

Essay

POLICING POLITICS AT SENTENCING

Stephanos Bibas, Max M. Schanzenbach,** & Emerson H. Tiller****

INTRODUCTION.....	1371
I. THE STRUCTURE AND POLITICS OF SENTENCING	1374
A. <i>Sentencing Guidelines as Political Compromise</i>	1374
B. <i>The Structure of the Federal Guidelines</i>	1374
C. <i>Judicial Politics and Obeying the Guidelines Bargain</i>	1377
II. SUPREME COURT DOCTRINE: IMPLICATIONS FOR THE EVOLVING REASONABLENESS STANDARD	1380
A. <i>Reasonableness Review of Within-Guidelines Sentences</i>	1381
B. <i>Appellate Review of Departures</i>	1385
C. <i>Policy Preferences as Grounds for Departures and Variances</i>	1388
III. POLITICAL AND INSTITUTIONAL REFORM	1391
A. <i>Political Alignments in Judicial Hierarchy</i>	1391
B. <i>Regulating Appeal Waivers</i>	1394
CONCLUSION.....	1396
APPENDIX	1397

INTRODUCTION

Sentencing guidelines are far better than the indeterminate sentencing that preceded them. Guidelines, enforced by appellate review, discipline sentencing courts and ensure that they consider the legislature's or sentencing commission's policy choices. The threat of reversal is a key component of guidelines. Not only do reversals alter sentences, but more importantly the threat of reversal constrains sentencing courts *ex ante*. This discipline and consistency make criminal sentencing more transparent and legitimate. The Federal Sentencing Guidelines, we acknowledge, may be too harsh. However, undermining the Guidelines system because sentences seem too

* Professor, University of Pennsylvania Law School. Email: stephanos *dot* bibas *at* gmail *dot* com.

** Benjamin Mazur Professor of Law, Northwestern University School of Law. Email: m-schanzenbach *at* law *dot* northwestern *dot* edu.

*** J. Landis Martin Professor of Law and Business, Northwestern University School of Law. Email: tiller *at* law *dot* northwestern *dot* edu.

long is counterproductive. Binding guidelines still leave room for substantial judicial discretion. Without them, Congress may rely more on mandatory minimum sentences, which can be harsher and even more rigid.

Though sentencing guidelines have many advantages, they are in jeopardy. Beginning with *Apprendi*¹ in 2000, the Supreme Court has interpreted the Sixth Amendment as limiting judicial guideline sentencing. Although juries must find guilt beyond a reasonable doubt, until recently guideline sentencing compelled judges to adjust sentences based on facts that they found by only a preponderance of the evidence. *Apprendi* required that juries, not judges, find the facts needed to increase a defendant's sentence beyond the statutory maximum. *Blakely* extended *Apprendi*, striking down a mandatory state sentencing guideline because it raised sentences based on judicial findings of fact.² In the Court's view, judges may not raise sentencing ranges based on their own factual findings, because doing so effectively punishes defendants for crimes more serious than those found by juries. A year later, *Booker* solved the same problem with the binding Federal Guidelines by making them "effectively advisory."³ In other words, the Federal Guidelines no longer compel judges to give higher sentences based on facts not found by juries, but simply advise them to do so.

But *Booker* left unclear exactly how loose appellate review must be to satisfy the Sixth Amendment. For example, if appellate courts review within-Guidelines sentences more deferentially than outside-Guidelines sentences, then the Federal Guidelines retain at least some binding force. Risk-averse sentencing judges might use within-Guidelines sentences as relatively safe harbors. Too much deference to the Guidelines would in effect make them mandatory again.

In the post-*Booker* cases of *Rita*,⁴ *Gall*,⁵ and *Kimbrough*,⁶ however, the Court continues to equivocate about appellate review. On the one hand, it mandates deferential review for all sentences. Yet, on the other hand, it accords the Guidelines some weight as starting points from which to measure deviation.⁷ Larger deviations from guidelines may receive greater scrutiny,

¹ *Apprendi v. New Jersey*, 530 U.S. 466, 490 (2000) (holding that the Sixth Amendment requires prosecutors to prove any fact that increases a defendant's sentence beyond the statutory maximum, except recidivism, to a jury beyond a reasonable doubt).

² *Blakely v. Washington*, 542 U.S. 296, 303–05 (2004) (extending *Apprendi* to require proof to a jury beyond a reasonable doubt of all facts that raise a defendant's maximum sentence under sentencing guidelines).

³ *United States v. Booker*, 543 U.S. 220, 258–65 (2005) (Breyer, J., remedial majority opinion) (remediating the *Blakely* problem with the Federal Sentencing Guidelines not by requiring juries to find sentencing facts but by making the Guidelines advisory rather than mandatory and replacing de novo appellate review with review for "unreasonableness").

⁴ *Rita v. United States*, 127 S. Ct. 2456 (2007).

⁵ *Gall v. United States*, 128 S. Ct. 586 (2007).

⁶ *Kimbrough v. United States*, 128 S. Ct. 558 (2007).

⁷ See *infra* Part I.B (discussing these three cases).

but how far can one take this principle without violating the Sixth Amendment?

In this Essay, we evaluate the Court's evolving guideline jurisprudence and suggest the best direction for it. We focus on evidence from the Federal Sentencing Guidelines, but most of our lessons and recommendations should translate to state guideline systems as well.⁸ Our approach is pragmatic. In Part I, drawing on positive political theory and empirical evidence of judicial sentencing practices, we contend that sentencing judges should defer significantly to sentencing guidelines. A key benefit of guidelines is that they create sentencing structures that make meaningful appellate review possible. Highly structured, detailed guidelines make decisions more transparent by requiring sentencing judges to provide clear, explicit reasons for increasing or decreasing sentences from prescribed baselines. Guidelines, coupled with a politically diverse judiciary, limit unwarranted policy discretion and make criminal sentencing more uniform and politically legitimate.

In Part II, we consider what sentencing could look like under a system of advisory guidelines. The Court has thus far provided little guidance, apart from requiring appellate courts to review the reasonableness of within-Guidelines sentences. The problem is that for guidelines to have any meaning, within-guideline sentences must receive greater deference than those that depart from guidelines. We suggest a solution to this doctrinal puzzle. Courts could improve guidelines by closely reviewing fact-bound adjustments, which are the most manipulable part of the Federal Guidelines. When sentencing judges use factual findings to adjust sentences significantly, appellate scrutiny should increase. Stricter review of sentencing judges ensures that they defer to guidelines and avoid injecting political ideology or other bias into their factfinding. Less defensible, however, is the Court's approval of expanding Guidelines departures based on policy grounds. Giving sentencing courts policy discretion bespeaks too rosy a picture of the judiciary's competence and makes sentences harder to review by removing objective benchmarks for departures.

Finally, Part III considers more radical reforms. Positive political theory of courts suggests some potentially new structures for appellate review of sentences. As we discuss below, politically mixed appellate panels or sentencing review courts could have a moderating influence by keeping ideology in check. Likewise, restricting waivers of appeals could strengthen appellate review as a check on sentencing-judge variation.

⁸ We capitalize Guidelines when referring specifically to the Federal Sentencing Guidelines, but use lower case when referring more generally to the idea or practice of guideline sentencing in state as well as federal courts. We use the terms trial court, sentencing court, and district court interchangeably to refer to courts that find facts and impose sentence in the first instance, and trial judge, sentencing judge, and district judge to refer to the judges of those courts. Likewise, we use the terms appellate court and circuit court interchangeably to refer to courts that review sentences on appeal, and appellate judge or circuit judge to refer to members of those courts.

I. THE STRUCTURE AND POLITICS OF SENTENCING

A. Sentencing Guidelines as Political Compromise

The Federal Sentencing Guidelines embody political compromises on the appropriate sentencing ranges for various crimes and the factors that should aggravate and mitigate sentences. The Sentencing Reform Act of 1984⁹ acknowledged that politics and ideology influence criminal sentencing by requiring political balance on the United States Sentencing Commission: the Act set up a seven-member commission and mandated that no more than four members could come from the same political party.¹⁰ Knowing that criminal sentencing is highly ideological, Congress set up a mechanism to produce political compromise. The deal that Congress struck included sentencing guidelines that judges would have to apply. The compromise produced a system biased towards higher sentences, as the public and Congress prefer to err on the side of harshness. Implicit in the arrangement was the assumption that the courts would faithfully enforce the Guidelines.

The Guidelines improve criminal sentencing in a variety of ways. First, they set up a zone of decision legitimacy—sentencing ranges—for sentencing courts’ outcomes in individual cases. The recommended sentencing ranges and specified, weighted aggravating and mitigating facts are more legitimate than random or ad hoc case-specific factors.

Second, the Guidelines make appellate review of sentencing manageable. Structured, detailed guidelines are more transparent; sentencing judges must explain clearly why they are raising or lowering sentences. Sentencing facts fall into identifiable, recurring categories that higher courts can use to test sentences against the Guidelines’ goals. The detailed structure of the Guidelines better enables courts to enforce the deals made by Congress and the Sentencing Commission over sentencing policy.

B. The Structure of the Federal Guidelines

The Sentencing Table, reproduced in the Appendix, is the key to understanding Federal Guidelines sentencing.¹¹ This table is no longer mandatory, but even after *Booker*, a sentencing judge determines the recommended ranges in the table in the same way as before.¹² The sentenc-

⁹ Pub. L. No. 98-473, 98 Stat. 1987 (codified as amended in scattered sections of 18 U.S.C. and 28 U.S.C.).

