GENDER AND SENTENCING: WHAT WE KNOW AND DON’T KNOW FROM EMPIRICAL RESEARCH

Kathleen Daly*

Many claims are made about gender differences in sentencing. I review what we know and don’t know from empirical research on judicial paternalism and the more lenient treatment of women compared to men. I then challenge several arguments made by Ilene Nagel and Barry Johnson1 on equal treatment, special treatment, and gender-neutrality in sentencing policies. I conclude by suggesting the need to contemplate women as sentencing subjects.

1. Judicial Paternalism

When I began my dissertation research on gender and sentencing in 1979, I searched in vain for empirical research on judicial paternalism toward women, even though the scant literature suggested that female defendants received more lenient treatment than their male counterparts because of such paternalism. I learned that almost no one had actually interviewed judges to determine if they used different considerations in sentencing male and female defendants. Or, if such interviews had been done, the results were not available in published form.2

This lack of evidence inspired me in 1981-82 to conduct interviews with judges and other court officials (prosecutors, defense attorneys, and probation officers) in a western Massachusetts state court and to link those interviews with statistical analyses of pre-trial and sentencing outcomes in two state court jurisdictions, Seattle and New York City.3 From that body of work I found that judges and other court officials engaged in a form of paternalism designed to protect families and children rather than women per se. Court personnel assumed that male defendants received more lenient treatment than their male counterparts because of such paternalism. I learned that almost no one had actually interviewed judges to determine if they used different considerations in sentencing male and female defendants. Or, if such interviews had been done, the results were not available in published form.2

This lack of evidence inspired me in 1981-82 to conduct interviews with judges and other court officials (prosecutors, defense attorneys, and probation officers) in a western Massachusetts state court and to link those interviews with statistical analyses of pre-trial and sentencing outcomes in two state court jurisdictions, Seattle and New York City.3 From that body of work I found that judges and other court officials engaged in a form of paternalism designed to protect families and children rather than women per se. Court personnel assumed that familialized defendants (i.e., those who supported or cared for others) had greater informal social control in their lives than nonfamilied defendants; thus, familialized defendants were thought to be better probation risks. More importantly, by incarcerating familialized men or familialized women, court officials were concerned that they would “break up families” and “punish innocent family members.” I called their considerations the social costs of punishment, and I found that courts attached even greater social costs to removing caregivers than wage earners from families. Clearly, the seriousness of the offense and the defendant’s prior record were central sentencing elements; these were interwoven with and (depending on the seriousness of the case) would often be more determining than a defendant’s family and work situation.

I applied what I learned from the interviews to a statistical analysis of case outcomes, which included controls for the severity and type of offense and the defendant’s prior record. For the pretrial release and sentencing decisions,4 what initially appeared to be “sex effects” favoring women could be explained by the greater leniency return of having dependents for the familialized women than the familialized men. Put differently, while nonfamilied men and women were treated the same (in a statistical sense), familialized men and women were not. For the joint influences of gender, race-ethnicity, and family ties, I found that the mitigating effect of family ties was strongest for black women compared to white women and latinas whereas it was weakest for black men compared to white men and latinos.

My research raised questions for social policy, gender stereotyping and equal treatment. While family-based sentencing can be seen to reinforce notions of women’s (or men’s) “proper” place in family life, many female defendants do care for children (and male defendants support families). Not taking this reality into account would deny sources of connection and care we might wish to maintain. I also noted that practical questions were raised for the legal ideal of “equal treatment” when groups (i.e., the familialized defendants with responsibilities for dependents) versus individuals (the non-familied defendants without such responsibilities) were sentenced.

I concluded that, on balance, equal treatment of defendants whose responsibilities for others not only varied but differed by gender may be unjust.5 However, my conclusion differed from that of the U.S. Sentencing Commission. The Commissioners reasoned that by taking family ties into account, white, middle-class, employed defendants with intact families would be advantaged over others.6 While commentators have since challenged the wisdom of this stance,7 as I discuss in parts (3) and (4), the problem is even more complex than so far contemplated. The Commission’s formulation presumes men not women as subjects of sentencing.

