DOI: 10.1111/jels.12351

ORIGINAL ARTICLE

JOURNAL OF EMPIRICAL LEGAL STUDIES

How accurate are rebuttable presumptions of pretrial dangerousness? A natural experiment from New Mexico

Cristopher Moore[1] | **Elise Ferguson**[2] | **Paul Guerin**[2]

[1]Santa Fe Institute, Santa Fe, New Mexico, USA

[2]Institute for Social Research, Center for Applied Research and Analysis, University of New Mexico, Albuquerque, New Mexico, USA

Correspondence
Cristopher Moore, Professor, Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA.
Email: moore@santafe.edu

Abstract

In New Mexico and many other jurisdictions, judges may detain defendants pretrial if the prosecutor proves, through clear and convincing evidence, that releasing them would pose a danger to the public. However, some policymakers argue that certain classes of defendants should have a "rebuttable presumption" of dangerousness, shifting the burden of proof to the defense. Using data on over 15,000 felony defendants who were released pretrial in a 4-year period in New Mexico, we measure how many of them would have been detained by various presumptions, and what fraction of these defendants in fact posed a danger in the sense that they were charged with a new crime during pretrial supervision. We consider presumptions based on the current charge, past convictions, past failures to appear, past violations of conditions of release, and combinations of these drawn from recent legislative proposals. We find that for all these criteria, at most 8% of the defendants they identify are charged pretrial with a new violent crime (felony or misdemeanor), and at most 5% are charged with a new violent felony. The false-positive rate, that is, the fraction of defendants these policies would detain who are not charged with any new crime pretrial, ranges from 71% to 90%. The broadest legislative proposals, such as detaining all defendants charged with a violent felony, are little more accurate than detaining a random sample of defendants released under the current system, and would jail 20 or more people to prevent a single violent felony. We also consider detention recommendations based on risk scores from the Arnold Public Safety Assessment (PSA). Among released defendants with the highest risk score and the "violence flag," 7% are charged with a new violent felony and 71% are false positives. We conclude that these criteria for rebuttable presumptions do not accurately target dangerous defendants: they cast wide nets and recommend detention for many pretrial defendants who do not pose a danger to the public.

INTRODUCTION: BAIL REFORM AND REBUTTABLE PRESUMPTIONS

In this article, we measure the accuracy of various "rebuttable presumptions" that a defendant is a danger to the public, and therefore should be detained pretrial. Specifically, we measure what fraction of the defendants to whom these presumptions apply are charged with new crimes during their pretrial period, especially violent crimes, and what fraction are "false positives" who are not charged with any new crime. Some of the criteria we consider, singly and in combination, are whether the defendant:

- is charged with a violent felony or firearm-related crime,
- has prior felony convictions,
- has prior failures to appear or prior violations of conditions of release,
- receives a high-risk score from a risk assessment algorithm.

Our goal is to help policymakers and advocates better understand the benefits and costs of these presumptions, and especially whether they yield an improvement over a system of individualized hearings which places the burden of proof on the prosecution. We show that many such policies, including recent legislative proposals, are not accurate ways of identifying dangerous defendants: they are wide nets that recommend detention for many pretrial defendants who do not pose a danger to the public. We do not attempt to design an "optimal" policy: rather we measure the effect of policies similar to those used or proposed in various jurisdictions today.

Very recently, the Mayor of Chicago complained that judges release too many defendants charged with violent crimes on bond or electronic monitoring (Chicago Tribune, June 7, 2022). She argued that such a charge is sufficient evidence of public danger in itself: "We shouldn't be locking up nonviolent individuals just because they can't afford to pay bail. But, given the exacting standards that the state's attorney has for charging a case, which is proof beyond a reasonable doubt, when those charges are brought, these people are guilty." This is exactly the debate—between individualized hearings versus broad criteria for detention—that we hope to inform.

We begin with some historical context. Since the 1960s and before, reformers and civil rights activists have argued that imposing financial conditions of release ("money bail") on defendants—and the resulting sub rosa detention of many defendants who cannot afford the amount of bail set—is unjust and unconstitutional. It is also a major source of bias in the criminal justice system against low-income defendants, and due to the correlations between race and class in our society, also contributes to bias against defendants who are people of color (e.g., Schlesinger, 2005). This has led to multiple waves of bail reform at the Federal and state level (Hegreness, 2013; Schnacke et al., 2010).

At the same time, a number of states have laws or clauses in their constitutions that identify certain defendants as being ineligible for pretrial release under any conditions. Originally most states followed Pennsylvania in guaranteeing a right to bail except for those accused of capital offenses. However, as part of an omnibus crime act passed during the Nixon administration, judges were given the ability to deny bail for noncapital crimes in the District of Columbia on the basis of danger to the public. This model was taken up by several state legislatures, and by Federal courts in the Bail Reform Act of 1984.

Where they exist, these laws create "rebuttable presumptions" that certain defendants are a danger to the public and therefore have no right to bail. In essence, they predict that a defendant will commit a serious crime if they are released between arrest and trial. These presumptions place the burden of proof on the defense to argue that the defendant can be released, perhaps under strict conditions of supervision, without posing a danger to the community—rather than requiring the prosecution to prove they cannot.

Since defendants in the United States are presumed innocent until proven guilty, these laws raise profound constitutional questions (e.g., Berger, 2023; Durham, 2021; Overbeck, 1986). However, in 1987, the Supreme Court upheld the facial constitutionality of the Bail Reform Act of 1984, as long as pretrial detention is "carefully limited." Specifically, the Court held[1] that a judge can detain a defendant pretrial if they hold a hearing[2] and find there is "clear and convincing evidence" that "no condition or combination of conditions will reasonably assure the safety of any other person and the community."

In addition to this general criterion, the 1984 Bail Reform Act created certain rebuttable presumptions of dangerousness for defendants in Federal courts. These consist of criteria involving the charge the defendant is currently accused of, whether the current charge occurred during pretrial supervision, and whether the defendant has past convictions for serious felonies.

At the state level, the analogous criteria differ widely. A defendant is presumed dangerous in Maryland[3] if they are charged with being a "drug kingpin"; in North Carolina[4] if their charge is associated with being a street gang, they have prior convictions, and are currently under supervision; in the District of Columbia[5] if they were armed with a firearm or had one readily available when arrested; and in New Jersey[6] if they are accused of murder or another crime

[1] *United States v. Salerno,* 481 U.S. 739 (1987).

[2] This hearing is not a trial and may not focus on whether the defendant committed the crime of which they are accused, although evidence that they did can be taken as evidence of danger to the public. The defense is entitled to counsel and may cross-examine witnesses, if any. However, the standard rules of evidence typically do not apply: for instance, proffers and "reliable hearsay," such as from the arresting officer, are allowed.

[3] Maryland Criminal Procedure Code Ann. §5–202 (2020).

[4] 2015 North Carolina General Statutes §15A-533, Right to pretrial release in capital and noncapital cases.

[5] Code of the District of Columbia §23–1322, Detention prior to trial.

[6] Rule Governing the Courts of the State of New Jersey. Rule 3:4A, pretrial detention.

punishable by life imprisonment. In California, state law briefly contained[7] a rebuttable presumption for defendants who are "assessed as high risk," a broad phrase that caused great concern in the age of algorithmic risk assessment (Koepke & Robinson, 2018). While this language was enacted in 2018, it was put on hold and repealed by a referendum in 2020.

While many of these rebuttable presumptions are intuitive and politically popular, there is no general agreement on what criteria should apply, or on best practices, if any, for using them. We argue that this is, in part, due to the lack of scientific data regarding their predictive accuracy: that is, the probability that these defendants will in fact commit additional crimes if released. Some argue that judges implicitly try to estimate this probability and impose pretrial detention if they believe this probability is above some threshold. But it is not clear what this threshold is or should be (Mayson, 2018).

To put this more sharply, for every violent crime we prevent by detaining someone, we detain a certain number of defendants who would not have posed a danger to the public if released. Detention has many collateral consequences. It causes defendants to lose jobs, housing, families, and relationships; makes it harder for them to mount a strong defense; pressures them to accept plea bargains even if they are innocent, or are guilty of a lesser charge; and makes them more likely to commit crime in the future (Digard & Swavola, 2019). Even a few days of detention can have significant effects on housing, employment, and health. As the American Bar Association (2007) states, "Deprivation of liberty pending trial is harsh and oppressive, subjects defendants to economic and psychological hardship, interferes with their ability to defend themselves, and, in many instances, deprives their families of support."

