Vanderbilt Law Review

Volume 59 | Issue 2 Article 2

3-2006

The Impact of Joinder and Severance on Federal Criminal Cases: An Empirical Study

Andrew D. Leipold

Hossein A. Abbasi

The Impact of Joinder and Severance on Federal Criminal Cases: An Empirical Study

Andrew D. Leipold
Hossein A. Abbasi

59 Vand. L. Rev. 349 (2006)

It is widely assumed that criminal defendants who face multiple charges in a single trial have a harder time prevailing than those who face several trials of one count each. Conventional wisdom also has it that a defendant who is joined for trial with other suspects is worse off than one who stands trial alone. Until now, these assumptions have never been tested empirically. Looking at nearly 20,000 federal criminal trials over a five-year period, this Article asks if the traditional beliefs are true, and if so, tries to measure the impact on trial outcomes of joining counts and defendants.

It turns out that joinder has a significant prejudicial effect on trials, although not quite the same effect that is usually assumed. Using statistical models that control for a range of variables, the authors discovered that trial defendants who face multiple counts are roughly 10% more likely to be convicted of the most serious charge than a defendant who stands trial on a single count. Surprisingly, however, joining co-defendants in a single trial had virtually no impact on the likelihood of conviction, at least in the aggregate. Using the results of the empirical study, the Article then reconsiders the competing policy issues that underlie the joinder and severance doctrine, and explores the implications of the findings.

***

The Impact of Joinder and Severance on Federal Criminal Cases: An Empirical Study

Andrew D. Leipold[*]

Hossein A. Abbasi[**]

* Professor, University of Illinois College of Law. Thanks to Wayne LaFave, Richard McAdams, Jacqueline Ross, Margareth Etienne, Patrick Keenan, and the faculty of the University of Utah, S.J. Quinney College of Law for their helpful comments on drafts of this paper.

** Ph.D candidate, Department of Economics, University of Illinois at Urbana-Champaign.

I. INTRODUCTION

Dave is in trouble. It was bad enough to be arrested for bank robbery; now he has learned that the prosecutor plans to join the current charge with three other, unrelated bank robberies and present all four counts in a single trial. To his priest and to his lawyer, Dave admits that he committed the first and the second robberies, but he did not commit the third or fourth. Dave is smart enough to realize, however, that once the jury starts hearing evidence of some of the crimes—all of which will sound quite similar—his ability to cast doubt on the remaining charges will be dimmed. And Dave's lawyer is smart enough to know that once the charges are joined, the chances of splitting them apart are relatively small.

It is widely assumed that criminal defendants who face multiple charges in a single trial have a harder time prevailing than those who face several trials of one count each. Conventional wisdom also has it that a defendant who is joined for trial with other suspects is in a worse position than one who stands trial alone. These assumptions have never been tested empirically; this Article tries to fill the gap. Looking at nearly 20,000 federal criminal trials over a five-year period, the Article asks if the traditional beliefs are true and, if so, tries to measure the impact on trial outcomes of joining counts and defendants.

The effect of joinder on criminal cases is part of a larger debate about how best to manage a growing criminal docket while still providing individual justice. The battle lines are easy to describe: courts and prosecutors typically want joined proceedings, defendants usually don't. Courts believe that consolidated proceedings play a "vital role"[1] in the administration of justice; defendants believe that they are a source of great prejudice. The problem is that both sides are right.

The details of joinder and severance law are dry, even boring, and perhaps as a result, the impact of consolidated trials has received

1. Richardson v. Marsh, 481 U.S. 200, 209 (1987).

little scholarly attention.[2] But the consequences are widespread: more than half of all federal defendants are charged with multiple counts, roughly one-third are joined with other defendants, and an overlapping one-quarter face both—a single proceeding with multiple charges plus one or more co-defendants.[3] If joinder makes a conviction significantly more likely, it should have a bearing on prosecutors' charging decisions, judicial rulings on severance motions, and defense decisions on whether to plead or stand trial. More importantly, understanding the dimensions of any prejudice should tell us something important about the tradeoffs we make between fairness to the accused and efficiency in processing criminal cases.[4]

Part II provides some background and describes the risks created by joinder. Part III(A) sets forth some working hypotheses and then offers an original empirical case for the prejudicial impact of joinder on the defense. Part III(B) then tests the empirical case with statistical models, trying to control for various features besides joinder that might explain the differences in trial outcome. To preview the results: it turns out that there is a measurable and significant prejudicial effect from joining multiple *counts* in a single trial, but the impact of joining multiple *defendants* is far less clear. With the results of the empirical test in hand, Part IV argues for a reconsideration of the competing interests and poses questions for future study.

---

2. The best analysis of joinder and severance remains Robert O. Dawson, *Joint Trials of Defendants in Criminal Cases: An Analysis of Efficiencies and Prejudices*, 77 MICH. L. REV. 1379 (1979). *See also* James R. Lucas, *Criminal Joinder and Severance*, 57 INTER ALIA 17 (1992); John F. Decker, *Joinder and Severance in Federal Criminal Cases: An Examination of Judicial Interpretation of the Federal Rules*, 53 NOTRE DAME LAWYER 147 (1977); Note, *Joint and Single Trials Under Rules 8 and 14 of the Federal Rules of Criminal Procedure*, 74 YALE L.J. 553, 560 (1965).

3. *See infra* Tables A, B, and C in Part III.A.1 (reporting statistics for federal criminal defendants between 1999 and 2003).

4. Although this paper will address only joinder in federal cases, it is worth noting that many states follow the federal formulation of the rules, at least with respect to joinder of offenses. 4 WAYNE R. LAFAVE ET AL., CRIMINAL PROCEDURE 593 (2d ed. 1999) ("The overwhelming majority of states have adopted the federal language [in Rule 8(a)] or something very close to it." (footnote omitted)). With respect to joinder of defendants, "[t]he statutes and court rules in over a third of the states utilize [Rule 8(b)'s] language, while some others either add to that language or else employ wording of about the same specificity permitting either somewhat broader or somewhat narrower joinder than in the federal system" (footnotes omitted). *Id.* at 603–04.

II. BACKGROUND

A. The Rules and Rationales

There are two types of joinder in criminal cases: combining multiple charges against a defendant in a single case ("joinder of offenses") and charging multiple defendants in a single proceeding ("joinder of defendants"). The two types can be combined, of course, so that multiple counts can be brought against multiple defendants in one consolidated case.[5]

The initial decision to join charges or defendants is made by the prosecutor. She drafts the indictment that, when approved by the grand jury, will name the defendants charged and specify the counts against each. Presumptively, all counts and all defendants named in one indictment will be resolved in a single trial. This initial joinder decision is discretionary: the prosecutor "may" charge multiple offenses or multiple defendants in one indictment but is not required to do so.[6] If the government originally files separate indictments against different defendants or separate charges against a single accused, the prosecutor can later combine them for a single trial, although now she needs permission from the court.[7] In theory, a defendant can request that all charges against him be consolidated or that his trial be combined with those of other defendants, but as we will see, joinder provides so few benefits to the accused that motions of this type are rare.[8]

Charges and defendants cannot be combined indiscriminately; there must be some nexus among the pieces to justify a single trial. Rule 8(a) of the Federal Rules of Criminal Procedure provides that the prosecutor can charge a defendant with multiple counts in a single

5. Unless the text or context suggests otherwise, we will use the term "joinder" and its derivatives to refer to all types of joined proceedings, involving both charges and defendants. At times the term "consolidation" will also be used, referring to the same concept as joinder.

6. FED. R. CRIM. P. 8(a)-(b). The ability to join or not join charges is, however, subject to limits imposed by the Double Jeopardy Clause of the Fifth Amendment. *See infra* note 135 and accompanying text (explaining how the charging of several overlapping crimes may result in the dismissal of some charges pursuant to the Double Jeopardy Clause of the Fifth Amendment).

7. *See* FED. R. CRIM. P. 13 ("The court may order that separate cases be tried together as though brought in a single indictment or information if all offenses and all defendants could have been joined in a single indictment or information."). *Cf.* U.S. v. Burkley, 591 F.2d 903, 919 (D.C. Cir. 1978) ("[T]he wording of [Rule 13] is permissive; all offenses which could be tried together thereunder need not be.").

8. *See infra* Part II.B.

trial in one of three settings.[9] First, joinder is allowed if the alleged crimes are based on the "the same act or transaction." Thus, the defendant who robs a liquor store and discharges a weapon can be required to face both charges in one trial.[10] Second, joinder is permitted if the alleged crimes are part of a "common scheme or plan," such as when a middleman buys drugs from a supplier then sells them to a distributor.[11] Finally, charges can be joined if they are of the "same or similar character." A defendant who buys stolen goods in June and then again in October may face both counts at once, even if the two charges are distant in time and location and even if they are not part of an overarching criminal plan.[12]

The rules are similar hut not identical for joined defendants. Despite Justice Jackson's observation that "guilt is personal,"[13] trials often are not. Under Rule 8(b), multiple defendants can be charged in a single indictment if they are alleged to have participated in the same act or transaction, or the same series of acts or transactions, that constituted a crime.[14] Once defendants are consolidated, the counts against each *individual* defendant must meet the requirements for joinder of offenses, but since each defendant need not be charged with each count in the indictment,[15] it is possible that a great deal of evidence that will be heard at trial will have nothing to do with some of the defendants. Thus, Defendant X who is charged only with

---

9. FED. R. CRIM. P. 8(a). Rule 8 was intended to codify the law as it existed when the Federal Rules became effective in 1946, and its language has not changed since then. 1A CHARLES ALAN WRIGHT, FEDERAL PRACTICE & PROCEDURE: CRIMINAL 3 (1999). For an extensive discussion of the history of Rule 8 and of federal practice prior to the adoption of the Rule, see Lester B. Orfield, *Joinder in Federal Criminal Procedure*, 26 F.R.D. 23 (1961).

10. *See, e.g.*, United States v. Woody, 55 F.3d 1257, 1267 (7th Cir. 1995) (holding that possession of stolen mail and assault on postal inspector were part of same "transaction" for joinder purposes where assault was an attempt to avoid arrest for the illegal possession). The court in *Woody* noted that it used an "expansive interpretation of what constitutes a 'transaction' under Rule 8(a)." *Id.*

11. *See, e.g.*, United States v. Johnson, 130 F.3d 1420, 1427 (10th Cir. 1997) (finding joinder of felon-in-possession count with drug distrihution count was proper, since "the handgun was arguably related to and part of Johnson's drug trafficking scheme").

12. *See, e.g.*, United States v. Tyndall, 263 F.3d 848, 850 (8th Cir. 2001) (finding sexual attacks on two different victims two years apart were of the "same or similar character" for Rule 8 purposes).

13. *Cf.* Korematsu v. United States, 323 U.S. 214, 243 (1944) (Jackson, J., dissenting) ("Now, if any fundamental assumption underlies our system, it is that guilt is personal and not inheritable.").

14. FED. R. CRIM. P. 8(b). Note that in contrast to the joinder of offenses, multiple defendants who merely engage in "similar acts" may not be joined for trial unless those acts were part of the same transaction or series of transactions. *See* United States v. Frost, 125 F.3d 346, 389 (6th Cir. 1997) (stating that joinder of defendants under Rule 8(b) is proper "only if each of the counts of the indictment arises out of the same act or transaction or series of transactions").

15. FED. R. CRIM. P. 8(b) ("[A]ll defendants need not be charged in each count.").

conspiracy may be forced to sit through a trial where the jury hears evidence not only of the conspiracy between Defendants X and Y, but evidence of how Y also committed murder and extorted money from widows and orphans, crimes for which X was not charged and was not even allegedly involved.[16]

The advantages of joint trials to courts, prosecutors, and society are obvious.[17] Trials are expensive, time consuming, and burdensome on witnesses and victims, so if a matter can be resolved with one trial instead of two or three, the saved resources are significant. Often the evidence for each count and against each defendant will be substantially the same. Thus, it would seem wasteful to require witnesses to attend and give the same testimony three times, to require three sets of jurors to consider identical evidence, and to require the lawyers and judges to go through multiple, identical performances. And of course, courts are always pressed for time: at the end of 2004 there were over 60,000 criminal cases pending in the federal district courts, an increase of over 100% since 1995.[18] In a world with few enough resources to conduct a single trial correctly, holding three very similar proceedings seems like an extravagance.

Efficiency is the cardinal justification for the joinder rules but not the only one. A secondary justification is the desire to avoid the "scandal and inequity of inconsistent verdicts."[19] If two defendants are charged with the same crime and face identical evidence, a single trial reduces the chance that a jury will acquit one and convict the other. It would be embarrassing, after all, to have different juries reach different results when presented with the same facts, especially if there is nothing to suggest that one jury was inclined to acquit against the evidence. Consolidated trials reduce the risk that a jury

16. *See, e.g.*, United States v. Flores-Rivera, 56 F.3d 319, 325 (1st Cir. 1995) (finding that joinder was proper where defendant was alleged to have participated in the overall scheme, even though he was named in fewer than ten percent of the overt acts in furtherance of the conspiracy).

17. For a discussion of the advantages of joint trials, see Richardson v. Marsh, 481 U.S. 200, 209–10 (1987); *see also* Dawson, *supra* note 2, at 1382–89 (discussing the "supposed efficiencies" of consolidated trials).

18. Administrative Office of the United States Courts, Federal Judicial Caseload Statistics, Judicial Caseload Indicators, Calendar Years 1995, 2000, 2003, and 2004, http://www.uscourts.gov/caseload2004/front/judbus03.pdf (last visited Mar. 24, 2006). At the end of 1995 the number of cases pending was 28,091; by 2004, the number was 62,857, an increase of 124% over nine years. During 2004 there were 70,746 new criminal cases filed. *Id.*

19. *See Marsh*, 481 U.S. at 210.

will act in a way that looks to outsiders like capricious decisionmaking.[20]

Because of these interests, there is a "preference in the federal system for joint trials of defendants who are indicted together."[21] But courts and other observers also have long understood that the benefits of efficiency and consistency may come at a high cost.

B. The Risk of Prejudice

A suspect facing joined counts is worried about three types of risk. His first concern is that the jury will look at the list of accusations against him and infer that he has a criminal disposition, which could erode the presumption of innocence.[22] Holmes observed that "it cannot be said that by common experience the character of most people indicted by a grand jury is good,"[23] and the more counts in the indictment, the quicker the jury may be to assume that the accused must be guilty of *something*.[24]

A second worry is that multiple counts require multiple blocks of evidence, and the jury (or judge) might confuse which block of evidence relates to which crime.[25] Particularly when the charges are

---

20. For an insightful discussion of inconsistent jury decisionmaking, see Eric L. Muller, *The Hobgoblin of Little Minds? Our Foolish Law of Inconsistent Verdicts*, 111 HARV. L. REV. 771 (1998).

21. *See* Zafiro v. United States, 506 U.S. 534, 537 (1993).

22. *See* United States v. Werner, 620 F.2d 922, 929 (2d Cir. 1980) (asserting that "[w]hile the mere fact that 'juries are apt to regard with a more jaundiced eye a person charged with two crimes than a person charged with one'" does not require severance, trial judges must be alert to the possibility that juries will cumulate the evidence of the separate charges and convict); *see also* United States v. Foutz, 540 F.2d 733, 736 (4th Cir. 1976) (noting that although "other crimes" are normally inadmissible to prove a criminal disposition, "[o]ne inevitable consequence of a joint trial is that the jury will be aware of evidence of one crime while considering the defendant's guilt or innocence of another. If the rationale of the 'other crimes' rule is correct, it would seem that some degree of prejudice is necessarily created by permitting the jury to hear evidence of both crimes." (citations omitted)).

