Appendix: Guilt and Guilty Pleas

Andrew T. Little* Hannah K. Simpson†

*Associate Professor of Political Science, UC Berkeley. andrew.little@berkeley.edu.

†Assistant Professor of Political Science, Texas A&M University. hannah.simpson@tamu.edu.

Contents

A Proofs

Proof of proposition 1

Proof. For part (i) Mas-Colell et al. (1995, Proposition 6.C.2) implies $c_G(\alpha)$ is strictly decreasing in α. Further, implicitly differentiating equation (**??**) gives that $\frac{\partial c_1}{\partial \pi_1} < 0$. By a similar analysis, $\frac{\partial c_0}{\partial \pi_0} < 0$. Combining these with equations (**??**) and (**??**) gives the result.

For part (ii), now writing the certainty equivalents explicitly as a function of the probability of being convicted, the guilty accept a wider range of plea deals if and only if:

$$
\begin{aligned}
&\bar{x}_{P,1} - \bar{x}_{P,0} > 0 \\
&b + c_0(\alpha, \pi_0) > c_1(\alpha, \pi_1)
\end{aligned}
$$

With constant absolute risk aversion, $c_1(\alpha, \pi) = b + c_0(\alpha, \pi)$ for any fixed π, and since c_0 is strictly decreasing in π the inequality holds. With increasing absolute risk aversion $c_1(\alpha, \pi) < b + c_0(\alpha, \pi)$ for any fixed π, so by the same argument a wider range of pleas are accepted.

For part (iii) with any interior π_1, the certainty equivalent must be on $(y_0 + b - x_K, y_0 + b)$ and so $\bar{x}_{P,1} \in (0, x_K)$. By a similar argument, $\bar{x}_{P,0} \in (0, x_K)$. As $\pi_1 \to 1$, going to trial when guilty results in conviction with certainty, and so $c_1(\alpha) \to u(y_0 + b - x_K)$, and $\bar{x}_{P,1} \to x_K$ and hence $\bar{x}_{P,1} > \bar{x}_{P,0}$. As $\pi_0 \to 0$, going to trial when innocent never results in conviction, and so $c_1(\alpha) \to u(y_0)$, and $\bar{x}_{P,0} \to 0$, again implying $\bar{x}_{P,1} > \bar{x}_{P,0}$. □

Proof of proposition 2 Follows directly from the argument in the main text that $C(b, \alpha)$ is strictly increasing in b and strictly decreasing in α.

Proof of proposition 3 First consider the probability of pleading guilty when innocent, which is:

$$
Pr(P = 1 | G = 0, A = 1) = \frac{Pr(\alpha > \hat{\alpha}_0, b < \hat{b}(\alpha))}{Pr(\alpha > \hat{\alpha}_0, b < \hat{b}(\alpha)) + Pr(\alpha < \hat{\alpha}_0, b < \hat{b}(\alpha))},
$$

When p_0 is small, π_0 does not affect the expected utility to either committing a crime or not, and so $\frac{\partial \hat{b}}{\partial \pi_0} \approx 0$. As a result the only part of this expression affected by π_0 is α_0, which is decreasing in π_0. So $Pr(P = 1|G = 0, A = 1)$ is increasing in π_0, and further it is zero when $\pi_0 = 0$ and is 1 when $\pi_0 = 1$.

$Pr(P = 1|G = 1, A = 1)$ depends on π_1 in a more complex manner since it affects the pool of individuals who commit crimes. Still, it also equals 0 when $\pi_1 = 0$ and 1 when $\pi_1 = 1$, and is continuous in π_1.

By the value in the limits of π_G and the continuity of both in π_G, what remains is showing that if $\pi_0 = \pi_1$ and individuals have constant or decreasing absolute risk aversion, then $Pr(P = 1|G = 0, A = 1) > Pr(P = 1|G = 1, A = 1)$.

