Do Higher Defendant Reversal Rates Imply Appellate Court Plaintiphobia?

Bharat Bhole*,† and Bríd Gleeson Hanna*

Abstract

The objective of this paper is to show that, contrary to what Eisenberg and Heise (2009) claim, a pro-defendant case-mix at the trial level (due to, for example, higher defendant stakes) is consistent with the higher plaintiff trial win rate, and the higher defendant appeal and reversal rates that are observed in the Supplemental Survey of Civil Appeals (2001). This is true even when (i) the defendants are not more selective about which cases they push on appeal, and (ii) neither the trial court nor the appellate court is biased. Thus, the conclusion of appellate court "plaintiphobia" drawn by Eisenberg and Heise (2009) does not necessarily follow from the data.

1. INTRODUCTION

The Supplemental Survey of Civil Appeals (2001), made available by the Bureau of Justice Statistics and the National Center for State Courts, is a rich data set that follows a comprehensive cohort of state court cases from trial through their conclusion at the appellate level. This data set shows that plaintiffs have a higher trial win rate compared to defendants and that defendants have higher appeal and reversal rates compared to plaintiffs. Eisenberg and Heise (2009) have conducted the first rigorous empirical analysis of state appellate activity using this data set. Their main objective is to explain the significantly higher reversal rate of defendants; they suggest that defendants most likely have a higher reversal rate because the appellate courts are pro-defendant. They claim that the appellate courts are pro-defendant because these courts believe that the trial court adjudicators are pro-plaintiff and they attempt to correct for this perceived bias. Eisenberg and Heise (2009) refer to this phenomenon as "plaintiphobia" on the part of the appellate court.

*Department of Economics, Rochester Institute of Technology, 92 Lomb Memorial Drive, Rochester, NY 14623-5604, United States.

†Corresponding Author. Email: bharat.bhole@rit.edu

The conclusion reached by Eisenberg and Heise (2009), if true, says that the appellate courts' decisions are biased, whereas the trial courts' decisions are not.[1] The implications of this are significant. It means that the appellate courts, whose purpose is to correct the trial court errors, are in some sense worse than the trial courts in decision-making. This casts a doubt on the validity of the current adjudication process. Given the significance of this, we believe it is important to think carefully about the other plausible explanations of the observed pattern in the data.

In this paper we propose a pro-defendant case-mix at the trial level as a possible alternative explanation of the observed pattern in the data. Priest and Klein (1984) show that if defendant stakes are higher than plaintiff stakes then relatively more defendant than plaintiff victories ought to be observed in disputes that are litigated.[2,3] The types of trial court cases on which the Supplemental Survey of Civil Appeals (2001) is based suggest a case-mix where defendant stakes are likely to be higher than plaintiff stakes. For example, in tort cases such as product liability, premises liability, and malpractice cases, there is a good reason to believe that defendant stakes are higher than plaintiff stakes. As Priest and Klein (1984) note, "a firm accused of manufacturing a blatantly defective product may face substantial loss in future sales if a jury returns an adverse civil liability verdict." Even in contract cases, such as employment discrimination cases, a loss may require a defendant firm to change existing practices or may encourage further lawsuits, causing defendant stakes to be greater than plaintiff stakes.

In addition to the above reasoning, the hypothesis of pro-defendant case-mix is

[1]Based on studies such as Lempert (1998), Saks (1998) and Vidmar (1998), Eisenberg and Heise (2009) consider it unlikely that the trial court adjudicators suffer from a pro-plaintiff bias.

[2]"Defendant stakes" refers to the total loss, in addition to any litigation related expenditure, suffered by a defendant when he loses the case. By "plaintiff stakes" we mean the gain of the plaintiff when she wins the case, gross of any litigation related expenditure. For example, if the defendant stands to lose only the damages awarded by the court then the defendant stakes are equal to the plaintiff stakes. This is referred to as the case of symmetric stakes. If, in addition to the damages awarded, the loss also affects the defendant's reputation and, thus, future sales then his stakes will be higher than those of the plaintiff's. This is referred to as a case of asymmetric stakes.