¹⁰ 28 U.S.C. § 991(a) (2006).

¹¹ Post-*Booker*, this basic structure has survived but is no longer mandatory. We discuss the post-*Booker* world in greater detail below, but because our data are pre-*Booker*, we also discuss the pre-*Booker* framework.

¹² For example, Justice Breyer was careful to note that judges must consider the Sentencing Guidelines range and continue to make fact findings in light of the Guidelines, conduct sentencing hearings, and justify departures. *United States v. Booker*, 543 U.S. 220, 259–60, 263 (2005).

ing judge uses the Sentencing Commission's detailed regulations to calculate the defendant's numeric offense level. The crime of conviction determines the base offense level. The judge then applies adjustments, adding or subtracting levels based on various factors, such as the use of a gun, the offender's role in the crime, and his acceptance of responsibility. The offender's previous crimes determine his criminal history category. These two factors—offense level and criminal history—yield a sentencing range expressed in months. These determinations are then subject to appellate review.¹³ The top of the sentencing range is roughly twenty-five percent above the bottom of the range.¹⁴ Under the mandatory system used before *Booker*, appellate courts lacked jurisdiction to review the sentence if a sentencing judge had properly calculated the range and sentenced within that range.¹⁵

Before *Booker*, the Sentencing Reform Act authorized judges to depart from the calculated sentencing range if there was an “aggravating or mitigating circumstance of a kind, or to a degree, not adequately taken into consideration by the sentencing commission in formulating the Guidelines that should result in a sentence different from that described.”¹⁶ Under the Act, a judge must justify a departure by making a statement in open court or in a written opinion.¹⁷ Prosecutors could appeal downward departures and defendants could appeal upward departures.¹⁸ Factors “not ordinarily relevant” could still be considered in departure decisions if they were present to such a degree that the case fell outside “the heartland of the Guidelines.”¹⁹ In effect, a sentencing court was at greater risk of reversal if it used the more extreme departure mechanism to lengthen or shorten the sentence.

These details reflect the political compromise that Congress struck. Congress allocated sentencing discretion among itself, the Sentencing Commission, and the courts.²⁰ The courts retained discretion at three differ-

¹³ 18 U.S.C. § 3742(a) (2006).

¹⁴ While the Sentencing Commission set the offense levels and criminal history categories, the Sentencing Reform Act mandated that the top of each sentencing range be no more than six months or twenty-five percent above the bottom of the Guidelines range, whichever is greater. 28 U.S.C. § 994(b)(2) (2006) (providing also that the maximum may be life where the minimum is at least thirty years).

¹⁵ 18 U.S.C. § 3742 (2006).

¹⁶ *Id.* § 3553(b)(1); *see also* U.S. SENTENCING GUIDELINES MANUAL § 5K2.0 (2008) (grounds for departure policy statement).

¹⁷ 18 U.S.C. § 3553(c) (1994) (before amendment by PROTECT Act).

¹⁸ *Id.*

¹⁹ U.S. SENTENCING GUIDELINES MANUAL, *supra* note 16, § 5K2.0 (grounds for departure policy statement) (as amended Nov. 1, 1994). The Supreme Court later endorsed the heartland concept. *See* *Koon v. United States*, 518 U.S. 81, 97, 109 (1996).

²⁰ Using sentencing guidelines as an example, Steven Shavell offers a formal treatment of optimal discretion within a framework of binding constraints. Steven Shavell, *Optimal Discretion in the Application of Rules*, 9 AM. L. & ECON. REV. 175 (2007). Shavell views the zone of discretion as a tradeoff between the desire to constrain agents (courts) to behave in conformity with the wishes of the principal (Congress) and the benefits of individualized treatment, which require the delegation of discretion. *Id.*

ent levels. First, sentencing judges retained absolute discretion within the sentencing range. Second, sentencing judges could use adjustments, preapproved by the Commission or Congress, to change sentencing ranges. These adjustments were based largely on findings of fact and so received deferential appellate review. Finally, departures let judges apply their traditional sentencing discretion by departing from the Guidelines framework. They could depart, however, only if they found, and the appellate court agreed, that the unique facts fell outside the heartland of typical cases envisioned by Congress and the Commission in the Guidelines.

Booker upended this compromise by declaring the Guidelines advisory while still preserving their constitutionality. Post-*Booker*, sentencing courts are no longer limited to modifying sentences through specified adjustments within the Guidelines or departures regulated by the Guidelines. They may instead simply offer some other reason outside the Guidelines' structure to vary from the Guidelines. In addition, sentences within properly calculated Guidelines ranges are now subject to appellate review. But *Booker* raised more questions than it answered. How should appellate courts review within-Guidelines sentences? How should they review departures and adjustments? *Booker* answered that the new standard was reasonableness in light of the Guidelines but offered little illustration.²¹ The Court added only a little clarity thereafter in *Rita*, *Gall*, and *Kimbrough*.

Rita v. United States held that courts of appeals may presume that within-Guidelines sentences are reasonable because they reflect the judgment and expertise of both the sentencing judge and the Sentencing Commission.²² But did this mean that appellate courts should review out-of-range sentences less deferentially? *Gall v. United States* answered that question, holding that courts of appeals must review sentences both within and without the Guidelines under a deferential abuse-of-discretion standard. Courts should use the Guidelines as their starting point and benchmark, and they may take the degree of variance from the Guidelines into account. Appellate courts may not, however, require extraordinary circumstances or apply rigid mathematical ratios when reviewing deviations from the Guidelines. Doing so would erect an impermissible presumption of unreasonableness and thus make the Guidelines too binding.²³ In sum, the Court equivocated, on the one hand mandating deferential review for all sen-

at 176–78. Thus, the cost to Congress of permitting individualized treatment is that the courts may express policy preferences within the zone that properly belongs to Congress.

²¹ *United States v. Booker*, 543 U.S. 220, 261–63 (2005) (Breyer, J., remedial majority opinion).

²² 127 S. Ct. 2456, 2462, 2464–65 (2007).

²³ *Gall v. United States*, 128 S. Ct. 586, 594–97 (2007). In addition, *Kimrough v. United States* held that a district court could disagree with the policy choices embedded in the Federal Sentencing Guidelines' crack-cocaine guidelines because the Commission itself had not relied on empirical data or experience in promulgating it, had criticized this guideline as overly harsh, and had tried to change it. 128 S. Ct. 558, 567–69, 574–75 (2007).

tences, yet on the other hand according to the Guidelines some weight as starting points from which to measure deviation.

C. *Judicial Politics and Obeying the Guidelines Bargain*

In disrupting the Guidelines system, the Court seemed unconcerned with the problem of interjudge disparity. Congress authorized, and the Commission promulgated, these sentencing rules to reduce the influence of judges' idiosyncrasies on sentences. As Congress recognized, judges matter a great deal in determining final sentences. Judges' sentencing philosophies vary, and these philosophies often reflect their political ideologies and demographic characteristics. These philosophies and characteristics correlate with widely divergent sentences.²⁴

Though unwarranted disparities remain,²⁵ studies suggest that the Fed-

²⁴ A considerable amount of research suggests that judges have different sentencing philosophies and that those philosophical differences influence outcomes. See ANTHONY PARTRIDGE & WILLIAM B. ELDRIDGE, FEDERAL JUDICIAL CENTER, *THE SECOND CIRCUIT SENTENCING STUDY: A REPORT TO THE JUDGES OF THE SECOND CIRCUIT* 36 (1974) (finding differences among judges in the hypothetical sentences they would impose on identical offenders); John S. Carroll et al., *Sentencing Goals, Causal Attributions, Ideology, and Personality*, 52 J. PERSONALITY & SOC. PSYCHOL. 107 (1987) (demonstrating how a judge's ideology is reflected in how she thinks about the causes of crime and the goals of sentencing); Shari S. Diamond & Hans Zeisel, *Sentencing Councils: A Study of Sentence Disparity and Its Reduction*, 43 U. CHI. L. REV. 109, 114 (1975) ("[I]t is reasonable to infer that the judges' differing sentencing philosophies are a primary cause of the disparity."); Brian Forst & Charles Wellford, *Punishment and Sentencing: Developing Sentencing Guidelines Empirically from Principles of Punishment*, 33 RUTGERS L. REV. 799, 813 (1981) (finding that judges oriented toward utilitarian goals of incapacitation and deterrence gave sentences that were at least ten months longer than judges motivated by other goals). See generally Paul Hofer et al., *The Effect of the Federal Sentencing Guidelines on Inter-Judge Sentencing Disparity*, 90 J. CRIM. L. & CRIMINOLOGY 239, 250 (1999) ("[L]iberals' tend to believe that factors external to the offender are responsible for criminal behavior. Rehabilitation is more of a sentencing goal for these judges, leading to greater reliance on probation and less concern with retribution. 'Conservatives' believe that offenders choose to commit crimes. They are more punishment-oriented and tend to impose longer prison terms.").