If $\pi_0 = \pi_1$, then the guilty only differ in that they have a higher baseline consumption level. As a result, with constant or decreasing absolute risk aversion the difference between the lottery of a trial and the baseline consumption level b is increasing in b (Mas-Colell et al., 1995, Proposition 6.C.3, Part iii). This implies:

$$\begin{aligned}&\pi u(y_0 + b - x_K; \alpha) + (1 - \pi)u(y_0 + b; \alpha) - (y_0 + b - x_P)\\&\geq \pi u(y_0 - x_K; \alpha) + (1 - \pi)u(y_0; \alpha) - (y_0 - x_P)\end{aligned}$$

That is, the guilty get a higher relative payoff from going to trial and hence go to trial for a wider range of α: $\hat{\alpha}_0 \leq \hat{\alpha}_1(b)$ for all $b > 0$. Then we can write the probabilities of pleading guilty conditional on being guilty as:

$$\begin{aligned}Pr(P = 1|G = 1, A = 1) &= \frac{Pr(\alpha > \hat{\alpha}_1(b), b > \hat{b}(\alpha))}{Pr(b > \hat{b}(\alpha))}\\&\leq \frac{Pr(\alpha > \hat{\alpha}_0, b > \hat{b}(\alpha))}{Pr(b > \hat{b}(\alpha))}\\&= \frac{Pr(\alpha > \hat{\alpha}_0)Pr(b > \hat{b}(\alpha)|\alpha > \hat{\alpha}_0)}{Pr(b > \hat{b}(\alpha))}\\&< Pr(\alpha > \hat{\alpha}_0)\end{aligned}$$

In words, the guilty are less likely to accept a plea deal than a general member of the public is to have a risk aversion parameter where they would accept a guilty plea if innocent (unconditional on the crime decision). But it also must be the case that:

$$Pr(P = 1|G = 0, A = 1) = \frac{Pr(\alpha > \hat{\alpha}_0)Pr(b < \hat{b}(\alpha)|\alpha > \hat{\alpha}_0)}{Pr(b < \hat{b}(\alpha))} > Pr(\alpha > \hat{\alpha}_0))$$

Where the inequality follows because $\hat{b}(\alpha)$ is increasing in α.

Combining, when $\pi_0 = \pi_1$ and with constant or decreasing absolute risk aversion $Pr(P = 1|G = 0, A = 1) > Pr(P = 1|G = 1, A = 1)$. Since each of these probabilities is continuous in the π_Gs, starting at any $\pi = \pi_1 = \pi_0$ this inequality will continue to hold for some range as π_1 increases or π_0 decreases.

Proof of proposition 4 The optimal offers are proven in text. For part i, the probability of a guilty plea for both kinds of defendants is 1.

In part ii, we can compare the probability of pleading guilty when by comparing the CDFs of the evidence conditional on guilty status:

$$Pr(P = 1|G = 0) = F_e(e^T|G = 0) > F_e(e^T|G = 1) = Pr(P = 1|G = 1)$$

and so the innocent are more likely to plead guilty. □

Proof of proposition 5 The offers are proven in the main text. There is no sorting in the sense that $Pr(P = 1|G) = 1$ for both $G \in \{0, 1\}$.

Proof of proposition 6 First we derive the condition for when a screening offer is made, which depends on the ordering of the e threholds. First observe that the critical level of evidence e^{TS} that leads to indifference between screening and going to trial solves:

$$u_T(e^{TS}) - u_S(e^{TS}) = w\pi_1(e^{TS}) - \kappa = 0 \tag{1}$$

The key question to ask is whether or not, at this critical level of evidence, the prosecutor prefers making either of these offers to the plea-inducing offer. To answer this, plug e^{TS} into $u_S(e) - u_P(e)$ and determine when the resulting equation is larger than zero:

$$u_S(e^{TS}) - u_P(e^{TS}) = \frac{q(e^{TS})}{1 - q(e^{TS})} x_K(\pi_1(e^{TS}) - \pi_0(e^{TS})) + \pi_0(e^{TS})w - \kappa \geq 0.$$

From Equation 1, we know that at e^{TS}, $\kappa = w\pi_1(e^{TS})$. Plugging this value in for κ, the equation above simplifies to when:

$$q(e^{TS}) \geq \frac{w}{w + x_K}. \tag{2}$$

In other words, $e^{SP} < e^{TS}$ only if the benefit to trial wins is a small enough fraction of the total benefit the prosecutor obtains from a trial conviction, relative to the probability of guilt at evidentiary threshold e^{TS}.

What remains is proving the claims about the probabilities of pleading guilty given the guilt status. Part i follows from an identical argument as part ii of proposition **??**.

For part ii, the innocent plead guilty when $e \leq e^{TS}$, while the guilty plead guilty when $e \leq e^{SP}$. Perverse sorting happens when:

$$F_e(e^{TS}|G = 0) - F_e(e^{SP}|G = 1) > 0$$

we know that

$$F_e(e^{SP}|G = 0) - F_e(e^{SP}|G = 1) > 0$$

and so this will hold if, e.g., $F_e(e^{TS}|G = 0) - F_e(e^{SP}|G = 0)$ is sufficiently small. Since e follows a continuous distribution, this will hold if $e^{SP} - e^{TS}$ is sufficiently small.