[3]This idea that litigants with higher stakes experience a higher win rate in adjudicated cases is referred to as the *selection hypothesis*. For example, as Priest and Klein (1984) explain, "(w)here defendants stand to lose more from adverse verdicts than plaintiffs stand to gain, the relative calculus of the parties with respect to litigation and settlement changes. Defendants in general will be willing to offer greater amounts to settle disputes, but this effect will be relatively more pronounced for disputes in which defendants face a greater chance of losing and relatively less pronounced for disputes in which defendants face a greater chance of winning. As a consequence, relatively more disputes with likely plaintiff verdicts will be settled and relatively more disputes with likely defendant verdicts will be litigated."

also suggested by the findings in Waldfogel (1995). According to the structural parameter estimates of that study, about 30% of filed tort cases and 74% of contract cases would yield a plaintiff victory at trial. Of the 7,862 tort and contract cases in the Supplemental Survey of Civil Appeals (2001), 5,451 (69.3%) are tort and 2,411 (30.7%) are contract cases (see Appendix B).[4] Using these proportions as weights and the above estimates of plaintiff victory in Waldfogel (1995), the weighted probability of plaintiff victory is approximately 43.5 percent. This suggests that the trial case-mix is pro-defendant.[5,6]

This is in contrast with Eisenberg and Heise (2009), who claim that the selection hypothesis cannot explain the higher defendant reversal rate. Their rationale is that if there are asymmetric stakes and the selection hypothesis is true then the observed higher plaintiff win rate at the trial level suggests a pro-plaintiff case-mix and, therefore, higher stakes of plaintiffs than of defendants. Furthermore, they claim that these higher stakes, which make plaintiffs more selective about which cases they push at trial, should also make them more selective about which cases they push on appeal. This should result in a higher plaintiff reversal rate. Since this is not the case in the data, Eisenberg and Heise (2009) conclude that the selection hypothesis and asymmetric stakes cannot explain what is happening at the appeals stage.

This reasoning is, however, not necessarily correct. As we argue below, a higher observed plaintiff trial win rate is not necessarily inconsistent with a pro-defendant case-mix at the trial level; and, if the case-mix at the trial level is pro-defendant, that can result in higher defendant appeal and reversal rates, as observed in the data.

How is a higher plaintiff trial win rate possible with a pro-defendant case-mix if the trial court does not have a pro-plaintiff bias? The answer has to do with the variable that measures plaintiff wins in the data. Any trial in which the defen-

[4] We ignore property cases because (i) they are only 2 percent of all the trial cases; and (ii) within the property cases category, Waldfogel (1995) provides a win probability for only Intellectual Property cases and not for property cases as a group.

[5] The estimates in Waldfogel (1995) were obtained using federal civil cases from the Southern District of New York. There may be systematic differences between federal civil cases and civil cases in the state courts. However, if the differences are not large, then applying these estimates to the state court trial data of the Supplemental Survey of Civil Appeals can give a reasonably good idea of the case-mix.

[6] Daughety and Reinganum (2000b) present some reasons, other than the selection hypothesis, for a pro-defendant case-mix. They explain, for example, how lower evidence sampling costs of the defendants relative to the plaintiffs and asymmetry in the sampling distribution of the evidence can result in a pro-defendant case-mix.

dant is deemed liable and has to pay damages is recorded as a plaintiff win.[7] This does not always accurately reflect who really won the trial. In disputes that are mainly about the level of damages and not about liability (henceforth, these are referred to as "damages" cases), even if the defendant is correct and the trial court rules according to the defendant's expectations, the trial court record would show a plaintiff win.[8] The presence of such disputes can skew the recorded trial win rate in favor of the plaintiffs, even if the proportion of the cases in which the defendants' position is correct and, hence, the proportion of cases in which they "win" is larger.[9]

