

COMMENTARY

ADVERSARIAL BIOGRAPHY: REFLECTIONS
ON THE SENTENCING OF MICHAEL MILKEN

Stanton Wheeler

FSR asked Stanton Wheeler, the Ford Foundation Professor of Law and the Social Sciences at Yale Law School, to comment on the Milken case. Beginning in 1976, Wheeler served as Director of a federal research program to study white collar crime. His 1988 book, *Sitting in Judgment: The Sentencing of White-Collar Criminals*, co-authored with Kenneth Mann and Austin Sarat, offers—in our view—the fullest analysis of federal judges' thinking processes and perspectives on white collar sentencing prior to sentencing guidelines.

The sentencing of Michael Milken evoked the most public commentary and newspaper coverage of any sentence in memory. For weeks prior to sentencing, journalists offered their thoughts on how Judge Kimba M. Wood should sentence Milken. In the days following, there was a torrent of reportage and opining. Is there anything left to be said, then, about the sentence itself?

Social scientists are leery of making too much of special cases, and there is little doubt that the sentencing of Michael Milken was a unique event. But the case has one particular virtue that commends it for detailed analysis: It is perhaps the last of the great pre-guidelines sentences. Ninety-four U.S. attorneys still need to wind up the conviction and sentencing of federal offenders whose crimes occurred prior to November 1987, but there can be none quite like this. Though the Milken case is atypical, it arguably represents the best the pre-guidelines system could offer.

Consider the parties: On the government's side, we have the U.S. Attorney for the Southern District of New York, by any account a premier office in the sanctioning of white collar crime since the days of Robert Morgenthau. On the defense side, we have the prestigious firm of Paul, Weiss, Rifkind, Wharton & Garrison, and no less a lead attorney than Arthur Liman. Liman's background as a private litigator, and even more so his public service as Chief Counsel to the Attica Commission, and in the Iran-Contra hearings, mark him as special indeed. We can expect nothing but the best in terms of lawyering for both prosecution and defense.

Then consider the defendant. Virtually everyone agreed that here was a genius of Wall Street, the most creative financier of the 80's, the man who gave junk bonds a good name. He could be counted upon to work with legal counsel with the same degree of intensity he had shown in his business practices.

Finally, Judge Kimba Wood. Here we meet the only possible departure from the model. She arrived on the federal bench in 1988 only shortly before being assigned this historic case. But from her Harvard Law School education, her handling of prior cases, and her no-nonsense dealings with

earlier stages of this case we have every reason to anticipate an intelligent and thoughtful decision. Thus the case provides an unparalleled opportunity to examine the sentencing process as it should be functioning at its best prior to the guidelines.

With the stage thus set, let us begin with the two prime documents presenting the views of the adversaries: defense counsel's sentencing memorandum, and that of the government. (Judge Wood had other documents at hand, including also the pre-sentence investigation report by the Federal Probation Department, and the *Fatico* hearing documents.)

The defense memorandum sounds its theme on page one. "Michael Milken's crimes stand in sharp contrast to the characteristics and values he has displayed throughout his life." Over the next 70 pages we are offered a distinctive portrait of that life: Milken's childhood and adolescence, his education, his introduction to Drexel, the establishment of the High Yield department. We see him as a devoted family man, with concern for friends and co-workers. We see in particular his commitment to charity and community service. The portrait is crafted with quotations from letters of employees, friends, business associates, and those who work with him in his charitable pursuits. No summary can do justice to the richness of these materials, but two examples give the flavor. Said a seventh grade teacher:

... he always took time out to be kind and considerate of those less successful and popular than he. This trait of genuine sensitivity and concern for others made him someone special to everyone.

From an observer of his math classes at a 24-hour emergency shelter in Los Angeles for abandoned, neglected and abused children:

Every Saturday morning magic happens at MacLaren. . . . to watch Michael teach these special children is like watching a miracle happen . . . They come alive! He stimulates these youngsters in a way in which they have never been stimulated before.

The defense memorandum devotes 24 pages to Milken's offenses. After arguing that his sentence should be based solely on the counts to which he pleaded guilty, the defense examines each count, taking pains to show the limit of culpability. It argues, for example, that "he engaged in thousands of trades . . . [but] . . . the trades that are the subject of Michael Milken's plea constituted a tiny fraction. . . of the Drexel - [redacted] business, which was, except for a few instances, wholly legitimate."