When $\pi_0(e) = \pi_1(e) = 0$, $e^{SP} = e^{TS}$, and so this condition holds.

To see that this condition also holds when $\pi_1(e) - \pi_0(e)$ is "small," define:

$$d = \max_e \pi_1(e) - \pi_0(e)$$

As $d \to 0$, $u_S(e) - u_P(e) \to \pi_0 w$, and since $u_T(e) - u_S(e) = w\pi_1(e)$ the difference between these utility differences is $wd \to 0$. As a result, the solution to the crossing of the equations must also go to zero, i.e., $e^{SP} \to e^{TS}$, proving the result. $\square$

B Robustness

B.1 The single crossing condition

Recall that $u_T(e) - u_S(e) = w\pi_1(e) - \kappa$ which is monotone in e, and so the single crossing condition always holds for this comparison. Further, when $w > \kappa$, there is a unique crossing for some interior e^{TS}.

However, the two comparisons involving the plea offer are not necessarily monotone in e. In particular, if $\pi_0(e)$ increases rapidly for some values of e, this increases the utility associated with pleading out all defendants $\pi_0(e)x_K$ more rapidly than the other two strategies. Inspection reveals that as this derivative is not too positive for any e, then $u_S(e) - u_P(e)$ and $u_T(e) - u_P(e)$ are both increasing in e, which ensures a single crossing.

The implications of this condition being violated depend on whether $w < \kappa$ or $w > \kappa$.

$w < \kappa$ In this case the only relevant comparison is between u_S and u_P. In this case, when $w \to 0$ the difference becomes:

$$u_S(e) - u_P(e) = \frac{q(e)}{1 - q(e)} x_K(\pi_1(e) - \pi_0(e)) - \kappa$$

Since $\pi_G(e) \to 1$ but $\pi_1(e) > \pi_0(e)$, this difference will tend to decrease when e is close to 1, and because of the $q(e)/(1 - q(e))$ term this decline will tend to make the overall expression negative. The upshot is that when w is low it is possible for there to be two crossings, which means there can be a range of e where the screening offer is made. This means that when w is lower than κ there may now be correct sorting, which is consistent with our overall message that high w leads to perverse sorting.

$w > \kappa$ Increasing w increases $\frac{\partial u_S(e)-u_P(e)}{\partial e}$, and so tends to make the single crossing condition easier to meet. However, for some π_G functions it might still not hold. If so, the only snag to the proof of proposition **??** is that even if the utility for pleading out all defendants is higher than screening at the point where screening and going to trial meet (which again must be unique), the lack of a single crossing means that screening might also be a superior strategy to pleading out all defendants at some lower value of evidence. However, if we assume that not only are $\pi_0(e)$ and $\pi_1(e)$ are close, but their derivatives are close as well, this will preclude the possibility of multiple crossings. So, the general point that if the probabilities of conviction conditional on the evidence are "similar" this tends to lead to perverse sorting as well.

B.2 Accuracy Concerns

We limit our consideration of the role of accuracy concerns to the case in which $w > \kappa$. Start with the special case where $\pi_1(e) = \pi_0(e)$. Here, the prosecutor cannot engage in screening, so his only choice is between making the offer that all defendants accept, $x_P = \pi(e)x_K$, or taking all to trial by making some offer $x_P > \pi(e)x_K$. The former choice now comes with an accuracy cost $-(1-q(e))\psi$ that reflects the fact that some innocent people are pleading guilty. The latter choice now comes with the additional costs of wrongfully convicting some innocent defendants, $-(1-q(e))\pi_0(e)\psi$, *and* failing to convict some guilty defendants, $-q(e)(1-\pi_1(e))(1-\psi)$.