Once the case-mix at the *trial* level is pro-defendant a higher defendant reversal rate at the appellate level becomes possible even when (i) defendants are *not* more selective about which cases they push on appeal (an assumption made by Eisenberg and Heise (2009)); (ii) neither the trial court nor the appellate court is biased; (iii) both the defendants and the plaintiffs have the same cost of appeal; and finally, (iv) both the defendants and plaintiffs are equally biased about the socially optimal outcome being in their favor (i.e., about whether 'defendant win' or 'plaintiff win' is the socially optimal outcome).[10]

The reason why a pro-defendant case-mix at the *trial* level can result in a higher defendant reversal rate at the *appellate* level is the following. As the case-mix becomes pro-defendant, of all the cases that the plaintiffs lose, the proportion of cases in which they deserve to lose goes up. Since the appeals follow the losses, of all the plaintiff appeals, the proportion in which they deserve to lose goes up. This, given our assumption that the appellate court renders the correct decision in at least 50% of the cases, results in a decrease in the plaintiff reversal rate. Thus, in contrast with the conclusion drawn by Eisenberg and Heise (2009), we show that a pro-defendant case-mix at the trial level, either due to their higher stakes or some

[7]We confirmed this through personal communication with Paula Hannaford, Director of the Center for Jury Studies at the National Center for State Courts.

[8]For example, suppose there is a dispute in which both plaintiffs and the defendants are clear that the defendant is liable. However, while the plaintiff thinks that the damages should be $500,000, the defendant thinks they should only be $100,000. It is this disagreement over damages that results in both parties going to the court. Even if the court awards damages of only $100,000 in this case, the data would record a plaintiff win because the defendant was deemed liable and had to pay damages when, in fact, the defendant effectively won the trial.

[9]This problem does not occur at the appellate stage. In an appeal by the defendant about the level of damages, if the appellate court reduces the damages to a smaller but positive level, the record would reflect a reversal on defendant appeal despite the positive level of damages.

[10]We assume that the objective of both the trial court and the appellate court is to render the socially optimal verdict.

other reason, may indeed explain the pattern observed in the data.

The paper proceeds as follows. In Section 2 we present our model. In Section 3 we derive the expressions for the trial win rates, appeal rates and the reversal rates implied by our model. In Section 4 we use these expressions to show that a pro-defendant case-mix and the presence of "damages" cases can result in the pattern observed in the data even if the courts are unbiased. The proofs are in Appendix A and the summary statistics are in Appendix B.

2. THE MODEL

We adapt the model introduced in Bhole and Gleeson Hanna (2009) to address the issues considered in this paper.[11] Consider a situation where risk neutral litigants go to a trial. Assume that in any particular trial there are only two possible verdicts: a "correct" and an "incorrect" verdict. We should clarify that the correctness of the verdict may not always correspond to the truth in a particular case. For example, in a particular case it may be socially optimal to find a defendant liable only if at least two of the three pieces of evidence — e_1, e_2 and e_3 — are available.[12] In a case of this type, if only one piece of evidence, say e_1, is available then by a correct verdict/outcome we mean the verdict of "not liable", even if the defendant is guilty. Let S denote the state of nature. If the correct outcome is that the defendant be deemed not liable (resp. liable) or that the damages awarded be low (resp. high), then $S = S_0$ (resp. $S = S_1$). Let r (resp. $1 - r$) denote the fraction of trial court disputes in which $S = S_0$ (resp. $S = S_1$). If the litigants' stakes are asymmetric, and that affects their decision to settle or go to trial, then this will be reflected in the value of r. For example, if the defendants have higher stakes, this would result in a pro-defendant case-mix, causing $r > 1/2$. Let α denote the fraction of disputes

[11]The basic framework in the two papers is similar, but they focus on different issues. This leads to certain crucial differences. First, the trial, appeal and reversal rates are derived here such that they correspond to the way these rates are recorded in the data. In Bhole and Gleeson Hanna (2009) these rates are derived to reflect the actual rates. As noted in the Introduction, there can be a discrepancy between the actual rates and the recorded rates when some cases are "damages" cases and some are "liability" cases. No such distinction between case types exists in Bhole and Gleeson Hanna (2009). Second, in this paper we allow for asymmetric stakes, which is not allowed for in Bhole and Gleeson Hanna (2009).