ERRORS AND ACCURACY

Legislators considering proposals for rebuttable presumptions need to face, not just constitutional questions, but questions of accuracy. They need to consider, as quantitatively and objectively as possible, both types of error inherent in any pretrial detention policy: "false positives" who are detained unnecessarily, and "false negatives" who are released and commit a serious crime during the pretrial period.[8] But there is a fundamental asymmetry in society's awareness of these two types of error. The defendants whose lives are disrupted due to pretrial detention are largely invisible, and are rarely reported on. In contrast, when

[7]California Penal Code §1320.20, 2019 version.

[8]We use these terms rather than opaque phrases like "Type I errors" and "Type II errors" partly because the reader may be familiar with them due to the current pandemic. If danger to the public is a disease, we think of a given rebuttable presumption as a diagnostic test, and pretrial detention as a treatment which may or may not be worth the harm it causes. In fact the phrase "false positive" in this context goes back to the debate around the Bail Reform Act of 1984: see for example (Himsell, 1986).

someone is arrested and released and then convicted of a terrible crime, the press reports their name, the name of their victim, and the nature of their act. This raises understandable questions about why they were not behind bars given their arrest, and produces a sense that pretrial supervision is not working: that law enforcement is being forced to "catch and release" dangerous criminals, creating a "revolving door" of crime.[9] The salience of these high-profile cases creates political pressure to broaden the conditions for pretrial detention, and—especially during periods where crime is increasing—makes it easy to portray policymakers who oppose rebuttable presumptions as "soft on crime." In a number of states and cities, this has cast doubt on the overall project of bail reform.[10]

This makes it all the more important to assess the accuracy of rebuttable presumptions—the fraction of the people they detain who would commit a serious crime if released, and the fraction who would not. If the stated goals of a pretrial detention policy are to prevent serious crime while preserving constitutional liberties, then it is important to know to what extent it succeeds in meeting these goals. Even if pretrial detention is constitutional in principle, sufficient inaccuracy can call the constitutionality of a policy into question: laws must have a rational connection with their intended effect, especially when they impose restrictions on a class of people. Thus if a policy has a high false-positive rate, and thus detains a large number of people for each serious crime it prevents, we should question whether it succeeds in balancing individual liberty with society's interest in preventing crime (e.g., Angel et al., 1971).

Of course, for many people—including many elected officials and their constituents—what matters is not rates or fractions, but the absolute number of crimes. As we show below, the fraction of pretrial defendants who are charged with an additional high-level felony during pretrial supervision is quite low. Moreover, only a small fraction of felony crimes in the community are committed by pretrial defendants. But as a spokesperson for the Governor of New Mexico said[11] in response to our report, "This very study references nearly 100 defendants who were charged with new violent felonies while on pre-trial release—that's nearly 100 victims and families who were subjected to violence

[9]"In Wake of Violent Weeks, Mayor's Metro Crime Initiative Develops Strategies for Shutting Revolving Door," https://www.cabq.gov/mayor/news/in-wake-of-violent-weeks-mayor2019s-metro-crime-initiative-develops-strategies-for-shutting-revolving-door; Greater Albuquerque Chamber of Commerce, "Unmonitored defendants are a serious safety problem in BernCo," Opinion in the Albuquerque Journal, Oct. 3rd 2021. https://www.abqjournal.com/2434481/unmonitored-defendants-are-a-serious-safety-problem-in-bernco.html.

[10]US News "New Mexico Poised to Reconsider Bail Reform Amid Crime Wave," https://www.usnews.com/news/best-states/new-mexico/articles/2021-08-31/new-mexico-poised-to-reconsider-bail-reform-amid-crime-wave One recent instance of movement in the opposite direction is Virginia, who repealed rebuttable presumptions along party lines in 2021 (Senate Bill 1266). See Daily Press February 6, 2021, "Virginia lawmakers scrap 'presumption' that those charged with certain crimes don't get bail."

[11]"Study says jailing violent defendants would barely move crime rate" Santa Fe New Mexican. Aug 30, 2022 Updated Sep 2, 2022, https://www.santafenewmexican.com/news/local_news/study-says-jailing-violent-defendants-would-barely-move-crime-rate/article_25c31ee0-287e-11ed-8602-ab7bcc4b5885.html.

due to an offender's violence not being given appropriate consideration." The victims of these crimes and their families have every right to ask why the offender was released, and whether these crimes could have been prevented by detaining them—just as, in the medical context, an anguished family might ask why their loved one was misdiagnosed and failed to receive a lifesaving treatment. But as with a proposed new diagnostic test, we should only adopt a detention policy if it is sufficiently accurate. What counts a "sufficient" depends both on the relative costs of false negatives and false positives—which is a matter for stakeholders to determine—and on the rate with which the policy would make these mistakes.

But how can we measure the accuracy of pretrial detention? Just as a bank only knows if a borrower would return a loan if it is approved, and a university only knows how a student would perform if they are accepted, we do not know how a defendant will behave unless they are released.

One approach is to measure jail populations and rearrest rates before and after a pretrial detention policy is established, repealed, or modified; for an early attempt to do this for the 1984 Bail Reform Act, see (Kingsnorth et al., 1987). However, crime rates and law enforcement are in constant flux for many reasons, due to changes in policy, procedures, demographics, social and economic conditions, the availability of weapons, and so on. This makes it difficult to assign credit or blame for decreases or increases in jail populations or crime to any one factor. Studies of this type for the effectiveness of risk assessment algorithms have produced mixed results (Stevenson, 2018; Viljoen et al., 2019).

Instead of a before-and-after study, we would ideally like to measure the accuracy of a given policy by looking at outcomes in two parallel worlds: one where the policy is in place, and one where it is not. These "natural experiments" are not always available to us.[12]

Here we carry out a natural experiment using a database of over 15,000 felony defendants who were released pretrial in Bernalillo County, New Mexico (where Albuquerque is located) over the course of 4 years. As discussed below, New Mexico's current system allows pretrial detention for any felony, but places the burden on the prosecutor to prove by clear and convincing evidence that the defendant is dangerous. Rebuttable presumptions would lower this threshold, and would detain additional defendants except when judges ruled in favor of the defense that the presumption has been overcome. By looking at outcomes for defendants released under the

[12]One natural experiment occurred during the litigation leading to *Schall v. Martin*. In 1981, a federal judge enjoined the preventive detention of juvenile offenders in New York State, creating a writ of habeus corpus that applied for 3 years until the Supreme Court found in 1984 that the New York law does not violate the Fourteenth Amendment. This led to the release of 74 juvenile defendants. Although this is a very small sample, it offered an opportunity to measure the accuracy of judicial decisions (Fagan and Guggenheim 1996).

current system, we can measure the accuracy of a variety of rebuttable presumptions—both those rejected by the New Mexico legislature in the 2022 legislative session, and those already in force or under discussion in other states and cities. These data lets us revisit the 50-year-old debate over rebuttable presumptions and pretrial detention.

LAWS AS ALGORITHMS

There is an important connection between our study and the growing debate on algorithmic risk assessment. These data on felony defendants are available to us because the New Mexico Administrative Office of the Courts undertook a local revalidation study (Ferguson, De La Cerda, Guerin, & Moore, 2021) of the Arnold Public Safety Assessment (PSA), a risk assessment tool used in many jurisdictions around the country. Indeed, one of the policies we assess is the use of this tool to recommend detention, as many jurisdictions such as New Jersey[13] and San Francisco[14] do. This use of the PSA, while not recommended by the PSA's designers,[15] is effectively an algorithmically-driven rebuttable presumption. Thus, it can be examined for accuracy using the same techniques that we apply to legislatively-driven policies.

It is our hope that the growing demand for transparency and accountability in algorithms will renew and strengthen the same demands for all aspects of decision-making in criminal justice, both human and automated. Many laws are algorithms by another name: a policy of the form "if [condition] then detain" might as well be a line of computer code. These lines of code, whether created by programmers, machine learning algorithms, or legislatures, must be subject to scrutiny and quantitative study, so that all stakeholders can make a clear-eyed assessment of their benefits to the public and their costs to liberty.[16]

[13]New Jersey Pretrial Release Recommendation Decision Making Framework (DMF), https://www.njcourts.gov/courts/assets/criminal/decmakframwork.pdf.

[14]Alissa Skog and Johanna Lacoe, *Validation of the PSA in San Francisco*. California Policy Lab https://www.capolicylab.org/wp-content/uploads/2021/08/Validation-of-the-PSA-in-San-Francisco.pdf.

[15]Currently, Arnold Ventures and their project Advancing Pretrial Policy and Research state that the PSA "doesn't make recommendations about whether to release or detain a person pretrial" and should only be used to recommend conditions of release, see https://advancingpretrial.org/psa/factors/release-conditions-matrix/ . However, it was initially presented as a tool for both purposes (e.g., https://www.arnoldventures.org/newsroom/laura-and-john-arnold-foundation-launches-psa-to-improve-risk-based-decision-making/) and many jurisdictions still interpret the highest risk scores as a recommendation to detain.