With accuracy concerns, the payoff to the prosecutor from pleading a defendant out at evidence e is

$$\pi(e)x_K - (1-q(e))\psi$$

and the payoff from taking a defendant to trial at evidence e is

$$\pi(e)(x_K + w) - \kappa - (1-q(e))\pi(e)\psi - q(e)(1-\pi(e))(1-\psi)$$

Comparing the prosecutor's utility from each strategy, he now derives greater value from taking a defendant to trial if

$$\pi(e)w + (1 - \pi(e))\,[\psi - q(e)] \geq \kappa.$$

For e sufficiently large, this is always met as before, because as $e \to 1$ the first term on the left-hand side approaches w, which is strictly larger than κ, while the second term approaches 0. However, for e sufficiently small, it may still sometimes hold: the first term approaches 0 as $e \to 0$, but the second term approaches $\psi \in (0, 1)$. Assuming a single-crossing condition similar to the one above, either the prosecutor prefers to try all defendants at all levels of e (when κ is very small) or he engages in perverse sorting, trying defendants above a threshold level of e, e_A^T, which solves

$$\pi(e_A^T)w + (1 - \pi(e_A^T))\psi - (1 - \pi(e_A^T))q(e_A^T) = \kappa,$$

and pleading them out below that threshold. Comparing to the threshold value e^T in the main text that solved:

$$\pi(e^T)w = \kappa,$$

this new threshold may be higher or lower. Specifically, $e_A^T < e^T$ if

$$(1 - \pi(e^T))(\psi - q(e^T)) > 0.$$

Intuitively, the evidentiary threshold for trial, given accuracy concerns, is lower (alleviating perverse sorting) when the posterior probability of guilt is low, the probability of conviction at trial is low, and the prosecutor's level of concern for wrongful convictions is high. Under other conditions, the threshold for trial may be higher (for example, if $\pi(e^T)$ and ψ are low, but $q(e^T)$ is high, meaning the prosecutor is most concerned with wrongly acquitting the guilty).

Now briefly consider the case where $\pi_1(e) > \pi_0(e)$. Intuitively, in this case, accuracy concerns increase the appeal of the screening offer relative to the other two, because (relative to the other options) the screening offer minimizes the likelihood of acquitting the guilty and maximizes the likelihood of acquitting the innocent. Assuming that each of the value comparisons we considered above changes signs at most once, we can broadly show that the conditions under which the screening offer is made grow broader: there is now a region of e where a screening offer is made if at the critical evidentiary level e_A^{TS} where a prosecutor concerned with accuracy is indifferent between trial and screening,

$$\frac{q(e_A^{TS})}{1-q(e_A^{TS})}x_K + \psi - w + \frac{1-\pi_1(e_A^{TS})}{\pi_1(e_A^{TS})-\pi_0(e_A^{TS})} \geq 0$$

A comparison to condition 2 confirms that, fixing e, this inequality holds for larger w. In other words, the benefit to trial wins must be larger here in order for the screening offer never to be optimal. If the condition does not hold, as before, the prosecutor takes defendants to trial if the value of trial exceeds the value from making an offer all defendants accept:

$$q(e)\left[w + x_K - \psi\right](\pi_1(e) - \pi_0(e)) + \pi_0(e)w - q(1-\pi_1(e)) + \psi(1-\pi_0(e)) - \kappa \geq 0.$$

In this case, however, it may now sometimes be the case that the prosecutor takes all defendants to trial, and there is no sorting. When the condition does hold, then there is correct sorting in the region of the parameter space where the prosecutor screens.

B.3 Combined Model

In this section we briefly examine how the emergence of perverse sorting in a model of strategic prosecution might change if defendants are allowed to vary in risk aversion. We consider this in the context of a fixed pool of defendants where the more risk-averse defendants are also more likely to be innocent, and also ask how the prosecutor behavior given a pool of defendants affects incentives to commit crime.

We return to defining the defendant's utility function as $u(y, \alpha)$, where y is the final consumption and α is the risk aversion. Here we abstract away from the monetary benefits of committing the crime and hence write the final consumption as $y_0 - x$, where x is the consumption equivalent of any sentence imposed (in a plea or at trial)

The prosecutor's utility function is unchanged from the main text, and the game proceeds as before. We maintain the assumption that the prosecutor does not observe the defendant's guilt but that both players observe the evidence e against the defendant, where we assume that e is drawn from a conditional distribution F_e such that $f_e(e|G)$ satisfies the strict monotone likelihood ratio condition (SMLRP).

B.3.1 Risk aversion known

First consider the case where the pool of defendants is fixed, both players observe the defendant's risk aversion α, and that α is likewise drawn from a conditional distribution F_α such that $f_\alpha(\alpha|G)$ satisfies the SMLRP (higher α implies lower risk aversion, and a higher probability of guilt). This assumption is primarily to simplify the analysis, and is broadly meant to capture the idea that experienced prosecutors have a decent sense of what kinds of pleas different defendants will accept based on observable characteristics.