23. Greer v. United States, 245 U.S. 559, 561 (1918).

24. Social scientists have found in laboratory experiments that the "halo effect" of multiple charges can lead mock jurors to find the accused less believable, less likeable, and more dangerous than people who are charged with only a single count, with the result that "a defendant is more likely to be convicted by mock jurors of any one charge when that offense is combined with another at trial." Edith Greene & Elizabeth F. Loftus, *When Crimes Are Joined at Trial*, 9 LAW & HUM. BEHAVIOR 193, 194, 197–98 (1985). The researchers are careful to note, however, that their experiment could not clearly determine whether the mock jurors' attitude about the defendant's character was a cause or effect of their conclusion about the defendant's guilt. They could only conclude that "jurors' sentiments toward the defendant is [sic] different when they know that he is charged with multiple crimes." *Id.* at 198.

25. The social science research supports the notion of jury confusion, although its impact on verdicts is less clear. One laboratory study found that jurors were likely to become confused by the evidence in a joined-offense trial, and that this inability to compartmentalize correlated to

of a "same or similar character,"[26] a jury might, for example, be so impressed with the evidence on counts one and two that it fails to notice that there was insufficient evidence on the very-similar count three.[27] The danger is compounded when evidence of wrongdoing that would not be admitted in separate trials is admitted in a single trial, despite the presence of instructions reminding jurors that certain facts should only be considered with respect to certain counts.[28]

Finally, for the accused who has a defense to some charges but not others, a multi-count trial can raise a thorny strategic choice: if he takes the stand to defend himself on Count 1 – a crime he did not commit – he opens himself up to cross-examination on Count 2, a crime he did commit. On the other hand, if the defendant stays off the stand entirely to protect himself on Count 2, his chances of acquittal on Count 1 diminish.[29] In short, the all-or-nothing aspect of taking

---

guilty verdicts. *See* Kenneth S. Bordens & Irwin A. Horowitz, *Information Processing in Joined and Severed Trials*, 13 J. APP. SOC. PSYCHOLOGY 351, 369 (1983) ("It is clear from the cognition and memory data that jurors in a joined trial situation cannot keep the two charges separate and arrive at independent verdicts."). Other studies have found evidence of jury confusion, but concluded that this confusion did not lead to biased verdicts. *See* James Farrin, Note, *Rethinking Criminal Joinder: An Analysis of the Empirical Research and Its Implications for Justice*, 52 L. & CONTEMP. PROB. 325, 328–29 (1989) (summarizing social science research on this issue). One other study found no evidence of confusion caused by multiple blocks of evidence, although the researchers acknowledged that this conclusion may be the result of giving the mock jurors relatively little evidence to remember. *See* Greene & Loftus, *supra* note 24, at 205–07 (finding little evidence of "memory effect" or "spillover phenomenon" when evidence involving two offenses was presented, in varying order, to mock jurors).

26. *See, e.g.*, Drew v. United States, 331 F.2d 85, 93–94 (D.C. Cir. 1964) (noting the danger of jury confusion over evidence on two distinct but similar crimes).

27. *See* Bean v. Calderon, 163 F.3d 1073, 1083–86 (9th Cir. 1998) (granting a habeas petition because, inter alia, the trial judge should have severed the counts). The court noted:

> [N]ot only did the trial court join counts for which the evidence was not cross-admissible, but the State repeatedly encouraged the jury to consider the two sets of charges in concert, as reflecting the modus operandi characteristic of Bean's criminal activities. Thus, the jury could not "reasonably [have been] expected to 'compartmentalize the evidence' so that evidence of one crime [did] not taint the jury's consideration of another crime."

Id. at 1084 (citation omitted).

28. The fact that a joint trial would admit evidence that would be inadmissible in separate trials can be grounds for severance, but it is not necessarily so. As long as the otherwise-inadmissible evidence is "simple and distinct," so that the jurors can keep straight in their minds which evidence relates to which alleged crime, joinder is permitted. *See, e.g.*, United States v. Alexander, 135 F.3d 470, 477–78 (7th Cir. 1998) (concluding that severance was not required where evidence on each of fourteen counts was relatively straightforward and the whole trial lasted only four days).

29. The leading case on this issue is *United States v. Cross*, 335 F.2d 897, 989 (D.C. Cir. 1964). *See also* United States v. Fenton, 367 F.3d 14, 22 (1st Cir. 2004) ("A defendant seeking a severance for the purpose of testifying on one of several counts must make a threshold showing that he has salient testimony to give anent one count and an articulable need to refrain from giving testimony on the other(s)." (citation omitted)).

the stand at trial can prevent the accused from putting on his best defense for each of the charges.[30]

In some cases, there can also be advantages to the defendant from having all the allegations against him resolved in one sitting.[31] For defendants who plead guilty, having the charges handled by a single prosecutor in a single package might facilitate a guilty plea that reflects a "volume discount." For those who plan to go to trial, a single proceeding can save money and trauma for the suspect and his family.[32] More importantly, a single trial means that the defendant may only have to cast doubt on the prosecutor's case once. Just as the jury may infer a criminal disposition from a long list of charges, it might also infer a weak government position on all charges if the accused can undermine the case on a critical point.

Thus, when considering the impact of joined charges, it is important to remember that sometimes the defendant will have agreed to the single proceeding (or would have agreed had he been asked).[33] But on balance, we should not be distracted by this fact. Although in theory the defendant can move to join cases that were originally brought separately,[34] the (admittedly impressionistic) sense of the reported cases is that it is prosecutors who want to consolidate charges and defendants who resist, not the other way around.[35] Nonetheless, the possibility of defense acquiescence to this type of joinder is a noteworthy qualification on the conclusions drawn below.

Joining defendants in a single trial presents similar risks but adds none of the advantages of joint charging. An accused who sits at the defense table with other suspects risks being found guilty by

30. The decision whether to testify is the most common strategic problem created by joinder, but it is not the only one. A defendant who wishes to represent himself on one charge but have appointed counsel for the other faces a similar choice. The Ninth Circuit recently considered and rejected a claim that having to choose between these two constitutionally-protected rights required severance. Cooks v. Newland, 395 F.3d 1077 (9th Cir. 2005).

31. For a discussion of some of the potential benefits, see 2 AMERICAN BAR ASSOCIATION, STANDARDS FOR CRIMINAL JUSTICE § 13-2.1, Commentary (2d ed. 1980).

32. *Id.*; *see also* Farrin, *supra* note 25, at 327 (noting that a defendant "may prefer the convenience of a single trial, may have his case disposed of more quickly, or may benefit from concurrent sentencing").

33. If the prosecutor brings all charges in a single indictment and defendant *prefers* this arrangement, the defense has no occasion to signal his acquiescence, making it hard to know how often this fact pattern occurs.

34. Federal Rule 13 provides that "the court may order" joinder of separate cases, without restriction on which party may ask the court to do so. FED. R. CRIM. P. 13.

35. *Cf.* 1A WRIGHT, *supra* note 9, § 221, at 472 (observing that while either party can move for severance, "[s]ince the government is usually in a controlling position with respect to the drafting of indictments and the joinder of defendants and offenses, there would seem ordinarily to be little point in permitting a motion by the prosecutor," although noting that some have been granted).

association, especially if his co-defendants have stronger evidence offered against them.[36] In a joint trial, the jury will be exposed to "spillover evidence," which might consist not only of information about the co-defendants' crimes (which is bad enough),[37] but also evidence of the defendant's own other bad acts, which now come to the jury's attention through the case against the co-defendants. In addition, joinder of multiple defendants introduces the risk of confusing jurors as to which evidence applies against which suspect, a risk that grows worse as the number of defendants increases.[38] Co-defendants may present antagonistic defenses[39] or make different decisions about whether to testify,[40] making distinctions among the defendants stark. Co-defendants who take the stand are often sorely tempted to point the fingers at others, confronting a defendant with an additional layer of accusation he would not have faced had he been tried separately.[41]

---

36. *See* Krulewitch v. United States, 336 U.S. 440, 454 (1949) (Jackson, J., concurring) (noting that in a joint conspiracy trial, "[t]here generally will be evidence of wrongdoing by somebody. It is difficult for the individual to make his own case stand on its own merits in the minds of jurors who are ready to believe that birds of a feather are flocked together.").

37. *See, e.g.*, United States v. Baker, 98 F.3d 330, 335 (8th Cir. 1996) (finding that defendant was prejudiced by spillover evidence of co-defendants' participation in a conspiracy and that although evidence was relatively straightforward, the risk to the defendant who was not part of the conspiracy was sufficiently high to require separate trials); *cf.* United States v. Johnson, 297 F.3d 845, 855–56 (9th Cir. 2002) (holding that "spillover evidence" of co-defendants' crimes was not "manifestly prejudicial" because it would have been admissible in a separate trial, "was inconsequential in the grand scheme of things," or was cured through limiting jury instructions).

38. *See* Zafiro v. United States, 506 U.S. 534, 539 (1993) ("[E]vidence of a codefendant's wrongdoing in some circumstances erroneously could lead a jury to conclude that a defendant was guilty. When many defendants are tried together in a complex case and they have markedly different degrees of culpability, this risk of prejudice is heightened.").

39. *See id.* at 538–39 (recognizing the risk of antagonistic defenses, but holding that they are not *per se* grounds for severance).

40. *See, e.g.*, United States v. Hodges, 502 F.2d 586, 587 (5th Cir. 1974) (holding that denial of severance was proper even though co-defendants took the stand and co-defense counsel commented on his clients' willingness to testify).

41. As the Ninth Circuit has observed:

> Defendants who accuse each other bring the effect of a second prosecutor into the case with respect to their codefendant. In order to zealously represent his client, each codefendant's counsel must do everything possible to convict the other defendant. The existence of this extra prosecutor is particularly troublesome because the defense counsel are not always held to the limitations and standards imposed on the government prosecutor.

United States v. Tootick, 952 F.2d 1078, 1082 (9th Cir. 1991). *But cf.* United States v. Flores, 362 F.3d 1030, 1039–40 (8th Cir. 2004) ("The mere fact that one defendant tries to shift blame to another defendant does not mandate separate trials . . . as a codefendant frequently attempts to 'point the finger,' to shift the blame, or to save himself at the expense of the other" (internal quotation marks and citations omitted)).

But unlike the joinder of charges, there is almost no advantage to the accused in being tried jointly.[42] The ability of the joined defendants to pool their resources and hire common counsel may provide a slight benefit, but even here, the risk of creating a problematic conflict of interest is great. In short, if given the choice between a trial with other defendants and being tried alone, it appears that a high percentage of suspects would choose the latter.

C. The Legal Response and Remedy

An accused who wants to avoid the risks of joinder has two possible lines of attack. He can initially argue to the trial court that the joinder fails to satisfy the requirements of Rule 8; if he is correct, this constitutes "misjoinder" and the judge is required to sever.[43] But even if the initial joinder was proper, a defendant can ask the court for severance. Federal Rule 14 provides that if joinder "appears to prejudice a defendant or the government, the court may order separate trials of counts, sever the defendants' trials, or provide any other relief that justice requires."[44] The text makes clear that even if a defendant is harmed by the consolidation, severance is not the only, or even the preferred, option. Probably the most common species of "other relief" is curative jury instructions – an admonishment from the court to make sure that there is sufficient evidence for each separate count, and a warning not to confuse or attribute the evidence on one charge or against one defendant to any other.[45]

---

42. *See* Dawson, *supra* note 2, at 1388–89 (noting and then discounting the argument that there are advantages to defendants when they are tried with others). *But see infra* Part IV.C.2.a (suggesting an interpretation of the data that would make it advantageous for certain defendants to be joined at trial with others).

43. If the trial court concludes that joinder is proper and defendant appeals the issue after conviction, his claim will be subject to harmless error review. *See, e.g.*, United States v. Macking, 315 F.3d 399, 412 (4th Cir. 2003) (stating that such error raises no constitutional problems and is reviewed for harmlessness); *cf.* 1A WRIGHT, *supra* note 9, § 221 at 465 ("If the original joinder was improper, Rule 14 confers no discretion. The trial court must order severance, and if it fails to do so, reversal is likely, although in many cases the error may be regarded as harmless.").

44. FED. R. CRIM. P. 14(a).

45. For example, the pattern jury instruction in the Seventh Circuit provides:

> Even though the defendants are being tried together, you must give each of them separate consideration. In doing this, you must analyze what the evidence shows about each defendant[, leaving out of consideration any evidence that was admitted solely against some other defendant or defendants]. Each defendant is entitled to have his/her case decided on the evidence and the law that applies to that defendant.

See PATTERN CRIMINAL FEDERAL JURY INSTRUCTIONS FOR THE SEVENTH CIRCUIT, INSTRUCTION 4.05, *available at* http://www.ca7.uscourts.gov/Rules/pjury.pdf. The Drafting Committee Comment then adds: "In cases involving more than one count, it will be necessary for the court to instruct both as to the separate consideration for each defendant and also with regard to separate consideration of charges." *Id.*

Not surprisingly, courts have adopted a balancing test in applying Rule 14, weighing the savings that come from a single trial against the possible harm to the defense.[46] Sometimes courts find the balancing quite easy. For example, virtually all courts agree that if the evidence in a joint trial would still be admissible against the defendant in a separate trial, the accused gains nothing from severance and the interest in efficiency prevails.[47]

In other cases the balancing is (or should be) harder, but the outcome is usually the same. Defendants are entitled to relief if it "appears" they will be harmed, yet most courts admit that some degree of prejudice from joinder may be inevitable and, in any event, is an acceptable price to pay for the efficiency gains. One of the best known articulations of this view came a few years before the enactment of the Federal Rules of Criminal Procedure, when the Second Circuit said, "Congress has authorized consolidation [of trials] in the belief that public considerations of economy and speed outweigh possible unfairness to the accused."[48]

In the intervening decades, courts have echoed this view with varying degrees of vigor. The D.C. Circuit has noted that whether it is joinder of charges or joinder of defendants, "some prejudice almost necessarily results," but that "Rule 8(a) permits the first sort of prejudice and Rule 8(b) the second."[49] The Second Circuit has agreed that the dangers of prejudice from joinder are "inevitabl[e]"[50] and that Rule 8(a) "necessarily recognizes the adverse effect on the defendant by a joinder of counts, but considers this to be outweighed by gains in trial economy when one of the criteria of the rule are [sic] met."[51]

---

46. *E.g.*, United States v. Bledsoe, 674 F.2d 647, 655 (8th Cir. 1982) ("[I]n evaluating joinder . . . the trial court must balance its obligation to avoid prejudice that may result from joining multiple defendants against a policy favoring maximum trial efficiency.").

47. *See, e.g.*, United States v. Soto-Beníquez, 356 F.3d 1, 29–30 (1st Cir. 2003); United States v. Hart, 273 F.3d 363, 370 (3d Cir. 2001) (explaining that severance is very difficult to obtain in conspiracy cases because "acts committed by one [co-conspirator] in furtherance of the conspiracy [are] admissible against the other" co-conspirator).

48. United States v. Smith, 112 F.2d 83, 85 (2d Cir. 1940). More recent cases have quoted or cited the *Smith* language with approval. *E.g.*, United States v. McClellan, No. 93-4084, 1994 WL 589497, at *3 (6th Cir. Oct. 25, 1994) (quoting *Smith*); United States v. Turoff, 853 F.2d 1037, 1043 (2d Cir. 1988) (same).

49. Cupo v. United States, 359 F.2d 990, 993 (D.C. Cir. 1966).

50. *Turoff*, 853 F.2d at 1043.

