Judicial Errors, Crime Deterrence and Appeals: Evidence from U.S. Federal Courts

Roee Sarel*

Draft version: December 4, 2018

Abstract

This paper investigates how different outcomes of criminal appeals impact crime deterrence. My panel data analysis (1997-2013) indicates that crime rates decrease when convictions are affirmed by the appellate court, as long as conviction rates are not too high. Crime rates are further found to decrease when convictions are reversed but increase when cases are remanded back to the district courts. The findings imply that judicial errors are detrimental to deterrence, while their review on appeal is Janus-headed: reversals provide remedy whereas remands are counter-productive. I therefore offer ways to overcome the negative effects of remands, conditional on potential problem sources: uncorrected errors, discounting of sanctions and cost asymmetry between guilty and innocent defendants.

JEL: K

Keywords: Judicial errors, wrongful convictions, crime deterrence, crime rates, appeals

1. Introduction

The economic analysis of crime assumes that individuals decide whether to commit crimes based on a cost-benefit analysis. In their decision-making process, these individuals compare the expected net benefit from crime on one hand and the expected opportunity cost of crime on the other hand, and then choose the more profitable option. Among else, this calculation requires an estimation of the expected sanctions,

*Frankfurt School of Finance and Management, D-60322 Frankfurt Am Main, Germany. Email: r.sarel@fs.de. I am grateful to Yun-Chien Chang, Robin Christmann, Darwyn Deyo, Dhammika Dharmapala, Eberhard Feess, Hartmut Kliemt, Alessandro Melcarne, Murat Mungan, Avraham Tabach, Eric Talley, Ronald Wintrobe and Ansgar Wohlschlegel for useful comments. I also thank participants of the following conferences: American Economic Association (Chicago, Jan. 2017), European Law & Economics association (Bolonga, Sep. 2016), German Law & Economic association (Budapest, Jul. 2016), Conference on Empirical Legal Studies in Europe (Amsterdam, Jun. 2016), European Public Choice Society (Freiburg, Mar. 2016), Public Choice Society (Fort Lauderdale, Mar. 2016) and the law faculty seminars at the Hebrew University of Jerusalem and Tel Aviv University (Nov. 2018).

i.e. the product of the (1) probability of apprehension and conviction and (2) the sanction size. For the calculation to be accurate, all possible scenarios must be considered - including those in which a trial court decides to convict, but the conviction is later overturned on appeal. Hence, appeal outcomes should, at least in theory, matter for crime deterrence.

Nonetheless, existing analyses mostly abstract away from appeals and focus only on the compounded probability of receiving a sanction, without refining the exact components of this probability. An exception to this approach is a recent scholarly debate surrounding the effect of judicial errors on crime deterrence. The mainstream view (e.g. Png, 1986), which I shall refer to as the ”Extended Becker Model” (“EBM” for short; referencing the canonical Becker (1968) model), attributes type I errors (wrongful convictions) to the opportunity cost side of crime and argues that such errors reduce the payoff from innocence. Therefore, ceteris paribus, a higher probability to be wrongfully convicted increases the incentive to commit a crime. Respectively, type II errors (wrongful acquittals) belong to the benefit side and represent the prospect of “getting away” with crime. Thus, ceteris paribus, a higher probability to be wrongful acquitted also increases the incentive to commit a crime. Critics of this approach, however, claim that judicial errors may be beneficial (e.g. by causing individuals to be “extra careful”; see Craswell and Calfee, 1986; Kaplow, 2011b) or inconsequential (e.g. because wrongful convictions for someone else’s crime can occur irrespective of whether the individual commits another crime; see Lando, 2006) to deterrence. Although the theoretical results diverge, existing analyses all share the common assumption that each individual self-selects into a specific level of criminal activity (which may be dichotomous, i.e. “innocent” vs. “guilty”, or continuous) which then corresponds to a conditional expected sanction.

Extending these theories to incorporate a second stage - in which the judicial errors may be corrected by an appellate court - seems straightforward: if an error has been made, but later corrected, then the (ex-ante) impact of appeal outcomes is simply a reduction in the relevant expected sanction. However, three main challenges stand in the way of such an exercise. First, appeal outcomes are typically not dichotomous but rather combine one or more categories, such as (1) affirmance, (2) reversal, and (3) remand to a lower court. Therefore, one must carefully consider what these different categories actually imply for expected sanctions. Second, the defendant’s type is always unobservable, as we simply do not know which trial decisions are erroneous and which are not - if we would, the errors would probably not occur in the first place. As a result, one can usually observe only changes in the unconditional probability for the realization of each outcome. Third, as the accuracy of both trial courts and appellate courts is unobservable (for the same reason), it may become difficult to disentangle whether changes stem from a shock to the trial courts' accuracy (e.g. the trial court reduces its error rate, leading to less reversals) or to the appellate courts' accuracy (e.g. the trial court's error rate did not change, but the appellate court nonetheless reverses more).

These challenges notwithstanding, an empirical assessment of how appeal outcomes impact deterrence is important in order to evaluate whether appellate courts achieve their declared goals. Specifically, as judicial errors entail a welfare loss - among else due to potential infringement of deterrence - social planners typically establish appellate courts in order to intervene where errors have been detected. If this intervention does not always achieve its goal, then the justification for the existence of appellate courts becomes weaker.

I develop a stylized theoretical model which allows to generate clear predictions on the effects of appeal outcomes on crime deterrence. The model builds on the key assumption that appellate courts are more accurate than trial courts, such that a change in the rate of a given outcome is generally correlated with a correct identification of errors rather than the creation of "new errors" on appeal. However, my analysis also assumes that the public's belief on the degree of appellate court accuracy may vary with other factors.

Several empirical research methods are then employed in order to assess whether potential offenders respond to changes in appeal outcomes. I construct a new panel dataset which tracks the frequency of each outcome - affirmance, reversal and remand - in the U.S. federal appellate courts over a period of 17 years (1997-2003), alongside corresponding crime rates. This allows me to check how the change in the rate of each outcome correlates with crime, while controlling for unobservables using fixed effects (and other relevant controls). Additionally, I use an instrumental variable approach ("IV") in order to tackle potential endogeneity, where lagged values and ideological distances between judges in different court levels are used as an instrument. Ideological distances are known predictors of judicial outcomes in the federal appellate courts (e.g. Hettinger and Lindquist, 2012; Iaryczower, Lewis, and Shum, 2013; Boyd, 2015; Borochoff, 2008; Sarel and Demirtas, 2017) but are also exogenous, as they typically change due to reasons unrelated to contemporaneous crime rates (e.g. retirement or death of judges).

My analysis yields two key findings. First, I find that crime rates are lower when affirmance rates increase, as long as conviction rates are not too high. A negative effect of affirmance rates on crime is consistent with the EBM, as more affirmances typically imply that fewer errors occurred at trial. The dependency of the effect on conviction rates, however, captures the change in the aforementioned public's belief: if conviction rates are too high, an increase in the affirmance rate is no longer interpreted as less errors at trial but rather as arbitrary decision making on appeal. This entails both an information effect (i.e. affirmance rates are perceived to be less correlated with the underlying rate of errors) and a trust effect (i.e. the justice system is perceived as unreliable).

Second, I find that crime rates are lower when the Reversal-Remand spread (i.e. reversal rate minus remand rate) increases, which implies that reversals decrease crime while remands increase crime. A negative effect of reversals is again consistent with the EBM, as more reversals imply that when errors do occur - they are corrected (thus, ex-ante, the compounded probability of a wrongful conviction decreases). A positive effect of remands on crime rates is similarly consistent, as remands imply that (1) wrongful convictions are not directly reversed and (2) correct convictions are not directly affirmed. For innocent defendants, remands are particularly discouraging if the lower courts are unlikely to change their original outcome of conviction.

The paper's contribution is then two-fold. First, to the best of my knowledge, this paper is the first to provide (observational) empirical evidence on how appeal outcomes impact crime deterrence, which are important for the evaluation of appellate court performance, especially given the conflicting theories on the effects of judicial errors. Second, I identify a negative impact of remands on crime deterrence, discuss possible explanations - uncorrected errors, time-preferences, and litigation costs - and propose solutions.

The paper is related to several streams of literature. First, the paper builds on deterrence models with type I and II errors (the EBM), whose origin is usually attributed to Png (1986) (for earlier models which explicitly consider judicial errors, see also Harris, 1970; Posner, 1973).[1]

This approach argues that type I errors reduce deterrence by decreasing the opportunity cost of crime and has been widely adopted in the literature (see, e.g. Kaplow, 1994; Dari-Mattiacci and Deffains, 2007; Polinsky and Shavell, 2000, 2001, 2007; Mungan, 2015b). Conversely, other papers argue that type I errors *increase* deterrence. e.g. if they are (mistakenly) perceived as correct convictions (Ehrlich, 1982) or when the action space is continuous and over-compliance breeds a higher payoff (Craswell and Calfee, 1986). A similar line of argument views type I errors as creating over-deterrence, by chilling benign behavior (Kaplow, 2011a,b, 2012, 2014) or by diluting the stigma which accompanies convictions (Epps, 2015; Mungan, 2015a). Other papers argue that type I errors have *no effect* on deterrence, e.g. because mistakes of identity may occur irrespective of whether the defendant chooses to commit a (different) crime (Lando, 2006; Lando and Mungan, 2017). Garoupa and Rizzolli (2012) argue conversely, that identity errors indirectly reduce deterrence, as they typically imply that the actual perpetrator is not punished (i.e. a type

[1] The original Becker model allows the probability of apprehension and conviction to be lower than 100% and therefore implicitly considers wrongful acquittals. Wrongful convictions are not explicitly addressed, although some aspects of the opportunity cost of crime are captured by the portmanteau variable u_j (see Becker, 1968, pp. 10 footnote 15).

II error). Epps (2015, pp. 1126-1128) notes that such 'type II errors' stem from police behavior, rather than adjudication. Obidzinski and Oytana (2017) argue that identity errors induce regulators to set a higher burden of proof, thereby reducing deterrence.[2] These miscellaneous papers mostly focus on judicial errors and do not explicitly consider appeals.

Second, the paper is related to the vast empirical literature on crime (see Chalfin and McCrary, 2017, for a review). Many papers use proxies for the probability of apprehension and conviction, such as police manpower (e.g Levitt, 2002; Evans and Owens, 2007), police tactics (e.g. Weisburd, Morris, and Groff, 2009), arrests (e.g. Machin and Meghir, 2004) and judicial vacancies (Yang, 2016). Criminolgy papers follow a similar path, but usually investigate how perceived (rather than actual) probabilities impact deterrence (e.g. Bachman, Paternoster, and Ward, 1993; Pogarsky and Piquero, 2003). Some papers consider the opportunity cost of crime, by exploring the effects of (un)employment (e.g. Anderberg et al., 2016; Raphael and Winter-Ebmer, 2001; Ihlanfeldt, 2007; Lin, 2008) and wages (e.g. Gould et al., 2002), but do not explicitly measure type I errors. Papers on time preferences and deterrence (e.g. Dalla Pellegrina, 2008; Soares and Sviatschi, 2010; Dušek, 2015; Lee and McCrary, 2017) mostly discuss court delays and yield mixed results as to whether discounting of sanctions leads to less deterrence.

Experimental research on deterrence has been equally extensive (for a review, see Engel, 2016). Experiments which consider judicial errors typically assign subjects the role of thieves who can steal money and then vary the error probabilities either explicitly (e.g. Rizzolli and Stanca, 2012; Rizzolli and Tremewan, 2018) or implicitly, using other subjects in the role of judges (e.g Feess et al., 2018; Baumann and Friehe, 2015) (see also Markussen, Putterman, and Tyran (2016) for a public goods game setting). One experiment (Lewisch, Ottone, and Ponzano, 2015) touches upon the appeal system's effect on deterrence and indicates that adding a second instance may increase deterrence, but only allows for a binary appeal outcome.

Third, the paper is related to empirical papers on judicial errors, outside the context of deterrence. These papers use indirect measurements such as agreement rates of judges and juries (e.g. Spencer, 2007; Kim et al., 2013), prison surveys (Loeffler, Hyatt, and Ridgeway, 2018), expert panels (e.g. Gould et al., 2014) or exonerations (Gross and O'Brien, 2008; Gross and Shaffer, 2012; Gross et al., 2014; Olney and Bonn, 2014; Bjerk and Helland, 2017) - but do not consider deterrence effects.

[2]Other aspects of the EBM have also been challenged, e.g. the EBM's argument that type I and II errors have an *equal marginal effect* on deterrence (Png, 1986; Rizzolli and Saraceno, 2013). Rizzolli and Stanca (2012) and Nicita and Rizzolli (2014) argue that type I errors are more detrimental to deterrence than type II errors, due to risk aversion; rank dependent utility; loss aversion *á la* prospect theory (Tversky and Kahneman, 1979, 1992); and indignation, instigated by wrongful convictions. Contrarily, Lando and Mungan (2017) assert that the scope of the EBM's symmetrical result of error types is limited to some special cases.

Finally, my analysis is related to papers discussing the reasons for wrongful convictions (e.g. Kaplow and Shavell, 1994; Bar-Gill and Gazal Ayal, 2006; Gould and Leo, 2010; Gould et al., 2014), the process of error correction on appeal (e.g. Gennaioli and Shleifer, 2008; Griffin, 2000), the behavior of judges (e.g. Kuersten and Songer, 2014; Epstein, Landes, and Posner, 2013; Rose, Casarez, and Gutierrez, 2018) and prosecutors (King and Heise, 2018) in the federal courts, and information dissemination among criminals (e.g. Bebchuck and Kaplow, 1992; Johnson and Payne, 1986; Garoupa, 1999; Buechel, Feess, and Muelheusser, 2018).

The rest of the paper is organized as follows: in section 2, I develop a stylized EBM-like theoretical model, which is then extended into hypotheses in Section 3. Section 4 describes the data, variables, and methodology. Section 5 presents the basic results. Section 6 reviews robustness tests and includes analyses of additional data and instrumental variables. Section 7 discusses implications and concludes.

2. Model

Consider a continuum of individuals who need to decide whether to commit a crime. Let b_j be the benefit from crime of individual j. Without loss of generality, I assume that b_j is uniformly distributed on the interval $(0, \bar{b})$. Each individual faces a dichotomous choice $D \in \{I, G\}$, between remaining innocent ($D = I$) and becoming guilty ($D = G$) by committing a crime. For each action, the individual is exposed to some potential sanction f (e.g. a monetary fine or a prison sentence), which is incurred with the conditional probability p_D.[3] The expected sanction is thus $E(f|D) = p_D * f$.[4] The expected payoffs are $b_j - p_G * f$ from crime and $-p_I f$ from innocence. The individual therefore commits the crime if and only if:

$$b_j > b^* \equiv E(f|G) - E(f|I) = (p_G - p_I)f \tag{1}$$

where b^* reflects the deterrence threshold, above which it is profitable to commit the crime. Note that, in line with the EBM's approach, both types of judicial errors reduce deterrence, as $\frac{\partial b^*}{\partial p_I} < 0$ implies that wrongful sanctions reduce deterrence whereas $\frac{\partial b^*}{\partial p_G} > 0$ implies that correct sanctions increase deterrence – such that wrongful acquittals reduce deterrence.

I proceed by introducing appeals into the model. Suppose that the overall probability of being sanctioned (p_D) can be broken-down into different scenarios, where a convicted defendant may appeal his conviction.

[3] One may want to additionally assume that the probability of being sanctioned is higher when choosing to commit the crime, such that $p_G > p_I$. However, this assumption is not required for the results to hold.

[4] Allowing the sanction f to be conditional on the decision, as in Sarel (2018), does not qualitatively affect the model's results.

The appellate court reviews the case and can reach one of three outcomes: affirm the conviction (i.e. the sanction remains in place), reverse the conviction (i.e. there is no sanction) or remand the case back to a lower court (i.e. there is an additional process, which may end in the sanction remaining in place or canceled). Let c_D be the probability of conviction by the trial court and the probabilities for each possible appeal outcomes be: a_D for affirmance, r_D for reversal, and m_D for remand. As mentioned in the introduction, I assume that appellate courts reach more accurate outcomes than trial courts, such that (without loss of generality) $r_G = 0$ and $a_I = 0$, i.e. correct convictions are not reversed and incorrect convictions are not affirmed. In other words, when the appellate court decides a case of a guilty defendant, the conviction may be affirmed or remanded; whereas in a case of an innocent defendant, the conviction may be reversed or remanded.

If the case is remanded, the expected sanction is lower, which is expressed by a discount factor $0 < \delta_D < 1$. This discount factor captures three aspects: first, the possibility that the trial court will change its original verdict after a remand and decide to acquit (i.e. that the probability of receiving f is lower than 1). Second, the possibility that the trial court will maintain the conviction but reduce the sentence. Third, the impact of time preferences, in case that the sentence will be 'stayed' pending the remand process (I discuss this feature below in the hypotheses section). Figure 1 summarizes the decision tree that individuals face, where the upper (lower) branch is the decision (not) to commit a crime. For brevity, trial acquittals (pre-appeal), which occur with probability $1 - c_D$ and yield a payoff of zero, are omitted from the figure.

Rewriting the deterrence threshold yields:

$$b^* = [c_G(a_G + m_G\delta_G) - c_I m_I \delta_I] \times f \tag{2}$$

The effects of the different appeal outcomes are then summarized in the following proposition (all proofs are in Appendix A):

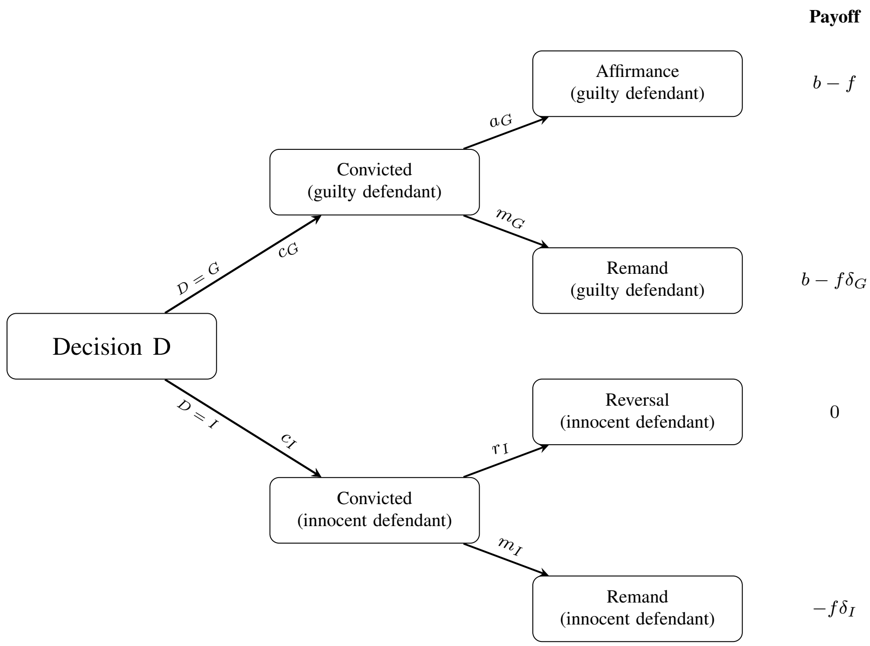
Proposition 1 *:*

(i) Deterrence increases in the probability of affirmance (of correct convictions) (i.e $\frac{\partial b^*}{\partial a_G} > 0$*)*

(ii) Deterrence increases in the probability of reversal (of incorrect convictions) (i.e $\frac{\partial b^*}{\partial r_I} > 0$*)*

(iii) Deterrence decreases in the probability of remand (of any conviction) (i.e $\frac{\partial b^*}{\partial m_D} < 0, \forall D \in I, G$*)*

Figure 1: Decision tree



For parts (i) and (ii), note that both affirmances and reversals have a positive impact on deterrence (which may sound strange at first, given that affirmance means the defendant lost the appeal whereas reversal means the opposite) but refer to different types of defendants: affirmances increase the expected sanction when committing the crime, whereas reversals decreases the expected sanction when remaining innocent. In this sense, the two effects are simply the mirror image of the EBM's claim regarding judicial errors, such that fixing the (deterrence-reducing) errors improves deterrence.