The main theme of the defense is that "although inexcusable, Michael's offenses were neither venal nor motivated by greed. He did not personally profit from them. . . . Nor did these offenses infect the core of his business." It quotes noted securities lawyer Joseph H. Flom, who writes the court to the effect that many of Milken's offenses were technicalities for which historically criminal prosecution was rare.

The third section of the defense memorandum is called "principles of sentencing." Here, in addition

COMMENTARY

to urging a sentence of probation, the defense fights hard to keep Milken from being singled out:

The values of proportionality of punishment and consistency in sentencing assume singular importance in this case, where public debate has taken on the appearance of a heresy trial. In our system, it is the sentencing judge, as Judge Lasker stated, who bears the ultimate responsibility "to protect a defendant against unreasonable passions of public opinion," to consider each case and each defendant "on the merits" and "to weigh the favorable and the unfavorable circumstances before it." We ask that this Court follow these principles, and sentence Michael Milken as an individual, not as a symbol."

The Court is also urged to abide by Southern District of New York precedents. Appendix A lists every pre-guideline criminal prosecution for securities fraud in S.D.N.Y. over the last 10 years. Even the worst of the offenders rarely got more than a 2 or 3 year sentence.

A concluding section devotes 20 pages to the defense sentencing proposal: a community service sentence to work in a drug prevention program in Los Angeles. After describing program benefits in detail, the defense shows how the proposed sentence will satisfy the goals of sentencing.

For retribution/punishment: He has been pilloried since 1986. He has suffered acutely. Though he is one of the most innovative and influential financial thinkers, he has been banished for life from his industry. "We submit that society's need to punish Michael has been fully vindicated in this case."

For "general deterrence:" This case has already harvested more than its share. Brokers have learned that acts formerly punished civilly or administratively now get punished criminally. Moreover, Milken has already paid \$200 million in criminal fines and \$400 million in restitution to satisfy a societal interest in general deterrence.

Having disposed of punishment and deterrence, the memorandum indicates how much Milken and the world have to gain from a community service sentence.

The government's sentencing memorandum stands in stark contrast, a hardly surprising outcome when we consider the nature of the adversary system. Still, it is disconcerting to see a person drawn in such dramatically different terms. About the only thing the two sides agree upon is that this is not, technically, a "guidelines" sentence case.

Just as the defense trumpeted its theme on page 1, so too does the government come out swinging.

Milken's crime forms a pattern of calculated fraud, deceit and corruption of the highest magnitude. In a highly sophisticated and systematic way, Milken endeavored to enhance his power and enlarge his wealth by willingly cheating clients, manipulating the markets, and evading the laws and rules designed to safeguard the entrusting public. . . .

The charges to which he pleaded guilty "provide but a sampling of the larger pattern of criminal activity that permeated Milken's operation . . . systematic and pervasive, not aberrational and isolated."

The memorandum is divided into three parts: (1) the government's view of sentencing and the principles to be applied; (2) the conduct to which Milken pleaded guilty; and (3) his other criminal conduct. Unlike the defense memorandum, there is hardly a mention of Milken's past. The section on sentencing principles (27 pages) is devoted primarily to the justification for considering *all* relevant conduct, not just that to which Milken pleaded guilty, and to his failure to mitigate punishment by genuinely helpful assistance to the government, or with genuine expressions of remorse. The second section (22 pages) highlights the crimes that Milken acknowledged, just as the defense had sought to minimize them.

The final section (70 pages) is devoted to "Milken's other crimes," those allegations for wrongful and culpable behavior to which he did not plead guilty.

The result of the sentencing memoranda is an adversarial portrait that rarely seems to describe the same person. The defense portrays Michael Milken as a misunderstood man whose offenses were often treated as minor violations in the past, and who is being punished for what he symbolizes, not for what he did. The government portrays Michael Milken as an intelligent but devious and crafty person who manipulated the system to his advantage, whose greed knew no bounds, and whose criminal acts wreaked havoc in and on the financial world.

About the only place they join issue is in reference to greed and money. The defense memorandum (page 11):

as everyone who knows him confirms, the accumulation of money and materials goods has never motivated Michael.

The government memorandum (page 20, footnote 7):

Former Drexel employees are skeptical of Milken's claim of indifference to personal profit. As one former employee put it, "Milken never let go of a nickel except to get a better grip."