2. Greater Leniency Toward Women

What does the research literature tell us about the treatment of women offenders? A colleague and I undertook a systematic review of all published studies of gender and non-capital sentencing, which used statistical methods, were written in English, and were published through mid-1990.8 Our analysis contained 50 cases, with a total of 249 outcomes.9 Of the 249 outcomes, only two showed differences favoring men.

* Fulbright Scholar at Law Program, Research School of Social Sciences, Australian National University; Senior Lecturer, Department of Criminology, University of Melbourne. My thanks to Nicola Lacey for her comments on an earlier draft.
Compared to a similar study of race and non-capital sentencing, we found that "sex effects" were more frequent than "race effects." As shown in Table 1, 45% of studies found sex effects favoring women, whereas 9% found race effects favoring whites.

Table 1: Comparison of Gender with Race in Sentencing Research (excludes studies of the death penalty; includes studies which control for prior record)

<table>
<thead>
<tr>
<th></th>
<th>Results from Kleck's (1981) review (race): (N=23)</th>
<th>Results from Daly &amp; Bordt's (1995) review (sex): (N=38)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effects favor whites</td>
<td>yes 9%</td>
<td>Effects favor women 45%</td>
</tr>
<tr>
<td>mixed</td>
<td>35%</td>
<td>mixed 29%</td>
</tr>
<tr>
<td>no</td>
<td>56%</td>
<td>no 26%</td>
</tr>
</tbody>
</table>

NOTE: "Mixed" refers to cases where one-third to one-half the outcomes showed effects favoring whites (Kleck) or women (Daly & Bordt).

The Daly & Bordt study (1995) showed that sex effects were as likely in those cases with more recent court disposition data (1976 to 1985) as in those with data collected before 1976. Sex effects were more likely in the decision to incarcerate (the "in-out" decision) than in the length of incarceration. Statistical procedures mattered: multivariate analyses that controlled for a variety of variables, especially prior record, attenuated sex effects. Yet even with such controls, sex effects remained. In general, when a research study focused on felonies, on felony courts, a study I recently completed that combines a statistical analysis with detailed case studies, I examined the sentencing practices of the New Haven felony court during 1981-86. My multivariate statistical analysis of a "wide sample" of cases showed an initial gender gap of 17 percentage points for the in-out decision. I then selected a deep sample of cases: 40 pairs of men and women who were accused and convicted of the same (or nearly the same) statutory offenses. The gender gap for the deep sample shrank to 10 percentage points. Further, upon reading the details of the matched pair cases more closely, I identified only one pair out of the 40 that evinced unwarranted gender disparity. The findings of the pair-wise analysis, when set alongside those of the statistical analysis, provided two stories of gender disparity in that court. While one would have concluded from the statistical analysis that women were sentenced more leniently than men, the pair-wise analysis suggested negligible (though not absent) gender disparity. Compared with my earlier research, the New Haven felony court research suggested that a defendant's familial

Recourse to statistical studies alone does not permit us to determine which option is correct. The merits of these options can be judged only after a closer look at the character and content of men's and women's cases in criminal courts. Only then can we tell whether a gender gap may or may not be warranted, may or may not reflect unmeasured sources of variation in men's and women's cases, or an amalgam of the two.

I became aware of problems in comparing "like crimes" for men and women while analyzing a dataset gathered by Stanton Wheeler and his colleagues on convicted defendants sentenced for eight "white-collar" offenses in seven urban federal courts during 1976-78. The character and content of men's white-collar crimes were more serious than women's, men more often worked in groups and used organizational resources, a higher fraction of men's cases involved indictments against corporations or businesses, and most employed men were managers or administrators while most employed women were clerical workers. Examining bank embezzlement, I found that 51% of the convicted male embezzlers were bank officers or managers, while 60% of the convicted female embezzlers were bank tellers. These and other features of gender-based variation in occupational position, role in the offense, and motive were not among the statistical controls in Wheeler and colleagues' analysis of sentencing. Their analysis focused on class-based variation and found that despite the many statistical controls they introduced, men were 25% more likely to receive an incarcerative sentence than women. This gender gap is on the high end; my review finds that the gap ranges from a low of 8 to a high of 26 percentage points. Let me return to the problem posed and discuss a study I recently completed that combines a statistical analysis with detailed case studies. I examined the sentencing practices of the New Haven felony court during 1981-86. My multivariate statistical analysis of a "wide sample" of cases showed an initial gender gap of 17 percentage points for the in-out decision. I then selected a deep sample of cases: 40 pairs of men and women who were accused and convicted of the same (or nearly the same) statutory offenses. The gender gap for the deep sample shrank to 10 percentage points. Further, upon reading the details of the matched pair cases more closely, I identified only one pair out of the 40 that evinced unwarranted gender disparity. The findings of the pair-wise analysis, when set alongside those of the statistical analysis, provided two stories of gender disparity in that court. While one would have concluded from the statistical analysis that women were sentenced more leniently than men, the pair-wise analysis suggested negligible (though not absent) gender disparity. Compared with my earlier research, the New Haven felony court research suggested that a defendant's familial
circumstances, though important, were secondary to pronounced gender differences in the crime and in the defendant's prior record.

