Technical Appendix for "Selecting Among Acquitted Defendants"

Andrew F. Daughety and Jennifer F. Reinganum

April 2015

Proof of Proposition 1

Suppose that one were to permit D to choose whether he will be tried under the standard (two-outcome) verdict or the Scottish (three-outcome) verdict. Moreover, suppose that this choice is made before P makes the plea offer. Then we argue that it is an equilibrium for both the G-type and the I-type to choose the Scottish verdict. To see why, suppose that both types choose the Scottish verdict; thus the mixture among those choosing the Scottish verdict is the same as the prior mixture (that is, a fraction λ are I-types and $1 - \lambda$ are G-types). Then their anticipated payoffs are given by the (three-outcome) equilibrium wherein the I-type rejects the plea offer for sure and the G-type does so with probability ρ_G^{D0}. Now consider what happens if a D deviates to choosing the standard verdict (an out-of-equilibrium move). If P and Θ believe that this deviation comes from both the G-type and the I-type in the prior mixture, then D's anticipated payoffs are given by the (two-outcome) equilibrium wherein the I-type rejects the plea offer for sure and the G-type does so with probability ρ_G^{D0} (because the G-type uses the same probability of rejection in both verdict regimes). Holding beliefs constant this way means that this deviation would be attractive to the G-type but not to the I-type. Updating these "base" beliefs marginally in the direction of the G-type would leave this preference ordering unchanged. No matter how much these beliefs were updated in the direction of the G-type, the I-type would not find the deviation attractive. Indeed, if these beliefs were updated sufficiently in the direction of the G-type, even the G-type would not find the deviation attractive. In particular, updating beliefs to place all of the probability on type G would deter G from making this deviation. This is because (following the deviation) P would make the plea offer[1] of $S_b = S_c(1 - F_G(\gamma_g)) + k^D$, which is larger than the plea offer in the three-outcome regime, $S_b = S_c(1 - F_G(\gamma_g)) + k^D + r^D\mu(G \mid g)(1 - F_G(\gamma_g)) + r^D\mu(G \mid ng)F_G(\gamma_{ng}) + r^D\mu(G \mid np)(F_G(\gamma_g) - F_G(\gamma_{ng})) - r^D$. On the other hand, updating the "base" beliefs marginally in the direction of the I-type would also leave this preference ordering unchanged. No matter how much these beliefs were updated in the direction of the I-type, the G-type would find the deviation attractive. Indeed, if these beliefs were updated fully in the direction of the I-type, then even the I-type would find the deviation attractive (as P would ultimately drop the case and Θ would impose no sanctions).

However, whenever an I-type would be willing to make this deviation, a G-type would strictly prefer to make the deviation. The D1 equilibrium refinement then requires that out-of-equilibrium beliefs attribute this deviation to type G. Basically, if both types are (in equilibrium) choosing the Scottish verdict, then deviating to the standard verdict is a clear signal of type G (under the refinement), which is then met with harsh informal sanctions and a higher plea offer. This is sufficiently disadvantageous to deter type G from deviating to the standard verdict from the Scottish verdict. Even though G prefers that all D-types be subject to the standard verdict, he will not unilaterally choose it if the I-type is not compelled to choose it as well. Thus, there is a (refined) equilibrium wherein both types of D choose the Scottish verdict.

---

[1] Technically, the plea offer is: $S_b = S_c(1 - F_G(\gamma_g)) + k^D + r^D\mu(G \mid g)(1 - F_G(\gamma_g)) + r^D\mu(G \mid ng)F_G(\gamma_g) - r^D$. However, if P and Θ believe that D is type G with probability 1 following this deviation to the standard verdict, then the last three terms sum to zero.

On the other hand, there cannot be a similar pooling equilibrium at the standard verdict. To see why, suppose that both types choose the standard verdict; thus the mixture among those choosing the standard verdict is the same as the prior mixture (that is, a fraction λ are I-types and $1 - \lambda$ are G-types). Then their anticipated payoffs are given by the (two-outcome) equilibrium wherein the I-type rejects the plea offer for sure and the G-type does so with probability ρ_G^{D0}. Now consider what happens if a D deviates to choosing the Scottish verdict (an out-of-equilibrium move). If P and Θ believe that this deviation comes from both the G-type and the I-type in the prior mixture, then their anticipated payoffs are given by the (three-outcome) equilibrium wherein the I-type rejects the plea offer for sure and the G-type does so with probability ρ_G^{D0} (because the G-type uses the same probability of rejection in both verdict regimes). Holding beliefs constant this way means that this deviation would be attractive to the I-type but not to the G-type. Updating these "base" beliefs marginally in the direction of the I-type would leave their preference ordering unchanged. No matter how much these beliefs were updated in the direction of the I-type, the I-type would still find the deviation attractive. Indeed, if these beliefs were updated sufficiently in the direction of the I-type, even the G-type would find the deviation attractive. In particular, updating beliefs to place <u>all</u> of the probability on type I would encourage type G to make this deviation. This is because (following the deviation) P would either drop the case immediately (if allowed) or else make a plea offer that is sure to induce rejection and would then drop the case (because of P's belief that puts probability 1 on type I). This is clearly better for a G-type than his equilibrium payoff in the standard verdict regime.

