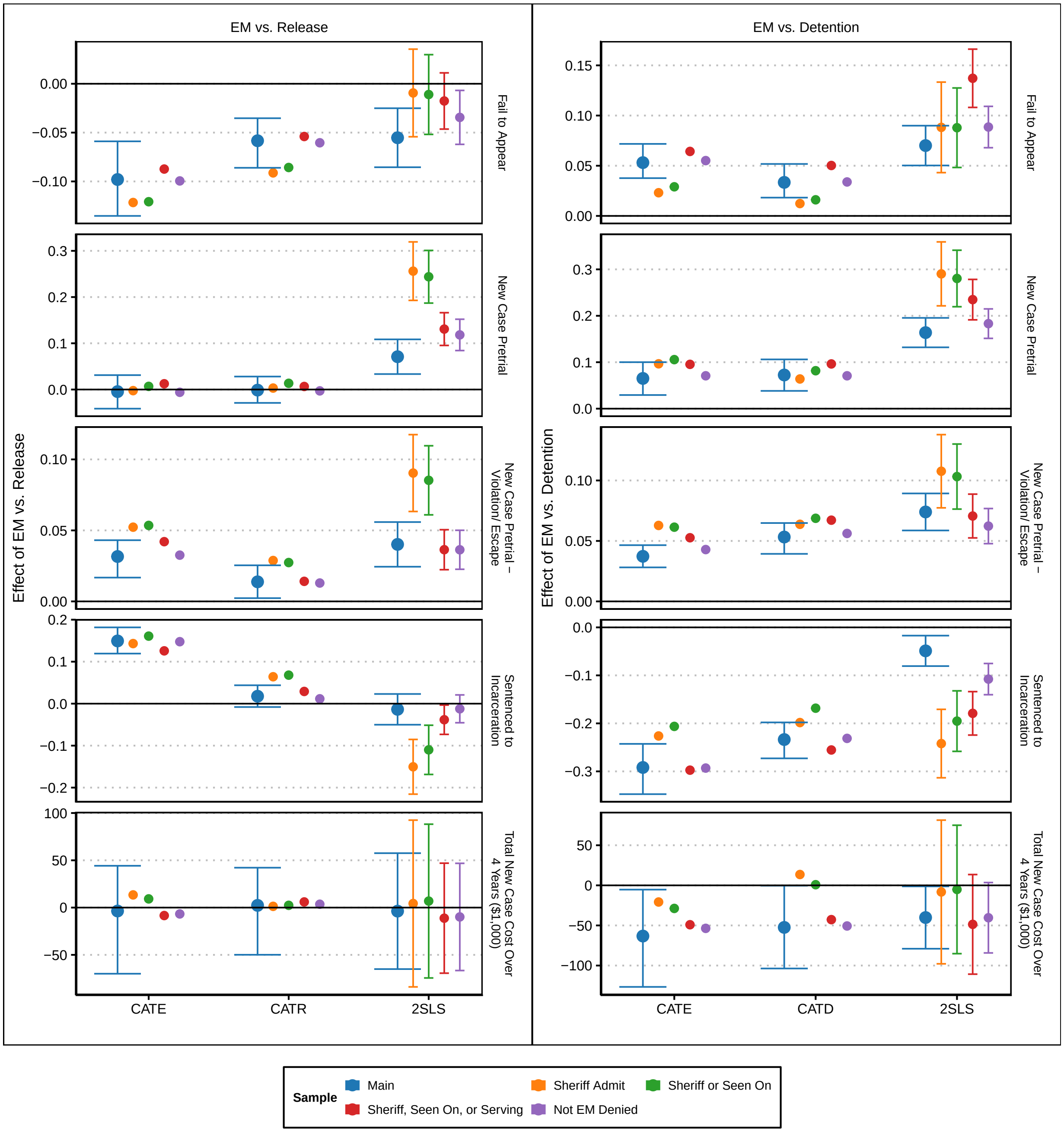
Note: Figure displays implied amounts of pretrial crime costs for the average defendant and detained defendant across potential values for the probability of detection under EM. Guilty charges refers to counting the crime cost only on charges which has a guilty finding. Low murder refers to using a reduced cost of murder in crime cost computations. Line types correspond to weights being used are for CATE (even weights across defendants and common support) or CATD (higher weights for detained defendants across common support). 95% confidence intervals are computed using 200 bootstrap runs.

Figure 12: Treatment Effects for Main Robustness Tests



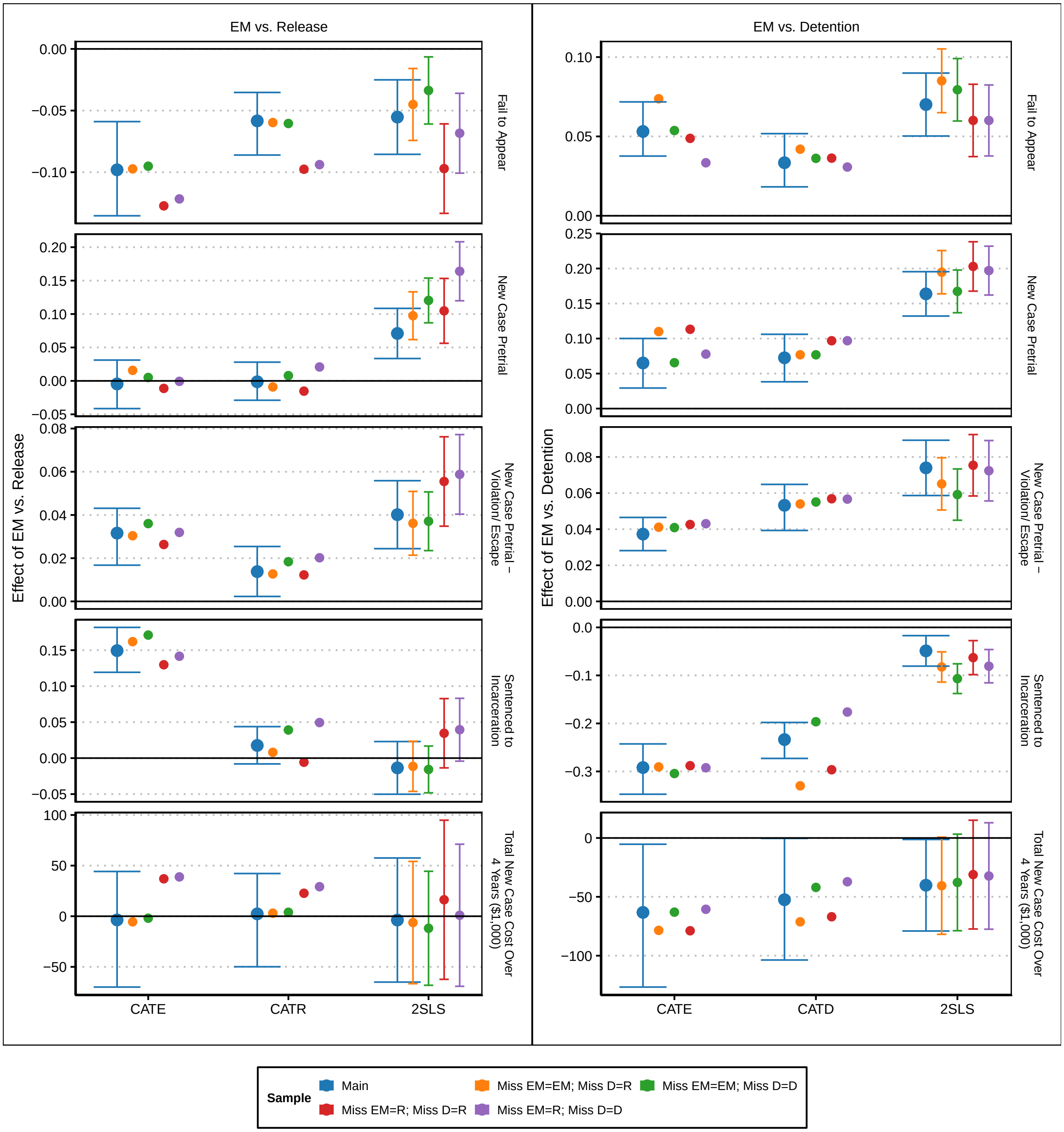
Note: Figure displays the CATE, CATR, CATD, and 2SLS estimates for various robustness checks for the effect of EM relative to Release (left) and EM relative to Detention (right). CATE(R/D) are constructed from MTEs that are estimated semiparametrically with equation (6) where Φ_s are 3rd degree polynomials for all samples unless otherwise specified. Poly d=2 (3rd degree polynomial for Φ_s) and Normal (normally distributed errors) are fully parametric and treatment parameters are constructed over full support. 95% confidence intervals of the estimates are computed using 200 bootstrap runs for non-2SLS, while 2SLS confidence intervals are constructed from standard errors clustered at the defendant level.

Figure 13: Treatment Effects for Recoded Treatments using Disposition Codes



Note: Figure displays the CATE, CATR, CATD, and 2SLS estimates for various robustness checks for the effect of EM relative to Release (left) and EM relative to Detention (right) under different re-codings of pretrial treatments using disposition codes to determine if a defendant was assigned to EM. "Sheriff Admit" means the defendant was explicitly noted to have been admitted into the sheriff's EM program; "Sheriff or Seen On" allows for if the defendant was explicitly noted to be on EM; "Sheriff, Seen On, or Serving" allows for if the defendant was explicitly noted to be serving a monitoring program; and "Not EM Denied" includes "Sheriff, Seen On, or Serving" defendants but excludes any defendant explicitly noted to not be admitted to EM (with bail set to stand). CATE(R/D) are constructed from MTEs recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s), unless otherwise specified. 95% confidence intervals of the estimates are computed using 200 bootstrap runs for non-2SLS, while 2SLS confidence intervals are constructed from standard errors clustered at the defendant level.

Figure 14: Treatment Effects for Recoded Missing Treatments



Note: Figure displays the CATE, CATR, CATD, and 2SLS estimates for various robustness checks for the effect of EM relative to Release (left) and EM relative to Detention (right) under different re-codings of pretrial treatments for cases with missing jail data. CATE(R/D) are constructed from MTEs recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s), unless otherwise specified. 95% confidence intervals of the estimates are computed using 200 bootstrap runs for non-2SLS, while 2SLS confidence intervals are constructed from standard errors clustered at the defendant level.

Table 1: Summary Statistics for Branch 1 Cases by Pretrial Treatment

	Release (1)	EM (2)	Detain (3)
Median Days in Jail	0	-	58
Median Days on EM	0	41	-
Median Case Duration	33	55	122
Bond Amount	15160.05	26699.44	71615.68
Defendant Demographics			
Def. Black	0.63	0.76	0.77
Def. Hispanic	0.2	0.12	0.12
Def. White	0.17	0.12	0.1
Def. Male	0.85	0.83	0.92
Def Age	33.14	36.75	32.97
Case Characteristics			
Charge - Felony	0.45	0.86	0.77
Charge - Felony Violent	0.01	0.01	0.11
Charge - Felony Drug Poss.	0.28	0.57	0.28
Charge - Felony Drug Deliv.	0.06	0.16	0.14
Charge - Felony Property	0.06	0.13	0.13
Charge - Felony Weapon	0.03	0.01	0.14
Charge - Misdemeannaor	0.58	0.23	0.36
Charge - Misd. Property	0.05	0.02	0.04
Charge - Traffic	0.19	0.16	0.12
Charge - Other	0.06	0.05	0.05
Case History (since 2000)			
Past Cases	5.95	7.86	8.91
Case within Year	0.34	0.41	0.53
Past Failure to Appear	1.3	1.82	2.02
Past Felonies	1.48	2.53	2.69
Past Felonies - Violent	0.1	0.09	0.21
Past Guilty	0.74	1.31	2.07
Past Guilty Felonies	0.26	0.55	0.91
Past Guilty Felonies - Violent	0.02	0.03	0.08
N Obs	34937	22714	26764
Share of Obs	0.41	0.27	0.32

Note: Table displays summary statistics by pretrial treatment for defendants with one observation per defendant and arrest, meaning if there is a felony and a misdemeanor case against a defendant both sets of charges are aggregated to one observation. Variables beginning with 'Charge' are binary variables indicating any charge of a specific type.

Table 2: Summary Statistics of Outcomes

	Release (1)	EM (2)	Detain (3)
Failure to Appear	0.1420	0.0761	0.0360
Any Guilty Felony Charge	0.221	0.451	0.589
New Case Pretrial	0.145	0.178	0.110
New Case Post-Trial within 4 Years	1.51	1.56	1.67
New Felony Case Post-Trial within 4 Years	0.477	0.616	0.597
New Violent Felony Case Post-Trial within 4 Years	0.0552	0.0427	0.0694

Note: Table displays summary statistics of outcomes by pretrial treatment for defendants with one observation per defendant and arrest, meaning if there is a felony and a misdemeanor case against a defendant both sets of charges are aggregated to one observation.

Table 3: Testing for Violation of Judge Assignment

Est.	(1)	(2)	(3)
LK / LK(1)	1	0.999	0.495
Pseudo-R2	0	0	0.51
Def. Chars		X	
DoW + YM Fes			X

Note: Table displays results from a multinomial logistic regression with the judge assigned to a specific case as the outcome variable to test if defendant observables are predictive of judge assignment. Column (1) is the base model including no regressors, Column (2) only includes case level characteristics, Column (3) contains day of week and year-month fixed effects. LK/LK(1) is the ratio of likelihoods of the model to that of Column (1). Pseudo-R2 is McFadden (1974)'s measurement of explained variation in the model.

Table 4: Relevance Tests

	(1)	(2)	(3)
Outcome = Release			
$\hat{Z}^R$ (1 SD)	0.087***	0.087***	0.087***
	(0.00167)	(0.00155)	(0.00152)
Outcome = Detain			
$\hat{Z}^D$ (1 SD)	0.097***	0.097***	0.097***
	(0.00152)	(0.00139)	(0.00136)
Controls		X	X
Month + DoW FEs			X

Note: Table displays the relevance of the standardized instruments in predicting whether or not the defendant is assigned to Release or Detention. Column (1) contains no controls, Column (2) adds case and defendant controls, and Column (3) adds time fixed effects. Standard errors are clustered at the defendant level.

Table 5: 2SLS Results for Main Sample

	EM vs. Release	EM vs. Detain
	(1)	(2)
Pretrial Misconduct		
Fail to Appear	-0.055***	0.07***
	(0.015)	(0.01)
New Case Pretrial	0.071***	0.16***
	(0.019)	(0.016)
New Case Pretrial - Serious	0.016	0.056***
	(0.012)	(0.011)
New Case Pretrial - Low-Level	0.035**	0.086***
	(0.017)	(0.014)
New Case Pretrial - Violation/ Escape	0.04***	0.074***
	(0.008)	(0.0078)
Case Outcomes		
Any Guilty Felony Charge	0.027	-0.016
	(0.021)	(0.014)
Sentenced to Incarceration	-0.014	-0.049***
	(0.019)	(0.016)
Post-Trial New Cases		
Total New Cases Post Trial within 4 Years	0.16	0.098
	(0.11)	(0.1)
Total New Cases Post Trial within 4 Years - Felony	0.017	-0.016
	(0.053)	(0.044)
Total Case Cost (Pre and Post) ($1,000)		
Total New Case Cost Over 4 Years	-3.7	-40**
	(31)	(20)
Total New Case Cost Over 4 Years (Low Murder)	0.12	-0.071
	(2.6)	(1.9)
Total New Case Pretrial Cost	-18*	-8.3
	(9.5)	(6.6)
Total New Case Pretrial Cost (Low Murder)	0.15	2.6***
	(0.93)	(0.72)
Total New Case Post-Trial Over 3 Years Cost	20	-17
	(26)	(15)
Total New Case Post-Trial Over 3 Years Cost (Low Murder)	-0.03	-1.8
	(2.1)	(1.5)
Min. N	51,327	51,327

Note: Table displays the results of 2SLS regressions for the main felony sample for the effect of EM vs. Release and EM vs. Detention. Includes case level controls and quarter and weekday fixed effects. Standard errors clustered at the defendant level are in parantheses. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 6: ATE, ATR, and ATD Results for Common Support

	EM vs. Release		EM vs. Detain	
	CATE	CATR	CATE	CATD
	(1)	(2)	(3)	(4)
Pretrial Misconduct				
Fail to Appear	-0.0981***	-0.0583***	0.0531***	0.0334***
	[-0.135 , -0.059]	[-0.0861 , -0.0353]	[0.0376 , 0.0718]	[0.0182 , 0.0518]
New Case Pretrial	-0.0045	-0.0013	0.0651***	0.0725***
	[-0.0414 , 0.0311]	[-0.0289 , 0.028]	[0.0294 , 0.1]	[0.0383 , 0.106]
Serious Charge	-0.003	-0.0035	0.0002	0.0035
	[-0.0263 , 0.0166]	[-0.0244 , 0.0126]	[-0.0228 , 0.0229]	[-0.0187 , 0.0243]
Low-Level Charge	-0.0298*	-0.0185	0.0382**	0.0307*
	[-0.066 , 0.0019]	[-0.0443 , 0.0105]	[0.0071 , 0.0661]	[-0.0009 , 0.0576]
Violation/ Escape Charge	0.0316***	0.0138**	0.0373***	0.0532***
	[0.0168 , 0.0431]	[0.0023 , 0.0254]	[0.0281 , 0.0465]	[0.0393 , 0.0648]
Total Cost of Pretrial New Cases ($1,000)				
Total Pre (Full Murder Cost)	-10.6	-2.47	-4.2	-9.57
	[-40.9 , 10.7]	[-18.8 , 7.34]	[-29.8 , 12.3]	[-27.8 , 5.7]
Total Pre (Low Murder Cost)	-1.47	-0.442	-0.465	-0.851
	[-3.63 , 0.416]	[-1.74 , 0.712]	[-2.6 , 1.63]	[-2.7 , 0.661]
Case Outcomes				
Any Guilty Felony Charge	0.0607**	-0.0278	-0.305***	-0.222***
	[0.0145 , 0.112]	[-0.0688 , 0.0098]	[-0.359 , -0.25]	[-0.266 , -0.186]
Sentenced to Incarceration	0.149***	0.0176	-0.292***	-0.234***
	[0.119 , 0.182]	[-0.008 , 0.0438]	[-0.347 , -0.243]	[-0.273 , -0.198]
Total Post-Trial New Cases within 4 Years				
Total Cases Post within 4 Years	0.389***	0.351***	-0.0166	0.281***
	[0.207 , 0.593]	[0.202 , 0.501]	[-0.239 , 0.245]	[0.111 , 0.468]
Total Felony Cases Post within 4 Years	0.134***	0.106***	0.01	0.143***
	[0.0364 , 0.216]	[0.033 , 0.184]	[-0.084 , 0.118]	[0.067 , 0.234]
Total Cost of Post-Trial New Cases within 3 Years ($1,000)				
Total Post (Full Murder Cost)	6.04	7.48	-44.4*	-31.2
	[-60.8 , 41.8]	[-36.6 , 45.5]	[-84.2 , 0.414]	[-72.3 , 10.7]
Total Post (Low Murder Cost)	2.86	3.02**	-2.28	0.704
	[-1.34 , 6.42]	[0.0141 , 5.94]	[-5.39 , 1.73]	[-2.64 , 4.22]
Total Cost of New Cases within 4 Years of Bond Court ($1,000)				
Total Pre and Post (Full Murder Cost)	-3.62	2.32	-63.2***	-52.4**
	[-69.9 , 44.2]	[-49.9 , 42.2]	[-127 , -5.32]	[-104 , -0.337]
Total Pre and Post (Low Murder Cost)	2.31	2.32	-5.73**	-1.41
	[-2.45 , 7.08]	[-1.05 , 5.77]	[-11.1 , -0.0579]	[-5.55 , 2.89]

Note: Table displays the common support average treatment effect (CATE) and average treatment effect on the released and detained (CATR, CATD) estimates averaging over the MTEs. 95% confidence intervals are displayed below the estimate. 95% CIs and p-values are computed through 200 bootstrap runs, with each run calculating the effect, then the lower bound and upper bound are the bootstrap treatment effects at the 1st, 2.5th, 5th, 95th, 97.5th, and 99th percentile. MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s), unless otherwise specified. ***p < 0.01; **p < 0.05; *p < 0.1

Table 7: Breakdown of Relative Counterfactual Costs ($1,000)

	EM vs. Release		EM vs. Detention	
	ATE	ATR	ATE	ATD
(Direct) Jail/EM	0.44	0.23	-7.19	-8.96
	[0.4 , 0.47]	[0.2 , 0.26]	[-8.17 , -6.46]	[-9.46 , -8.54]
Failure to Appear	-0.1	-0.06	0.05	0.03
	[-0.14 , -0.06]	[-0.09 , -0.04]	[0.04 , 0.07]	[0.02 , 0.05]
Pre-Trial Crime	-46.31	-19.41	-24.99	-41.76
	[-147.06 , 26.1]	[-72.36 , 14.17]	[-106.29 , 28.05]	[-105.04 , 6.83]
Pre-Trial Crime (Low Murder)	-12.08	-8.05	-8.74	-10.62
	[-18.98 , -6.07]	[-11.67 , -4.67]	[-15.38 , -1.99]	[-16.12 , -6.14]
Sentencing	2.99	0.35	-5.84	-4.67
	[2.39 , 3.63]	[-0.16 , 0.88]	[-6.95 , -4.85]	[-5.46 , -3.96]
Post-Trial Crime (3 Years)	20.12	24.94	-147.9	-103.93
	[-202.7 , 139.37]	[-121.93 , 151.62]	[-280.61 , 1.38]	[-240.83 , 35.6]
Post-Trial Crime (3 Years) (Low Murder)	9.52	10.05	-7.6	2.35
	[-4.47 , 21.39]	[0.05 , 19.79]	[-17.97 , 5.78]	[-8.79 , 14.07]
Total	-22.86	6.06	-185.87	-159.3
	[-347.11 , 169.51]	[-194.34 , 166.9]	[-401.99 , 18.18]	[-360.77 , 29.99]
Total (Low Murder)	0.77	2.53	-29.31	-21.88
	[-20.79 , 19.36]	[-11.67 , 16.22]	[-48.44 , -7.46]	[-39.81 , -4.51]

Note: Table displays the relative costs (positive) and savings (negative) per defendant under counterfactual policies, where the average defendant is moved from release to EM or detention to EM (ATE), where the average released or detained defendant is moved to EM (ATR or ATD). Estimates are constructed using common support analogs of MTRs with costs or scaling applied and weighted by the relevant treatment parameter weights. I use average costs of $15 per defendant-day for EM and $150 defendant-day for jail and apply a 40% marginal cost. FTAs are valued at $1,000. Crime costs are based on aggregated arrest costs assuming a 30% cost-weighted detection rate under release, post-detention release, or post-trial, and a 60% cost-weighted detection rate under EM. Pretrial crime costs include punishment costs. Sentencing costs are assumed to be $20,000 per defendant sentenced to incarceration, based on Mai and Subramanian (2017) estimates of average prison cost per defendant in Illinois in 2015 of $33,507 and applying a 60% marginal cost discount. This also assumes defendants sentenced to incarceration are serving one year on average, which may be an underestimate (over 75% of individuals sentenced to prison or jail are sentenced to prison). 95% confidence intervals are displayed below the estimate. 95% CIs are computed through 200 bootstrap runs.