51. United States v. Werner, 620 F.2d 922, 929 (2d Cir. 1980); *see also* United States v. Flores, 362 F.3d 1030, 1039 (8th Cir. 2004) ("Once defendants are properly joined under Rule 8, there is a strong presumption for their joint trial."); United States v. Winter, 663 F.2d 1120, 1139 (1st Cir. 1981) ("A joinder of offenses, or of defendants, involves a presumptive possibility of prejudice to the defendant." (quoting King v. United States, 355 F.2d 700, 703 (1st Cir. 1966))); United States v. Satterfield, 548 F.2d 1341, 1347 (9th Cir. 1977) ("We recognize that even permissible joinder will often result in some prejudice to a defendant. The purpose of rule 8(b) is

Given that some harm is inevitable, courts have naturally concluded that cases should be severed only if the defendant can show that the prejudice is "severe" or "compelling,"[52] even though these qualifications do not appear in Rule 14. Or as one court put it, "Since Rule 8 anticipates some prejudice, severance under Rule 14 is limited to cases of 'substantial prejudice.'"[53]

This tolerance of harm to the accused makes courts disinclined to find misjoinder[54] and miserly in granting severance,[55] a practice that has been sharply criticized. Charles Alan Wright put it best when he said it "seems strange that one presumably innocent may be made to undergo something less than a fair trial, or that he may be prejudiced in his defense if the prejudice is not 'substantial,' merely to serve the convenience of the prosecution and the efficiency of the judicial system."[56]

It does indeed seem strange that courts are so willing to tolerate this apparent unfairness, strange enough that we should dig deeper. The best potential explanation for this casual judicial attitude would be that, despite their language, judges have abiding faith in

---

to limit joinder to those cases where considerations of trial efficiency clearly outweigh a defendant's interest in a separate trial."); United States v. Orefice, No. 98 CR. 1295(DLC), 1999 WL 349701, at *5 (S.D.N.Y. May 27, 1999) ("The policy behind Rule 8(a) is that considerations of economy and speed outweigh possible unfairness to the defendant so as to justify joinder.").

52. United States v. Liveoak, 377 F.3d 859, 864 (8th Cir. 2004). The court also states that "[t]o grant a motion for severance, the necessary prejudice must be severe or compelling . . . ." *Id.* at 693 (internal quotation marks and citation omitted). *Cf.* 1A WRIGHT, *supra* note 9, § 223 at 489–90 ("The burden is put on the defendants to make a strong showing of prejudice in order to obtain the relief permitted by Rule 14. With such a test, it is hardly surprising that in most cases relief has been denied.").

53. United States v. Alverado, No. 92 Cr. 728 (LMM), 1994 WL 669968, at *3 (S.D.N.Y. Nov. 30, 1994); *see also* United States v. Sarracino, 340 F.3d 1148, 1165 (10th Cir. 2003) ("[S]everance is not required even if prejudice is shown; 'tailoring of the relief to be granted, if any, [is left] to the district court's sound discretion.'" (citation omitted)). For an additional discussion of this point, see Andrew D. Leipold, *How the Pretrial Process Contributes to Wrongful Convictions*, 42 AM. CR. L. REV. 1123 (2005) (discussing how joinder and severance rules have the potential to "expose[] the jury to evidence that is unfairly prejudicial").

54. *See, e.g.*, *Liveoak*, 377 F.3d at 864 (stating that joinder requirements should be liberally construed); United States v. Johnson, 130 F.3d 1420, 1427 (10th Cir. 1997) ("Although alleged misjoinder under Rule 8 is a question of law subject to de novo review, we construe Rule 8 broadly to allow liberal joinder to enhance the efficiency of the judicial system.").

55. *See* 4 LAFAVE ET AL., *supra* note 4, § 17.1(f) at 602 ("[D]efendants generally have not fared very well under rules and statutes which permit them to obtain a severance of offenses only upon proof of prejudice."); 1A WRIGHT, *supra* note 9, at 4–5 (noting the "restrictive and unrealistic construction courts have put on Rule 14," and the "reluctance of trial and appellate courts to grant separate trials").

56. 1A WRIGHT, *supra* note 9, § 141, at 6; *see also* Dawson, *supra* note 2, at 1381 ("[C]ourts have greatly exaggerated the supposed efficiencies of joint trials while grossly underestimating the impediments joint trials pose to fair and accurate determinations of individual guilt or innocence.").

their own ability to spot the cases of genuine prejudice and to take corrective action. The risks of joinder are well known, and if judges truly can spot the riskiest cases and take remedial steps, perhaps the prejudice is more hypothetical than real.

What is missing from the discussion is an empirical measure of the impact of joinder on criminal cases. The next Part provides a first attempt to quantify what we will call the "joinder effect."

III. THE PREJUDICIAL IMPACT OF JOINDER

A. The Threshold Empirical Case

The first step in evaluating the impact of joinder is to ask some basic questions. How often do defendants face multiple charges rather than a single count? How often are they joined with other defendants, and how often are both conditions (multiple defendants and multiple counts) present? We will refer to these various combinations of joined counts and joined co-defendants as the accused's "joinder status."

Second, we will ask if the outcomes of the cases differ depending on the defendant's joinder status. Here we will examine outcomes in two ways: (i) for those defendants who go to trial, the likelihood that they will be convicted; and then secondarily, (ii) the likelihood that a defendant will plead guilty rather than go to trial.

Our working hypothesis will track the assumptions that courts and commentators make about consolidation. Namely:

- defendants who are charged with multiple counts are more likely to be convicted at trial than defendants who are charged with a single count;
- defendants who stand trial with other defendants are more likely to be convicted than those who stand trial alone; and
- defendants who face multiple charges and stand trial with multiple defendants are more likely to be convicted than any other group.

To these three we can add a fourth, conditional prediction. *If* those who face multiple counts and/or who are joined for trial with others are more likely to be convicted, we would expect to see defendants in those groups pleading guilty more often than those who are not similarly joined. We assume that a defendant who faces a reduced chance of winning at trial is less likely to risk a trial than

others and is more willing to take whatever benefits are offered in return for a plea.[57]

To answer these questions and test these hypotheses, we calculated original statistics from relatively raw government data.[58] We began by looking at the records of over 375,000 formally charged federal defendants whose cases terminated during fiscal years 1999 through 2003.[59]

1. The Frequency of Joinder

Over the five-year period of this study, nearly half of the defendants were charged with a single count, the other half with more than one count. Of those charged with multiple counts, roughly half again (26% of all defendants) were charged with two counts, and a generally diminishing percentage were charged with an increasing number of offenses, as shown in Table A.

---

57. We assume that a higher percentage of guilty pleas based on joinder status means that these defendants collectively are worse off, even though in individual cases they may not be. Often a suspect will plead guilty in exchange for a concession from the prosecutor, and in cases where the evidence of guilt is strong, the defendant may be in a much better position than if he had gone to trial. However, if certain *groups* of defendants face a higher chance of conviction at trial than others, their bargaining power is surely reduced, putting them in a relatively worse position than they would he in the absence of joinder.

58. Unless otherwise indicated, the numbers and statistics presented in the balance of this paper are original calculations derived from the records compiled by the Federal Judicial Center, tbe research and education agency of the federal judicial system. *See generally* Federal Judicial Center Home Page, http://www.fjc.gov/ (last visited Mar. 24, 2006) (providing access to FJC puhlications, information on federal judicial history, and links to educational programs and materials). The data are made available through the Inter-University Consortium for Political and Social Research (ICPSR), located in Ann Arbor, Michigan. *See* Inter-University Consortium for Political and Social Research (ICPSR) Front Page, http://www.icpsr.umich.edu (last visited Mar. 24, 2006) (providing the ability to search for specific data holdings from its front page). The 1999 and 2000 data sets are found in Study Number ICPSR 8429, Datasets (DS) 112 and 113. The 2001 database is found at Study Numher ICPSR 3415, DS1. The 2002 data are found at Study Number 4059, DS3; and the 2003 data at Study Number 4026, DS5. For each of the datasets, one record equals one defendant, which is the unit of measurement used in this paper.

59. To avoid double counting across the years, we only used cases for which there had been a final resolution during fiscal years 1999–2003 (i.e., the variable STATUSCD = L). The number of defendants whose cases met this description was 393,225. Of this number, we eliminated 17,505 cases (4.5% of the database) where the disposition of the most serious count (variable DISP1) was "final" for court reporting purposes, but was not necessarily a final resolution of the criminal charges against that defendant. Thus, we did not include 11,885 defendants for whom the most serious charge against them was dismissed without prejudice. In addition, we eliminated cases where the charges resulted in a pre-trial diversion (1,926 defendants); where the case was transferred to another judicial district for disposition (3,531 defendants); cases that ended in a mistrial (132 defendants); and cases that resulted in some other non-final resolution (31 defendants). As modified, the final dataset contained 375,720 records.

TABLE A

Joinder of Federal Criminal Charges 1999-2003		
Number of Counts	**Number of Defendants**	**Percent of Defendants**
1	183,540	49%
2	99,321	26%
3	38,396	10%
4	22,759	6%
5 or more	31,704	8%
Total	**375,720**	**100%**

During the same five-year period, nearly two-thirds (65%) of federal defendants were charged alone, while the remaining one-third were joined with others, either in the original indictment or via Rule 13. Many of the joined defendants were charged with only one co-defendant, but a non-trivial number (14%) were charged with four or more others, as shown in Table B.[60]

TABLE B

Joinder of Federal Criminal Defendants 1999-2003		
No. Defendants in Case	**No. of Defendants**	**Percent of Defendants**
1	245,575	65%
2	43,089	11%
3	21,754	6%
4	13,017	4%
5 or more	52,285	14%
Total	**375,720**	**100%**

These two categories overlap. If we simplify our categories to cases where defendants: (a) face one charge or more than one charge; and (b) stand trial alone or with at least one other co-defendant, we can see the relationships in a simple matrix.

60. Tables A and B (as well some of the other Tables and Figures *infra*) provide data for a category "5 or more" charges or defendants. Although both categories have the same label, there is a slight difference between them. While some number of defendants undoubtedly faced more than five charges, the datasets only collect information on the five most serious counts; they give no indication how many defendants faced more than five counts or of the disposition of those later counts. In contrast, the dataset allows us to determine precisely how many defendants were joined in a particular case, and for administrative convenience, we created the category "5 or more." We do not believe this categorization has any effect on the statistics or the conclusions.

TABLE C

Frequency of Joinder, 1999-2003			
	One Defendant	Multiple Defendants	Total
One Count	40% (n = 150,276)	9% (n = 33,264)	49% (n = 183,540)
Multiple Counts	25% (n = 95,299)	26% (n = 96,881)	51% (n =192,180)
Total	65% (n = 245,575)	35% (n = 130,145)	100% (n = 375,720)

These figures confirm that consolidation is quite common. Thus, if even a small amount of prejudice attaches to joinder, it will affect a substantial number of defendants.

Importantly, these large numbers do not tell us the status of defendants who actually went to trial. Notice that in the discussion of prejudice in Part II(B), nearly all of the presumed harms occur at the trial itself. So while even the expectation of prejudice at trial may influence a decision to plead guilty (more on this later[61]), it is important to know how often the risk of prejudice is presented, a figure that requires us to measure how often joinder occurs among *trial* defendants.

Here, we need to define the terms carefully because, for our purposes, joinder among trial defendants is different than joinder in general. In the previous three tables, we measured the extent to which a defendant was *charged* with multiple crimes or with other defendants—the tables told us how popular joinder is. But now we are interested in how often a defendant is faced with multiple counts or sits at counsel table next to co-defendants at the trial itself. Here we do not care if a defendant was originally charged with two counts if he pleads guilty to one before trial; in our calculations, he will be treated as having faced only a single count. Likewise, even if Defendant A was originally indicted along with three others, we will not count him as being joined with other defendants for trial if the co-defendants B, C, and D all pled guilty, had their cases dismissed, or otherwise washed out of the system.

61. *See infra* Part IV.C.2 (discussing whether joinder of defendants causes prejudice).

Only 5 percent of federal criminal defendants go to trial, so naturally this group is much smaller.[62] Among trial defendants, the joinder percentages are different than the "all defendants" group described above, particularly for multiple defendants.

TABLE D

Trial Defendants, 1999-2003					
Joinder of Charges			Joinder of Defendants		
No. Counts	No. Trial Ds	% Trial Ds	No. Ds	No. Trial Ds	% Trial Ds
1	10,325	52%	1	15,146	76%
2	4,614	23%	2	2,350	12%
3	1,957	10%	3	1,068	5%
4	1,115	6%	4	448	2%
5 or more	1,797	9%	5 or more	796	4%
Total	**19,808**	**100%**	**Total**	**19,808**	**100%**

Finally, we should ask how often our various joinder status combinations occur. Here we can return to the basic matrix used above:

TABLE E

Trial Defendants, Frequency of Joinder			
	One Defendant	Multiple Defendants	Total
One Count	45% (n = 8,986)	7% (n = 1,339)	52% (n = 10,325)
Multiple Counts	31% (n = 6,160)	17% (n = 3,323)	48% (n = 9,483)
Total	76% (n = 15,146)	24% (n = 4,662)	100% (n = 19,808)

These figures allow us to make general observations about what trials look like. Nearly half of all trial defendants (45%) stand trial alone on a single count. If our working hypotheses are correct, these are the most-favored defendants because they face no prejudice from joinder. Conversely, roughly one in six defendants (17%) face both varieties of prejudice by standing trial with co-defendants on multiple counts. The remaining suspects (31% + 7% = 38%) have some

62. Between 1999 and 2003, there were 375,720 federal criminal defendants and 19,808 defendants who stood trial on at least one count (19,808 ÷ 375,720 = 0.0527 = 5.3%).

combination of joinder and non-joinder, which, assuming for the moment the existence of some prejudicial impact, should allow us to isolate and identify the type of joint trial (multi-count or multi-defendant) that has the greater effect.

2. Case Outcomes

Looking at all federal criminal defendants who stood trial over a five-year period, we see that on average, slightly more than three-quarters (77%) were convicted while the rest (23%) were acquitted.[63] The picture is brighter, however, for those in our hypothesized most-favored group—the single defendant who stands trial on a single count. Here only 67% of the defendants are convicted, 10 percentage points lower than the overall average.[64]

We can make an early test of our working hypotheses by comparing this group to those who face one type of joinder but not both. But before we do so, we should sidestep an obvious trap. It would hardly be surprising to find that a defendant charged with four crimes is more likely to be convicted of *something* than a defendant charged with only one offense; in theory, the prosecutor's chances of getting at least one charge to stick should increase with each joined offense. So to make sure we are comparing apples to apples, we will measure how often the defendant was convicted of the *most serious* crime charged (designated "Count 1" in the datasets). By limiting our attention to the conviction rates on the "top count," we can better isolate the incremental effect of additional charges.

And in fact, the conviction rate does rise when joinder enters the picture. For defendants who face multiple counts in a single trial, the conviction rate on the most serious charge jumps dramatically, from 68% for a single-count defendant to 82% when a second count is added. The conviction rate for Count 1 then rises again, to 88%, when Count 3 is added, and then remains steady as other charges are joined. The conviction percentages are shown in Figure 1:[65]

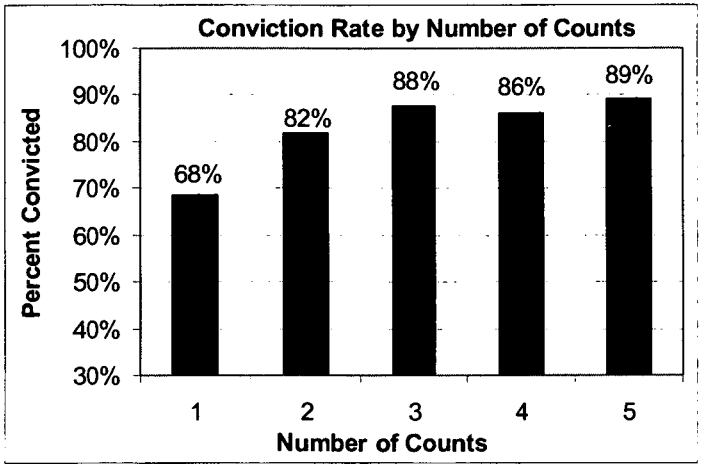
---

63. Of the 19,808 trial defendants identified in Table D, 19,057 went to judgment on the most serious charge. Of these, 14,589 were convicted (14,589 ÷19,057 = .7655) and 4,468 were acquitted (4,468 ÷19,057 = .2345). In addition, there were 751 defendants who either pled guilty to the most serious charge or had the top count dismissed, yet nevertheless went to trial on some of the lesser charges. For consistency, we excluded this small group of defendants and only included those who went to trial on the most serious count in the calculations.