For part (iii), remands have a negative impact on deterrence both when the defendant is guilty (m_G) and when the defendant is innocent (m_I). The first effect occurs because for guilty defendants, remands come at the expense of less affirmances. Then, instead of receiving a full sanction, (guilty) defendants receive a discounted sanction. The second effect occurs because for innocent defendants, remands come at the expense of reversals. Thus, even though the sanction is discounted, it is still higher than no sanction at all.[5]

Note that these effects clearly carry over even when relaxing the simplifying assumption of zero incorrect affirmances and zero incorrect reversals, as long as appeal outcomes are sufficiently correlated with the defendant's type. In other words, even if there are some incorrect reversals/affirmances, as long as affirmances (reversals) are sufficiently correlated with the defendant being guilty (innocent), the theoretical

[5] This result is reinforced if one assumes that $\delta_I > 1$, e.g. because of litigation costs which enhance the sanction above and beyond the fine/sentence itself.

results remain similar.

3. Hypotheses development

The stylized model above predicts that deterrence will increase in the affirmance rate, increase in the reversal rate, and decrease in the remand rate. Albeit these predictions are useful as a benchmark, the empirical setting introduces some challenges for estimating the predicted effects.

First, classifying appeal outcomes to match the categories discussed in the model (affirmance, reversal, and remand) may be difficult, as cases may belong to more than one category (e.g. cases that are reversed and remanded) or to other categories (e.g. a dismissal).[6] To overcome this challenge, I exploit the existing classification of the federal courts' database, which includes the three categories of the model explicitly, alongside two other categories (dismissal and "other"). In my robustness tests, I exploit a similar classification system (with some additional sub-categories such as "partial affirmance") used by the U.S. Sentencing Commission ("USSC").

Second, one may question whether appellate courts are more accurate than trial courts, as assumed by the model. For example, trial courts are directly exposed to the evidence and witnesses and may have expertise in evaluating evidence (so called "institutional superiority"; see Hessick (2012)), whereas appellate courts are usually limited to examining the legal questions as raised in the trial court's decision. Nonetheless, the assumption that appellate courts are indeed more accurate is common in the literature and is based on the notion that appellate judges are either more qualified or better informed than trial judges. For example, it has been argued that the appellate process harvests private information from litigants (e.g. Shavell, 1995), with some papers going so far as to assume a perfectly accurate appellate court (e.g. Levy, 2005; Shavell, 2006). Furthermore, even if appellate courts sometimes err - the random assignment of appellate judges to cases (e.g. Hettinger and Lindquist, 2012, pp. 128) may prevent systematic errors. However, even with superior ability, target-functions of appellate courts must also be considered. For example, appellate courts may wish to promote legal coherence, reach a certain ideological outcome or minimize effort cost (for reviews, see Sarel and Demirtas, 2017; Knight and Epstein, 2017). While these could work against accuracy maximization, many models assume that appellate judges either inherently aim at reaching correct decisions or are incentivized to do so given a disutility from reversal by a higher court (for a review, see Feess and

[6] Dismissals function similarly to affirmances, as they imply the appeal is rejected. However, as the appellate court does not explicitly declare that no error occurred, it is hard to interpret whether dismissals occur in cases where the defendant is guilty or innocent. As dismissals carry no error correction effect, I remain agnostic about their effect in this paper.

Sarel, forthcoming). Thus, assuming superior accuracy at the appeal level seems plausible.

Third, even if appellate courts are more accurate, individuals must perceive them as such in order for appeal outcomes to set incentives as depicted in the model. While it seems plausible to assume that potential criminals generally do *perceive* appellate courts as more accurate, their belief may vary with other factors. Namely, I assume that potential criminals use conviction rates (at trial) as a signal on whether the identification of errors by the appellate court is accurate. This mechanism works as follows: if conviction rates are low, individuals will attribute this to proper screening by the trial court judges. Thus, if a conviction is affirmed, the appellate court is perceived as fair and accurate. Conversely, if conviction rates become too high, individuals expect some convictions to be erroneous. Then, a higher affirmance rate would imply that the appellate court either failed to identify the error - or has allowed judicial preferences to determine the outcome irrespective of the case merits. In other words, if practically everyone is convicted and then also everyone lose their appeal, affirmances would bear a different interpretation. This can then entail two effects: an information effect, where affirmances are perceived as simply less correlated with $E(f|G)$, and a (mis)trust effect which may nudge individuals toward crime commission (e.g. by imposing a perceived cost for innocence, i.e. an increase in $E(f|I)$).[7] Hence, I hypothesize as follows:

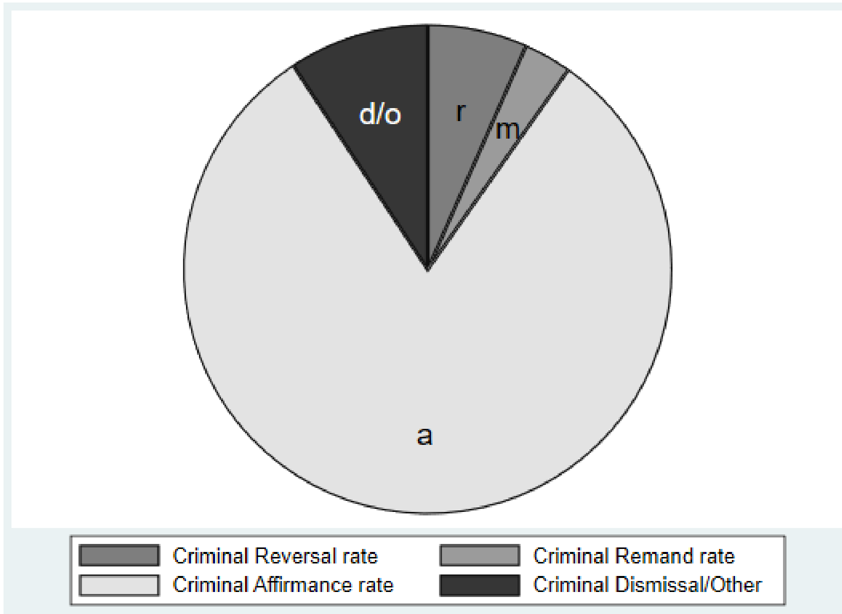
H1: Affirmance rates will be negatively correlated with crime rates.

H2: The effect size of affirmance rates will be smaller when conviction rates are high.

Another, fourth, challenge entails the construction of independent variables. Ideally, one would need to observe how a change in only one probability (of affirmance, reversal, or remand) is correlated with crime rates. However, as these probabilities are all "shares of the same pie", it makes little sense to check for the effect of one probability while keeping all other shares constant. Instead, the variables should reflect the two main phenomena of interest: error *identification* and error *correction*. The former is already captured by the affirmance rate (H1 and H2). The latter should then entail a measure of the mix of reversals and remands, in order to see how the appellate court's choice of correction tool is correlated with crime. There are two straightforward variables which may fulfill this criteria: the relative share of reversals among cases which have not been affirmed (nor dismissed), or the difference between the shares of reversals and remand, i.e.

[7] The possibility that mistrust may nudge individuals toward crime has been raised in various forms. For example, Nicita and Rizzolli (2014) argue that the feeling of indignation as a result of a wrongful conviction may encourage crime commission. The intuition is somewhat similar to the "broken windows theory" (see e.g. Corman and Mocan, 2005), where the mere appearance of civil disorder is sufficient to facilitate crime.

Figure 2: Illustration: appeal outcomes distribution



the *Reversal-Remand spread*. To illustrate, consider Figure 2 which shows an example for a distribution of appeal outcomes: affirmances ("a"), reversals ("r"), remands ("m") and dismissal/other ("d/o").

Clearly, one cannot change one share while keeping all shares equal. However, if "a" and "d/o" are kept constant the mix of "r" and "v" can change. The relative share of reversals, i.e. $\frac{r}{r+m}$ and the spread $r - m$, then both capture the same tradeoff. For convenience, I opt for the using the Reversal-Remand spread for the analysis.[8] The results remain largely the same when using the share of reversals.[9]

Note that the Reversal-Remand spread increases either when the reversal rate increases OR when the remand rate decreases, which implies that deterrence should increase when the spread increases. Hence, I hypothesize that:

H3: The Reversal-Remand spread will be negatively correlated with crime rates.

However, there are again several challenges worth noting. While the reversal rate is assumed to relate to $E(f|I)$ only, the observable remand rate is the overall unconditional rate for both guilty and innocent defendants (i.e. $m_G + m_I$). Thus, albeit a higher probability of remand reduces deterrence for both defendant types under the model, the effect size may differ. For example, recall that remands are assumed to lead to a

[8] Among else, using the spread circumvents a technical problem which arises when, in some specific court and year, there were no reversals and no remands (such that the denominator of the reversal share would be zero).

[9] In particular, both the basic results and the instrumental variables approach yield similar results. Results in robustness checks using sentencing cases somewhat differ in significance levels. All results using the share of reversals are available upon request.

lower sanction, but the source of reduction was captured in the model under one, unspecific, discount factor. If expected sanctions are lower because of the trial court's probability to change its mind, such a scenario seems more likely for innocent defendants.[10] Conversely, if the sanction is lower because of resentencing, it is not clear ex-ante whether innocent defendants will receive larger reductions (for a discussion of punishing the innocent, see Sarel, 2018). Furthermore, if the source of reduction is discounting due to time preferences, two questions emerge: first, are sanctions in the remanded cases delayed ("stayed") during the remand process? If the answer is no, there should be no discounting, as the defendant already received the sanction (e.g. has begun serving a prison sentence).[11] Second, how large is the discount factor? While defendants may generally have heterogeneous time preferences, some argue that guilty defendants discount sanctions more heavily (Davis, 1988; Silverman, 2004; Beraldo, Caruso, and Turati, 2013; Mastrobuoni and Rivers, 2016) or even hyperbolically (Lee and McCrary, 2017).[12] However, a delayed process can also – by itself – create mistrust in the judicial system (Dalla Pellegrina, 2008) or cause evidence quality to deteriorate (Torre, 2003, pp. 102) and reduce error corrections, which means that a prolonged process may have a negative impact on deterrence irrespective of the defendant's type.

Prolonged proceedings following a remand may also lower deterrence through *litigation costs*, but only if such costs are extremely high for innocent defendants. Namely, higher litigation costs are similar to a monetary fine, which would decrease deterrence only if it increases the expected sanctions for innocents ($E(f|I)$). One reason why litigation costs may indeed be higher for innocents is that substantiating a strong, but costly, defense is worthwhile only if the defendant can actually prove his innocence. In other words, guilty people may invest less, as their chances of "proving" innocence are low (Johnson, 2016, pp. 264). An additional challenge is that not all remands are 'born equal'. For example, remands may include instructions to lower courts, which then encourage a correction (see, e.g. Beery, 2002; Boyd, 2015). Conversely, if remands are made back to the same judicial quorum that reviewed the case originally, the risk of a possible

[10] In some circumstances, remanding may be obligatory. For example, 18 U.S. Code 3742(f) states that if the sentence is in violation of law, the case should be remanded. A violation of law may be more likely for innocents.

[11] Note that two combinations can lead to less deterrence: (1) the defendant is innocent and the sentence is *not* stayed, such that time spent in custody pending remand reduces the payoff from innocence (see Domenech and Puchades, 2014, for a similar argument on wrongful arrests), OR (2) the defendant is guilty and the sentence *is* stayed. A stay of sentence depends, inter alia, on the type of sanction: death sentences are automatically stayed. Imprisonment is stayed if a request is granted. Fines are stayed if a request is granted, but a stay might include a condition of depositing money as a guarantee. See rule 38(a) of the Federal Rules of Criminal Procedure; Rule 8 of the Federal Rules of Appellate Procedure. As defendants seeking a stay must generally show that they are likely to succeed on the merits (see *Nken v. Holder*, 556 U.S. 418 (2009); *Leiva-Perez v. Holder*, 640 F.3d 962 (9th Cir. 2011)), sentences of innocent defendants may be more likely to be stayed.

[12] Strong discounts by guilty defendants can explain why some papers find a positive relationship between crime rates and court delays. See Torre (e.g. 2003) for a model and Soares and Sviatschi (2010); Dušek (2015) for (mixed) empirical evidence. Findings from Italy (Dalla Pellegrina, 2008) also indicate that trial delays reduce deterrence but appeal delays do not.

re-enactment of the previous result seems higher.[13]

Similarly, reversals may also differ in whether the person is formally exonerated. While it is tempting to restrict attention to exonerations, their relationship with the EBM is ambiguous. First, exonerations typically occur many years after the conviction, such that they do not actually reflect the contemporaneous probability of a type I error at the time of crime commission. Second, exonerations may reflect cases where even the appellate process did not yield a correct result. Third, exonerations may suffer from a severe selection bias, where only a specific sub-set of cases is considered. For example, exonerations relying on DNA evidence may correspond mainly to identity errors, where the actual culprit left DNA at the scene. Similarly, exonerations may exclude defendants with insufficient resources for access to representation (Olney and Bonn, 2014). Fourth, wrongful convictions due to a legal error (rather than factual errors) may be under-represented in exonerations.[14] In light of the above, treating reversals homogeneously seems superior to considering only exonerations.

Note that these various issues regarding remands and reversals mostly relate to their effect size, whereas the direction of the effect remains clear: remands (reversals) are expected to decrease (increase) deterrence.

4. Data, variables and methodology

4.1. Sources

Data on U.S. federal courts (1997-2013) was gathered from the "Judicial Business of the U.S. courts" reports. The data contains information on proceedings in each district court: prosecuted defendants, conviction rates;[15] bench/jury trials; punishment type (fine/imprisonment), and rates of appeal outcomes. Data on sentencing cases was derived from USSC databases, as published by the Inter-university Consortium for Political and Social Research ("ICPSR"). Additional data was extracted from the "Supreme Court Database" (Spaeth, Epstein et al., 2016) and online data tools of the Bureau of Justice Statistics. I exclude three territories (Virgin Islands, Northern Mariana Islands, Guam), due to a low quality of data.[16] Ideological distances were derived from "judicial common space scores" (see section 6.6 below), as downloaded from publicly

[13] I discuss the problem of 'same-panel' remands in detail below, in section 7 below.

[14] Albeit legal errors do not completely overlap with factual errors, those errors that are reviewed by appellate courts in criminal cases are likely to serve as a good proxy for type I errors, as defined by the EBM. Namely, as long as legal errors identified by appellate courts (1) facilitate an appealed conviction and (2) are more likely to occur when the defendant is innocent – such errors will reduce the opportunity cost of crime and will thus be conceptually equivalent to type I errors.

[15] The conviction rate in Oregon for 2002 is missing because it is omitted in the reports for unknown reasons.

[16] A large portion of the data was missing for these three territories and using other, publically available, data (e.g. from statistical digests) yields inconsistent estimates. For example, the "rape" crime included forcible rape for some years but not others.

available websites. Appendix B describes the court structure and provides data citations.

Crime rates were derived from two sources: (1) the FBI's Uniform Crime Reports ("UCR") and (2) the National Crime Victimization Survey ("NCVS"). UCR rates are measured using index crimes (violent and property crimes), that are traditionally reported to the authorities.[17] NCVS similarly includes violent and property crimes, but is derived from a survey of crime victims (for a review of the differences between UCR and NCVS, see Fay and Diallo, 2015). As NCVS rates are generally unavailable at the state-by-year level, the analysis mostly focuses on UCR rates. However, I utilize NCVS rates for some robustness checks.

4.2. Variables

Rates of appeal outcomes are measured as shares of overall appeals decided on the merits, multiplied by 100 (i.e. measured in percentage points). Independent variables are then (1) the *Affirmance rate* and (2) the *Reversal-Remand spread*. The dependent variable ($ln\_cp100_{i,t}$) is the natural logarithm of total crime (violent and property) per capita in the state of district court i, at year t, also multiplied by 100.[18] This log-linear model thus measures the effect of marginally increasing the relevant rate by 1 percentage points.

In order to avoid a potential omitted variable bias, it is important to control for factor which are correlated with both appeal outcomes and crime rates. While many variables identified in the literature as determining crime presumably have no correlation with appeals (e.g. unemployment rates), the events which occur throughout the stages of the legal process - prosecution, trial, conviction, sentencing and appeal filing - may be related to the outcomes of appeals. Namely, agents in these stages may respond to changes in both sides of the chain - either crime rates or appeal results. For example, prosecutors who care about their success rate (e.g. due to career concerns; see Perry, 1998) or explicit guidelines which limit discretion[19] may respond to a higher reversal rate (=higher rate of losing) by focusing only on "winner cases", eventually leading to a higher conviction rate. Similarly, police officers may avoid passing-on cases to the prosecutors in order to avoid "wasting time" on arrestees who will anyway avoid prosecution later. This may cause the police to refer more guilty people to prosecutors, again leading to more convictions. In an opposite chain reaction, judicial errors may decrease deterrence, causing crime to rise and police officers to have less free resources

[17] Violent crime includes, for example, murder, aggravated assault and forcible rape. Property crime includes, for example, burglary, larceny and motor vehicle theft.

[18] This is the same linear transformation as in the independent variables and is made for convenience only.

[19] The United States Attorneys Manual (USAM, 1997) dictates that federal prosecution should be initiated only if the evidence is sufficient to sustain a conviction. See also Yang (2016, pp 293).

(Epps, 2015, pp. 1097-1098).[20] Investigations may then become less thorough, resulting in more innocent people arrested and, eventually, in less convictions, or more appeals. In addition, the chances of success on appeal naturally change the incentive to file appeals ex-ante, so that defendants' actions may be impacted.

To reduce concerns of any potential bias due to reactions of agents in the legal process, I add two key controls: *conviction rate* and *appeal-filing rate*.[21] These two variables are the typical outcomes of behavior adjustments by police officers, prosecutors, and defendants - and controlling for these should overcome concerns on this issue.[22] For the sake of caution, I also include a long line of additional controls:

- *Arrest rate per capita*, which captures the (unconditional) probability of apprehension.
- *Prosecution rate (share of prosecuted arrestees)*, which captures both prosecutorial behavior and "type II-like" errors which result from failures to prosecute guilty defendants.
- *Share of UCR index crimes in sentences,* as defendants accused of UCR crimes can also be prosecuted in state-courts, I control for the share of sentences in the federal courts relating to UCR crimes.
- *Shares of jury and bench trials*, to control for divergences due to the decision maker's identity.[23]
- *Average sentence length, rates of Imprisonment and Fine-only verdicts and Prison population in state prisons*:[24] these variables capture the type and gravity of sanctions, which may be correlated with type I errors (e.g. due to "compromise verdicts", where sanction size is a function of the guilt-probability). Imprisonment rates also account for possible crime-reducing effects of incapacitation.[25]
- *Shares of "conviction-only" and "sentence+conviction" appeals,* controlling for appeal types.
- *Shares of minority defendants (Black, Hispanic and female)*, as minorities may be more likely to be wrongfully convicted (Olney and Bonn, 2014).
- *Rate of dismissals and ("other") appeal outcomes*, which are correlated with the Affirmance rate and Reversal-Remand spread, but may also relate to unobservables.
- *Supreme Court ("SC") appeal rates,* which are important for two reasons. First, criminals may care

[20] A different bias can arise if the assignment of judges to courts depends on crime (e.g. better judges, who err less, self-select into "safer" areas with lower crime rates).

[21] The appeal filing rate is calculated as $\frac{Appeals filed}{Convicted Defendants}$. Note that rule 4(b)) of the Federal Rules of Appellate Procedure dictates that appeals must generally be filed within 14 days of conviction.