A significant finding from the New Haven research was that men and women who were accused and convicted of statutorily similar crimes did not commit crimes of similar seriousness. For the 40 pairs, I judged 48% to be comparably serious; however, 40% of men's offenses were more serious than women's, and 12% of women's crimes were more serious than men's. In my analysis of the presentence reports, I found that women's biographies more often than men's displayed "blurred boundaries" between victimization and criminalization. Such a construction rendered some women's crimes less blameworthy; they seemed more serious than men's. In my analysis of sentencing transcripts, I also found that judges expressed more optimism about women's potential for reform. In both instances, judges and other court officials were responding to gender differences that were partly "real" and partly amplified.

Based on the empirical research to date, I reach two conclusions. First, with the best statistical controls or matching procedures, we should expect to find a gender gap of about 10 percentage points in the likelihood of incarceration for more serious (or felony) cases. With more information about the cases or judicial reasons for departing from sentence guidelines, that gap may grow smaller. Thus, I would caution policymakers and scholars against using statistical evidence alone in evaluating sentencing practices. If a study finds gender differences apparently favoring women, the authors should have the additional burden of demonstrating, by case analysis or in another way, that the statistical analysis is correct. Optimally, research studies should permit one to move between statistical analyses of aggregates and the details of cases.

Second, we need to document and understand how gender-linked determinations enter into court decisions, and whether these are just. By gender-linked I refer to those elements of cases and defendants' circumstances which can be seen to be linked to gender, but are not based solely on gender.

Most sentencing criteria are, in fact, gender-linked. For example, defendants with a previous record of violent arrests and convictions are (all else being equal) sentenced more harshly. More men than women have such records. Hence prior record is a gender-linked sentence determinant, although disparities that arise from using it are viewed as warranted. More contentious is whether the circumstances and conditions of defendants' lives or of justice system institutions (such as prison overcrowding) may be considered warranted sources of disparity. In considering certain other sex- and gender-linked criteria in sentencing (such as pregnancy, the care of children, and battering and male domination) as they relate to the prosecuted offense, we need not assume that men's and women's natures differ even though this was the basis of U.S. juvenile and criminal justice policy in the late nineteenth century and continuing through the 1950s. Rather, we are assuming that some features of men's and women's crimes and circumstances may differ, which could be acknowledged in sentencing.

In sum, my reading of the literature leads me to conclude that option (c) is the most sound. Unwaranted sources of gender disparity may exist, but they are likely to be far less prevalent than statistical studies lead us to think.

3. Equal Treatment, Special Treatment, and Gender Neutrality

Nagel and Johnson suggest that "despite the criticisms [concerning the methods of statistical studies], an impressive number of methodologically sound, multivariate studies report that women offenders receive preferential treatment" (p. 186). They argue that the "pattern of preferential treatment [of women] in sentencing" is at odds with "the value feminists attach to equality of treatment in wide variety of social and legal contexts" (p. 190). In addition, they say, such a pattern appears to fly in the face of the "primary goal" of the Sentencing Reform Act, i.e., "to eliminate unwarranted sentencing disparity" (p. 191). Nagel and Johnson remind us that "Congress instructed the Commission to 'assure that the guidelines and policy statements are entirely neutral as to the race, sex, national origin, creed, and socioeconomic status of offenders'" (p. 191, emphasis added by Nagel and Johnson). The implication of reforms such as the federal sentencing guidelines, which focus on the "offense" rather than "offender characteristics," is that the guidelines "if successful, will reduce the favorable treatment previously afforded female offenders, by increasing both their incarceration rate and the length of their sentence" (p. 182).