The only way to sustain the standard verdict as an equilibrium would be to attribute deviations to the Scottish verdict as coming from a G-type with a <u>sufficiently higher</u> probability (than the prior), as type I is willing to tolerate even a somewhat upward-revised weight on type G. Thus, we have shown that type I is willing to deviate to choosing the Scottish regime for a strictly larger set of beliefs (and ensuing best responses by P and Θ) than is type G. The D1 equilibrium refinement then requires that out-of-equilibrium beliefs attribute this deviation to type I. Basically, if both types are (in a putative equilibrium) choosing the standard verdict, then deviating to the Scottish verdict is a clear signal of type I (under the refinement), which is then met with no informal sanctions and a dropped case. But this is so advantageous that type G will also deviate to the Scottish verdict. Thus, it cannot be a (refined) equilibrium for both types to choose the standard verdict.

Other possible patterns of regime selection are easily shown to be impossible as part of an equilibrium to the overall game. For instance, consider a candidate for equilibrium wherein the I-type chooses the Scottish verdict and the G-type chooses the standard verdict. If this were the pattern, then the choice of the Scottish verdict is a clear signal of innocence and P would ultimately drop the case. But then G would do better by choosing the Scottish verdict, so this candidate cannot be part of an equilibrium. The same argument eliminates a candidate wherein the G-type chooses the standard verdict and the I-type mixes between the standard verdict and the Scottish verdict, as a realized choice of the Scottish verdict remains a clear signal of innocence. A somewhat perverse candidate involves thr I-type choosing the standard verdict and the G-type choosing the Scottish verdict. In this pattern, the choice of the standard verdict is a clear signal of innocence and P would ultimately drop the case. But then G would do better by choosing the standard verdict, so this candidate cannot be an equilibrium. Another candidate involves the I-type choosing the Scottish verdict and the G-type mixing between regimes. But now the realized choice of the standard verdict is a clear signal of guilt, which will be met with harsh beliefs by P and Θ and a higher plea offer in a regime that is worse for G. This is actually the worst possible outcome for G, so G will defect to choosing the Scottish verdict for sure and thus this candidate cannot be part of an overall equilibrium in the game.

The only remaining pattern would involve both the I-type and the G-type mixing between regimes. But we argue that it is not possible to make both types indifferent between the two verdict regimes at the same time. To see why, recall that the probability that a randomly-drawn D is type I is given by λ, and the probability that a randomly-drawn D is type G is given by 1- λ. Now suppose that D is allowed to choose between the standard and the Scottish verdict; moreover, suppose that both defendant types choose the standard verdict with the same probability, denoted σ. Then the probability that a randomly-drawn D, from among those that chose the standard verdict, is innocent (resp., guilty) is still λ (resp., 1 - λ), and similarly for the Scottish verdict. This leads to the same value of ρ_G^{D0} in both regimes.

In this putative equilibrium, both the G-type and the I-type receive payoffs equal to their expected losses from trial (the I-type always go to trial, whereas the G-type sometimes goes to trial and sometimes accept a plea offer, but these yield equal payoffs). We have previously shown that, for the same value of ρ_G^{D0} in both regimes, the G-type faces a lower expected loss from trial in the standard verdict regime, whereas the I-type faces a lower expected loss from trial in the Scottish verdict regime. Thus, beginning from a common fraction σ of both types choosing the standard verdict, it is not true that both types are indifferent between the two regimes.