²⁵ See generally DOUGLAS C. McDONALD & KENNETH E. CARLSON, *SENTENCING IN THE FEDERAL COURTS: DOES RACE MATTER?* 177 (1993); Celesta A. Albonetti, *Sentencing Under the Federal Sentencing Guidelines: Effects of Defendant Characteristics, Guilty Pleas, and Departures on Sentence Outcomes for Drug Offenses, 1991–1992*, 31 LAW & SOC'Y REV. 789 (1997); David B. Mustard, *Racial, Ethnic, and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts*, 44 J.L. & ECON. 285 (2001) (finding that, no matter how one cuts the data—within-Guidelines sentences, departures, drug crimes, etc.—substantial unexplained racial and gender disparities exist); Max M. Schanzenbach & Michael L. Yaeger, *Prison Time, Fines and Federal White-Collar Criminals: The Anatomy of Racial Disparity*, 96 J. CRIM. L. & CRIMINOLOGY 757 (2006); Max M. Schanzenbach, *Racial and Sex Disparities in Prison Sentences: The Effect of District-Level Judicial Demographics*, 34 J. LEGAL STUD. 57 (2005) (using district-level variation in judicial characteristics and finding some evidence that minority and female judges sentence differently). Studies on the effect of judicial characteristics on sentencing at the state level have demonstrated that judicial demographic characteristics matter for sentencing. See, e.g., Darrell Steffensmeier & Chris Hebert, *Women and Men Policymakers: Does the Judge's Gender Affect the Sentencing of Criminal Defendants?*, 77 SOC. FORCES 1163 (1999) (finding, in a study of Pennsylvania criminal cases, that female judges tended to sentence offenders more severely, were more likely to incarcerate minorities, and were less likely to incarcerate women).

eral Sentencing Guidelines have at least slightly reduced interjudge disparities.²⁶ Recent work on state guidelines systems also suggests that mandatory guidelines have significantly reduced sentencing variation.²⁷ In two recent articles, Schanzenbach and Tiller considered how judges' political characteristics interact with the Federal Guidelines' structure. One article analyzed a very large dataset, using changes in the ratio of Democratic appointees versus Republican appointees in each district to identify ideological effects on sentencing.²⁸ Another analyzed a smaller dataset that actually identified the judge's political affiliation by looking at the party of each judge's appointing President.²⁹ Both studies came to remarkably similar conclusions. For serious crimes—robbery, firearms, drug trafficking, extortion, and murder—Democratic appointees' sentences were, on average, seven to eight months lower than Republican appointees' sentences. This discrepancy amounted to roughly ten percent of the average sentence.³⁰

Those studies also considered the effect of appellate review on district court sentencing and drew two conclusions. First, Democratic appointees to district courts were more likely to adjust sentences downward than were Republican appointees, regardless of the politics of the reviewing court. Conversely, Republican-dominated districts were more likely to adjust sentences upward than were districts dominated by Democratic appointees, again regardless of the circuit court's politics. In other words, the preferences of the circuit did not influence offense-level adjustments (within-Guidelines sentencing). Democratic- and Republican-appointed district court judges used offense levels to adjust prison sentences to roughly the same degree whether they were in majority-Democratic or majority-Republican appointed circuits. Judges most likely follow this pattern be-

²⁶ See John M. Anderson et al., *Measuring Interjudge Sentencing Disparity: Before and After the Federal Sentencing Guidelines*, 42 J.L. & ECON. 271, 303–04 (1999) (finding a decrease in interjudge disparities in sentence length after the Guidelines, but cautioning that the advent of mandatory minimum sentences might have contributed to the decline); Hofer et al., *supra* note 24, at 289 (concluding that the Guidelines slightly decreased interjudge sentence disparities.).

²⁷ John F. Pfaff, *The Continued Vitality of Structured Sentencing Following Blakely: The Effectiveness of Voluntary Guidelines*, 54 UCLA L. REV. 235 (2006). Pfaff's findings indicate that voluntary guidelines, such as those adopted in some states, can reduce sentencing variations. The results suggest, however, that reduction in variance for violent and property crimes is much greater for mandatory guidelines than for voluntary ones, though drug crimes may not follow this pattern. See *id.* at 263–67 & tbls.3–4. Though suggestive, Pfaff's study cannot support strong inferences, as only two states adopted voluntary guidelines during the time frame of his sample. *Id.* at 256–57.

²⁸ Max M. Schanzenbach & Emerson H. Tiller, *Strategic Judging Under the Sentencing Guidelines: Positive Political Theory and Evidence*, 23 J.L. ECON. & ORG. 24 (2007).

²⁹ Max M. Schanzenbach & Emerson H. Tiller, *Reviewing the Sentencing Guidelines: Judicial Politics, Empirical Evidence, and Reform*, 75 U. CHI. L. REV. 715 (2008).

³⁰ The studies did not find any differences for white-collar offenses such as fraud, tax, embezzlement, and antitrust crimes. But because of low offense levels and low sentences, there were technical problems that prevented us from reaching a conclusion about less serious crimes.

cause fact-oriented adjustments received more deferential appellate review.³¹

On the other hand, Democratic district court appointees in majority-Democratic circuits departed to a degree greater, in terms of length of sentence, than Democratic appointees in majority-Republican circuits. In addition, we found some evidence that departures were more likely in these cases as well. This result was in line with our theoretical prediction: because appellate courts review law-oriented departures more stringently than fact-oriented adjustments, the reviewing court's political alignment was more important for departures. The converse position did not hold: Republican-appointed district court judges did not depart upwards more frequently in Republican-dominated circuits.³² This result makes sense because the Guidelines encourage upward adjustments and have high presumptive sentences to begin with, so usually there is little need to depart upwards. In addition, this reasoning explains why Republican-appointed district court judges did not have to give lower sentences in Democratic-dominated circuits: they could rely on the high presumptive sentences, upward adjustments, and the safe harbors provided by the Guidelines. Note that upward departures were rare before *Booker*,³³ suggesting that judges who wanted to raise sentences found it easy to do so within the Guidelines. The asymmetrical ease of increasing sentences within the Guidelines has some important implications, which we will explore later.

The studies above illustrate how vertical political alignment between courts—alignment between lower and higher courts—can increase or decrease transparency and sentencing consistent with the Guidelines. Non-sentencing research also suggests that horizontal political alignment on three-judge appellate panels can influence obedience to legal doctrine. Specifically, Cross and Tiller found that circuit court panels were more likely to ignore the *Chevron* doctrine in administrative law cases when (1) the three judges were all appointed by Presidents of the same political party, and (2) the doctrine ran counter to their political preferences. When the panels were diverse, judges obeyed legal doctrine more, whether or not they favored the resulting outcomes.³⁴

Thus, where there is a clear rule, a political minority can invoke it to produce the outcome intended by the rulemaker. But where there is no clear rule or only a vague standard, the majority can more easily dominate the outcome with its own preferred policy choice. The political minority can

³¹ Schanzenbach & Tiller, *supra* note 28; Schanzenbach & Tiller, *supra* note 29.

³² Schanzenbach, & Tiller, *supra* note 28, at 47–52.

³³ In fiscal years 1998 through 2002, before the PROTECT Act took effect, the rate of upward departures varied between 0.6% and 0.8%. U.S. SENTENCING COMM'N, 2002 SOURCEBOOK OF FEDERAL SENTENCING STATISTICS fig.G (2002), available at <http://www.ussc.gov/ANNRPT/2002/fig-g.pdf>.

³⁴ Frank B. Cross & Emerson H. Tiller, *Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Court of Appeals*, 107 YALE L.J. 2155 (1998).

do little more than engage in hard-knuckled political compromises, log-rolling decisions, or potentially disruptive dissents.

The same logic would likely apply to the sentencing policy compromises hammered out in the Sentencing Guidelines. The appellate panel is more likely to check sentencing court abuses if (1) the sentencing rules are clear, and (2) the appellate court has at least one member whose politics differ from those of the sentencing court—in other words, if the panel itself is politically diverse. If the sentencing or appellate judge were to slant the Guidelines toward his political preferences, at least one appellate judge with different political preferences could act as a whistleblower. Political diversity thus checks Democratic appointees who would stretch the Guidelines to impose lower sentences and Republican appointees who would stretch the Guidelines to impose higher sentences.³⁵

For guidelines to deter evasion and manipulation, however, they must bind both sentencing and appellate courts. The more discretion appellate courts give sentencing courts to apply sentencing guidelines, the less guidelines will ensure consistency and legitimacy. Put differently, ideological forces and judicial discretion will push sentences to high and low extremes, especially when there is no political diversity to keep judges honest.

We now turn to the Supreme Court's post-*Apprendi* sentencing jurisprudence. The next Part analyzes the Court's treatment of adjustments and departures after *Booker*. The Court's reasonableness review standard for within-Guidelines sentences may improve the Guidelines, but the Court's increasing acceptance of departures is likely to undermine the legitimacy of the Guidelines. Most distressing is the new policy-grounds departure, which circuit courts will find hard to review.

II. SUPREME COURT DOCTRINE: IMPLICATIONS FOR THE EVOLVING REASONABLENESS STANDARD

To enforce the Guidelines' political compromise and reduce sentencing disparities, courts need structures that foster robust appellate review. The main issue raised by *Booker* and its progeny is precisely how stringently appellate courts should review departures and within-Guidelines sentences. The Court's answer so far is hard to decipher. *Booker*, *Rita*, and especially *Gall* signal that sentencing courts should have greater discretion in sentencing, which means that they should be freer to depart from the Guidelines. However, these cases require courts of appeals to review sentences for reasonableness even if sentencing courts have calculated sentences properly within the Federal Guidelines. Thus, the Court has both expanded sentencing courts' discretion to depart from the Guidelines and regulated their formerly absolute discretion to sentence anywhere within the Guidelines.