A Appendix A: Additional Analyses and Background

A.1 Background

The judge can also link I and D bonds with supervised release requirements, though the main role of the bond is to determine if they can leave the custody of the Sheriff (i.e., exit jail pretrial). In this sense, EM can also be coupled with D-bonds (D-EM), which required the defendant to stay in jail until they paid 10% of the bond amount and were then released onto EM, which accounted for 17% of D-bonds between 2008 and 2012 (Federation (2017)). However, it is unclear how many defendants were actually released onto EM from D-EMs during the period. Prior to 2012, little data on EM usage in Cook County was available (Dizikes and Lightly (2015)). Figure B.3 displays these trends across the sample period.

In 2012, disputes began between the Court and the Sheriff (who runs the jail and most of the EM releases) over the overcrowding of the Cook County Jail and EM usage began. As a result, in November 2012, judges functionally stopped using D-EM bonds which further contributed to jail overcrowding (Federation (2017)), though they were occasionally used during the period of study. This sparked the introduction of the IEM bond discussed in the paper, though IEM bonds are referred to as "Electronic monitoring with D-bonds" in CGL and Appleseed (2022). The IEM bond offered an attractive solution to judges: release defendants from jail and avoid overcrowding but have them monitored by the Sheriff, who bears responsibility for any failures. The initial appeal was increased following reforms in September 2013 which urged a reduction in defendants forced to stay in jail due to lack of money (Federation (2017)). A report by the Civic Federation, using different data, also indicates that in September 2013, IEM was about 25% of dispositions (Federation (2017)).

A significantly less common EM program was "Curfew EM" which requires defendants to be in their homes between specified hours, usually 7 pm to 7 am. These programs also co-exist with other monitored release programs by the Chief Judge's office, GPS home confinement, which is primarily used for domestic violence cases (Federation (2020)). See here for a discussion of differences between EM programs.

Recently, Cook County has adopted GPS monitoring systems instead, though these GPS ankle or wrist bracelets operate in a similar capacity, simply without a home unit, and tracks all of the subject's movements (Sheriff (2020)).

A.2 Data

The court data has large numbers of cases, first starting in 1984, but also contains sporadic records dating back to the 1930s. Linking is done using individual record numbers as well as personally identifiable characteristics, such as name, birth date, race, gender, and home address. As a single booking can result in multiple cases (generally 2 if it is a felony case), cases can be linked within individuals using central booking numbers common to both cases (RD numbers if CBs are missing). For linking jail data to court data, I connect defendant identities using individual record numbers, identifiable information (names), and case/detention information.

The CB and RD numbers associated with cases are used to connect court/jail profiles to Chicago-specific arrests and reported crimes. While I have additional information for Chicago arrestees, I do not require this filter, though 92.37% of the data are reported to have been arrested by the CPD — 87.32% can be linked to a specific arrest, while 73.47% can be linked to a specific crime.

Importantly, not all cases can be linked to a reasonable jail spell, which means that the individual follows a quick timeline of beginning a jail spell (i.e., defendant is reported to have entered jail), having a case opened against them, and proceeding to bond court. This lack of linkage is possible due to some individuals never entering the jail system due to immediately going to bond court and being released or due to a linking error — I test the sensitivity of my results to these unlinked cases in Section 6. Cases with I-bonds or EM-bonds are much more likely to be unlinked to a jail spell (61.32% and 25.08%, respectively) relative to all other bond types (which averages 13.81%), which supports the former case. Interestingly, there is no pattern in missing rates for increasing bond amounts. If this were largely due to immediate release from EM, we would expect higher bonds to imply fewer missing. For instrument construction, I include all cases; for treatment construction in the main specification, non-I-bond and non-EM cases which are unlinked to jail spells are dropped in the main sample, while unlinked I-bond and EM bond cases are kept. A defendant can also be classified as "detained" if they technically exit jail due to a transfer or are sent to alternative detention (e.g., prison).

For final filters, I exclude cases that went the branch 1 within 2 days of being opened. I also remove a small subset of individual-booking observations with irregular case patterns and those which do not have resolutions within the court system. I remove 1.94% if they contain more than 3 unique cases, if there are multiple cases and the difference between the minimum and maximum case initialization date is more than 120 days, if the defendant had more than 60 past cases, and if there were more than 6 individual-bookings corresponding to that defendant within the 2 year period. I drop cases within this time period that are transferred outside of the regular system (1.61%), have short case histories without resolution (3.19%), or end with a warrant being issued (0.24%).

Some cases do not have final disposition dates but end with the case being dropped and contain a guilty, not guilty, stricken, or dropped disposition code (4.91%), I use the last event date as the final disposition date. Lastly, I remove a small number of cases that had murder or felony sex charges or resulted in bond denial, and I drop cases without a categorizable treatment (which includes missing jail spells for defendants without I or EM bonds). In Section 6, I test the sensitivity of the results to alternate classifications of the dropped cases due to missing jail information.

In July 2015, the court introduced a public safety assessment system that guided judges on release decisions using a scoring system (Federation (2020)). However, it is not clear in the data if this influenced judge behavior in any way. After 2015 following defendants into future cases becomes more difficult due to the data ending in 2019, and there were significant changes in Chicago in 2016. As a result, I limit the data to July 2013 through 2015.

A.3 Tests for Interpreting 2SLS Results as LATEs

As shown in Bhuller and Sigstad (2022), under certain conditions, we can recover a LATE for one treatment level versus an adjacent one (EM versus release and EM versus detention) with 2SLS. In particular, there are two sets of assumptions relevant for this paper.

Linearity of Predicted Treatments Assuming a single index crossing multiple thresholds model, as presented in Section 4, is correct, then we require that predicted treatment probabilities are linear functions of each other — for example, $\mathbb{E}[Z^1|Z^3]$ is linear in Z^3. Figure B.6 displays the results for the informal test of this linearity assumption. It shows that, by comparing a linear fit to the local polynomial fit, there are some mild deviations from linearity in the full sample, but in the felony sample the fit is near-linear for the entire distribution. However, the assumption applies globally to Z^1 and Z^3 and so even the minor deviations from the linear fit in the felony sample at the tails constitute a small violation.

Average Monotonicity and No Cross Effects Without additional assumptions on selection into treatment, we require two conditions for 2SLS to recover a LATE: average monotonicity (e.g., there is a positive correlation between $Release_{ic}$ and Z_1) across sub-samples and no cross effects (e.g., there is a no correlation between $Release_{ic}$ and Z_3, conditional on Z_1). We can test for violations of average monotonicity across sub-samples and of no cross effects. Table B.6 displays the results of these tests. Across sub-samples, the endogenous treatments are strongly and positively correlated with their respective instrument, and across sub-samples the relationship of a treatment with the unrelated instrument, conditional on the correct instrument, is either statistically or economically insignificant (e.g., the coefficient for Z_a is at a minimum 6 times larger in magnitude than that of Z_b with treatment a as the dependent variable). Nevertheless, these do constitute violations of the no cross effect assumptions.

Overall, these results do constitute violations of the assumptions, but in both cases they are relatively minor. However, as of this writing, the tests in Bhuller and Sigstad (2022) are not formalized, nor do we know to what degree minor violations bias results. With this, I cautiously proceed with the interpretation of the 2SLS results as a LATE.

A.4 Construction of Treatment Parameters

Because I assume MTRs are additively separable in u and x, the construction of the observable and unobservable components are done separately, using a uniform grid of 99 points ($u \in \{0.01, 0.02, ..., 0.99\}$), following the construction of the average treatment effect on the treated (for ATR, ATD) and ATE from Cornelissen et al. (2016).

$$
\begin{aligned}
ATE_{2,1} &= \frac{1}{N}\sum_{i=1}^{N} X_i(\beta_2 - \beta_1) + \frac{1}{n_{p_1}} \sum_{u=100\times\underline{p_1}}^{100\times\overline{p_1}} E[\omega_2 - \omega_1 | u = u] \\
ATE_{2,3} &= \frac{1}{N}\sum_{i=1}^{N} X_i(\beta_2 - \beta_3) + \frac{1}{n_{p_2}} \sum_{u=100\times\underline{p_2}}^{100\times\overline{p_2}} E[\omega_2 - \omega_3 | u = u] \\
ATR &= \frac{1}{N}\sum_{i=1}^{N} \frac{p_{1,i}}{\hat{p_1}} X_i(\beta_2 - \beta_1) + \frac{1}{n_{p_1}} \sum_{u=100\times\underline{p_1}}^{100\times\overline{p_1}} \frac{\Pr(p_1 > 1 - \frac{u}{100})}{\hat{p_1}} E[\omega_2 - \omega_1 | u = u] \\
ATD &= \frac{1}{N}\sum_{i=1}^{N} \frac{p_{2,i}}{\hat{p_2}} X_i(\beta_2 - \beta_3) + \frac{1}{n_{p_2}} \sum_{u=100\times\underline{p_2}}^{\overline{p_2}} \frac{\Pr(p_2 > \frac{u}{100})}{\hat{p_2}} E[\omega_2 - \omega_3 | u = u]
\end{aligned}
$$

where N is the number of observations (cases), $p_1 = 1 - \pi_1$, $p_2 = \pi_2$, $\overline{x}$ and $\underline{x}$ refer to the upper and lower limits of common support, n_x is the number of points between the upper and lower limits, and $\hat{x}$ is the average over the range of common support.

A.5 Comparison with Prior Methods

Figure B.24 displays MTEs within common support ("Main"), polynomial MTEs with full support using sieve-style estimation ("Poly"), and Normal method MTEs for a subset of outcomes. Noticeably, when the MTEs are linear with Main and Poly (e.g., in the case of FTAs or sentence to incarceration for EM versus release), they match the Normal MTEs closely for most values of u, for which the Normal MTEs are linear by construction. At more extreme values of u, however, the Normal MTEs' pre-determined shape leads to more extreme values.

The patterns diverge more significantly when the MTEs for Main and Poly are non-linear (in most cases), leading the Normal MTEs to extrapolate too far different results or mask underlying heterogeneity. For example, using EM versus detention: for the likelihood of having a new case pretrial, Main and Poly produce a parabolic shape, while the Normal method produces a flat line; for the likelihood of being sentenced to incarceration and total case cost over 4 years, Main and Poly have a parabolic shape resulting in near zero effects on high u defendants while the Normal estimates continue moving negative implying large effects for high u defendants.

A.6 Additional Robustness Checks

A.6.1 Alternative Specifications

Results are displayed in Figure B.25.

Alternative π Computation The main results use π_1 and π_2 based on equation (1) using a probit model. I test the robustness of this using two modifications. First, because I do not exclude defendants' observations from the data I use to construct their judge-specific predicted propensities (πs), we may be concerned that defendant unobservables could be biasing the instruments. Though this is unlikely due to the fact that each judge sees thousands of cases and defendants, I test the robustness of the results with respect to this concern by re-calculating $\pi_{1,i}$ and $\pi_{2,i}$ using only observations excluding defendant i, then predicting out of sample, similar to a jackknife instrument (Angrist, Imbens, and Krueger (1999)) used in Dobbie, Goldin, and Yang (2018).

Second, Goldsmith-Pinkham, Hull, and Kolesár (2022) show that judge-instrument first stages can suffer from contamination bias. One recommended solution is to interact judges with fixed effects for the level of randomization (e.g., court rooms). I this paper, judges are randomly assigned conditional on time fixed effects. While not directly applicable, as I use a probit first stage, not OLS, and interact judges with defendant observables, I add judge-time interactions to my first stage (equation (1)) to test the robustness of my results. The results for both of these tests labeled 'Out-Sample' and 'Interact FEs', respectively, and the results are largely the same.

Alternate Specifications The main results use 3rd degree polynomials for Φ_s. To test the sensitivity of the results with respect to this specification, I re-estimate the MTRs and construct MTEs under 3 alternative specifications for Φ_s: 3rd degree B-splines with 4 degrees of freedom, 4th degree B-splines with 6 degrees of freedom, and a 5th degree polynomial. For each of these alternative specifications, the results are generally similar to the main results.

A.6.2 Alternative Samples

Results are displayed in Figure B.26.

Bonds as Treatments It is useful to know how sensitive the results are to mismeasuring the treatment. So, rather than determining treatment based on a day-in-jail cutoff which uses both court and jail data, I recode the treatments to be based on bonds: release is now only I-Bonds and low D-bonds ($< \$20,000$), EM is all EM bonds, and detention is all D-bonds $\geq \$20,000$. The results are generally consistent with the main results. For new case pretrial, the pattern is more exaggerated with larger increases for low-resistance and a modest decrease for mid-level defendants for EM versus release, and for detention versus EM the curve is shifted downward, displaying a larger incapacitation effect of detention. for Sentences to incarceration for detention versus EM the pattern is flat and sloping upwards (but still negative).

CPD Rearrests Instead of using court records to determine new pretrial cases, I can use CPD arrest records instead. For this, I limit the sample to the CPD arrestees (making this sample similar to that of housing-secure and CPD-only samples). The MTEs for CPD rearrests are shown in the same figure as new cases pretrial. The results for EM versus release are highly similar to the main results, though the MTE for detention versus EM is shifted upwards for a net negative effect.

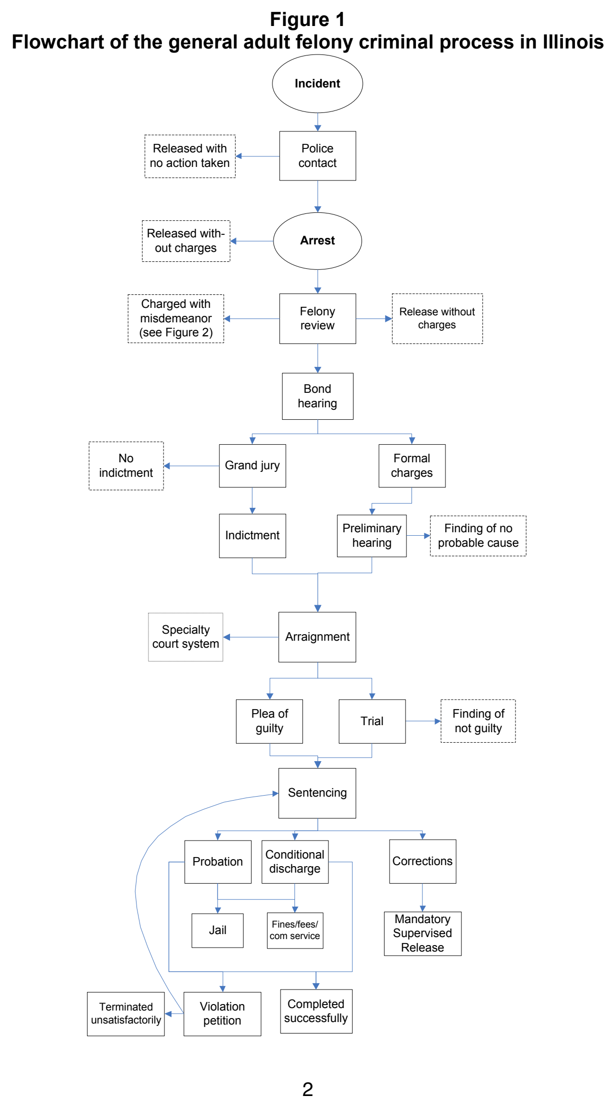
Chicago Arrests Only The main results may be altered because I only have data on arrests and crimes within Chicago which allows for improved matching earlier on in the data cleaning, which is only a part of Cook County. To see if this is significantly influencing my results, I subset my analysis to cases that were initiated by arrests by the Chicago Police Department (92.74% of the felony cases). The results are similar to the main results.

Excluding Non-Home Addresses There is also the potential that the treatment contains selection bias because individuals without stable living conditions cannot be released on EM or from jail if they have no residence to go to — alternatively, individuals without stable living environments may also be forced to violate their EM conditions. In order to determine if this is influencing my results, I subset the data to individuals for whom I have address information for their case. This keeps cases with addresses within the bounds of Chicago and excludes cases if the address is associated with a homeless shelter, recovery facility, or halfway house. This leaves 77.15% of the full sample observations. Overall, for almost all outcomes, the results tract very closely to the CPD-only results and are generally consistent with the patterns and conclusions of the main sample.

Including Misdemeanor Cases Rather including only felony cases, I include misdemeanor cases and D-EM bonds in estimating the MTEs. The misdemeanor sample by itself lacks significant support between treatments, so these results are suggestive at best. Overall, the results are similar to the main sample result, with a few exceptions, particularly for EM versus release: the effect of EM relative to release on FTAs is larger, the new case pretrial pattern is flipped (large negative effects for low resistance defendants which moves positive), and the effect on new case cost is downward sloping.

B Appendix B: Additional Figures and Tables

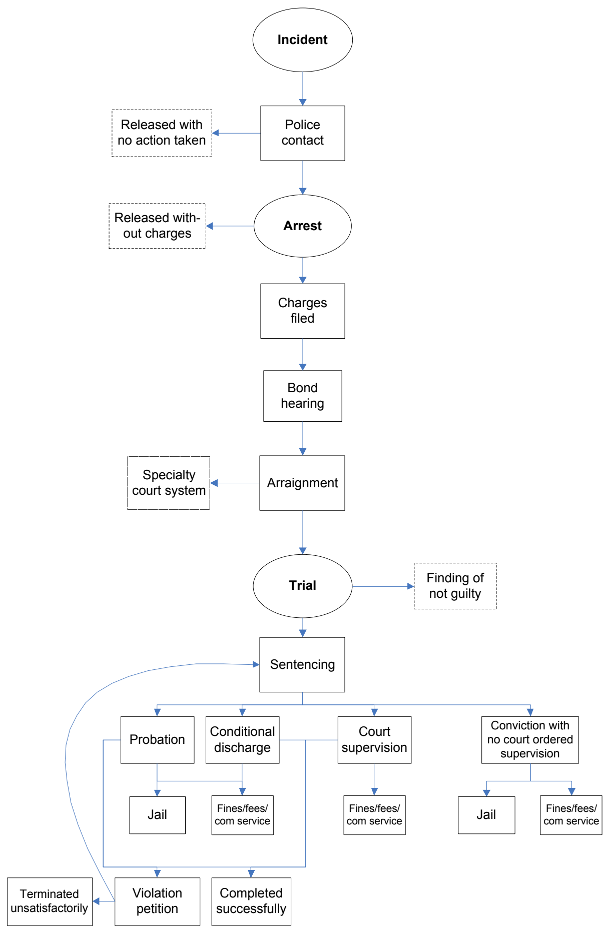
Figure B.1: Felony Case Flow Chart



Note: Figure displays sequence of events for felony cases within Cook County — though document is meant for the entire Illinois criminal justice system more broadly. Source: Afeef et al. (2012), page 2.