[22] The literature usually controls for detection/arrests and seldom on convictions (Weatherburn, 2012).

[23] Differences in conviction rates may arise either due to different accuracies of judges and juries or to self-selection of defendants into jury and bench trials. For example, innocents may opt for bench trials whereas the guilty choose jury trials, hoping for a "noisy jury" to wrongfully acquit them (Gay et al., 1989). Using a similar logic, type I errors may be less common in jury trials.

[24] Prison population in federal courts is a state-invariant variable, and thus captured by the year fixed effects.

[25] Previous literature struggles with isolating the effects of incapacitation and deterrence (see Chalfin and McCrary, 2017; Listokin, 2003; Owens, 2009; Kessler and Levitt, 1999), usually when trying to identify sentence severity. Namely, as longer sentences reduce crimes committed by recidivists through incapacitation, it is hard to distinguish this from deterrence. However, the problem is less severe when looking directly at imprisonment rates, rather than using only prison population.

about the prospect of overturning convictions at the SC. Second, appellate judges may opt for certain outcomes to avoid SC reversal. Given the scarcity of SC criminal appeal decisions per court and year, I use instead rates for all cases originating from each circuit court in a given year (criminal and other). Presumably, this captures the SC's specific temper towards each circuit court.[26]

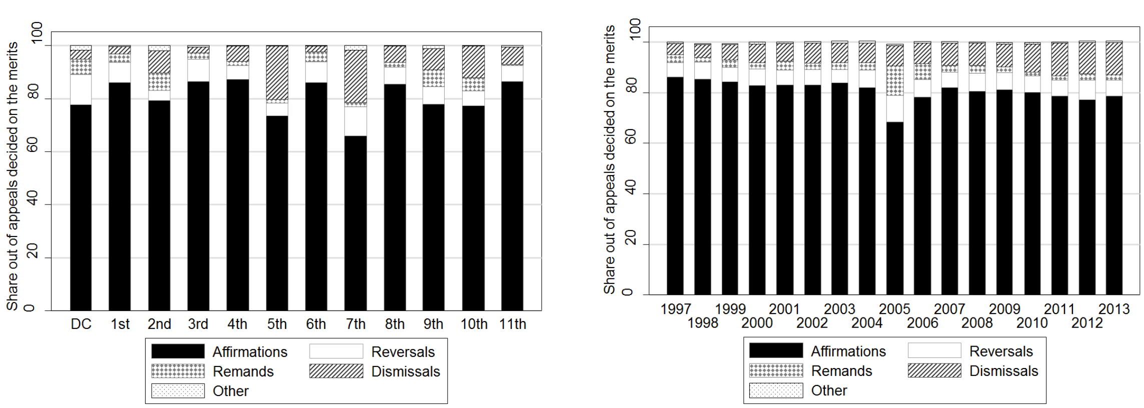
4.3. Summary Statistics - Federal courts data

Table 1 summarizes the main variables of interest (summaries of controls and observations are provided in Tables C2 and C3 in Appendix C). Affirmance rates are fairly high (mean of 80.9%); reversals and remands span a wide range (0-1% to 24-31%), mainly due to inter-court variation. Table C1 in Appendix C presents correlations between appeal outcomes, and shows that remands are negatively correlated with both affirmances and reversals - which is in line with the theoretical model.[27] Figure 3 shows that the distribution of appeal outcomes by court and year. Part (1) shows that outcome frequencies differ greatly by court, which justifies the use of court-fixed effects. Part (2) similarly shows some differences (which justifies year fixed-effects), where one conspicuous jump occurs in 2005 (with lower affirmance rates). This jump likely results from the "Booker Case" (*United States v. Booker*, 543 U.S. 220 (2005); See also *Apprendi v. New Jersey*, 530 U.S. 466 (2000) and *Blakely v. Washington*, 542 U.S. 296 (2004)), which allowed judges to deviate from the strict federal sentencing guidelines and reopen cases on appeal (Catterson, 2006; Heise, 2009).

Figure 3: APPEAL RESULTS - MEANS

(1) Appeal results by district

(2) Appeal results by year



[26]As the SC barely remands without also reversing/affirming, remands are categorized as reversal, affirmance or partial affirmance.

[27]As guilty defendants presumably constitute a higher share of the defendants pool, the model also predicts that remands are mainly an alternative to affirmances. This is supported by the relatively higher negative correlation (-0.4) between remands and affirmances, compared to the correlation between remands and reversals (-0.18).

Table 1 DESCRIPTIVE STATISTICS

Panel A: crime rates	UCR				NCVS			
	mean	sd	min	max	mean	sd	min	max
Total crime per capita	3849.23	1044.03	1581.71	9839.13				
Property crime per capita	3401.38	904.80	1351.31	7814.93				
Violent crime per capita	447.84	202.42	66.91	2024.20				
NCVS rate total					61.74	15.75	26.88	127.47
NCVS rate property					52.10	13.47	23.67	105.17
NCVS rate violent					9.64	2.80	3.22	22.30
Panel B: Appeal rates	mean	sd	min	max				
Criminal Affirmance rate	80.93	8.91	54.75	93.68				
Criminal Reversal rate	6.61	3.25	0.68	24.07				
Criminal Remand rate	3.17	4.44	0.00	31.04				
Criminal Dismissal rate	8.72	7.36	0.00	36.49				
Criminal Remainder rate	0.57	1.21	0.00	14.29				
Criminal Reversal-Remand spread	3.43	5.97	-24.83	20.37				
Observations	1547				1260			
Panel C: Main controls	mean	sd	min	max				
Conviction rate	89.50	6.64	46.55	99.01				
Appeal-filing rate	19.35	8.48	1.95	51.41				
Observations	1424							

NOTE.–This table presents descriptive statistics for the main variables of interest used in the analysis. The figures refers to the year range 1997-2013. Appeal results are at the appeal court level. UCR rates are measured by crime per 100,000 persons, whereas NCVS rates are per 1,000 households.

4.4. *Empirical analysis: regression model and methodology*

The hypotheses are tested using the following regression model:

$$ln\_cp100_{i,t} = \beta_0 + \beta_1 spread_{i,t} + \beta_2 aff_{i,t} + \beta_3 conrate_{i,t} + \beta_4(aff_{i,t} \times conrate_{i,t}) + \beta_5' X' + \varepsilon_{i,t}$$

where: $ln\_cp100_{i,t}$ is the logged crime rate; $spread_{i,t}$ is the reversal-remand spread;[28] $aff_{i,t}$ is the affirmance rate; $conrate_{i,t}$ is the conviction rate; X' is a vector of (varying) control variables, including fixed effects for year and court; and $\varepsilon_{i,t}$ is the error term.

Estimating the model requires a careful choice of methodology (see Cornwell and Trumbull, 1994; Scott, 2006). I thus use statistical tests to rule out inappropriate methods. These tests indicate that (1) fixed effects are preferable to random effects, (2) heteroskedasticity is present and (3) the residuals are auto-correlated.[29]

[28] $spread_{i,t} = rev_{i,t} - rem_{i,t}$, where rev is the reversal rate and rem the remand rate.

[29] A Hausman-test (Hausman, 1978), rejects the null hypothesis of random effects, $p.value < 0.0001$). The null hypothesis of homoskedasticity is rejected in a modified Wald-test ($p.value < 0.0001$), implemented by xttest3 command in Stata (Baum et al., 2001). The null hypothesis of no-autocorrelation was rejected in a Wooldridge test for autocorrelation in panel data ($p.value < 0.0001$), implemented by the xtserial command in Stata (Drukker et al., 2003).

Additionally, there is indication of cross-sectional dependence ("contemporaneous correlation").[30] Cross-sectional dependence may stem from judicial behavior, e.g. if judges compete with other judges in neighboring districts (or districts that serve as professional benchmarks) by altering their rulings, such that rates become correlated. Alternatively, regional shocks to crime incentives in neighboring districts may cause rates to become correlated. I address these issues by (two-way) clustering of the standard errors by court and year, which allows for arbitrary correlations between courts/years and does not impose strict assumptions (Cameron and Miller, 2015). The robustness tests below include some alternative methods as well.

5. Basic Results

Table 2 presents regression results. All columns include the main variables of interest: affirmance rate, conviction rates, their interaction, and the Reversal-Remand spread. Column (1) excludes controls and S.E are not clustered. Column (2) adds court fixed effects and S.E are clustered by court. Column (3) include court and year fixed effects and S.E are, accordingly, two-way clustered. Column (4) adds all other controls, as specified above (S.E are again two-way clustered).[31] The coefficients of interest are all significant once both court and year fixed effects are included (columns 3-4). The basic results thus reflect two key findings:

Finding 1 *The Reversal-Remand spread is negatively associated with crime rates*

Finding 2 *Affirmance rates are negatively associated with crime rates when conviction rates are low.*

Finding 1 can be directly seen from the table, where the coefficient of the spread is negatively significant. Taking, for example, the specification with full controls (column 4), the coefficient can be interpreted as follows: an increase of 1 percentage point in the spread is associated with a decrease of crime rates per capita of 0.19%. Finding 2 requires some additional comparisons. The significant coefficient of the interaction term implies that the effect of affirmance rates on crime depends on conviction rates. However, the coefficient of the interaction term is positive, but the coefficient of the main term - the affirmance rate - is negative.

[30] Since the number of panels (N) is larger than the number of periods (T), one cannot use the standard Breusch-Pagan test (xttest2 command in Stata). Instead, I test the residuals using the user-developed command xtcd2 (Eberhardt, 2011). A Pesaran test (Pesaran, 2015) then indicates cross-sectional dependence ($p.value < 0.1$). To ensure that the cross-sectional dependence is not a technical by-product of using the state level crime rates for all courts within that state, I also test each independent variable separately and find evidence of significant cross-sectional correlation in each series of independent variables ($p.value < 0.01$), indicating that the correlation is not driven by the dependent variable alone.

[31] Recall that these controls are: arrest rate, prosecution rate, share of UCR crimes in sentences, share of jury trials, share of bench trials, Average sentence length, rate of imprisonment, rate of fine-only sentences, prison population in state prisons, share of conviction-only appeals, share of conviction+sentence appeals, shares of black, hispanic and female defendants, rate of dismissal, rate of "other", and the following SC rates: reversal, (full) affirmance and partial affirmance.

Table 2
BASIC RESULTS

	(1)	(2)	(3)	(4)
Criminal Reversal-Remand spread	-0.09	-0.67***	-0.16***	-0.19***
	(0.12)	(0.10)	(0.04)	(0.06)
Criminal Affirmance rate	-1.37	-0.28	-1.31***	-1.01**
	(1.23)	(0.75)	(0.28)	(0.37)
Conviction rate	-1.95*	-1.49**	-1.18***	-0.91**
	(1.12)	(0.66)	(0.25)	(0.32)
Criminal Affirmance rate × Conviction rate	0.01	0.01	0.01***	0.01**
	(0.01)	(0.01)	(0.00)	(0.00)
R-squared	0.03	0.82	0.95	0.95
Court fixed effects	No	Yes	Yes	Yes
Year fixed effects	No	No	Yes	Yes
Controls	No	No	No	Yes
S.E clustered by Court	No	Yes	Yes	Yes
S.E clustered by Year	No	No	Yes	Yes
Number of observations	1546	1546	1546	1424

NOTE.– This table presents regression results of the logged crime rate $ln_c p100$ on the affirmance rate, conviction rate, their interaction, the Reversal-Remand spread, and varying controls. Fixed effects and controls are included in those columns where it is specified so at the lower part of the table. S.E are in parantheses. Controls (in column 4) are: arrest rate, prosecution rate, share of UCR crimes in sentences, share of jury trials, share of bench trials, Average sentence length, rate of imprisonment, rate of fine-only sentences, prison population in state prisons, share of conviction-only appeals, share of conviction+sentence appeals, shares of black, hispanic and female defendants, rate of dismissal, rate of "other", and the following SC rates: reversal, (full) affirmance and partial affirmance. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

To interpret this, note that the coefficient of the main term is the effect when the conviction rate is equal to *zero*.[32] Albeit an effect in a world with a zero conviction rate bears little practical meaning, it provides an indication that affirmance rates may have a significant effect when conviction rates are sufficiently low.

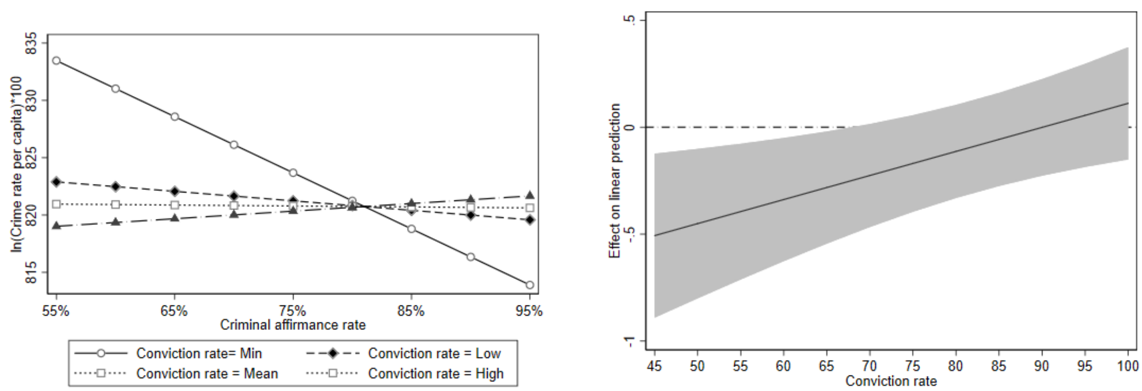
Exploring this possibility, Figure 2 plots the dependencies between the effects of affirmance and conviction rates (coefficients are taken from column 4, with full controls). Part (1) of the figure presents the relationship between affirmance rates (x-axis) and predicted crime rate (y-axis), by varying conviction rates to four representative values: minimum, mean, low (=mean minus one s.d.) and high (=mean plus one s.d.).

[32] If the variables were centered, the interpretation would be different. To see this, compare the coefficients in column 4 of Table 2 with those in column 1 of Table 3, which includes an identical specification except that the the affirmance rate and conviction rates are centered. The main term in that table is insignificant, because centering changes the interpretation: the main term is instead the effect when the conviction rate is at its *mean* (see Afshartous and Preston, 2011; Balli and Sørensen, 2013), which is no longer a low value such as zero.

This illustrates the negative slope of affirmance rates when conviction rates are low and the changes in the slope as conviction rates increase. For these particular values, the effect of affirmance rates is only significant when conviction rates are set to their minimum (hence, the positive slope of the line at the bottom is in fact insignificant). Part (2) of the figure presents how the average marginal effect (hereinafter: "AME") of affirmance rate changes with conviction rates, in order to emphasize the existence of a cut-off point.

Figure 4: MARGINAL EFFECTS OF AFFIRMANCE RATES ON CRIME RATES

(1) Linear prediction (2) Average marginal effects



The Y-axis of part (2) shows the coefficient of affirmance rates, as a function conviction rates on the X-axis. The gray area represents 95% confidence intervals. This graph then shows that the coefficient of the affirmance rate is negative and significant when conviction rates are low and becomes smaller in size as conviction rates rise. The effect remains significant up to the cut-off point of approximately 70%. Hence, the basic results suggest that affirmance rates are negatively associated with crime rates when conviction rates are sufficiently low, but that the effect size decreases in conviction rates (Finding 2)). For an interpretation of the coefficients of control variables, see Appendix E.

Summarizing the basic results, I find clear evidence in favor of H3 (a negative effect of the Reversal-Remand spread) and H2 (smaller effect size of affirmance rates when conviction rates are high). I also find some evidence in favor of H1 (a negative effect of affirmance), but the effect seems to be limited to the case where conviction rates are low. This is consistent with the aforementioned information effect: when conviction rates are too high, the relationship between affirmances are crime weakens, as affirmances are perceived as more arbitrary.

6. Robustness checks

6.1. Data constraints

The basic results are subject to several limitations, given some inherent data constraints. First, UCR rates are published for a full calendar year (January - December) whereas court rates are published for October to September. Hence, the independent variables are already (somewhat) lagged. However, this time lag is also advantageous, assuming that potential criminals take time to update their expectations. Indeed, it is unclear how potential criminals obtain their information. Some (more sophisticated) criminals may hire statisticians to collect data; some may ask representatives to ”sit in” on appeal hearings; others may only get information from newspapers. Informed and naive defendant may also have different perceived probabilities of detection (Buechel, Feess, and Muelheusser, 2018), also due to past-exprience (Sah, 1991), and may rely on heuristics (see, e.g. Johnson and Payne, 1986). Moreover, under the EBM *any person* is in fact a potential criminal, where the deciding factor is only whether, at a given time, the benefit from crime exceeds the expected sanction. Once changes occur in either of these (also due to judicial errors), some non-criminals may become criminals (and vice versa), making it hard to establish an encompassing rule on how types of people receive their information. However, a significant effect of appeals on crime indicates that potential criminals do respond to appeal outcomes. This response need not be due to a direct exposure to court statistics, but can also arise due to network effects (Sparrow, 1991), endogenous decisions to acquire information (Dalla Pellegrina, 2008), or diffusion of deterrence (Bergin, 2011). For example, individuals may interpret the change in appeal outcomes as a general policy change which exceeds the specific case, and then respond *as-if* the relevant change applies to them (even if they are considering to commit a different offense type and even if they expect to be indicted in other courts).

Second, crime rates are measured at the state (rather than district) level. This constraint is, however, irrelevant for the majority of the sample, as most states contain only one district court.[33] In the other states, crime might differ from region to region, but individuals must still rely on the same appeal rates. Thus, deterrence effects of appeal outcomes may be similar for all sub-regions.

Third, the category of “affirm” includes also appeals where the conviction is reversed in part.[34] I overcome

[33]The following 28 states have only one district court: Alaska, Arizona, Colorado, Connecticut, District of Columbia, Delaware, Hawaii, Idaho, Kansas, Massachusetts, Maryland, Maine, Montana, North Dakota, Nevada, New Hampshire, New Jersey, New Mexico, Oregon, Puerto Rico, Rhode Island, South Carolina, South Dakota, Utah, Vermont and Wyoming.

[34]Starting from 2007, this is explicitly stated in the federal courts’ reports.

this problem using alternative data (see section 6.5 below), which separates full from partial affirmances.

A fourth constraint is that the data used for the basic results pools “conviction only” appeals with sentencing appeals. I overcome this by analyzing the said additional data on sentencing appeals.

A fifth constraint concerns granularity. While the federal courts do publish data on the *amount* of appeals filed and terminated per district court, appeal results are only published in aggregate numbers, such that one can only extract the overall rate of an appellate court but not individual rates for each district court.[35] This aggregation is, however, of little concern as long as no trial court within a given circuit errs systematically more than others. A stable disparity in trial court accuracy also seems unlikely, as criminal cases are generally decided by a jury, i.e. by a random group of people.[36] Furthermore, judges can counter biased juries by issuing a “directed verdict” or deciding whether to admit certain evidence. Thus, even if juries decide non-randomly, judges can intervene. Indeed, one may then argue that some judges systematically err, but there are checks and balances which should prevent such distortions as well.[37] Thus, if no court errs systematically on average, no bias will occur. Nonetheless, I address this constraint below by using the said additional data, which does not suffer from granularity issues.

A final constraint is the need to separate type I and II errors. In the U.S., only defendants are allowed to appeal their convictions whereas the prosecution cannot (as a rule) appeal acquittals due to the ”double jeopardy” rule.[38] Therefore, only type I errors are reviewed on appeal. The restriction to type I errors is advantageous for identification, as appeals clearly correspond to convictions and not acquittals. Type II errors are partially captured through the conviction rate, but may constitute a relevant omitted variable - which I overcome using an IV approach.

The following sections present the robustness checks used to overcome these various issues. Control variables in all the checks below are the same as those used in the basic results, unless stated otherwise.