³⁵ For a formal theoretical treatment of appeals in a discretionary system, see Shavell, *supra* note 20, at 187–89.

Indeed, Justice Stevens's concurrence in *Rita* carefully noted that the review of within-Guidelines sentences, while deferential, should be meaningful and substantive, not purely procedural.³⁶ How can one possibly make sense of this doctrinal tangle? Why would the Court emphasize judicial discretion in one sphere while seemingly constricting it in another?

One might perhaps view the Court's apparent inconsistencies as an effort to address the asymmetries of the pre-*Booker* sentencing regime. Before *Booker*, most offense-level adjustments pushed sentences upwards, and base offense levels often started out high. As a result, the Sentencing Guidelines made it pretty easy for sentencing courts to increase sentences via offense-level adjustments and harder for appellate courts to review these adjustments. In addition, the sentencing court's refusal to grant a downward departure was basically unreviewable.

Apprendi and *Booker* reflected the Court's unease with this prosecution asymmetry. The constitutional problem, after all, is with sentencing courts' raising sentences based on facts not proven to juries. Because appellate review is so important, as discussed earlier, expanding appellate review to include within-Guidelines sentences makes some sense. We are much more skeptical, however, of reducing the scrutiny of departures, in particular the creation of fuzzy, hard-to-review policy departures.³⁷

A. Reasonableness Review of Within-Guidelines Sentences

The Federal Guidelines pre-*Booker* made it very easy to raise sentences via adjustments or departures but left less room to reduce them. Almost all departures explicitly permitted in the *United States Sentencing Guidelines Manual* are upward departures, such as departures based on dangerousness, a crime resulting in death, or inadequacy of criminal history.³⁸ Downward departures are generally discouraged, as are departures based on factors already accounted for by the offense level, such as acceptance of responsibility.³⁹ In addition, the vast majority of adjustments allowed by the Guidelines are upward adjustments.⁴⁰ These factors combine with the Federal Guidelines sentencing table, which increases sentences exponentially by offense level, as the following graph shows:

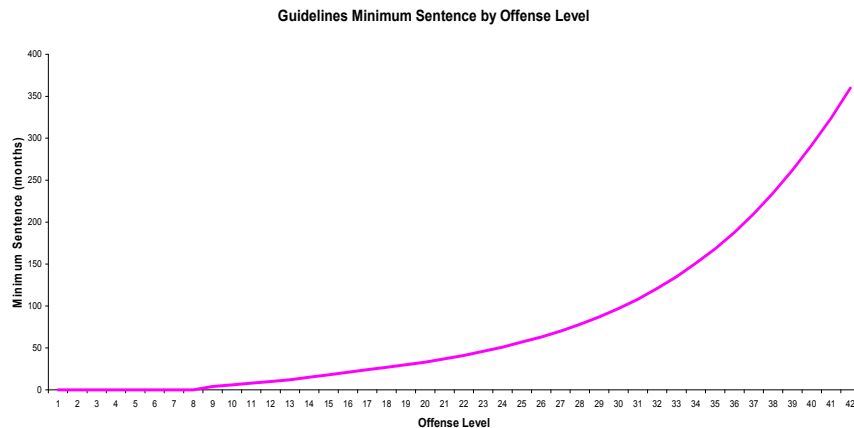
³⁶ *Rita v. United States*, 127 S. Ct. 2456, 2473–74 (2007) (Stevens, J., concurring).

³⁷ Granted, appeal waivers in plea bargains threaten to undercut appellate review's efficacy as a guarantor of consistency. We grapple with this problem and possible solutions to it later in this Essay. See *infra* Part III.B.

³⁸ U.S. SENTENCING GUIDELINES MANUAL, *supra* note 16, §§ 4A1.3, 5K2.1, 5K2.6. Courts may also depart downward for overstatement of criminal history, but unlike upward departures, these downward departures face a lower bound. See *id.* § 4A1.3(b)(2)–(3).

³⁹ *Id.* §§ 5K2.0(2), 5K2.20.

⁴⁰ See, e.g., *id.* ch. 3 (listing ten upward adjustments, including the rules for increasing sentences for multiple counts, but only two downward adjustments, for minor or minimal role in the offense (§ 3B1.2) and acceptance of responsibility (§ 3E1.1)).



One of the strongest substantive objections to the Federal Guidelines is to their asymmetry. This asymmetry may have been at play in *Apprendi* and *Blakely*. Indeed, those cases limited both upward departures and adjustments but did not affect sentence reductions. In practice, however, these cases likely did not constrain prosecutors much, as creative prosecutors can construct multicount indictments to raise defendants' statutory and guidelines maxima. In addition, the higher penalties associated with higher offense levels increased judicial discretion to raise sentences. Judges inclined to raise offenders' sentences may do so through upward adjustments, which, as detailed above, are harder to review on appeal. This factor may partly explain why upward departures are more rare than downward departures. Instead of departing upward, a judge may instead choose to adjust upward.

A judge's ability to raise sentences by using adjustments, which are not subject to substantial appellate scrutiny, will increase sentencing disparities. Indeed, our studies found that Republican and Democratic appointees differed consistently on adjustments and that appellate courts did not affect these differences much.⁴¹ Of course, in theory the pre-*Booker* Guidelines compelled judges to adjust upward when the necessary facts were proven by a preponderance of the evidence. In practice, however, many cases left the sentencing judge leeway to make factual findings that faced only limited appellate scrutiny, especially in close cases.

Appellate courts, we suggest, should interpret *Booker* and *Rita*'s reasonableness review of within-Guidelines sentences as a directive to review adjustments a little more rigorously. This is a little different from determining whether the within-Guidelines sentence is reasonable. The advantage of reviewing adjustments more rigorously is that specific guideline factors guide appellate review. In addition, instead of the pre-*Booker* approach of adjustment-by-adjustment review, the reasonableness standard may permit

⁴¹ Schanzenbach & Tiller, *supra* note 28, at 45–48.

appellate courts to review sentences based on the cumulative effect of adjustments. For example, imagine that in a fairly standard drug trafficking case a court piles on adjustments for a large drug quantity, major role in the offense, and other potentially relevant factors. Because the adjustments lead to a much higher sentence, both the factual and legal findings underpinning the adjustments deserve a close look on appeal.

In addition to reviewing adjustments cumulatively, appellate courts should pay special attention to adjustments that raise or lower sentences substantially. Among these are adjustments for career offenders and adjustments for substantial monetary losses or drug amounts.⁴² In addition, departures that are encouraged by the Guidelines, which have previously received more deferential review,⁴³ should also get a closer look. Key examples are departures based on the inadequacy of the criminal history or because the crime caused a death not otherwise reflected in the sentence.⁴⁴ These high-stakes departures generally increase sentences and can do so dramatically. Because the stakes are so high in these cases, appellate courts ought to take a closer look instead of deferring.

A closer appellate look in these instances should counteract the political extremes discussed earlier. Three-judge panels will be less likely to reflect extreme views than a single judge. A randomly chosen three-judge panel may well have a mix of judges with different perspectives. This diversity will moderate both the panel's deliberation and the district court's idiosyncrasies, especially if the two courts differ in their politics or ideology. While appellate courts inevitably defer somewhat to the sentencing judge who heard the evidence, they may still check decisions and improve consistency. Because adjustments and encouraged departures more often increase sentences, reasonableness review of within-Guidelines sentences will encourage judges to tie enhancements more closely to the facts and the crime of conviction. Thus, stricter review of adjustments may best address *Blakely's* and *Booker's* constitutional concerns.

⁴² See, e.g., U.S. SENTENCING GUIDELINES MANUAL, *supra* note 16, § 4B1.1 (providing for a career offender adjustment that creates a new offense level based on the statutory maximum); *id.* § 2T4.1 (defining tax loss adjustments that vary offense levels from six to thirty-six); *id.* § 2B1.1 (defining loss adjustments for frauds and thefts that vary offense levels from six to thirty-six); *id.* § 2D1.1 (varying drug offense levels from six to thirty-eight depending on quantity of drugs).

⁴³ In *United States v. Koon*, 518 U.S. 81 (1996), the Court clearly specified a lower standard of review for encouraged factors, holding that if “the special factor is an encouraged factor, the court is authorized to depart if the applicable Guideline does not already take it into account.” *Id.* at 96. However, if the factor is a discouraged factor, “the court should depart only if the factor is present to an exceptional degree.” *Id.*

⁴⁴ For example, in *United States v. Mayle*, 334 F.3d 552 (6th Cir. 2003), defendant Mayle was convicted of various counts of wire fraud and forgery relating to his cashing of social security checks. The original sentencing range was fifteen to twenty-one months, but the sentencing judge attributed three murders for which he was not convicted (two in furtherance of his fraud). The judge departed upward 339 months based both on inadequacy of the criminal history and death caused by the crime. *Id.* at 555–57.