[12]This may be socially optimal, for example, because considering only one piece of evidence sufficient to render a liable verdict may result in many actually innocent defendants being deemed liable and the likelihood of only one piece of evidence being present when the defendant is actually liable may be relatively low.

where it is clear to the litigants that the defendant will be deemed liable, but they cannot agree on the magnitude of damages. Henceforth, we refer to these cases as the "damages" cases. In the remaining $(1-\alpha)$ of the disputes either liability or both liability and damages are at issue. For convenience, we refer to these as the "liability" cases. As noted in the introduction, in the damages cases (i.e., in the fraction α of all the cases), the trial record always reflects a plaintiff win.

The Trial Court: The objective of the trial court is to render the correct decision. Let V_T denote the trial court verdict. With some abuse of notation, let $V_T = S_0$ (resp. $V_T = S_1$) when the trial court renders a verdict in favor of the defendant (resp. plaintiff). Let $t_{ij} = \Pr(V_T = S_i | S = S_j)$, $(i, j \in \{0, 1\})$, denote the probability that an *unbiased* trial court renders a verdict of S_i when the state of nature is S_j. Note that $t_{10} = 1 - t_{00}$ and $t_{01} = 1 - t_{11}$. We assume that

(A.1) $t_{00} > 1/2$ and $t_{11} > 1/2$
(A.2) $t_{00} = t_{11}$.

Assumption $t_{ii} > 1/2$, $i \in \{0, 1\}$ says that an unbiased trial court's accuracy is greater than what it would be if it made a decision based solely on, say, a toss of a coin. Assumption $t_{00} = t_{11}$ says that the probability that an unbiased trial court decides in favor of the defendant, when that is the correct outcome, is the same as the probability that the trial court finds in favor of the plaintiff, when that is the correct outcome.

The Appellate Court: As in the case of trial court, the appellate court's objective is to render the correct decision. Assume that the appellate court engages in a *de-novo* review of the trial court decision.[13] That is, it makes its decision solely based on the evidence; it does not consider the trial court's verdict. Let V_A denote the appellate court verdict. With some abuse of notation, let $V_A = S_0$ (resp. $V_A = S_1$) when the appellate court renders a verdict in favor of the defendant (resp. plaintiff). Let $a_{ij} = \Pr(V_A = S_i | S = S_j)$, $(i, j \in \{0, 1\})$. We assume that the accuracy of the appellate court's decision is such that

(A.3) $a_{00} > 1/2$, $a_{11} > 1/2$
(A.4) $a_{00} = a_{11}$.

These assumptions are similar to those made in the case of the trial court.

[13]This assumption is also made by Shavell (1995) and Daughety and Reinganum (2000a). Further, it can be shown that a Bayesian model of the appellate court (where the appellate court updates its beliefs about the state of nature, S, based on both its own signal about S and the trial court's decision) would give the same results as we obtain in the paper. Proof is available from authors upon request.

The Litigants: We only consider the litigants' decision after the trial court renders its verdict. We do not explicitly model the decision to go to trial, but account for the possibility that asymmetry in stakes or another factor affects case-mix by adjusting r.

It is assumed that the losing litigant appeals whenever the expected gain from appeal exceeds the cost of appeal. We now derive the expected gain from appeal, for which we first need to define some notation. Let q^i_D (resp. q^i_P) denote $P_D(S_0|S_i)$ (resp. $P_P(S_0|S_i)$). That is, q^i_D (resp. q^i_P) denotes the prior belief of the defendant (resp. plaintiff) that the state is S_0, when it is, in fact, $S_i\,(i \in \{0,1\})$.[14] Hence, $q^0_D \equiv P_D(S_0|S_0)$, $q^0_P \equiv P_P(S_0|S_0)$, $q^1_D \equiv P_D(S_0|S_1)$ and $q^1_P \equiv P_P(S_0|S_1)$). The assumption of differences in beliefs across plaintiffs and defendants is a standard one in divergent expectations theories of litigation.[15]

We assume

(A.5) $q^0_D > q^1_D$ and $q^0_P > q^1_P$.