[16]For more on the analogies between law and code, see Lessig (2006). Benjamin (2019) argues that algorithms reproduce historical structures of racism and discrimination. Some other recent contributions, including Kleinberg et al. (2020) and Bell et al. (2021) take a more optimistic view: namely that if algorithms are sufficiently transparent, they can be used to detect discrimination in human decisions, and thus create fairer systems overall. This is a vital debate. Our claim here is that both algorithms and human policies should be held accountable to their stated goals, and independently assessed to see if they meet these goals.

THE NEW MEXICO SETTING

New Mexico has a unique history of bail reform (Siegrist et al., 2020). In 2014, the New Mexico Supreme Court considered a case where a defendant accused of murder was detained for 2 years and 4 months because he could not afford a $250,000 bail bond even though uncontroverted evidence was presented that he would not pose a danger to the public if he were released under a strict set of conditions, including monitoring by a GPS device, living with his father, making regular contact with the pretrial services program, and maintaining employment at a local restaurant. The Court found that this detention violated the defendant's right to bail in the New Mexico constitution, as well as New Mexico's rules of criminal procedure that a defendant be released "on the least restrictive conditions necessary to reasonably assure both the defendant's appearance in court and the safety of the community."[17]

This case began a process that led to an amendment to the state constitution, approved by New Mexico voters in 2016. This amendment ostensibly prevents defendants from being detained simply because they cannot afford bail, but it also allows judges to detain any felony defendant if the prosecutor proves that they are a danger to the public. The new language in Article II, Section 13 of the New Mexico constitution reads[18]:

Bail may be denied by a court of record pending trial for a defendant charged with a felony if the prosecuting authority requests a hearing and proves by clear and convincing evidence that no release conditions will reasonably protect the safety of any other person or the community.

A person who is not detainable on the grounds of dangerousness nor a flight risk in the absence of bond and is otherwise eligible for bail shall not be detained solely because of financial inability to post a money or property bond.

Like a number of other states, New Mexico's movement away from money bail was coupled with an increased interest in risk assessment tools, also known as instruments or algorithms, in the hope that they would make release decisions more consistent and evidence-based. Since June 2017, the Bernalillo County Metropolitan Court and the Second Judicial District Court have used the Arnold PSA for felony defendants. Since then, the state has started pilot projects in several other jurisdictions.

When a jurisdiction adopts the PSA, local stakeholders construct a "release conditions matrix" that translates risk scores into recommendations to the judge for conditions of release. These conditions might include additional notifications of upcoming hearings, required check-ins with pretrial supervision officers, drug testing, or tracking with a GPS device. In Bernalillo County, the

[17]State v. Brown, 2014-NMSC-038, 338 P.3d 1276.

[18]The 2016 amendment also repealed a 1980 "three strikes"-like amendment which allowed detention for felony defendants with two or more prior felony convictions, or those accused of a felony involving a deadly weapon who have one or more prior felony convictions.

recommendation in the highest score category is for the judge to hold a hearing to see if the constitutional conditions for detention are met as per the 2016 amendment, and if not, to release the defendant with maximum conditions (see Figure 1 below).

Thus, in the current system, the burden of proof is on the prosecutor to bring a motion and show that detention is necessary, even for defendants with high-risk scores and/or serious charges.[19] While judges take PSA scores into account when considering these motions, they are not dispositive in either direction. The prosecutor can argue that the PSA's risk assessment is too low, for instance if the defendant has boasted of their intent to kill a potential witness. Similarly, the defense can argue that a defendant should be released despite a high PSA score, for instance if they have completed rehabilitation programs and shown a pattern of complying with conditions of release.

In recent years, judges in Bernalillo County have agreed with roughly half the motions to detain brought by prosecutors. However, due partly to high-profile cases where serious crimes were committed by pretrial defendants and partly to an overall concern about rising crime rates, a wide range of New Mexico policymakers from both parties—including the local District Attorney, the Mayor of Albuquerque, state legislators, and the Governor of New

		New Criminal Activity Scale					
		NCA 1	NCA 2	NCA 3	NCA 4	NCA 5	NCA 6
Failure to Appear Scale	FTA 1	(A) ROR	(B) ROR				
	FTA 2	(C) ROR	(D) ROR	(E) ROR-PML 1	(F) ROR-PML 3	(G) ROR-PML 4	
	FTA 3		(H) ROR-PML 1	(I) ROR-PML 2	(J) ROR-PML 3	(K) ROR-PML 4	(L) Detain or Max Conditions
	FTA 4		(M) ROR-PML 1	(N) ROR-PML 2	(O) ROR-PML 3	(P) ROR-PML 4	(Q) Detain or Max Conditions
	FTA 5		(R) ROR-PML 2	(S) ROR-PML 2	(T) ROR-PML 3	(U) Detain or Max Conditions	(V) Detain or Max Conditions
	FTA 6				(W) Detain or Max Conditions	(X) Detain or Max Conditions	(Y) Detain or Max Conditions

FIGURE 1 The release conditions matrix used in Bernalillo County and the Second Judicial District to interpret the results of the Arnold Public Safety Assessment (PSA). [Color figure can be viewed at wileyonlinelibrary.com]

[19]Like most states, New Mexico's constitution excludes from the right to bail those accused of "capital offenses when the proof is evident or the presumption great." However, the death penalty was repealed in New Mexico in 2009 so there are currently no capital offenses in New Mexico law. In 2018 the State Supreme Court ruled that this exclusion does not apply to formerly capital crimes such as first degree murder, referring instead to judges' ability to find a danger to the public under the 2016 amendment. State v. Ameer, No. S-1-SC-36395 slip op. (N.M. S. Ct. April 23, 2018).

Mexico—have called for more defendants to be detained pretrial. In the 2021 and 2022 legislative sessions, bills were introduced that would have created rebuttable presumptions of dangerousness based on the current charge, past felony convictions, or past violations of conditions of release. These bills failed due to a combination of constitutional, financial, and evidence-based concerns, but they are likely to reappear in some form.

DATA

Our study is based on the 15,134 felony defendants arrested in the 48-month period from July 1st, 2017 through June 30th, 2021, whose cases originated in Bernalillo County Metropolitan Court and were closed by the end of this period, who were assessed with the Arnold PSA, and who were released pretrial for all or part of the intervening period—either because the prosecutor did not bring a motion to detain them, or because the judge ruled against such a motion. For each defendant, our data include the following:

- the original charge filed by the District Attorney,
- past felony convictions, past failures to appear, and past violations of conditions of release,
- the scores and recommendations produced by the PSA,
- whether the prosecutor brought a motion to detain,
- new charges filed from crimes committed during pretrial supervision, if any,
- failure to appear at court hearings during pretrial supervision, if any.

See (Advancing Pretrial Policy and Research, 2022) for a review of the PSA, including its input factors, weights, and cutpoints.

All charges are classified by type (property, drug, DWI, public order/other, and violent) and severity (petty misdemeanor, misdemeanor, and first through fourth-degree felonies). Note that all the original charges are felonies, since Bernalillo County only uses the PSA for felony defendants. In the tables below, we classify rearrests by violent versus nonviolent and felonies versus misdemeanors for simplicity. We exclude most traffic charges, all parking statutes, and city ordinances.

It is worth noting that very few of these defendants were charged with new high-level felonies. A total of 15 defendants in our dataset, or one out of every thousand, were charged with a new first-degree felony; 11 of these were violent. A total of 141 were charged with a new second-degree felony, 70 of which were violent—about one-half of 1%. We consider criteria for detention based on the original charge, including violent felonies, Serious Violent Offenses (SVO) (see Tables A1 and A2), and firearm- and weapons-related charges (Tables A3 and A4); prior convictions for felonies and violent crimes, and prior violations of conditions of release, including failure to appear; and scores and recommendations derived from the PSA. We also consider various

combinations of these similar to proposals recently considered in the New Mexico legislature.

For each criterion or combination of criteria, we measure how many additional defendants in our sample would have been detained. We then ask what fraction of these defendants would have been arrested pretrial for crimes of various types and severity (including violent vs. nonviolent and felonies vs. misdemeanors) and what fraction would not have been rearrested at all. This lets us measure the false-positive rate of a given detention policy, as well as the number of people who would have to be detained to prevent a single violent rearrest.

Our analysis assumes that the effect of a rebuttable presumption is to detain all the defendants to which it applies, in addition to those detained under the current system, and otherwise leave the system unchanged. This is simplistic in several respects. First, in some cases, the judge will find that the defense has overcome the presumption, and release the defendant: if judges are accurate in these rulings it would partly ameliorate the false-positive rate. Second, there may be substitution effects where some defendants detained under the current system would be released because the presumption does not apply to them, or because judges attempt to compensate for their loss in discretion by becoming more lenient. There could also be changes in prosecutor behavior, for instance to raise the severity of the charge in order to trigger the presumption, which may create additional false positives.

