Brian Forst and Shawn Bushway have done an excellent job of bringing together perspectives from various disciplines on discretion in criminal sentencing. That itself is a great service, and there is much that I agree with. What I will attempt to do in my comments is to continue building a more complete picture of the role discretion plays in the criminal justice system from my own slightly different disciplinary perspective of law and economics. First, I will fill out some of the picture of theoretical economic work provided by the authors. The most important theoretical contributions by economists in the area of crime have been on the costs and benefits of incarceration. However, to my knowledge none of that work has incorporated judicial discretion, prosecutorial discretion, or sentencing guidelines even though they are of obvious relevance to deterrence and incapacitation. I argue that we have learned some important lessons on sentencing from theoretical work, but that more work can be done because prior forays do not adequately reflect the institutional picture. Second, I will discuss briefly the challenge of reducing and measuring “unwarranted disparity” in sentencing. Reducing disparity was an important goal of the state and federal Guidelines, and there are substantial challenges to actually identifying the magnitude of unwarranted disparity and the success of the Guidelines in mitigating it.

The Economic Approach

When economists write about crime, both in theoretical work and empirical work, the focus has been largely on optimal sanctions. The theoretical work assumes that a criminal act will result, with some probability, in a sanction and the criminal incorporates this expected sanction into his or her calculus about whether and which crime to commit. The work generally assumes that policymakers can manipulate the size of the sanction by setting penalties or the probability that the sanction is applied by investing in enforcement.

In this deterrence work, economists have generally not studied directly how a sanction is actually determined, taking the judicial system as something of a “black box.” The probability of detection varies based on investments in enforcement, and the level of sanction is set by policymakers. Once detection is achieved, a pre-determined sanction is then imposed. Economic theorists have in general been uninterested in the effect of judicial or prosecutorial discretion. Forst and Bushway argue correctly that discretion is a crucial component of understanding what sanction is actually imposed, and a theoretical exploration of the black box combined with a deterrence model may yield important insights to deterrence theory.
The rare times in which economists have considered judicial discretion, it is generally used in empirical papers to aid identification of a causal impact. For example, Kling (2004) uses random assignment to a judge and differences between judges in sentencing to determine the effect of longer incarceration periods on later earnings and employment. As such, he can mimic a random experiment, and his results suggest that even large increases in incarceration length will have little impact on later job market performance. This is an excellent use of random assignment and shows that economists are not entirely unaware of issues around discretion.

One approach worth more discussion is that of Shavell (2007) discussing optimal judicial discretion (cited by the authors). Although the model is generally applicable, Shavell takes his motivation the Sentencing Guidelines. Basically, Shavell sets up a principal (Congress) and agent (judge) model, in which the principal recognizes a trade-off between individualized justice and consistency in sentencing. If judges observe relevant facts that cannot always be prescribed by statute, we want judges to retain the ability to enhance or reduce sentences accordingly. On the other hand, an individual judge’s preferences may be different from Congress’s or society’s. If a judge has too much discretion, he may act on his own preferences and undermine the social goals in sentencing. Thus, the Guidelines prescribe a sentencing range for relevant conduct. Some judges may act according to their preferences within this range, but that is a price that the principal (Congress) is willing to pay for particularized justice. This may help explain why we have no examples of fully determinative sentencing regimes, but almost all assign a sentencing range. Moreover, in the federal system, departures from the guidelines are reviewed more strictly than adjustments. Adjustments rest on findings of fact approved by Congress or the Sentencing Commission, whereas departures are less grounded and provoke greater concern. An additional layer not considered by Shavell is that of the Sentencing Commission itself, which as a quasi-administrative agency plays an important role not just in setting the original Guidelines, but also altering them as circumstances warrant. Bibas et al. (2008) suggest (informally) that federal Guidelines were a compromise in which Congress apportioned sentencing discretion between itself, judges, and the Commission.

