Technical Appendix for "Sentencing Guidelines, Judicial Discretion and Plea Bargaining"

Jennifer F. Reinganum*

Further Variations and Robustness Checks

Imperfect Information

Now consider the case in which information is merely imperfect rather than asymmetric. That is, suppose that the defendant and the prosecutor both know only the prior distribution $F(\cdot)$ of the random variable X. If a case should go to trial, the court will still observe a public signal which will accurately reveal the realized value x with probability $1 - \epsilon$, and will otherwise impose x^T (which will be x^J in the case of discretion and x^G in the case of guidelines). In this model with symmetric but imperfect information, the defendant and the prosecutor should always settle.

If the defendant makes the plea offer, the equilibrium sentence offer under imperfect information, denoted S^I, will be $S^I = (1 - \epsilon)\hat{x} + \epsilon x^T - k^P$. Under judicial discretion, a case coming to trial (for which the public signal is uninformative) will receive the sentence $\hat{x}$; a case coming to trial is an out-of-equilibrium event, but there is no information on which to update the court's beliefs. Thus, under judicial discretion, the equilibrium plea bargain will be $S^I = \hat{x} - k^P$; the defendant will benefit from the prosecutor's concern with saving trial costs and the lack of coordination between the prosecutor and the court.

Now consider the sentencing commission's decision. The expected social loss will only involve errors in plea bargain sentences, since no trials will occur in equilibrium.

$$ESL(x^T) = \alpha ESE^B(x^T) = \alpha \int (x - (1 - \epsilon)\hat{x} - \epsilon x^T + k^P)^2 f(x)dx,$$

where the integral is taken over the domain $[\underline{x}, \overline{x}]$. Minimizing this expression yields $x^G = x^B = \hat{x}$ +

k^P/ϵ, which makes the equilibrium plea bargain exactly $\hat{x}$.

Under the same assumption of imperfect information, if the prosecutor makes the plea offer then again all cases should settle, but now at $s^I = (1 - \epsilon)\hat{x} + \epsilon x^T + k^D$. Again judicial discretion would result in $x^T = x^J = \hat{x}$, but this now implies an equilibrium plea bargain of $\hat{x} + k^D$. In this case the prosecutor takes advantage of the defendant's concern with saving trial costs. The sentencing commission's objective of the expected social loss will now be given by

$$ESL(x^T) = \alpha ESE^B(x^T) = \alpha \int (x - (1 - \epsilon)\hat{x} - \epsilon x^T - k^D)^2 f(x)dx,$$

where the integral is taken over the domain $[\underline{x}, \overline{x}]$. Minimizing this expression yields

$x^G = x^B = \hat{x} - k^D/\epsilon$, which again makes the equilibrium plea bargain exactly $\hat{x}$.

Thus, in a regime of imperfect (but symmetric) information, the sentencing commission can induce the prosecutor to achieve the *ex ante* efficient outcome (i.e., all cases settle at $\hat{x}$) despite his private incentives to settle for too little (when the defendant makes the offer) or to demand too much (when the prosecutor makes the offer). This is not possible under asymmetric information. Although the sentences arising from both modes of conviction (plea bargains and trials) can be affected by x^T, reducing the expected social loss from sentencing error arising from one mode increases the expected social loss from sentencing error arising from the other mode. Since both modes occur with positive probability in equilibrium, a compromise must occur in the case of asymmetric information.

Alternative Prosecutorial Objectives

An interesting potential extension would be to modify the prosecutor's objective function in order to examine alternative motivations and their impact on the relationship between the sentencing commission and the judiciary. For instance, one could assume that the prosecutor was more motivated by expected sentences and less concerned about conserving trial costs. Formally, this would be equivalent to <u>lowering</u> k^P. This definitely results in a higher equilibrium plea offer function, since $S^*(x; x^T) = (1 - \epsilon)x + \epsilon x^T - k^P$. Moreover, lowering k^P also raises the equilibrium probability of trial, since $\rho(x) = 1 - \exp\{-(1 - \epsilon)(\overline{x} - x)/K\}$. Consequently, both x^J and x^B will be affected, but x^G will still lie between them as long as $\alpha \in (0,1)$, and hence the incentive to constrain judicial discretion (from below) will still exist. This is true even in the limiting case wherein the prosecutor doesn't care at all about saving trial costs (i.e., when $k^P = 0$). In this case, $S^*(x; x^T) = (1 - \epsilon)x + \epsilon x^T$ and $\rho(x) = 1 - \exp\{-(1 - \epsilon)(\overline{x} - x)/k^D\}$.