Figure B.2: Misdemeanor Case Flow Chart

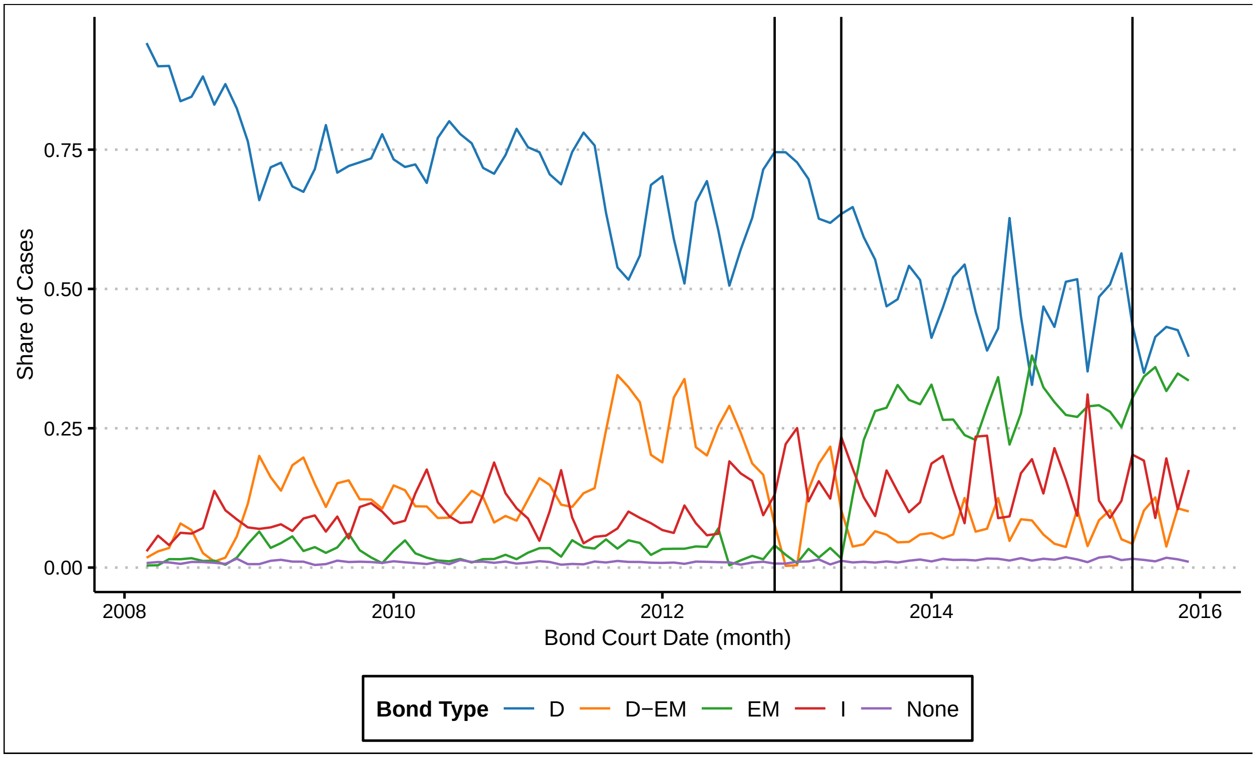
Figure 2
Flowchart of the general adult misdemeanor criminal process in Illinois



3

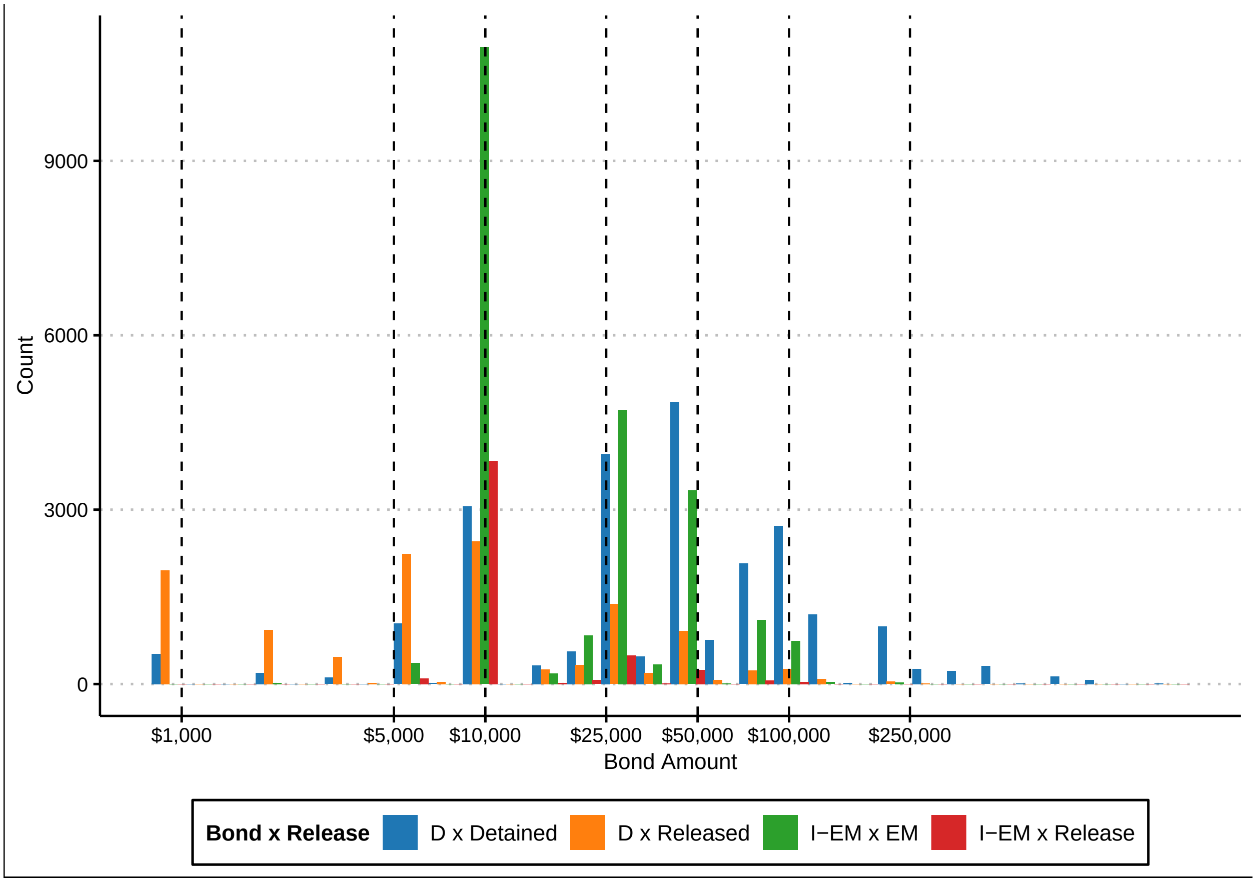
Note: Figure displays sequence of events for misdemeanor cases within Cook County — though document is meant for the entire Illinois criminal justice system more broadly. Source: Afeef et al. (2012), page 3.

Figure B.3: Bond Time Trends in Cook County Court



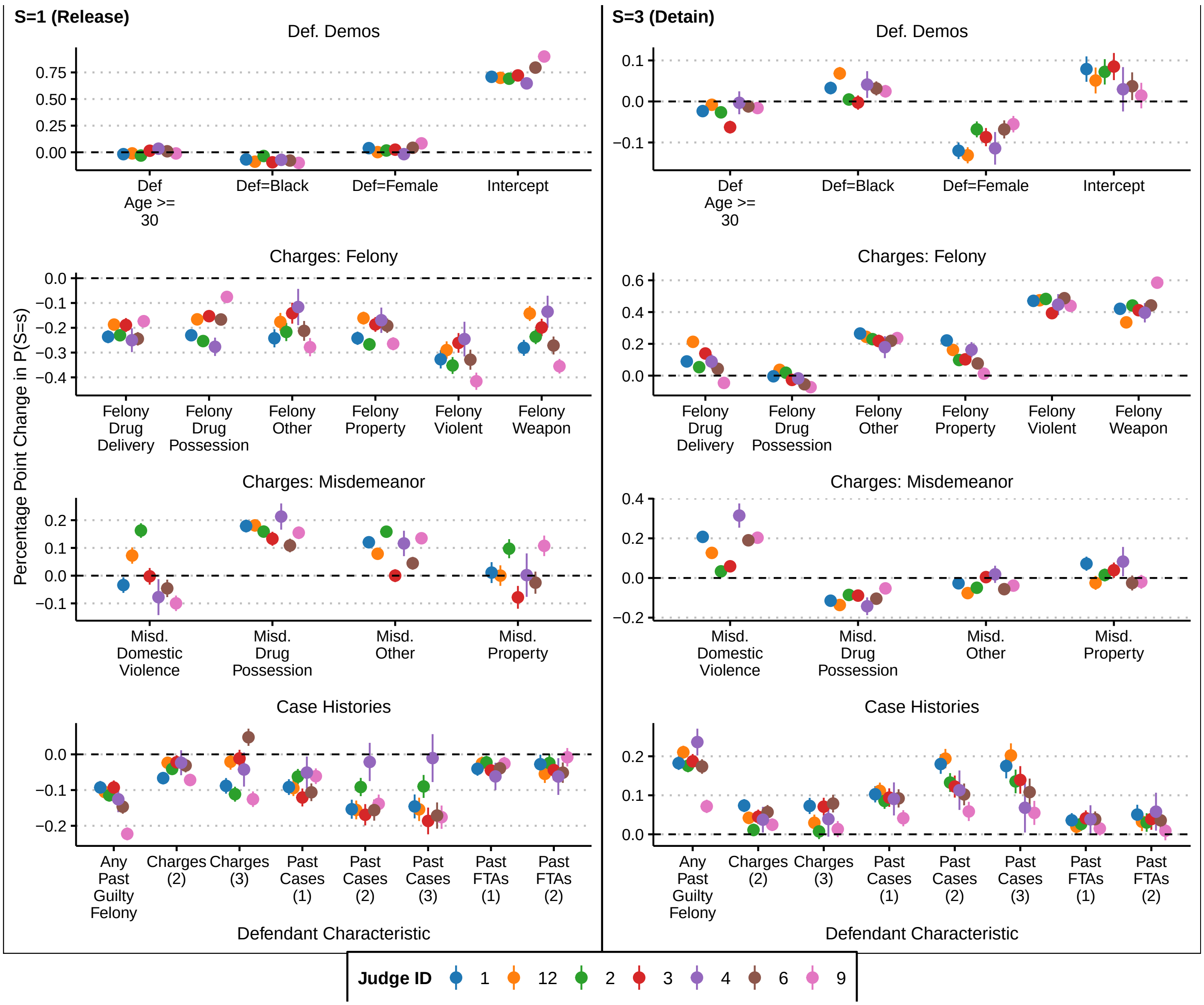
Note: Figure displays the composition of bond types within the sample between 2008 and 2015 aggregated by year-month of bond date. D-EM bond refers to D-bonds coupled with EM release, and EM refers to I-bonds coupled with EM (IEM). The first vertical line denotes November 2012, when D-EM bonds stopped being issued temporarily, and the second vertical line denotes the introduction of IEM bonds in June 2013.

Figure B.4: Distribution of Bond Amounts



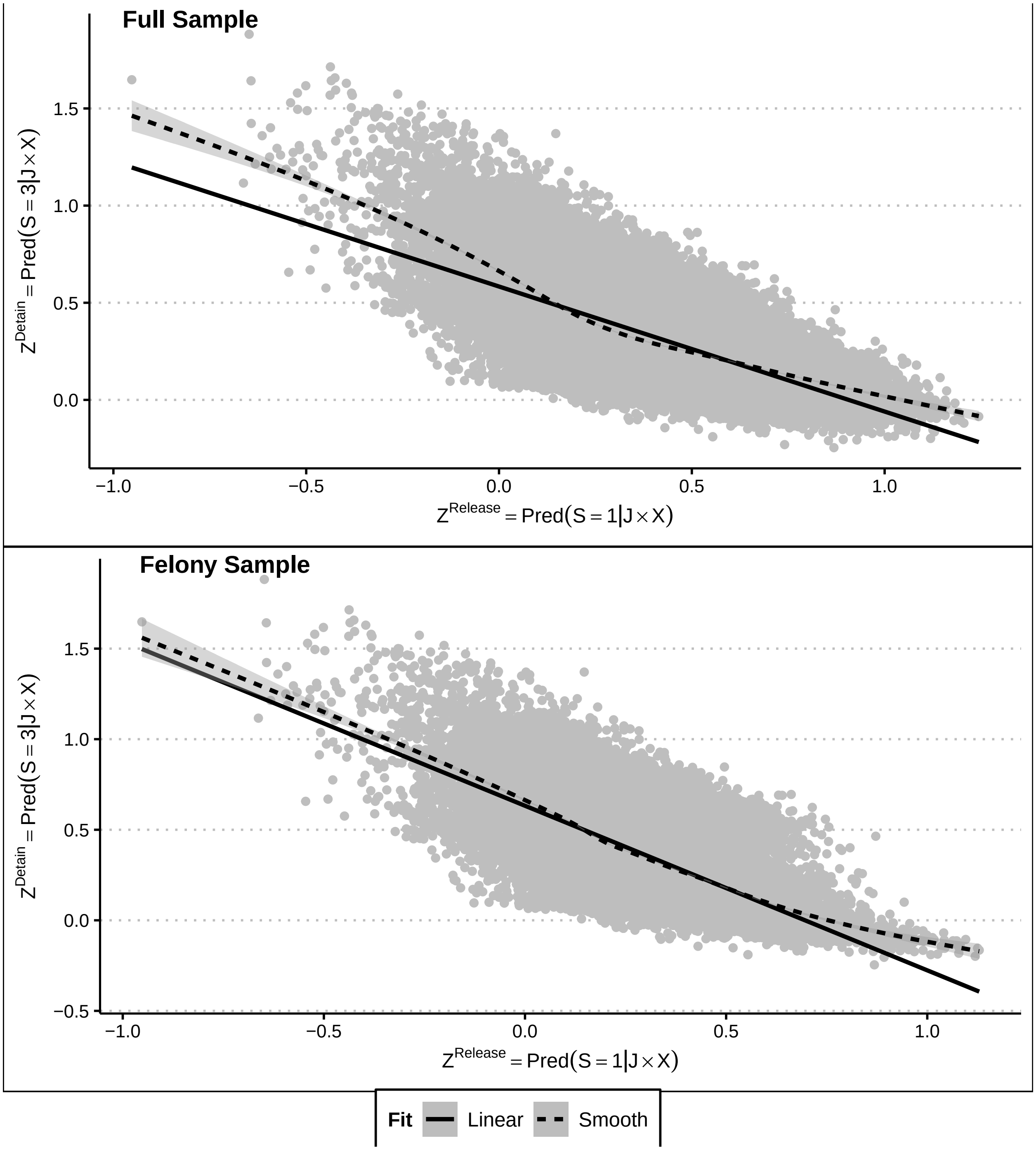
Note: Figure displays histogram of bond amounts (price of release being 10% of the bond amount) for IEM and D bonds for each pretrial treatment during the sample period. X-axis is scaled by the log of bond amount.

Figure B.5: Judge Preferences over Defendant Observables



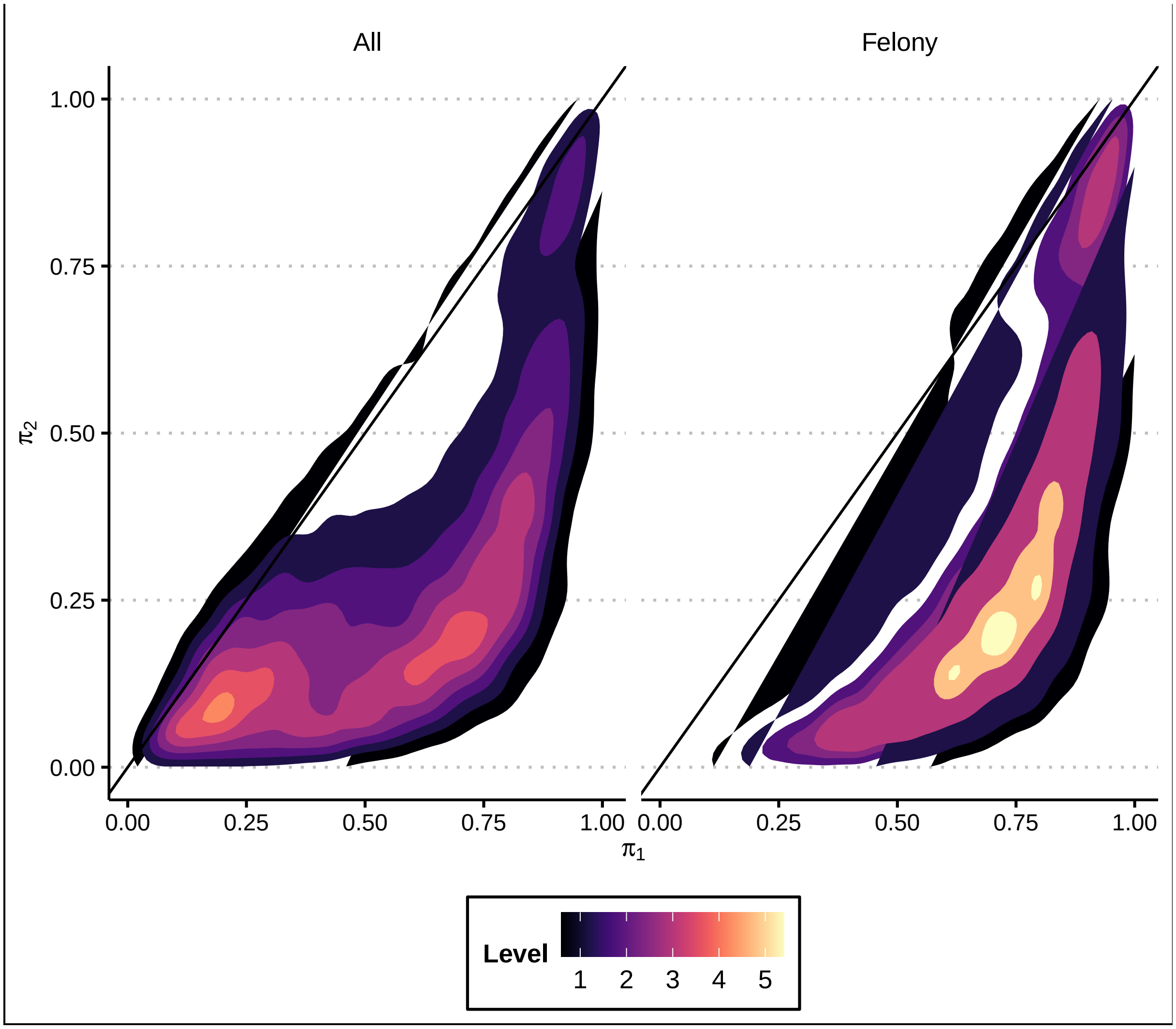
Note: Figure displays OLS estimates of judge-specific coefficients (points, with 95% confidence intervals) for each defendant characteristic (binary variables) recovered by estimating an LPM where pretrial treatment being either Release (S=1) or Detention (S=3) is the outcome, including month and day of week fixed effects (equation (4)). Intercept is the baseline probability of the treatment by the judge. Case history variables involve bins: Charges (2) refers to between 3 and 10 charges total, and Charges (3) refers to more than 10 charges total; Past Cases (1) has between 1 and 4 past cases, (2) has between 5 and 12, and (3) has more than 12; Past FTAs (1) has either 1 or 2 FTAs in past cases, and (2) has more than 2.

Figure B.6: Testing for Linearity Between Predicted Treatments



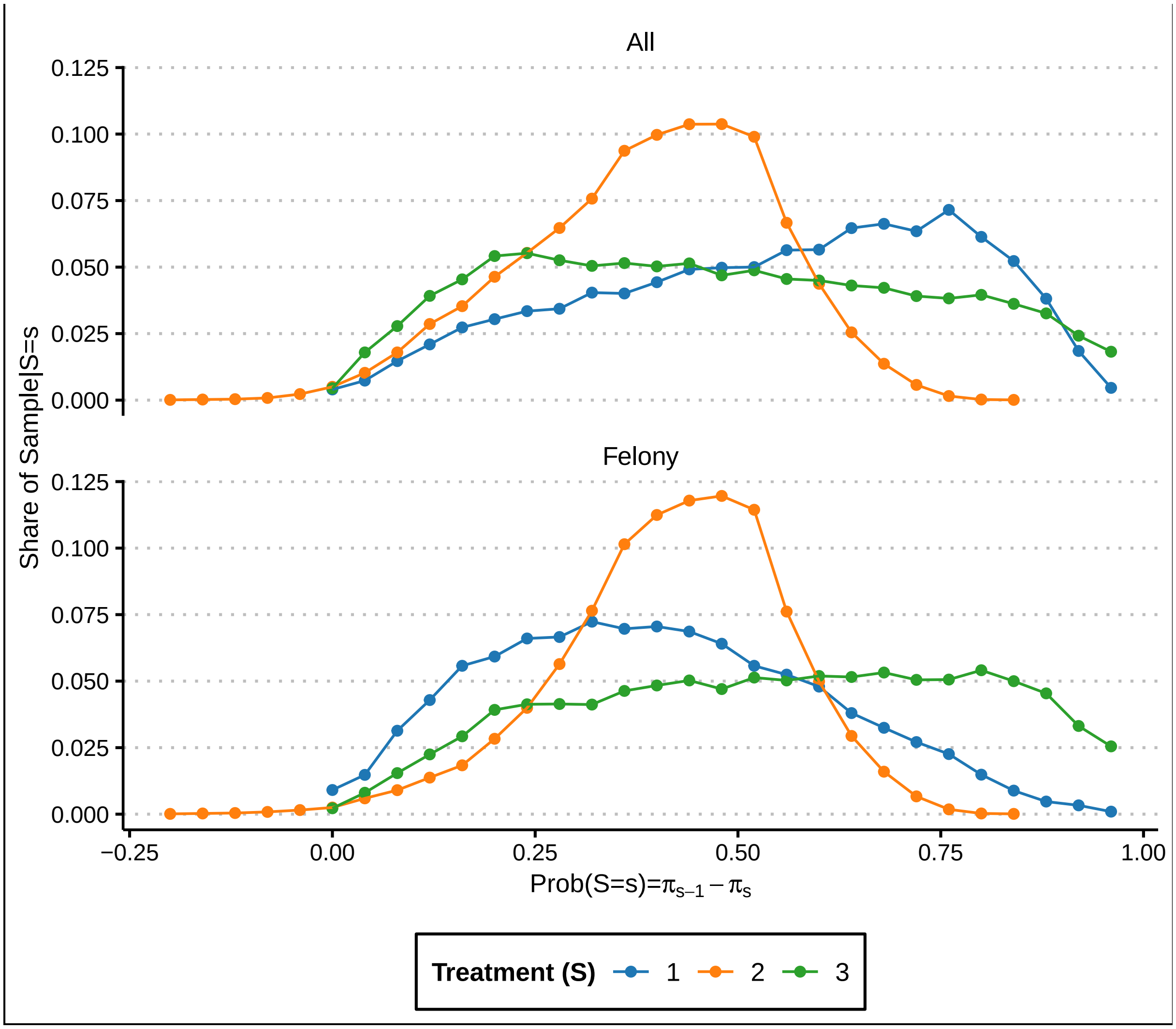
Note: Figure displays scatter plot of observations by values of instruments for Release and Detention as well as their relationship with a linear fit (solid line) and nonlinear fit (dashed line), for both full (top) and main felony (bottom) samples. Nonlinear fit is computed using 'loess' local polynomial regression.