[35] Suppose, for example, that two appeals are filed from two courts. A reversal rate of 50% may occur either (1) when both appeals from one court are reversed and no appeals from the other court are reversed; or (2) when one appeal from each court is reversed.

[36] Jury pools may, however, under-represent some groups from the population (Rose, Casarez, and Gutierrez, 2018).

[37] For example, trial judges may learn from appeal results and adjust their behavior to avoid future mistakes. Problematic judges may be assigned ”easier” cases, or in extreme circumstances - impeached and replaced by better judges (see Pfander, 2007).

[38] The double jeopardy rule forbids the state from placing a citizen twice under the risk of conviction, including via a prosecutorial appeal upon acquittal (see 18 U.S. Code 3731). The rule does not preclude the government from appealing sentences, but such appeals are infrequent (see Commission et al., 2012, part B pp. 38).

6.2. *Alternative Regression Methods*

As a first robustness check, I contrast the basic results with alternative regression models, all of which use fixed effects but differ in the standard errors ("S.E."). Appendix D compares these methods and provides examples for previous papers on either crime rates or courts which use the method.[39] In order to make the coefficient of the affirmance rate more interpretable, both affirmance and conviction rates were centered prior to interacting them. Table 3 compares regressions results of each method (see Table E1 in Appendix E for an extended table). Column (1) reiterates the basic results, by using two-way clustered S.E. Columns (2) and (3) use "Panel corrected standard errors" ("PCSE", Beck and Katz, 1995), assuming either an AR1 or PSAR1 disturbance.[40] This method is robust in the presence of contemporaneous correlation, which is problematic if some - but not all - panels are correlated (Worrall and Pratt, 2004). Column (4) uses Driscol-Kraay standard errors ("SCC", Driscoll and Kraay, 1998; Hoechle, 2007), which similarly accounts for cross-sectional correlation. This first check supports the basic results: the spread is negatively and significantly correlated with crime rates, and affirmance rates are negatively correlated with crime rates when conviction rates are low, as can be seen in Figure 5. Marginal effects for the two-way clustered S.E (column 1) are identical to the base results and thus omitted from the figure. Note that the x-axis now refers to the centered (demeaned) conviction rate, where the mean was 89.5% (e.g. "-10" means 79.5%).

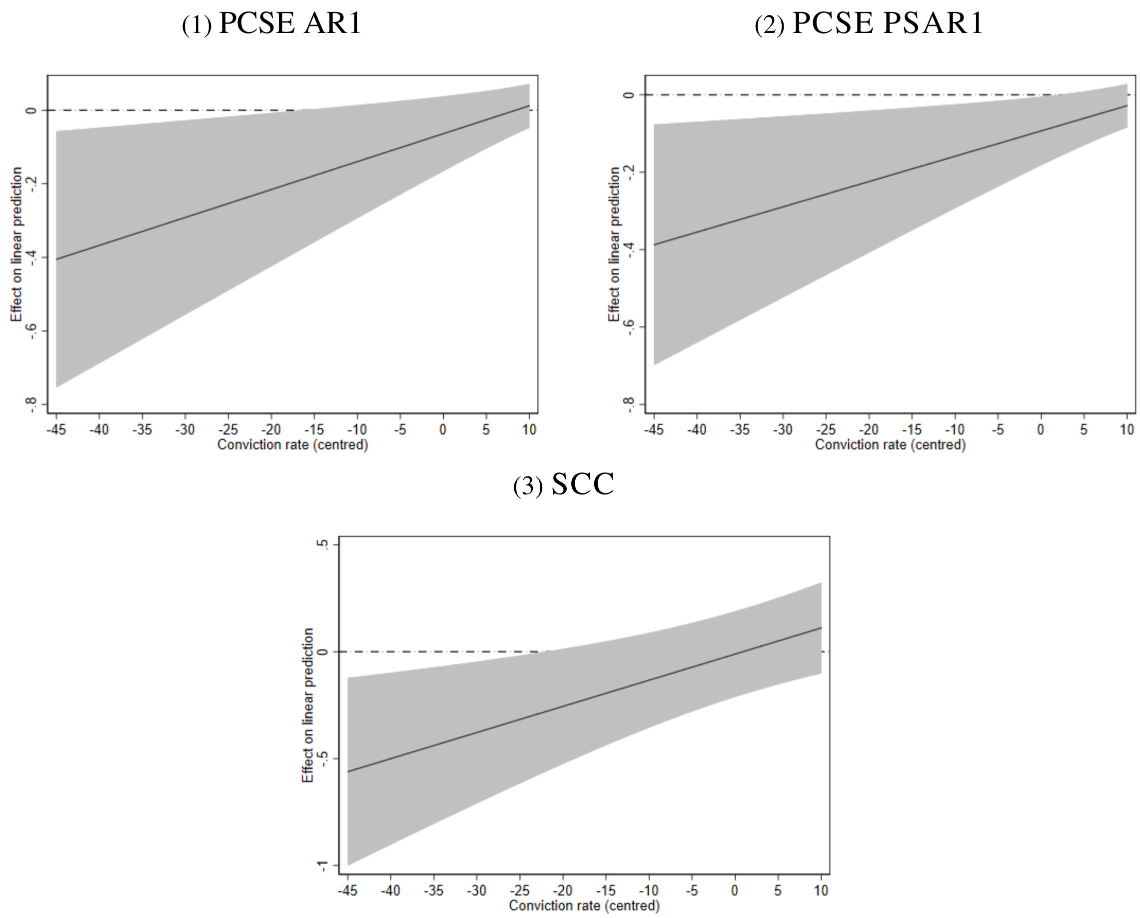
Table 3
BASE RESULTS - COMPARISON OF METHODS

	(1) Two-way cluster	(2) PCSE (AR1)	(3) PCSE (PSAR1)	(4) SCC
Criminal Reversal-Remand spread	-0.192***	-0.085***	-0.060**	-0.192***
	(0.059)	(0.026)	(0.028)	(0.041)
Criminal Affirmance rate (centered)	-0.011	-0.064	-0.093**	-0.011
	(0.112)	(0.052)	(0.045)	(0.102)
Conviction rate (centered)	-0.004	0.036	0.023	-0.004
	(0.070)	(0.043)	(0.045)	(0.065)
Criminal Affirmance rate (centered) × Conviction rate (centered)	0.012**	0.008***	0.007**	0.012***
	(0.005)	(0.003)	(0.003)	(0.004)
Court and Year Fixed Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Number of observations	1424	1424	1424	1424

NOTE.– This table compares results of the basic model using 5 different methods: (1) Two-way clustered S.E. (by court and year); (2) Panel-corrected S.E. (AR1 disturbance) ; (3) Panel-corrected S.E. (panel-specific AR1 disturbance); (4) Driskol-Kraay standard errors, also known as 'SCC'; (5) Poisson with roust S.E. Constant and FE coefficients are not reported. Note that Poisson regression uses a different dependent variable. $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

[39]Estimating the total number of crimes with a fixed-effects Poisson regression (see Osgood, 2000) yields similar results.
[40]PCSE regressions with fixed effects do not produce an interpretable R-squared (see Blackwell III, 2005, pp. 250-251).

Figure 5: AME OF AFFIRMANCE RATES (95% CI)



6.3. Controlling for polynomial effects and spurious significance

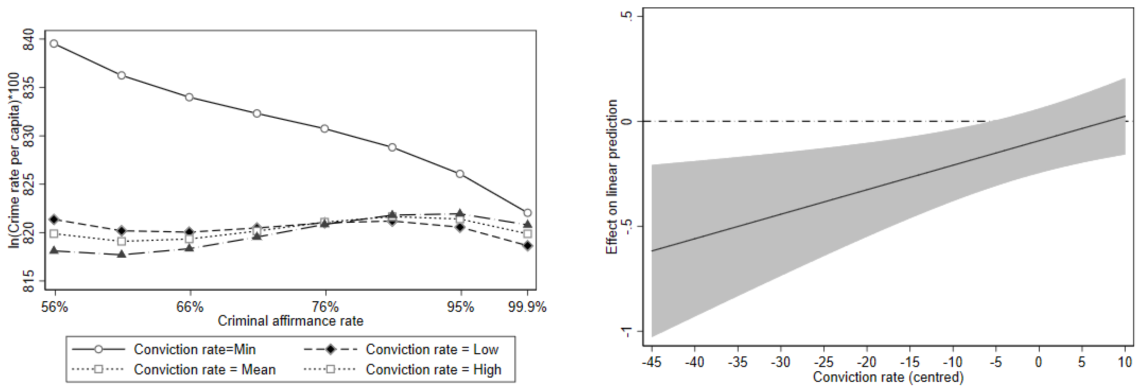
Next, I test the interaction term by adding quadratic and cubic terms of the interacted variables.[41] Results are reported in Table C4 in Appendix C; marginal effects of affirmance rates are presented in Figure 6. The results are again qualitatively similar to the basic results. When conviction rates are low, the effect of affirmance rates is almost linear, i.e. polynomials do not change the results much.

[41]The intuition behind this test is the fear of a 'spurious significance' of the interaction term, due to correlation between the interacted variables (Afshartous and Preston, 2011; Balli and Sørensen, 2013). The need to control for cubic terms is conveniently described by Wooldridge (2012, pp. 303) in the context of conviction rates: *"The presence of the quadratics makes interpreting the model somewhat difficult. (...) We might conclude that there is little or no deterrent at all at lower values of [conviction rates]; the effect only kicks in at higher prior conviction rates. We would have to use more sophisticated functional forms than the quadratic to verify this conclusion."*.

Figure 6: AME OF AFFIRMANCE RATES (WITH POLYNOMIALS)

(1) Linear prediction

(2) AVERAGE MARGINAL EFFECTS



6.4. *Comparing UCR and NCVS, property and violent crimes*

Next, I alternate between (1) UCR/NCVS and (2) Property/Violent crime/Both. The validity of this check relies on an assumption that appeal rates are representative for both crime types, which I cannot test due to the aggregated nature of the data. Nonetheless, as previous papers find different elasticities for violent and property crime (Lee and McCrary, 2017; Chalfin and McCrary, 2017), I explore this estimation as well. To attain state-level estimates of NCVS, I interpolate a 3-year average of the 'Small Area Estimates' (Fay and Diallo, 2015) into a representative yearly average for the sub-sample (1999-2013).[42] Results are reported in Table 4; AME of affirmance rates are presented in Figure 7.[43]

The coefficient of the spread is negatively significant in all but the UCR violent crime specification. The effect of affirmance rates is qualitatively similar to the base results, with the exception of NCVS violent crime, in which the affirmance rate has a positive AME (significant at the 10% level) when conviction rates are high. The overall effect on NCVS remains, however, similar to the base results. Given the differences in how UCR and NCVS are measured (see Fay and Diallo, 2015, pp. 9-10) and the aforementioned aggregation issues, these findings should be taken with caution. Yet, for the purpose of evaluating the strength of the base results, it seems that the effects are not merely driven by the choice of using UCR.

[42] The SAE contain 3-year averages. Thus, each year is repeated in the sample at least *twice*. For example, the year 2006 appears in the range 2005-2007 and 2006-2008. Thus, I add up each adjacent 3-year average and divide by 6, to extract an estimate for each year. The dependent variables for the NCVS are derived from "Appendix table E1" in Fay and Diallo (2015), which lists "Rates of victimization by 1000 Persons or Households".

[43] Note that 90% rather than 95% CI are used, in order to demonstrate that some coefficients are indeed significant at the 10% level. Coefficients for UCR violent and property crimes are presented separately on the right side of the figure, to increase readability).

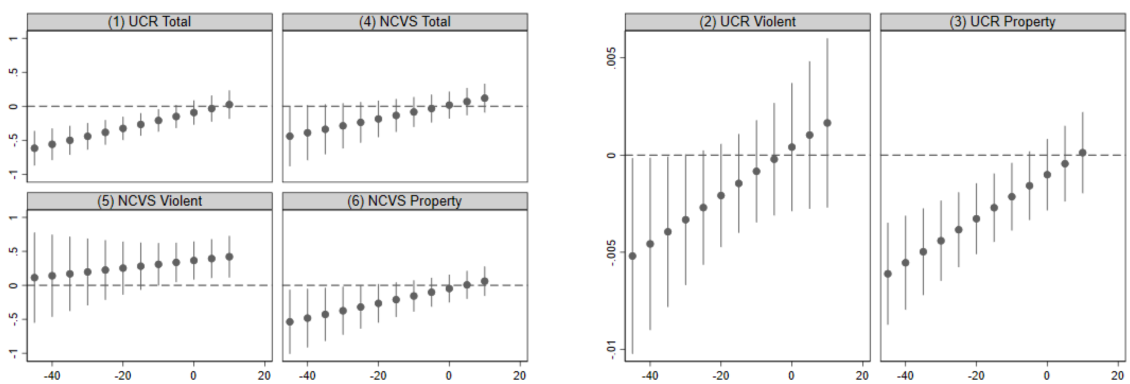
Table 4

ROBUSTNESS CHECK: COMPARISON OF VIOLENT AND PROPERTY CRIME

	UCR			NCVS		
	(1) Total	(2) Violent	(3) Property	(4) Total	(5) Violent	(6) Property
Criminal Reversal-Remand spread	-0.229***	-0.002	-0.002***	-0.205**	-0.229*	-0.197**
	(0.062)	(0.001)	(0.001)	(0.083)	(0.116)	(0.083)
Criminal Affirmance rate (centred)	0.048	0.002	0.000	-0.004	0.189	-0.037
	(0.100)	(0.003)	(0.001)	(0.124)	(0.179)	(0.126)
Criminal Affirmance rate (centred) squared	-0.016**	-0.000*	-0.000**	-0.008	0.018	-0.013
	(0.006)	(0.000)	(0.000)	(0.009)	(0.012)	(0.009)
Criminal Affirmance rate (centred) cubic	-0.001**	-0.000	-0.000*	0.000	0.001	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Criminal Affirmance rate (centered) $\times$ Conviction rate (centered)	0.012**	0.000	0.000**	0.010*	0.006	0.011*
	(0.005)	(0.000)	(0.000)	(0.005)	(0.008)	(0.006)
Conviction rate (centered)	0.055	0.002	0.000	-0.132	-0.073	-0.149
	(0.080)	(0.002)	(0.001)	(0.106)	(0.153)	(0.110)
Conviction rate (centred) squared	-0.003	-0.000	-0.000	-0.013	0.029	-0.022*
	(0.008)	(0.000)	(0.000)	(0.012)	(0.018)	(0.012)
Conviction rate (centred) cubic	-0.000	-0.000	-0.000	-0.000	0.001	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Court and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.945	0.932	0.942	0.916	0.874	0.911
Number of observations	1424	1424	1424	1245	1245	1245

NOTE.– This table compares regression results for violent, property and total crime as measured by the Uniform crime reports (columns 1-3) and the National Victimization Survey (columns 4-6). S.E. are two-way clustered by court and year. Control variables are identical to those in Table 3, but are centered. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 7: VARYING CRIME MEASUREMENTS: AME OF AFFIRMANCE RATES; 90% CI



6.5. *Disentangling appeal results using sentencing appeals*

I proceed by addressing three additional concerns:

1. Do the base results replicate when using a district-level measurement?
2. Is the effect of affirmances driven by government-filed appeals?
3. Is the effect of affirmances driven by "partial affirmances"?

To do so, I use an alternative measure for appeal results, derived from the USSC databases. These rates do not suffer from granularity issues, as they are measured at the district-court level and can be separated into defendant-filed, government-filed and cross-appeals, with affirmances separated into "full" and "partial". At the same time, the additional data suffers from its own share of limitations. Namely, only *sentencing appeals* are included, i.e. where the sentence was appealed (with or without appealing the conviction). The database thus neglects outcomes of conviction-only appeals (the share of conviction-only appeals is still included, as a control variable). The restriction to sentencing appeals however bears advantages, such as the exclusion of interlocutory appeals, which are less relevant for estimating final expected sanctions.

A second caveat arises since it is unclear whether sentencing appeal rates were public information in real time. Some version of the database seems to have been available since 2006, in the form of an online data-tool (covering 2006-2013), but previous years were excluded. The 2006-2013 range happens to overlap with the post-Booker-case range, i.e. when sentences could be reopened and appeals possibly became salient.

Table C5 in Appendix C presents summary statistics of the USSC dataset. Sentencing appeal rates somewhat differ from the appellate court mean, especially for defendant and government appeals. For example, recall that the affirmance rate mean was 80.9% in the large sample. The defendant-appeals affirmance rate is a bit lower (76.2%), but the government-appeals affirmance rate is much lower (29.08%). Furthermore, reversal rates are higher than in the large sample: 8.7% for defendant-filed and 64.5% for government-filed appeals. A higher reversal rate for government-filed appeals is consistent with findings of Hettinger and Lindquist (2012, pp. 140), who suggest that a selection effect - where the government hand-picks cases expected to end in reversal - may be driving this difference (see also King and Heise, 2018). The combined sentencing affirmance rate (full and partial) is, however, not statistically different from the average in the larger sample, supporting the assumption that appeal-level data is in fact representative of each district court.

Table 5 compares regression results for the full sample and the sub-range 2005-2013.[44] In this check, rates are pooled for both appellant types (defendant/government). The outcome of this robustness test supports the basic results. The (sentencing) Reversal-Remand spread has a significantly negative coefficient in all specifications. Both the full and the partial affirmance rate is negatively correlated with crime, which suggests that partial affirmances do not confound the effect. The interaction term is (positively) significant only when using PCSE and in the sub-sample. Hence, most specifications show no interaction effect, i.e.

[44] Independent variables are centered to reduce collinearity with the interaction terms. For example, interactions are correlated with the main terms: 0.92 for full affirmances and 0.99 for partial affirmances (each with their interaction with conviction rates). The year 2005 is included, as the Booker case was already decided during that year.

affirmance rates are just always negatively associated with crime (which supports H1).

Table 5
SENTENCING DATABASE REGRESSIONS (FULL SAMPLE V. POST-2005)

Sample range:	Full sample		2005-2013	
Method	(1) Two-way cluster	(2) PCSE-AR1	(3) Two-way cluster	(4) PCSE-AR1
Reversal-remand spread (sentence) (centered)	-0.073**	-0.028***	-0.114***	-0.089***
	(0.027)	(0.008)	(0.027)	(0.025)
Full Affirmance rate (centred)	-0.060	-0.041***	-0.109**	-0.088***
	(0.039)	(0.009)	(0.036)	(0.021)
Full Affirmance rate (centred) squared	-0.001	-0.000	0.003	0.002*
	(0.002)	(0.000)	(0.002)	(0.001)
Full Affirmance rate (centred) cubic	-0.000	-0.000	0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Partial Affirmance rate (centred)	-0.014	-0.046**	-0.074	-0.070
	(0.074)	(0.019)	(0.062)	(0.050)
Partial Affirmance rate (centred) squared	-0.017	-0.008*	-0.015	-0.007
	(0.016)	(0.005)	(0.017)	(0.013)
Partial Affirmance rate (centred) cubic	0.000	0.000*	0.001	0.000
	(0.001)	(0.000)	(0.001)	(0.001)
Conviction rate (centered)	0.060	0.049	-0.052	-0.015
	(0.084)	(0.032)	(0.091)	(0.065)
Full affirmance rate (senetence) (centered) $\times$ Conviction rate (centered)	0.004	0.002	0.005	0.006***
	(0.003)	(0.002)	(0.004)	(0.002)
Partial affirmance rate (sentence) (centered) $\times$ Conviction rate (centered)	0.001	0.002	0.005	0.008
	(0.008)	(0.005)	(0.010)	(0.005)
Conviction rate (centred) squared	-0.004	-0.000	0.005	0.005
	(0.011)	(0.004)	(0.008)	(0.006)
Conviction rate (centred) cubic	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Court and year fixed effects	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes
Number of observations	1424	1424	801	801

NOTE.– This table compares regression results of two-way clustered and panel-corrected S.E. (AR1) for the full range (1997-2013) and post-Booker range (2005-2013). Control variables are identical to those used in Table 3. Constant and FE coefficients are not reported. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

The next check separates the rates by the appellant type: defendants, government and cross-appeals.[45] Results are reported in Table 6. The spread is significantly negative for both defendant-filed and government-filed appeals. The coefficient of cross-appeals spread is negative as well, but statistically insignificant.