These and other effects could alter jail populations, conviction rates, and so on in complicated ways. We set these effects to the side because we our goal is simply to assess the predictive accuracy of rebuttable presumptions, and to measure whether the defendants they target are in fact more dangerous than the typical defendant released under the current system.

Before we proceed we consider three caveats to our data and how it is coded.

Charges versus crime

We equate crime with arrests that lead to filed charges. In particular, "new criminal activity" (NCA) means that a charge was filed for a crime committed during the pretrial period. (Our data include new charges filed anywhere in New Mexico, and indeed about 8% of new charges in our data are from jurisdictions outside Albuquerque.) While this is the only measure we have, it makes errors in both directions: not all crime is reported and not all reported crime leads to arrest and prosecution, and not all arrests or charges correspond to actual crimes.

Based on the National Crime Victimization Survey, the estimated fraction of crimes reported to police is 52% for aggravated assault, 47% for robbery, and

34% for rape and sexual assault (Pew Research Center, 2020). Some writers have speculated that these rates are higher for crimes committed by pretrial defendants, either because they are under pretrial supervision or because the fact that they were caught once suggests that they are less adept at evading the law (Angel et al., 1971) but we know of no studies to that effect.

In the other direction, we know of very few studies that measure pretrial crime in terms of convictions rather than charges. A study of the PSA in Los Angeles found that less than one-third of pretrial rearrests led to new convictions, although they caution that this may be an underestimate due to reporting delays and cases still pending at the end of the study period (County of Los Angeles, 2021).

Thus, with misgivings, and the usual sheepish reply that "it's the data we have," we use filed charges as a signal of crime in both directions: we give a detention policy credit for having prevented a crime if a detainee would have had new charges filed against them during their pretrial period, and we call a detainee a false positive if they would not have been. We hope to revisit this in the future with other forms of data. But even though charges are both noisy and biased as signals of crime, we believe they yield compelling results in this context.

Definitions of serious violent offense

Several legislative proposals have focused on defendants charged with an SVO. New Mexico law defines these as in Table A1. However, there is an additional set of charges, shown in Table A2, which a court can find to be SVOs as well. Thus, in several places, below we consider both a narrow definition based just on Table A1, and a broad one that includes both tables. In practice, judicial discretion would fall somewhere in between these two.

Firearm-related crimes

The involvement of firearms in a crime can be defined in many ways, ranging from discharging a firearm, to brandishing one, to having one "readily available." Since our data only give us the filed charge, we do not always know which cases involved a firearm: for instance, aggravated assault with a deadly weapon includes cases both with and without the use of a firearm. In addition, the only appearance of brandishing a firearm in the current New Mexico code is an enhancement for other felonies (N.M. Stat. § 31-18-16).

As a proxy for firearm use, we focus on charges listed in Table A3 that clearly indicate the use of a firearm or a deadly weapon. In Table A4, we give a longer list of charges that indicate the use of a deadly weapon which may or

may not be a firearm. We hope to improve this in the future by looking at crime reports and sentencing hearings. We claim, however, that broadening the definition to include possession of a firearm, the presence of a firearm in easy reach (say, in the glove compartment when the defendant's car is searched), and so on, is likely to lead to higher false-positive rates. Wider nets tend to be less accurate.

HOW MUCH CRIME IS COMMITTED BY PRETRIAL DEFENDANTS?

Since the stated purpose of rebuttable presumptions is to prevent crime, we start by considering their potential benefits to the public in terms of reducing crime rates. Of all felonies, what fraction is committed by pretrial defendants? To what extent do these defendants drive a "revolving door" of crime?

To answer this, we compare the number of felony cases filed against released defendants (the numerator) with the total number of felony cases filed in the jurisdiction during the study period (the denominator). As discussed above, not all felony crimes are reported or lead to arrest and prosecution. But we assume this affects the numerator and denominator equally, that is, that the probability a felony leads to filed charges is the same for crimes committed by pretrial defendants as for other defendants. In order to compare apples with apples, in the denominator we only count felonies for which—like those contributing to the numerator—the defendant was in custody for their release decision, the PSA was assessed, and the case was closed by the end of the 48-month period. Since this reduces the denominator, it provides an upper bound on the fraction of felony charges in the jurisdiction where the defendant was on pretrial supervision at the time.

Tables 1 and 2 show that, by this measure, a small fraction of felonies are committed by pretrial defendants. Defendants who are accused of violent felonies commit about 2% of all felonies, about 3% of violent felonies, and about 2% of first-degree felonies. For instance, there were a total of 8366 violent felony cases filed and closed during the study period. Four hundred fifty-six of these cases were filed against felony defendants under pretrial supervision, and of those 228 were filed against defendants accused of a violent felony. Thus, assuming a simple relationship between detention and prevention, detaining all defendants accused of a violent felony would prevent about 3% of violent felonies (228 out of 8366) and about 2% of felonies overall (444 out of 21,745).

As a result, detaining all defendants accused of a violent felony would only decrease felony crime by a few percent. Even detaining *all* felony defendants, which to our knowledge no one has proposed, would prevent about 8% of all felonies, about 5% of violent felonies, and about 6% of first-degree felonies. Thus, going beyond the current system and detaining additional defendants

based on rebuttable presumptions would decrease the rate of serious crime only slightly.

Released defendants also commit some misdemeanors, including violent misdemeanors, as we show below. However, since our data only concerns felony defendants, we do not know the total number of misdemeanor cases filed during the study period. Lacking that denominator, we cannot calculate the fraction of misdemeanors committed by pretrial defendants. But we see no reason why this fraction would be very different than it is for felonies.

We assume here, as stated above, that detaining someone who would have been charged with a new crime is synonymous with preventing that crime. In addition to the assumption that being charged with a crime is equivalent to committing one, this ignores second-order effects, such as whether a higher level of pretrial detention would deter some potential criminals, or (in the other direction) whether preventing one perpetrator from committing a crime creates an opportunity for someone else to do so. Our point is that even if we assume a simple, direct relationship between detention and crime prevention, rebuttable presumptions would have a relatively small effect on the rate of serious crime.

RESULTS: CRITERIA BASED ON CHARGES AND PRIORS

Since rebuttable presumptions would prevent at most a small fraction of serious crimes, their benefits to society are modest. The rest of this paper focuses on their costs—including the number of people they detain unnecessarily—and whether they are accurate enough to justify these costs.

TABLE 1 The fraction of felonies committed by pretrial felony defendants released under the current hearing-based system.

Felony type	Felony cases filed against pretrial defendants accused of any felony	Felony cases filed against pretrial defendants accused of a violent felony	Total number of closed felony cases during study period
Violent	456 (5%)	228 (3%)	8366
Drug	567 (8%)	104 (1%)	6972
Property	630 (11%)	100 (2%)	5503
DWI	5 (2%)	2 (1%)	235
Public order/other	49 (7 ± 1%)	10 (1%)	669
Total	1707 (8%)	444 (2%)	21,745

Note: Except where shown, standard errors are less than 1%.

TABLE 2 The same data as in Table 1, but classified by severity rather than type.

Felony degree	Felony cases filed against pretrial defendants accused of any felony	Felony cases filed against pretrial defendants accused of a violent felony	Total number of closed felony cases during study period
First degree	15 (6 ± 1%)	5 (2%)	262
Second degree	141 (8%)	34 (2%)	1869
Third degree	276 (6%)	94 (2%)	4751
Fourth degree	1275 (9%)	311 (2%)	14,863
Total	1707 (8%)	444 (2%)	21,745

Note: Except where shown, standard errors are less than 1%.

We consider policies that recommend detention for various classes of defendants. Within each such class, what fraction of defendants in fact pose a danger to the public if released, measured by new charges filed against them during the pretrial period? And what fraction are false positives who would receive no new charges if released?

Since New Mexico's current system already allows the prosecutor to present evidence that a defendant is dangerous to the public, and allows the judge to detain if they find this evidence to be clear and convincing, the reader should feel free to assume that at least some defendants with strong evidence of dangerousness are already being detained. Our question is whether it is accurate to go beyond this hearing-based system and recommend detention for broad classes of defendants even though the judge does *not* find the prosecutor's evidence to be clear and convincing—or, equivalently, to direct the judge to treat these criteria in and of themselves as sufficient evidence of public danger.

We consider presumptions based on a number of criteria, including the type or severity of crime of which the defendant was originally accused (Table 3); their criminal record, including prior convictions or prior violations of conditions of release (Table 4); and their risk score according to the Arnold PSA (Table 5). In the interests of readability, we round all rates to the nearest 1%, and report uncertainty in a rate p only in the few cases where the standard error $\sqrt{p(1\text{-}p)/n}$ exceeds 1%. As the reader may know, the standard error is not the best way to determine confidence intervals when p or n is small, but it suffices for our purposes.