How exactly would a system of robust adjustment review work? In *McMillan*, an early state sentencing case, the Supreme Court suggested in dicta that a higher standard of review would apply to factual findings when they become the “tail wagging the dog of the substantive offense.”⁴⁵ Before *Booker* and *Blakely*, the Third Circuit had adopted a “clear and convincing” standard of review for fact findings that dramatically increased the overall sentence.⁴⁶ A few other circuits toyed with the same idea,⁴⁷ but most circuits clearly rejected this view.⁴⁸ *Blakely* considered the *McMillan* approach when dealing with state guideline sentencing enhancements, but rejected it in favor of a bright-line approach rejecting any mandatory enhancements to sentences based on facts not determined by a jury.⁴⁹ The Third Circuit has recently overruled its clear-and-convincing standard based on the Supreme Court’s reasoning in *Blakely*.⁵⁰

We think, however, that the Third Circuit erred in interpreting *Blakely*, especially in light of *Booker* and *Rita*. *Blakely* was concerned with a mandatory system of sentencing, and there the Court rejected shifting standards of review as a basis for preserving a mandatory system. The Court was not addressing appellate review in the current system of Guidelines, which are no longer mandatory but must carry some weight.

Perhaps appellate courts are uncomfortable scrutinizing fact-intensive adjustments closely. The Third Circuit’s pre-*Booker* clear-and-convincing approach may not have caught on for this very reason. Searching appellate review of adjustments may seem difficult, if not impossible, because sentencing judges are uniquely positioned to find facts and weigh their relevance and credibility. But appellate courts have not been so reluctant to review fact-bound determinations in other areas of law. For example, in administrative law, courts engage in sweeping review of the facts deter-

⁴⁵ *McMillan v. Pennsylvania*, 477 U.S. 79, 88 (1986).

⁴⁶ In *United States v. Kikamura*, 918 F.2d 1084 (1990), the Third Circuit held that an extreme increase in sentence, from thirty months to thirty years, required proof by “clear and convincing” evidence.

⁴⁷ See, e.g., *United States v. Hopper*, 177 F.3d 824, 833 (9th Cir. 1999) (suggesting that proof by clear and convincing evidence may be necessary “when a sentencing factor has an extremely disproportionate effect on the sentence relative to the offense of conviction”); *United States v. Townley*, 929 F.2d 365, 369–70 (8th Cir. 1991) (discussing but not applying the clear-and-convincing standard because the evidence at hand failed an even lower standard).

⁴⁸ See *United States v. Graham*, 275 F.3d 490, 517 n.19 (6th Cir. 2001) (reasoning that judges have discretion within *Apprendi* limits, so no greater fact-based review is required); *United States v. Cordoba-Murgas*, 233 F.3d 704, 708–09 (2d Cir. 2000) (rejecting clear-and-convincing evidentiary standard at sentencing for uncharged murders); *United States v. Valdez*, 225 F.3d 1137, 1143 n.2 (10th Cir. 2000) (rejecting clear-and-convincing evidence standard at sentencing even though relevant conduct dramatically increased sentence).

⁴⁹ *Blakely v. Washington*, 542 U.S. 296, 307–08 (2004).

⁵⁰ *United States v. Grier*, 449 F.3d 558, 570 (3d Cir. 2006).

mined by expert agencies,⁵¹ but defer more to agency findings of law under *Chevron*.⁵² The administrative law emphasis is quite the opposite of current sentencing review. Of course, the constitutional implications of criminal sentencing and administrative agency review are quite different. But institutional competency—the ability of a higher court to delve into the lower court’s or agency’s factfinding—does not dictate automatic deference to findings of fact in criminal sentencing.

To date, appellate courts have not taken advantage of Justice Stevens’s invitation to review within-Guidelines sentences substantively. Indeed, it was more than two and a half years after *Booker* before an appellate court reversed a within-Guidelines sentence as substantively unreasonable.⁵³ Perhaps reversals are so rare because the Court has provided too little guidance for substantive reasonableness review. However, greater appellate review of adjustments provides clear guidance and will help to make sentencing more consistent by policing overly long sentences.

In short, we suggest more aggressive appellate review of fact-based adjustments, particularly those that raise or lower sentences substantially. Our proposed approach has the benefit of preserving much of the political compromise that the Guidelines represent, while at the same time expanding the protection that appellate review provides. We do not know how many sentences our proposed review standard would affect. However, given the tens of thousands of federal criminal sentences since *Booker*, surely more than one merited reversal as an unreasonable within-Guidelines sentence.

B. Appellate Review of Departures

The basic structure of the Federal Guidelines requires appellate courts to scrutinize departures more closely than within-Guidelines sentences. If one wishes to transfer discretion to sentencing courts, one must give them greater leeway to depart. As suggested by the *Gall* Court, one way to do this is to review departures more deferentially. Another way is to increase

⁵¹ The Administrative Procedures Act requires application of the “substantial evidence test” in review of formal adjudications and formal rulemakings, 5 U.S.C. §§ 554, 556, 557 (2006), and de novo review in very limited circumstances, *id.* § 706(2)(E)–(F). Informal adjudications and informal rulemakings are subject to the arbitrary-and-capricious standard of review, which courts have generally interpreted as hard-look review. See *Motor Vehicle Mfrs. Ass’n v. State Farm Mut. Auto. Ins. Co.*, 463 U.S. 29, 51–52 (1983); *Ass’n of Data Processing Serv. Orgs. v. Bd. of Governors of the Fed. Reserve Sys.*, 745 F.2d 677, 683 (D.C. Cir. 1984).

⁵² *Chevron U.S.A., Inc. v. Natural Res. Def. Council, Inc.*, 467 U.S. 837 (1984).

⁵³ *United States v. Paul*, 239 F. App’x 353, 354–55 (9th Cir. 2007) (unpublished per curiam); Douglas A. Berman, *Ninth Circuit Reverses Within-Guideline Sentence as Substantively Unreasonable!*, *Sentencing Law and Policy*, Aug. 17, 2007, http://sentencing.typepad.com/sentencing_law_and_policy/2007/08/ninth-circuit-r.html (“A full 31 months and five days since *Booker*, today has brought what I believe constitutes the very first appellate reversal of a within-guideline sentence as substantively unreasonable.”).

the costs of manipulating adjustments within the Guidelines, by for example increasing the risk of appellate reversal. This insight helps to explain the presumption-of-reasonableness battle in *Rita*. Justice Breyer and the pro-Guidelines faction may have been happy to save the Guidelines by making departures a bit easier. Nonetheless, stricter appellate review of departures than of within-Guidelines sentences preserves much of the structure of the Guidelines.

The Court's combined approach, reducing scrutiny of departures while increasing scrutiny of within-Guidelines sentences, blurs the distinction between trial courts' sentencing instruments. On the one hand, sentencing courts can no longer rely as easily on Guidelines factors in sentencing—though in practice they continue to do so. On the other hand, sentencing courts are less likely to be reversed for departures under *Gall*'s deferential abuse-of-discretion standard.⁵⁴ Although courts of appeals may presume within-Guidelines sentences to be reasonable, they may not presume outside-Guidelines sentences unreasonable.⁵⁵ Thus, sentencing courts will rely more on departures and less on adjustments than they used to. That does not mean that they will use departures a great deal more or that adjustments are now as risky as departures. It simply means that *Booker*, *Rita*, and *Gall* changed the relative cost of using one or the other. In addition, the political context may amplify these effects. Political diversity in the courts will lose its moderating influence on sentences, as appellate courts cannot enforce the zone of decision legitimacy established by the Guidelines' sentencing ranges. The power of political diversity in a judiciary depends on clear legal boundaries such as guidelines. At best, political diversity without guidelines means hard-knuckled, case-by-case, ideological compromises. Guidelines give dissenters a tool to compel compliance with the law, curbing politicized deviations from the law.

This state of affairs is troubling. When appellate courts relax review of departures, they allow more variance in otherwise normal guidelines cases. To be sure, post-*Booker* empirical evidence suggests that federal sentencing practices have not changed much. Average sentence length and probability of imprisonment actually increased slightly. However, judge-induced downward departures and variances have increased by nearly fifty percent, from 8.6% to 12.5% of all sentences.⁵⁶ Upward departures doubled, from 0.8% to 1.6%.⁵⁷

⁵⁴ *Gall v. United States*, 128 S. Ct. 586, 596 (2007) (“[T]he abuse-of-discretion standard of review applies to appellate review of all sentencing decisions—whether inside or outside the Guidelines range.”).

⁵⁵ *Id.* at 595 (quoting *Rita v. United States*, 127 S. Ct. 2456, 2467 (2007)).

⁵⁶ The effect of *Booker* depends very much on what the relevant comparison period is. The PROTECT Act was effective in the year and a half before *Booker*, and it restricted the use of downward departures and encouraged prosecutors to appeal them more often. The Sentencing Commission's statistics indicate that in the eight months before the PROTECT Act, the average judge-induced departure rate was 8.6%; the rate fell to 5.5% while the PROTECT Act was in effect, but increased to 12.5% (compris-

These facts are consistent with two different explanations. First, although average sentences changed little, interjudge variance may have increased. Those judges who are increasing sentences may offset those who are decreasing sentences. In other words, average sentences have not changed because both upward and downward departures increased. Similarly, circuits that adopt strong presumptions of reasonableness may effectively treat the Guidelines as safe harbors. In aggregate statistics, these cautious circuits may mask important changes in other circuits.

The second possible explanation is that district judges' sentencing patterns have not changed much post-*Booker*. However, because the Court has blurred the distinction between sentencing instruments, district court judges merely choose to use departures more often. In other words, it may now be easier and less risky for district court judges to get where they want to go using departures instead of adjustments. Of course, we would have to do a much more detailed analysis before we could conclude that nothing significant has changed. More importantly, the decisions in *Rita* and *Gall* are too recent to be reflected in the data. Those who deplored the Guidelines for their harshness, however, can take no comfort from the data so far. Sentences have not decreased, and upward departures have increased.