Prosecutorial discretion has been studied in greater detail; in particular plea bargaining has been studied from a game theoretic perspective with regard to whether and how prosecutors might screen cases for prosecution. Baker and Mezzetti (2001) model how constraints on prosecutorial resources could affect the conviction rates of the guilty, and conclude that (under the right circumstances) more resources can increase accuracy. By contrast, Reinganum (1988) presents a game theoretic model of plea bargaining and determines that at times prosecutorial restrictions can result in better sorting of guilty from the innocent.
More recently, Reinganum (2001) suggests that sentencing guidelines, which are seen to limit judicial discretion and perhaps increase the discretion of prosecutors, will not necessarily change the plea rate. The intuition is that if prosecutors largely care about the number of trials as well as the sanction imposed, then increasing sanctions or giving prosecutors more control of the sanctions may not yield fewer trials because prosecutorial resources have not changed. Indeed, Reinganum observes that around the time of the Guidelines’ adoption, plea rates barely budged. My own cursory glance at plea rates around the time of Booker and its later progeny suggest that a move to advisory guidelines has likewise not changed the plea rate in any appreciable fashion. The plea rate was 95.2% the judicial term prior to Booker, 94.6% the judicial term of Booker, and 95.5% for the two full terms after Booker. In short, I am highly suspicious of the claim that guidelines are responsible for the disappearing criminal trial. Empirical evidence suggests they are not, and theory provides some clues as to why the Guidelines may not be the most important determinants of the trial rate. A much more likely source of the increasing plea rate is the exploding federal criminal caseload and limited resources at the state level. Perhaps more empirical work on the introduction of state guidelines regimes could provide us with further evidence of the effect of guidelines on pleas.

To sum up, sentencing is hardly mechanistic, even under the most rigid Guidelines in place today, but economists modeling deterrence have largely treated it as such. At times, this may be a justifiable abstraction because discretion may not be easy to incorporate into models, and my point does not undermine the insights provided by many of these models. Some of the work referenced above (my literature review is by no means extant), demonstrates that economists have considered prosecutorial and judicial discretion in more limited contexts of game theory and principal-agent models. However, theoretical work in the economics of deterrence is a well-traversed area, and I hope that future work may begin to build more realistic institutional designs in to these models. For example, perhaps discretion increases uncertainty, and this may itself be a deterrent if criminals are risk averse. Or when we think criminals are particularly risk-loving or incapacitation has high benefits, as may be the case in the most horrible offenses of violence, then the case for removing discretion may be strongest. Indeed, this may mimic some of the Guidelines system in place, in which there is an overlay of mandatory minimums for repeat offenders or crimes of great harm; whereas for crimes of a more calculating nature, less costly deterrence may be achieved by some degree of uncertainty in sentencing.

Measuring Disparity

One justification for sentencing guidelines is to reduce “unwarranted disparity” in sentencing. Basic principles of justice and equality before the law argue in favor of such a system. There are possible justifications for sentencing guidelines from a narrower social welfare perspective too. Excess punishment is costly, insufficient
punishment could undermine the benefits of deterrence or incapacitation. Of course, all of these arguments, deontological and consequential, were made during the debate over the federal Guidelines.

Unfortunately, economists, criminologists, and political scientists have had a difficult time determining the extent to which unwarranted disparity persists after the adoption of the federal guidelines. As Forst and Bushway document, studies have consistently found that unexplained racial and gender disparities persist after the Guidelines adoption. Well-known pre-Guidelines studies on state and federal sentencing practices have found similar evidence of racial and gender disparities. Schanzenbach (2005) could not tie racial or gender sentencing disparities to judicial demographics (race, gender, or political party) in federal sentencing, with a few exceptions. Fischman and Schanzenbach (2010a) confirm these results over a longer time span. This could be because unobserved factors explain most of the racial or gender disparity, but it is also consistent with a judicial selection process that filters for judges who have the same biases as their peers regardless of race or gender. Other work on state guidelines has suggested that disparity may vary with judge characteristics (see e.g., Steffensmeier and Herbert [1999]).

Identifying the source of an unwarranted disparity is by itself a challenge. How do we know that judges are responsible for disparity, rather than prosecutors, public defenders, or the Guidelines structure itself? In an important study that has not received as much attention as it deserves, Glaeser and Sacerdote (2003) found that punishments for homicide harsher when victims are white or female, not unlike the results of Baldus et. al (1990) and others regarding the death penalty. However, to answer criticism that unobserved victim provocation or blameworthiness could explain the disparity, the authors explored a subsample of vehicular homicides on the theory that in such cases victims were not typically blameworthy or provocative, and the authors found that the disparities in victim race and gender persist. Perhaps such a result is not surprising to many people, but it certainly undermines theories of optimal deterrence.