For comparison, in the top row of each table, we show the rates for all felony defendants released under the current system. We find that 82% of all released defendants do not receive any new charge during the pretrial period; these would become false positives if detained. We find that 18% of all released defendants receive some new charge pretrial. This includes 5% who are charged

with a new violent crime, which in turn includes 3% who are charged with a new violent felony.[20]

In Table 3, we see that defendants whose original charge is a violent felony actually have a slightly lower rate of rearrest overall. When they are rearrested, it is somewhat more likely that the new charge will be violent rather than nonviolent. Nevertheless, for all of these charge-based criteria, the rate of new charges for violent crime is at most 7%, including 5% who are charged with violent felonies. The false-positive rates of these criteria range from 77% to 90%. For instance, detaining all 136 people charged with a first-degree felony would have prevented 6 charges for violent crimes, 4 of which were felonies, while also detaining 121 defendants who would not have received any new charge if released.

In Table 4, we see similar results for presumptions based on past convictions or violations of conditions of release, including past failures to appear. Among defendants with three or more prior convictions for violent crimes (felonies or misdemeanors), 8% are charged with a new violent crime, including 5% who are charged with a violent felony. Most of these presumptions again have a false-positive rate of at least 77%. For defendants with two or more failures to appear in the past 2 years the false-positive rate is 71%, but just 7% of them are charged with new violent crimes including just 4% with violent felonies.

RESULTS: ALGORITHMIC CRITERIA FOR DETENTION

Table 5 shows criteria based on Arnold PSA scores. As discussed above, while jurisdictions that use risk assessments to detain do not typically call them rebuttable presumptions (except, briefly, California) they have much the same effect. Thus, algorithmic criteria deserve the same scrutiny as those written by legislators.

The PSA produces two scores on a 6-point scale, one for new criminal activity (NCA) and one for failure to appear (FTA). It also produces a "violence flag" which is either present or absent. All of these are derived from the defendant's criminal record; the only demographic variable the PSA uses is age. We consider the criterion that the NCA score is 6, and also the "Detain if Constitutional Conditions are Met" category from the local conditions of release matrix, where either the NCA or FTA score is 6 or if both scores are 5 (see Figure 1). We consider each of these criteria alone and in combination with the violence flag. When we include the flag, 11–12% of these defendants are charged with a

[20]There are a handful of cases where a defendant is simultaneously charged with a nonviolent felony and a violent misdemeanor. We count these as "violent felonies" even though the highest charge is nonviolent. This has no signficant effect on our results.

TABLE 3 Criteria based on the original charge of which the defendant is accused.

Criterion (original charge)	No new charge	New charge: nonviolent misdemeanor	New charge: nonviolent felony	New charge: violent misdemeanor	New charge: violent felony	Total
All released felony defendants	12,388 (82%)	744 (5%)	1252 (8%)	295 (2%)	455 (3%)	15,134
First-degree felony	121 (89 ± 3%)	4 (3 ± 1%)	5 (4 ± 2%)	2 (2 ± 1%)	4 (3 ± 1%)	136
First- or second-degree felony	1046 (83 ± 1%)	64 (5%)	105 (8%)	13 (1%)	27 (2%)	1255
Violent felony	5137 (86%)	232 (4%)	217 (4%)	172 (3%)	228 (4%)	5986
Violent first-degree felony	108 (90 ± 3%)	4 (3 ± 2%)	2 (2 ± 1%)	2 (2 ± 1%)	4 (3 ± 2%)	120
Violent first or second-degree felony	426 (83 ± 2%)	30 (6 ± 1%)	32 (6 ± 1%)	5 (1%)	21 (4%)	514
Narrow SVO (Table A1)	1835 (86%)	80 (4%)	70 (3%)	61 (3%)	81 (4%)	2127
Broad SVO (Table A1 or A2)	4351 (85%)	203 (4%)	189 (4%)	147 (3%)	202 (4%)	5092
Firearm-related charge (Table A3)	315 (77 ± 2%)	18 (4 ± 1%)	50 (12 ± 2%)	9 (2%)	16 (4 ± 1%)	408
Deadly weapon- related charge (Table A3 or A4)	2248 (83%)	122 (4%)	148 (5%)	72 (3%)	124 (5%)	2714

Note: Criteria include high-level or violent felonies, Serious Violent Offenses (SVOs) as defined in New Mexico law (see Tables A1 and A2), and charges that indicate the use of a firearm or deadly weapon (see Tables A3 and A4). These rates are for pretrial defendants released under New Mexico's current hearing-based system. Except where shown, standard errors are less than 1%.

TABLE 4 Criteria based on prior convictions and prior violations of conditions of release.

Criterion (prior convictions and violations)	No new charge	New charge: Nonviolent misdemeanor	New charge: Nonviolent felony	New charge: Violent misdemeanor	New charge: Violent felony	Total
All released felony defendants	12,388 (82%)	744 (5%)	1252 (8%)	295 (2%)	455 (3%)	15,134
Prior felony conviction	4401 (77%)	340 (6%)	608 (11%)	128 (2%)	210 (4%)	5687
Prior violent conviction	3518 (78%)	251 (6%)	426 (9%)	115 (3%)	187 (4%)	4497
3 or more prior violent convictions	745 (77 ± 1%)	67 (7%)	80 (8%)	32 (3%)	46 (5%)	970
Prior violations of COR	8273 (79%)	611 (6%)	1039 (10%)	223 (2%)	370 (4%)	10,516
FTA in the past 2 years	3802 (74%)	388 (8%)	620 (12%)	128 (2%)	207 (4%)	5145
2 or more FTAs in the past 2 years	1932 (71%)	251 (9%)	346 (13%)	72 (3%)	113 (4%)	2714

Note: Criteria include prior violations of conditions of release (COR) or recent failures to appear (FTA). These rates are for pretrial defendants released under New Mexico's current hearing-based system. Note that "prior violent convictions" includes both violent felonies and violent misdemeanors. Except where shown, standard errors are less than 1%.

TABLE 5 Algorithmically-driven criteria based on the Arnold Public Safety Assessment (PSA).

Criterion (PSA scores and violence flag)	No new charge	New charge: Nonviolent misdemeanor	New charge: Nonviolent felony	New charge: Violent misdemeanor	New charge: Violent felony	Total
All released felony defendants	12,388 (82%)	744 (5%)	1252 (8%)	295 (2%)	455 (3%)	15,134
Detain/max PSA	1691 (71%)	202 (8%)	308 (13%)	75 (3%)	111 (5%)	2387
Detain/max PSA + violence flag	514 (72 ± 2%)	49 (7 ± 1%)	63 (9 ± 1%)	37 (5%)	47 (7%)	710
NCA score of 6	837 (71 ± 1%)	102 (9%)	147 (12 ± 1%)	37 (3%)	56 (5%)	1179
NCA score of 6 + violence flag	323 (71 ± 2%)	36 (8 ± 1%)	43 (9 ± 1%)	21 (5 ± 1%)	30 (7 ± 1%)	453

Note: Criteria include the highest NCA score, the violence flag, and the "Detain if Constitutional Conditions are Met" category from the release conditions matrix used in Bernalillo County and the Second Judicial District (Figure 1). These rates are for pretrial defendants released under New Mexico's current hearing-based system. Except where shown, standard errors are less than 1%.

new violent crime, including 7% who are charged with a violent felony. Their false-positive rates are 71%–72%.

THE ACCURACY OF RECENT LEGISLATIVE PROPOSALS

Table 6 shows results for various combinations of criteria based on recent bills proposed in the New Mexico legislature. House Bill 80, introduced in the 2021 session, would have created a rebuttable presumption of dangerousness for defendants who are (1) charged with a first-degree felony or SVO and who also (2) have a prior felony conviction or prior violations of conditions of release. As before, we consider both the narrow and the broad definition of SVO. In both cases, the fraction of defendants who would be charged with a new violent felony is 5%. An additional 3%–4% would be charged with a violent misdemeanor, and an additional 10%–12% would be charged with a nonviolent crime. The false-positive rate in both cases is 81%–82%. Thus, four out of five defendants targeted by this law would not be charged with any new crime pretrial.

Even broader proposals were introduced in the 2022 session. House Bill 5 lacks HB80's requirement of a past felony conviction. Among other things, it regards being charged with an SVO, or using or brandishing a firearm, as prima facie proof of dangerousness. House Bill 27 appears to go even farther by changing the "and" of HB80 to an "or." That is, it would create a rebuttable presumption for defendants who are charged with an SVO or used or brandished a firearm, *or* who have prior FTA on felony cases, or "a pattern of failure to follow conditions of release."