There is some room for greater use of departures within the framework we have laid out. One reason to depart is because the punishment mandated by the Guidelines does not seem to fit the crime. This may occur when a series of adjustments results cumulatively in an unusually harsh sentence. Indeed, we suspect that such circumstances lie behind many departures, both pre- and post-*Booker*. Before *Booker*, sentencing courts had to shoe-horn such cases into a permitted or discouraged ground. Prior doctrine did not permit departures solely because the sentence seemed unreasonable. For example, courts may have invented the aberrant-act departure to reduce sentences for defendants who seemed neither especially blameworthy nor dangerous but faced long Guidelines sentences. *Booker* and its progeny may have had such cases in mind. Thus, a reasonable new ground for departure may be one in which the facts compel a large number of upward adjustments, rendering the Guidelines range inappropriate. This ground is much more limited than the reasonableness standard the Court lays out, however, because it still anchors sentences to the Guidelines. In addition, it would not permit the judge to inject his policy preferences into the sentence overtly, an issue to which we now turn.

ing 3.2% downward departures plus 9.3% *Booker* variances) in the year following *Booker*. U.S. SENTENCING COMM'N, FINAL REPORT OF THE IMPACT OF *UNITED STATES V. BOOKER* ON FEDERAL SENTENCING app. E, at E-1 (2006).

⁵⁷ *Id.*

C. Policy Preferences as Grounds for Departures and Variances

While fact-based adjustments and departures deserve some leeway, as discussed above, policy-based ones do not. An individual judge's policy preferences should not influence sentencing. Those idiosyncratic policy preferences are not relevant to society's decision to punish, incapacitate, or deter criminals. In addition, courts lack the institutional competence to make systemic policy choices. Congress has established an agency, the Sentencing Commission, to collect data and the views of various constituencies in formulating policies and rules. Congress, not courts, can hear expert testimony about the dangers and harms of various crimes and the best ways to address them. Congress, not courts, sets budgets and has to balance priorities such as funding for prisons and law enforcement. Most importantly, Congress has democratic legitimacy; courts do not.

The Guidelines did not completely constrain judges' ability to act on their own policy preferences in sentencing. Judges' policy preferences may still influence adjustments and discretion within sentencing ranges, but they are constrained by the facts, the Guidelines framework, and appellate review. In addition, the Guidelines range, formerly a safe harbor, specified the zone of influence a judge was to have in a standard case. Thus, even under the Guidelines, the ultimate sentence may still reflect a judge's policy preferences. The rationale for these moving parts, however, was that no set of rules could foresee and address the myriad individual characteristics that may arise. Congress did not authorize judges to follow their policy preferences, but was willing to pay some of that price so that judges could give some individualized justice.

Unfortunately, *Rita* and especially *Kimbrough* give sentencing courts substantial latitude to incorporate their own policy preferences at sentencing. *Kimbrough* embraced the government's concession that "courts may vary [from Guidelines ranges] based solely on policy considerations, including disagreements with the Guidelines."⁵⁸ *Kimbrough* was an exceptional case, where the Sentencing Commission itself had tried to back away from the 100:1 sentencing ratio for crack versus powder cocaine. Nevertheless, *Rita* embraced this principle more generally. It asserted that a sentencing court may entertain "argu[ments] that the Guidelines reflect an unsound judgment, or, for example, that they do not generally treat certain defendant characteristics in the proper way."⁵⁹

The Guidelines tried to harmonize divergent views on whether, for example, youth or family circumstances should aggravate or mitigate sentences or be irrelevant. Some judges thought young offenders more sympathetic; others thought them more dangerous and likely to recidivate.

⁵⁸ *Kimbrough v. United States*, 128 S. Ct. 558, 570 (2007) (quoting Brief for United States at 16, *Kimbrough*, 128 S. Ct. 558 (No. 06-6330)).

⁵⁹ *Rita v. United States*, 127 S. Ct. 2456, 2468 (2007).

The Commission split the difference, deciding that age should be irrelevant to sentencing except in extraordinary cases.⁶⁰

Rita destroys this uniformity, removing these policy choices from a central, transparent body that may possess some expertise and democratic legitimacy. The result will likely be a greater dispersion of sentences due solely to judges' varying policy preferences. Sentencing-court discretion makes sense when judges need to tailor rules to case-specific facts, but it makes much less sense for recurring policy issues susceptible to rules, or at least rules of thumb. Indeed, *Rita* is far too sanguine about sentencing-court discretion for discretion's sake, and not careful to guard against the idiosyncratic policy preferences of a single unelected judge.

Moreover, in returning these policy choices to hundreds of judges, the Court discouraged them from being completely candid and transparent. Sentencing courts must articulate reasons for deviating from the Guidelines, but at the same time are discouraged from adopting clear rules. Indeed, the Court in *Kimbrough* praised the sentencing court for rejecting the Commission's 100:1 ratio but not announcing an alternative ratio.⁶¹ The Court likewise discouraged appellate courts from being transparent about how they review the size of a departure. *Gall* rejected mathematical proportionality as a guidepost but recognized in much fuzzier terms that large departures need weightier justifications.⁶² The upshot reminds us of the Supreme Court's hand-waving in the *Grutter* and *Gratz* affirmative action cases: a school may give race significant or decisive weight as an admissions factor, but only if it does not transparently quantify that weight.⁶³ This lack of transparency hinders appellate review, public scrutiny, and criticism.

Sentencing policy discretion might make more sense in the hands of sentencing juries, which may represent local popular opinions. The *Apprendi* line of cases started as a ringing vindication of the Sixth Amendment right to a criminal jury trial against sentencing-court encroachment. We are not fans of *Apprendi* or jury sentencing, but at least we understand this argument. By the time the Court reached *Booker*, however, its shifting coalitions turned that jury right into a vindication of sentencing-court discretion.

⁶⁰ U.S. SENTENCING GUIDELINES MANUAL, *supra* note 16, § 5H1.1 (policy statement regarding age).

⁶¹ *Kimrough*, 128 S. Ct. at 575.

⁶² *Gall v. United States*, 128 S. Ct. 586, 594–97 (2007) (“We also reject the use of a rigid mathematical formula that uses the percentage of a departure as the standard for determining the strength of the justifications required for a specific sentence,” but “[w]e find it uncontroversial that a major departure should be supported by a more significant justification than a minor one.”).

⁶³ *Compare* *Grutter v. Bollinger*, 539 U.S. 306, 334–37 (2003) (upholding the University of Michigan's affirmative action program because it treated race as a plus within a flexible, holistic evaluation rather than giving it a clear weight or quota), *with* *Gratz v. Bollinger*, 539 U.S. 244, 271–72 (2003) (striking down the University of Michigan's racial preference in undergraduate admissions because it automatically gave every underrepresented minority applicant twenty out of one hundred points needed for admission, rather than allowing for individualized consideration).

In doing so, it gave the power back to the very officials who supposedly threatened juries. Instead of upholding or striking down the Guidelines, the Court excised provisions about which sentencing judges had long complained, most notably the Guidelines' binding force.⁶⁴ As a policy matter, this result is preferable to Justice Scalia's reading of the Sixth Amendment. He would invalidate all substantive appellate review and limit appeals to procedural review.⁶⁵ The Court, however, has rewritten the Sentencing Reform Act as it pleased without responding effectively to Justice Scalia's challenge. Why does the Sixth Amendment permit deferential abuse-of-discretion appellate review, but not the de novo appellate review that Congress specified in the PROTECT Act?⁶⁶ Why is the Sixth Amendment case-law now about sentencing versus appellate court discretion, instead of judges versus juries? The Court has never provided a satisfactory answer to either question.

Policy departures will inevitably undermine appellate review, Congress and the Commission's policymaking authority, and thus consistency. How is an appellate court to review a sentencing judge's disagreement with policy? When are policy disagreements reasonable or unreasonable?

Of course, judges have always had some wiggle room to express their policy preferences. Even under the extraordinarily detailed Federal Guidelines, the final range itself allowed a twenty-five percent variance in sentence. As long as the sentencing court had calculated the final offense level properly, the sentencing court's discretion within that range was unreviewable. Sentencing judges had further discretion to impose adjustments when the facts were close or at least could survive clear-error review. In addition, sentencing judges had discretion to depart, even though appellate courts reviewed departures more carefully. We hope courts will not take the dicta in *Rita* too seriously and will limit *Kimbrough* to its facts, namely the extreme 100:1 crack-cocaine sentence disparity.

* * *

We have criticized the Court's guidelines jurisprudence, which seems to be an unintended outgrowth of a series of fragile and shifting coalitions. It upends valuable federal and state sentencing reforms, yet has only tenuous roots in the Sixth Amendment.⁶⁷ The Court's relaxed standard of review for departures is troubling, as it will likely increase disparity. We are particularly concerned that sentencing courts now have new leeway to inject their policy preferences into their sentences. In contrast, our suggested ap-

⁶⁴ *United States v. Booker*, 543 U.S. 220, 259 (2005) (Breyer, J., remedial majority opinion).

⁶⁵ *Rita v. United States*, 127 S. Ct. 2456, 2482–84 (2007) (Scalia, J., concurring in the judgment).