An alternative modification would be to assume greater alignment between the objectives of the prosecutor and the sentencing commission. If the prosecutor is also concerned about the expected squared sentencing error and trial costs, then given a plea offer of S and the prosecutor's beliefs $b(S; x^T)$, the prosecutor chooses the probability of rejection r to minimize his expected loss

$$\Pi^P(r, S; b(S; x^T)) = (1 - r)(b(S; x^T) - S)^2 + r[b(S; x^T) - (1 - \epsilon)b(S; x^T) - \epsilon x^T]^2 + rk^P.$$

If the offer S is not rejected (which occurs with probability 1 - r), then the prosecutor's loss is the squared difference between the sentence he believes is ideal, $b(S; x^T)$, and the sentence he accepts, S. If the offer S is rejected (which occurs with probability r), then the prosecutor's loss is the squared difference between the sentence he believes is ideal and the sentence he anticipates at trial, plus trial costs k^P. In this case, the differential equation characterizing $r^*(S; x^T)$ becomes very complex. However, the differential equation is readily solved for the special case in which the prosecutor does not care about saving trial costs, but only wishes to minimize the expected loss from sentencing error, which is given by

$$\Pi^P(r, S; b(S; x^T)) = (1 - r)(b(S; x^T) - S)^2 + r[b(S; x^T) - (1 - \epsilon)b(S; x^T) - \epsilon x^T]^2.$$

Incorporating the notion of consistent beliefs, the prosecutor will be willing to randomize if and only if

$(x - S)^2 = [x - (1 - \epsilon)x - \epsilon x^T]^2$. This equation has two roots, each of which is associated with a revealing equilibrium: $S^*(x; x^T) = (1 - \epsilon)x + \epsilon x^T$ and $S^{**}(x; x^T) = (1 + \epsilon)x - \epsilon x^T$. The first equilibrium involves the same plea offer function $S^*(x; x^T)$ and probability of rejection function $\rho(x) = 1 - \exp\{-(1 - \epsilon)(\overline{x} - x)/k^D\}$ that occur in the case of a pure expected-sentence-maximizing prosecutor. However, this equilibrium does not survive refinement using, for example, D1 or universal divinity. The second equilibrium (which does survive refinement) involves a plea offer function that gives the prosecutor the same expected loss as $S^*(x; x^T)$, but now the offer function $S^{**}(x; x^T) = (1 + \epsilon)x - \epsilon x^T$ is a <u>decreasing</u> function of x^T. The associated probability of rejection function is $r^{**}(S; x^T) = 1 - \{[k^D - 2\epsilon(\overline{S} - x^T)/(1 + \epsilon)]/ [k^D - 2\epsilon(S - x^T)/(1 + \epsilon)]\}^{(1+\epsilon)/2\epsilon}$, which now depends on x^T, making further analysis considerably more complex. However, since the equilibrium probability of trial is now given by $\rho(x) = 1 - \{[k^D - 2\epsilon(\overline{x} - x^T)]/[k^D - 2\epsilon(x - x^T)]\}^{(1+\epsilon)/2\epsilon}$, it is still true that the more severe cases are more likely to settle by plea bargain. Now, in addition, an increase in x^T (the anticipated sentence at trial in the event of an uninformative public signal) reduces the equilibrium probability that any defendant type goes to trial. Although a complete characterization of x^J and x^G for this case is beyond the scope of this paper, as long as the posterior is different from the prior, it follows that the sentencing commission will find it optimal to constrain judicial discretion in one direction or the other. The details follow.

<u>Analysis of the Prosecutor Who Minimizes Expected Squared Sentencing Error</u>

The prosecutor's objective is to choose r to minimize

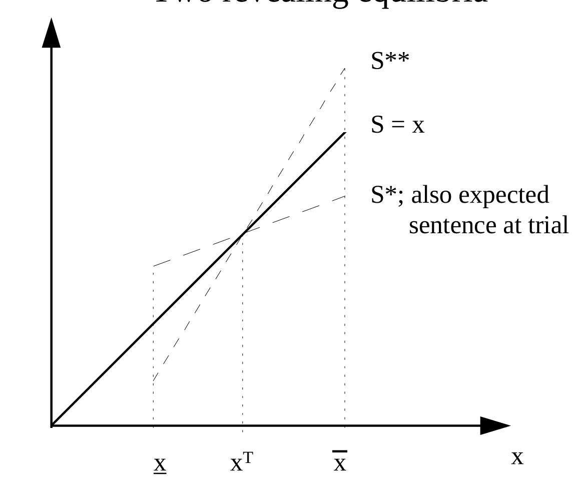
$$\Pi^P(r, S; b(S; x^T)) = (1 - r)(b(S; x^T) - S)^2 + r[b(S; x^T) - (1 - \epsilon)b(S; x^T) - \epsilon x^T]^2.$$