We approximate the results of these broader proposals as well as possible given our data. As before, we consider both the broad and narrow definitions of SVO. We interpret even a single violation of COR as "a pattern of failure," with the caveat that a judge may or may not agree depending on the circumstances or severity of these violations. While our data does not tell us which past failures to appear (FTA) are for felony cases, any FTA is also a violation of COR.

These proposals would significantly increase jail populations, detaining a significant fraction of defendants released under the current system. But a glance at Table 6 shows that, despite the differences between these proposals, none of them are particularly accurate. For HB5 and HB27, the rate of new charges among the defendants they would detain is essentially the same as the rate among all released defendants. The rate of new violent felony charges among defendants targeted by these bills is 4% as opposed to 3% among all released defendants; the rate of new violent misdemeanor charges is 2%–3% as opposed to 2%; and the false-positive rate is 77%–85% as opposed to 82%. In other words, these legislative proposals are roughly equivalent to detaining a random

TABLE 6 Criteria based on recent legislative proposals from the New Mexico legislature.

Criterion (legislation)	No new charge	New charge: nonviolent misdemeanor	New charge: nonviolent felony	New charge: violent misdemeanor	New charge: violent felony	Total
All released felony defendants	12,388 (82%)	744 (5%)	1252 (8%)	295 (2%)	455 (3%)	15,134
HB80: (1st deg felony or narrow SVO) and (prior felony or COR)	934 (82 ± 1%)	57 (5%)	55 (5%)	41 (4%)	55 (5%)	1142
HB80: (1st deg felony or broad SVO) and (prior felony or COR)	2373 (81%)	145 (5%)	155 (5%)	101 (3%)	152 (5%)	2926
HB5: Narrow SVO or firearm charge	2038 (85%)	93 (4%)	114 (5%)	66 (3%)	92 (4%)	2403
HB27: Broad SVO, firearm charge, or past FTA or COR	7603 (80%)	538 (6%)	776 (8%)	234 (2%)	350 (4%)	9501
HB27: Narrow SVO, firearm charge, or past FTA or COR	5681 (77%)	469 (6%)	725 (10%)	184 (3%)	288 (4%)	7347

Note: House Bill 80, introduced in the 2021 session, would have created a rebuttable presumption of dangerousness for pretrial defendants charged with a first-degree felony or Serious Violent Offense (SVO) who also have a prior felony conviction or prior violation of conditions of release (COR). House Bill 5, introduced in the 2022 session, would have created a presumption if a defendant committed an SVO from Table A1, or if they discharged or brandished a firearm. House Bill 27 would have created a presumption for defendants who are charged with an SVO, or have past failures to appear (FTA) in felony cases, or violations of conditions of release (COR). Except where shown, standard errors are less than 1%.

sample of defendants who are currently released; they do not identify a set of defendants that are more dangerous than others.

HOW ACCURATE ARE MOTIONS TO DETAIN?

In jurisdictions that already allow judges to consider evidence of public danger, calls for rebuttable presumptions are essentially calls to reduce judicial discretion. What would happen if the decision to detain were up to the prosecutor rather than the judge? How would this change the accuracy, and false-positive rate, of the current system?

Our data tell us in which cases the prosecutor brought a motion to detain. When judges deny or dismiss these motions and release these defendants, we can measure how much of a danger they pose to the public. Perhaps prosecutors and judges are both good at identifying dangerous defendants when they agree. But when they disagree, how accurate are the prosecutors?

Our results are shown in Table 7 and are consistent with earlier studies in New Mexico (Denman et al., 2021; Ferguson et al., 2020). The rate at which these defendants receive new charges is slightly above the rate among all released felony defendants: 4% versus 3% for violent felonies, and 3% versus 2% for violent misdemeanors. Their false-positive rate is 78%. As with the broadest legislative proposals examined above, detaining these defendants is not much more accurate than detaining a random sample of all defendants released by the current system.

Since these numbers depend both on judicial and prosecutorial discretion, and specifically on how prosecutors allocate their resources to bring motions to detain, we do not claim that they will necessarily generalize to other jurisdictions. We report them here in the hope that other jurisdictions will undertake similar studies.

DISCUSSION

We have shown that a wide variety of rebuttable presumptions have poor accuracy and a high false-positive rate. Despite the presumed intentions of policymakers, these proposals do not accurately target the small fraction of defendants who will be charged with new serious crimes if released pretrial. Instead, they cast a wide net, recommending detention for a large number of defendants who would not receive any new charges during the pretrial period.

We agree there are some defendants for whom the evidence of danger is so clear that pretrial detention is warranted. But as we have emphasized, New Mexico's current system already allows the prosecutor to present this evidence to a judge. Going beyond this individualized system and using legislation or

TABLE 7 Rearrest rates for defendants where prosecutors brought a motion to detain which the judge denied.

	No new charge	New charge: Nonviolent misdemeanor	New charge: Nonviolent felony	New charge: Violent misdemeanor	New charge: Violent felony	Total
All released felony defendants	12,388 (82%)	744 (5%)	1252 (8%)	295 (2%)	455 (3%)	15,134
Defendants for whom a PTD motion was brought but denied	1097 (78 ± 1%)	75 (5%)	139 (10%)	39 (3%)	61 (4%)	1411

Note: These are rearrested rates for defendants where a motion for pretrial detention (PTD) was brought, but where the judge denied or dismissed this motion and released the defendant. Except where shown, standard errors are less than 1%.

actuarial tools to target defendants for pretrial detention only makes sense if there are broad classes of defendants who are in fact dangerous. Our results show that, even using the type and severity of the current charge, past convictions, past violations of conditions of release, and past FTO, it is difficult or impossible to identify such a class.

Number needed to treat

In the medical literature, the number of people we need to give a treatment to prevent one bad outcome is called the "number needed to treat" or NNT. This is an easily interpreted measure of the cost of pretrial detention to individual liberties (and to the taxpayer) as compared with its benefits to crime prevention. For a typical legislative proposal from Table 6, we would need to detain about 5 people to prevent a single new charge, 11 or 12 to prevent a single violent charge, and at least 20 to prevent a single violent felony charge.[21]

Cost ratios

A high false-positive rate or NNT does not necessarily make a policy unacceptable. This depends on the ratio between the costs of false negatives and false positives, which is a matter for policymakers and stakeholders to determine. If policymakers believe that false negatives are 10 times worse than false positives—namely, that it is 10 times worse to release someone who then commits a serious crime than it is to incarcerate someone who would not have committed a crime—then a policy "breaks even" even if it detains 10 false positives for every crime it prevents, allowing a false-positive rate of about 91%.[22]

But the less accurate a policy is, the higher this ratio needs to be to justify adopting it—and if this ratio is high enough, it justifies detaining everyone. For instance, consider House Bill 27 from the last row of Table 6. Focusing on violent felonies and misdemeanors and equating them for simplicity, it would be justified if false negatives are about 12 times worse than false positives. If we raise this ratio to about 16, then it becomes justified to detain all felony defendants, which would presumably violate the constitution.

[21]These numbers are calculated by dividing 100% by 20% (the total rate of new charges), 8%–9% (the total rate of new violent charges including both felonies and misdemeanors) and 4%–5% (the rate of new violent felony charges).

[22]Note that this is the inverse of Blackstone's ratio: his maxim that it "is better that ten guilty persons escape than that one innocent suffer" suggests that false positives are more than 10 times worse than false negatives. A New Mexico opinion, State v. Chambers, 86 N.M. 383 (N.M. Ct. App. 1974) set this ratio at 99, in which case the tolerable false positive rate would be 1%. For a delightful review of different values of this ratio, see Volokh (1997).

Whether this utilitarian approach is appropriate in criminal justice is up for debate. Lawrence Tribe (1970) wrote, "the very enterprise of formulating a tolerable ratio of false convictions to false acquittals puts an explicit price on an innocent man's liberty and defeats the concept of a human person as an entity with claims that cannot be extinguished, however great the payoff to society."

On the other hand, the fact that our system makes errors of both kinds suggests that society, for better or worse, finds some finite ratio tolerable. What matters is that we be clear-eyed about the difficulty of this tradeoff, and cognizant of the costs to both victims and defendants.

Uses and misuses of the PSA

It is notable that, of all the criteria we examined, the most accurate is one based on the Arnold Public Safety Assessment, presumably because the PSA was designed to maximize the relationship between NCA scores and rearrest. But even among released defendants with an NCA score of 6 and the violence flag—those identified by the PSA as among the most dangerous—only 7% are charged with a new violent felony, so a policy that detains these currently released defendants would have an NNT of about 14. While this is superior to the legislative proposals we studied, we agree with the PSA's designers, and legal scholars such as Schnacke (2017), that the PSA should not be used on its own to recommend detention or to create a rebuttable presumption as in California's short-lived policy on defendants who are "assessed as high risk."