Figure B.7: Density of π_1 and π_2 by Sample



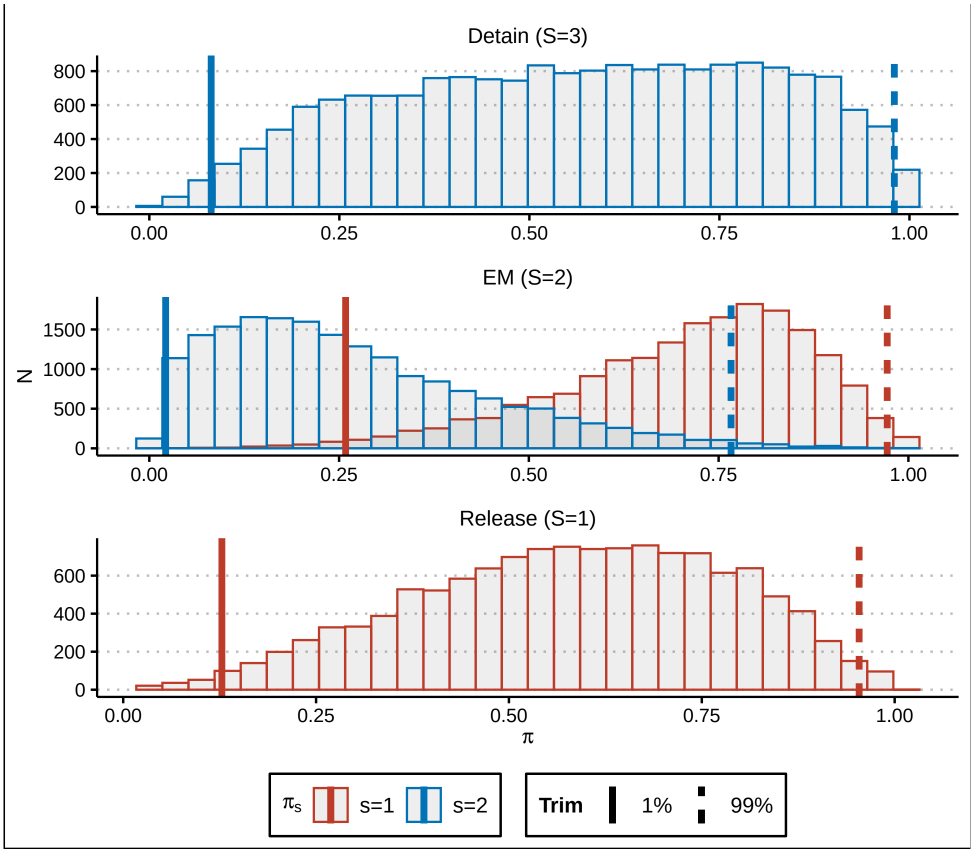
Note: Figure displays density of fitted values for π_1 and π_2 for the full and main felony samples.

Figure B.8: Density of $\Pr(S = s) = \pi_{s-1} - \pi_s$



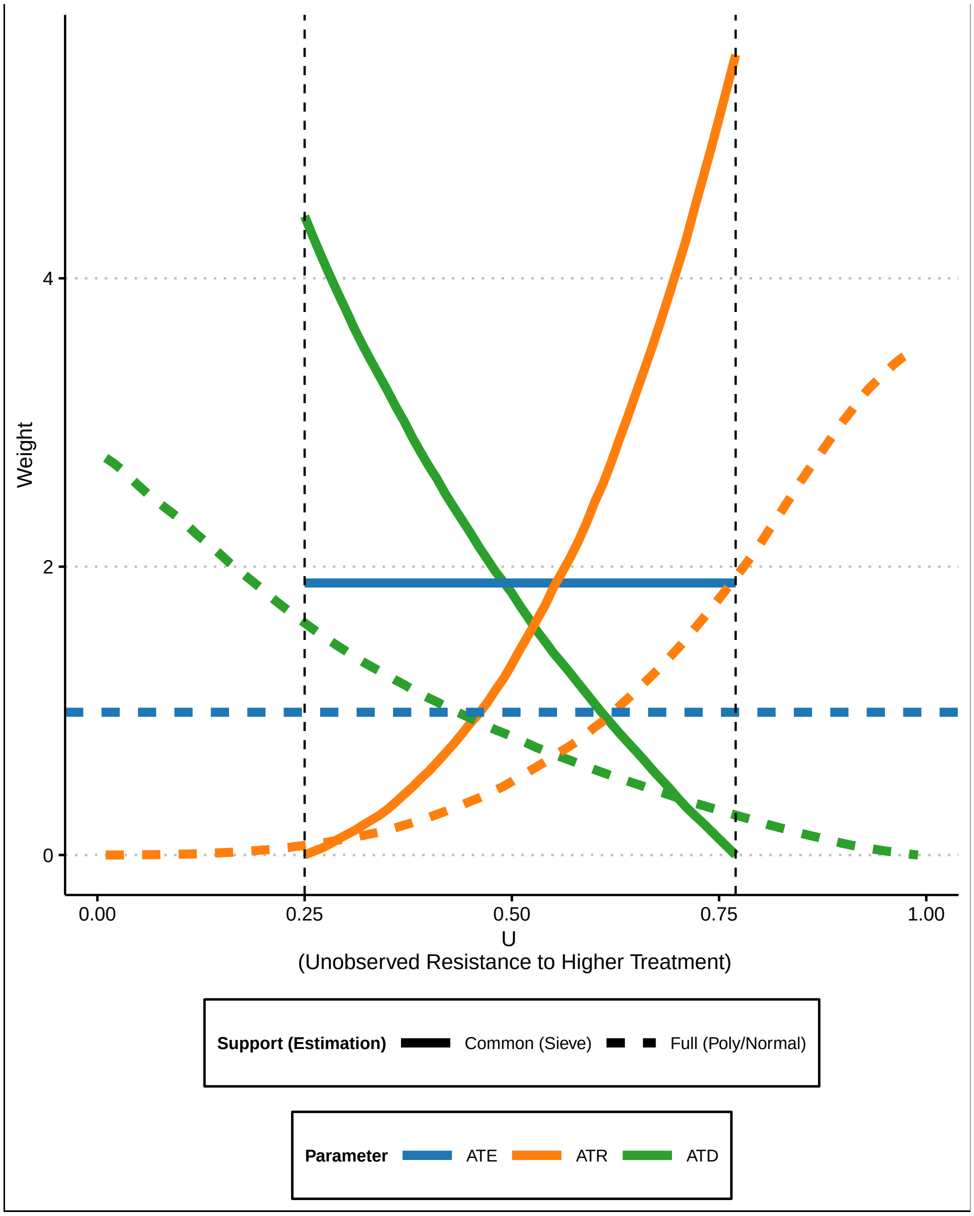
Note: Figure displays densities, within each treatment, of the predicted likelihood the defendant is assigned to that treatment for both full and main felony samples.

Figure B.9: Common Support Across Treatment Levels



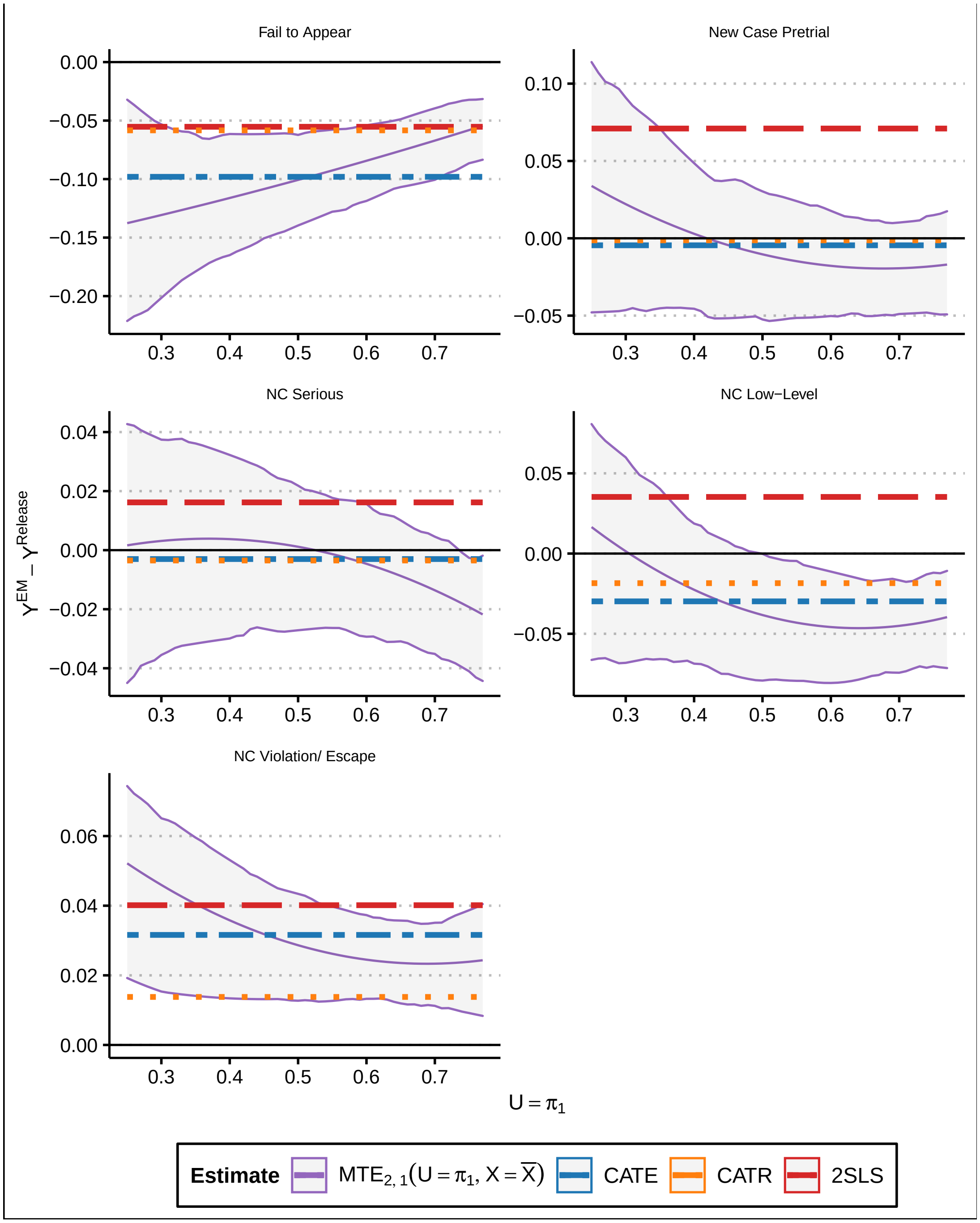
Note: Figure displays histograms of relevant π_s and π_{s-1} values for each treatment level (s). Trims indicate the 1st percentile and 99th percentile of the treatment-specific sample for each relevant π value.

Figure B.10: Treatment Effect Weights



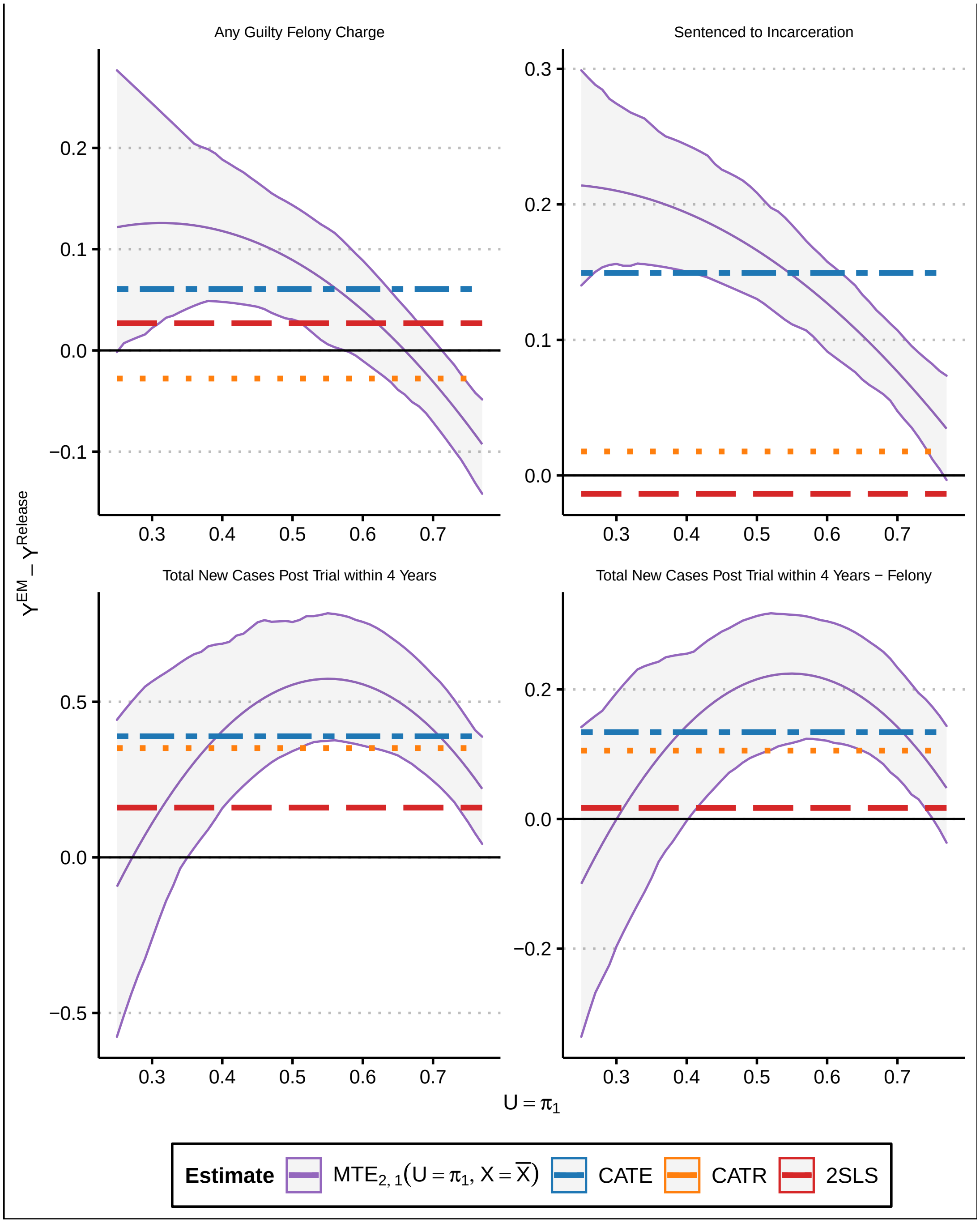
Note: Figure displays the weights used to sum over unobserved resistance to treatment ($U = \pi_1, \pi_2$ depending on the treatment margin) to compute each treatment effect. ATR(D) means average treatment effect within common support on the released (detained) defendants and thus applies a higher weight to higher (lower) resistance to higher treatment. The ATE applies equal weights as u is distributed uniformly.

Figure B.11: EM vs. Release MTEs for Pretrial Misconduct



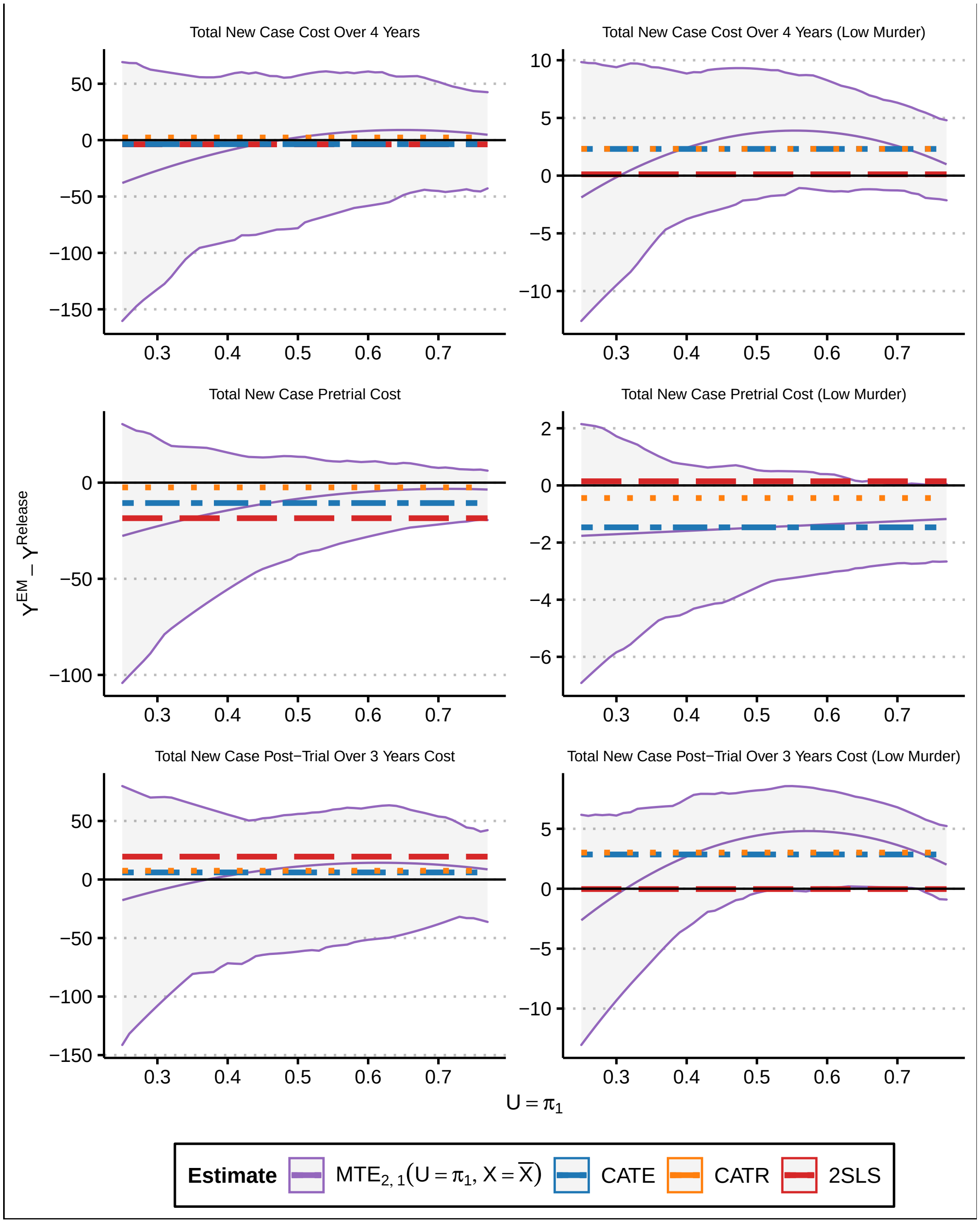
Note: Figure displays the marginal treatment effects (MTEs) across the distribution of defendant types (higher U means more unobservably resistant to higher treatment) for the effect of EM relative to Release for main felony sample cases. Outcomes consist of variables relating to pretrial misconduct. Horizontal lines denote the corresponding 2SLS estimates and CATE, which averages the MTEs over the common support with equal weights, and the CATR, which averages the MTEs placing higher weights on defendant observably and unobservably more likely to be released. MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs.

Figure B.12: EM vs. Release MTEs for Case and Post-Trial Outcomes



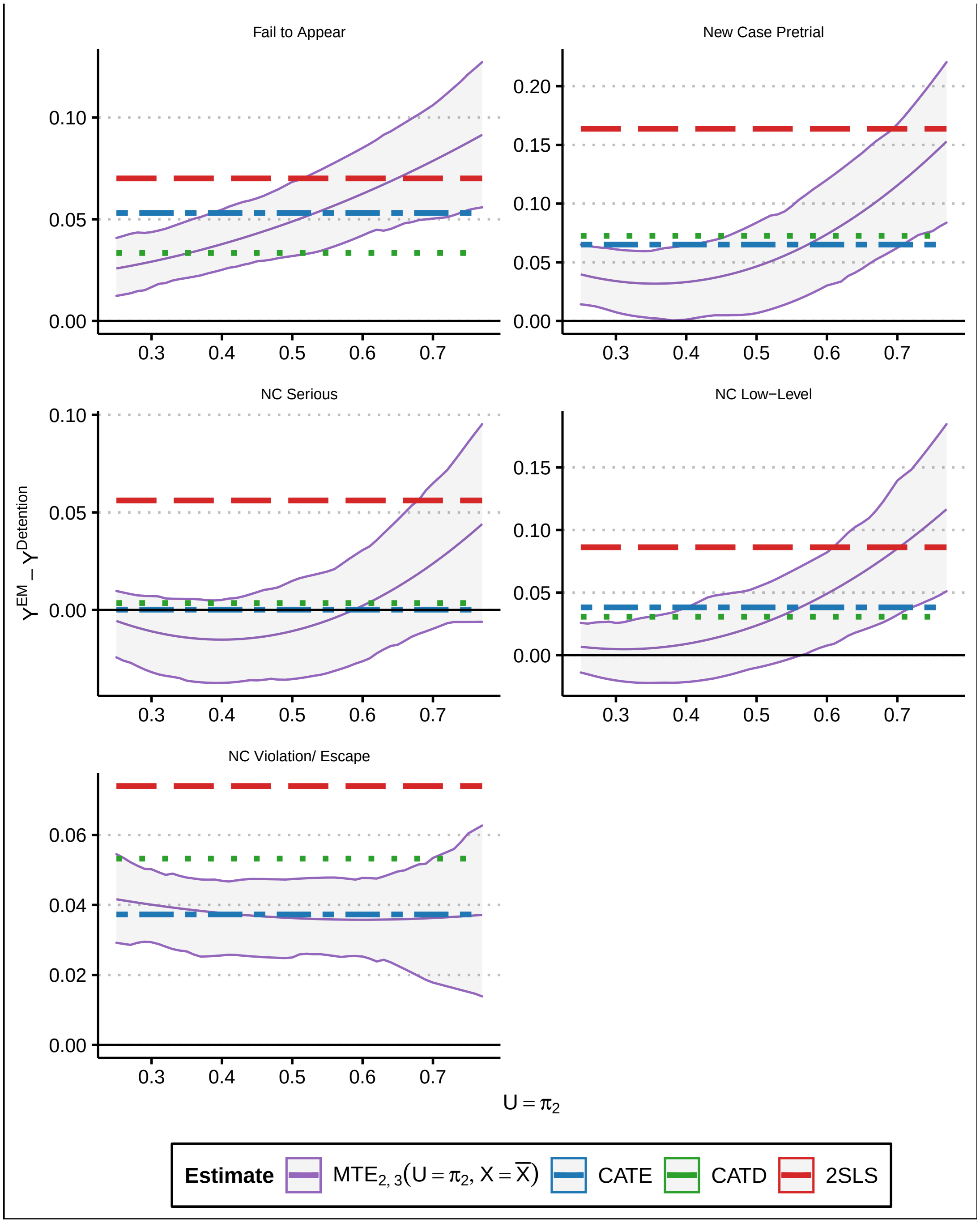
Note: Figure displays the marginal treatment effects (MTEs) across the distribution of defendant types (higher U means more unobservably resistant to higher treatment) for the effect of EM relative to Release for main felony sample cases. Outcomes consist of variables relating to case outcomes and post-trial recidivism. Horizontal lines denote the corresponding 2SLS estimates and CATE, which averages the MTEs over the common support with equal weights, and the CATR, which averages the MTEs placing higher weights on defendant observably and unobservably more likely to be released. MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs.