In a revealing equilibrium, the function $r(S; x^T)$ must be decreasing and thus the prosecutor's decision rule will involve randomization. The prosecutor will be willing to randomize if only if he is indifferent between the settlement and court outcomes. That is, if and only if

$(b(S; x^T) - S)^2 = [b(S; x^T) - (1 - \epsilon)b(S; x^T) - \epsilon x^T]^2$.

This equation has two solutions: $S^* = (1 - \epsilon)b(S; x^T) + \epsilon x^T$ and $S^{**} = (1 + \epsilon)b(S; x^T) - \epsilon x^T$. In equilibrium, beliefs must be correct, so $S^*(x; x^T) = (1 - \epsilon)x + \epsilon x^T$ and $S^{**}(x; x^T) = (1 + \epsilon)x - \epsilon x^T$. These are displayed below along with the ideal sentence, which lies along the 45° line.

Two revealing equilibria



The first solution involves the defendant offering exactly the sentence he expects to receive at trial. Notice that those with $x < x^T$ tend to "overpay" at trial (i.e., they receive more than the ideal sentence associated with their crime) and those with $x > x^T$ tend to "underpay" at trial (i.e., they receive less than the ideal sentence associated with their crime). The second solution involves the defendant offering not the same sentence he expects to receive at trial, but an alternative sentence which is equivalent from the prosecutor's point of view. In this case, the prosecutor knows the defendant with $x < x^T$ will overpay at trial (i.e., receive a longer than ideal sentence) by the amount $\epsilon(x^T - x)$. Thus, in the plea offer function $S^{**}(x; x^T)$, the defendant offers to underpay at settlement by the amount $\epsilon(x^T - x)$. Similarly, the defendant of type $x > x^T$ will underpay at trial (i.e., receive a shorter than ideal sentence) by the amount $\epsilon(x - x^T)$, and offers to overpay at settlement by the amount $\epsilon(x^T - x)$. The prosecutor is indifferent between these two outcomes, and hence is willing to randomize.

Claim 1. For the prosecutor's alternative objective function

$$\Pi^P(r, S; b(S; x^T)) = (1 - r)(b(S; x^T) - S)^2 + r[b(S; x^T) - (1 - \epsilon)b(S; x^T) - \epsilon x^T]^2,$$

the following strategies and beliefs form a revealing equilibrium.

(i) The equilibrium plea offer function is $S^*(x; x^T) = (1 - \epsilon)x + \epsilon x^T$ for $x \in [\underline{x}, \overline{x}]$,

with $\underline{S} = (1 - \epsilon)\underline{x} + \epsilon x^T$ and $\overline{S} = (1 - \epsilon)\overline{x} + \epsilon x^T$.

(ii) The equilibrium probability of rejection is $r^*(S; x^T) = 1 - \exp\{-(\overline{S} - S)/k^D\}$ for $S \in [\underline{S}, \overline{S}]$,

with $r^*(S; x^T) = 1$ for $S < \underline{S}$

and $r^*(S; x^T) = 0$ for $S > \overline{S}$ until S reaches $(1 + \epsilon)\overline{x} - \epsilon x^T$;

for $S > (1 + \epsilon)\overline{x} - \epsilon x^T$, $r^*(S; x^T) = 1$.

(iii) The beliefs are $b^*(S; x^T) = [S - \epsilon x^T]/(1 - \epsilon)$ for $S \in [\underline{S}, \overline{S}]$,

with $b^*(S; x^T) = \overline{x}$ for $S > \overline{S}$ and $b^*(S; x^T) = \overline{x}$ for $S < \underline{S}$.

Proof. First, note that the defendant offers exactly what he expects to receive as a sentence at trial. Since the prosecutor has consistent beliefs, for $S \in [\underline{S}, \overline{S}]$, the prosecutor is indifferent between settlement and trial and is therefore willing to randomize using $r^*(S; x^T)$. An offer of $S > \overline{S}$ is believed to have come from a defendant of type $\overline{x}$; since this defendant type underpays at trial, it is optimal to accept this offer unless $S > (1 + \epsilon)\overline{x} - \epsilon x^T$ (in which case the defendant would be overpaying more in settlement than he would underpay at trial, so the prosecutor must reject such an S). An offer of $S = \underline{S} - \delta$ (for positive δ) is also believed to have come from a defendant of type $\overline{x}$, which causes the prosecutor to reject it, since $[\bar{x} - (\underline{S} - \delta)]^2 > [\bar{x} - ((1 - \epsilon)\bar{x} + \epsilon x^T)]^2$.