This assumption says that both the defendant and the plaintiff will have a greater belief that the true state is S_0 when it is, in fact, S_0 (see footnote 14). After the trial court renders its decision, both plaintiffs and defendants update their beliefs about the state S. Let $\hat{q}^{ij}_D \equiv \hat{P}_D(S_0|V_T = S_i, S = S_j)$ (resp. $\hat{q}^{ij}_P \equiv \hat{P}_P(S_0|V_T = S_i, S = S_j)$) denote the posterior belief of the defendant (resp. plaintiff) that the state is S_0, when the trial court's verdict is $V_T = S_i$, and the true state is $S = S_j$, where $i, j \in \{0,1\}$. Assuming Bayesian updating and that the litigants are aware of the trial court's average accuracy, the defendant's posterior belief, $\hat{q}^{1j}_D$, is

$$\hat{q}^{1j}_D = \frac{q^j_D t_{10}}{q^j_D t_{10} + (1 - q^j_D) t_{11}}, \tag{1}$$

[14] Prior beliefs refer to beliefs before the trial court renders its verdict. While the litigants do not know the true state with certainty, the circumstances are likely to be different in the two states and, hence, the beliefs, which are likely to be formed on the basis of these circumstances, may be different. For example, consider a situation in which Teddy and Dolly are involved in an accident. Dolly is not one hundred percent certain whether Teddy is at fault. But it seems reasonable to assume that her belief that Teddy is at fault would be greater if Teddy is legally at fault than when he is legally not at fault. This could be because circumstances, such as witnesses who claim that Teddy is at fault, are more likely when he is legally at fault.

[15] See, for example, Priest and Klein (1984).

where $j \in \{0,1\}$.[16,17] Similarly, the plaintiff's posterior belief $\hat{q}_P^{0j}$ is

$$\hat{q}_P^{0j} = \frac{q_P^j t_{00}}{q_P^j t_{00} + (1 - q_P^j) t_{01}}, \tag{2}$$

where $j \in \{0,1\}$. From both (1) and (2), it follows that if the trial court rules in the plaintiff's favor then the defendant's posterior belief that the state is S_0, is smaller than his prior belief. Similarly, if the trial court rules in the defendant's favor then the plaintiff's posterior belief that the state is S_0, is larger than her prior belief.

Let H_0 (resp. H_1) denote the damages awarded by the trial court when it decides in the defendant's (resp. the plaintiff's) favor. Clearly, $H_0 < H_1$. Let $\Delta H = H_1 - H_0$. To allow for asymmetry in stakes, let K denote the amount by which the defendant's stakes exceed those of the plaintiff. That is, if the defendant loses, his loss is $H_1 + K$, whereas the plaintiff's gain is H_1. Also, assume that the litigants are aware of the average accuracy of the appellate court's decisions, which we have referred to as a_{ij} $(i, j \in \{0,1\})$. From these assumptions, it follows that if the state is S_j $(j \in \{0,1\})$ and the plaintiff loses in the trial court, then her expected gain from an appeal is

$$EG_P^j = (\hat{q}_P^{0j} a_{10} + (1 - \hat{q}_P^{0j}) a_{11}) \Delta H. \tag{3}$$

Similarly, if the state is S_j $(j \in \{0,1\})$ and the defendant loses in the trial court then, with symmetric stakes, his expected gain from an appeal is

$$EG_D^j = (\hat{q}_D^{1j} a_{00} + (1 - \hat{q}_D^{1j}) a_{01}) \Delta H. \tag{4}$$

With asymmetric stakes, the defendant's expected gain from appeal is

[16] The assumption of Bayesian updating is not necessary for our results. We can obtain similar results if we simply assume instead that the defendant's (resp. plaintiff's) posterior belief that the state is S_0 is smaller (resp. larger) than his prior belief when the trial court rules against the defendant (resp. plaintiff).