Figure B.13: EM vs. Release MTEs for New Case Costs ($1,000)



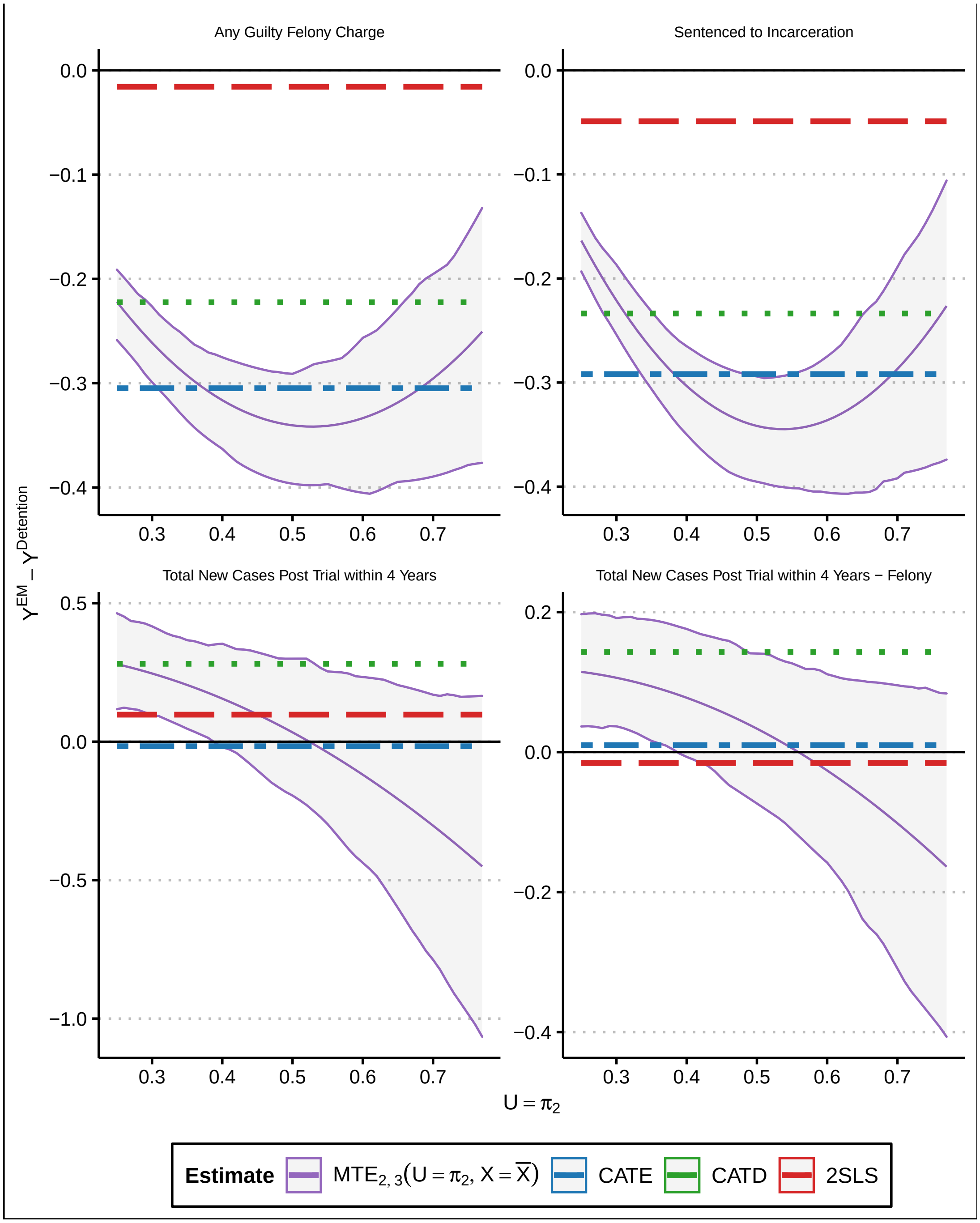
Note: Figure displays the marginal treatment effects (MTEs) across the distribution of defendant types (higher U means more unobservably resistant to higher treatment) for the effect of EM relative to Release for main felony sample cases. Outcome consist of total new case costs pre and post-trial using full and low-cost values for murder based on incidence costs from Miller et al. (2021). Horizontal lines denote the corresponding 2SLS estimates and CATE, which averages the MTEs over the common support with equal weights, and the CATR, which averages the MTEs placing higher weights on defendant observably and unobservably more likely to be released. MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs.

Figure B.14: EM vs. Detention MTEs for Pretrial Misconduct



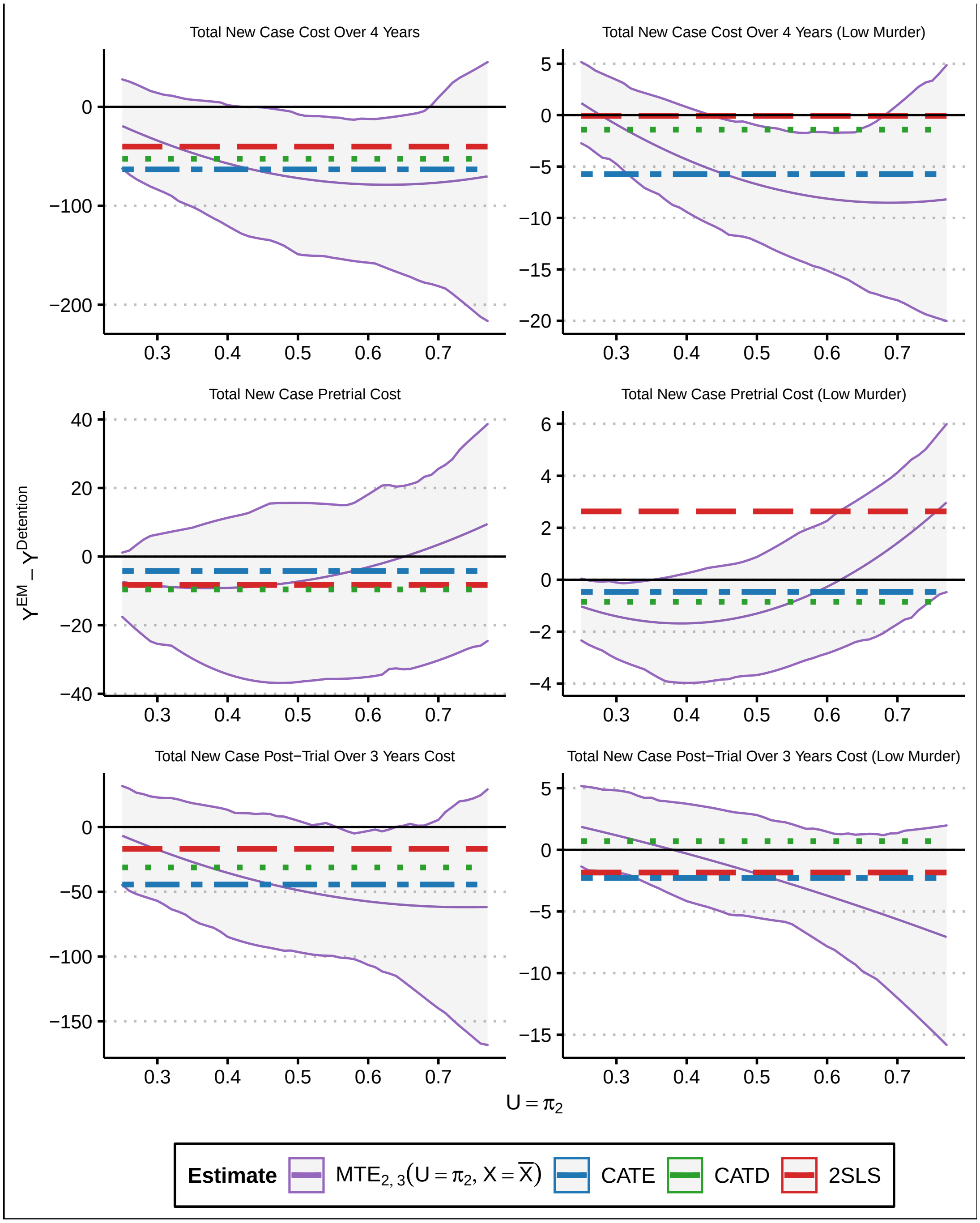
Note: Figure displays the marginal treatment effects (MTEs) across the distribution of defendant types (higher U means more unobservably resistant to higher treatment) for the effect of EM relative to Detention for main felony sample cases. Outcomes consist of variables relating to pretrial misconduct. Horizontal lines denote the corresponding 2SLS estimates and CATE, which averages the MTEs over the common support with equal weights, and the CATD, which averages the MTEs placing higher weights on defendant observably and unobservably more likely to be detained. MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs.

Figure B.15: EM vs. Detention MTEs for Case and Post-Trial Outcomes



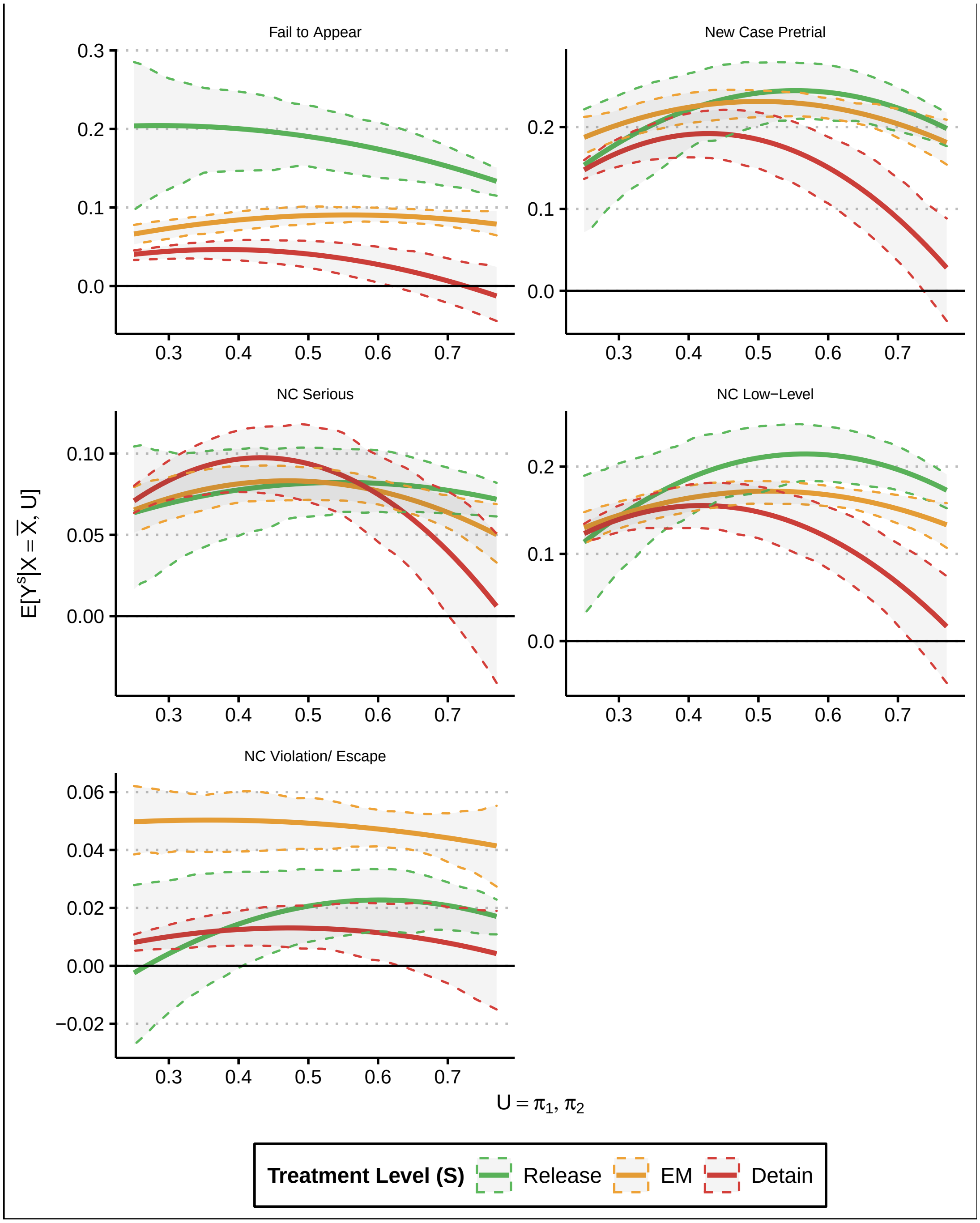
Note: Figure displays the marginal treatment effects (MTEs) across the distribution of defendant types (higher U means more unobservably resistant to higher treatment) for the effect of EM relative to Detention for main felony sample cases. Outcomes consist of variables relating to case outcomes and post-trial recidivism. Horizontal lines denote the corresponding 2SLS estimates and CATE, which averages the MTEs over the common support with equal weights, and the CATD, which averages the MTEs placing higher weights on defendant observably and unobservably more likely to be detained. MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs.

Figure B.16: EM vs. Detention MTEs for New Case Costs ($1,000)



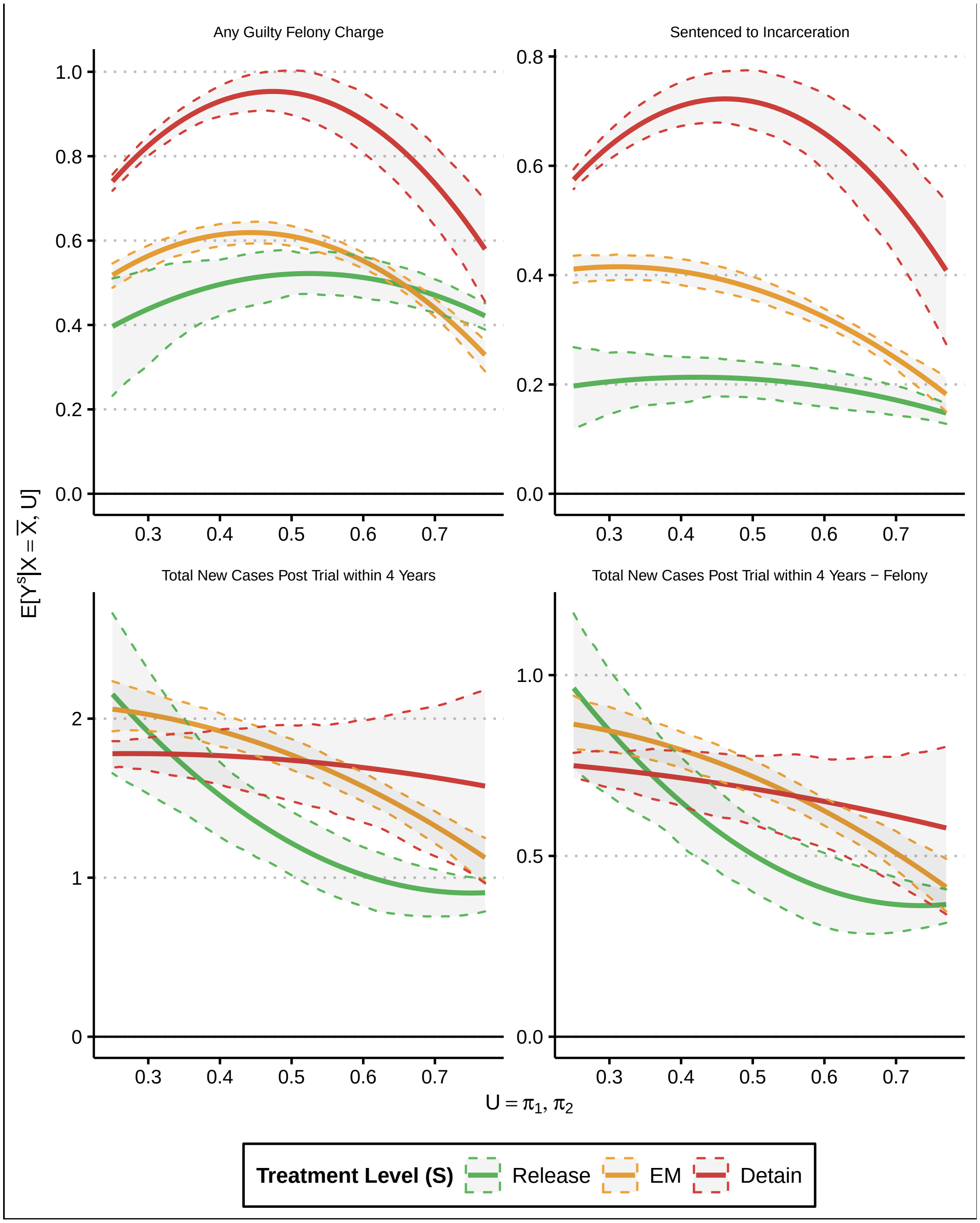
Note: Figure displays the marginal treatment effects (MTEs) across the distribution of defendant types (higher U means more unobservably resistant to higher treatment) for the effect of EM relative to Detention for main felony sample cases. Outcome consist of total new case costs pre and post-trial using full and low-cost values for murder based on incidence costs from Miller et al. (2021). Horizontal lines denote the corresponding 2SLS estimates and CATE, which averages the MTEs over the common support with equal weights, and the CATD, which averages the MTEs placing higher weights on defendant observably and unobservably more likely to be detained. MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs.

Figure B.17: MTRs for Pretrial Misconduct



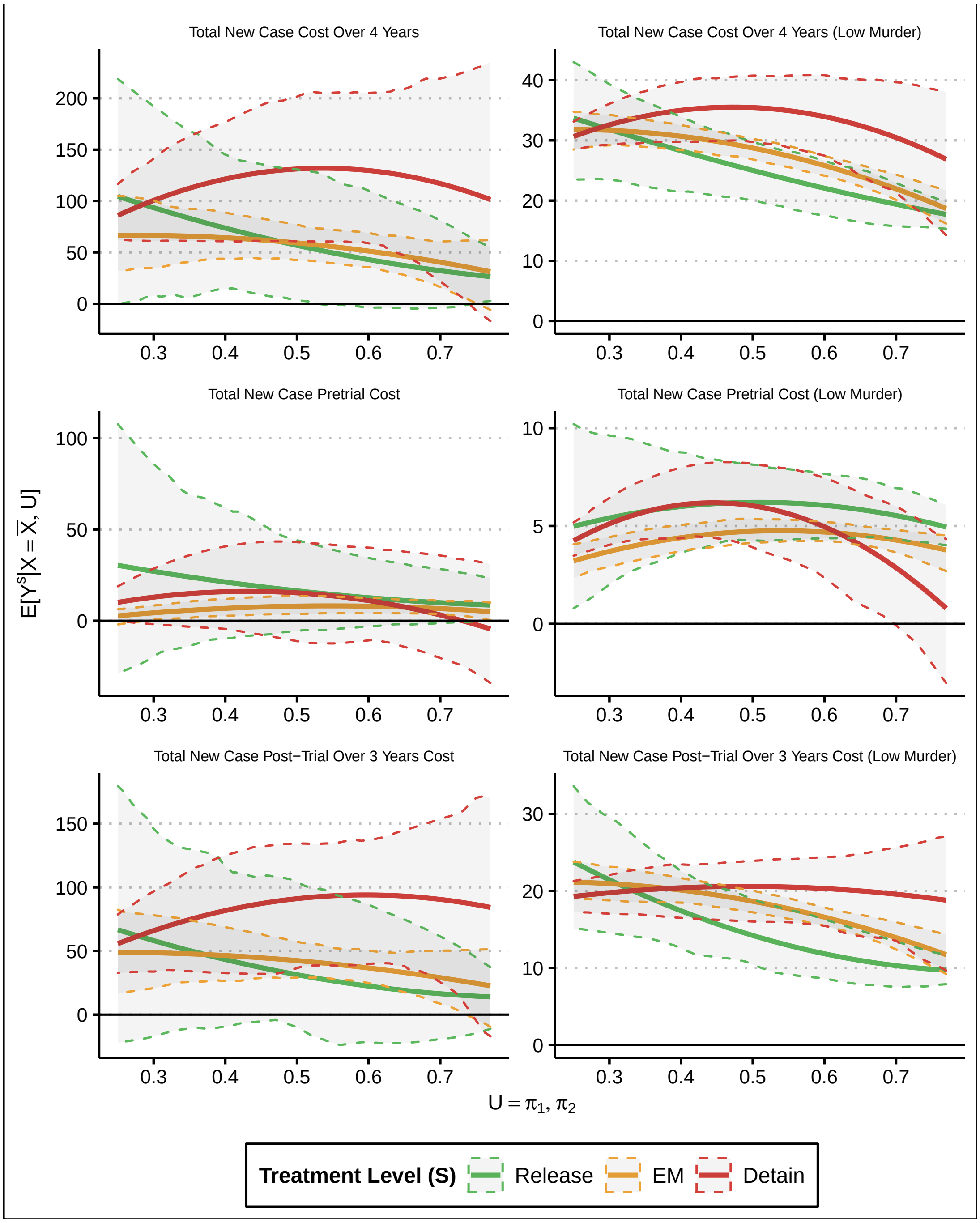
Note: Figure displays the marginal treatment response functions (MTR) across the distribution of defendant types (higher U means more unobservably resistant to higher treatment) for the expected response in each treatment level for main felony sample cases for the relevant outcomes. MTRs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs.