⁶⁶ *See Booker*, 543 U.S. at 259, 261 (Breyer, J., remedial majority opinion).

⁶⁷ *See, e.g., Stephanos Bibas, Rita v. United States Leaves Open More than It Answers*, 20 FED. SENT'G REP. 28 (2007).

proach would preserve much of the Court's jurisprudence and better address the substantive concerns about the Federal Guidelines. Thus, it would constrain overpunishment and produce more consistency.

III. POLITICAL AND INSTITUTIONAL REFORM

In our account thus far, appellate review has been the linchpin of the sentencing regime, holding sentencing judges in line and ensuring consistency. As we have argued, appellate review must retain enough bite to prevent sentencing judges' biases from skewing sentences. We propose additional reforms that could add more teeth to appellate review as a check on sentencing judges' discretion. Section A below considers how ensuring political balance on courts of appeals would moderate ideological extremes and ensure fidelity to clear sentencing rules. Section B then briefly notes the problem of waivers of appeals and why the law might restrict them.

A. Political Alignments in Judicial Hierarchy

Doctrine and judicial review matter in sentencing. Previous studies show that appellate courts constrain sentencing courts' law-based departures but leave them freer to use fact-based adjustments.⁶⁸ Political alignment between the lower court and the appellate court relaxes the constraints sentencing courts face when they want to depart. Democratic appointees depart more in Democratic circuits, while Republican appointees behave about the same regardless of alignment.⁶⁹ These findings fit the Guidelines' structure. Federal sentences are generally high, and adjustments, which are hard to review, exponentially increase them.⁷⁰ While Republicans can raise sentences within this system, Democrats can lower them only by departing and only with the acquiescence of sympathetic, usually Democrat-controlled, courts of appeals.

A corollary is that the lack of political diversity within the judicial hierarchy produces more sentencing disparity because sentencing judges do not fear being reversed on appeal. Current random assignment of federal circuit judges to panels sometimes produces ideologically unbalanced panels, with three appointees of the same party controlling the outcome. We refer to that unity as horizontal political alignment. When such unified panels review a sentencing judge who holds the same political orientation as the panel (vertical political alignment), there is little check on the sentence's severity or lenience. Thus, a district judge sitting in a circuit that is strongly tilted toward ideological alignment feels less vertical constraint in choosing a severe or lenient sentence. For example, the Fifth Circuit currently com-

⁶⁸ Schanzenbach & Tiller, *supra* note 28; Schanzenbach & Tiller, *supra* note 29.

⁶⁹ Schanzenbach & Tiller, *supra* note 28, at 47–52.

⁷⁰ Of course, because upward departures are far rarer, there may be political differences in those too that we cannot detect because our sample size is too small.

prises twelve Republican appointees and only four Democratic appointees, so a Republican-appointed district judge is unlikely to fear reversal.

The political structure of the judicial hierarchy and horizontal relationships within a court profoundly influence how consistently courts apply doctrines to similar cases. Cross and Tiller have shown that a politically diverse panel of judges is more likely to apply established doctrine faithfully than a panel of judges from the same party. The presence of political minority members on three-judge federal appellate panels led to more consistent application of the *Chevron* doctrine in administrative law cases.⁷¹ Likewise, Cass Sunstein found that judges on like-minded panels appear to vote far less moderately than judges on divided panels.⁷² Richard Revesz reached similar findings in his study of environmental law cases.⁷³ One possible explanation for this moderation is that judges within panels learn from one another's ideas and worldviews, blow the whistle on extreme or disingenuous positions, and otherwise check abuses of discretion. Thus, divided panels are likely to decrease sentencing disparity by deterring extreme positions.

To reduce sentencing disparity, we should acknowledge the politics and ideology at work and consider structural changes to judicial review. In particular, we suggest institutionalizing political diversity in sentencing cases by ensuring that appellate courts not share the sentencing court's political orientation. At a minimum, one could require that each appellate panel include a panel member who does not share the sentencing court's political orientation. The simplest way to do this would be to use the ideology of the appointing President for each judge. There would be occasional errors; a few Democratic appointees are very conservative and a few Republican appointees are very liberal. These outliers would not matter much, however, because most judges are true to political type (at least for sentencing issues) and there are many opportunities for sentencing review.

There are various ways to achieve political diversity in sentencing. First, one could limit three-judge appellate panels to having at most two judges who share the same political orientation. This ideological cap would modify the normal random assignment of appellate panels, thus always ensuring both vertical and horizontal political diversity.⁷⁴ A second mechanism would be to set up a separate court of appeals for criminal cases or criminal sentencing and to ensure that the court is politically diverse. The court could comprise current federal circuit judges, brought together in an alternative forum. There is precedent for specialized courts of appeals: the

⁷¹ Cross & Tiller, *supra* note 34, at 2172 & tbl.3.

⁷² CASS R. SUNSTEIN, WHY SOCIETIES NEED DISSENT 169 (2003).

⁷³ See Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717, 1752-56 (1997).

⁷⁴ Emerson H. Tiller & Frank B. Cross, *A Modest Proposal for Improving American Justice*, 99 COLUM. L. REV. 215 (1999).

Court of Appeals for the Federal Circuit has exclusive appellate jurisdiction over patent appeals from federal district courts.⁷⁵

Other government agencies use political ideology or partisan identifiers to ensure balance. Many independent regulatory commissions limit partisan excess through split party arrangements. Examples include the Surface Transportation Board (formerly the ICC), the National Transportation and Safety Board, the Securities and Exchange Commission, and the Nuclear Regulatory Commission.⁷⁶ Congress designed these agencies as institutions filled with experts who have some quasi-judicial functions, but with partisan safeguards to protect against ideological excess by one party. Most telling is the design of the United States Sentencing Commission itself, which by law must have seven voting members, no more than four of whom may be members of the same political party.⁷⁷

Some have criticized proposals to restructure courts to minimize political bias. Former D.C. Circuit judge Patricia Wald claims that using political orientation to check panel make-up would make judges act more partisan, as role-players.⁷⁸ While that is possible, the political diversity of the mixed panel should nonetheless offset the extremes of either group. In other words, a highly partisan yet mixed panel is more likely to sentence consistently than a less charged group of like-minded partisans randomly assigned to the same panel.

Professor Eric Posner has also criticized split-panel solutions. He argues that while political diversity is good, spreading minority judges out means that they will never wind up being in the majority. Random assignment, in contrast, would guarantee that the minority would sometimes be in the majority.⁷⁹ This argument misses the point about the power of a single minority member to blow the whistle, however. The problem is not the voting power of an ideological majority, but ideological unanimity on a panel or between the sentencing and appellate court. Unanimity makes it easier for ideologues to disobey doctrine and manipulate sentencing guidelines.⁸⁰ A single panel member can dissent, questioning the honesty of the majority's reasoning and publicly flagging and criticizing the manipulation. Diluting minority votes by spreading them across many panels is more tolerable than lacking any whistleblower at all on politically unified panels. An alternative would be to increase the probability of drawing a minority

⁷⁵ 28 U.S.C. §§ 1295(a)(1), 1338(a) (2006).

⁷⁶ STAFF OF S. COMM. ON GOV'T OPERATIONS, 95TH CONG., STUDY ON FEDERAL REGULATION: THE REGULATORY APPOINTMENTS PROCESS 2 (Comm. Print 1977).

⁷⁷ 28 U.S.C. § 991(a) (2006).

⁷⁸ Patricia M. Wald, *A Response to Tiller and Cross*, 99 COLUM. L. REV. 235 (1999).

⁷⁹ Eric A. Posner, *Does Political Bias in the Judiciary Matter? Implications of Judicial Bias Studies for Legal and Constitutional Reform*, 75 U. CHI. L. REV. 853 (2008).

⁸⁰ Cross & Tiller, *supra* note 34, at 2172.

judge in certain cases that are likely to be politically charged, such as sentencing cases.

We admit that this discussion is somewhat of a thought experiment. It shows, however, the relevance of appellate review and the need for sentencing rules that foster meaningful review. Even in a circuit that has become largely Republican or Democratic, there is still a diversity of ideologies within each party. This diversity itself may have a whistleblowing effect and may increase consistency in sentencing. Contrast this situation with sentencing in the trial court, where there is only one judge and no check. In addition, the selection process for circuit judges is more rigorous than for district judges. District-court appointments are rarely blocked, but circuit nominees' records and personal opinions receive careful scrutiny, and a minority of senators can block them.⁸¹

B. Regulating Appeal Waivers

Up to now, we have spoken of appellate review as if it were an automatic check on sentencing-judge variation. But appellate courts can, of course, review only those sentences that are appealed. If the parties choose not to appeal, then appellate courts can do nothing to police trial courts. For this reason, the widespread use of appeal waivers is troubling.

In an excellent empirical study, Nancy King and Michael O'Neill found that defendants waived their rights to appeal in nearly two-thirds of all plea-bargained cases.⁸² The authors also found evidence that defendants who accepted appeal waivers received lower sentences, suggesting that the threat of appellate review influences sentencing.⁸³ It is impossible to know, however, if the size of the sentencing discount for accepting an appeal waiver is usually proportional to the chance of reversal on appeal. If the discounts are roughly proportional, then the threat of appellate review constrains parties even if appeal waivers are widespread. Nevertheless, appeal waivers could undermine guideline sentencing by reducing the information produced by appellate opinions or by fostering collusive bargains between prosecutors and defendants.⁸⁴ Thus, many of the arguments in favor of appellate review under the Guidelines also support the regulation of appeal waivers.