[17] The numerator in equation (1) is the product of (i) the defendant's prior belief that the state is S_0 when the true state is, in fact, S_j, and (ii) the defendant's belief that the trial court will rule in the plaintiff's favor, conditional on his belief that the state is S_0. The denominator is the defendant's belief that the trial court will rule in the plaintiff's favor when the true state is S_j. Equation (2) is similarly obtained.

$$EG_D^{jK} = EG_D^j + \delta_j K, \tag{5}$$

where $\delta_j = (\hat{q}_D^{1j} a_{00} + (1 - \hat{q}_D^{1j}) a_{01})$ denotes the defendant's belief in state S_j that the appellate court will rule in its favor. The losing litigant compares the expected gain from appeal with the cost of appeal and files an appeal whenever the former is higher. From assumptions (A.3) and (A.5), it follows that $EG_P^0 < EG_P^1$ and $EG_D^0 > EG_D^1$.

To focus on the effect of a pro-defendant case-mix at the trial level, we assume that there are no differences in other factors that may influence the appeal and reversal rates. Specifically, we assume that the cost of appeal is the same across plaintiffs and defendants and that they are equally biased about the state being in their favor when it is, in fact, not in their favor. We denote by $F(\cdot)$ the distribution of the litigants' cost of appeal. Then the probability of appeal given a defendant (resp. plaintiff) loss is $F(EG_D^0)$ (resp. $F(EG_P^0)$) when the state is S_0; and $F(EG_D^1)$ (resp. $F(EG_P^1)$) when the state is S_1. To express the equality in their bias we assume

(A.6) $q_D^0 = 1 - q_P^1$ and $q_D^1 = 1 - q_P^0$.

Assumption (A.6) says that the prior belief of the defendant that the state is S_0, when it is in fact S_0 (resp. S_1), is the same as the prior belief of the plaintiff that the state is S_1 when it is in fact S_1 (resp. S_0).

3. IMPLICATIONS FOR TRIAL WIN RATES, APPEAL RATES AND REVERSAL RATES

In this section, we derive the expressions for the trial win rates, appeal rates and reversal rates. In deriving the expressions for various rates we should keep in mind that, as we discussed in the introduction, there can be a discrepancy between the actual win rates and the recorded win rates of the two parties. Since we want to compare the trial, appeal and reversal pattern implied by our model with that in the data, the expressions for the various rates are derived such that they correspond to the way these rates are measured in the data.

Trial Win Rates: The defendants' trial win rate is

$$\Pr(\text{D win at trial}) = (1 - \alpha)(rt_{00} + (1 - r)t_{01}). \tag{6}$$

As explained in Section 2, fraction α of disputes are about damages. The defendant is deemed liable in these cases and, hence, they are recorded as plaintiff wins. Of the remaining fraction $(1-\alpha)$ of the disputes, where liability is also an issue, in fraction r (resp. $1-r$) the correct outcome is in the defendants' (resp. plaintiffs') favor; of these cases, the trial court decides fraction t_{00} (resp. t_{01}) in the defendants' favor. Thus, of all the cases, fraction $(1-\alpha)(rt_{00}+(1-r)t_{01})$ are recorded as having been decided in the defendant's favor. The plaintiffs' trial win rate is given by

$$\Pr(\text{P win at trial}) = \alpha + (1-\alpha)(rt_{10} + (1-r)t_{11}). \tag{7}$$

In the "damages" cases (fraction α of the cases) plaintiffs are always recorded as having won the case. In the remaining "liability" cases, the plaintiff trial win rate is determined in a similar manner as for the defendants.