Perhaps this is all we can expect from adversarial biography. We have two sides going at it, each trying to counter the other's portrayal.

In this system it is the Judge, of course, who has the discretion to set the sentence and who can choose to accept either party's characterization of the defendant. At the sentencing hearing, the Judge heard first from Mr. Liman (32 pages in transcript), then extremely briefly from Mr. Milken (4 lines), and briefly from the government (4 pages). The oral arguments are not essentially different from the sentencing memoranda, but two points are noteworthy. First, Mr. Liman begins by reminding the court that sentencing is different than all the criminal proceedings that precede it:

COMMENTARY

Criminal proceedings inevitably focus only on the ill events in a person's life, and I sometimes think that it's like judging a painter on his unsuccessful painting or a writer on a novel or novels that are critical failures. A sentencing, though, is different. A sentencing is a retrospective in which a person's entire works, good as well as bad, and a person's entire life, and not just that thin slice that is placed under the microscope in a criminal proceeding, have to be weighed and evaluated by the court, and that's what the court has been doing.

Here is a personal appeal to change the frame—to get away from the government's insistent focus only on criminal wrongdoing, and to get the Court thinking about the good and charitable deeds of the past.

Second, speaking for the government, Assistant U.S. Attorney Jess Bardella doesn't deny Michael Milken's good qualities and considerable talents.

No one disputes that he has family and friends to whom he has been loyal and loving. . . It is acknowledged too that Mr. Milken has engaged in substantial charitable contributions and works.

Yet fairness means these positive facts must be placed in perspective, and he alludes to the sentencing guidelines which, while they don't govern, are instructive.

They say that family ties and responsibilities and community ties are not ordinarily grounds for giving special consideration to a defendant.

The government seeks to turn Milken's good qualities to its advantage. After arguing that under the guidelines, such qualities do not really count, Mr. Bardella turns the argument on its head:

That the defendant was blessed with intelligence, energy, education and the support of family and friends makes his choice to engage in persistent violation of the law all the more inexcusable.

There follow the principal substantive arguments. The defendant profited from his transactions. He obstructed justice. He increased his wealth at the expense of his victims. He demonstrated contempt for laws designed to insure sound markets. His prominence as a financier means that this is a case in which an appropriate sentence can make a difference. The sums he has agreed to pay in fines and restitution reflect substantial deterrence and retribution for his crimes, but "enormous monetary penalties standing alone . . . are perceived as a cost of doing business. . . ." Thus only substantial imprisonment will send a message, will satisfy society's need for retribution, and will reaffirm the integrity of our markets.

We come then to Judge Wood's courtroom presentation. A glimmer of her approach is evident at the outset, for she devotes the first few pages of transcript to the legitimacy of sentencing Mr. Milken on the basis of all his conduct, not just that to which he pled guilty. She goes into some detail about the *Fatico* hearing and why she opted for it in this case.

Although the *Fatico* hearing, in her view, established neither version, it did show two instances in which Milken obstructed justice, and one in which he deliberately withheld information from clients. These factors add to the wrongdoing to which he pled guilty.

Judge Wood then sentences Milken to a total of 10 years in prison followed by 3 years community service, 1800 hours per year.

Since this is arguably the most highly publicized sentence in the history of white collar crime, it is fair to ask a number of questions about it. How did Judge Wood arrive at these particular numbers? Is there authority for them in the practices of other federal judges faced with white collar crime sentencing? Finally, what can this tell us, if anything, about the process of pre-guideline sentencing?

Sentencing is anything but an exact science, and different commentators will approach these questions differently. My own observations are informed by the effort my colleagues and I made to understand white collar sentencing by interviewing in depth and detail some 50 federal district court judges.¹

Surely there is a basis in the common sense wisdom of sitting judges for sanctioning heavily those who occupy special positions of authority and responsibility and who abuse those positions by committing crimes, especially when the crimes are sophisticated and of substantial duration. The argument for stiff sentences in such cases goes both to the gravity of the offense and the importance of sending a deterrent message. Judges feel strongly about violations of trust, and about violations that have the capacity to affect the fabric of financial transactions.