64. *See infra* note 67 and Table E for the complete statistics of conviction rates hy joinder status.

65. *See supra* Table D for the number of defendants per count category.

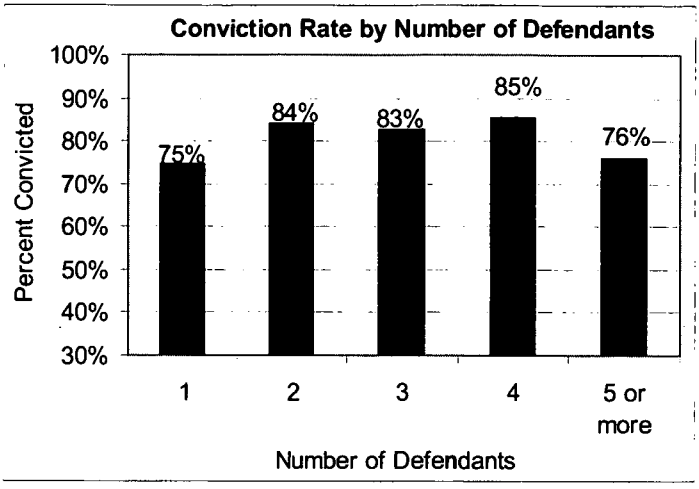
FIGURE 1



For those who faced the other joinder prong—standing trial with co-defendants—the conviction rate also jumps, although not as much. The addition of a second defendant increased the chances of conviction on Count 1, from 75% when the defendant stands trial alone to 84% when a second defendant is added. The addition of a third and fourth defendant has much less of an impact, with the conviction percentage on Count 1 staying in the mid-80% range, and then dropping back to 76% when trials get really large, involving five or more co-defendants.[66] The conviction rates are shown in Figure 2.

66. *See supra* Table D for the number in each category of joined defendants.

FIGURE 2



Having isolated the impact of adding counts and adding defendants, we now can combine the two types of joinder to measure the impact of joinder status on the trial conviction rate. Table F shows the rates for our four groups.[67]

TABLE F

Conviction Rate by Joinder Status			
	One Defendant	Multiple Defendants	Total
One Count	67%	77%	**68%**
Multiple Counts	85%	85%	**85%**
Total	**75%**	**83%**	**77%**

This Table is largely consistent with our working hypothesis and provides us with a prima facie case for the prejudicial impact of joinder. Table F does, however, reveal one surprise. Joining additional charges increases the conviction rate, as does joining

67. There were 8,365 defendants who stood trial alone on a single count, 5,605 (67%) of whom were convicted. There were 6,068 single defendants who faced multiple counts, 5,167 of whom were convicted (85%). For those who stood trial with co-defendants, there were 1,317 who faced a single count, 1,016 of whom were convicted (77%), and 3,307 who faced multiple counts, of whom 2,891 were convicted (85%).

additional parties with a single-count defendant. As between individual defendants facing multiple counts and joint defendants facing multiple counts, however, the conviction rate does not change. This suggests that when a defendant faces multiple counts, joining or not joining other defendants makes no real difference.

This last point is sufficiently important to warrant a closer look. Table G uses a Probit Model[68] to isolate the effects of the types of joinder. In the first model, the only explanatory variable is "counts" (and a constant). It shows the positive effect of counts on the probability of conviction. The same result is shown in Model 2 for number of defendants. But considering *both* counts and defendants (Model 3) shows that only the number of counts has significant effect on the probability of conviction. The comparison of Model 3 with Model 1 shows that the number of defendants in fact has no significant effect on the conviction rate.[69]

TABLE G

Probit Model of guilty/not guilty at trial on the most serious charge (dependent variable: Guilty(1))			
Variable	**Model 1**	**Model 2**	**Model 3**
Counts	0.558 (0.0207)* [0.000]**	-	0.533 (0.0214) [0.000]
Number of defendants	-	0.273 (0.0245) [0.000]	0.111 (0.0256) [0.000]
Constant	0.478 (0.0133) [0.000]	0.663 (0.0113) [0.000]	0.464 (0.0137) [0.000]
Number of observations	19057	19057	19057
Pseudo R-Squared	0.0360	0.0062	0.0369
Log likelihood	-10005.1	-10314.5	-9995.6

* Standard errors in parentheses

** Numbers in brackets show the probability of coefficient being outside the 95% confidence interval.

(1) It is equal to one if the defendant is convicted, zero otherwise.

68. In econometrics, whenever the dependent variable is discrete, regular models such as Ordinary Least Square (OLS) cannot be used. For binary dependent variables (which, in our case, are set to either "conviction" or "acquittal"), the Probit Model is the widely used alternative. Using the Probit Model, we estimate the parameters of the model (the coefficients of the explanatory variables) assuming the distribution of error term to be standard normal. The final result (the probability of conviction for each of the explanatory variables) is achieved by calculating the cumulative standard normal distribution function for that point using the estimated coefficients. For a complete explanation of these models, see G.S. MADDALA, LIMITED-DEPENDENT AND QUALITATIVE VARIABLES IN ECONOMETRICS 22–27.

69. Note that in this comparison the log likelihood and the Pseudo R-Squared (which measure the reliability of the model) do not change much when the new variable is added.

This model allows us to draw a modestly significant conclusion. If it really is the addition of counts that matters, then we should not worry about those types of prejudice that are specific to the joinder of defendants. For example, the fear that defendants will be harmed by being forced to sit at counsel table with other alleged criminals is not problematic, nor, apparently, is it especially harmful for the jury to hear information about co-defendants' misdeeds, at least to the extent the evidence is limited to the co-defendants' conduct.[70] In short, the risk of guilt by association seems overblown.[71]

What does seem prejudicial is the jury being exposed to other *evidence* that implicates the accused, regardless whether it comes in directly, as proof against the defendant, or through evidence against co-defendants that happens to include the defendant. Just as with multiple count cases, multi-defendant trials can expose the jury to the defendant's other bad conduct, and perhaps it is this spillover that causes the rise in the conviction rates. This explanation is not perfect, of course; if joining other parties lets incriminating information leak in against the defendant, it is unclear why we see a prejudicial effect when defendants face a single count but not when they face multiple charges.[72] Nevertheless, the statistical evidence indicates that, contrary to our working hypothesis, the prejudicial impact seems directly tied to the evidence, not to the appearances of a multi-defendant trial.[73]

The impact of joinder on trial defendants is the primary topic of this study but not the only one. An overwhelming majority of suspects will plead guilty rather than stand trial, and a large percentage of this group will also face multiple counts or be joined with others.[74] Since

---

70. *See supra* note 37 and accompanying text (discussing whether "spillover" evidence has a prejudicial effect).

71. Interestingly, nearly all of the social science experiments performed to date measure the impact of joinder by testing the impact of joining counts; none were discovered that measure the impact of joinder based on the effects of joining defendants. *See* Farrin, *supra* note 25, at 325–26 (finding, as of 1997, that "[a] large amount of empirical data exists for joinder of offenses, while no empirical research has been done directly on joinder of defendants").

72. Recall that Table F, *supra*, shows that defendants charged with one count face a significant increase in conviction rate as co-defendants are added, but those facing multiple counts do not. If the prejudicial effect flows from the evidence introduced against co-defendants, we would expect to see a similar conviction rate regardless of the number of charges filed against the defendant. The explanation may be that defendants facing multiple charges already suffer the prejudice associated with overlapping evidence, thus the similar evidence introduced in a co-defendant's case has no effect on the jury. Under this view, evidence that indirectly implicates the defendant has an effect only when the jury is not confronted with other, more direct evidence of the defendant's wrongdoing.

73. For a potentially important qualification on this conclusion and a counter-hypothesis that might explain the data, see *infra* Part IV.C.2.

74. *See supra* Table C and accompanying text.

the decision to plead is typically made with an eye toward the expected outcome of a trial,[75] we would expect to see the presence of joined charges or co-defendants influence the decision of whether to admit guilt. By hypothesis, if all other factors are equal, a defendant's willingness to plead guilty should increase as the odds of conviction at trial increase. We therefore would predict an increase in the guilty-plea rates that roughly parallels the conviction rates shown in Table F.

In calculating the guilty-plea rates, we ignored the cases where the defendant was successful in having the charges dismissed,[76] and those where the disposition was something other than trial or a guilty plea.[77] For the remaining defendants, the guilty-plea rates correlate to the joinder status as follows:[78]

TABLE H

Guilty-Plea Rate by Joinder Status		
	One Defendant	Multiple Defendants
One Count	88%	86%
Multiple Counts	91%	87%

The inference drawn from Table H is that joinder has virtually no impact on the decision to plead guilty, with our most-favored group (one count, one defendant) being just as likely to plead as our presumptive least-favored group (multiple counts, multiple defendants). Perhaps this means that we are looking at the wrong

75. Even this seemingly simple observation has its critics. For a useful discussion of the complexity of the plea decision, see Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 HARV. L. REV. 2463 (2004); *see also* William J. Stuntz, *Plea Bargaining and Criminal Law's Disappearing Shadow*, 117 HARV. L. REV. 2548 (2004) (discussing Bibas's article).

76. Roughly 7 percent of all defendants considered in this study had their cases dismissed (25,186 dismissed defendants / 375,720 total defendants = 0.067).

77. During the period of the study, there were 11,885 defendants whose cases were dismissed "without prejudice." Although the cases were "final" for court-reporting purposes, prosecutors may well have refiled many of these charges. Therefore, unless otherwise indicated, these dismissals are not considered in the calculations. *See also supra* note 59.

78. Over the five years of this study, the number of defendants in each category and the number of guilty pleas are as follows. There were 150,276 single defendants charged with a single count, 132,299 of whom (88%) pled guilty. There were 95,299 single defendants charged with multiple counts, 86,544 of whom (91%) pled guilty. In the multiple-defendant cases, there were 33,264 who were charged with one count, 28,721 of whom (86%) pled guilty, and 96,881 who were charged with multiple counts, 83,913 of whom (87%) pled guilty.

variable, an issue that will be explored in the next Section. But for current purposes, the lack of correlation between guilty pleas and joinder is worth remembering. If it turns out that joinder does in fact influence trial outcomes, we should then return to the issue of guilty pleas and ask *why* the potential impact on the trial appears to play little or no role in this critical decision.

B. Controlling for Other Variables

The phrase "all things being equal" carried a lot of baggage in the last Section; now it is time to unpack it. Somewhat (but not entirely) consistent with our working hypotheses, we have seen that there is a large gap in conviction rates based on joinder, with the addition of counts appearing to be very important and the addition of defendants being much less so. We should now test these conclusions to see if they stand up when we control for other variables.

Our datasets allow us to generate several alternative theories for why certain trial defendants are more likely to be convicted than others, independent of the joinder effect. Four potential alternative explanations are: (1) the conviction rate is affected by whether the factfinder is a judge or the jury; (2) the conviction rate is affected by the type of crimes charged, both the nature of the offense and its seriousness; (3) the conviction rate is affected by the type of lawyer representing the defendant, that is, whether counsel is appointed or retained; and (4) the conviction rate is influenced by regional characteristics, and thus differs from one part of the country to another.

In considering each of these theories, we will first ask if there are significant differences in the frequency of convictions that correlate to the elements of the variable in question; we then will test more rigorously to see if the effect of joinder remains substantial and robust despite any correlations. As we consider each variable, we will add its impact to our model, so that by the end of the Section, we will have a cumulative measure of the joinder effect, one that controls for all of the tested variables.

1. Judge vs. Jury

One explanation is that the difference in conviction rates is explained primarily by the identity of the factfinder. On average, juries are more likely to convict than judges,[79] and perhaps the

79. For an extensive discussion of this point, see Andrew D. Leipold, *Why Are Federal Judges So Acquittal Prone?*, 83 WASH. U. L. Q. 151 (2005).

conviction patterns map onto the defendant's choice of decisionmaker.[80] A more refined theory posits that perhaps defendants are more likely to take multi-count, multi-defendant cases to a jury, leaving the simpler cases for the more expeditious bench trials.[81] The judges' lower conviction rates thus may correlate with joinder, without being a cause or effect of it.

It turns out that the identity of the factfinder does play a large role in conviction rates. Among defendants who go to trial, those tried by a jury were substantially more likely to be convicted, regardless of their joinder status.

TABLE I

Conviction Rate by Factfinder	
	Conviction Rate
Jury Trial (n = 13,835)	85%
Bench Trial (n = 5,222)	55%

This disparity is great enough that we should look at this variable more closely. In particular, we should look at bench trials and jury trials separately, to see if the joinder effect is still significant within each subgroup of cases.

Continuing with the Probit Model developed in Table G, we can control for the identity of the factfinder.[82]

80. To say that defendants have a "choice" of factfinder glosses over a potential technical problem. Although defendants have an absolute right to a jury in felony and serious misdemeanor cases, they do not have the absolute right to choose a bench trial. Rule 23(a) of the Federal Rules of Criminal Procedure requires agreement of the prosecutor and court for the defense to waive a jury, and thus it is possible that defendants' preference for a bench trial is vetoed in some cases. But as one of us (Leipold) noted in an earlier article, after conducting more than two dozen interviews with prosecutors and defense counsel about this matter, absolutely no evidence was found that prosecutors veto a defense preference for a bench trial in more than a tiny fraction of cases. It seemed obvious that defendants' choice of factfinder is largely unchecked by the government or tbe court. *See* Leipold, *supra* note 79, at 167–68 (noting that when defense council seeks to waive jury trial prosecutors rarely veto this choice).

81. Some support for this view can be found in the very high percentage of misdemeanor cases that are tried by the court rather than a jury, even when the right to a jury is present. *See infra* note 88 and accompanying text (observing that misdemeanors and pretty crimes are overwhelmingly tried by bench trials).

82. The correlation coefficient between the variables "guilty" and "factfinder" (a binomial variable equal to one if the factfinder is jury, zero otherwise) is 0.3070, which shows a considerable positive relationship between two variables. This requires us to look closely at the effect of the factfinder on our model.

TABLE J

Variable	Model 4
Counts	0.297
	(0.0232)*
	[0.000]**
Number of defendants	-0.048
	(0.0268)
	[0.069]
Factfinder	0.775
(Judge/Jury)	(0.0239)
	[0.000]
Constant	0.086
	(0.0179)
	[0.000]
Number of observations	19057
Pseudo R-Squared	0.0883
Log likelihood	-9462.4

* Standard errors in parentheses

** Numbers in brackets show the probability of coefficient being outside the 95% confidence interval.

(1) It is equal to one if the defendant is convicted, zero otberwise.

Table J shows that adding the new variable makes the coefficient of the variable "number of defendants" turn negative and significantly different from zero at a 90% confidence interval, but not at 95%, indicating that this variable is not very important in explaining the conviction rates. Adding the factfinder variable also decreases the effect of "counts" on the conviction rate, but the sign remains positive, and the significance is strong enough that it confirms the effect of "count" on conviction rates.

These figures allow us to calculate a revised joinder effect, one that controls for the judge / jury distinction. First, the conviction rates for each factfinder:

TABLE K

Jury Conviction Rates by Joinder Status		
	One Defendant	Multiple Defendants
One Count	81%	79%
Multiple Counts	88%	87%

TABLE L

Bench Conviction Rates by Joinder Status		
	One Defendant	Multiple Defendants
One Count	53%	52%
Multiple Counts	65%	63%

Here we see in more concrete form what the earlier models had predicted. Whether a defendant is tried by a jury or by the bench, adding counts has a significant effect on the conviction rate, while adding defendants without changing the number of counts has virtually no impact.