How might they be made indifferent? First, consider the I-type, who currently prefers the Scottish verdict. In order to make the I-type indifferent between the regimes, one would have to shift the composition of D-types in the Scottish verdict regime towards more guilty (and/or fewer innocent) defendants, which results in a corresponding shift in the composition of D-types in the standard verdict regime towards more innocent (and/or fewer guilty) defendants.[2] This shift acts to equalize I's payoffs in the two regimes because when there is an increase in the relative frequency of the G-type in the population going to trial, observers impose harsher informal sanctions, regardless of the trial outcome. But in order to make the G-type, who currently prefers the standard verdict, indifferent between the two regimes, one would have to shift the composition of D-types in the Scottish verdict regime towards more innocent (and/or fewer guilty) defendants, which results in a corresponding shift in the composition of D-types in the standard verdict regime towards more guilty (and/or fewer innocent) defendants. That is, the change in composition needed to make an I-type indifferent is <u>inconsistent</u> with the change in composition that is required to make a G-type indifferent. Thus we conclude that both types of defendant cannot be made indifferent between the two regimes at the same time, and therefore there cannot be an equilibrium in which both types of defendant mix between choosing the standard verdict and the Scottish verdict.

---

[2] This would allow a somewhat lower value of ρ_G^{D0} in the Scottish verdict regime and would require a somewhat higher value of ρ_G^{D0} in the standard verdict regime, in order to just maintain P's incentives to take the case to trial following a rejected plea offer.

Now suppose that D makes the choice between the standard verdict and the Scottish verdict just prior to trial. Notice that at this point in the game, the plea offer has been rejected and P has chosen trial, so the only way that D's payoff can be affected is through the outside observers' beliefs about the likelihood that D is type G, given D's choice of regime. Since Θ conjectures that the I-type rejects the plea offer for sure and the G-type rejects it with positive probability (this must be true in order to incentivize P to choose trial following a rejected plea offer), the choice of verdict regime can signal D's type. Let $\rho_G^{D\Theta}$ denote Θ's conjecture about the probability that the G-type rejects the plea offer.

First, we argue that there is an equilibrium wherein both types of D choose the Scottish verdict. Suppose that both types of D choose the Scottish verdict for sure; then Θ's conjecture about the fraction of G-types among those that choose the Scottish verdict is still $(1 - \lambda)\rho_G^{D\Theta}/[(1 - \lambda)\rho_G^{D\Theta} + \lambda]$. If a G-type could deviate to the standard verdict and have Θ's conjecture remain the same, then G would do so (because, for the same conjecture $\rho_G^{D\Theta}$, G prefers the standard verdict to the Scottish verdict). On the other hand, if Θ's conjecture upon seeing a choice of the standard verdict increases the relative weight on the G-type, then this reinforces the I-type's preference for the Scottish verdict and undermines the G-type's preference for the standard verdict. Indeed, if Θ's conjecture is updated all the way to probability 1 on the G-type following a choice of the standard verdict, then G would also be deterred from deviating to choosing the standard verdict (as this would imply an informal sanction of r^D for sure, rather than $r^D\{\mu(G \mid g)(1 - F_t(\gamma_g)) + \mu(G \mid np)\Delta_t + \mu(G \mid ng)F_t(\gamma_{ng})\}$). On the other hand, if Θ's conjecture upon seeing a choice of the standard verdict decreases the relative weight on the G-type, then this reinforces the G-type's preference for the standard verdict and undermines the I-type's preference for the Scottish verdict. If Θ's conjecture is updated all the way to probability 1 on the I-type following a choice of the standard verdict, then the I-type would also be induced to deviate to choosing the standard verdict (as this would result in no informal sanctions). But since whenever an I-type would be willing to deviate to the standard verdict, a G-type strictly prefers to deviate to the standard verdict, the D1 equilibrium refinement implies that a deviation to the out-of-equilibrium standard verdict should be met with the outside observers' belief that the deviation comes from type G. Thus, such deviations are deterred and the equilibrium in which both types choose the Scottish verdict survives.

On the other hand, there cannot be a similar pooling equilibrium at the standard verdict. To see why, suppose that both types choose the standard verdict for sure; thus the mixture among those choosing the standard verdict is the same as the mixture among those rejecting the plea offer. Then their anticipated informal sanctions are given by the (two-outcome) equilibrium wherein Θ's conjecture is that the I-type rejects the plea offer for sure and the G-type does so with probability $\rho_G^{D\Theta}$. Now consider what happens if a D deviates to choosing the Scottish verdict (an out-of-equilibrium move). If Θ believes that this deviation comes from both the G-type and the I-type in the same mixture as rejected the plea offer, then D's anticipated informal sanctions are given by the (three-outcome) equilibrium wherein Θ's conjecture is that the I-type rejects the plea offer for sure and the G-type does so with probability $\rho_G^{D\Theta}$. Holding beliefs constant this way means that this deviation would be attractive to the I-type but not to the G-type. Updating these base beliefs marginally in the direction of the I-type would leave this preference ordering unchanged. No matter how much these beliefs were updated in the direction of the I-type, the I-type would still find the deviation attractive. Indeed, if these beliefs were updated sufficiently in the direction of the I-type, even the G-type would find the deviation attractive. In particular, updating beliefs to place all of the probability on type I would encourage type G to make this deviation. This is because (following the deviation) Θ would impose no sanctions (because of Θ's belief that puts probability 1 on type I). This is clearly better for a G-type than his equilibrium payoff in the standard verdict regime.