Figure B.18: MTRs for Case and Post-Trial Outcomes



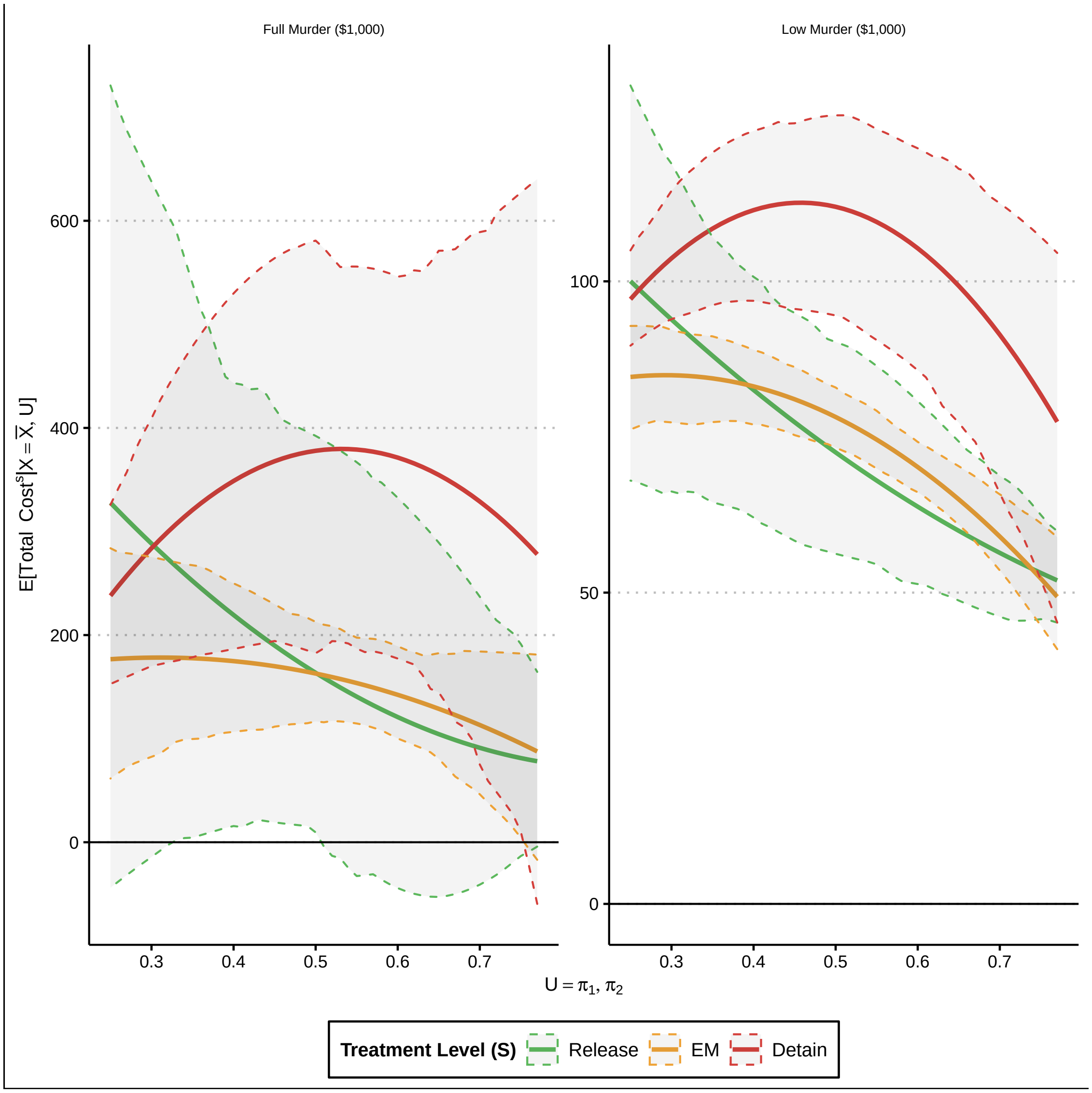
Note: Figure displays the marginal treatment response functions (MTR) across the distribution of defendant types (higher U means more unobservably resistant to higher treatment) for the expected response in each treatment level for main felony sample cases for the relevant outcomes. MTRs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs.

Figure B.19: MTRs for New Case Costs ($1,000)



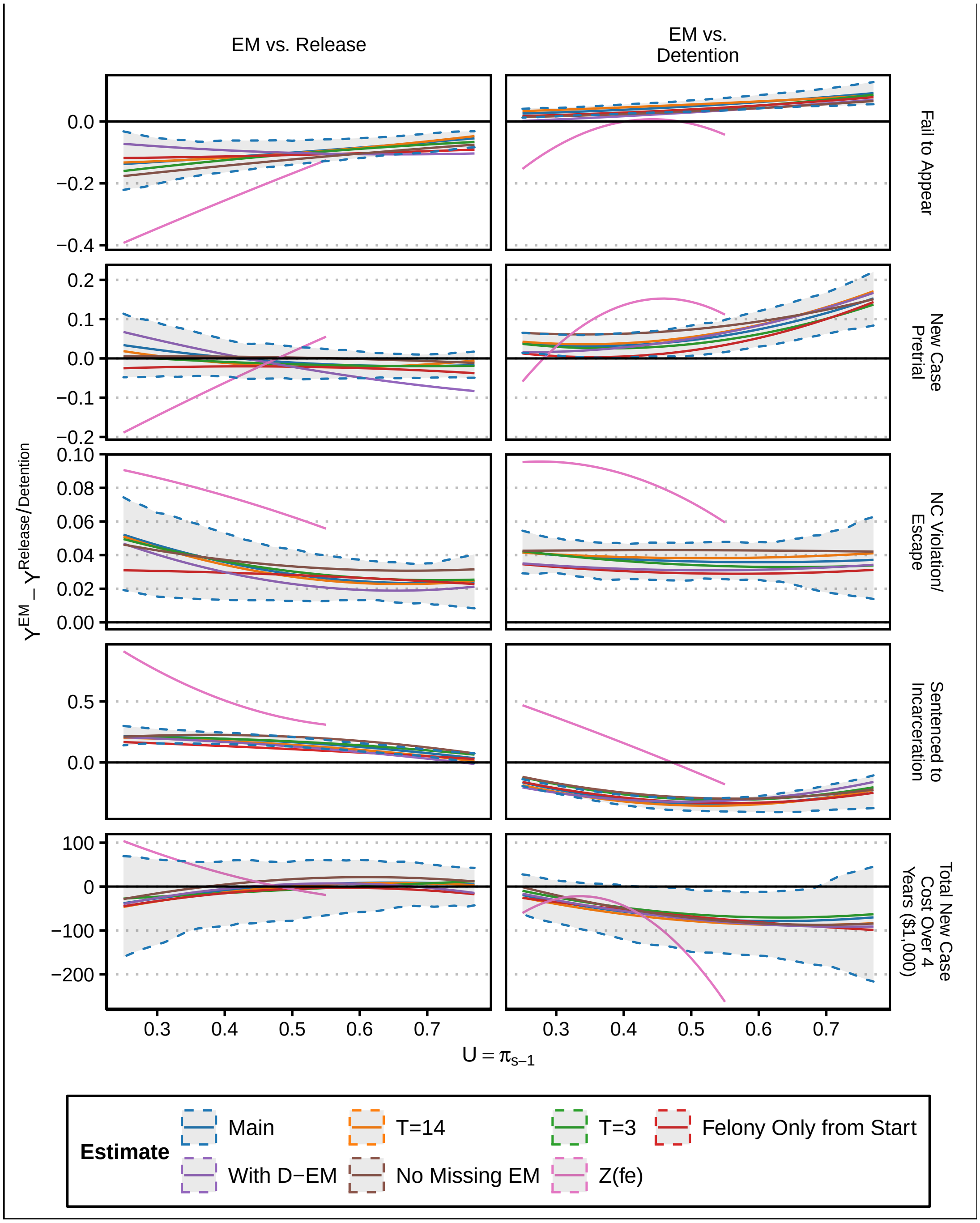
Note: Figure displays the marginal treatment response functions (MTR) across the distribution of defendant types (higher U means more unobservably resistant to higher treatment) for the expected response in each treatment level for main felony sample cases for the relevant outcomes. MTRs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s). 95% confidence intervals are computed using 200 bootstrap runs.

Figure B.20: MTRs for Total Costs ($1,000)



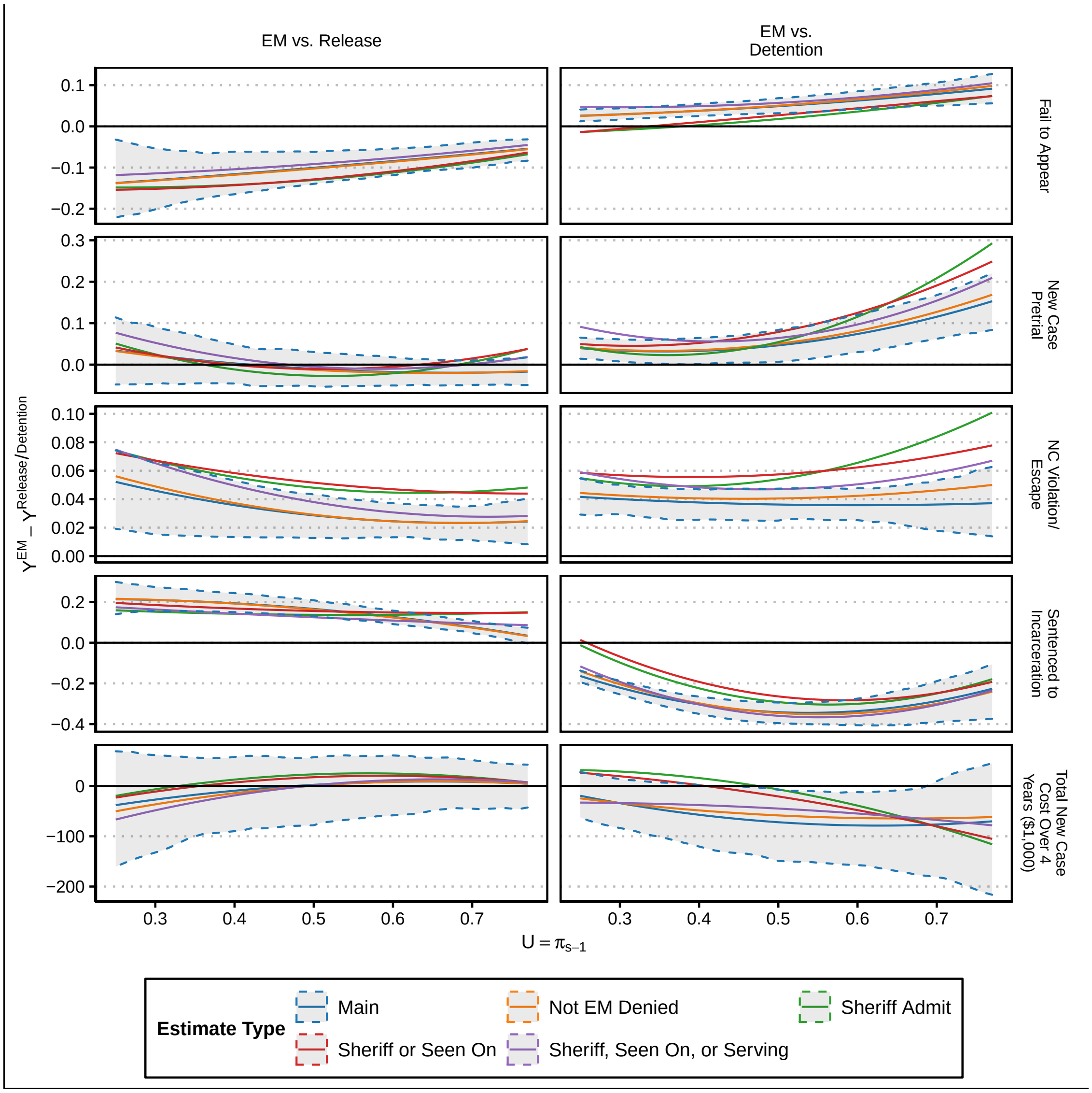
Note: Figure displays the marginal treatment response functions (MTR) across the distribution of defendant types (higher U means more unobservably resistant to higher treatment) for the expected total costs (including pre- and post-trial crime, sentencing, failures to appear, and direct costs) in each treatment level for main felony sample cases. MTRs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s) with weights and costs applied as discussed in Section 5.4. 95% confidence intervals are computed using 200 bootstrap runs.

Figure B.21: MTEs for Main Robustness Tests



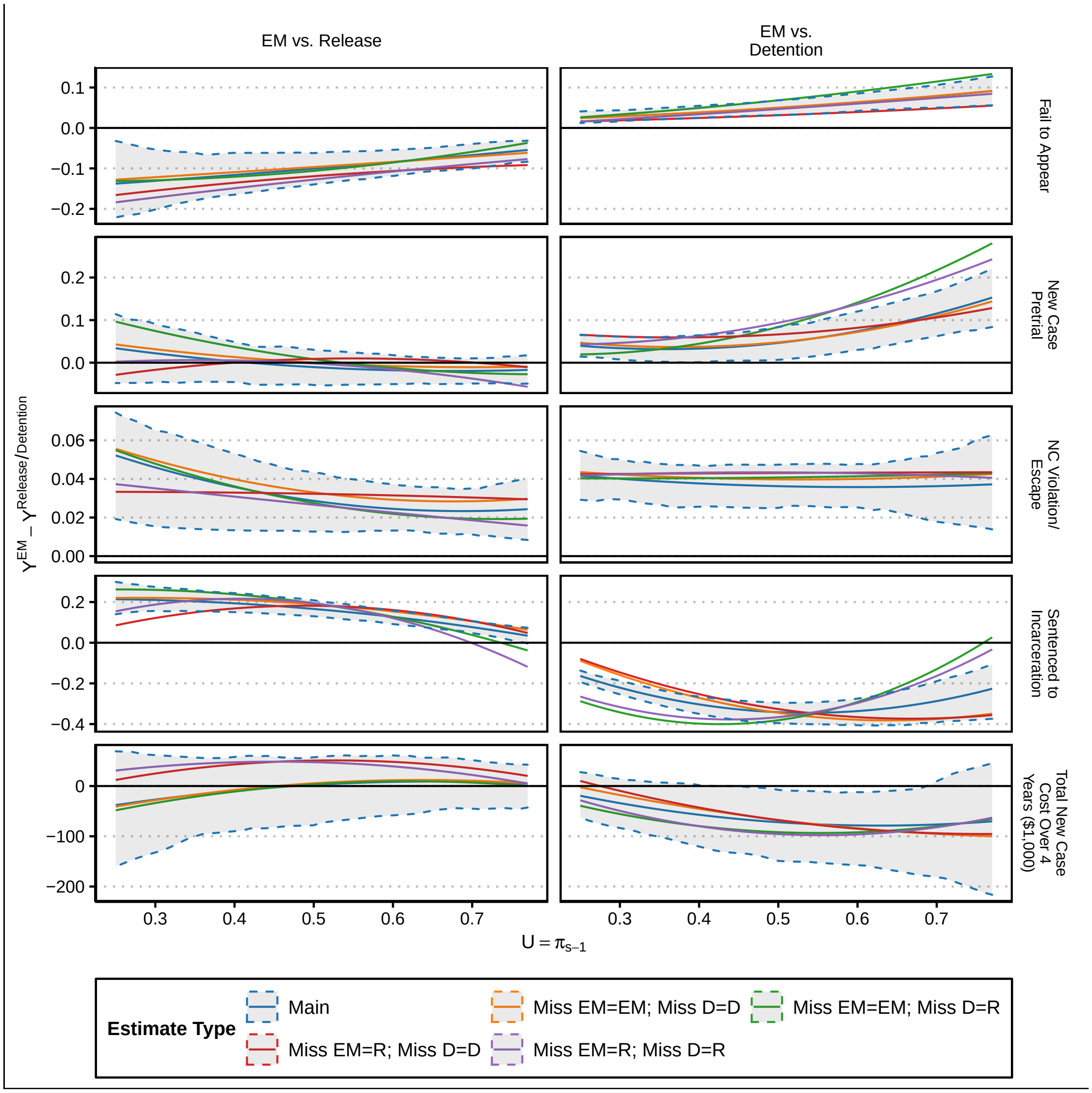
Note: Figure displays the marginal treatment effects (MTEs) for various robustness checks for the effect of EM relative to Release (left) and EM relative to Detention (right). MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s), unless otherwise specified. 95% confidence intervals of main estimates are computed using 200 bootstrap runs.

Figure B.22: MTEs for Recoded Treatments using Disposition Codes



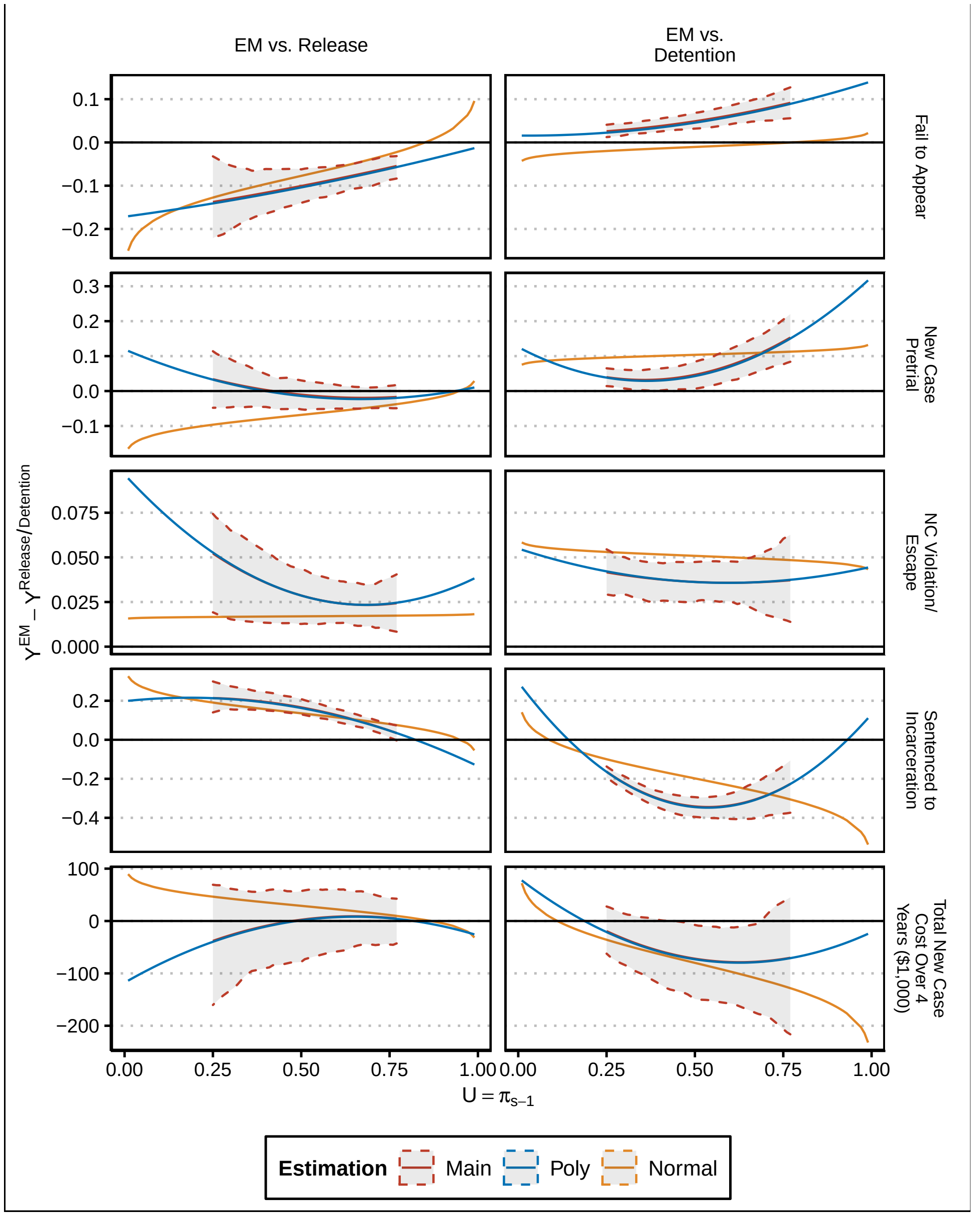
Note: Figure displays the marginal treatment effects (MTEs) for various robustness checks for the effect of EM relative to Release (left) and and EM relative to Detention (right) under different re-codings of pretrial treatments using disposition codes to determine if a defendant was assigned to EM. "Sheriff Admit" means the defendant was explicitly noted to have been admitted into the sheriff's EM program; "Sheriff or Seen On" allows for if the defendant was explicitly noted to be on EM; "Sheriff, Seen On, or Serving" allows for if the defendant was explicitly noted to be serving a monitoring program; and "Not EM Denied" includes "Sheriff, Seen On, or Serving" defendants but excludes any defendant explicitly noted to not be admitted to EM (with bail set to stand). MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s), unless otherwise specified. 95% confidence intervals of main estimates are computed using 200 bootstrap runs.

Figure B.23: MTEs for Recoded Missing Treatments



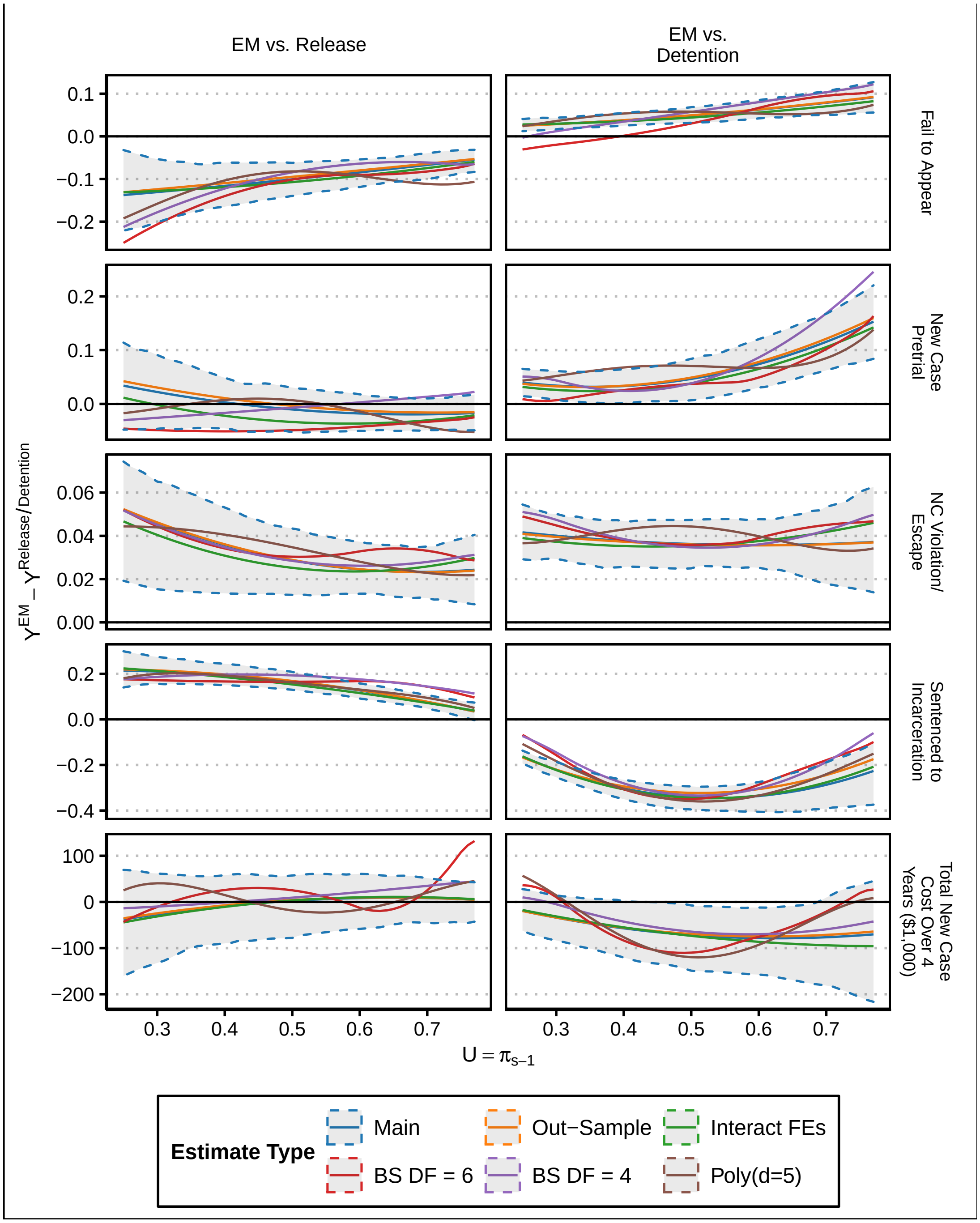
Note: Figure displays the marginal treatment effects (MTEs) for various robustness checks for the effect of EM relative to Release (left) and and EM relative to Detention (right) under different re-codings of pretrial treatments for cases with missing jail data. MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s), unless otherwise specified. 95% confidence intervals of main estimates are computed using 200 bootstrap runs.