Figure 8 plots AME of the different affirmance rates (numbers in parentheses correspond to column numbers in Table 6). Part (1) shows that for defendant-filed appeals, the effect of full affirmances is visually

[45] As government-appeals and cross-appeals are fairly rare, missing values are replaced by 1-year lagged values, assuming that criminals rely on earlier rates as a proxy when no government/cross appeals were filed in a specific year.

Table 6
ROBUSTNESS CHECK: DEFENDANT V. GOVERNMENT APPEALS

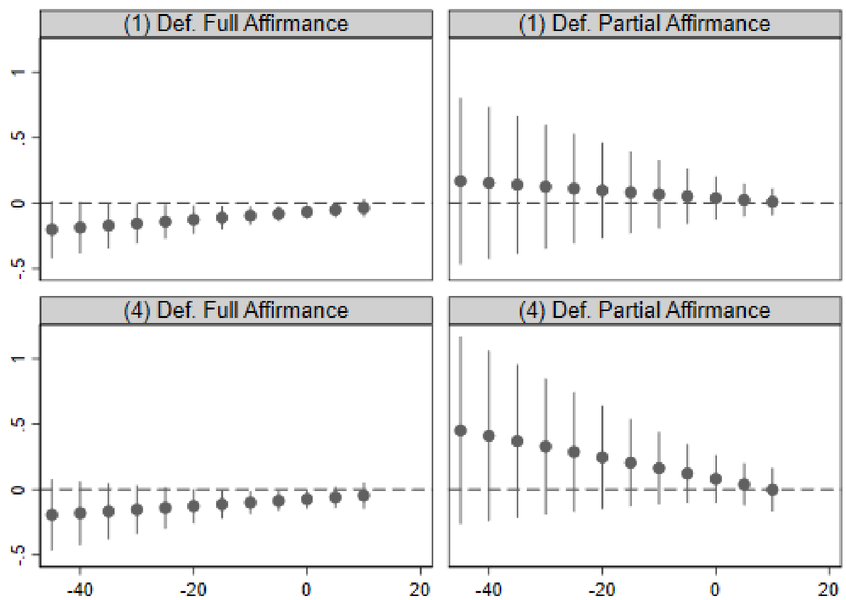
	(1) Defendant appeals	(2) Government appeals	(3) Cross appeals	(4) All appeals
Coefficients of Defendant rates				
Defendant Reversal-Remand spread	-0.072** (0.030)			-0.068** (0.027)
Defendant Partial Affirmance rate	0.000 (0.067)			0.041 (0.083)
Defendant Full Affirmance rate	-0.058 (0.037)			-0.062 (0.039)
Defendant Full Affirmance × Conviction rate	0.003 (0.003)			0.003 (0.005)
Defendant Partial Affirmance × Conviction rate	-0.003 (0.009)			-0.008 (0.010)
Defendant Full Affirmance squared	-0.001 (0.002)			0.000 (0.002)
Defendant Full Affirmance cubic	-0.000 (0.000)			-0.000 (0.000)
Defendant Partial Affirmance squared	-0.024 (0.019)			-0.030 (0.023)
Defendant Partial Affirmance cubic	0.000 (0.001)			0.001 (0.001)
Coefficients of Government rates				
Government Reversal-Remand spread		-0.067* (0.032)		-0.069* (0.037)
Government Partial Affirmance rate		-0.175 (0.151)		-0.166 (0.168)
Government Full Affirmance rate		-0.052 (0.039)		-0.070 (0.045)
Government Full Affirmance × Conviction rate		-0.001 (0.001)		-0.001 (0.002)
Government Partial Affirmance × Conviction rate		-0.002 (0.002)		-0.002 (0.002)
Government Full Affirmance squared		0.001 (0.001)		-0.000 (0.002)
Government Full Affirmance cubic		-0.000 (0.000)		0.000 (0.000)
Government Partial Affirmance squared		0.005 (0.006)		0.003 (0.006)
Government Partial Affirmance cubic		-0.000 (0.000)		-0.000 (0.000)
Coefficients of Cross-appeals rates				
Cross-Appeals Reversal-Remand spread			-0.015 (0.010)	-0.010 (0.010)
Cross-Appeals Full Affirmance rate			-0.093 (0.061)	-0.049 (0.063)
Cross-Appeals Partial Affirmance rate			-0.007 (0.046)	0.013 (0.047)
Cross-Appeals Full Affirmance × Conviction rate			0.000 (0.002)	0.001 (0.002)
Cross-Appeals Partial Affirmance × Conviction rate			0.002 (0.001)	0.002 (0.001)
Cross-Appeals Full Affirmance squared			-0.003** (0.001)	-0.002 (0.001)
Cross-Appeals Full Affirmance cubic			0.000* (0.000)	0.000 (0.000)
Cross-Appeals Partial Affirmance squared			0.001 (0.001)	0.001 (0.001)
Cross-Appeals Partial Affirmance cubic			-0.000 (0.000)	-0.000 (0.000)
Conviction rate (centered)	0.059 (0.085)	-0.024 (0.062)	0.018 (0.091)	-0.087 (0.075)
Conviction rate (centred) squared	-0.004 (0.010)	-0.010 (0.009)	-0.008 (0.011)	
Conviction rate (centred) cubic	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	
R-squared	0.943	0.944	0.954	0.957
Court and Year fixed effects	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes
Number of observations	1424	1337	1198	1148

NOTE.– This table shows results for two-way clustered (by court and year) regression, using separate variables for defendant-filed, government-filed and cross appeals. Column (1) includes only defendant appeals. Column (2) includes only government appeals. Column (3) includes only cross-appeals. Column (4) includes all rates together. Appeals results are at the district court level. Control variables are the same as the previous table (7). All regressions include court and year fixed effects. Constant and fixed effect coefficients are not reported. Dependent variable: ln_cp100. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

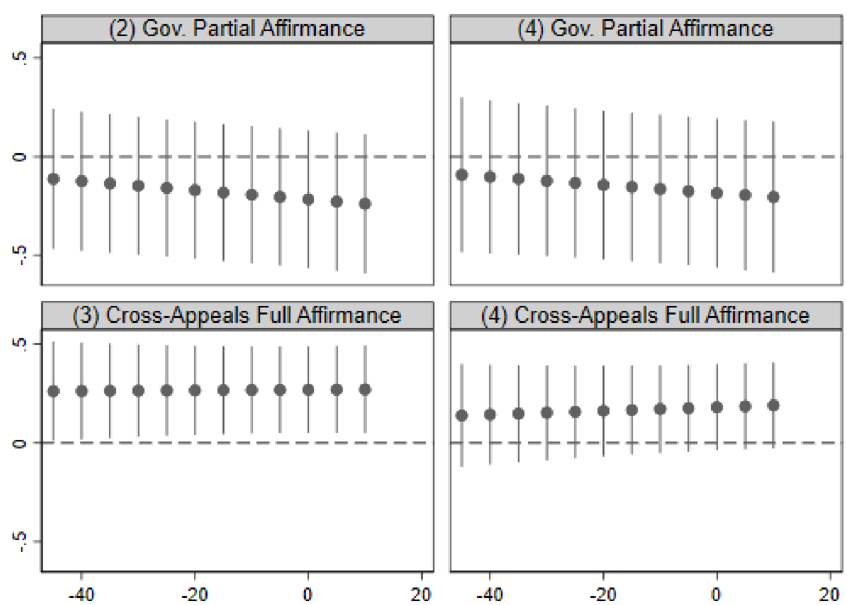
similar to the base results. However, the moderation effect of conviction rates is in fact insignificant (i.e. full affirmances have a negative effect which does not significantly decrease in size with conviction rates). Conversely, the effect of (defendant) partial-affirmance is insignificant, i.e. these outliers do not impact the results.

Figure 8: AME OF AFFIRMANCE RATES; 95% CI:

(1) Defendant appeals



(2) Government and Cross-Appeals



Part (2) shows AME for government (partial) affirmance, and cross-appeals (full) affirmance (For brevity, I omit government full affirmance and cross-appeals partial affirmances, which are insignificant). These coefficients are marginally insignificant once all rates are included, but the figure seems to signify that cross-appeals full affirmances may in fact have a positive effect on crime. As such an effect goes against the basic results, it clearly demonstrates that my results are not driven by these outliers as well.

As government appeals are about sentences - i.e. about f rather than p - an increase in affirmance rates (i.e. the government loses more appeals) implies a lower f. This may explain the positive AME of cross-appeals affirmances, where the government is also an appellant. Alternatively, as the defendant is also an appellant in cross-appeals, an affirmance may imply the opposite, i.e. an increase in f. Some argue that higher punishments may decrease deterrence (see discussion in Appendix E), but a straightforward interpretation can also be derived from the EBM. Namely, if the government is prone to (cross-)appeal sentences that accompany wrongful convictions, then the affirmance of the conviction implies an increase in $E(f|I)$, which should decrease deterrence. A tendency to appeal the sentences of wrongfully convicted defendants may emerge, e.g. if such sentences are generally lower (e.g. due to "compromise verdicts" of judges and juries; see Lundberg, 2016; Sarel, 2018). Alternatively, these effects may be driven by the trust channel: when courts infrequently overturn punishments , the system may be perceived as unfairly harsh.

6.6. An Instrumental variables approach

An additional check addresses potential endogeneity, using an instrumental variables ("IV") approach.[46] The concern of endogeneity due to omitted variable bias or reversed causality (judges responding to crime rates) seems stronger for the Reversal-Remand spread and less so for affirmance rates – for two reasons. First, reversals and remands are possibly correlated with the type of error – legal or factual – which is unobservable in the data. Then, if reversals capture legal errors while remands reflect factual errors, their opposite effect on deterrence is congruent with the theoretical argument that errors concerning the legal standard increase deterrence while errors concerning the act decrease deterrence (Mungan and Lando, 2015). Second, changes in crime rates might cause appellate judges to alternate between reversals and remands, making the choice of tool correlated with the regression's error term. Conversely, affirmance rates seem

[46] In an IV approach, an additional variable ("an instrument") is used, which is: (1) "exogenous," i.e. uncorrelated with the regressions' error term, and (2) "relevant", i.e. sufficiently correlated with the endogenous variable. The instrument is first used to predict the endogenous independent variable *as if* it was unrelated to the error term (in a "first stage" regression). Predicted values then substitute original values in the main regression ("second stage" regression). For previous papers which use IV for other stages of the legal chain, see e.g. Cornwell and Trumbull (1994); Levitt (1996, 1997, 2002).

less likely to be affected by crime rates, as appellate judges are constrained in their ability to manipulate affirmances by declaring that an error took place where it did not (or vice versa), e.g. due to the potential for Supreme Court review. The focus of this robustness test is therefore a possible endogeneity of the spread.

My IV approach exploits variations in judicial ideology, as captured by Judicial Common Space scores (”JCS”) (Epstein et al., 2007) which build on the identity of appointing political actors to measure ideological preferences on a continuous scale between ”conservative” (1) and ”liberal” (-1).[47] These scores allow to construct time-varying “ideological distances” between court-levels (see e.g. Sarel and Demirtas, 2017), which arguably fulfill the necessary conditions for an instrument: exogeneity and relevance. Namely, variation in ideological distances is *exogenous* to crime rates, as (1) the assignment of appellate judges to cases is generally random and (2) variations stem from judges leaving/joining the court, which depends mainly on promotions, pensions and death (See Yang (2016); See also Di Tella and Schargrodsky (2013) for another application of ideologies as an instrument for decisions in criminal cases).

Ideological distances are also *relevant*, as they are known predictors of the tendency of appellate judges to reverse and remand (see references in the introduction). I thus use two distances as instruments:

1. $Dist_{SC-AC}$: the distance between the median JCS scores of the Supreme Court and the appellate court, measured in absolute value.

2. $Dist_{AC-DC}$: the distance between the median JCS scores of the appellate court and the district court, measured in absolute value.

In addition to the ideological distances, I also consider (2-year and 3-year) lagged values of the Reversal-Remand spread. Using lagged values is a standard choice, which is effective if the lags themselves do not directly impact the outcome but are sufficiently correlated with the instrumented variable (see Reed, 2015).

The IV approach is implemented in two steps, estimating the following two equations:

$$\widehat{Spread} = \gamma_0 + \gamma_1 Dist_{SC-AC} + \gamma_2 Dist_{AC-DC} + \gamma_3' Z' + \epsilon_2 \tag{3}$$

$$ln\_cp100 = \beta_0 + \beta_1 \widehat{Spread} + \beta_2' X' + \epsilon_3. \tag{4}$$

where Z' is a vector which may include the lagged spread variables, $spread_{i,t-2}$ and $spread_{i,t-3}$; and X'

[47]Scores range between -1 (fully liberarl) and 1 (fully conservative). Supreme Court JCS score lies on the same scale, but are a transformation of Martin-Quinn scores (Martin and Quinn, 2002), which are based on judicial votes.

is a vector of additional variables, which may include the affirmance and conviction rates, their interaction and polynomials, and other controls.

Results are presented in Table 7. Column (1) excludes controls and other variables and uses only the ideological distances and 2-year lagged spread as instruments. Columns (2) adds a 3-year lagged spread as an instrument. Column (4) adds the affirmance and conviction rates, interactions and polynomials. Column (5) adds controls. All regressions include court fixed-effects, but time-fixed effects are excluded, due to their correlation with the instruments. The coefficient of the spread is negatively significant in all specifications, supporting the basic results. Standard econometric tests are used to verify the validity of instruments.[48] These tests indicate that instruments are exogenous and sufficiently relevant.[49]

The effect of affirmance rates slightly differs from the base results, where the coefficient of affirmance-squared (column (4)) is significant, but the AME is insignificant. One explanation for this difference is that the instruments mainly capture judicial preferences to reverse vs. remand, irrespective of the case merits. This then allows to measure the local average treatment effect (LATE) of "compliers" (see Wooldridge, 2002, Chapter 18), i.e. those courts that respond to the instruments. It is possible that error identification - unlike error correction - does not often comply to changes in ideology. Then, the prediction relates to a counter-factual in which an affirmance implies that no error occurred, yet no-affirmance implies that the outcome is arbitrarily decided according to preferences - yielding a different coefficient for affirmances. In other words, while judicial ideology is arguably exogenous precisely because it predicts appellate judges' behavior irrespective of the case merits, it may be useful to instrument for the affirmance rate as well.

[48]The following tests are used:

- 'Overidentifying restrictions test' (Hansen, 1982), to ensure that instruments are jointly valid. Note that the null hypothesis is that of validity, thus a failure to reject the null indicates that instruments are indeed valid.
- 'Endogeneity test', to check whether there is a need to treat the spread as endogenous. The endogeneity test is implemented using the Stata command xtivreg2, endog() (Schaffer, 2012). And an "orthogonality test" (which is simply the flip-side of the endogeneity test, see Baum, Schaffer, and Stillman (2007, pp. 15)), to verify that all other control variables that are suspicious of endogeneity - including affirmance - can be treated as exogenous.
- Underidentification LM test (Kleibergen and Paap, 2006), where a rejection of the null hypothesis means that the model is not under-identified.
- Kleibergen-Paap F-statistic,(Kleibergen and Paap, 2006) which is the heteroskedasticity robust version of the Cragg-Donald statistic (Cragg and Donald, 1993). Instruments are not weak if this statistic exceeds a threshold (Stock and Yogo, 2005) of a relative bias compared to OLS. A 10% threshold means that IV estimator bias is smaller by at least 90 percentage points than OLS bias.
- Tests for significance of endogenous regressors, as implemented by Stata's "Weakiv" command (Anderson and Rubin (1949) test, K-test and K-J test). A rejection of the null hypothesis means that inference can be made even if instruments are only weakly relevant.

[49]The Hansen's J test has $p.value > 0.1$, suggesting that the null hypothesis of exogeneity cannot be rejected. F-tests are above the Stock-Yogo threshold, i.e. instruments are not weak (the highly significant Anderson-Rubin, CLR, K and K-J tests, support the same conclusion).

Table 7

INSTRUMENTAL VARIABLES: IDEOLOGY AS AN INSTRUMENT

	(1)	(2)	(3)	(4)
Criminal Reversal-Remand spread	-0.420* (0.238)	-0.590*** (0.191)	-0.817*** (0.140)	-0.588** (0.256)
Criminal Affirmance rate (centered)			-0.167 (0.142)	-0.030 (0.185)
Criminal affirmance rate (centered) squared			-0.006 (0.009)	-0.022*** (0.009)
Criminal affirmance rate (centered) cubic			0.000 (0.000)	-0.001* (0.000)
Criminal affirmance * Conviction (both rates centered)			-0.003 (0.009)	0.003 (0.007)
Conviction rate (centered)			-0.502*** (0.154)	-0.064 (0.122)
Conviction rate (centered) squared			-0.030* (0.018)	-0.009 (0.014)
Conviction rate (centered) cubic			-0.001 (0.000)	-0.000 (0.000)
Criminal Dismissal rate			-0.928*** (0.144)	-0.495*** (0.145)
Criminal Remainder rate			-1.308** (0.571)	-1.816*** (0.377)
Hansen J (p.value)	0.207	0.291	0.300	0.801
Endogeneity test (p.value)	0.95	0.14	< 0.001	0.49
Kleibergen-Paap LM (underidentification)	< 0.001	< 0.001	< 0.001	< 0.001
F-statistic (Kelibergen-Paap)	18.901	15.510	13.540	11.022
Stock-Yogo F-statistic cutoff (10% bias)	9.08	10.27	10.27	10.27
Anderson-Rubin (p.value)	.154	.069	.01	.264
CLR test	.154	.039	.002	.054
K test	.162	.042	.002	.044
K-J test	.196	.052	.002	.054
Ideological distances as instruments	Yes	Yes	Yes	Yes
2-year lagged spread as instrument	Yes	Yes	Yes	Yes
3-year lagged spread as instrument	No	Yes	Yes	Yes
Controls	No	No	No	Yes
Clustered S.E.	Yes	Yes	Yes	Yes
Observations	1365	1274	1273	1245

NOTE.– This table shows the IV regression results. Appeal rates are taken from the appellate-court level, as in the basic results. Controls in column (4) are the same as those used in the basic results. S.E. are clustered at the court level. A two-step GMM estimator is implemented in the regressions. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Thus, as a final robustness test, I run additional IV regressions using the sentencing appeal rates (which avoid granularity problems), while simultaneously instrumenting for the spread and affirmance rates - but using different instruments. The different instrument set consists of appeal result rates *in non-criminal cases*, relying on those general factors that plausibly carry over from criminal cases. For example, remands used to restrain trial judges (see Boyd, 2015)) seem equally likely in all proceeding types. I include three reversal-remand spreads and three affirmance rates, of: (1) administrative (2) private-civil and (3) other-civil cases.[50] In light of the affirmance$\times$conviction interaction, I interact also the instruments with conviction rates (for the reasoning behind interacting instruments with the exogenous variable of the interaction, see Bun and Harrison (2018); Nizalova and Murtazashvili (2016)).