Although consideration of the factfinder nicely supports our working hypothesis, there are some mildly surprising figures in Tables K and L. In particular, we would have predicted a more noticeable joinder effect in jury trials than in bench trials. It is an article of legal faith that judges are better than juries at ignoring irrelevant and prejudicial evidence, and there is some feeling that they are better at keeping complicated information straight. Yet the Tables suggest otherwise: adding counts produces a ten or eleven percentage point increase in conviction rates for bench trials, while additional counts add only seven or eight percentage points to conviction rates in the case of jury trials. This effect becomes less surprising, however, when we consider how infrequently bench trials occur: more than 70% of all trial defendants are tried by a jury.[83] The relatively small number of bench trials quite likely contributes to the larger variations in conviction rates.

Because we are measuring joinder effects across all types of cases, we will conclude this sub-section by re-combining bench and jury trials. We do so, however, using a weighted average that accounts for the heavy proportion of jury cases.[84] Having thus controlled for the impact of the factfinder on the conviction rates, we

83. Of the 19,057 trial defendants studied, 13,835 (72.6%) had their cases decided by a jury. *See supra* Table I (showing that only 5,222 of the defendants studied had bench trials, whereas 13,835 had their cases tried by a jury).

84. Instead of putting one and zero (for bench and jury trial) and getting two separate tables for two groups, we used the weighted average of factfinder variable to get a single table that shows the probability of conviction for different joinder types after considering the effect of factfinder. The weighted average is (13835/19057)*1 + (5222/19057)*0 = 0.72598.

can now generate a new Table, one that shows a diminished but still substantial joinder effect:

TABLE M

Trial Conviction Rates by Joinder Status (controlling for identity of the factfinder)			
	One Defendant	Multiple Defendants	Difference (in percentage points)
One Count	74%	73%	-1%
Multiple Counts	83%	82%	-1%
Difference (in percentage points)	9%	9%	—

We will add the judge / jury variable to our model as we test the impact of other factors. In a small gesture of mercy to the reader, only the conclusions and simple frequency tables are set forth for the remaining variables; the balance of the statistical models are set forth in the Appendix.

2. Type of Crime

Not all types of crimes end in conviction at the same rate. Weapons cases may be harder or easier for the prosecutor to win than immigration cases; juries may be more harsh or lenient when the defendant is charged with a misdemeanor instead of a felony. It also may be that joinder is more common in some types of offenses than others (expansive drug conspiracy trials, for example), and perhaps this is the correlation that causes the gap in conviction rates, rather than the joinder effect.

As these alternative theories show, "crime type" can be considered along two dimensions. First, we can characterize offenses by their seriousness, asking about the conviction rates and joinder patterns for felonies, misdemeanors, and petty offenses.[85] Second, we can look at the nature of the alleged criminal act, putting the most serious charge into one of seven categories: (a) violent crimes; (b)

85. In general, a felony is a crime for which more than one year imprisonment is authorized, and a Class A misdemeanor is a crime for which six months to one year imprisonment is authorized. *See* 18 U.S.C. § 3559(a)(1)–(6) (2006) (defining generally the grades of criminal offenses). A petty offense, which includes Class B and Class C misdemeanors as well as infractions, includes crimes for which less than six months imprisonment or only a fine is authorized. *Id.* §§ 3559(a)(7)–(9).

property crimes; (c) drug crimes; (d) regulatory crimes; (e) public order crimes; (f) weapons offenses; and (g) immigration crimes.[86]

The conviction rate fluctuates widely depending on whether the most serious trial count ("Count 1") is a felony, misdemeanor, or petty offense.

TABLE N

Crime Level	Conviction Rate
Felony (n = 14,790)	82%
Misdemeanor (n = 3,077)	42%
Petty (n = 1,190)	96%

The Table shows that nearly all top counts that are petty offenses end in convictions, while fewer than half of the misdemeanor charges do. But it turns out there is much less to these variations than meets the eye. We also see that the overwhelming share of Count 1s (78%)[87] charge a felony, leaving a relatively small number of misdemeanor and petty offense cases. More importantly, drilling down into the data reveals that there is a nearly complete overlap between misdemeanor and petty offense cases and bench trials: an astonishing 93% of cases where a misdemeanor is the most serious charge and 95% of the cases where a petty crime is the most serious are tried by the judge alone.[88] Since we have already controlled for a bench versus jury trial effect in our model, this should largely account for the differences in conviction rates shown here.

In fact, when we control for crime seriousness in our model, we see that the joinder effect remains, and is nearly unchanged. As

86. With one exception, these are the offense classifications used by the Administrative Office of the U.S. Courts. *See* Offense Classification Table, http://fjsrc.urban.org/noframe/dd/cross/aofjsp2002.pdf (last visited Mar. 25, 2006) (demonstrating this in the "FJSP Primary Offense" column). Tbe one exception is that we took the category "Public Order Offenses" —which covers an extremely broad array of crimes—and broke it down into two subcategories: "regulatory" offenses and "other" public order crimes. *See id.* (same). The former category includes civil rights crimes, customs crimes, social security crimes and, postal offenses. The "other" public order crimes include racketeering, extortion, criminal income tax evasion, and traffic offenses. *Id.*

87. Of the 19,057 defendants who were tried during the period of the study, 14,970 faced a felony as the most serious charge (14,970 / 19,057 = 0.776).

88. Of the 3,077 defendants who went to trial facing a misdemeanor as the most serious charge, 2,854 elected a bench trial. Similarly, of the 1,190 defendants who faced a petty offense, 1,131 either elected a bench trial or were required to have one because they were not entitled to a jury. In contrast, only 8 percent of those charged with a felony in Count 1 (1,237 out of 14,790) had a bench trial.

shown in the Probit Model in the Appendix, joinder of counts still has a positive and significant impact on the conviction rate, and joinder of defendants does not.[89] As a result, when controlling for the identity of the decisionmaker and the seriousness of the crime, we see the following revised joinder effect:

TABLE O

Conviction Rate by Joinder Status (controlling for factfinder and crime seriousness)			
	One Defendant	Multiple Defendants	Difference
One Count	75%	73%	-2%
Multiple Counts	84%	83%	-1%
Difference	9%	10%	-

We therefore can say with some confidence that the differences in conviction rates based on joinder status are not explained in any meaningful way by the seriousness of the offense charged.

Controlling for the *type* of crime shows similar results. Once again we begin with the basic conviction rates:

TABLE P

Crime Type	Conviction Rate
Violent (n = 1,245)	79%
Property (n = 3,443)	80%
Drug (n = 6,171)	85%
Weapon (n = 2,210)	80%
Immigration (n = 644)	85%
Regulatory (n = 384)	63%
Public Order (n = 4,940)	62%

89. The Appendix contains an expanded version of the Probit Model shown *supra* in Tables G and J. The seriousness of the crime is first analyzed in Appendix Model 5.

The outlying categories are regulatory and public order offenses, but these are unlikely to have much explanatory power for the conviction gap. The "regulatory" group is small, making up only 2% of the trial defendants. And while the public order group is larger (26% of defendants), further investigation shows that an usually large percentage of public order trials (74%) are bench trials, where the conviction rates for all types of offenses are lower.[90] Again, because our model has already controlled for the identity of the factfinder, we would not expect this subset of offenses to materially change our joinder hypotheses.

When we add this variable to our expanding model—now controlling for factfinder, crime seriousness, and crime type—we see virtually no change from the prior figures.[91]

TABLE Q

Conviction Rate by Joinder Status (controlling for decisionmaker, crime seriousness and crime type)			
	One Defendant	Multiple Defendants	Difference
One Count	75%	73%	-2%
Multiple Counts	84%	83%	-1%
Difference	9%	10%	-

3. Lawyer Type

Another variable we should consider is the type of lawyer the defendant had. Here the data allow us to make three relevant distinctions.[92] We can tell from the records when defendants were

90. Of the 4,940 defendants whose most serious charge was a public order offense, 3,653 elected bench trials (3,653 / 4,940 = 0.739). The conviction rate for public order bench trials was 53%, which is consistent with the lower conviction rates for bench trials across all categories of crimes. *See supra* Table I (calculating conviction rate by factfinder).

91. The expanded Probit Model on which this table is based is set forth in the Appendix *infra* at Model 6.

92. The data also allow us to examine cases where the defendant had no lawyer, either because he was not entitled to one, *see, e.g.*, Scott v. Illinois, 440 U.S. 367, 373 (1979) (holding that the right to appointed counsel only applies to cases where defendant could be sentenced to imprisonment), or because he waived counsel and proceeded pro se. Because of the small number and the often unusual nature of these cases, defendants without counsel were not considered in this part of the study.

represented by retained counsel, when they were represented by full-time defense counsel who work for a public defender's office, and when they were represented by court appointed lawyers from the private bar ("panel" attorneys).[93]

Even speculative differences among the three groups are not obvious and require some high-level generalizations. Perhaps, like judges and juries, one group's clients have a much higher conviction rate, one that explains the different outcomes better than joinder. It may also be that different groups of lawyers approach joinder problems differently. Those represented by retained lawyers may be more likely to fight for severance because they have greater resources to object to every prosecutorial move. Perhaps panel attorneys are more likely to accept joinder because they find it easier to enter joint defense agreements with co-counsel, thereby minimizing the (literal) prisoner's dilemma problem. Or perhaps public defenders, as institutional repeat players, are more likely to fall into patterns on joinder and severance issues, ones that over time become informal practice norms within the office.

When we turn to the data, however, we see that while there is some difference in conviction rates based on lawyer type, the gap is not huge:

TABLE R

Lawyer Type	Conviction %
Retained (n = 6,852)	81%
Public Defender (n = 3,309)	78%
Panel Attorney (n = 6,292)	86%

Given this, we are unsurprised to see that our working hypothesis is supported when we control for counsel type. Holding this and all prior variables constant, our Probit Model shows that no matter who represents the accused, adding counts increases the likelihood of conviction, while joining defendants has little

93. A "panel" attorney is appointed by the district court from a list of private attorneys for a particular case. *See* 18 U.S.C. § 3006A(b) (2006) (setting out the procedures for appointment of counsel from a panel); *see also* CAROLINE WOLF HARLOW, BUREAU OF JUSTICE STATISTICS SPECIAL REPORT: DEFENSE COUNSEL IN CRIMINAL CASES 2 (2000), *available at* http://www.ojp.usdoj.gov/bjs/pub/pdf/dccc.pdf (defining panel attorneys as private attorneys chosen by the district court from a list on a case by case basis).

significance.[94] As a result, our revised conviction matrix changes very little:

TABLE S

Conviction Rate by Joinder Type (controlling for factfinder, crime seriousness, crime type, and defense counsel)			
	One Defendant	Multiple Defendants	Difference
One Count	76%	73%	-3%
Multiple Counts	85%	83%	-2%
Difference	9%	10%	-

4. Geography

A final variable worth testing is regional influences. Here we will use, as a very rough proxy, the federal judicial circuit where the trial court is located. Each circuit has its own local rules, jury pools, enforcement priorities, and case mix. Perhaps these variations lead to different conviction rates and perhaps even to different local practices that influence the joinder decision.

The trial conviction rates among the circuits are only mildly revealing:

TABLE T

Trial Conviction Rates by Circuit					
CADC (n = 135)	84%	CA4 (n = 2,496)	85%	CA8 (n = 1,150)	80%
CA1 (n = 1,369)	92%	CA5 (n = 2,065)	79%	CA9 (n = 2,115)	83%
CA2 (n = 1,163)	83%	CA6 (n = 1,307)	78%	CA10 (n = 734)	77%
CA3 (n = 889)	82%	CA7 (n = 882)	92%	CA11 (n = 4,752)	57%

Here we see relatively modest variations in the conviction rates,[95] with one outlier.[96] Yet even with these variations, our

94. *See* Appendix Model 7 (analyzing conviction rates relating to the type of defense counsel).

95. The highest conviction rate is in the First and Seventh Circuits, at 92%, and the lowest rate *other* than the Eleventh Circuit (discussed *infra* note 96) is the Tenth Circuit at 77%. The

conclusions remain the same. Almost without regard to the location of the trial, the joinder of counts significantly increases the likelihood of conviction, and the addition of defendants has little or no effect.[97]

* * * * *

We now have a full model, one that controls for factfinder, crime seriousness, crime type, lawyer type, and geography, with only the judge / jury distinction having a significant impact on the conviction rate. Holding all these variables constant, we can calculate the joinder effect more precisely:

TABLE U

Conviction Rate by Joinder Status (Final)			
	One Defendant	Multiple Defendants	Difference
One Count	76%	74%	-2%
Multiple Counts	85%	83%	-2%
Difference	9%	9%	-

From this model we can now draw some conclusions.

C. Conclusions

1. Joinder of counts has a significant impact on the outcome of trials. A defendant who stands trial on a single count is roughly 9 percentage points less likely to be convicted than defendants who face multiple counts. Stated differently, a defendant's chances of conviction *increase by more than 10%* if he stands trial on more than

unweighted mean rate of conviction, including the Eleventh Circuit, is 81%; the standard error for conviction rates is 0.089.

96. The Eleventh Circuit has a conviction rate of only 57%, the lowest by twenty percentage points. As explained in more detail elsewhere, *see* Leipold, *supra* note 79, at 192, the oddness of the Eleventh Circuit is almost entirely a function of the extremely high number of bench trials (and bench acquittals) in the district courts there. It turns out that the number of bench trials is driven to a large extent by an unusual arrangement between two military bases and the district courts of the Middle District of Georgia, under which violations that take place on the military bases, including traffic offenses, are tried in the civilian courts. *See id.* at 192 n.142 (explaining the details of the arrangement between the district and the military bases).

97. The impact of geography on the conviction rates is set forth in Appendix *infra* at Model 8.

one count.[98] The impact of adding counts is present whether a defendant stands trial alone or with other defendants.

2. Although the models test only the distinctions between one count and more than one count, the frequency data discussed above in Part IlI(A)(2)[99] suggest that the conviction rate rises with the addition of Count 2 and then again with Count 3, but after that remains relatively stable for additional charges.

3. Somewhat surprisingly, the joinder effect of counts was more pronounced in bench trials than in jury trials. However, the unusual nature of the bench trial docket (disproportionately public order crimes and misdemeanors), coupled with the relatively small number of bench trials, makes it difficult to draw firm conclusions from the data.

4. Joinder of defendants standing alone seems to have little effect, despite the conventional wisdom that it does. In fact, holding other variables constant, our model predicts no effect or even a slight beneficial effect to the defense from the addition of a co-defendant. The lack of impact is seen whether a defendant is charged with one count or multiple counts.

5. As a result, and contrary to our hypothesis, those who face both types of joinder are not the worst off. Although they are 7 percentage points more likely to be convicted than our most favored group, they are no worse off than single defendants who face multiple counts.

6. The joinder effect is quite robust. It is significant and reliable in each instance when we analyze other variables, including variables such as the identity of the factfinder, that are strongly correlated to conviction rates.

7. Despite the prejudicial effect of joinder, defendants facing multiple charges are not significantly more likely to plead guilty than those who face only one charge.

D. Weaknesses: What's Missing

As with most empirical studies, the validity of the prior Section is hostage to the limits on the data. In particular, there are two

98. The increase of nine percentage points, from the mid-70% range to the mid-80% range, translates into a greater than 10% increase in the chances of conviction. Thus for defendants standing trial alone, their chances go from 76% to 85% when counts are added, an increase of almost 12% ((.85 - .76) / .76 = 0.118).

99. *See supra* note 65 and Figure 1 (showing a 68% to 82% jump when a second count is added, 88% for a third count, and then the chance remaining steady for additional counts).

variables we did not measure that could influence the conviction rates and therefore alter our conclusions.