Let us consider these claims in light of what we know and don't know from empirical research, and then, from the feminist literature on equal treatment. From the empirical literature, we saw it is not certain that a pattern of preferential treatment of women exists. To be sure, there are gender sentencing gaps in about half the statistical studies, but these may largely reflect unmeasured sources of variation in men's and women's cases, coupled with warranted forms of disparity. Nagel and Johnson, however, have interpreted the research literature with option (a) in mind. They have concluded, I think wrongly, that there is a pattern of unwarranted favoritism toward women that needs to be corrected. They have offered a picture of past sentencing practices as affording leniency to women, whereas in the current regime of equal treatment, such leniency will diminish and
We could achieve equal treatment (in the form of women as the norm. More difficult is the meaning whether to emphasize formal or substantive equality ways: as having full moral agency by being punished women to adopt a male standard, while special what is wrong. Feminist scholars argue that equal women. I find it remarkable that in the name of bargaining, substantial assistance departures were not gender-neutral. blameworthiness of offenders, and in the conse
quences of punishment. The guidelines, however, are
beyond the dualism is apparent in Nagel and Johnson's discussion of sentencing policy. They imagine that women can be treated in one of two ways: as having full moral agency by being punished the same as men ("equal treatment") or as having an inferior moral agency, that of infants ("special treatment"). Surely, there must be a third option. The problem is that men's lives and the character of their crimes are taken as the norm, against which women are to measure up. Thus, the only way that women can be seen to achieve adult status is by "raising" their punishment to the male standard. This makes sense only if one assumes that men's lives and crimes— not women's—ought to be the norm. That assumption is at the heart of "equal treatment" and so-called "gender neutrality" in current sentencing policy.

Even if we assume that there are gender disparities in need of correction, why must we use men's average penalties as the norm? Why not women's? We could achieve equal treatment (in the form of similar statistical averages) just as easily by using women as the norm. More difficult is the meaning and implementation of gender neutral sentencing policies. By gender neutral I want to refer to policies that have considered the varied ways in which sex and gender matter in the social organization of crime, in the particular contexts in which crimes emerge, in the seriousness of crime, in the culpability and blameworthiness of offenders, and in the consequences of punishment. The guidelines, however, are not gender-neutral.

My claim is supported by Nagel and Johnson's assessment of guideline sentencing practices. They found that for drug offenses, women more often received downward departures for substantial assistance, downward departures for atypical facts or circumstances, and sentences from the low end of the sentencing range. Based on another study of plea bargaining, substantial assistance departures were interpreted as resulting from "prosecutors' and judges' greater sympathy for female offenders," and prosecutors' views that female defendants "deserved a less harsh penalty than that shown in the guidelines" (p. 218). For Nagel and Johnson, this is evidence that "special treatment [of women] . . . persists" (221). Alternatively, one could argue that because the guidelines have not been fashioned with women or variations in their lawbreaking in mind, legal officials are redressing this omission in their practices. That is, in the face of a male-normed sentencing scheme, some officials are trying to do justice to women.

What, one wonders, would the guidelines look like if women were the presumptive subjects of sentencing? Would the penalties be different? Would there be different concerns? Might women-normed guidelines be more humane, more defensible?

4. Women as Sentencing Subjects

Like many other commentators, Nagel and Johnson suggest that the reason women have only recently been visible in criminology and criminal justice policy is the "gender-skewed demographics of criminal behavior" (p. 182), that is, that about 80% of those arrested for crime are men; and they constitute an even higher share of arrests for violent crime (about 90%) and of those imprisoned (about 94%). Although common, the numbers argument is not persuasive. For example, while a very small number of people are on death row, there is a vast literature on the death penalty. We rightfully devote much attention to the death penalty because it raises moral questions about why we punish, what is humane punishment, and how we exercise discretion in the criminal justice system.