Given the probability of rejection function $r^*(x; x^T)$ specified above, a defendant of type x always prefers the offer $S^*(x; x^T) = (1 - \epsilon)x + \epsilon x^T$ to any $S > \bar{S}$ or $S < \underline{S}$, since these surely result in trial (and a disutility of $(1 - \epsilon)x + \epsilon x^T + k^D$) while S* has at least a probability of being accepted. Choosing S from $[\underline{S}, \bar{S}]$ so as to minimize $\Pi^D(S, x; r^*(S; x^T))$ yields

$$\partial\Pi^D(S, x; r^*(S; x^T))/\partial S = 1 - r^*(S; x^T) + [\partial r^*(S; x^T)/\partial S][(1 - \epsilon)x + \epsilon x^T + k^D - S] = 0.$$

The unique solution to this equation is $S^*(x; x^T) = (1 - \epsilon)x + \epsilon x^T$. Moreover, the second order condition for a minimum, $\partial^2\Pi^D(S, x; r^*(S; x^T))/\partial S^2 > 0$, is also satisfied at $S^*(x; x^T) = (1 - \epsilon)x + \epsilon x^T$. Finally, the beliefs are consistent: $b^*(S^*(x; x^T); x^T) = x$ for all $x \in [\underline{x}, \bar{x}]$.

Recall that the beliefs assign any out-of-equilibrium offer to the defendant of type $\bar{x}$. Although these out-of-equilibrium beliefs support these strategies as a revealing equilibrium, they do not survive standard refinements such as D1 or universal divinity. To see why, characterize the <u>minimum probability of rejection</u>, denoted $\theta(x; S)$, which is necessary to deter a defendant of type x from defecting from his equilibrium strategy $S^*(x; x^T)$ to some S just below $\underline{S}$. This value is the smallest θ such that

$$(1 - \theta)S + \theta[(1 - \epsilon)x + \epsilon x^T + k^D] \geq (1 - \epsilon)x + \epsilon x^T + \rho(x)k^D,$$

where the r.h.s. is the type x defendant's equilibrium payoff $\Pi^D(S^*(x; x^T), x; r^*(S^*(x; x^T); x^T))$.

Thus, $\theta(x; S) = 1 - k^D(1 - \rho(x))/[(1 - \epsilon)x + \epsilon x^T + k^D - S]$. The denominator is positive for all $x \in [\underline{x}, \bar{x}]$, since S is assumed to be less than $\underline{S}$. It is straightforward to show that $\partial\theta/\partial x < 0$; that is, it is easier (i.e., requires a lower probability of rejection) to deter higher defendant types from defecting to S just below $\underline{S}$ than to deter lower defendant types from defecting. The type which is therefore most willing to defect is $\underline{x}$. Refinements such as D1 and universal divinity require that the out-of-equilibrium beliefs at such an offer S must be $\underline{x}$. But in this case, the prosecutor wants to accept S for sure (since the defendant of type $\underline{x}$ overpays at trial, and overpays an equal amount in settlement, settling at a sentence just below $\underline{S}$ would reduce the social loss). But if such an S is accepted, then all defendant types will defect to offering S, and the equilibrium unravels.

Claim 2. Assume that $2\epsilon(\overline{x} - x^T) \leq k^D$ (a stronger sufficient condition based only on exogenous parameters is that $2\epsilon(\overline{x} - \underline{x}) \leq k^D$)). For the prosecutor's alternative objective function

$$\Pi^P(r, S; b(S; x^T)) = (1 - r)(b(S; x^T) - S)^2 + r[b(S; x^T) - (1 - \epsilon)b(S; x^T) - \epsilon x^T]^2,$$

the following strategies and beliefs form a revealing equilibrium.

(i) The equilibrium plea offer function is $S^{**}(x; x^T) = (1 + \epsilon)x - \epsilon x^T$ for $x \in [\underline{x}, \overline{x}]$,

with $\underline{S} = (1 + \epsilon)\underline{x} - \epsilon x^T$ and $\overline{S} = (1 + \epsilon)\overline{x} - \epsilon x^T$.