First, we cannot measure the relative strength of the evidence across the different categories of cases. Perhaps defendants are more likely to be acquitted in one-defendant, one-count cases because the prosecutor had so little evidence of wrongdoing that she could only muster a single charge against a single defendant. Stated differently, the number of counts and defendants may increase as the strength of the evidence increases, meaning that case quality, not prejudice from a consolidated trial, is the explanatory variable.[100] This reasoning is undermined somewhat by the plea bargaining statistics, which show that there is virtually no difference in the guilty-plea rates based on joinder status.[101] If there really were differences in case strength, we would expect those with the weaker cases—by hypothesis, our single defendant, single count group—to be disproportionately willing to take their chances at trial, but the data show otherwise. Ultimately, however, the only reliable indicator of case strength for a dataset this size is the final outcome, and so this variable must remain an unquantified limit on the conclusions.

Second, we could not account for individual defendant characteristics that might have influenced the trial outcomes. The government data that are publicly available redact information about defendants' race, age, gender, education, and perhaps most importantly, criminal history. It is easy to imagine stories that would help explain why certain defendants are joined and convicted more often, explanations that have only a marginal intersection with the impact of joint trials. And so while it would be improper to assume that these characteristics *have* an effect on conviction rates, we cannot rule it out, leaving the calculations unfortunately incomplete.

IV. ANALYSIS

Increasing a defendant's chance of conviction by more than 10% is a high price to pay to secure the societal benefits of joined trials. Armed with the data, we now can reconsider and perhaps re-weigh the interests between the government and the accused.

100. Thanks to our colleague Steve Beckett for helping us develop this point.

101. *See supra* Table H and accompanying text (determining that the plea rates stay within 86% and 91% regardless of the joinder status).

A. Rethinking the Defendant's Interest

The threshold question is why defendants' interests are entitled to so little weight. We assume, as we should, that judges are neutral when it comes to ruling on pretrial matters and would not deliberately refuse to sever cases if they thought joinder would unfairly harm the accused. So the fact that so many defendants are worse off for what appears to be precisely that reason requires some explanation.

We might conclude that courts are just not very good at spotting problematic joinders. Motions to sever are typically made pretrial, before the evidence is fully developed and before the judge has immersed herself in the case. Defendants bear the burden of proving prejudice,[102] and perhaps defendants can rarely do better than speculate about the harms that they fear will flow from a joint trial.[103] But while the inability to prove harm almost certainly contributes to the problem—especially if defense counsel fail to move for severance as often as they should—it is doubtful that this is the primary culprit. As noted above, the risks of prejudice arise frequently, are well known, and are largely undisputed.[104]

A more plausible explanation is that courts recognize the risks but believe that jury instructions are genuinely curative. By training and necessity, judges may believe that juries really do listen to, understand, and implement the command that they compartmentalize the evidence in their minds, applying each bit of information only to

---

102. *See* United States v. Evans, 272 F.3d 1069, 1084 (8th Cir. 2001) (explaining that there is a strong presumption against severing cases that were joined properly); United States v. LiCausi, 167 F.3d 36, 48 (1st Cir. 1999) (holding that the defendant seeking severance has the burden to show that he would be severely prejudiced without it).

103. Wayne LaFave and his co-authors have nicely described the defendant's problems in proving prejudice:

> [I]t is very difficult for the trial judge to make a finding on the prejudice issue before trial, for it involves speculation about many things which may or may not occur. Also, judges are understandably reluctant to make a finding of prejudice during trial, after the prosecution has put in most or all of its proof. And if a trial judge denies defendant's severance motion on the ground that the showing of prejudice has not been made, experience has shown that it is virtually impossible for the defendant to prevail on appeal.

4 LAFAVE, *supra* note 4, at 602; *see also* United States v. Novaton, 271 F.3d 968, 988–89 (11th Cir. 2001) (analyzing defense argument that severance was required so defendant could obtain the testimony of a co-defendant whom he claimed would testify on defendant's behalf in a separate trial, but would refuse to testify in a joint trial). The court of appeals upheld the trial judge's refusal to sever, finding that the affidavits submitted in support of his motion lacked "specific and exonerative facts" needed to show prejudice. *Id.* at 989–90.

104. *See supra* Part II.B (laying out the three types of risk about which a suspect facing joint counts is concerned).

the defendant and the charge to which it relates.[105] When the use of jury instructions is coupled with the other devices that are available to soften the harshest forms of prejudice,[106] courts might believe that they can still buy the benefits of consolidated trials at a relatively low cost to the accused.

But ultimately there is more at work than inattention or inadequate remedial efforts. It is impossible to miss the message of the case law that some amount of prejudice is tolerable, and that efficient trials are important enough to justify some genuine harm. In the face of legislative silence, the Supreme Court has spoken clearly about both the desirability of consolidated proceedings and the limited class of harms that should lead to severance,[107] and most lower courts have embraced this attitude without dissent. Whether judges are in fact aware that the "prejudice premium" exceeds 10% is another matter, but under current doctrine, it seems likely that judges would

---

105. Social science research on the effectiveness of jury instructions in curing the prejudice of joint trials is not conclusive but is also not encouraging. *See, e.g.*, Greene & Loftus, *supra* note 24, at 205 ("Our data showed that no matter when [the jury instruction] was presented, the instruction was ineffective in reducing guilt estimates [of those who judged defendants charged with multiple counts] to the level obtained from subjects who judged only one offense."). These findings were consistent with most, but not all, of the social science research on the topic. *See, e.g.*, Sarah Tanford & Steven Penrod, *Social Inference Processes in Juror Judgments of Multiple-Offense Trials*, 47 J. PERSONALITY & SOC. PSYCHOL. 749, 761 (1984) (reporting that in a mock juror experiment, "[a] very strong set of judges' instructions designed to eliminate joinder effects had no influence on verdict judgments, or on any other dependent measures"). *But see* Sarah Tanford et al., *Decision Making in Joined Criminal Trials: The Influence of Charge Similarity, Evidence Similarity, and Limiting Instructions*, 9 LAW & HUM. BEHAV. 319, 333–34 (1985) (finding that limiting instructions reduced the joinder effect, while noting that "[t]his finding runs counter to previous joinder research[,] and indeed, much of existing research in other domains which indicates that limiting instructions tend to be ineffective" (citations omitted)).

106. Thus, for example, courts can require that prosecutors redact the pretrial statement of a non-testifying co-defendant to eliminate any direct identifying reference to any of the joined defendants who did not make the statement. *See* Gray v. Maryland, 523 U.S. 185, 197 (1998) (allowing such a redacted statement where the defendant's name is replaced by a symhol). This procedure is designed to address the so-called "*Bruton* problem." *See* Bruton v. United States, 391 U.S. 123, 123–24 (1968). In *Bruton*, the Court found that when a non-testifying co-defendant makes a pretrial statement that implicates the accused, it is unrealistic to believe that the jury can limit its consideration of the statement to the co-defendant (against whom it would be admissible), and not consider the statement against the accused (for whom it would be inadmissible). *Id.* at 126.

107. *See, e.g.*, Zafiro v. United States, 506 U.S. 534, 537 (1993) (reaffirming the importance of consolidated trials and, arguably, narrowing the type of harm that required severance). Although the question presented in the case was whether antagonistic defenses required severance as a matter of law (it does not), the Court stated that "a district court should grant a severance under Rule 14 only if there is a serious risk that a joint trial would compromise a specific trial right of one of the defendants, or prevent the jury from making a reliable judgment about guilt or innocence." *Id.* at 539. As examples, the Court offered the *Bruton* problem, *see supra* note 106, and cases where a co-defendant would provide exculpatory testimony but is entitled under the Fifth Amendment not to take the stand in a joint trial. *Id.*

not be surprised to learn that joinder makes a significant difference in trial outcomes.

Still, there are reasons to think that courts have systematically undervalued the defendant's interest in a separate proceeding. No phrase better captures the current judicial mind-set than the oft-repeated "defendants are not entitled to severance merely because they may have a better chance of acquittal in separate trials."[108] While the phrase has a sensible ring to it—defendants want all kinds of things they are not entitled to—it can inadvertently mask an important distinction between severance motions and other types of pretrial requests.

Courts and legislatures routinely evaluate the accused's interest in certain procedures far lower than defendants would like, but they do so because of a suspicion that what the defendant is really seeking is either illegitimate or highly susceptible to abuse. For example, defendants often want more pretrial disclosure from the government than the rules allow, plausibly claiming that greater discovery will reduce the chances of surprise and therefore lead to more accurate trials.[109] No one disagrees with this argument in isolation, but discovery is deliberately made restrictive because of the fear that defendants will misuse the information, either to concoct a better lie to offer in defense or to intimidate witnesses.[110] Likewise, defendants who fail to abide by certain deadlines can be precluded from presenting an alibi witness, raising an insanity defense, changing counsel, or withdrawing a guilty plea,[111] even though in

108. The phrase is most prominently featured in *Zafiro*, 506 U.S. at 540, and has been repeated by numerous lower courts. *See, e.g.*, United States v. Rodriguez-Marrero, 390 F.3d 1, 26 (1st Cir. 2004); United States v. Flores, 362 F.3d 1030, 1040 (8th Cir. 2004) (finding that a defendant's possible better chance for acquittal in a separate trial is not sufficient to justify severance); United States v. Cassano, 132 F.3d 646, 651 (11th Cir. 1998) (citing *Zafiro* for the idea that a defendant's increased chance of acquittal if tried alone is not enough to justify severance).

109. For perhaps the best known articulation of this view, see William J. Brennan, Jr., *The Criminal Prosecution: Sporting Event or Quest for Truth?*, 1963 WASH. U. L.Q. 279, 287 (1963); *see also* H. Lee Sarokin & William E. Zuckerman, *Presumed Innocent? Restrictions on Criminal Discovery in Federal Court Belie This Presumption*, 43 RUTGERS L. REV. 1089, 1090 (1990-1991).

110. One prominent example of the struggle between fairness to the accused and the fear of misuse concerns the pretrial disclosure of witness statements, with the Supreme Court initially finding that equity required pretrial disclosure, *see* Jencks v. United States, 353 U.S. 657, 671–72 (1957), and Congress quickly reversing this judgment with the Jencks Act, 18 U.S.C. § 3500(a) (2006). *See also* FED. R. CRIM. P. 26.2 (incorporating substance of the Jencks Act).

111. *See* FED. R. CRIM. P. 12.1(e) (permitting the exclusion of alibi witnesses who are not disclosed to the government within ten days of prosecutor's request for the names of such witnesses); FED. R. CRIM. P. 12.2(a) (providing that a defendant who fails to timely notify the government of her intent to raise an insanity defense is barred from raising the defense at trial); FED. R. CRIM. P. 11(d)–(e) (precluding a defendant from withdrawing a guilty plea after sentence, and requiring leave of court to withdraw after the plea has been accepted but before sentencing).

many instances it could plausibly be argued that allowing the request would lead to more fair and accurate results. Once again, the belief is that, in the absence of restrictions, defendants will deliberately engage in strategic behavior to sandbag the prosecution, facilitate perjury, or otherwise manipulate the system. And while advocates quarrel over the size of the risk, the difficulty in sorting the legitimate requests from the illegitimate ones means that even appropriate defense requests end up being discounted.

But motions for severance are different because it is hard to think of how greater use of severance could somehow undermine the process. Defendants *gain* virtually nothing from a separate trial.[112] They are not entitled to keep out evidence in a separate proceeding that would otherwise be admissible. There is no strategic game they can play to hide evidence from the government or confuse the jury about their actions or responsibility. The only benefit to a sole trial on a single count is that it *avoids the prejudice* of joinder: inadmissible evidence stays out, the jury keeps the evidence straight, no inference of criminal disposition is created, and the defendant's trial choices (e.g., to testify or not) are preserved.

A skeptic would argue that defendants always have a motive for filing any colorable motion because they always have an incentive to make the prosecutor's life miserable. The government, like the defense, faces case-load pressures, and a successful severance motion may lead to a more attractive plea offer from a prosecutor trying to manage her now slightly-larger docket.[113]

But this argument proves too much because *any* procedural challenge that is made available to the defense creates this same concern. The question is not whether defendants have a motive to file the severance motion—we assume that they do—and not whether the increased use of severance would lead prosecutors to change their behavior, as it probably would. Instead, the issue is whether granting

112. Saying that defendants "gain nothing" is slightly overbroad, given the occasional benefits of joint defense agreements as well as the other resource savings that accrue from joinder. *See supra* notes 31–32, 42 and accompanying text (detailing some possible benefits from joint trial).

113. A more refined argument would be that a motion to sever could influence the judge's ruling on the admissibility of "other crimes" evidence. The government may have a good argument under Federal Rule of Evidence 404(b) for why evidence of crime B should be admissible to prove crime A, but judges typically dislike mini-trials on uncharged offenses, especially if they're time-consuming. So if defense counsel can persuade the judge to sever the cases, the judge might very well decide that it would be wasteful to prove offense B twice—namely once in the trial of offense A and a second time in the severed proceeding on that count. As a result, the judge might deny a Rule 404(b) motion by the government to prove offense B in the trial of offense A. In such cases, defendants have nothing to lose by routinely moving to sever.

the motion would create the opportunity for defense mischief, would deprive the prosecution of the chance to fairly prove its case, or would otherwise distort the trial process—fears commonly used in other contexts to strike the balance against the accused. Because the defendant's separate trial should look virtually identical to the joint trial, *other than* the prejudicial features that come from a single proceeding, it is hard to understand why courts should be as tolerant of prejudice here as they are in other, dissimilar, pretrial settings.

This brings us to an important conclusion. Now we can say not only that joinder "prejudices" the defendant, but that joinder *unfairly* causes prejudice. Unlike many other harms about which defendants complain, the effects of spillover evidence, inference of a criminal disposition, and jury confusion do not themselves further any policy goals, do not avoid risks of defendant manipulation, and do not even plausibly make trials more accurate or more fair – quite the contrary. So to say a defendant is not entitled to a separate trial just because his chances of acquittal are better is to say that a defendant should not be granted a severance merely to avoid a significant (~10%) risk that his conviction will be influenced by improper factors.[114]

Framed in this way, it makes the current balance seem badly struck. Outcomes based on impermissible factors immediately raise the specter of wrongful convictions, and it is hard to argue that abstract notions of efficiency are worth an appreciable increase in the risk of sending innocent people to jail. The last few years have revealed a rash of demonstrably false convictions, stories that can easily drown out claims of scarce dollars, crowded dockets, and witness inconvenience.[115]

---

114. The restrictive severance rules are also distinctive because, unlike other circumstances where we force defendants down a less desirable litigation path, there is not countervailing value to the accused. We know, for example, that many suspects foolishly decide to waive their right to a conflict-free lawyer, or waive counsel entirely and represent themselves, even though no neutral observer would believe that this is in the defendant's best interests, and, indeed, can easily lead to less accurate results. *Cf.* FED. R. CRIM. P. 44(c)(2) (requiring trial judges to inquire into cases of joint representation, although not requiring that each defendant be represented by counsel who has no conflict of interest). In these cases, we accept the predictable harm that flows to the defendant because we value some other legitimate interest *that is also of value to the defense,* namely, autonomy and the right to make decisions about one's own interest. There is no comparable value to the accused in favoring joint trials over single trials.

115. It is important not to oversell this point. Many of the factors present in the false conviction cases, such as bad eyewitness identifications and false confessions, are just as likely in a single count, single defendant case as in any other, and would be unaffected by severance. For extensive discussions of the wrongful conviction problem and its causes, see Samuel L. Gross et al., *Exonerations in the United States, 1989 Through 2003*, 95 J. CRIM. L. & CRIMINOLOGY 523, 523–27 (2005) (analyzing the results of a study compiling exonerations from wrongful convictions); BARRY SCHECK ET AL., ACTUAL INNOCENCE: WHEN JUSTICE GOES WRONG AND HOW TO MAKE IT RIGHT 361 (2000) (showing factors leading to wrongful convictions).