This is not to say that the PSA does not have predictive value. Like the evidence offered by prosecutors and the defense, PSA scores can and should be taken into account by a judge who is considering a motion to detain.[23] While some studies question the predictive value of risk assessment tools (Dressel & Farid, 2018) others have found that these tools are more accurate than human clinical judgments (Lin et al., 2020). Moreover, since we only have data on the behavior of released defendants, it is entirely possible that the PSA is good at distinguishing high-risk from low-risk defendants in general, even if it is only moderately good at doing so among those defendants who are currently released. But if so, our data suggest that judges are already doing a reasonable job of taking the PSA into account, and detaining the highest-risk defendants. While we are not constitutional scholars, we question whether using PSA scores to override judicial discretion would add

[23]New Mexico judges currently consider PSA scores as part of their release decisions. However, we would urge that judges receive quantitative information on the risks associated with those scores as measured in the local population, classified by the type and severity of new charges, as opposed to simply an abstract score like "5" or a color-coded category like "Orange.".

enough accuracy to counteract concerns about due process and individualized justice.[24]

Predictable crimes

There may be types of serial crime that are easier to predict. A machine learning approach to predicting rearrest for domestic abuse achieved an accuracy of 21%, that is, 21% of the defendants it identified were rearrested during the pretrial period (Berk et al., 2016). This algorithm was devised after consulting stakeholders on the relative costs of different kinds of false negatives and false positives. For instance, they set the cost of a pretrial rearrest for domestic violence with physical injury, or an attempt or threat of injury, to 10 times that of a false positive where this type of rearrest is predicted but no rearrest occurs. Tuning the algorithm based on this cost ratio gave a false-positive rate of 77%. We have not examined algorithms or legislation specific to domestic abuse.

Bias and fairness

We have focused on the low accuracy of rebuttable presumptions over the entire population, but it is also important to consider bias and disparities between demographic groups. Since pretrial detention has negative consequences for both case outcomes and social outcomes, it can perpetuate and amplify racial and ethnic inequalities.

Judicial discretion is one potential source of bias in pretrial detention. Release decisions are often made quickly with limited information, a setting in which bias can play a strong role. If judges apply stereotypes and overestimate the risk of defendants in a certain group, they will unnecessarily detain more of those defendants (Arnold et al., 2018; Mizel, 2018) although McIntyre and Baradaran (2013) argue that racial bias largely disappears when we control for the probability of rearrest for a violent felony. Moreover, if they release only very low-risk defendants from a group, presumptions may have a higher false-positive rate in that group.

Another potential source of bias is prosecutorial discretion. If prosecutors give more severe charges to some defendants than others even when the facts of the crime are similar, such as applying felony versus misdemeanor charges or labeling a crime as firearm-related due to more aggressive searches, then more

[24]For a valuable counterpoint, see Slobogin (2021), who argues that a jurisprudence of risk, informed by well-validated and regularly audited risk assessment algorithms, would be fairer than relying on human judges at many stages of the criminal justice system, including pretrial detention: for instance, that human judgments are less individualized than we like to think, and that much of the current level of incarceration can be traced to elected judges responding to political pressures. See also Goel et al. (2021).

defendants in that group will meet the criteria for presumptions of dangerousness. If the actual risk of these defendants is similar, this would again increase the false-positive rate in that group.

We hope to examine demographic differences in accuracy and false-positive rates in future work, but our preliminary explorations suggest that in our dataset they are broadly similar across groups except for a lower rate of new violent charges among women.

Evaluation versus design

Finally, we have focused on evaluating existing policies and proposals for rebuttable presumptions to inform the current debate over bail reform and pretrial detention. We have not attempted to *design* a rebuttable presumption policy. It is possible, and indeed likely, that by combining PSA scores, current charges, and other factors, one could develop a policy which is slightly more accurate than any of the particular ones we considered here. However, we would be very surprised if the false-positive rate of any such policy was significantly less than 70%, or if its NNT was less than 10; after all, this design exercise would essentially recapitulate the research behind the PSA. In particular, our data show that many of the input factors that a designer might be tempted to use—such as the severity of the current charge or prior failures to appear—have little predictive value among currently released defendants.

CONCLUSION

The National Association of Pretrial Services Agencies (NAPSA, 2020) states that "All defendants should have a statutory presumption of release on personal recognizance with the requirement that the defendant attend all court proceedings and not commit any criminal offense while released." The burden is on the government to overcome this presumption of release, not on the defense to overcome a presumption of dangerousness: "Pretrial detention should be authorized but limited only to when the court finds by clear and convincing evidence that a detention-eligible defendant poses an unmanageable risk of committing a dangerous or violent crime during the pretrial period or willfully failing to appear at scheduled court appearances."

Rebuttable presumptions relax this standard. They reduce judicial discretion by requiring judges to regard large classes of defendants as dangerous by default, rather than demanding that prosecutors prove this individually. Their proponents argue that they prevent a large amount of crime with a minimal impact on civil liberties. We have shown that this is not the case, both because a

small fraction of crime is committed by pretrial defendants, and because presumptions detain many defendants for each crime they prevent.

Our results should not be taken as an endorsement of the current system. Since we do not know how detainees would behave if they were released, the decisions of judges to detain are largely unfalsifiable (Goldkamp, 1985). We do not know what fraction of people currently detained are actually dangerous, and we are not saying that judicial discretion is a panacea.

What we have done is compare New Mexico's current system with one where presumptions are added to it, detaining additional defendants who are currently released. By measuring how dangerous these defendants are, we have shown that presumptions would not add accuracy to the system. That does not mean that the system could not be reformed in other ways.

The justice system makes many tragic mistakes: crimes that should have been prevented, and defendants that should not have been detained. Judges, algorithms, and legislators all try to prevent these mistakes by identifying dangerous defendants and releasing the rest. But crime is hard to predict, and basing pretrial detention on inaccurate predictions is neither effective nor just. Unless and until some much more accurate method of prediction appears, we should continue to require that prosecutors prove, case by case, that detention is justified.

ACKNOWLEDGMENTS

We thank Robert Berk, Sam Bowles, Kelly Bradford, Sharad Goel, Stephen E. Henderson, Elizabeth Hinton, Sandra Mayson, Timothy Schnacke, Rajiv Sethi, Cosma Shalizi, Christopher Slobogin, and Megan Stevenson for comments and feedback on a draft of this article; they may or may not agree with our conclusions. Cristopher Moore is funded in part by the National Science Foundation through grant BIGDATA-1838251. A preliminary and partial version of these results appeared in a report to the New Mexico Sentencing Commission (Ferguson, De La Cerda, O'Connell, & Guerin, 2021). We are grateful to Alise O'Connell and Helen De La Cerda for research assistance.

REFERENCES

Advancing Pretrial Policy and Research. (2022). *About the Public Safety Assessment*. https://advancingpretrial.org/psa/factors/

American Bar Association. (2007). *Standards for criminal justice: Pretrial release* (3d ed.) https://www.americanbar.org/content/dam/aba/publications/criminal_justice_standards/pretrial_release.pdf

Angel, A. R., Green, E. D., Kaufman, H. R., & Van Loon, E. E. (1971). *Preventive Detention: An Empirical Analysis*, 6 Harvard Civil Rights-Civil Liberties Law Review 300–396, and Forward by Sam J. Ervin Jr., Preventive Detention—A Step Backward for Criminal Justice.

Arnold, D., Dobbie, W., & Yang, C. S. (2018). Racial bias in bail decisions. *Quarterly Journal of Economics*, *133*, 1885–1932.