We assume that the probability of conviction at trial is solely a function of the strength of the evidence, e, however, the prosecutor's posterior belief in the likelihood of the defendant's guilt, $q(e, \alpha)$ is now increasing in e and decreasing in α.

The defendant pleads guilty if the plea deal offered by the prosecutor, x_P, yields her a utility at least as great as the utility she expects from trial, given her guilt G. As in the crime choice model, it is convenient to represent this by defining the certainty equivalent of going to trial as a function of risk aversion and the probability of conviction $c(\alpha, \pi)$.

As before, this certainty equivalent is decreasing in both α and π. By definition, the defendant will accept an offer x_P if and only if $y_0 - x_P \geq c(\alpha, \pi)$, or

$$x_P \leq y_0 - c(\alpha, \pi) \equiv \bar{x}_P(\alpha, \pi)$$

Notice that as before, $\pi_1(e) \geq \pi_0(e)$; consequently, fixing α and e, the guilty always accept weakly worse plea deals than the innocent. As before, the prosecutor has three possible strategies: screening out the guilty, pleading out all defendants, or taking all defendants to trial.

For simplicity we look at the special case examined in the main text where $\pi_1(e) =$

$\pi_0(e) = \pi(e)$. Recall from the main text that here, for a given e and α, the prosecutor has only two choices: he can make a plea offer that both types of defendant accept, or he can make a plea offer that forces both types of defendant to go to trial. The prosecutor now prefers to induce a plea if:

$$\bar{x}_P(\alpha, \pi(e)) \geq \pi(e)(x_K + w) - \kappa \tag{3}$$

When the defendant is risk-neutral, this inequality simplifies to inequality (9) in the main text.

As α increases (as the defendant becomes more risk-averse), the maximum plea deal the prosecutor can offer increases (since the certainty equivalent of going to trial decreases, and $\bar{x}_P$ is decreasing in $c(\alpha, \pi)$). Since α does not affect the prosecutor's payoff from trial, this implies that for more risk-averse defendants, the prosecutor tends to prefer to use plea agreements more. And so if the innocent are more likely to be more risk-averse, this will lead to more perverse sorting above and beyond that driven by the mechanism driven by evidence quality.

Formally, analogous to the case in the main text, the prosecutor prefers to plead a defendant out for values of e such that equation 3 holds.

when $e = 1$ and hence $\pi(e) = 1$ this simplifies to

$$0 \geq w - \kappa,$$

which, since we've assumed here that $w > \kappa$, always holds. So for sufficiently high e, the prosecutor always wants to take the defendant to trial.

Similarly, at $e = 0$ this simplifies to

$$0 \geq -\kappa,$$

such that at any α, the prosecutor never wants to take the defendant to trial for low enough e. Defining the utility from trial minus the utility from a plea as $G(e) = \pi(e)(x_K + w) - \bar{x}_P(\alpha, \pi(e)) - \kappa$, we can show that so long as we assume that $G(e)$

only changes signs once (such that the prosecutor does not flip back and forth between preferring a plea and preferring trial as the evidence increases), for any α there is a unique e^* above which the prosecutor wants to take a defendant to trial and below which he wants to plead the defendant out, and moreover, this e^* is decreasing in α, such that the higher α (the more risk-seeking the defendant) the less evidence is required to take the defendant to trial. To see this, note that evaluated at e^*, $G'(e)$ must be increasing, since at $e = 0$, $G(e) < 0$ while at $e = 1$, $G(e) > 0$, which implies that $G(e)$ approaches 0 from below.

By the implicit function theorem,

$$\frac{\partial e^*}{\partial \alpha} = - \left. \frac{\frac{\partial G}{\partial \alpha}}{\frac{\partial G}{\partial e}} \right|_{e=e^*}.$$

From above, we know that $G'(\alpha) > 0$. Moreover we know that $G'(e)|_{e=e^*} > 0$. Consequently, $\frac{\partial e^*}{\partial \alpha} < 0$, i.e., the threshold level of evidence e^* above which the prosecutor prefers to take the defendant to trial is decreasing in the defendant's risk acceptance, or to rephrase, increasing in the defendant's risk aversion. The reason is that the greater the defendant's risk aversion, the larger the plea sentence the prosecutor can induce her to accept, and the higher the benefit to the prosecutor from inducing a plea.

Again, because α tends to be higher when the defendant is actually guilty, this suggests that the prosecutor will require a lower standard of proof to take the likely-guilty to trial than he will to take the likely-innocent to trial.