The problem, of course, is that the strength of the defendant's interests is only one side of the balance. Before condemning the injustice of the current practices, we should also re-examine the government's interests.

B. Rethinking the Government's Interest

No matter what the data show, at some level there is an undeniable interest in joined trials, and thus genuine weight on the government's side of the scales. But once again, there is reason to think that the correlation between joint trials and efficiency is less than is sometimes believed.

1. The Conventional Interests

The criminal caseload in the federal district courts has increased over 50% since 1995 (and almost 15% just since 2000),[116] while the number of authorized judgeships during that time rose only slightly.[117] Most of the additional load has been absorbed by plea bargaining; in fact, the pace of guilty pleas has picked up, leading to a decrease in both the percentage of cases that go to trial[118] and in the absolute number of trial defendants.[119] But even extra plea bargains take time, and so it is unsurprising that institutional pressures to clear the docket have seeped into the severance decisions.

Whether the caseload pressure would be significantly compounded by rules that encouraged more frequent severance is another matter. Robert Dawson pointed out over a quarter-century

---

116. *See* Administrative Office of the United States Courts, Federal Judicial Caseload Statistics, Judicial Caseload Indicators, Calendar Years 1995, 2000, 2003, and 2004, www.uscourts.gov/caseload2004/front/judbus03.pdf (last visited Mar. 25, 2006) (providing caseload statistics for U.S. federal courts, the Federal Probation System, and pretrial services). The number of criminal cases filed increased 56.7% between 1995 and 2004, while the number of defendants against whom charges were filed increased 47.2%. The increase between 2000 and 2004 was 15.5% for the number of cases filed, and 13% for the number of defendants. *Id.*

117. *See* Administrative Office of the United States Courts, Judicial Facts and Figures, tbl 3.1, Criminal Cases and Defendants Filed, Terminated, Pending (Includes Transfers), http://www.uscourts.gov/judicialfactsfigures/table3.01.pdf (last visited Mar. 25, 2006) (showing that the number of authorized judgeships in the U.S. district courts rose from 649 in 1995 to 679 in 2004, an increase of less than 5%).

118. Between 1994 and 2003, the percentage of defendants who stood trial dropped from slightly over 9% of defendants to slightly over 4%. *See* Bureau of Justice Statistics, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS ONLINE, tbl. 5.22, http://www.albany.edu/sourcebook/pdf/t522.pdf (last visited Mar. 25, 2006) (reporting, by type of disposition, the number of defendants disposed of in U.S. district courts from 1945 to 2003).

119. Between 1994 and 2003, the number of federal defendants who had their case disposed of at trial dropped by more than one-third, from 5,526 defendants in 1994 to only 3,463 in 2003. *Id.*

ago the fallacy of assuming that a greater use of severance would result in a proportional increase in the number of trials.[120] A defendant who successfully *moves* for three separate trials may not *stand trial* three times; if the first trial ends in a conviction (as most do), the defendant may quickly plead guilty to the remaining counts.[121] Conversely, if the first trial ends in an acquittal, the prosecutor might decide to cut her losses and either offer an attractive deal to the remaining defendants or even *nol pros* the remaining charges. While the ultimate impact of fewer consolidated trials is an empirical question, this type of Bayesian reasoning by the parties could dramatically lower the costs of increased severance.[122]

It also is worth noting that while the federal caseload has steadily increased, the frequency of joinder has remained constant. If we look at the percentage of all defendants (not just trial defendants) who face multiple charges or who are joined with co-defendants year-by-year, we see virtually no change in the charging behavior of the government.

TABLE V

Percentage of Cases Joined, by Year[123]						
	1999	**2000**	**2001**	**2002**	**2003**	**Total**
Multiple Counts	50%	50%	52%	52%	50%	**51%**
Multiple Defendants	35%	35%	35%	35%	33%	**34%**

Given the incentives the government has to consolidate proceedings, these figures suggest that prosecutors have maximized their joinder authority and cannot obtain more consolidated

120. *See* Dawson, *supra* note 2, at 1389–91 (discussing how severance actually may lead to the disposition of a second case through settlement rather than through a trial).

121. A defendant is particularly likely to plead guilty to the remaining counts following a conviction on the first if he believes that he has a chance of receiving a reasonable sentence reduction. *See* UNITED STATES SENTENCING GUIDELINE MANUAL § 3E1.1 (2004) (permitting a two to three level reduction in the offense level if the defendant "clearly demonstrates acceptance of responsibility for his offense").

122. *See* Dawson, *supra* note 2, at 1389–91 (discussing how severance actually may lead to the disposition of a second case through settlement rather than through a trial). In support of his argument, Professor Dawson noted a 1973 study in Vermont, where defendants had the right to separate trials if they were charged with serious felonies. *Id.* at 1390. Although joint trials in these cases became "virtually unheard of," most of the time the first trial was the only trial; after the verdict in the first case was known, the rest of the cases were resolved out of court. As a result of this (admittedly limited) study, the state judges and prosecutors concluded that separate trials were more economical than joint trials. *Id.* at 1390–91 (discussing Peter Langrock, *Joint Trials: A Short Lesson From Little Vermont*, 9 CRIM. L. BULL. 612 (1973)).

123. The number of defendants considered in each of the years were: 71,738 in 1999; 72,997 in 2000; 73,463 in 2001; 76,381 in 2002; and 81,141 in 2003, for a total of 375,720 defendants.

indictments without significantly changing their case mix. In any event, the figures indicate that the increased case load is being absorbed only in proportional terms by joint proceedings.

The other articulated government interest—avoiding the "scandal of inconsistent verdicts"—is easier to dismiss because it has always felt like a makeweight argument. While courts genuinely care about the perception of fairness arising from trials, institutionally the system has a thick skin; it does not embarrass easily.[124] The claim of scandal seems particularly overdone in the context of jury verdicts, where courts have repeatedly upheld verdicts that were internally inconsistent in the *same case*.[125] More pointedly, the awkwardness of having juries make inconsistent decisions in separate trials pales in comparison to the scandal of wrongful convictions, which, at the margins, joint trials seem to foster.

2. The Interest in Convictions

There is a third government interest at stake, one that is rarely discussed in the cases or literature. There will be times when the prosecutor has enough evidence to bring an additional charge or indict an additional defendant, but because of resource constraints will *only* do so if the new charge or defendant can be joined with others. The

---

124. Compare, for example, the Supreme Court's earlier view that the Fourth Amendment's exclusionary rule was justified in part by the independent need to prevent courts from becoming "accomplices in the willful disobedience of a Constitution they are sworn to uphold," *see* Elkins v. United States, 364 U.S. 206, 223 (1960), with its later, singular, focus on whether the exclusionary rule served as a deterrent. *See, e.g.*, United States v. Janis, 428 U.S. 433, 458 n.35 (1978) (concluding that "judicial integrity rationale" was subsumed by the inquiry into whether exclusion of evidence deterred unlawful police conduct). Consider also *United States v. Alvarez-Machain*, 504 U.S. 655 (1992), where the defendant was kidnapped in a foreign country and brought to the U.S. to stand trial. The Court rejected a claim that the charges should be dismissed because of the illegal method used to bring defendant before the court, saying:

> Respondent . . . may be correct that [his] abduction was "shocking," and that it may be in violation of general international law principles. . . . We conclude, however, that respondent's abduction was not in violation of the Extradition Treaty between the United States and Mexico. . . . The fact of respondent's forcible abduction does not therefore prohibit his trial in a court in the United States for violations of the criminal laws of the United States.

Id. at 669–70.

125. Federal courts have repeatedly ruled that inconsistent verdicts in the same case are not grounds for overturning a defendant's conviction. *See* United States v. Dotterweich, 320 U.S. 277, 279 (1943) ("Whether the jury's verdict was the result of carelessness or compromise or a belief that [one defendant] should suffer the penalty [and not the other] . . . is immaterial. Juries may indulge in precisely such motives or vagaries."); *see also* Dawson, *supra* note 2, at 1391–92 (rejecting the premise of the "inconsistent verdict" rationale that separate trials against similar co-defendants really are identical: "[a]lmost always the evidence admissible in different trials against alleged coparticipants in an offense will differ enough to permit different treatment by the jury," particularly when co-defendants are represented by different attorneys).

charge may be minor enough or the additional defendant peripheral enough to the primary criminality that it would be too expensive, too burdensome on the participants, or simply not worth the trouble to bring but for the availability of joint trials.

More defense-favorable severance rules could undermine this interest because the net results would be fewer charges and fewer convictions, presumptively leading to more unpunished crime.[126] Whether these foregone prosecutions ultimately represent a cost or a benefit depends on whether we view this type of resource constraint as a valuable check on prosecutorial zealousness or as an unfortunate opportunity cost.[127] But even if we assume that the marginal loss of prosecutions is a social cost, the waters are still muddy. Once we look at the data, alternate explanations emerge for bringing the extra charge against the extra defendant.

One lurking question about the government's charging conduct is whether prosecutors are unfairly using the prejudicial impact of joinder to improve their hand at trial. Here the worry is *not* that prosecutors are simply seeking an additional conviction by adding an extra name or charge to the indictment; instead, the worry is that prosecutors might bring additional charges or indict additional defendants simply to gain the prejudicial effect of joinder to help convict on the existing charges. Stated differently, prosecutors might be using the joinder effect "offensively," adding less serious charges to the indictment just to bias the factfinder and help them convict on the more serious charges.[128]

---

126. The percentage of foregone prosecutions that actually resulted in unpunished crime under this scenario is unclear because some of the activity not charged by the federal authorities presumably would be prosecuted by the states.

127. It may be that the mere possibility of dropped charges also gives defendants an incentive to file severance motions in every case, a point discussed in notes 112–113 and accompanying text. While a more generous severance policy will of course provide some incentive for more motions, the likelihood that defendants will be able to use this fact strategically seems small. A prosecutor who worries that a court will sever will presumably take that into account before seeking the indictment. Once the prosecutor has accounted for the possibility of severance in her charging decision, however, the only increase in the number of severance motions *granted*—and thus, the only increase in the defendant's interest to act strategically—should be in cases where the prosecutor misjudged what the court would do. In such a case, even if the prosecutor would not have filed the smaller matter in the first instance, she may continue to prosecute the separate charge or defendant anyway, given the significant strategic and political disincentives to dismiss a case once it is filed.

128. It is possible that prosecutors might also join additional defendants who would not otherwise be charged, simply to get the prejudicial effects of multi-defendant trial. On balance, however, this risk seems smaller than the possibility of adding weak counts. Not only is the joinder effect of adding defendants quite small, the need to fully prove at least one charge against each defendant should be a deterrent to adding marginal parties. So for purposes of this sub-section, the discussion will be limited to the risk of compounding charges.

This risk is speculative because there are many legitimate reasons why a prosecutor would want to charge defendants with as many counts as the evidence will bear. The Constitution may require the joinder of all charges,[129] or the prosecutor may worry that an acquittal in the first trial may collaterally estop a later proceeding.[130] Juries are often characterized as unpredictable,[131] and so giving a full range of choices on which to convict, including lesser included offenses, is usually prudent. Sometimes the defendant will have committed ten crimes but there is only admissible proof of five, and so charging all five will help the punishment reflect the defendant's actual criminal behavior. But even allowing for these reasons, the joinder effect creates a risk that additional counts will be filed or defendants will be joined *despite* the relative weakness of the evidence to support these new charges, simply because the prosecutor wants to paint a picture for the jury of a criminally-disposed defendant.

Courts are not blind to this risk. In a recent opinion upholding a refusal to sever, the Fifth Circuit observed:

> We are keenly aware that the claimed "efficiency" of a joint trial can be a surrogate for the reality that a joint trial of multiple defendants is simply to the advantage of the government. It is the potential presence of prosecutorial advantage distinct from the expense of duplicating efforts that draws our attention.[132]

---

129. In general, if the various charges the prosecutor wants to bring have the same elements, or one is a lesser included offense of the other, both charges must be brought in the same proceeding. *See* United States v. Dixon, 509 U.S. 688, 696, 705 (1993) (concluding that "where the two offenses for which the defendant is punished or tried cannot survive the 'same-elements' test, the double jeopardy bar applies" and, in accord, that "prosecution for a greater offense . . . bars prosecution for a lesser included offense").

130. Collateral estoppel will prevent a prosecutor from bringing a charge in a second trial that was "actually and necessarily" decided in the defendant's favor in the first. *See* Shiro v. Farley, 510 U.S. 222, 236 (1994) (stating that there is no collateral estoppel effect when a jury fails to return a verdict unless the record shows that the issue was actually and necessarily decided in the defendant's favor); Ashe v. Swenson, 397 U.S. 436, 444 (1970) (stating that where there is a general acquittal verdict, the court must examine the entire record to determine whether a rational jury could have based that verdict on an issue other than that which the defendant seeks to collaterally estop). Because of the general nature of jury verdicts, however, it is often impossible to tell what was "necessarily" resolved in the first trial, making the doctrine quite narrow. *See generally* 4 LAFAVE, *supra* note 4, §17.4(a).

131. But note that no matter how unpredictable juries appear in individual cases, collectively their behavior is extremely stable. Between 1989 and 2002, federal juries always convicted between 80% and 85% of the defendants who appeared before them. *See* Leipold, *supra* note 79, at 164–65 (discussing the conviction rates for juries as compared to those of judges).

132. United States v. Simmons, 374 F.3d 313, 318 (5th Cir. 2004). Interestingly, the Fifth Circuit's observations appear to have been raised *sua sponte*; the defendant in *Simmons* did not raise the issue of strategic governmental behavior in his briefs. *See* Brief for Appellant at 35, United States v. Simmons, 374 F.3d 313 (5th Cir. 2004) (No. 03-10515), 2003 WL 23857777 (arguing that the failure to severe was an abuse of discretion because of specific facts in a unique situation).

Admittedly, the impact of adding charges to the indictment can push prosecutors in both directions. On the one hand there is the simple but powerful fact that juries are more likely to convict on the most serious count if there are other crimes alleged.[133] It would take a courageous or oblivious prosecutor not to consider this fact when making the charging decision. On the other hand, there are risks: adding a weak charge may undermine the government's credibility with the jury, possibly infecting the other, stronger charges. Adding counts also means spending more scarce resources and may complicate the case presentation in undesirable ways.

Not surprisingly, there is no direct evidence of how often prosecutors bring charges with the hopes of gaining a trial advantage. We can, however, track the disposition of these additional charges. Our dataset ranks the counts according to their severity, allowing us to compare the outcome of the top count with the outcome of the less serious charges.

TABLE W

Disposition	**Count 2**			**Count 3**			**Count 4**			**Count 5**		
Count 1	Convict	Acquit	Dismiss	Convict	Acquit	Dismiss	Convict	Acquit	Dismiss	Convict	Acquit	Dismiss
Convict 77% (*n* =14,589)	78%	7%	12%	61%	8%	25%	49%	8%	36%	41%	8%	45%
Acquit 23% (*n* = 4,468)	0%	85%	12%	0%	72%	24%	0.2%	63%	31%	0.3%	56%	37%

Here we see some intriguing numbers. If a defendant is convicted at trial of the top count (which is 77% likely), he has a nearly identical chance (78%) of being convicted of the next most serious count. But as more charges are added, the chances of being convicted of each new count drop dramatically. At Count 3, a full quarter of those charges end up being dismissed. At Count 4, more than one-third of the charges are being dismissed, and by Count 5, almost one-half (45%) of the charges end in dismissal. Moreover, at Count 5, more than half of the final charges are being resolved

133. *See supra* Table U and accompanying text (discussing conviction rates based on joinder of counts and defendants).

favorably to the defense, through a combination of dismissal and acquittal.[134]

It is important to put these numbers in context. It is quite likely that a large number of dismissed charges following a conviction on Count 1 are the byproduct of the prosecutorial practice of charging several overlapping crimes—including lesser included charges – for a single criminal act. Because a defendant can normally be punished only once for a single crime,[135] lesser and overlapping charges that follow a conviction of the top count will often be dismissed as a matter of course. This could falsely inflate the impression created by the dismissal statistics that the additional counts are weak.