Figure B.24: MTEs for Comparing Main, Jointly Normal, and Polynomial Models



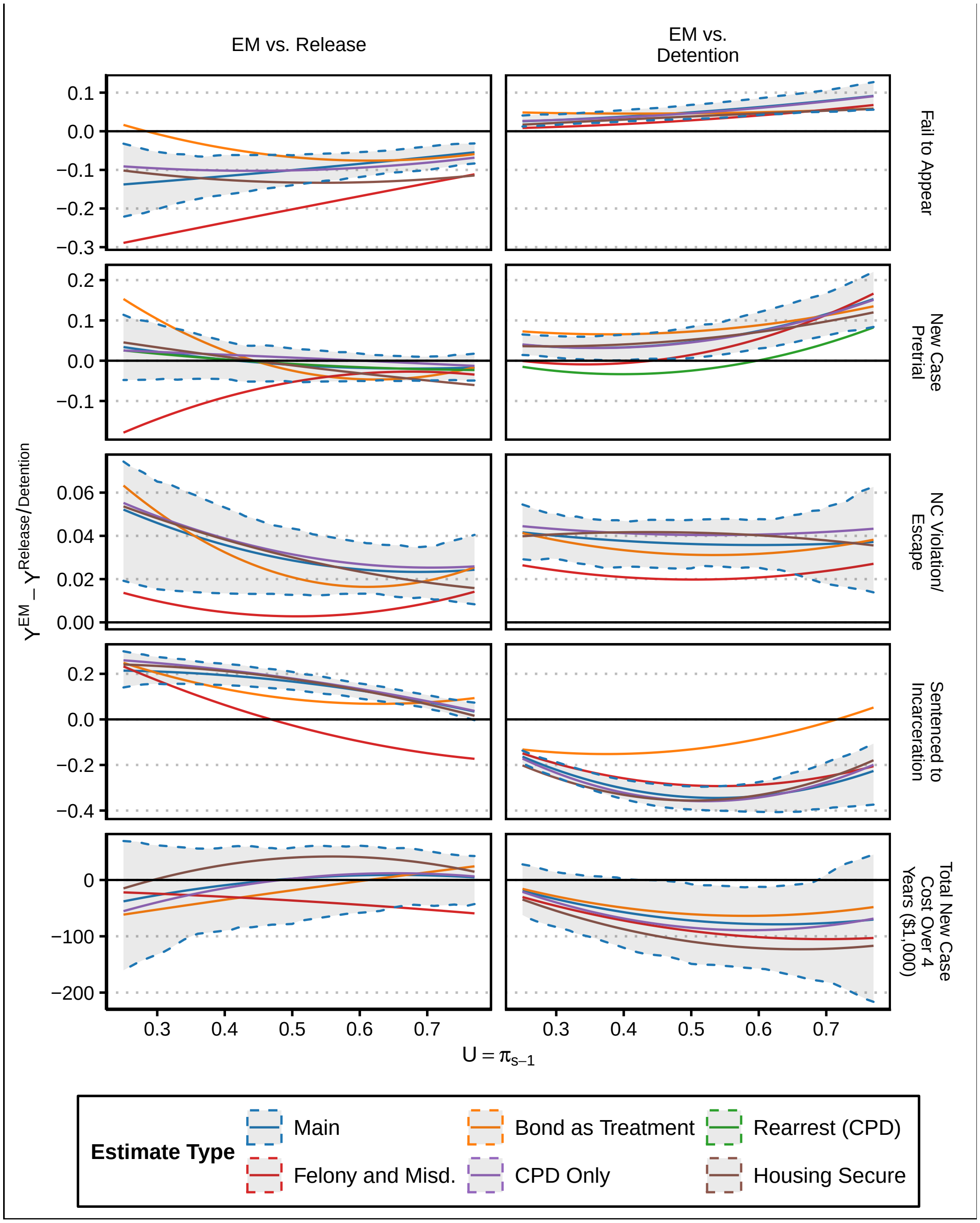
Note: Figure displays the marginal treatment effects (MTEs) across the distribution of defendant types (higher U means more unobservably resistant to higher treatment) assuming unobservables are drawn from a jointly normal distribution ('Normal'), following Cornelissen et al. (2018), the 'Main' specification over the common support (with confidence intervals), and a 3rd degree polynomials ('Poly') for Φ_s as a fully parametric model with full support.

Figure B.25: MTEs with Alternative Specifications



Note: Figure displays the marginal treatment effects (MTEs) for various robustness checks for the effect of EM relative to Release (left) and EM relative to Detention (right) for different specifications. MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s), unless otherwise specified. 95% confidence intervals of main estimates are computed using 200 bootstrap runs.

Figure B.26: MTEs with Alternative Samples



Note: Figure displays the marginal treatment effects (MTEs) for various robustness checks for the effect of EM relative to Release (left) and EM relative to Detention (right) for sub-samples of defendants. MTEs are recovered using the main semiparametric estimation method (equation (6) estimated with 3rd degree polynomial for Φ_s), unless otherwise specified. 95% confidence intervals of main estimates are computed using 200 bootstrap runs.

Table B.1: Summary Statistics for Branch 1 Cases by Pretrial Treatment for Main Felony Sample

	Release (1)	EM (2)	Detain (3)
Median Days in Jail	1	-	92
Median Days on EM	0	42	-
Median Case Duration	68	55	167
Bond Amount	20655	26914	88423
Defendant Demographics			
Def. Black	0.63	0.78	0.81
Def. Hispanic	0.17	0.1	0.1
Def. White	0.19	0.12	0.09
Def. Male	0.81	0.82	0.92
Def Age	35.85	37	32.62
Case Characteristics			
Charge - Felony	1	1	1
Charge - Felony Violent	0.03	0.02	0.15
Charge - Felony Drug Poss.	0.63	0.66	0.35
Charge - Felony Drug Deliv.	0.11	0.19	0.18
Charge - Felony Property	0.15	0.15	0.17
Charge - Felony Weapon	0.07	0.01	0.19
Charge - Misdemeannaor	0.22	0.18	0.24
Charge - Misd. Property	0.02	0.01	0.02
Charge - Traffic	0.16	0.1	0.08
Charge - Other	0.05	0.05	0.04
Case History (since 2000)			
Past Cases	5.38	8.01	9.09
Case within Year	0.29	0.42	0.54
Past Failure to Appear	1.18	1.85	2.03
Past Felonies	1.61	2.69	3
Past Felonies - Violent	0.07	0.09	0.22
Past Guilty	0.52	1.31	2.28
Past Guilty Felonies	0.19	0.57	1.07
Past Guilty Felonies - Violent	0.01	0.03	0.1
N Obs	13039	19523	18808
Share of Obs	0.25	0.38	0.37

Note: Table displays summary statistics for the main felony sample by pretrial treatment for defendants with one observation per defendant and arrest, meaning if there is a felony and a misdemeanor case against a defendant both sets of charges are aggregated to one observation. Variables beginning with 'Charge' are binary variables indicating any charge of a specific type.

Table B.2: Summary Statistics of Outcomes for Main Felony Sample

	Release (1)	EM (2)	Detain (3)
Failure to Appear	0.1220	0.0732	0.0292
Any Guilty Felony Charge	0.345	0.453	0.701
New Case Pretrial	0.171	0.187	0.109
New Case Post-Trial within 4 Years	1.09	1.56	1.58
New Felony Case Post-Trial within 4 Years	0.458	0.656	0.634
New Violent Felony Case Post-Trial within 4 Years	0.0328	0.0405	0.0594

Note: Table displays summary statistics of outcomes for the main felony sample by pretrial treatment for defendants with one observation per defendant and arrest, meaning if there is a felony and a misdemeanor case against a defendant both sets of charges are aggregated to one observation.

Table B.3: Outcomes Means for Main Felony Sample

Outcome	All	Release (S=1)	EM (S=2)	Detain (S=3)
Fail to Appear	0.069	0.122	0.073	0.029
New Case Pretrial	0.154	0.171	0.187	0.109
New Case Pretrial - Serious	0.058	0.061	0.063	0.051
New Case Pretrial - Low-Level	0.122	0.147	0.136	0.091
New Case Pretrial - Violation/ Escape	0.021	0.013	0.042	0.005
Any Guilty Felony Charge	0.517	0.345	0.453	0.701
Sentenced to Incarceration	0.363	0.139	0.284	0.601
Total New Cases Post Trial within 4 Years	1.450	1.090	1.560	1.580
Total New Cases Post Trial within 4 Years - Felony	0.598	0.458	0.657	0.634
Total New Case Cost Over 4 Years	58.100	44.900	49.600	75.900
Total New Case Cost Over 4 Years (Low Murder)	24.200	19.900	25.000	26.300
Total New Case Pretrial Cost	7.190	8.740	5.580	7.780
Total New Case Pretrial Cost (Low Murder)	3.580	4.190	3.680	3.050
Total New Case Post-Trial Over 3 Years Cost	38.700	28.300	35.400	49.400
Total New Case Post-Trial Over 3 Years Cost (Low Murder)	15.500	12.400	16.400	16.700

Note: Table displays summary statistics of all outcomes for the main felony sample for all treatments and each pretrial treatment for defendants.

Table B.4: Additional Instrument Tests

	Release	Detain
Hansen Over-Identification		
J-Statistic	445.088	446.462
P-value	<0.001	<0.001
Sanderson-Windmeijer F-Test		
F-statistic	28.67	43.28
P-value	<0.001	<0.001
Montiel Olea - Pflueger Weak Instrument		
Eff. Fstat	35.793	54.304
2SLS Critical Value (5% of Worst Case Bias)	22.593	22.692
Anderson-Rubin Weak Instrument		
F-statistic	5.2	
P-Value	<0.001	

Note: Table displays results for additional instrument tests. Eff. Fstat refers to the effective F-statistic from the Montiel Olea and Pflueger (2013) weak instruments test.

Table B.5: Reduced Form

	$\hat{Z}^R$ (1 SD)	$\hat{Z}^D$ (1 SD)	F-Stat
	(1)	(2)	(3)
Pretrial Misconduct			
Fail to Appear	0.0042***	-0.0079***	51.88
	(0.00135)	(0.00111)	
New Case Pretrial	-0.0072***	-0.018***	51.069
	(0.00169)	(0.00179)	
New Case Pretrial - Serious	-0.0018*	-0.0062***	14.472
	(0.00106)	(0.00118)	
New Case Pretrial - Low-Level	-0.0036**	-0.0095***	18.212
	(0.00154)	(0.00158)	
New Case Pretrial - Violation/ Escape	-0.004***	-0.0081***	43.918
	(0.00072)	(0.00086)	
Case Outcomes			
Any Guilty Felony Charge	-0.0022	0.0018	2.502
	(0.00185)	(0.00152)	
Sentenced to Incarceration	0.0015	0.0054***	4.808
	(0.00165)	(0.00178)	
Post-Trial New Cases			
Total New Cases Post Trial within 4 Years	-0.014	-0.01	1.055
	(0.00999)	(0.01094)	
Total New Cases Post Trial within 4 Years - Felony	-0.0014	0.0018	0.216
	(0.00466)	(0.00486)	
Total Case Cost (Pre and Post) ($1,000)			
Total New Case Cost Over 4 Years	610	4400**	2.159
	(2701.76516)	(2174.83872)	
Total New Case Pretrial Cost	1600**	870	2.03
	(824.39783)	(725.40152)	
Total New Case Post-Trial Over 3 Years Cost	-1500	1900	1.141
	(2218.11026)	(1637.11968)	

Note: Table displays the results of reduced form regression, regressing the outcome on the two instruments, controlling for case observables and quarter and day of week fixed effects. Columns (1) and (2) contain the point estimate for the respective standardized instrument, and Column (3) contains the f-statistic from the projected model which already accounts for the influence of case observables and fixed effects. Standard errors clustered at the defendant level are in parantheses. ***p $<$ 0.01; **p $<$ 0.05; *p $<$ 0.1

Table B.6: Average Monotonicity Test

	Release	Detain	Release	Detain
	(1)	(2)	(3)	(4)
Full Sample				
$Z^{Release}$	1***	0	1***	0
	(0.00954)	(0.00918)	(0.01938)	(0.01616)
Z^{Detain}	0	1***	0	1***
	(0.00862)	(0.00816)	(0.01595)	(0.01611)
Felony Sample				
$Z^{Release}$	0.73***	0.15***	1***	0.081***
	(0.0142)	(0.01408)	(0.02327)	(0.01897)
Z^{Detain}	-0.081***	1.1***	-0.00016	1.1***
	(0.00942)	(0.00983)	(0.01678)	(0.01856)
No Felony Sample				
$Z^{Release}$	1.1***	-0.16***	1.1***	-0.13***
	(0.0265)	(0.02308)	(0.03439)	(0.03002)
Z^{Detain}	0.013	0.73***	-0.047	0.86***
	(0.02839)	(0.02543)	(0.0395)	(0.03571)
Drug Charge Sample				
$Z^{Release}$	0.78***	0.14***	1***	0.1***
	(0.01994)	(0.01811)	(0.02773)	(0.02171)
Z^{Detain}	-0.028	1.1***	0.06***	1.1***
	(0.01784)	(0.01789)	(0.02259)	(0.02276)
Non-Black Sample				
$Z^{Release}$	0.96***	-0.07***	1***	-0.051*
	(0.02024)	(0.01724)	(0.03553)	(0.02881)
Z^{Detain}	0.0039	0.84***	0.026	0.85***
	(0.01871)	(0.01649)	(0.04073)	(0.03724)
Controls			X	X
Month + DoW Fes	X	X	X	X

Note: Table displays results of regressing the endogenous treatment (release or detain) on both instrumented treatment assignments with fixed effects and with controls in Columns (3) and (4) across predetermined subsamples. Standard errors clustered at the defendant level are in parantheses. ***p < 0.01; **p < 0.05; *p < 0.1

C Appendix C: General Results for MTEs

Marginal treatment effects (MTE) provide a structure for understanding how treatments affect individuals differently based on their unobservable characteristics. Generally, this involves estimating the effect of a treatment for defendants ranked according to their unobservable ‘resistance’ to treatment. In the binary case, this is straightforward, and there has been significant work on estimating MTEs for binary treatments.

When multiple ordered treatments are introduced, identification becomes more complicated theoretically and more difficult to estimate empirically. Prior work tends to make simplifying assumptions about the joint distribution of unobservables; following Heckman, Urzua, and Vytlacil (2006) (HUV) and Heckman and Vytlacil (2007) (HV), this generally a jointly normal distribution (as in Cornelissen et al. (2018)) which results in MTEs with predetermined shapes (linear for most values of u). Rose and Shem-Tov (2021) is an exception, as they extend Mogstad, Santos, and Torgovitsky (2018)’s bounding method to the case of ordered treatments by using discrete cutoffs as instruments to construct candidate marginal treatment response functions (MTRs) complying to specific shape restrictions, with the goal being bounding average treatment effects. In contrast, I focus on a general case, extending HUV, to point-identify TSMTEs and provide a method to estimate MTRs semiparametrically, relying on an interval of common support.

This section will focus on the case of estimating marginal treatment effects for ordered multi-valued treatments. The ordering of the treatments can be either cardinal — such as a dose-response function — or ordinal — where the treatments increase in intensity but have no clear quantitative distance. This section follows the work of HUV and HV closely. Building on their work, I relax one of HUV and HV’s identification assumptions. The main contribution is to provide a tractable method for semiparametric estimation of MTRs.

C.1 Set up

Consider an individual (i, though individual subscripts are suppressed for brevity) either choosing between or being assigned to one of $\bar{S}$ different 'levels' of treatment which can be ordered by their intensity and given ranks such that $S \in \mathcal{S} = \{1, ..., \bar{S}\}$. For example, in the context of this paper, the individual is a defendant and the treatment levels are pretrial release ($S = 1$), EM ($S = 2$), and detention ($S = 3$), so $\bar{S} = 3$, but the levels could correspond to medicine dosage, years of schooling, or intensity of a social service intervention. Each individual has some set of observable (to the econometrician) characteristics X, as well as unobservable features ($\{\omega_1, ..., \omega_{\bar{S}}, V\}$). The unobservable factor V is observed by the agent who determines treatment.

Potential Outcomes For some outcome of interest, if the individual were assigned to treatment level $S = s$, their *potential* outcome is $Y_s = \mu_s(X, \omega_s)$. μ_s is some treatment-specific function, and X and ω_s is the individual-specific observable and unobservable components which contributes to Y_s. Let $D_s = 1$ if the defendant received treatment s, and $D_s = 0$ otherwise. From this, we know the observed outcome for an individual is:

$$Y = \sum_{s=1}^{\bar{S}} Y_s \times D_s = \sum_{s=1}^{\bar{S}} \mu_s(X, \omega_s) \times D_s.$$

Selection into Treatment The treatment level received by an individual is determined by a single index crossing a series of thresholds. The index, $T(Z, X)$, can be interpreted as the individuals 'net benefit' from a higher level of treatment (in the eyes of the agent assigning their treatment, possibly the individual themselves). The individual receives a treatment level higher than $S > s$ (e.g., $s + 1, s + 2, \dots$) if and only if $T(Z, X) > C^s(W)$, where $C^s(W)$ is the cutoff value that is the highest level of $T(Z, X)$ which will result in being assigned to treatment level s. As before, X denotes observables that influence potential outcomes and possibly selection, while Z and W are observable instruments which do not directly influence outcomes but do influence selection either through the index (Z) or cutoffs (W).

From this, an individual is assigned to treatment level s if and only if:

$$D_s = 1(S = s) = 1[C^s(W) \geq T(Z, X) > C^{s-1}(W)]$$

where $C^{\bar{S}} = +\infty$ and $C^0 = -\infty$, because no one can receive a treatment higher than the highest level or lower than the lowest level (1), and $C^{s-1}(W) \leq C^s(W)\ \forall s$.

I assume that the single index can be decomposed into $T(Z, X) = \tau(Z, X) - V$ where $\tau(Z, X)$ is the individual's 'benefit' from a higher level of treatment and V is the individual's resistance to (or cost of) a higher level of treatment. The single index assumption has multiple implications. First, this separable form implies monotonicity, because for all individuals going from $\tau(Z, X)$ to $\tau(Z', X)$ will shift $T(\cdot, X)$ in the same way (similarly for if Z is fixed and X changes), and thus (weakly) move S in the same direction (HUV). Second, the existence of a single index, V, that determines treatment (conditional on observables) means that all factors can be reduced down to a single dimension in determining which treatment is optimal (to the agent deciding).

With this model, we seek to identify the effect of being assigned to treatment $s + 1$ relative to treatment s for an individual with resistance level v and observables x. This is the transition-specific marginal treatment effect as coined in HV (TSMTE):

$$MTE_{s+1,s}(x, v) = \mathbb{E}[Y_{s+1} - Y_s | V = v, X = x]$$

C.2 Identification

In order to identify transition-specific marginal treatment effects, I assume the following assumptions from HUV (denoted HUV 1-6, though called OC 1-6 in the original paper):

Assumption 1. *HUV1:* $(\omega_s, V) \perp (Z, W)$ *for all* $s \in \mathcal{S}$ *conditional on* X*.*

Assumption 2. *HUV2:* $\tau(Z, X)$ *is a non-degenerate random variable conditional on* X *and* W*.*

Assumptions 1 and 2 ensure that the instruments are valid and relevant after conditioning on regressors.

Assumption 3. *HUV3: The distribution of V is absolutely continuous conditional on X.*

From Assumption 3, we can use probability integral transformation to get $U = F_V(V|X = x)$ which is uniformly distributed $U \sim Unif[0,1]$, conditional on X:

$$\begin{aligned} D_s &= 1[C^s(W) \geq T(Z,X) > C^{s-1}(W)] \\ &= 1[F_V(\tau(Z,X) - C^s(W)) \leq F_V(V) < F_V(\tau(Z,X) - C^{s-1}(W))] \\ &= 1[F_V(\tau(Z,X) - C^s(W)) \leq U < F_V(\tau(Z,X) - C^{s-1}(W))] \end{aligned}$$

Let $\pi_s(Z,X,W) = F_V(C^s(W) - \tau(Z,X)) = \Pr(S > s|Z,X,W)$. By construction, $\pi_0(Z,X,W) = 1$ and $\pi_{\bar{S}}(Z,X,W) = 0$. Then the selection equation becomes: $D_s = 1[\pi_s(Z,X,W) \leq U < \pi_{s-1}(Z,X,W)]$. With this, we can redefine the TSMTE with the selection unobservable being in terms of U (with a known distribution) rather than V (with an unknown distribution):

$$MTE_{s+1,s}(x,u) = \mathbb{E}[Y_{s+1} - Y_s|U = u, X = x]$$

Assumption 4. *HUV4:* $E|Y_s| < \infty \ \forall s \in S$

Assumption 5. *HUV5:* $0 < \Pr(S = s|X) < 1 \ \forall s \in S$

Assumption 6. *HUV6: The distribution of $C^s(W)$ conditional on X and Z and other $C^{s'}$ is non-degenerate and continuous $\forall s \in \{1, ..., \bar{S} - 1\}$.*

With these assumptions, HUV and HV show that the TSMTE is identified by taking: $\frac{\partial \mathbb{E}[Y|\pi(Z,X,W)=\pi, X=x]}{\partial \pi_s} = MTE_{s+1,s}(u,x) = \mathbb{E}[Y_{s+1} - Y_s|U = u, X = x]$, where $\pi = [\pi_1, ..., \pi_{\bar{S}-1}]$.