The only way to sustain the standard verdict as an equilibrium would be to attribute deviations to the Scottish verdict as coming from a G-type with a sufficiently higher probability (than the prior), as type I is willing to tolerate even a somewhat upward-revised weight on type G. Thus, we have shown that type I is willing to deviate to choosing the Scottish regime for a strictly larger set of beliefs (and ensuing best response by Θ) than is type G. The D1 equilibrium refinement then requires that out-of-equilibrium beliefs attribute this deviation to type I. Basically, if both types are (in a putative equilibrium) choosing the standard verdict, then deviating to the Scottish verdict is a clear signal of type I (under the refinement), which is then met with no informal sanctions. But this is so advantageous that type G will also deviate to the Scottish verdict. Thus, it cannot be a (refined) equilibrium for both types to choose the standard verdict.

Other possible patterns are also easily eliminated. For instance, there cannot be an equilibrium wherein the I-type chooses the Scottish verdict and the G-type chooses the standard verdict. If this were the pattern, then the choice of the Scottish verdict is a clear signal of innocence and Θ would impose no informal sanctions. But then G would do better by choosing the Scottish verdict, so this candidate cannot be part of an equilibrium. The same argument eliminates a candidate wherein the G-type chooses the standard verdict and thr I-type mixes between the standard verdict and the Scottish verdict, as a realized choice of the Scottish verdict remains a clear signal of innocence. A somewhat perverse candidate involves the I-type choosing the standard verdict and the G-type choosing the Scottish verdict. In this pattern, the choice of the standard verdict is a clear signal of innocence and Θ would impose no informal sanctions. But then G would do better by choosing the standard verdict, so this candidate cannot be an equilibrium. Another candidate involves the I-type choosing the Scottish verdict and the G-type mixing between regimes. But now the realized choice of the standard verdict is a clear signal of guilt, which will be met with harsh beliefs by Θ and an informal sanction of r^D for sure (which is higher than in the Scottish verdict), so G will defect to choosing the Scottish verdict for sure and thus this candidate cannot be part of an overall equilibrium in the game.

The only remaining pattern would involve both the I-type and the G-type mixing between regimes. But we argue that it is not possible to make both types indifferent between the two verdict regimes at the same time. To see why, recall that at this point in the game, Θ's conjecture is that all of the I-types and a fraction $\rho_G^{D\Theta}$ of the G-types rejected the plea offer. Now suppose that D is allowed to choose between the standard and the Scottish verdict; moreover, suppose that both defendant types choose the standard verdict with the same probability, denoted σ. Then Θ's beliefs about the probability that a randomly-drawn D is type G is the same, whether that draw is from among those that chose the standard verdict or from among those that chose the Scottish verdict.

In this putative equilibrium, the G-type faces lower expected informal sanctions in the standard verdict regime, whereas the I-type faces lower expected informal sanctions in the Scottish verdict regime. Thus, beginning from a common fraction σ of both G- and I-types choosing the standard verdict, it is not true that both types are indifferent between the two regimes.

How might they be made indifferent? First, consider the I-type, who currently prefers the Scottish verdict. In order to make the I-type indifferent between the regimes, one would have to shift the composition of D-types in the Scottish verdict regime towards more guilty (and/or fewer innocent) defendants, which results in a corresponding shift in the composition of D-types in the standard verdict regime towards more innocent (and/or fewer guilty) defendants. This shift acts to equalize the I-type's payoffs in the two regimes because when there is an increase in the relative frequency of G-types in the population going to trial, observers impose harsher informal sanctions, regardless of the trial outcome. But in order to make the G-type, who currently prefers the standard verdict, indifferent between the two regimes, one would have to shift the composition of D-types in the Scottish verdict regime towards more innocent (and/or fewer guilty) defendants, which results in a corresponding shift in the composition of D-types in the standard verdict regime towards more guilty (and/or fewer innocent) defendants. That is, the change in composition needed to make an I-type indifferent is <u>inconsistent</u> with the change in composition that is required to make a G-type indifferent. Thus we conclude that both types of defendant cannot be made indifferent between the two regimes at the same time, and therefore there cannot be an equilibrium in which both types of defendant mix between choosing the standard verdict and the Scottish verdict.