At the same time, there is little working out of the precise reasoning behind the numbers. Federal judges often talk about the harm caused by an offense, the blameworthiness of the offender, and the consequence the sentence is likely to have, especially the consequence for general deterrence. Words reflecting each of these concerns are found in Judge Wood's remarks, but in the end we have little idea of how much of the 10 years is due to the seriousness of the offense, how much to Milken's own culpability, and how much to the capacity of a long sentence to deter others. Her commentary suggests that the latter two factors weighed more heavily than the first, but how much more and with what relative weight between them seems impossible to discern.

Indeed, the number 10 itself remains a mystery. Cannot one read all of her words and come up with a different number, like 5 or 6? Six years would have been twice as long, for example, as the Boesky sentence, and might have been a means of giving some weight to the acknowledged prior good deeds of Mr. Milken.

Here there is a telling feature of Judge Kimba Wood's sentencing. The judges we interviewed returned again and again to a common theme we came to call the paradox of leniency and severity, or the dilemma of praiseworthy conduct and blameworthy status. On the one hand, many white collar offenders have exemplary pasts. They have led

COMMENTARY

upright lives, have contributed to their communities and deserve some credit for these past good deeds. At the same time, they have an advantaged social status and often occupy positions of responsibility, and therefore ought to be held to a higher standard and deserve more punishment when they fall.

Many judges feel genuine anguish over the balancing of these two contrary pulls at sentencing. Judge Kimba Wood clearly resolved them on the side of severity. She uses the words from both sides of the paradox—she has “taken into account” his significant community service, his encouragement of colleagues to do the same, the emotional support he has provided and the positive image many colleagues and competitors held of him. One can ask, then, if Judge Wood was taking account of these factors, what might Michael Milken’s sentence have been if she had not taken them into account? Twelve years? Fifteen years? Twenty?

To raise this question is to raise another point about the sentencing. Broadly speaking, the culture of federal sentencing, pre-guidelines, is to use only a modest portion of the total possible range, and to sentence concurrently rather than consecutively. In departing from these pre-guidelines norms, in referring to the total legal range of 0-28 years and in sentencing consecutively, Judge Wood may have been attempting to give the Milken sentence a different cast than others, setting it apart because of its special nature. An alternative speculation is that, had she been accustomed to “pre-guidelines” thinking for 15 or 20 years, she would have internalized the standards many other judges were using, and the Milken sentence might have turned out differently.

One other feature of Judge Wood’s sentence for which there appears little basis in sentencing prior to the guidelines is her special emphasis on Mr. Milken’s committing his crimes in such a way as to escape detection. Other judges surely have given voice to the added culpability that stems from sophisticated plans, detailed schemes, and the like. But rarely have judges argued the special culpability of “stepping just over to the wrong side of the law” instead of engaging in cruder, and presumably less blameworthy, offenses. Perhaps we will need a new jurisprudence of cleverness. Does this reasoning hold only for complex frauds? Is the larcenist who steals from a store where he works more blameworthy than the burglar who breaks and enters? In any event, the special stress placed on the culpability of those who commit hard-to-detect crime is a creative addition to the common sense wisdom of pre-guideline sentencers. It is certainly consistent with models of economic rationality that would make the penalty go up as the detectability goes down.

Finally, what is there to say overall about this last of the great pre-guideline sentences. Surely we can see in the variety of interpretations that can be given Judge Kimba Wood’s reasoning, and in the lack of precise justification of the number, many of the reasons for wanting to move to a system in which the sentencing judge is more constrained by rules. But

this particular sentence represents something in addition, for in the end it is not a pre-guideline sentence after all. The guidelines tend to rule out the prior moral career of the defendant (as the prosecution reminded us) and apparently, so did Judge Wood. The sentencing guidelines give enormous weight to whether the defendant is of substantial assistance to the government, and Judge Wood gives substantial weight to that consideration. How much of the 10 years reflects Milken not yet having given full cooperation, and how much will it be reduced if he does?

Judge Wood, like the government, found it instructive to know what the guideline sentence for Michael Milken would be, and requested both parties to prepare their own estimates. (The defense estimate—*supra* page 163—is 21 to 27 months; the government’s estimate—*supra* pages 164 to 166—is 46 to 57.) And sure enough, depending upon what the parole board does, Milken’s sentence is not grossly out of line with what the government thinks the guidelines sentence would be. So this may be a pre-guidelines sentence, but it is certainly sentencing in the shadow of the guidelines. Perhaps this is not surprising, for although this is a pre-guidelines case, Judge Wood is not a “pre-guidelines” judge. Her confirmation as a judge took place six months after the guidelines went into effect.