Administrative appeals are presumably exogenous, as they are unlikely to be affected by crime, especially by the types of offenses in the sample (violent and property crime). However, administrative appeals typi-

[50] Adding the ideological distances in addition does not qualitatively change these results.

cally originate from administrative agencies rather than district courts (see e.g. rule 13 of the Federal Rules of Appellate Procedure, governing appeals from tax courts.) and thus may miss factors relating to trial courts. Civil appeals, conversely, originate from district courts, providing higher predictive power while maintaining a distance from crime related issues. Unfortunately, it cannot be ruled out that some civil litigation that involves damages caused by crimes may render civil appeals imperfectly exogenous. However, the main concern for all non-criminal spreads is the possibility that these are correlated with some unobserved factor, such as type II errors. For example, administrative appeals are a manifestation of a conflict between a private entity and a public authority, where an incorrect ruling in favor of the administrative petitioner can bear some resemblance to a wrongful acquittal. Thus, to ensure validity, I enlarge the instrument set using the Lewbel (2012) approach, which generates valid instruments (from heteroskedastic errors) and is useful when external instruments are unavailable (Mishra and Smyth, 2015, pp. 167). The generated instruments also serve as distinct variation sources for affirmance rates.[51]

Results are presented in Table 8.[52] The tests indicate that the instruments are relevant[53] and when all instruments are included (column 2) - also jointly exogenous.[54] The IV estimations indicate once more that crime rates decrease in both the Reversal-Remand spread and affirmance rates - full and partial. The interaction term coefficient is positively significant for partial affirmances but negatively significant for full affirmances. However, the high p.value of the endogeneity test indicate that there is no significant efficiency gain from these IV estimations compared to the previous results.

Summing up, the various robustness tests largely support the basic results, especially for defendant-filed appeals - which are the theoretically relevant kind for testing type I errors. The evidence suggests that the effect of the Reversal-Remand spread and the affirmance rate is negative, with some mixed evidence regarding an affirmance$\times$conviction interaction effect.

[51]One limitation of this approach is its reliance on the assumption that heteroskedastic covariates are exogenous (see also Baum et al., 2012). Some control variables do seem likely to satisfy the exogeneity assumption. For example, previous studies find that higher crime rates do not necessarily lead to higher conviction rates (Jensen and Heller, 2003, pp. 83), indicating that convictions do not suffer from reversed causality. Subsequently, interaction terms with conviction rates are plausibly exogenous as well, as the endogenous variable is controlled for. (see Bun and Harrison, 2018; Nizalova and Murtazashvili, 2016).

[52]Polynomial terms are excluded from the regression, given the difficulty of combining the Lewbel approach with such terms. In principal, one could construct a three-step procedure, where (1) the affirmance rate is predicted using the instruments (2) a polynomial of the predicted value ($\widehat{a}ff$) is constructed (e.g. $(\widehat{a}ff))^2$) and (3) these constructed polynomials are used as instrument for the endogenous polynomials (e.g. $(\widehat{a}ff)^2$ serves as an instrument for aff^2) - as suggested by Wooldridge (2002, section 9.5, pp. 236-237). However, it is unclear whether this is applicable when using generated instruments.

[53]As the regressions include multiple endogenous variables, I also report on F-tests for individual regressors (Angrist and Pischke (2008) F-test, as modified by Sanderson, Windmeijer et al. (2016)). These F-tests can be compared to Stock-Yogo critical values.

[54]The Hansen J test in column 1 is significant at the 10% level, indicates that generated instruments alone do not clearly satisfy the overidentification restrictions. However, in column 2 the test is insignificant, indicating exogenity.

Table 8
IV: HETEROSKEDASTICITY-BASED & EXTERNAL INSTRUMENTS

	(1) Generated instruments only	(2) All instruments
Defendant Reversal-Remand spread	-0.134*** (0.017)	-0.091*** (0.010)
Defendant Full Affirmance	-0.011 (0.014)	-0.038*** (0.007)
Defendant Partial Affirmance	-0.242*** (0.042)	-0.168*** (0.023)
Defendant Full Affirmance × Conviction	-0.001 (0.002)	-0.003** (0.001)
Defendant Partial Affirmance × Conviction	0.007* (0.003)	0.007*** (0.002)
Hansen J (p.value)	0.086	0.309
Endogeneity test (p.value)	0.7943	0.4987
F-statistic (Kelibergen-Paap)	26.041	62.266
Stock-Yogo F-statistic cutoff (10% bias)	10.78	10.76
Control variables	Yes	Yes
Court fixed effects	Yes	Yes
Year fixed effects	No	No
Number of observations	1424	1424

Weak-IV F-test (Sanderson-Windmeijer)		
Defendant Reversal-Remand spread	Defendant Full Affirmance rate	Defendant Partial Affirmance rate
79.62872	176.515	149.7631

NOTE.– This table shows the additional IV regression results, where three variables are instrumented: Defendant reversal-remand spread, Defendant Full affirmance and Defendant partial affirmance. In column 1, only heteroskedasticity-based generated instruments are included. Column 2 presents results using all instruments: (I) heteroskedasticity-based instruments (II) administrative reversal-remand spread and affirmance rate (III) private civil appeals reversal-remand spread and affirmance rates (IV) other civil appeals reversal-remand spread and affirmance rate; and (V) the interactions of conviction rates with (II), (III) and (IV). Instruments are taken from the appeal court level. Control variables are the same as in the basic results, except that variables are all centered by court to include court fixed effects. Time fixed effect are excluded. S.E. are clustered at the district court level. A two-step GMM estimator is implemented in all regressions. The regressions are run using the user-developed command ivreg2h (Baum and Schaffer, 2014) in Stata. $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

7. Conclusion

The results are mostly in line with the theoretical model and seem to support the (so far debated) EBM. Hence, the findings suggest that judicial errors should, in fact, be corrected in order to increase deterrence. The polar effects of reversals and remand, however, create a twist: the way of correction matters. Under my model, this is attributed to assignment of these outcomes to different defendant types (guilty and innocent), but as mentioned - other explanations, such as time preferences, litigation costs and trust, may also partially drive the effect. From a research perspective, the fact that the coefficients of appeal outcomes are consistently significant implies that future empirical research should combine these variables in the analysis of crime.

From a policy perspective, the impact of appeals on deterrence should be integrated into a cost-benefit analysis, noting that errors should only be avoided or corrected when the benefit exceeds the social cost. As the effect size is relatively small, appellate judges should perhaps feel free to decide as they see fit, knowing that deterrence effects are negligible anyway.

Furthermore, even if the additional cost of crime justifies a policy change, the problem of deterrence and remands should be approached with caution. Remands are usually justified by either (1) "institutional superiority" of the first instance to correct errors cheaper; (2) A narrow jurisdiction of appellate courts to directly correct errors; or (3) avoiding an unnecessary binding precedent (Hessick, 2012). The deterrence effects of remands, which thus far have been neglected in the literature, roughly correspond to three issues: (1) uncorrected errors, (2) delayed sanctions and (3) litigation cost-asymmetry for the guilty and the innocent.

If the problem is uncorrected errors (i.e. remands substituting reversals of wrongful convictions), appellate judges who suspect the defendant may be innocent may wish to (always) reverse, rather than remand. However, it is difficult to draw clear lines as to the optimal threshold for choosing reversals over remands, also because the public may mistakenly misinterpret such reversals as releasing guilty defendants. One possible intervention may be in strengthening the appellate courts' role in giving guidance to lower courts. For example, providing more specific instructions can reduce ambiguity (Hessick, 2012, pp. 5-6) and encourage error correction (Boyd, 2015). If the problem lies in remands back to the same panel, appellate courts should perhaps be restricted from doing so. Hannibal and Worth (2012) note, that the authority to enforce reassignment to another judge (28 U.S.C 2106) is reluctantly used, as to avoid implicit votes of no-confidence in the trial judge. Nonetheless, some courts have developed rules or criteria that dictate if and when cases are reassigned (Scheinfeld and Bagley, 2013). Unfortunately, my results would suggest that these measures are perhaps insufficient to prevent declines in deterrence.

If delayed sanctions are the problem, one may consider either speeding up the legal process (which can be costly) or improving the scrutiny of petitions to stay sentences, as to reduce the frequency of guilty defendants with postponed sanctions. Finally, if a cost-asymmetry explains the effect, additional compensation for wrongful convictions following a remand process may be considered. A cost reimbursement for acquitted defendants can balance the effect of judicial errors (Fon and Schäfer, 2007), but has been argued to be ineffective in the U.S. (Kahn, 2010). Improving reimbursement mechanisms for post-remand acquittals may thus provide relief (see also Mungan and Klick (2016) for a similar argument regarding plea bargains of innocents).

There exists also one extreme measure: abolishing the remand process altogether. The implications of such a step go far beyond the issue at hand, but even when focusing on deterrence alone, the costs of "not remanding" may be too high. Namely, removing the possibility to remand would force judges to alternate between affirmance, reversal and dismissal. Judges will be reluctant to switch remand-worthy cases to affirmance due to the constraints of (not) declaring an error, but some might then turn to reversal as the only available measure to correct errors. This might affect deterrence in two channels: first, the reversal category would become ambiguous. Second, the payoff from crime may increase, as factual mistakes that would usually require remand may also occur when one is guilty. Furthermore, prosecutors may be reluctant to prosecute where questions of evidence are pivotal, given a higher chance of reversed convictions. Conversely, incentivizing prosecutors to focus on strong cases may prevent wrongful prosecutions, thereby reducing wrongful convictions. Alternatively, judges may dismiss appeals that would otherwise be remanded. This may deteriorate faith in the system, as convictions will be less frequently overturned. Prosecutors will then have a higher incentive to prosecute weak cases, possibly leading to more wrongful convictions. One must therefore carefully consider these complex effects of appellate courts on deterrence, particularly given the countervailing influences of reversals and remands.

The results of this paper are subjected to some limitations. Measurement errors may occur due to aggregations of crime rates and appeal results; the sample may be imperfectly random, as only years with available data are included; ideological scores may be imperfectly measured (see Bonica et al., 2016); and dynamic effects are not considered. The use of the Reversal-Remand spread as a main independent variable also bears some disadvantages (e.g. as it imposes a symmetrical effect of reversals and remand). Additionally, focusing on the federal courts rather than state courts (for which, unfortunately, high-quality data is scarce) is naturally expected to yield a relatively small effect (as most defendant are prosecuted in state courts). Nonetheless, the evidence provided in this paper suggest that potential criminals do respond to appeal outcomes in the federal courts, as deterrence theory would indeed predict. I address most concerns throughout the paper, but further research is needed to address remaining issues. Nonetheless, my results are a first step in the empirical analysis of the effects of appellate courts on deterrence.

References

Afshartous, David and Richard A Preston. 2011. "Key results of interaction models with centering." *Journal of Statistics Education* 19 (3):1–24.

Anderberg, Dan, Helmut Rainer, Jonathan Wadsworth, and Tanya Wilson. 2016. “Unemployment and Domestic Violence: Theory and Evidence.” *The Economic Journal* 126 (597):1947–1979.

Anderson, Theodore W and Herman Rubin. 1949. “Estimation of the parameters of a single equation in a complete system of stochastic equations.” *The Annals of Mathematical Statistics* 20:46–63.

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Bachman, Ronet, R. Paternoster, and S. Ward. 1993. “The rationality of sexual offending: Testing a deterrence/rational choice conception of sexual assault.” *Law and Society Review* 26 (2):343–372.

Balli, Hatice Ozer and Bent E Sørensen. 2013. “Interaction effects in econometrics.” *Empirical Economics* 45 (1):583–603.

Bar-Gill, Oren and Oren Gazal Ayal. 2006. “Plea bargains only for the guilty.” *The Journal of Law and Economics* 49 (1):353–364.

Baum, Christopher F, Arthur Lewbel, Mark E Schaffer, Oleksandr Talavera et al. 2012. “Instrumental variables estimation using heteroskedasticity-based instruments.” *UK Stata Users Group Meetings* :13–14.

Baum, Christopher F and Mark E Schaffer. 2014. “IVREG2H: Stata module to perform instrumental variables estimation using heteroskedasticity-based instruments.” *Statistical Software Components* .

Baum, Christopher F, Mark E Schaffer, and Steven Stillman. 2007. “Enhanced routines for instrumental variables/GMM estimation and testing.” *Stata Journal* 7 (4):465–506.

Baum, Christopher F et al. 2001. “Residual diagnostics for cross-section time series regression models.” *Stata Journal* 1 (1):101–104.

Baumann, Florian and Tim Friehe. 2015. “Proof Beyond a Reasonable Doubt: Laboratory Evidence.” Working Paper No. 181 DICE Discussion Paper. Duesseldorf, Germany. URL: `http://www.dice.hhu.de/fileadmin/redaktion/Fakultaeten/Wirtschaftswissenschaftliche_Fakultaet/DICE/Discussion_Paper/181_Baumann_Friehe.pdf`.

Bebchuck, Lucian Arie and Louis Kaplow. 1992. *The Journal of Legal Studies* 21:365–370.

Beck, Nathaniel and Jonathan N Katz. 1995. “What to do (and not to do) with time-series cross-section data.” *American Political Science Review* 89 (03):634–647.

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

Beery, Brendan. 2002. “Be Careful What You Ask for: Navigating a Remand after You've Won the Appeal.” *Michigan Bar Journal* (May, 2002).

Beraldo, Sergio, Raul Caruso, and Gilberto Turati. 2013. “Life is now! Time preferences and crime: Aggregate evidence from the Italian regions.” *The Journal of Socio-Economics* 47:73–81.

Bergin, Tiffany. 2011. “How and why do criminal justice public policies spread throughout US states? A critical review of the diffusion literature.” *Criminal Justice Policy Review* 22 (4):403–421.

Bjerk, David and Eric Helland. 2017. “Using a Ratio Test to Estimate Racial Differences in Wrongful Conviction Rates.” *IZA Discussion Papers, No. 10631. URL:* `https://ssrn.com/abstract=2940621` .

Blackwell III, J Lloyd. 2005. “Estimation and testing of fixed-effect panel-data systems.” *Stata Journal* 5 (2):202–207.

Bonica, Adam, Adam S Chilton, Jacob Goldin, Kyle Rozema, Maya Sen, and John F Kennedy. 2016. “Measuring Judicial Ideology

Using Law Clerk Hiring." *American Law and Economics Review* :ahw013.

Borochoff, Elise. 2008. "Lower court compliance with Supreme Court remands." *Touro Law Review* 24:849–906.

Boyd, Christina L. 2015. "The Hierarchical Influence of Courts of Appeals on District Courts." *The Journal of Legal Studies* 44 (1):113–141.

Buechel, Berno, Eberhard Feess, and Gerd Muelheusser. 2018. "Optimal Law Enforcement with Sophisticated and Naive Offenders." Cesifo working paper 7106. URL:`http://www.cesifo-group.de/DocDL/cesifo1_wp7106.pdf`.

Bun, Maurice J. G. and Teresa D Harrison. 2018. "OLS and IV estimation of regression models including endogenous interaction terms." *Econometric Reviews* URL `https://doi.org/10.1080/07474938.2018.1427486`.

Cameron, A Colin and Douglas L Miller. 2015. "A practitioner's guide to cluster-robust inference." *Journal of Human Resources* 50 (2):317–372.

Catterson, Cathy. 2006. "Changes in Appellate Caseload and Its Processing." *Arizona Law Review* 48:287–299.

Caulkins, Jonathan P. 1993. "Zero-tolerance policies: do they inhibit or stimulate illicit drug consumption?" *Management Science* 39 (4):458–476.

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

Commission, US Sentencing et al. 2012. "Report on the Continuing Impact of United States v. Booker on Federal Sentencing." *Washington, DC: US Sentencing Commission* .

Corman, Hope and Naci Mocan. 2005. "Carrots, sticks, and broken windows." *The Journal of Law and Economics* 48 (1):235–266.

Cornwell, Christopher and William N Trumbull. 1994. "Estimating the economic model of crime with panel data." *Review of Economics and Statistics* 26 (2):360–366.

Correia, Sergio et al. 2015. "REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects." *Statistical Software Components* .

Cotte Poveda, Alexander. 2011. "Socio-economic development and violence: an empirical application for seven metropolitan areas in Colombia." *Peace Economics, Peace Science and Public Policy* 17 (1):1–23.

Cragg, John G and Stephen G Donald. 1993. "Testing identifiability and specification in instrumental variable models." *Econometric Theory* 9 (02):222–240.

Craswell, Richard and John E Calfee. 1986. "Deterrence and uncertain legal standards." *Journal of Law, Economics, and Organization* 2:279–303.

Dalla Pellegrina, Lucia. 2008. "Court delays and crime deterrence." *European Journal of Law and Economics* 26 (3):267–290.

Dari-Mattiacci, Giuseppe and Bruno Deffains. 2007. "Uncertainty of law and the legal process." *Journal of Institutional and Theoretical Economics (JITE)* 163 (4):627–656.

Davis, Michael L. 1988. "Time and punishment: an intertemporal model of crime." *Journal of Political Economy* :383–390.

Di Tella, Rafael and Ernesto Schargrodsky. 2013. "Criminal recidivism after prison and electronic monitoring." *Journal of Political Economy* 121 (1):28–73.

Dickens, William T. 1986. "Crime and punishment again: the economic approach with a psychological twist." *Journal of Public Economics* 30 (1):97–107.

Domenech, Gabriel and Miguel Puchades. 2014. “Better that ten innocent persons suffer than that one guilty scape.” Working Paper. University of Valencia Law School, Valencia, Spain. URL: `https://www.uv.es/gadopas/2014.standards.evidence.detention.pdf`.

Driscoll, John C and Aart C Kraay. 1998. “Consistent covariance matrix estimation with spatially dependent panel data.” *Review of Economics and Statistics* 80 (4):549–560.

Drukker, David M et al. 2003. “Testing for serial correlation in linear panel-data models.” *Stata Journal* 3 (2):168–177.

Dušek, Libor. 2015. “Time to punishment: The effects of a shorter criminal procedure on crime rates.” *International Review of Law and Economics* 43:134–147.

Eberhardt, Markus. 2011. “XTCD: Stata module to investigate Variable/Residual Cross-Section Dependence.” *Statistical Software Components* .

Ehrlich, Isaac. 1982. “The optimum enforcement of laws and the concept of justice: a positive analysis.” *International Review of Law and Economics* 2 (1):3–27.

Engel, Christoph. 2016. “Experimental Criminal Law. A Survey of Contributions from Law, Economics and Criminology.” Working Paper No. 2016/7 MPI Collective Goods Preprint, Bonn, Germany. URL: `https://ssrn.com/abstract=2769771`.

Epps, Daniel. 2015. “Consequences of Error in Criminal Justice, The.” *Harvard Law Review* 128:1065–1151.

Epstein, Lee, William M Landes, and Richard A Posner. 2013. *The behavior of federal judges: a theoretical and empirical study of rational choice*. Cambridge, MA.: Harvard University Press.

Epstein, Lee, Andrew D Martin, Jeffrey A Segal, and Chad Westerland. 2007. “The judicial common space.” *Journal of Law, Economics, and Organization* 23 (2):303–325.

Evans, William N and Emily G Owens. 2007. “COPS and Crime.” *Journal of Public Economics* 91 (1):181–201.

Fay, Robert E and Mamadou Diallo. 2015. “Developmental Estimates Of Subnational Crime Rates Based On The National Crime Victimization Survey.” *Report, NCJ 249238. URL: `http://www.ncjrs.gov/App/publications/abstract.aspx?ID=271379`.*

Feess, Eberhard and Roee Sarel. forthcoming. “Judicial Effort and the Appeal System: Theory and Experiment.” *The Journal of Legal Studies* .

Feess, Eberhard, Hannah Schildberg-Hörisch, Markus Schramm, and Ansgar Wohlschlegel. 2018. “The Impact of Fine Size and Uncertainty on Punishment and Deterrence: Theory and Evidence from the Laboratory.” *Journal of Economic Behavior and Organization* 149:58–73.

Fon, Vincy and Hans-Bernd Schäfer. 2007. “State liability for wrongful conviction: Incentive effects on crime levels.” *Journal of Institutional and Theoretical Economics (JITE)* 163 (2):269–284.

Garoupa, Nuno. 1999. “Optimal law enforcement with dissemination of information.” *European Journal of Law and Economics* 7:183–196.