(ii) The equilibrium probability of rejection function is

$r^{**}(S; x^T) = 1 - \{[k^D - 2\epsilon(\overline{S} - x^T)/(1 + \epsilon)]/ [k^D - 2\epsilon(S - x^T)/(1 + \epsilon)]\}^{(1+\epsilon)/2\epsilon}$, for $S \in [\underline{S}, \overline{S}]$,

with $r^{**}(S; x^T) = 1$ for $S < \underline{S}$ and $r^{**}(S; x^T) = 1$ for $S > \overline{S}$.

(iii) The beliefs are $b^{**}(S; x^T) = [S + \epsilon x^T]/(1 + \epsilon)$ for $S \in [\underline{S}, \overline{S}]$; any out-of-equilibrium beliefs support this equilibrium.

Proof. Recall that those defendants with $x < x^T$ overpay at trial (and hence offer to underpay at settlement by the same amount) and those with $x > x^T$ underpay at trial (and hence offer to overpay at settlement by the same amount). Given the prosecutor's beliefs, for $S \in [\underline{S}, \overline{S}]$ the prosecutor is indifferent between settlement and trial and is therefore willing to randomize using $r^{**}(S; x^T)$. Now consider out-of-equilibrium offers. If an offer of $S > \overline{S}$ is believed to have come from a defendant of type $x \in (x^T, \overline{x}]$, then it should be rejected since these defendant types already overpay in settlement (and underpay an equal amount at trial) and are now offering to overpay even more. If an offer of $S > \overline{S}$ is believed to have come from a defendant of type $x \in [\underline{x}, x^T]$, then it should be rejected since $S > \overline{S} = (1 + \epsilon)\overline{x} - \epsilon x^T$ is worse (from the prosecutor's viewpont) than the trial outcome $(1 - \epsilon)x + \epsilon x^T$. Thus out-of-equilibrium offers $S > \overline{S}$ should be rejected independent of beliefs. If an offer of $S < \underline{S} = (1 + \epsilon)\underline{x} - \epsilon x^T$ is believed to have come from a defendant of type $x \in [\underline{x}, x^T)$, then it should be rejected since these defendant types underpay in settlement (and overpay an equal amount at trial), and are now offering an even lower settlement. On the other hand, if an offer of $S < \underline{S}$ is believed to have come from a defendant of type $x \in [x^T, \overline{x}]$, then it should be rejected since $S < \underline{S} = (1 + \epsilon)\underline{x} - \epsilon x^T$ is worse (from the prosecutor's viewpoint) than the trial outcome $(1 - \epsilon)x + \epsilon x^T$. Thus out-of-equilibrium offers $S < \underline{S}$ should be rejected independent of beliefs.

Given the probability of rejection function $r^{**}(S; x^T)$ specified above, and given the parameter restriction $2\epsilon(\overline{x} - x^T) \leq k^D$, a defendant of type x always prefers the offer $S^{**}(x; x^T) = (1 + \epsilon)x - \epsilon x^T$ to any $S > \overline{S}$ or $S < \underline{S}$, since these surely result in trial (and a payoff of $(1 - \epsilon)x + \epsilon x^T + k^D$) while S^{**} has at least a probability of being accepted. Choosing S from $[\underline{S}, \overline{S}]$ so as to minimize $\Pi^D(S, x; r^{**}(S; x^T))$ yields

$\partial\Pi^D(S, x; r^{**}(S; x^T))/\partial S = 1 - r^{**}(S; x^T) + [\partial r^{**}(S; x^T)/\partial S][(1 - \epsilon)x + \epsilon x^T + k^D - S] = 0$. Use the relationship $x = b^{**}(S^{**}(x; x^T); x^T)$ to convert this to an ordinary differential equation in S only: x = (S

$+ \epsilon x^T)/(1 + \epsilon)$. This yields

$$1 - r^{**}(S; x^T) + [\partial r^{**}(S; x^T)/\partial S][\, k^D - 2\epsilon(S - x^T)/(1 + \epsilon)] = 0.$$

The unique solution to this equation is $S^{**}(x; x^T) = (1 + \epsilon)x - \epsilon x^T$. Moreover, the second order condition for a minimum, $\partial^2\Pi^D(S, x; r^{**}(S; x^T))/\partial S^2 > 0$, is also satisfied at $S^{**}(x; x^T) = (1 + \epsilon)x - \epsilon x^T$. Finally, the beliefs are consistent: $b^*(S^*(x; x^T); x^T) = x$ for all $x \in [\underline{x}, \overline{x}]$.