⁸¹ For a discussion, see John O. McGinnis & Michael Rappaport, *Supermajority Rules and the Judicial Confirmation Process*, 26 CARDOZO L. REV. 543 (2005).

⁸² Nancy J. King & Michael O'Neill, *Appeal Waivers and the Future of Sentencing Policy*, 55 DUKE L.J. 209, 231 (2005).

⁸³ *Id.* at 232–36.

⁸⁴ Plea bargaining is a sufficiently low-visibility process, riddled with agency costs, imperfect information, and psychological biases and heuristics, that there is serious reason to doubt that plea bargains closely track the merits or the outcomes that litigation would have achieved. See Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 HARV. L. REV. 2463 (2004).

Appeal waivers that restrict the right to appeal sentences that fall within a Guidelines range do not risk undercutting the consistent application of the Guidelines. But sweeping waivers of the right to appeal any sentence within the statutory range allow prosecutors and defense lawyers to collude. These sweeping waivers also signal to sentencing judges that they can sentence without regard to sentencing law or policy. To preclude this kind of collusion, appellate courts might hold broad waivers unenforceable as a matter of public policy. Likewise, the Department of Justice could issue a policy memorandum forbidding very broad appeal waivers.

Even if appellate courts and head prosecutors clamp down on appeal waivers, collusion remains a problem inherent in plea bargaining. Defendants, of course, always retain incentives to challenge high sentences, and invalidating appeal waivers will free them to do that. But line prosecutors who collude with defense counsel do not want the defendant to appeal because an appeal would create more work for the prosecutor and unravel the bargain. If the defendant declines to appeal, then it does not matter whether an appeal waiver is enforceable. The only solution, far from a perfect one, is for head prosecutors to restrict collusive bargains and to press for more appeals of below-Guidelines sentences. The Department of Justice has tried, with uncertain success, to restrict plea bargains that circumvent the Guidelines.⁸⁵ Likewise, Congress pressured prosecutors to report and appeal below-Guidelines sentences through the PROTECT Act.⁸⁶ These political checks are imperfect. But just as the Guidelines are an imperfect effort to harmonize judges' sentences, so too are prosecutorial policies an imperfect but useful effort to harmonize prosecutors' bargains.

To be fair, the case for centralization is far weaker for crimes that are often prosecuted locally, such as robberies and small drug sales. After all, local courthouse working groups make policy in light of the fallback option of state law. Rigid centralization ignores the divergence of federal and state sentences, making an enormous amount ride on the fortuity of whether a

⁸⁵ See, e.g., WILLIAM F. WELD, U.S. DEP'T OF JUSTICE, PROSECUTORS HANDBOOK ON SENTENCING GUIDELINES (1987) ("The Redbook"), reprinted in 6 FED. SENT'G REP. 333 (1994); Richard Thornburgh, U.S. Dep't of Justice, Plea Policy for Federal Prosecutors (Mar. 13, 1989) ("Thornburgh Bluesheet"), reprinted in 6 FED. SENT'G REP. 347 (1994); Memorandum from Attorney General John Ashcroft to All Federal Prosecutors, Department Policies and Procedures Concerning Sentencing Recommendations and Sentencing Appeals (July 28, 2003), available at [http://www.nacdl.org/public.nsf/legislation/ci_03_32/\\$FILE/AG_Guidance_Steg_Recs.pdf](http://www.nacdl.org/public.nsf/legislation/ci_03_32/$FILE/AG_Guidance_Steg_Recs.pdf); Memorandum from Attorney General John Ashcroft to All Federal Prosecutors, Department Policy Concerning Charging Criminal Offenses, Disposition of Charges, and Sentencing (Sept. 22, 2003), available at http://www.usdoj.gov/opa/pr/2003/September/03_ag_516.htm; Memorandum from Acting Deputy Attorney General George J. Terwilliger, III to Holders of United States Attorneys' Manual Title 9, Indictments and Plea Procedures Under Guideline Sentencing (Feb. 7, 1992) ("Terwilliger Bluesheet"), reprinted in 6 FED. SENT'G REP. 350 (1994). For a suggestion that these efforts do not "meaningfully restrain" prosecutors in practice, see Frank O. Bowman, III, *Beyond Band-Aids: A Proposal for Reconfiguring Federal Sentencing After Booker*, 2005 U. CHI. LEGAL F. 149, 193.

⁸⁶ Pub. L. No. 108-21, § 401(l), 117 Stat. 650, 674 (2003).

prosecutor takes a case federally.⁸⁷ The case for centralization is far stronger for crimes that are de jure or de facto primarily federal responsibilities. These include the many cases involving immigration, large-scale drug trafficking, tax evasion, terrorism, accounting fraud, and insider trading.⁸⁸ At least for these purely federal cases, vigorous oversight of line prosecutors and sentencing judges makes a great deal of sense. Even for local crimes, each prosecutorial office and court of appeals should at least ensure intradistrict or intrastate consistency.

CONCLUSION

This Essay applied positive political theory and recent empirical evidence to critique the Supreme Court's sentencing guidelines jurisprudence. We conclude that the Court's jurisprudence is deeply misguided. Binding guidelines and searching appellate review are needed to make sentencing more consistent and legitimate. In addition, the Guidelines reflect a political bargain struck by legislators and sentencing commissioners, a compact that judges should honor. The Court's new tack on sentencing undermines this process by reducing appellate court review for departures and creating a new ground for departures—the nebulous and hard-to-review policy departure. On the other hand, we endorse the trend toward more robust appellate review of sentencing adjustments, which judges can at times manipulate to alter sentences substantially. Stricter review of adjustments makes the review of sentences more symmetric and should rein in discretion. Of course, even this beneficial feature of *Booker* and its progeny could be undermined if judges are also freer to depart.

Much of the benefit of a guidelines system comes from introducing appellate review into the criminal sentencing process. We could improve this system by ensuring political balance on courts of appeals, so that there is always at least one judge of a different party to blow the whistle on ideological manipulation of sentencing rules. Guidelines coupled with politically diverse appellate review dampen the political and ideological tendencies inherent in judicial discretion. Whether judges agree with the forces that produced the guidelines, they must enforce the resulting legislative bargain faithfully. Legislatures and sentencing commissions have democratic mandates to channel judicial discretion with rules. Appellate judges need to check and balance both sentencing judges and their own appellate colleagues, and overly broad appeal waivers threaten that check. This system of institutional checks and balances would prevent judges of a single political ideology from unilaterally warping sentencing law to serve their own biases and preferences.

⁸⁷ Daniel Richman, *Federal Sentencing in 2007: The Supreme Court Holds—The Center Doesn't*, 117 *YALE L.J.* 1374, 1396–1407 (2008).

⁸⁸ *Id.* at 1407–11.

APPENDIX

Sentencing Table (in months of imprisonment)

Offense Level	Criminal History Category (Criminal History Points)					
	I (0 or 1)	II (2 or 3)	III (4, 5, 6)	IV (7, 8, 9)	V (10, 11, 12)	VI (13 or more)
1	0-6	0-6	0-6	0-6	0-6	0-6
2	0-6	0-6	0-6	0-6	0-6	1-7
3	0-6	0-6	0-6	0-6	2-8	3-9
4	0-6	0-6	0-6	2-8	4-10	6-12
5	0-6	0-6	1-7	4-10	6-12	9-15
6	0-6	1-7	2-8	6-12	9-15	12-18
7	0-6	2-8	4-10	8-14	12-18	15-21
8	0-6	4-10	6-12	10-16	15-21	18-24
9	4-10	6-12	8-14	12-18	18-24	21-27
10	6-12	8-14	10-16	15-21	21-27	24-30
11	8-14	10-16	12-18	18-24	24-30	27-33
12	10-16	12-18	15-21	21-27	27-33	30-37
13	12-18	15-21	18-24	24-30	30-37	33-41
14	15-21	18-24	21-27	27-33	33-41	37-46
15	18-24	21-27	24-30	30-37	37-46	41-51
⋮	⋮	⋮	⋮	⋮	⋮	⋮
25	57-71	63-78	70-87	84-105	100-125	110-137
26	63-78	70-87	78-97	92-115	110-137	120-150
27	70-87	78-97	87-108	100-125	120-150	130-162
28	78-97	87-108	97-121	110-137	130-162	140-175
29	87-108	97-121	108-135	121-151	140-175	151-188
30	97-121	108-135	121-151	135-168	151-188	168-210
31	108-135	121-151	135-168	151-188	168-210	188-235
32	121-151	135-168	151-188	168-210	188-235	210-262
33	135-168	151-188	168-210	188-235	210-262	235-293
34	151-188	168-210	188-235	210-262	235-293	262-327
35	168-210	188-235	210-262	235-293	262-327	292-365
36	188-235	210-262	235-293	262-327	292-365	324-405
37	210-262	235-293	262-327	292-365	324-405	360-life
38	235-293	262-327	292-365	324-405	360-life	360-life
39	262-327	292-365	324-405	360-life	360-life	360-life
40	292-365	324-405	360-life	360-life	360-life	360-life
41	324-405	360-life	360-life	360-life	360-life	360-life
42	360-life	360-life	360-life	360-life	360-life	360-life
43	Life	Life	Life	Life	Life	Life