Garoupa, Nuno and Matteo Rizzolli. 2012. “Wrongful convictions do lower deterrence.” *Journal of Institutional and Theoretical Economics (JITE)* 168 (2):224–231.

Gathmann, Christina. 2008. “Effects of enforcement on illegal markets: Evidence from migrant smuggling along the southwestern border.” *Journal of Public Economics* 92 (10):1926–1941.

Gay, Gerald D, Martin F Grace, Jayant R Kale, and Thomas H Noe. 1989. "Noisy juries and the choice of trial mode in a sequential signalling game: theory and evidence." *RAND Journal of Economics* 20 (2):196–213.

Gennaioli, Nicola and Andrei Shleifer. 2008. "Judicial Fact Discretion." *The Journal of Legal Studies* 37 (1):1–35.

Gneezy, Uri and Aldo Rustichini. 2004. "Incentives, punishment and behavior." *Advances in Behavioral Economics* :572–589.

Gould, Eric D et al. 2002. "Crime rates and local labor market opportunities in the United States: 1979–1997." *Review of Economics and Statistics* 84 (1):45–61.

Gould, Jon B, Julia Carrano, Richard A Leo, and Katie Hail-Jares. 2014. "Predicting Erroneous Convictions." *Iowa Law Review* 99:471–2299.

Gould, Jon B and Richard A Leo. 2010. "One hundred years later: Wrongful convictions after a century of research." *Journal of Criminal Law and Criminology* 100 (3):825–868.

Griffin, Lissa. 2000. "The Correction of Wrongful Convictions: A Comparative Perspective." *American University International Law Review* 16:1241.

Gross, Samuel R and Barbara O'Brien. 2008. "Frequency and predictors of false conviction: Why we know so little, and new data on capital cases." *Journal of Empirical Legal Studies* 5 (4):927–962.

Gross, Samuel R, Barbara OBrien, Chen Hu, and Edward H Kennedy. 2014. "Rate of false conviction of criminal defendants who are sentenced to death." *Proceedings of the National Academy of Sciences* 111 (20):7230–7235.

Gross, Samuel R and Michael Shaffer. 2012. "Exonerations in the United States, 1989–2012." University of Michigan Public Law Working Paper No. 277. URL: `https://ssrn.com/abstract=2092195`.

Hannibal, Kemerer and James A. Worth. 2012. "Judicial reassignment: a proposal." The National Law Journal (Nov. 19, 2012). URL: `https://www.law.com/nationallawjournal/almID/1202578501776/?slreturn=20181104091159`.

Hansen, Lars Peter. 1982. "Large sample properties of generalized method of moments estimators." *Econometrica* 50 (4):1029–1054.

Harris, John R. 1970. "On the economics of law and order." *Journal of Political Economy* 78:165–174.

Hausman, Jerry A. 1978. "Specification tests in econometrics." *Econometrica* 46:1251–1271.

Heise, Michael. 2009. "Federal Criminal Appeals: A Brief Empirical Perspective." *Marqutte Law Review* 93 (2):825–844.

Hessick, F Andrew. 2012. "The Cost of Remands." *Arizona State Law Journal* 44:1025–1036.

Hettinger, Virginia A and Stefanie A Lindquist. 2012. "Decision Making in the US Courts of Appeals." In *New Directions in Judicial Politics*, edited by Kevin T McGuire. Routledge, 126–143.

Hoechle, Daniel. 2007. "Robust standard errors for panel regressions with cross-sectional dependence." *Stata Journal* 7 (3):281–312.

Huck, Steffen and Michael Kosfeld. 2007. "The dynamics of neighbourhood watch and norm enforcement." *The Economic Journal* 117 (516):270–286.

Iaryczower, Matias, Garrett Lewis, and Matthew Shum. 2013. "To elect or to appoint? Bias, information, and responsiveness of bureaucrats and politicians." *Journal of Public Economics* 97:230–244.

Ihlanfeldt, Keith R. 2007. "Neighborhood drug crime and young males' job accessibility." *Review of Economics and Statistics* 89 (1):151–164.

Jensen, Erik Gilbert and Thomas C Heller. 2003. *Beyond common knowledge: empirical approaches to the rule of law*. Stanford, CA.: Stanford University Press.

Johnson, Eric and John Payne. 1986. "The decision to commit a crime: An information-processing analysis." In *The reasoning criminal: Rational choice perspectives on offending*, edited by Cornish B. Derek and Ronald Clarke V., chap. 11. New Brunswick, Canada: Transaction Publishers, 170–185.

Johnson, Joel S. 2016. "Benefits of Error: A Dynamic Defense of the Blackstone Principle in Criminal Law." *Virginia Law Review* 102:237–283.

Kahan, Dan M. 1996. "Between economics and sociology: The new path of deterrence." *Michigan Law Review* 95:2477–2497.

Kahn, Daniel S. 2010. "Presumed guilty until proven innocent: The burden of proof in wrongful conviction claims under state compensation statutes." *University of Michigan JL Reform* 44:123–168.

Kaplow, Louis. 1994. "The value of accuracy in adjudication: An economic analysis." *The Journal of Legal Studies* 23 (1):307–401.

———. 2011a. "On the Optimal Burden of Proof." *Journal of Political Economy* 119 (6):1104–1140.

———. 2011b. "Optimal proof burdens, deterrence, and the chilling of desirable behavior." *American Economic Review* 101 (3):277–80.

———. 2012. "Burden of Proof." *Yale Law Journal* 121:738–859.

———. 2014. "Likelihood Ratio Tests and Legal Decision Rules." *American Law and Economics Review* 16:1–39.

Kaplow, Louis and Steven Shavell. 1994. "Accuracy in the Determination of Liability." *The Journal of Law and Economics* 37 (1):1–15.

Kessler, Daniel and Steven D Levitt. 1999. "Using sentence enhancements to distinguish between deterrence and incapacitation." *The Journal of Law and Economics* 42 (S1):343–364.

Kim, Sangjoon, Jaihyun Park, Kwangbai Park, and Jin-Sup Eom. 2013. "Judge-Jury Agreement in Criminal Cases: The First Three Years of the Korean Jury System." *Journal of Empirical Legal Studies* 10 (1):35–53.

King, Nancy J and Michael Heise. 2018. "Appeals by the Prosecution." *Journal of Empirical Legal Studies* 15 (3):482–538.

Kleibergen, Frank and Richard Paap. 2006. "Generalized reduced rank tests using the singular value decomposition." *Journal of Econometrics* 133 (1):97–126.

Knight, Jack and Lee Epstein. 2017. "Strategic Accounts of Judging." In *Routledge Handbook of Judicial Behavior*. Routledge, 48–61.

Korobkin, Russell B and Thomas S Ulen. 2000. "Law and behavioral science: Removing the rationality assumption from law and economics." *California Law Review* 88 (4):1051–1144.

Kovandzic, Tomislav V and Lynne M Vieraitis. 2006. "The effect of county-level prison population growth on crime rates." *Criminology and Public Policy* 5 (2):213–244.

Kuersten, Ashlyn and Donald Songer. 2014. *Decisions on the US Courts of Appeals*. New York, NY: Routledge.

Lando, Henrik. 2006. "Does wrongful conviction lower deterrence?" *The Journal of Legal Studies* 35 (2):327–337.

Lando, Henrik and Murat C Mungan. 2017. "The Effect of Type-1 Error on Deterrence." *International Review of Law and Economics* 53:1–8. URL `http://dx.doi.org/10.1016/j.irle.2017.08.001`.

Lee, David S and Justin McCrary. 2017. "The deterrence effect of prison: Dynamic theory and evidence." In *Regression Disconti-*

nuity Designs: Theory and Applications. Emerald Publishing Limited, 73–146.

Levitt, Steven D. 1996. "The effect of prison population size on crime rates: Evidence from prison overcrowding litigation." *Quarterly Journal of Economics* 111 (2):319–351.

———. 1997. "Using electoral cycles in police hiring to estimate the effect of police on crime." *American Economic Review* 87 (3):270–290.

———. 2002. "Using electoral cycles in police hiring to estimate the effects of police on crime: Reply." *American Economic Review* 92 (4):1244–1250.

Levy, Gilat. 2005. "Careerist Judges and the Appeals Process." *RAND Journal of Economics* 36 (2):275–297.

Lewbel, Arthur. 2012. "Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models." *Journal of Business & Economic Statistics* 30 (1):67–80.

Lewisch, Peter, Stefania Ottone, and Ferruccio Ponzano. 2015. "Third-Party Punishment under Judicial Review: An Economic Experiment on the Effects of a Two-Tier Punishment System." *Review of Law & Economics* 11 (2):209–230.

Lin, Ming-Jen. 2008. "Does unemployment increase crime? Evidence from US data 1974–2000." *Journal of Human Resources* 43 (2):413–436.

Listokin, Yair. 2003. "Does more crime mean more prisoners? An instrumental variables approach." *The Journal of Law and Economics* 46 (1):181–206.

Livernois, John and Chris J McKenna. 1999. "Truth or consequences: enforcing pollution standards with self-reporting." *Journal of Public Economics* 71 (3):415–440.

Loeffler, Charles E, Jordan Hyatt, and Greg Ridgeway. 2018. "Measuring Self-Reported Wrongful Convictions Among Prisoners." *Journal of Quantitative Criminology* :1–28.

Lundberg, Alexander. 2016. "Sentencing Discretion and Burdens of Proof." *International Review of Law and Economics* :34–42.

Machin, Stephen and Costas Meghir. 2004. "Crime and economic incentives." *Journal of Human Resources* 39 (4):958–979.

Markussen, Thomas, Louis Putterman, and Jean-Robert Tyran. 2016. "Judicial error and cooperation." *European Economic Review* 89:372–388.

Martin, Andrew D and Kevin M Quinn. 2002. "Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis* 10 (2):134–153.

Mastrobuoni, Giovanni and David A Rivers. 2016. "Criminal Discount Factors and Deterrence." IZA Discussion Papers, No. 9769, URL: `http://ftp.iza.org/dp9769.pdf`.

Mishra, Vinod and Russell Smyth. 2015. "Estimating returns to schooling in urban China using conventional and heteroskedasticity-based instruments." *Economic Modelling* 47:166–173.

Mungan, Murat C. 2015a. "Stigma dilution and over-criminalization." *American Law and Economics Review* 18 (1):88–121.

Mungan, Murat C. 2015b. "Wrongful convictions and the punishment of attempts." *International Review of Law and Economics* 42:79–87.

Mungan, Murat C and Jonathan Klick. 2016. "Reducing False Guilty Pleas and Wrongful Convictions through Exoneree Compensation." *The Journal of Law and Economics* 59 (1):173–189.

Mungan, Murat C and Henrik Lando. 2015. "The Effect of Type-1 Error on Deterrence." *FSU College of Law, Public Law Research*

Paper 687 URL http://dx.doi.org/10.2139/ssrn.2441733.

Nicita, Antonio and Matteo Rizzolli. 2014. "In Dubio Pro Reo. Behavioral explanations of pro-defendant bias in procedures." *CESifo Economic Studies* .

Nizalova, Olena Y and Irina Murtazashvili. 2016. "Exogenous treatment and endogenous factors: Vanishing of omitted variable bias on the interaction term." *Journal of Econometric Methods* 5 (1):71–77.

Obidzinski, Marie and Yves Oytana. 2017. "Identity mistakes and the standard of proof." TEPP Working paper No. 2018 08, URL: https://halshs.archives-ouvertes.fr/halshs-01879555/document.

Olney, Maeve and Scott Bonn. 2014. "An Exploratory Study of the Legal and Non-Legal Factors Associated With Exoneration for Wrongful Conviction: The Power of DNA Evidence." *Criminal Justice Policy Review* 26 (4):1–21.

Osgood, D Wayne. 2000. "Poisson-based regression analysis of aggregate crime rates." *Journal of Quantitative Criminology* 16 (1):21–43.

Owens, Emily G. 2009. "More time, less crime? Estimating the incapacitative effect of sentence enhancements." *The Journal of Law and Economics* 52 (3):551–579.

Perry, HW. 1998. "United States Attorneys: Whom Shall They Serve?" *Law and Contemporary Problems* 61 (1):129–148.

Pesaran, M Hashem. 2015. "Testing weak cross-sectional dependence in large panels." *Econometric Reviews* 34 (6-10):1089–1117.

Pfander, James E. 2007. "Removing Federal Judges." *The University of Chicago Law Review* 74 (4):1227–1250.

Png, Ivan PL. 1986. "Optimal subsidies and damages in the presence of judicial error." *International Review of Law and Economics* 6 (1):101–105.

Pogarsky, Greg and Alex R. Piquero. 2003. "Can Punishment Encourage Offending? Investigating the "Resetting" Effect"." *Journal of Research in Crime and Delinquency* 40 (1):95–120.

Polinsky, A Mitchell and Steven Shavell. 2000. "The Economic Theory of Public Enforcement of Law." *Journal of Economic Literature* 38 (1):45–76.

———. 2001. "Corruption and optimal law enforcement." *Journal of Public Economics* 81 (1):1–24.

———. 2007. "The theory of public enforcement of law." *Handbook of Law and Economics* 1:403–454.

Posner, Richard A. 1973. "An economic approach to legal procedure and judicial administration." *The Journal of Legal Studies* 2 (2):399–458.

Raphael, Steven and Rudolf Winter-Ebmer. 2001. "Identifying the effect of unemployment on crime." *The Journal of Law and Economics* 44 (1):259–283.

Reed, William Robert. 2015. "On the practice of lagging variables to avoid simultaneity." *Oxford Bulletin of Economics and Statistics* 77 (6):897–905.

Rizzolli, Matteo and Margherita Saraceno. 2013. "Better that ten guilty persons escape: punishment costs explain the standard of evidence." *Public Choice* 155 (3-4):395–411.

Rizzolli, Matteo and Luca Stanca. 2012. "Judicial errors and crime deterrence: Theory and experimental evidence." *The Journal of Law and Economics* 55 (2):311–338.

Rizzolli, Matteo and James Tremewan. 2018. "Hard labor in the lab: Deterrence, non-monetary sanctions, and severe procedures." *Journal of Behavioral and Experimental Economics* 77:107–121.

Roberts, Julian V and Mike Hough. 2005. "The state of the prisons: exploring public knowledge and opinion." *The Howard Journal of Criminal Justice* 44 (3):286–306.

Rose, Mary R, Raul S Casarez, and Carmen M Gutierrez. 2018. "Jury Pool Underrepresentation in the Modern Era: Evidence from Federal Courts." *Journal of Empirical Legal Studies* 15 (2):378–405.

Sah, Raaj Kumar. 1991. "Social osmosis and patterns of crime: A dynamic economic analysis." *Journal of Political Economy* 99 (6):1272–1295.

Sanderson, Eleanor, Frank Windmeijer et al. 2016. "A Weak Instrument F-Test in Linear IV Models with Multiple Endogenous Variables." *Journal of Econometrics* 190 (2):212–221.

Sarel, Roee. 2018. "Do Wrongfully Convicted Defendants Receive Higher Punishments?" Working Paper. Frankfurt School of Finance & Management, Frankfurt am Main, Germany. URL: https://ssrn.com/abstract=3205036.

Sarel, Roee and Melanie Demirtas. 2017. "Delegation at the US Federal Appellate Courts: The Power to Remand As a Double-Edged Sword." Working Paper. Frankfurt School of Finance & Management, URL: https://ssrn.com/abstract=3094634.

Schaffer, Mark E. 2012. "xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models." *Statistical Software Components* .

Scheinfeld, Robert C and Parker H. Bagley. 2013. "Reassignment to a New Judge After Remand From the Federal Circuit." *The New York Law Journal* 249 (98).

Scott, Kevin M. 2006. "Understanding judicial hierarchy: Reversals and the behavior of intermediate appellate judges." *Law & Society Review* 40 (1):163–192.

Shavell, Steven. 1995. "The Appeals Process as a Means of Error Correction." *The Journal of Legal Studies* 24 (2):379–426.

———. 2006. "The Appeals Process and Adjudicator Incentives." *The Journal of Legal Studies* 35 (1):1–29.

Silverman, Dan. 2004. "Street crime and street culture." *International Economic Review* 45 (3):761–786.

Soares, Yuri and Maria Micaela Sviatschi. 2010. "Does Court Efficiency Have a Deterrent Effect on Crime? Evidence for Costa Rica." Unpublished Manuscript. Inter-American Development Bank and Columbia University. URL: http://www.inesad.edu.bo/bcde2012/papers/7.%20Sviatschi_Crime%20and%20Efficiency.pdf.

Spaeth, Harold J, Lee Epstein et al. 2016. "2016 Supreme Court Database, Version 2016 Release 1, URL: http://Supremecourtdatabase.org."

Sparrow, Malcolm K. 1991. "The application of network analysis to criminal intelligence: An assessment of the prospects." *Social Networks* 13 (3):251–274.

Spencer, Bruce D. 2007. "Estimating the accuracy of jury verdicts." *Journal of Empirical Legal Studies* 4 (2):305–329.

Stock, James H and Motohiro Yogo. 2005. "Testing for weak instruments in linear IV regression." In *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, chap. 5. Cambridge Univ Press, 80–109.

Torre, Andrew. 2003. "The impact of court delays on the prosecutor and the defendant: An economic analysis." *European Journal of Law and Economics* 16 (1):91–111.

Tversky, Amos and Daniel Kahneman. 1979. "Prospect theory: An analysis of decision under risk." *Econometrica: Journal of the Econometric Society* 47 (2):263–291.

———. 1992. “Advances in prospect theory: Cumulative representation of uncertainty.” *Journal of Risk and Uncertainty* 5 (4):297–323.

Weatherburn, Don. 2012. “The effect of arrest and imprisonment on crime.” *Crime and Justice Bulletin* 158:1–20.

Weisburd, David, Nancy A Morris, and Elizabeth R Groff. 2009. “Hot spots of juvenile crime: a longitudinal study of arrest incidents at street segments in Seattle, Washington.” *Journal of Quantitative Criminology* 25 (4):443–467.

Wooldridge, Jeffrey. 2012. *Introductory econometrics: A modern approach*. Boston, MA.: Cengage Learning.

Wooldridge, Jeffrey M. 2002. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT press.

Worrall, John L and Travis C Pratt. 2004. “Estimation issues associated with time-series-cross-section analysis in criminology.” *Western Criminology Review* 5:35–49.

Yang, Crystal S. 2016. “Resource Constraints and the Criminal Justice System: Evidence from Judicial Vacancies.” *American Economic Journal: Economic Policy* 8 (4):289–332.

A. Appendix A: Proofs

Proof 1 *Proof of Proposition 1 Part (i): Replacing $m_G = 1 - a_G$ and rewriting equation 2 yields:*

$$b^* = [c_G(a_G + (1 - a_G)\delta_G) - c_I m_I \delta_I] \times f \tag{5}$$

Thus:

$$\frac{\partial b^*}{\partial a_G} = c_G(1 - \delta_G)f > 0$$

as $\delta_G < 1$ by assumption.

Proof 2 *Proof of Proposition 1 Part (ii): Replacing $r_I = 1 - m_I$ and rewriting equation 2 yields*

$$b^* = [c_G(a_G + m_G\delta_G) - c_I(1 - r_I)\delta_I] \times f \tag{6}$$

Thus, $\frac{\partial b^*}{\partial r_I} = c_I \delta_I f > 0$.