By analogy, we could imagine using women offenders as our referent group when considering the moral questions of punishment and the fashioning of a just sentencing policy. To date, neither the United States Sentencing Commission nor the states have taken this path. Quite to the contrary, Nagel and Johnson document the many ways in which women have been excluded from consideration as sentencing subjects. Here is a sampling:

- "While Congress was clear in its prescription of gender neutrality, the legislative history contains no discussion of the potential consequences of this neutrality for the overall severity of the sentences of female offenders." (p. 194)
- "One especially difficult sentencing issue . . . is how pregnancy is to be taken into account . . . The guidelines do not specifically mention pregnancy." (p. 197)
- "Another . . . gender-related issue . . . is the treatment of parenthood, particularly single parenthood. . . . The Commission does not collect data on offenders' single parent status . . ." (p. 200)
- "The Commission has wrestled with the issue of the non-coercive domination of female offenders.
by co-defendant males... Unfortunately, to this point the Commission has been unable to devise a satisfactory way to take this factor into account... while remaining faithful to the gender neutrality requirement of the statute." (p. 212)

Eight years after the guidelines have been established, the Commission is just beginning to wrestle with "unresolved dilemmas" and "the special issues posed by female offenders" (p. 215). It is, however, paradoxical, perhaps even futile, to attempt to devise policies that do not have a "disparate impact on female offenders" while simultaneously devising "facially neutral guidelines" (p. 221) that are, in reality, male-normed. It cannot be done.

We need to bring women fully into the picture. We need to consider all the varied circumstances and contexts in which women commit crimes in light of gender differences as well as overlap. We need to consider variations in women's responses to sanctions or threats of sanctions. For example, based on profiles of women in jail and prison, we know that, on average, women are less likely to be repeat offenders and to return to prison. This should matter in devising sanctions and in fashioning sentencing policy.

Legal commentators have focused on the "special issues" that emerge when considering women as sentencing subjects: pregnancy, parenthood, fighting back abusive mates, coercion and domination of male partners, and the conditions of incarceration, among others. These issues, though important, may spiral us further into legal minutiae while we sidestep the more important general questions: what kinds of penalties are appropriate, and what should be the aims and purposes of a justice system? We can revisit these questions freshly by putting women in the center of our crime and justice imaginations.

As an initial step, we might imagine a form of gender neutrality that is female-normed and a punishment scheme that is based on women's more law-abiding standard. I am not arguing that an abstract "male standard" should be replaced by a "female standard." This would reinforce gender dichotomies, presume that essential qualities of women and men exist, and limit our analysis of the circumstances of their lawbreaking; we can then apply what we have learned to men's lives and crime. In so doing, we can disrupt the imagery of the presumptive male as sentencing subject; we may see some men's lawbreaking in a different, perhaps more sympathetic light; and we can revisit questions of what is just and humane punishment. It would not be the first time that men could benefit from hiding behind women's petticoats.

NOTES

1 See Ilene N. Nagel & Barry L. Johnson, The Role of Gender in a Structured Sentencing System: Equal Treatment, Policy Choices, and the Sentencing of Female Offenders under the United States Sentencing Guidelines, 85 J. Crim. L. & Criminology 181 (1994). While I use this article as a focus of my critique, I am indebted to the authors for articulating the issues and providing an insider's view of the problems the Commissioners faced in fashioning the federal guidelines.

2 The one reference to judicial interviews was reported in Rita J. Simon, Women and Crime 49 (1975). According to Nagel & Johnson, supra note 1, at 187 n.30, Simon's assessment came from a student's unpublished master's thesis.

3 See Kathleen Daly, Structure and Practice of Familial-Based Justice in a Criminal Court, 21 Law & Society Rev. 267 (1987), and Rethinking Judicial Paternalism: Gender, Work-Family Relations, and Sentencing, 35 Gender & Society 9 (1989) [hereinafter Daly 1989a] [reporting on interviews of judges in Massachusetts and New York courts]. The statistical analyses are presented in Discrimination in the Criminal Courts: Family, Gender, and the Problem of Equal Treatment, 66 Social Forces 152 (1987), and Neither Conflict Nor Labelling Nor Paternalism Will Succeed: Intersections of Race, Ethnicity, Gender, and Family in Criminal Court Decisions, 35 Crime & Delinquency 136 (hereinafter Daly 1989b)]. At the time of case dispositions and interviews, the states studied had neither restrictions on judicial discretion nor sentencing guidelines.