Yet even if this explains a significant portion of the dismissal statistics, it does not change the effect of the charging practices. For whatever reason, the percentage of counts that does not reach the jury almost quadruples between Count 2 and Count 5, such that by Count 5, the court (or prosecutor[136]) has concluded that the charges cannot or should not be pursued. Critically, however, the addition of these extra counts still has a *positive* correlation to the *likelihood of convicting on Count 1.*[137] Whether the later charges are alternative theories or stand-alone instances of criminality that warrant separate punishment, the data raise difficult questions about the wisdom of giving this degree of charging flexibility when it comes at such a high cost.

---

134. Interestingly, if the defendant is acquitted for the first count, the government's chances for success in any respect drop dramatically. Virtually all of the remaining charges end up being dismissed or end in acquittal. And so with acquittals, but not with convictions, it is fair to say that as goes the first count, so goes the rest of the case.

135. *See* Rutledge v. United States, 517 U.S. 292, 297 (1996) (presuming that the legislature does not intend to impose multiple punishments where more than one statutory provision proscribes the "same offense"); *see also* Missouri v. Hunter, 459 U.S. 359, 366 (1983) (stating that double jeopardy does nothing more than preclude a court from sentencing a defendant to a greater punishment than the legislature intended). For a highly insightful analysis of this area of the law, see George C. Thomas III, *A Blameworthy Act Approach to the Double Jeopardy Same Offense Problem*, 83 CAL. L. REV. 1027 (1995).

136. In this study "dismissal" includes cases where the case was dismissed on the defendant's motion or on the government's motion, including the government's decision to *nol pros* a case. *See* Federal Judicial Center, Federal Court Cases: Integrated Data Base Codebook at 21–22 (values for Termination Disposition Code 1 (DISP1)), *available at* www.icpsr.umich.edu/cgi-bin/bob/archive2?study=8429&path=ICPSR (documentation for ICPSR 8429).

137. *See supra* Figure 1 and accompanying text; *see also supra* Table U and accompanying text (discussing conviction rates based on joinder of counts and defendants). Recall that the models used to measure the joinder effect were calculating the change in conviction rate of Count 1.

C. Remedies and Questions

1. What Remedy?

The data provide modest support for the argument that some rebalancing of the interests—fairness to the accused versus efficiency of processing cases—is in order. As usual, identifying the problem does not tell us how to solve it, and this study is more concerned with the former than the latter. There are, however, some points that might inform any remedial efforts.

The simple and unrealistic solution would be to recast the rules and permit defendants to have separate counts or separate parties severed on demand. But in the modern age, this type of proposal seems highly improbable, at least in the federal system. Now and in the foreseeable future, criminal case processing will take place under resource pressures, making any proposal with the potential to significantly multiply the number of trials a likely non-starter.[138]

Alternatively, we might place a heavy thumb on the defense side of the balance. For example, we might say that a "colorable" or "prima facie" claim of prejudice by the defense is sufficient to shift the burden to the government to prove that joinder will not cause undue harm.[139] At a minimum, the doctrine could be modified to eliminate the judicial requirement of "substantial" or "compelling" prejudice before severance is ordered.[140] This would require some change in the

---

138. *But see* 2 ABA STANDARDS FOR CRIMINAL JUSTICE, *supra* note 31, §§ 13-3.1, 13-1.2 (providing for severance as of right for charges not based on the same conduct or on a common plan); 4 LAFAVE, *supra* note 4, § 17.1(f), at 602–03 & 2005 Supp. at 77 ("Pursuant to statutes and rules of court so providing, and in [two] instance[s] by virtue of a court decision, some states bave now given defendants an absolute right to a severance in certain circumstances." (footnotes omitted)).

139. *See* Dawson, *supra* note 2, at 1453–54 (discussing a similar proposal). A prosecutor might do this, for example, by showing that all of the evidence that would come in at a joint trial would in fact be admitted at separate trials. Although this is similar to the analysis courts now employ, *see supra* note 47 and accompanying text, simply shifting the burden of proof would be significant, since prosecutors typically have greater knowledge of what the evidence will show. The problems of forcing the defendant to speculate about the evidence would thus be reduced. *See supra* note 103 and accompanying text.

140. *See supra* notes 52–53 and accompanying text. The data suggest that courts should be particularly suspicious where a defendant faces joined counts that are of the "same or similar character." Not only are multi-count cases the ones where prejudice is most likely to arise, but charges that are *not* connected by a common plan or scheme or are not part of the same transaction seem to offer far fewer efficiency benefits. A defendant charged with two unrelated bank robberies will not face overlapping witnesses or other evidence, but will run the risk of jury confusion of the evidence that is by definition similar but unrelated. Unconnected charges are also the ones that are most likely to put a defendant in an untenable position of having to choose between testifying in his own defense or not, particularly if defendant is factually guilty of only

case law but not in the Federal Rules; Rule 14 merely indicates that a court should take action if a joinder "appears to prejudice a defendant."

Even if not essential, legislative attention would nonetheless be useful. Rule 14 has not been amended in relevant part since its enactment in 1946,[141] and a refreshed look at severance is past due. But while better rules would help, the solution ultimately must be found in the judiciary. Because joinder is clearly appropriate in many cases and clearly improper in others, some balancing is inevitable and judges are the only ones who can do it. Perhaps it will be sufficient for judges to keep in mind the size of the prejudicial effect of joinder. Perhaps the data presented above will be a caution against assuming too readily that jury instructions and other measures can adequately protect the defendant's interests. Or perhaps a simple judicial refocus on the interests at stake, realizing that defendants' interests may be greater and the government's interest smaller than conventional wisdom suggests, might help avoid the routine approval of consolidated cases, decisions that now often seem reflexive rather than nuanced.

2. Questions

There are two remaining questions that can contribute to the debate, or at least provide direction for future study.

a. Joinder of Defendants

One surprising finding of this study is that joinder of defendants has no measurable impact. We have suggested a partial explanation for it: perhaps the appearance of a joint trial – the fear of guilt by association from having multiple defendants sitting together at counsel table – matters very little to a jury, but hearing additional evidence matters a lot. Because adding defendants necessarily means adding more evidence, it may be that jury confusion about the facts is

---

one charge. For a discussion of the criticisms that have heen directed at the "same or similar character" joinder provision, see 1A C. Wright, *supra* note 9, § 143, at 33–43; *see also* United States v. Foutz, 540 F.2d 733, 738 (4th Cir. 1976) ("[T]he only real convenience served by permitting joint trial of unrelated offenses against the wishes of the defendant may be the convenience of the prosecution in securing a conviction.").

141. A 1966 amendment added what is now subsection (b) to Rule 14, which provides that before ruling on a motion to sever, the trial court can order the prosecutor to provide for *in camera* review any statements by the defendant that the government intends to use as evidence. FED. R. EVID. 14(b). Amendments made in 2002 were stylistic only. *See* Advisory Committee Notes to Rule 14, *reprinted in* 3C WRIGHT ET AL., *supra* note 9, App. C, at 228.

causing the prejudice, not the optics of a multi-party case.[142] If the joinder effect really is about the evidence rather than the appearances of a joint trial, the lack of impact caused by co-defendants is easier to understand.

But there is an alternative explanation. Perhaps the impact of this type of joinder varies according to the circumstances of the defendant. Perhaps if the accused is the *less or least* culpable member of the joined group, he is actually better off than if he were being tried alone because the jury will be most interested in the real bad actors and more inclined to give the small players the benefit of the doubt. The obverse is that *more* culpable defendants are in fact prejudiced by the joinder for the same reason. Thus, our statistical conclusions that show no impact of adding defendants may show nothing more than that there is no effect *on average*. Our analysis might mask the fact that half of the defendants (the less culpable half) could be better off because of joinder, while the more culpable half is worse off. We do not know if this theory is correct, but testing this hypothesis could be a productive area of additional study.

b. Plea Bargaining

Any analysis of the criminal process would be incomplete without some discussion of plea bargaining. Another puzzling feature of the data is that a defendant's joinder status does not appear to influence whether he takes a plea or goes to trial. As noted, whether a defendant has joined counts, joined co-defendants, both, or neither, seems to make no real difference: between 86% and 91% of defendants in all groups plead guilty, with those in the most favored group just as likely to plead as those facing multiple counts along with co-defendants.[143] This is counterintuitive: we wrongly hypothesized that as the chances of conviction increased, the incentive to plead would also increase as the expected value of a trial went down.

The fact that the plea rate remains constant may just be evidence that parties lack information. Perhaps defense lawyers assume there is a vague risk of prejudice from joint trials, but either underestimate the effect or do not sufficiently account for it in their decisionmaking. On the other hand, the guilty-plea data may show that the parties are working with perfect information. Perhaps both prosecutors and defendants know precisely how prejudicial joinder is, but the prosecutor adjusts her plea offer accordingly. As the likelihood

142. *See supra* notes 70–72 and accompanying text.

143. *See supra* Table H and accompanying text (examining the rate of guilty pleas based on joinder status).

of conviction at trial changes, the prosecutor increases or decreases the attractiveness of the inducement, fine-tuning the negotiations until collectively about 90% of all defendants plead. Given that prosecutors and many defense attorneys are the prototypical repeat players, and given the caseload pressures outlined above, this latter explanation seems quite plausible. But this is speculation only, and further research about plea negotiation practices might shed useful light on the topic.

V. CONCLUSION

Our study shows that the joinder of charges has a prejudicial effect on the defendant, increasing the chances of conviction of the most serious charge by more than 10%. Somewhat surprisingly, the data do not show a similar prejudicial effect for joinder of defendants, although more research is needed to determine if our figures are hiding potentially large differences in impact among sub-groups of defendants.

In addition to offering empirical findings, we have argued that courts currently miscalculate the defense and governmental interests when deciding whether to sever. In particular, we claim that judges undervalue the defendants' interests, in part because they may assume that defendants are trying to gain some tactical or unfair advantage from separate trials. This fear seems exaggerated; we could not identify any illegitimate reason why a defendant would seek severance. At the same time, the fear of increased trials in a world of more generous severance also seems overdone. More severance would undoubtedly lead to some increase in the number of proceedings, but given the predictive effect of the first trial outcome on the rest of the cases, it is questionable whether all, or even most, subsequent trials would find their way to the courtroom.

At a minimum, the existence of the joinder effect calls for a reexamination of the doctrine. A significant increase in the likelihood of conviction, one that appears to be based on factors that detract from the fairness of the proceedings and even the accuracy of the outcome, is a steep price to pay for more efficient proceedings. The critical need to reduce those costs without materially undermining the benefits of consolidation means that even dry and technical issues, like joinder and severance, warrant another look.

APPENDIX

This Probit Model is a continuation of the ones in Tables G and J in the main text. The individual models within the Appendix Table control for the cumulative effect of the variables discussed in Part III(B). Model 5 analyzes the impact of the seriousness of the crime, Model 6 the type of crime, Model 7 the type of defense counsel, and Model 8 the judicial circuit where the trial occurred. Each model includes the effects of all prior variables, including the number of counts, number of defendants, and factfinder.

Probit Model of guilty/not guilty at trial on the most serious charge (dependent variable: Guilty[1])				
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Counts	0.320 (0.0238)* [0.000]**	0.340 (0.0242) [0.000]	0.326 (0.0244) [0.000]	0.321 (0.0247) [0.000]
Number of defendants	-0.057 (0.0268) [0.033]	-0.049 (0.0276) [0.073]	-0.081 (0.0280) [0.004]	-0.063 (0.0283) [0.025]
Factfinder (=1 if jury)	0.848 (0.0353) [0.000]	0.855 (0.0361) [0.000]	0.800 (0.0367) [0.000]	0.808 (0.0371) [0.000]
Offence level dummy (=1 if misdemeanor)	-2.040 (0.0713) [0.000]	-2.048 (0.0712) [0.000]	-1.759 (0.0781) [0.000]	-1.597 (0.0856) [0.000]
Offence level dummy (=1 if felony)	-1.727 (0.0755) [0.000]	-1.871 (0.0796) [0.000]	-1.894 (0.0814) [0.000]	-1.722 (0.0880) [0.000]
Offence type dummy (=1 if violent)	-	-0.735 (0.3529) [0.037]	-0.728 (0.3558) [0.041]	-0.722 (0.3538) [0.041]
Offence type dummy (=1 if property)	-	-0.730 (0.3514) [0.038]	-0.662 (0.3545) [0.062]	-0.651 (0.3525) [0.065]
Offence type dummy (=1 if drug)	-	-0.610 (0.3511) [0.082]	-0.582 (0.3541) [0.100]	-0.546 (0.3522) [0.121]
Offence type dummy (=1 if regulatory)	-	-1.242 (0.3569) [0.001]	-1.165 (0.3599) [0.001]	-1.147 (0.3580) [0.001]
Offence type dummy (=1 if public order)	-	-0.823 (0.3518) [0.019]	-0.645 (0.3552) [0.070]	-0.616 (0.3533) [0.081]

Variable (*continued*)	Model 5	Model 6	Model 7	Model 8
Offence type dummy (=1 if weapon)	-	-0.733 (0.3518) [0.037]	-0.694 (0.3547) [0.051]	-0.695 (0.3528) [0.049]
Offence type dummy (=1 if immigration)	-	-0.333 (0.3554) [0.348]	-0.294 (0.3582) [0.412]	-0.254 (0.3564) [0.476]
Lawyer type dummy (=1 if retained)	-	-	0.525 (0.0518) [0.000]	0.394 (0.0549) [0.000]
Lawyer type dummy (=1 if public defender)	-	-	0.565 (0.0544) [0.000]	0.449 (0.0575) [0.000]
Lawyer type dummy (=1 if panel attorney)	-	-	0.786 (0.0536) [0.000]	0.626 (0.0570) [0.000]
Circuit dummy (=1 if CA1)	-	-	-	0.115 (0.1468) [0.435]
Circuit dummy (=1 if CA2)	-	-	-	-0.075 (0.1404) [0.596]
Circuit dummy (=1 if CA3)	-	-	-	-0.025 (0.1424) [0.860]
Circuit dummy (=1 if CA4)	-	-	-	0.256 (0.1371) [0.062]
Circuit dummy (=1 if CA5)	-	-	-	-0.113 (0.1370) [0.410]
Circuit dummy (=1 if CA6)	-	-	-	-0.091 (0.1390) [0.515]
Circuit dummy (=1 if CA7)	-	-	-	0.433 (0.1474) [0.003]
Circuit dummy (=1 if CA8)	-	-	-	-0.146 (0.1399) [0.296]

Variable (*continued*)	Model 5	Model 6	Model 7	Model 8
Circuit dummy (=1 if CA9)	-	-	-	0.017 (0.1374) [0.899]
Circuit dummy (=1 if CA10)	-	-	-	-0.209 (0.1433) [0.144]
Circuit dummy (=1 if CA11)	-	-	-	-0.236 (0.1357) [0.082]
Constant	1.734 (0.0647) [0.000]	2.547 (0.3578) [0.000]	1.957 (0.3642) [0.000]	1.932 (0.3875) [0.000]
Number of observations	19057	19057	19057	19057
Pseudo R-Squared	0.1525	0.1592	0.1707	0.1827
Log likelihood	-8795.6	-8726.2	-8607.1	-8482.1

* Standard errors in parentheses

** Numbers in brackets show the probability of coefficient being outside the 95% confidence interval.

(1) It is equal to one if the defendant is convicted, zero otherwise.