However, this introduces complications that make semiparametric or nonparametric estimation of such a model difficult, particularly with more than 3 treatments, for two main reasons. First, Assumption 6 effectively requires variation in the C^s's conditional on all other $C^{s'}$ $\forall s' \in \{1, ..., \bar{S} - 1\} \setminus s$ and observables wherever the TSMTE is to be identified. This means if there are $\bar{S} = 6$ treatment levels, then wherever one wishes to estimate $MTE_{4,3}(u)$, there must be variation in all 5 π_s's. This may very well not be the case, for example, at high u comparing $S = 4$ to $S = 3$, π_1 or π_2 may be degenerate (or effectively so in the data). In this scenario, Assumption 6 would not hold.

Second, the form $\frac{\partial \mathbb{E}[Y|\pi(Z,X,W)=\pi]}{\partial \pi_s}$ conditions on vector $\pi = [\pi_1, ..., \pi_{\bar{S}-1}]$ for estimation and taking a partial derivative of the function $\mathbb{E}[Y|\pi]$ with respect to the specific π_s of interest, and it does not allow us to recover treatment responses only treatment effects. Assuming both ω_s and U are drawn from a jointly normal distribution is a common method for estimation, though this fully parametric assumption leads to effectively linear TSMTEs for most values of u.

I provide an alternative identification method and a weaker assumption to replace Assumption 6, which improves upon both of these limitations. First, rather than recovering TSMTEs through a local-IV approach, we can recover TSMTEs as the difference between marginal treatment response (MTR) functions (Carneiro and Lee (2009), Brinch, Mogstad, and Wiswall (2017), Mogstad, Santos, and Torgovitsky (2018), Rose and Shem-Tov (2021)) at a fixed value of $X = x$ and $U = u$:[57]

$$MTE_{s+1,s}(u,x) = \mathbb{E}[Y_{s+1}|U = u, X = x] - \mathbb{E}[Y_s|U = u, X = x]$$

Identification of an MTR (e.g., $\mathbb{E}[Y_s|U,X]$) can be achieved relying solely on variation in adjacent π_s's (e.g., π_s and π_{s-1}). Specifically, I provide an weaker alternative to Assumption 6, which both reduces the required variation for identification of TSMTEs when $\bar{S} > 3$ and allows the identification of MTRs and simpler estimation using the separate approach:

Assumption 7. *For all $s \in \mathcal{S}\backslash\{1, \bar{S}\}$, the joint distribution of $\pi_s(Z,X,W)$ and $\pi_{s-1}(Z,X,W)$ conditional on X is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^2$. Furthermore, the joint distribution of $\pi_s(Z,X,W)$ and $\pi_{s-1}(Z,X,W)$ conditional on X is*

[57] The framework stems from the literature on estimating the marginal distributions of potential outcomes in Imbens and Rubin (1997) and Abadie (2003).

non-degenerate in the sense that its support cannot be reduced to a subset on $\mathbb{R}$.

Assumption 7 improves upon Assumption 6. First, because it only makes assumptions on the joint distribution of π_s and π_{s-1}, simply requiring that that they are not highly codependent conditional on X — in the language of Assumption 6, for $MTE_{s+1,s}$, C^s only need be non-degenerate and continuous conditional on C^{s+1}, C^{s-1}, and X. Second, this assumption lends itself to a simple semiparametric estimation approach, as will be discussed below, and thus is more feasible for applications in applied work. TSMTEs are then the difference between MTRs, and MTRs are identified under Assumptions (1)-(5) and (7):

Theorem 1. *Let Assumptions (1)-(5), and (7) hold. Then,*

$$\mathbb{E}[Y_1|U=u,X=x] = -\frac{\partial \mathbb{E}[Y \times D_1|\pi_1(Z,X,W)=\pi_1, X=x]}{\partial \pi_1}\bigg|_{\pi_1=u}$$

$$\mathbb{E}[Y_{\bar{S}}|U=u,X=x] = \frac{\partial \mathbb{E}[Y \times D_{\bar{S}}|\pi_{\bar{S}-1}(Z,X,W)=\pi_{\bar{S}-1}, X=x]}{\partial \pi_{\bar{S}-1}}\bigg|_{\pi_{\bar{S}-1}=u}$$

And, for all $s = 2, \ldots, \bar{S}-1$,

$$\mathbb{E}[Y_s|U=u,X=x] = \frac{\partial \mathbb{E}[Y \times D_s|\pi_{s-1}(Z,X,W)=\pi_{s-1}, \pi_s(Z,X,W)=\pi_s, X=x]}{\partial \pi_{s-1}}\bigg|_{\pi_{s-1}=u}$$

$$= -\frac{\partial \mathbb{E}[Y \times D_s|\pi_{s-1}(Z,X,W)=\pi_{s-1}, \pi_s(Z,X,W)=\pi_s, X=x]}{\partial \pi_s}\bigg|_{\pi_s=u}.$$

Proof of Theorem 1. Using Assumptions (1)-(5) and (7), write (suppressing W):

$$\begin{aligned}
&\mathbb{E}[Y \times D_s|Z=z,X=x,W=w] \\
&= \mathbb{E}[\mu_s(X,\omega_s)1\{\pi_{s-1}(Z,X,W) > U \geq \pi_s(Z,X,W)\}|Z=z,X=x,W=w] \\
&= \mathbb{E}[\mu_s(X,\omega_s)1\{\pi_{s-1}(Z,X,W) > U \geq \pi_s(Z,X,W)\}|X=x] \text{ by (1)} \\
&= \int_{\pi_s(z,x,w)}^{\pi_{s-1}(z,x,w)} \mathbb{E}[Y_s|U=u,X=x]du \text{ by (3)}
\end{aligned}$$

Then, because

$$\mathbb{E}[Y \times D_s | Z = z, X = x, W = w] = \int_{\pi_s(z,x,w)}^{\pi_{s-1}(z,x,w)} \mathbb{E}[Y_s | U = u, X = x] du,$$

taking the derivative of both sides with respect to π_{s-1} (when $s > 1$) gives:

$$\left.\frac{\partial \mathbb{E}[Y \times D_s | Z = z, X = x, W = w]}{\partial \pi_{s-1}}\right|_{\pi_{s-1} = u} = \mathbb{E}[Y_s | U = u, X = x],$$

and similarly taking the derivative of both sides with respect to π_s gives:

$$\left.\frac{\partial \mathbb{E}[Y 1(S = s) | Z = z, X = x, W = w]}{\partial \pi_s}\right|_{\pi_s = u} = -\mathbb{E}[Y_s | U = u, X = x].$$

Because $D_s = 1(S = s) = 1[\pi_{s-1}(Z, X, W) > U \geq \pi_s(Z, X, W)]$ and by Assumption 7, only variation in the adjacent π's are relevant. So:

$$\begin{aligned} &\left.\frac{\partial \mathbb{E}[Y \times D_s | Z = z, X = x, W = w]}{\partial \pi_{s-1}}\right|_{\pi_{s-1} = u} \\ = &\left.\frac{\partial \mathbb{E}[Y \times D_s | \pi_{s-1}(Z, X, W) = \pi_{s-1}, \pi_s(Z, X, W) = \pi_s, X = x]}{\partial \pi_{s-1}}\right|_{\pi_{s-1} = u} \end{aligned}$$

and

$$\begin{aligned} &\left.\frac{\partial \mathbb{E}[Y \times D_s | Z = z, X = x, W = w]}{\partial \pi_s}\right|_{\pi_s = u} \\ = &\left.\frac{\partial \mathbb{E}[Y \times D_s | \pi_{s-1}(Z, X, W) = \pi_{s-1}, \pi_s(Z, X, W) = \pi_s, X = x]}{\partial \pi_s}\right|_{\pi_s = u} \end{aligned}$$

Finally, we have for s with non-degenerate π_s and π_{s-1}:

$$\begin{aligned}
\mathbb{E}[Y_s|U=u,X=x] &= \left.\frac{\partial \mathbb{E}[Y \times D_s|\pi_{s-1}(Z,X,W)=\pi_{s-1},\pi_s(Z,X,W)=\pi_s,X=x]}{\partial \pi_{s-1}}\right|_{\pi_{s-1}=u} \\
&= -\left.\frac{\partial \mathbb{E}[Y \times D_s|\pi_{s-1}(Z,X,W)=\pi_{s-1},\pi_s(Z,X,W)=\pi_s,X=x]}{\partial \pi_s}\right|_{\pi_s=u}.
\end{aligned}$$

$\square$

Assumptions 2, 4, and 5 ensure values in the population are well-defined, while Assumption 7 ensures variation in adjacent π_s's.

With this, we can identify the TSMTE between levels s and $s+1$ by identifying the conditional means (MTRs) for s and $s+1$:

$$MTE_{s+1,s}(x,u) = \mathbb{E}[Y_{s+1} - Y_s|U=u,X=x] = \mathbb{E}[Y_{s+1}|U=u,X=x] - \mathbb{E}[Y_s|U=u,X=x]$$

with

$$\begin{aligned}
&\mathbb{E}[Y_s|U=u,X=x] \\
&= \left.\frac{\partial \mathbb{E}[Y_s \times D_s|\pi_{s-1}(Z,X,W)=\pi_{s-1},\pi_s(Z,X,W)=\pi_s,X=x]}{\partial \pi_{s-1}}\right|_{\pi_{s-1}=u} \\
&= -\left.\frac{\partial \mathbb{E}[Y_s \times D_s|\pi_{s-1}(Z,X,W)=\pi_{s-1},\pi_s(Z,X,W)=\pi_s,X=x]}{\partial \pi_s}\right|_{\pi_s=u}.
\end{aligned} \tag{7}$$

And this equality applies only to $1 < S < \bar{S}$ — so for intermediate treatment levels, $\mathbb{E}[Y_s|U=u,X=x]$ is over-identified.

While $\mathbb{E}[Y_s \times D_s|\pi_{s-1}(Z,X,W)=\pi_{s-1},\pi_s(Z,X,W)=\pi_s,X=x]$ can be estimated non-parametrically, in practice the data requirements make such estimation are rarely feasible, and semiparametric estimation is often the preferred approach in practice in the MTE literature.

For semiparametric estimation, I assume that $\mu_s(X,\omega_s)$ is composed of additively separable functions of X and ω_s, essentially that across all values of covariates, the effect of unobservables works the same and it allows for treatment effects on observables and unobservables separately (Andresen (2018)). This assumption (either directly or as a result of full independence) is common in the literature (Carneiro, Heckman, and Vytlacil (2011), Kline and Walters (2016), Brinch, Mogstad, and Wiswall (2017), Bhuller et al. (2018), Rose and Shem-Tov (2021)). Specifically, I assume: $Y_s = \beta_s X + \omega_s$ and $\mathbb{E}[Y_s|x,u] = \beta_s x + \mathbb{E}[\omega_s|U=u]$. Then, the marginal treatment effect of moving from one treatment to the next highest one (s to $s+1$) is:

$$\begin{aligned} MTE_{s+1,s}(U=u,X=x) &= \mathbb{E}[Y_{s+1} - Y_s|U=u,X=x] \\ &= (\beta_{s+1} - \beta_s)x + \mathbb{E}[\omega_{s+1}|U=u] - \mathbb{E}[\omega_s|U=u] \end{aligned}$$

C.3 Semi-Parametric Estimation of MTRs

The following section will provide simple functional form assumptions and an accompanying semiparametric estimation procedure. For notational purposes, I suppress W and allow it to be subsumed by Z, X, as in Cornelissen et al. (2018). In practice, when estimating π_s's there is no explicit distinction between index instruments (Z) and cutoff instruments (W). Assumptions for estimation are stronger than those for identification above (see Appendix C.3.3).

C.3.1 Estimation Form

With additive separability, we can recover $\mathbb{E}[Y_s|X=x,U=u]$, which we do not observe, by starting with $Y \times D_s$, which we do observe. Specifically:

$$\begin{aligned} &\mathbb{E}[Y \times D_s|\pi_{s-1}(Z,X) = \pi_{s-1}, \pi_s(Z,X) = \pi_s, X = x] \\ &= \beta_s x(\pi_{s-1} - \pi_s) + \Lambda_s(\pi_{s-1}) - \Lambda_s(\pi_s) \end{aligned} \tag{8}$$

where each $\Lambda_s(k) = \int_0^k \mathbb{E}[\omega_s|U=u]du$. In practice, we can approximate each $\Lambda_s(k)$ with sieves, such that $\Phi_s(k)$ is a vector of basis functions and ϕ_s is a vector of coefficients:

$$\Lambda_s(k) - \Lambda_s(k') \approx \phi_s'[\Phi_s(k) - \Phi_s(k')] = \sum_{j=1}^{J} \phi_{s,j}(k)[\vartheta_{s,j}(k) - \vartheta_{s,j}(k')].$$

Assume also that we have i.i.d. data $\{(Y_i, S_i, Z_i, X_i) : i = 1, \ldots, n\}$.

Then, from equation (7) and using the functional form assumption in equation (8), for all s such that $1 < s < \bar{S}$:

$$\mathbb{E}[Y_s|X = x, U = \pi_s \text{ or } \pi_{s-1}] = -(-\beta_s x - \frac{\partial}{\partial \pi_s}\Lambda_s(\pi_s)) = \beta_s X + \frac{\partial}{\partial \pi_{s-1}}\Lambda_s(\pi_{s-1})$$

and for $s = 1$:

$$\mathbb{E}[Y_1|X = x, U = \pi_1] = \beta_1 x + \frac{\partial}{\partial \pi_1}\Lambda_1(\pi_1)$$

while for $s = \bar{S}$:

$$\mathbb{E}[Y_{\bar{S}}|X = x, U = \pi_{\bar{S}-1}] = \beta_{\bar{S}} x + \frac{\partial}{\partial \pi_{\bar{S}-1}}\Lambda_{\bar{S}}(\pi_{\bar{S}-1})$$

C.3.2 Estimation Steps

Estimation is based on equation (8). We can approximate Λ_s using B-splines or a polynomial of π_s (and similarly for π_{s-1}). The main estimation procedure in this paper follows five steps:

- 1. Recover estimates of π_s ($\forall s \in \{1, ..., \bar{S} - 1\}$) as probabilities (i.e., $\hat{\pi}_s \in [0, 1]$) (for example, using separate probit or logistic regressions) by regressing the treatment level being higher than s on instruments and regressors:

$$1\{S_i > s\} = \beta^s[X_i, Z_i, W_i] + \epsilon_i^s$$

  Then predict $\hat{\pi}_s$ $\forall s \in \{1, ..., \bar{S} - 1\}$. In this paper, I use a probit specification regressing $s \in [EM, Detention]$ ($s > 1$) and $s = Detention$ ($s > 2$) on judge fixed effects interacted with observables and time fixed effects to get predicted values for $\hat{\pi}_1$ and $\hat{\pi}_2$, respectively.

- 2. For each $s \in S$, construct $\Phi_s(\pi_s)$ and $\Phi_s(\pi_{s-1})$ either as polynomials or B-splines, each being a vector of basis functions.

- 3. For each $s \in S$, regress

$$Y \times D_s = \beta_s X(\pi_{s-1} - \pi_s) + \phi_s(\Phi_s(\pi_{s-1}) - \Phi_s(\pi_s)).$$

  If $s = 1$ then exclude $\Lambda_s(\pi_{s-1})$, and if $s = \bar{S}$ then exclude $\Phi_s(\pi_s)$. ϕ_s is a vector of coefficients with each element corresponding to each basis function. For example, if we are using a 3rd degree polynomial, then $\Phi_s(k) = [k, k^2, k^3]$, so $\phi_s = [\phi_s^1, \phi_s^2, \phi_s^3]$ and $\Phi_s(\pi_{s-1}) - \Phi_s(\pi_s) = [\pi_{s-1} - \pi_s, \pi_{s-1}^2 - \pi_s^2, \pi_{s-1}^3 - \pi_s^3]'$.

- 4. Then compute the estimate for $\mathbb{E}[Y_s|x, u]$ as:

$$\mathbb{E}\widehat{[Y_s|x, u]} = \begin{cases} \textit{if } s = 1 & \hat{\beta}_s x + \hat{\phi}_s \Phi_s'(\pi_s),\ u = \pi_s \\ \textit{if } \ s > 1 & \hat{\beta}_s x + \hat{\phi}_s \Phi_s'(\pi_{s-1}),\ u = \pi_{s-1} \end{cases}$$

  In the $s = 1$ case, $\mathbb{E}\widehat{[Y_s|x, u]} = -\frac{\partial \mathbb{E}\widehat{[Y(S=s)|x}, u=\pi_s]}{\partial \pi_s} = -[-\beta_s x - \hat{\phi}_s \Phi_s'(\pi_s)] = \beta_s x + \hat{\phi}_s \Phi_s'(\pi_s)$.

- 5. For each value of u in the support of both s and $s+1$ and any value of $X = x$, compute $\widehat{MTE}_{s+1,s}(x, u) = \mathbb{E}\widehat{[Y_{s+1}|x}, u] - \mathbb{E}\widehat{[Y_s|x}, u]$.

C.3.3 Assumptions for Estimation

Assumption 8. *E1:* $(\omega_s, U) \perp (Z, X)$ *for all* $s \in \mathcal{S}$.

Assumption 9. *E3: The distribution of* U *is uniform on* $[0, 1]$.

Assumption 10. *E4:* $E|Y_s| < \infty \ \forall s \in S$

Assumption 11. *E5:* $0 < \Pr(S = s) < 1 \ \forall s \in S$

Assumption 12. *E6: For all* $s \in \mathcal{S} \setminus \{1, \bar{S}\}$, *the joint distribution of* $\pi_s(Z, X)$ *and* $\pi_{s-1}(Z, X)$ *is absolutely continuous with respect to the Lebesgue measure on* $\mathbb{R}^2$. *Furthermore, the joint distribution of* $\pi_s(Z, X)$ *and* $\pi_{s-1}(Z, X)$ *is non-degenerate in the sense that its support cannot be reduced to a subset on* $\mathbb{R}$.

Assumption 8 replaces V with U and W is subsumed by X, Z and imposes a full independence assumption (rather than conditional independence). From this, the other assumptions no longer condition on X (due to full independence), W is subsumed into X, Z, and U is used in place of V for expediency removing the need for assumptions on τ.

C.4 Common Support

Estimation requires variation in $\pi_{s-1}-\pi_s$, which provides variation in higher terms of $\Lambda_{s-1}-\Lambda_s$. This must be factored into determining where the MTRs are estimable, in addition to ensuring that the MTRs of two adjacent treatments exist. Common support also determines the region over which we are able to produce counterfactuals and treatment effects.

Given that $\pi_0 = 1$ and $\pi_3 = 0$ for all individuals, the variation of π_1 and π_2 are of most concern — though their correlation is high (0.67) this is to be expected given that if $s > 2$ then $s > 1$ as well. As shown in Figure B.7, below the 45 degree line (such that $\pi_2 < \pi_1$), there is variation in values of each π excluding very high values of π_1 and very low values of π_2. The right panel in Figure B.7 displays the same plot for the felony-only sample.

To estimate the MTEs however, we require common support for the π's and variation within a treatment level for which we will estimate a marginal treatment response function. Figure B.8 displays the distribution of $\pi_{s-1} - \pi_s$ for each treatment level (right panel displays the same for felony-only sample). For $s = 1, 3$, $\pi_{s-1} - \pi_s$ covers the entire unit interval, but for $s = 2$, it falls short, with strong support only between about $[0.05, 0.8]$.

Figure B.9 displays the supports for each relevant π by treatment type (i.e., π_1 for $S = 1, 2$, and π_2 for $S = 2, 3$) for the main sample, along with 1% trim vertical lines denoting the 1st and 99th percentiles within the treatment-specific sample. There is nearly full support for $S = 1, 3$. However, for $S = 2$, the support ranges from about $\pi_1 \in [0.25, 1]$ and $\pi_2 \in [0.1, 0.77]$. Since the overlap we require is where there is common support for π_1 and π_2 for $S = 2$, we can only compute the MTR for EM for $\pi_1, \pi_2 \in [0.25, 0.77]$, which limits the range we can compute TSMTEs for EM versus release and EM versus detention, as EM is the most limited in support.

C.5 Confidence Intervals

The main results use bootstrapped MTRs based on 200 runs to compute confidence intervals. For each run, data is sampled with replacement and MTRs are computed. π's are not recalculated each run. Then MTEs are the difference between MTRs. Each run provides MTE estimates for each value of π_1 and π_2 for EM versus release and EM versus detention, respectively, and the 95% confidence intervals are taken as the 97.5% and 2.5% (195th and 5th highest values) for each MTE point independently — meaning if bootstrap sample 1 corresponds to the 2.5% value for $MTE_{EM,R}(\pi_1 = 0.5)$, it does not mean the 2.5% value for $MTE_{EM,R}(\pi_1 = 0.51)$ is from bootstrap sample 1 as well. As a result, the confidence intervals are not symmetric, as would result if standard errors were computed using the distribution of estimates.

While the confidence intervals in the main results do not account for the fact that π_1 and π_2 are estimated objects, doing so has very little effect on the size of confidence intervals. I test this by using 400 bootstrap runs in which π_1 and π_2 are re-computed each run using the full (bootstrapped) sample then filtering for main felony sample observations. This full bootstrap only increase the 95% confidence interval sizes for the five main robustness outcomes on the parameters of interest (CATEs, CATR, and CATD) by an average of 1.04% with a maximum increase of 18%.