4 Statistical "sex effects" (suggestive of leniency toward women) are more often found in pretrial release and sentencing decisions than in other contexts, such as the decision to prosecute (or not), to plead guilty (or not), to enter other forms of guilty pleas. The bulk of empirical research is, however, focused on sentencing.

5 Daly 1989a, supra note 3, at 11-12.

6 Nagel & Johnson, supra note 1, at 200-201.


8 See Kathleen Daly & Rebecca L. Bordt, Sex Effects and Sentencing: An Analysis of the Statistical Literature, 12 Just. Q. 141 (1995).

9 The counting unit (or case) was the data set, not the article. See id. at 142-45 (methodology of study).


11 The few studies of race and ethnic differences in sentencing women do not generally show race effects favoring white women. For a review of this literature, see Daly 1989b, supra note 3, at 139-43.

See Kathleen Daly, Gender and Varieties of White-Collar Crime, 27 Criminology 769 (1989). My study may explain the “curious” finding, reported by Nagel & Johnson at 220, supra note 1, that men were somewhat more likely than women to receive substantial assistance departures in bank embezzlement cases. I found that men’s embezzlements were more likely to be coordinated with others than women’s (25% and 6%, respectively) and that men more often played a subordinate role. I further speculated that men’s workplace crime groups were more hierarchical in nature, reflecting a mix of managers and clerical workers (p. 787).

See Daly & Bordt, supra note 8, at 157-58 (listing studies and size of gaps).

See Kathleen Daly, Gender, Crime, and Punishment (1994) [hereinafter Daly 1994].

The selection of the deep sample was made without knowledge of the sentencing outcomes. The offenses included homicide, assault, arson, harm to a child, larceny, robbery, and drug law violations.

My study examined the interweaving of race and gender to a greater degree than reported here. Both gender and race prejudice were evident in the one matched pair that displayed unwarranted disparity. Race-ethnic differences in assigning blameworthiness and reform potential were more evident for the men than for the women.

It is useful to distinguish the terms culpability and blameworthiness. I use culpability to refer to a defendant’s precise role or actions in an offense. Blameworthiness refers to a broader linking of the defendant’s biography (social history and prior record) to the offense. By defining blameworthiness this way, we can see how a defendant’s social history imposes meaning on a crime. (As such I would argue that it is not possible to separate “offender characteristics” from “the crime” as many assume.) This concept of blameworthiness stems from interviews of federal judges in the pre-guidelines era, reported in Stanton Wheeler, Kenneth Mann, & Austin Sarat, Sitting in Judgment (1988).

For the less serious or non-felony cases, the gap may be smaller or non-existent in statistical analysis.

For assessments in the large literature, see, e.g., Martha Fineman, The Illusion of Equality (1991); Deborah Rhode, Justice and Gender (1989); Lise Vogel, Mothers at Work (1993).

See Daly 1994, supra note 15, ch. 1 & 12.


From my New Haven research, I find that there are several male standards operating, not just one: the more conventional, familial man, and the more disreputable, victimizing, and heavily criminalized man. Daly 1994 p. 254.


More departures for female than male defendants were also documented under the Pennsylvania Sentencing Guidelines. See Darrell Steffensmeier, John Kramer, & Cathy Streifel, Gender and Imprisonment Decisions, 31 Criminology 411 (1993).


Using state and federal data, at year end 1993, there were 2,716 prisoners under sentence of death; at year end 1994, there were 1,053,738 prisoners, including 64,403 women. See Bureau of Justice Statistics, Correctional Populations in the United States, 1993 at 2 (1995) (death penalty data). Bureau of Justice Statistics, Prisoners in 1994 at 1 & 6 (1995) (prison data).

See Daly 1994, supra note 15, ch. 3-8. I emphasize the need to understand variation in both men’s and women’s lawbreaking, including points of overlap. In this way, we can avoid fixed notions of “male” and “female” crime and consider the multiple social contexts (gender, class, race-ethnicity, age, family situation) that promote opportunities and rationales for crime.


See Coombs, supra note 27, at 703-12, for examples.