Proof 3 *Proof of Proposition 1 Part (iii): Replacing $a_G = 1 - m_G$ and rewriting equation 2 yields*

$$b^* = [c_G((1 - m_G) + m_G\delta_G) - c_I m_I \delta_I] \times f \tag{7}$$

Thus:

$$\frac{\partial b^*}{\partial m_G} = c_G(-1 + \delta_G)f < 0$$

as $\delta_G < 1$ by assumption.

and:

$$\frac{\partial b^*}{\partial m_I} = -c_I \delta_I f < 0$$

B. Appendix: Data

The data on proceedings in the federal courts relates to three different levels:

- *District (trial) courts* - 94 district courts serve as a first instance for criminal cases (18 U.S. Code, §3231). Each court is located within one state or a jurisdictional territory. Appeals are filed to a circuit appellate court. Exact definitions of jurisdiction can be found in 28 U.S.C 41, 81-131.
- *States and territories* - each state contains one or more district (trial) courts, but belongs to only one "circuit", which includes only one appellate court. In addition to the states, four "territories" are included in the judicial system.[55]
- *Circuit appellate courts* - 12 appellate courts review criminal appeals which originate from district courts.[56] Eleven courts (numbered 1st to 11th) review appeals from multiple district courts. The 12th court ("DC") reviews appeals from one district court only.

An updated map of the different judicial districts and further information on court structure can be found in the federal courts' website (http://www.uscourts.gov/about-federal-courts/court-role-and-structure.)

The following datasets have been combined in the analysis:

Supreme court database Harold J. Spaeth, Lee Epstein et al. 2016 Supreme Court Database, Version 2016 Release 1, URL: http://Supremecourtdatabase.org

Buro of Justice statistics

- Federal Criminal Case Processing Statistics.

  URL: http://www.bjs.gov/fjsrc/
- Corrections Statistical Analysis Tool (CSAT) - Prisoners.

  URL: http://www.bjs.gov/index.cfm?ty=nps

Federal courts statistics Judicial business of the U.S. courts (1997-2013).

URL: http://www.uscourts.gov/

[55] The four territories are: Puerto Rico, Virgin Islands, Guam and the Northern Mariana Islands.
[56] An additional judicial district is the "federal district", which does not review criminal appeals.

ICPSR database: appellate courts

- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 1993-1998 . ICPSR06559-v4. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2004-11-05.
  http://doi.org/10.3886/ICPSR06559.v4
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 1999 . ICPSR03105-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2001. http://doi.org/10.3886/ICPSR03105.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2000 . ICPSR03493-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2002. http://doi.org/10.3886/ICPSR03493.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2001 . ICPSR03494-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2002. http://doi.org/10.3886/ICPSR03494.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2001 . ICPSR03494-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2002. http://doi.org/10.3886/ICPSR03494.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2002 . ICPSR04108-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2005. http://doi.org/10.3886/ICPSR04108.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2003. ICPSR04632-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-03-20. http://doi.org/10.3886/ICPSR04632.v2
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2004. ICPSR04629-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-03-20. http://doi.org/10.3886/ICPSR04629.v2
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2005. ICPSR04627-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-03-21. http://doi.org/10.3886/ICPSR04627.v2
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2006. ICPSR20101-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-09-26. http://doi.org/10.3886/ICPSR20101.v2
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2007. ICPSR22624-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2011-03-08. http://doi.org/10.3886/ICPSR22624.v2

- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2008. ICPSR25425-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-03-04. http://doi.org/10.3886/ICPSR25425.v2
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2008. ICPSR25425-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-03-04. http://doi.org/10.3886/ICPSR25425.v2
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2009. ICPSR28601-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-09-24. http://doi.org/10.3886/ICPSR28601.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2010. ICPSR35337-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-09-24. http://doi.org/10.3886/ICPSR35337.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2011. ICPSR35340-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-11-25. http://doi.org/10.3886/ICPSR35340.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2012. ICPSR35343-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-12-03. http://doi.org/10.3886/ICPSR35343.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Convictions and Sentences: Appeals Data, 2013. ICPSR35346-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-12-03. http://doi.org/10.3886/ICPSR35346.v1

ICPSR database: district courts

- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 1987-1998. ICPSR09317-v5. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2012. http://doi.org/10.3886/ICPSR09317.v5
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 1999. ICPSR03106-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2001. http://doi.org/10.3886/ICPSR03106.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2000. ICPSR03496-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2002. http://doi.org/10.3886/ICPSR03496.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2001. ICPSR03497-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2005. http://doi.org/10.3886/ICPSR03497.v1

- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2002. ICPSR04110-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2004. http://doi.org/10.3886/ICPSR04110.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2003. ICPSR04290-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007. http://doi.org/10.3886/ICPSR04290.v3
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2004. ICPSR04633-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007. http://doi.org/10.3886/ICPSR04633.v2
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2005. ICPSR04630-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007. http://doi.org/10.3886/ICPSR04630.v2
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2006. ICPSR20120-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007. http://doi.org/10.3886/ICPSR20120.v2
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2007. ICPSR22623-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009. http://doi.org/10.3886/ICPSR22623.v2
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2008. ICPSR25424-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009. http://doi.org/10.3886/ICPSR25424.v2
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2009. ICPSR28602-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014. http://doi.org/10.3886/ICPSR28602.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2010. ICPSR35336-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014. http://doi.org/10.3886/ICPSR35336.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2011. ICPSR35339-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014. http://doi.org/10.3886/ICPSR35339.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2012. ICPSR35342-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014. http://doi.org/10.3886/ICPSR35342.v1
- United States Sentencing Commission. Monitoring of Federal Criminal Sentences, 2013. ICPSR35345-v1. Ann Arbor, MI:

Inter-university Consortium for Political and Social Research [distributor], 2014.
http://doi.org/10.3886/ICPSR35345.v1

JCS scores

- Prof. Lee Epstein's website: http://epstein.wustl.edu/research/JCS.html
- Prof. Christina Boyd's website: Boyd, Christina L. 2015. "Federal District Court Judge Ideology Data." available at: http://cLboyd.net/ideology.html
- In combination with Biographical Directory of Article III Federal Judges, 1789-present, available at: https://www.fjc.gov/history/judges

The federal justice center database on biography of judges provides information on which judges served in each court at a given year. This database is then crossed with JCS scores for appellate judges (Epstein's website) and district judges (Boyd's website) to achieve a court-by-year distance.

C. Appendix: Additional Tables and Figures

Table C1 CORRELATIONS: RATES OF OUTCOMES IN CRIMINAL APPEALS

	(1) Affirmance rate	Reversal rate	Remand rate	Dismissal rate	Remainder rate	Reversal-Remand spread
Affirmance rate	1.00					
Reversal rate	-0.27***	1.00				
Remand rate	-0.42***	-0.19***	1.00			
Dismissal rate	-0.80***	-0.00	-0.05*	1.00		
Remainder rate	-0.24***	-0.01	0.23***	-0.01	1.00	
Reversal-Remand spread	0.17***	0.68***	-0.85***	0.03	-0.18***	1.00

NOTE.– This table presents correlations between the rates of outcomes in criminal appeals for the whole sample (1997-2013) in the federal courts. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table C2 DESCRIPTIVE STATISTICS: CONTROL VARIABLES

	mean	sd	min	max
Conviction rate	89.50	6.64	46.55	99.01
Arrests per capita	26.33	47.83	2.70	567.01
Prosecution rate	83.70	23.81	17.84	349.02
Share of jury cases	4.25	2.59	0.00	18.34
Share of bench-trials	0.65	1.71	0.00	34.15
Fine-only rate	3.00	7.23	0.00	67.84
Imprisonment rate	80.39	11.45	14.93	98.12
Appeal-filing rate	19.35	8.48	1.95	51.41
State prison population	39885.34	44093.27	915.00	175512.00
Female defendants share	14.89	4.13	2.68	32.05
Black defendants share	30.84	18.20	0.00	78.03
Hispanic defendants share	28.20	21.78	0.68	96.25
UCR Violent crime sentences share	4.68	3.71	0.00	33.11
UCR Property crime sentences share	4.25	3.59	0.00	42.86
Conviction-only share	28.40	13.14	0.00	83.33
Sentence+Conviction share	20.31	9.56	0.00	68.18
Average imprisonment length (years)	6.50	4.43	1.02	45.97
SC full affirmance rate	30.81	28.09	0.00	100.00
SC reversal rate	64.52	28.74	0.00	100.00
SC partial affirmance rate	2.85	9.15	0.00	100.00
SC dismissal rate	1.41	4.86	0.00	25.00
Observations	1424			

NOTE.–This table presents descriptive statistics for additional control variables used in the analysis. The figures refers to the year range 1997-2013.

Table C3

SAMPLE SELECTION PROCESS

Database/Source	N	Cumulative observations	Years range
All (91 courts * 17 years)	1547	1547	1997-2013
Conviction rates (missing for OR 2002)	-1	1546	1997-2013
Arrest rates (year>1997 only)	-91	1455	1998-2013
Prosecution rate (missing for PA,W 1998,1999)	-2	1453	1998-2013
State prison population (missing for PR, DC)	-29	1424	1998-2013
Total		1424	1998-2013

Note - This table presents the origins of the observations used in the regressions, where some observations drop due to missing values.

Table C4

ROBUSTNESS CHECK - QUADRATIC AND CUBIC TERMS

	(1) Two-way cluster	(2) PCSE (AR1)	(3) PCSE (PSAR1)	(4) SCC
Criminal Reversal-Remand spread	-0.229***	-0.102***	-0.077***	-0.229***
	(0.064)	(0.025)	(0.025)	(0.041)
Criminal Affirmance rate (centered)	0.048	-0.049	-0.089	0.048
	(0.129)	(0.069)	(0.057)	(0.108)
Criminal Affirmance rate (centered) squared	-0.016**	-0.008***	-0.007***	-0.016***
	(0.007)	(0.003)	(0.002)	(0.004)
Criminal Affirmance rate (centered) cubic	-0.001*	-0.000**	-0.000**	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Criminal Affirmance rate (centered) squared	0.012**	0.008**	0.006**	0.012***
	(0.005)	(0.003)	(0.003)	(0.004)
Conviction rate (centered) squared	-0.003	0.000	0.001	-0.003
	(0.008)	(0.004)	(0.004)	(0.006)
Conviction rate (centered) cubic	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Criminal Dismissal rate	0.069	-0.005	-0.020	0.069
	(0.126)	(0.080)	(0.071)	(0.093)
Criminal Remainder rate	-0.498	-0.118	-0.133*	-0.498
	(0.526)	(0.099)	(0.071)	(0.472)
Conviction rate (centered)	0.055	0.056*	0.034	0.055
	(0.072)	(0.033)	(0.034)	(0.058)
Control variables	Yes	Yes	Yes	Yes
Court and Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	1424	1424	1424	1424

NOTE.–This table compares regression results of the different methods, with quadratic and cubic affirmance rates. Columns 1-5 include an interaction of centered (demeaned) affirmance and conviction rates. Control variables are identical to those used in Table 3. Constant and FE coefficients are not reported. $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Table C5

DESCRIPTIVE STATISTICS - SENTENCING COMMISSION DATABASE

	mean	sd	min	max
Pooled rates (defendant and government)				
Full affirmance rate (senetence)	75.35	13.98	0.00	100.00
Partial affirmance rate (sentence)	4.49	4.70	0.00	31.03
Reversal rate (sentence)	9.66	7.75	0.00	71.43
Remand rate (sentence)	2.10	5.05	0.00	45.95
Dismissal rate (sentence)	8.39	8.78	0.00	44.44
Reversal-remand spread (sentence)	7.56	8.29	-39.13	57.14
Defendant-filed				
Affirmance rate (defendant appeals)	76.45	14.08	0.00	100.00
Reversal rate (defendant appeals)	8.68	7.48	0.00	71.43
Partial affirmance rate (defendant appeals)	4.27	4.57	0.00	31.03
Remand rate (defendant appeals)	2.06	4.97	0.00	45.95
Dismissal rate (defendant appeals)	8.53	8.91	0.00	46.67
Reversal-Remand spread (defendant appeals)	6.62	7.92	-39.13	57.14
Government-filed				
Affirmance rate (government appeals)	29.48	41.49	0.00	100.00
Reversal rate (government appeals)	64.15	43.20	0.00	100.00
Partial affirmance rate (government appeals)	3.92	16.78	0.00	100.00
Remand rate (government appeals)	2.05	13.43	0.00	100.00
Dismissal rate (government appeals)	0.39	4.76	0.00	100.00
Reversal-Remand spread (government appeals)	66.20	42.65	0.00	100.00
Government appeals share	1.71	2.77	0.00	23.53
Observations	1438			

NOTE.– This table presents descriptive statistics for the sentencing appeals rates over 1997-2013. Appeal results are measured at the district court level.

D. Appendix: Comparison of Methodologies

Table compares the five different methods used in the paper, and provides citations to several papers on crime or court behavior which have previously employed the method:

Table C1
COMPARISON OF METHODS

Method	Advantages	Disadvantages	Examples
Fixed effects, two-way clustered standard errors ("**Two-way cluster**")	Robust for heteroskedasticity; Allows for correlation within-court and within-year; does not require strong assumptions	Weaker performance given spatial correlation and low number of clusters	Beraldo, Caruso, and Turati (2013); Gathmann (2008).
Fixed effects, Panel corrected standard errors ("**PCSE**")	Robust for heteroskedasticity, autocorrelation and contemporaneous correlation	Assumes very specific autoregressive process: **AR(1)** assumes one-period correlation disturbance, which is the same for all panels. **PSAR(1)** assumes panel-specific one-period auto correlation	Kovandzic and Vieraitis (2006); Scott (2006)
Fixed effects, Driscoll-Kraay standard errors ("**SCC**")	Robust for heteroskedasticity, autocorrelation, spatial correlation and large N (number of panels)	Weaker performance for panels with small T (number of periods, i.e. number of years) or little spatial correlation	Cotte Poveda (2011)

Note – This table compares the five different methods used in the base model regressions. Columns (2) and (3) briefly describe the advantages and disadvantages of each method. The last column lists examples for papers that have used the method. The Stata commands used in each methods are as follows: (1) FE two-way cluster: ivreg2 or reghdfe (Correia et al., 2015); (2) PCSE: xtpcse, c(ar1) *or* c(psar1). (3) ivreg2, dkraay(2) (Baum, Schaffer, and Stillman, 2007), which produces the same result as the xtscc command (Hoechle, 2007). Bandwidth of 2 was chosen by the rule of thumb $T^{0.25}$.

E. Appendix: Additional Analysis of Controls

Table E1 below is an extended version of Table 3, which includes the coefficients of control variables.

Table E1

BASE RESULTS - COMPARISON OF METHODS

	(1) Two-way cluster	(2) PCSE (AR1)	(3) PCSE (PSAR1)	(4) SCC
Main variables				
Criminal Reversal-Remand spread	-0.192***	-0.085***	-0.060**	-0.192***
	(0.059)	(0.026)	(0.028)	(0.041)
Criminal Affirmance rate (centered)	-0.011	-0.064	-0.093**	-0.011
	(0.112)	(0.052)	(0.045)	(0.102)
Conviction rate (centered)	-0.004	0.036	0.023	-0.004
	(0.070)	(0.043)	(0.045)	(0.065)
Criminal Affirmance rate (centered) $\times$ Conviction rate (centered)	0.012**	0.008***	0.007**	0.012***
	(0.005)	(0.003)	(0.003)	(0.004)
Additional Controls				
Criminal Dismissal rate	0.011	-0.028	-0.043	0.011
	(0.138)	(0.063)	(0.061)	(0.112)
Criminal Remainder rate	-0.411	-0.127	-0.169***	-0.411
	(0.498)	(0.098)	(0.056)	(0.470)
Arrests per capita	-0.034*	-0.021***	-0.018***	-0.034***
	(0.018)	(0.004)	(0.004)	(0.009)
Prosecution rate	-0.012	0.012	0.018**	-0.012
	(0.029)	(0.007)	(0.007)	(0.019)
Share of jury cases	-0.061	0.006	0.037	-0.061
	(0.172)	(0.049)	(0.048)	(0.148)
Share of bench-trials	-0.055	-0.190**	-0.116	-0.055
	(0.121)	(0.080)	(0.090)	(0.126)
Fine-only rate	0.016	-0.037	-0.044	0.016
	(0.088)	(0.056)	(0.053)	(0.068)
Imprisonment rate	0.068	0.037	0.026	0.068*
	(0.072)	(0.027)	(0.024)	(0.041)
Appeal-filing rate	-0.042	-0.019	-0.023	-0.042
	(0.052)	(0.024)	(0.025)	(0.043)
State prison population	0.000	0.000	0.000*	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Female defendants share	0.116	0.011	0.002	0.116**
	(0.092)	(0.030)	(0.033)	(0.058)
Black defendants share	0.033	0.012	0.008	0.033*
	(0.045)	(0.014)	(0.016)	(0.017)
Hispanic defendnats share	0.025	0.005	-0.004	0.025
	(0.068)	(0.020)	(0.024)	(0.029)
UCR Violent crime sentences share	0.167	0.035	-0.012	0.167
	(0.197)	(0.074)	(0.081)	(0.126)
UCR Property crime sentences share	-0.379**	-0.105**	-0.110**	-0.379***
	(0.146)	(0.044)	(0.045)	(0.110)
Conviction-only share	-0.005	0.003	0.007	-0.005
	(0.030)	(0.007)	(0.007)	(0.022)
Sentence+Conviction share	-0.007	0.013*	0.012	-0.007
	(0.034)	(0.007)	(0.009)	(0.024)
Average imprisonment length (years)	0.289***	0.139***	0.140**	0.289***
	(0.045)	(0.050)	(0.060)	(0.048)
SC full affirmance rate	-0.119***	-0.047***	-0.044**	-0.119***
	(0.034)	(0.014)	(0.021)	(0.022)
SC reversal rate	-0.097**	-0.046***	-0.045**	-0.097***
	(0.035)	(0.014)	(0.021)	(0.023)
SC partial affirmance rate	-0.147**	-0.060***	-0.054***	-0.147***
	(0.050)	(0.014)	(0.020)	(0.036)
SC dismissal rate	-0.093*	-0.053**	-0.057**	-0.093**
	(0.046)	(0.022)	(0.028)	(0.045)
Court and Year Fixed Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Number of observations	1424	1424	1424	1424

NOTE.– This table compares results of the basic model using 5 different methods: (1) Two-way clustered S.E. (by court and year); (2) Panel-corrected S.E. (AR1 disturbance) ; (3) Panel-corrected S.E. (panel-specific AR1 disturbance); (4) Driskol-Kraay standard errors, also known as 'SCC'; (5) Poisson with roust S.E. Constant and FE coefficients are not reported. Note that Poisson regression uses a different dependent variable. $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Coefficients of control variables are mostly insignificant, with a few exceptions:

- *Arrest rates:* negative coefficient, implying that crime decreases in apprehension probability.
- *Prosecution rates:* mixed signs in different regressions. A positive coefficient may occur, e.g. if prosecutors are perceived as instigators of wrongful convictions.
- *Average imprisonment length:* positive coefficients, which seem counter-intuitive, as more years in prison imply a higher f. However, several theories have argued that increasing sentence length can in fact lead to more crime. For example, longer imprisonments may facilitate a "school for crime" (Roberts and Hough, 2005, pp. 297) within prisons. Imprisonment may increase crime also by reducing prisoners' future wage prospects, leading to irregular employment and higher crime commission incentives (Johnson, 2016, pp. 281). For additional arguments suggesting that higher punishments may decrease deterrence, seeGneezy and Rustichini (2004); Caulkins (1993); Kahan (1996); Dickens (1986); Livernois and McKenna (1999); Huck and Kosfeld (2007).
- *Share of UCR property crime:* negative coefficient, possibly reflecting that property crime is more likely to end in conviction (and thus in sentencing).
- *Share of bench trials:* negative coefficient, reflecting that judges are perhaps perceived as less likely to err compared to juries.
- *SC-rates:* significant negative coefficients, slightly smaller than those of appellate courts. It seems surprising that the effect is not much smaller, given the low likelihood of SC review. A possible explanation may be that the salience of Supreme Court decisions leads to behavioral overweighting of their importance (Korobkin and Ulen, 2000, for a discussion on criminals' response to salient outcomes, see).