A more important question may be whether the introduction of guidelines reduced disparity. Attempts to look at the introduction of the Guidelines as a natural experiment are fraught with difficulty. At the federal level, the Commission will not release judge identifiers, which blocks perhaps the most straightforward way to study inter-judge disparity. Anderson et al. (1999), cited by the authors, had access to coded identifiers and provide some evidence of a decrease in disparity after the Guidelines introduction. However, the large increase in penalties, the uncertainty surrounding the viability of the Guidelines regime, and the overlay of mandatory minimums provide some confounding variation. Schanzenbach and Tiller (2007, 2008) and Fischman and Schanzenbach (2010b) suggest sentencing under the federal guidelines is sensitive to appellate review. When review is more relaxed or an appellate body is more
sympathetic to lower sentences, the differences between Republicans and Democrats on the federal bench increase. The authors take this as evidence that the Guidelines do in fact increase consistency, but it is not identical to a before and after comparison.

A point missed by scholars from many different disciplines is that the Guidelines provide an avenue for appellate review. Indeed, appellate review of sentences is the primary enforcement mechanism for guidelines sentences. Thus, one question that we may be able to answer is what happens when we increase or decrease judicial discretion within the guidelines strictures itself. A short, informal history of appellate review under the federal guidelines is as follows: (1) When the guidelines were first adopted, the decision to departure was evaluated under a fairly strict de novo standard; (2) in Koon v. United States (1996), the Supreme Court adopted a less-strict “abuse of discretion” standard; (3) the de novo standard was restored by Congress under the PROTECT Act (2003); and (4) the Supreme Court effectively restored an abuse-of-discretion standard in United States v. Booker (2005) which also declared the Guidelines to be advisory. Later Supreme Court cases (Rita and Gall [2007]) clearly encouraged district courts to depart from the Guidelines more often.

This back and forth transition between strict and more relaxed appellate review permits us to examine what happens to disparities as the guidelines become more or less binding. Figure 1 shows black-white odds ratios for departures with standard errors after adjusting for demographic and Guidelines variables. The values presented are the Race coefficients from the basic regression run for each Guidelines year controlling for offense type, district, criminal history, and base offense level. As the odds ratio gets closer to one, the black-white disparity decreases. The black-white odds ratio of a departure was basically flat in the years preceding Koon, and then increased substantially in the three years following Koon, though it trended downward. The move from the Koon abuse-of-discretion regime to the PROTECT Act seems to have had little effect, but there is a dramatic increase in the odds ratio after Booker.

Though departures are a key element of Guidelines sentencing, the substantively most important element is obviously prison sentences. Figure 2 reports prison sentences disparities. Here, there is a discernable decrease in black-white disparities after Koon, though disparities return to their pre-Koon levels at times. This is consistent with Figure 1, which shows that the black-white disparity in departures converged at this time as well. There is little or no effect immediately after Booker or from the PROTECT Act, but disparities appear to increase in 2007 and 2008, after Rita and Gall. For Hispanics, by contrast, there is no evidence of a persistent disparity or a change around Koon, the PROTECT Act, or Booker.

In short, it appears that increasing discretion of district court judges within the framework of the Guidelines does not necessarily increase disparities. In fact, judicial discretion can reduce disparity! This suggests that judges, when given greater
discretion, can offset sentencing disparity from other actors or from the structure of the Guidelines themselves.

On the other hand, the Commission claims that disparities have risen recently. That is reflected in the second year after Booker in prison sentences in Figure 2, though it does not really appear in departures in Figure 1. An increase in disparity may be attributable to the strong signals sent by the Court in Rita and Gall that the Guidelines were truly no longer binding. In such a case, the anchoring role that the Guidelines previously performed may be undermined.

**Conclusion**

Some discretion in criminal sentencing, arrests, and prosecution is necessary. The ultimate question is what are the costs and benefits of discretion in various contexts. This question has not been answered in a satisfactory way. Forst and Bushway have done a service by providing an integrated view of the multi-disciplinary work on discretion in sentencing. My hope is that the economics profession will pay closer attention to the institutional details of sentencing in their modeling efforts. Likewise, the usefulness and role of appellate review is still not well-understood, though from initial work it appears as though appellate review might be a useful tool for mitigating disparity.
Figure 1: Black-White Odds Ratio for Departure (with standard errors)
Figure 2: Prison Disparity (Months)

<table>
<thead>
<tr>
<th>Year</th>
<th>Black Prison Disparity (Months)</th>
<th>Hispanic Prison Disparity (Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995-96</td>
<td>Koon</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002-03</td>
<td>Protect</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005-06</td>
<td>Booker</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