Bell, K., Hong, J., McKeown, N., & Voss, C. (2021). The recon approach: A new direction for machine learning in criminal law. *Berkeley Technology Law Journal*, *36*, 821–860.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Wiley.
Berger, M. (2023). *The constitutional case for clear and convincing evidence in bail hearings* (Vol. 75). Stanford Law Review (forthcoming).
Berk, R. A., Sorenson, S. B., & Barnes, G. (2016). Forecasting domestic violence: A machine learning approach to help inform arraignment decisions. *Journal of Empirical Legal Studies*, *13*(1), 94–115.
County of Los Angeles. (2021). *Pretrial Risk Assessment Tool Validation*. http://file.lacounty.gov/SDSInter/probation/1109503_LosAngeles-UPDATEDCountyvalidation-FINAL7-1-21.pdf
Denman, K., Siegrist, E., Robinson, J., Maus, A., & Dole, J. (2021). *Bail reform: Motions for pretrial detention and their outcomes*. New Mexico Statistical Analysis Center.
Léon Digard and Elizabeth Swavola. (2019). *Justice denied: The harmful and lasting effects of pretrial detention*. Vera Institute of Justice. https://www.vera.org/downloads/publications/Justice-Denied-Evidence-Brief.pdf
Julia Dressel and Hany Farid. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, *4*, eaao5580.
Durham, R. (2021). Innocent until suspected guilty. *University of Cincinnati Law Review*, *90*, 644–664.
Jeffrey Fagan and Martin Guggenheim. (1996). Preventive detention and the judicial prediction of dangerousness for juveniles: A natural experiment. *Journal of Criminal Law and Criminology*, *86*, 415–448.
Ferguson, E., De La Cerda, H., & Guerin, P. (2020). *Failure to appear and new criminal activity: Outcome measures for preventive detention and Public Safety Assessments*. Institute for Social Research, University of New Mexico.
Ferguson, E., De La Cerda, H., O'Connell, A., & Guerin, P. (2021). *The Public Safety Assessment, preventive detention, and rebuttable presumptions in Bernalillo County*. Institute for Social Research, Center for Applied Research and Analysis, University of New Mexico.
Ferguson, E., De La Cerda, H., Guerin, P., & Moore, C. (2021). *Bernalillo county Public Safety Assessment validation study*. Institute for Social Research, University of New Mexico. https://isr.unm.edu/reports/2021/bernalillo-county-public-safety-assessment-validation-study.pdf
Goel, S., Shroff, R., Skeem, J. L., & Slobogin, C. (2021). The accuracy, equity, and jurisprudence of criminal risk assessment. In *Research handbook on big data law* (pp. 9–28). Edward Elgar Publishing.
Goldkamp, J. S. (1985). Danger and detention: A second generation of bail reform. *Journal of Criminal Law and Criminology*, *76*(1), 1–74.
Hegreness, M. J. (2013). America's fundamental and vanishing right to bail. *Arizona Law Review*, *55*, 909.
Himsell, S. D. (1986). Preventive detention: A constitutional but ineffective means of fighting pretrial crime. *Journal Criminal Law & Criminology*, *77*(2), 439–476.
Kingsnorth, R., Barnes, C., Davis, C., Hodgins, T., & Nicholes, C. (1987). Preventive detention: The impact of the 1984 bail reform act in the eastern Federal District of California. *Criminal Justice Policy Review*, *2*(2), 149–172.
Kleinberg, J., Ludwig, J., Mullainathan, S., & Sunstein, C. R. (2020). Algorithms as discrimination detectors. *Proceedings of the National Academy of Science*, *117*(48), 30096–30100.
Koepke, J. L., & Robinson, D. G. (2018). *Danger Ahead: Risk Assessment and the Future of Bail Reform*. 93 Washington Law Review. https://ssrn.com/abstract=3041622
Lessig, L. (2006). Code 2.0. http://codev2.cc/download+remix/
Lin, Z. ". J."., Jung, J., Goel, S., & Skeem, J. (2020). The limits of human predictions of recidivism. *Science Advances*, *6*, eaaz0652.
Mayson, S. G. (2018). Dangerous Defendants. *127 Yale Law Journal*, *127*(3), 490–568.
Frank McIntyre and Shima Baradaran. (2013). Race, prediction, and pretrial detention. *Journal of Empirical Legal Studies*, *10*(4), 741–770.

Mizel, M. (2018). *A plea for justice: Racial bias in pretrial decision making*. Ph.D. thesis. UCLA.
National Association of Pretrial Services Agencies (NAPSA). (2020). Standards on Pretrial Release. napsa.org
Overbeck, A. M. (1986). Detention for the dangerous: The bail reform act of 1984. *Cincinnati Law Review*, *1986*, 153–199.
Pew Research Center. (2020). *What the data says (and doesn't say) about crime in the United States*. https://www.pewresearch.org/fact-tank/2020/11/20/facts-about-crime-in-the-u-s/
Schlesinger, T. (2005). Racial and ethnic disparity in pretrial criminal processing. *Justice Quarterly*, *22*(2), 170–192.
Schnacke, T. R., Jones, M. R., & Brooker, C. M. (2010). *The history of bail and pretrial release*. Pretrial Justice Institute.
Schnacke, T. R. (2017). *"Model" bail Laws: Re-drawing the line between pretrial release and detention*. Center for Legal and Evidence-Based Practices. http://www.clebp.org/images/04-18-2017_Model_Bail_Laws_CLEPB_.pdf
Siegrist, E. J., Dole, J. L., Denman, K., Maus, A., Robinson, J., Dorsey, C., & White, G. (2020). *Implementing bail reform in New Mexico*. Institute for Social Research, University of New Mexico.
Slobogin, C. (2021). *Just algorithms*. Cambridge University Press.
Stevenson, M. (2018). Assessing risk assessment in action. *103 Minnesota Law Review*, *303*, 82.
Tribe, L. H. (1970). An ounce of prevention: Preventive justice in the world of John Mitchell. *Virginia Law Review*, *56*, 371–407.
Viljoen, J. L., Jonnson, M. R., Cochrane, D. M., Vargen, L. M., & Vincent, G. M. (2019). Impact of risk assessment instruments on rates of pretrial detention, postconviction placements, and release: A systematic review and meta-analysis. *Law and Human Behavior*, *43*(5), 397–420.
Volokh, A. (1997). n guilty men. *Pennsylvania Law Review*, *146*, 173–216.

How to cite this article: Moore, C., Ferguson, E., & Guerin, P. (2023). How accurate are rebuttable presumptions of pretrial dangerousness? A natural experiment from New Mexico. *Journal of Empirical Legal Studies*, *20*(2), 377–408. https://doi.org/10.1111/jels.12351

APPENDIX

Tables of charges in New Mexico law used in this study

TABLE A1 Serious Violent Offenses according to New Mexico Statute §33-2-34 (2018) subparagraphs L(4)(a–n). These constitute the "narrow" definition of SVO.

(a) Ssecond- degree murder
(b) Voluntary manslaughter
(c) Third-degree aggravated battery
(d) Third-degree aggravated battery against a household member
(e) First-degree kidnapping
(f) First- and second-degree criminal sexual penetration
(g) Second- and third-degree criminal sexual contact of a minor
(h) First- and second-degree robbery
(i) Second-degree aggravated arson
(j) Shooting at a dwelling or occupied building
(k) Shooting at or from a motor vehicle
(l) Aggravated battery upon a peace officer
(m) Assault with intent to commit a violent felony upon a peace officer
(n) Aggravated assault upon a peace officer

TABLE A2 Additional offenses that the court can find to be a serious violent offense (SVO) according to New Mexico Statute §33-2-34 (2018) subparagraph L(4)(o). The "broad" definition of SVO includes these charges as well as the charges in Table A1.

(1) Involuntary manslaughter
(2) Fourth-degree aggravated assault
(3) Third-degree assault with intent to commit a violent felony
(4) Fourth-degree aggravated assault against a household member
(5) Third-degree assault against a household member with intent to commit a violent felony
(6) Third- and fourth-degree aggravated stalking
(7) Second-degree kidnapping
(8) Second-degree abandonment of a child
(9) First-, second-, and third-degree abuse of a child
(10) Third-degree dangerous use of explosives
(11) Third- and fourth-degree criminal sexual penetration
(12) Fourth-degree criminal sexual contact of a minor
(13) Third-degree robbery
(14) Third-degree homicide by vehicle or great bodily harm by vehicle
(15) Battery upon a peace officer

TABLE A3 Firearm-related charges in New Mexico law.

Unlawful Carrying of a Firearm in Licensed Liquor Establishment
Receipt, transportation, or possession of a firearm, deadly weapon, or destructive device or destructive device by a prisoner or Felon
Negligent use of a deadly weapon (discharge)
Shooting at or from a vehicle
Shooting at dwelling or occupied building
Unlawful possession or discharge of weapons or firearms

TABLE A4 Additional deadly weapon-related charges in New Mexico law. This list is somewhat simplified: there are specific statutes for assault or battery on household members, health care workers, and peace officers. There are also additional statutes for attempting, conspiring, or soliciting to commit many of these felonies.

Aggravated assault (deadly weapon)
Aggravated battery—felony (great bodily harm or deadly weapon)
Aggravated burglary (armed after entering)
Aggravated stalking (possession of a deadly weapon)
Armed robbery with deadly weapon
Criminal sexual contact/minor (deadly weapon)
Criminal sexual contact in the fourth degree (deadly weapon)
Criminal sexual penetration in the second degree (armed with a deadly weapon)
Negligent use of a deadly weapon (intoxication, near dwelling or building, unsafe handling)
Negligent use of explosives
Unlawful carrying of a deadly weapon
Unlawful carrying of a deadly weapon on school premises
Unlawful carrying of a handgun by a person under age 19
Unlawful negligent use of weapons
Unlawfully carrying deadly weapons