The impact the guidelines will have on sentencing seems quite clear from the contrast between the adversarial biographies in the sentencing memoranda, and the adversarial computations in the guideline letters. In the former we see the whole person, good or evil depending on whose version we accept, but a whole person nonetheless. In the letter, the argument is reduced to whether the defense will agree that their client gets an additional 2 points for “more than minimal planning” and the government and the defense can argue over whether Milken’s role deserves a 2 point or a 4 point adjustment in the offense level. Adversarial biography produces no closure in outcome. The dimensions of adversarial biography, however, have been reduced by the guidelines to those few qualities that produce adjustments in the base level. For Judge Wood, the result might not have been so different had this been a guidelines sentence.

It is far too early to tell whether the net effect of the guidelines will be to reduce disparity, discrimination and inequality in sentencing, as many hope, or whether it will merely push these risks toward the prosecutor, as many fear. One such consequence, however, will be to reduce the adversarial search for the true nature of the offender—from a rendering of the whole person to arguments over a countable bundle of points.

FOOTNOTE

¹ Stanton Wheeler, Kenneth Mann, Austin Sarat, *Sitting in Judgment: The Sentencing of White Collar Criminals* (Yale University Press, 1988).

COMMENTARY

SUPREME COURT CONSIDERS THREE SENTENCING CASES

The Supreme Court granted certiorari in *Chapman v. U.S.*, 908 F.2d 1312 (7th Cir. 1990) No. 90-5744, a case which split the Seventh Circuit 6-5 in deciding to include the weight of the carrier medium in determining the appropriate guideline for selling an amount of LSD. The majority was written by Judge Frank Easterbrook, the main dissent by Judge Richard Posner.

The Court has already heard arguments in *Harmelin v. Michigan*, No. 89-7272 (November 5, 1990) (whether Michigan's mandatory life sentence for drug dealers violates due process or the Eighth Amendment) and *U.S. v. Burns* (D.C. Cir. 1990) No. 89-7260 (December 3, 1990) (whether a defendant has a constitutional or statutory right to prior notice of a sentencing judge's intention to depart from the guidelines). See 3 Fed. Sent. R. 111.

BLOCK—continued from page 117

CONCLUDING OBSERVATIONS

Karpoff and Lott, *infra*, find hidden in the stock market a possible rationale for why a dollar's worth of damage from fraud is not punished as severely as a dollar's worth of damage from environmental pollution in current practice. What transcends their specific finding is the general proposition that close examination of the pre-guidelines sentencing structure reveals substantial rationality and complexity in the way corporations are held accountable for various crimes. If one message comes across from all of the various drafts and all of the commentary it is how little we know about this existing structure of corporate sanctions.

If ever there is a case for anchoring sentencing guidelines in empirical reality, it is with guidelines for corporate defendants. There are now a respectable number of corporate cases to analyze. The Commission should use this empirical information to inform its setting of magnitudes for corporate guidelines.

FOOTNOTE

¹A second alternative is:

"(2) Gross pecuniary loss shall not be used for the determination of the guideline fine range if:

(A) the extent of the gross pecuniary loss was substantially greater than would have been anticipated by a reasonable person acting under the circumstances in which the [organization] [individual agent or agents of the organization who committed the offense] acted; or

(B) the offense conduct that triggered the organization's criminal liability involved neither intentional, knowing, nor reckless criminal conduct."

A third alternative is to revise §8C2.1(c)(1) to give courts greater flexibility in setting fines based on gross pecuniary loss, gross pecuniary gain, or the amount derived from the offense level table. For example, the following language might be used:

"(c) Select, from the following three amounts, the amount most appropriate to achieve the purposes of sentencing as set forth in 18 U.S.C. §3553(a)(2):

(1) the amount from the table below corresponding to the offense level determined under §8A1.2 (Application Instructions—Organizations);

(2) the gross pecuniary loss caused by the offense conduct of the defendant; or

(3) the gross pecuniary gain to the defendant from the offense."

Finally, the Commission might specify particular offenses or types of offenses for which a fine based on gross pecuniary loss would not ordinarily be appropriate to achieve the purposes of sentencing.