Crime choice Given this analysis, what can we say about the incentives to commit crime in the first place? In the case where $\pi_1(e) = \pi_0(e)$, the prosecutor will always either take all defenants to trial or make an offer which renders the defendant indifferent between accepting or going to trial. Further this offer only depends on the observed e and α, and not the perceived probability of guilt. As a result, citizens know that when committing a crime they will either not be caught, or will be caught and then receive a utility equal to what they would receive at trial. The utility from going to trial is decreasing in risk aversion since higher risk aversion decreases the certainty equivalent for going to trial. As a result, the risk averse are less apt to commit crime for two reasons: (1) they receive a worse schedule of offers as a function of the evidence againist them, and (2) for a fixed schedule of offers those who are more risk averse prefer the unrisky option of not committing a crime to the compound lottery of potentially being caught, and if caught receiving a distribution of offers given the evidence produced.

B.3.2 No information about risk aversion

Now consider the opposite case where the prosecutor does not observe the risk aversion of the defendant. Again assume that the probability of being convicted conditional on the evidence is not a function of guilt status.

What matters from the prosecutor perspective is the distribution of certainty equivalents for going to trial as a function of e; assume this follows a CDF $H(c; e)$, with density $h(c; e)$. An offer x_P is accepted if and only if $c \leq y_0 - x_P$, which happens with probability $H(y_0 - x_P; e)$.

This distribution changes as a function of e for two reasons: in addition to directly changing the probability of conviction, the prosecutor updates his belief about the likelihood of guilt and hence the distribution of risk aversion of the defendants.

The prosecutor expected utility for making an offer x_P is then:

$$Pr(x_P \leq y_0 - c(\alpha, \pi(e)))x_P + Pr(x_P > y_0 - c(\alpha, \pi(e)))\left[\pi(e)(x_K + w) - \kappa\right]$$

or

$$E[u_P(x_P, e)] = H(y_0 - x_P; e)x_P + (1 - H(y_0 - x_P; e))\left[\pi(e)(x_K + w) - \kappa\right]$$

So an optimal offer must solve

$$\frac{\partial E[u_P]}{\partial x_P} = -h(y_o - x_P, e)[x_P - (\pi(e)(x_K + w) - \kappa)] + H(y_0 - x_P; e) = 0 \tag{4}$$

While it is challenging to characterize how the solution to this equation changes as a function of e, there are some claims we can make relevant to the key question of when perverse sorting happens.

First, as $e \to 0$ (and hence $\pi(e) \to 0$) or $e \to 1$ (and hence $\pi(e) \to 1$), the outcome of the trial is certain, and hence the analysis is the same as in the case of no risk aversion. That is, for sufficiently weak evidence the prosecutor will make an offer all accept, and for sufficiently strong evidence he will take all to trial. Again, this will tend to lead to perverse sorting.

For intermediate e, risk aversion will affect the offer accepted, and there may be an interior solution solving equation (4). When such an offer is made, it will have the property in the first model where those who are more risk-averse are the ones who accept plea offers and those who are more risk-acceptant will go to trial. This will tend to lead to perverse sorting even for a fixed revelation of evidence, potentially strengthening the effect where the innocent/more risk-averse will tend to produce evidence that is more likely to lead to an attractive plea offer.

Crime choice Finally, in the version of the model where risk aversion is unobserved by the prosecutor, we can make some claims about what happens when we step back and consider how risk aversion affects the decision to commit a crime. If we assume the probability of being picked up when innocent is low, the payoff to not committing a crime is a certain y_0. Committing a crime now leads to a compound lottery where first a level of evidence is drawn, and then the prosecutor makes a plea offer based on that evidence, which will then lead to accepting the plea or a trial. If we assume the evidence generated is only a function of guilt (and hence there is no relationship between risk aversion and evidence conditional on guilt), then the distribution of plea offers when committing a crime does not depend on risk aversion. So, as in the first model, risk aversion will decrease the utility associated with committing a crime and accepting a plea deal, as well as the utility associated with committing a crime and going to trial. As mentioned in the discussion, there are some challenging signaling implications to work out if the prosecutor makes inferences about risk aversion based on their beliefs about the guilt of the defendant, but the general threshold structure of the citizen strategy must be the same in any equilibrium. In sum, the contention that the risk-averse are less likely to commit crimes should therefore hold up in this combined model.

References

Mas-Colell, Andreu, Michael Dennis Whinston, Jerry R Green et al. 1995. *Microeconomic theory.* Vol. 1 New York, NY: Oxford University Press.