Appeal Rates: The defendants' appeal rate (of the cases that the defendants lose, the fraction in which they appeal) is given by

$$\Pr(\text{D appeal}|\text{D loss}) = \frac{rt_{10}F(EG_D^0 + \delta_0 K) + (1-r)t_{11}F(EG_D^1 + \delta_1 K)}{\alpha + (1-\alpha)(rt_{10} + (1-r)t_{11})}. \tag{8}$$

The numerator is the joint probability of the defendant losing and appealing the case. A case is appealed by a defendant when (i) he *actually* loses the case (as opposed to just being recorded as having lost the case) and (ii) the expected gain from appeal exceeds the cost of appeal. The joint likelihood of these two occuring is $rt_{10}F(EG_D^{0K}) + (1-r)t_{11}F(EG_D^{1K})$, irrespective of whether the case is a "damages" case or a "liability" case. This is why α does not affect the numerator. The denominator, which is the probability of a defendant loss, is equal to the probability of plaintiff win and is given by expression (7). Similarly, the plaintiffs' appeal rate can be shown to be[18]

[18] In calculating the appeal rate, Eisenberg and Heise (2009) drop the observations in which the appeal is filed by the party that won the trial. To be consistent with their calculation, we only take "liability" cases into account in deriving the expression for plaintiff appeal rate (because in all of the "damages" cases the plaintiffs are recorded as having won the trial). Given this, the joint likelihood of the plaintiff losing and appealing the case is $(1-\alpha)(rt_{00}F(EG_P^0) + (1-r)t_{01}F(EG_P^1)$ and the likelihood of a plaintiff loss is $(1-\alpha)(rt_{00} + (1-r)t_{01})$. The $(1-\alpha)$ parts from these terms cancel out and we get the expression in equation (9).

$$\Pr(\text{P appeal}|\text{P loss}) = \frac{rt_{00}F(EG_P^0) + (1-r)t_{01}F(EG_P^1)}{rt_{00} + (1-r)t_{01}}. \tag{9}$$

Reversal Rates: The reversal rate of the defendants (the fraction of all defendant appeals that are reversed) is given by

$$\Pr(\text{Reversal}|\text{D Appeal}) = \frac{rt_{10}F(EG_D^0 + \delta_0 K)a_{00} + (1-r)t_{11}F(EG_D^1 + \delta_1 K)a_{01}}{rt_{10}F(EG_D^0 + \delta_0 K) + (1-r)t_{11}F(EG_D^1 + \delta_1 K)}. \tag{10}$$

The numerator is the joint probability of a defendant appeal and a reversal and the denominator is the marginal probability of defendant appeal. Similarly, the reversal rate for the plaintiffs is

$$\Pr(\text{Reversal}|\text{P Appeal}) = \frac{rt_{00}F(EG_P^0)a_{10} + (1-r)t_{01}F(EG_P^1)a_{11}}{rt_{00}F(EG_P^0) + (1-r)t_{01}F(EG_P^1)}. \tag{11}$$

Using these expressions we have the following benchmark case:

Proposition 1 (Benchmark Case). *If $\alpha = 0$, the case-mix is symmetric (i.e., $r = 0.5$) and stakes are symmetric then the trial win rates, appeal rates and the reversal rates are the same for both the plaintiffs and the defendants.*

When there is no difference in the proportion of cases in which the defendants and plaintiffs deserve to win, they are equally biased about the correct verdict being in their favor, and that their costs of appeal are identical, there is essentially no difference between the defendants and plaintiffs. Hence, the trial win rates, appeal rates and the reversal rates are the same across the defendants and plaintiffs.

4. PRO-DEFENDANT CASE-MIX AND THE PRESENCE OF "DAMAGES" CASES

In this section investigate how a pro-defendant case-mix (i.e., $r > 1/2$) and presence of damages cases ($\alpha > 0$) affect trial win rates, appeal rates and reversal rates. The results are then used to show that a pro-defendant case-mix and presence of damages cases can result in the same pattern as observed in the data: the higher trial win rate of the plaintiffs and the higher appeal and reversal rates of the defendants. By providing this alternative explanation of the observed pattern in the data, we show that the conclusion drawn by Eisenberg and Heise (2009) does not necessarily follow. Their paper concludes that appellate court judges rule more often in the defendants' favor because they believe that the trial court adjudicators are pro-plaintiff and the appellate court judges attempt to correct for this perceived trial court bias.

Proposition 2 below shows how r, α and K affect the trial win rates, appeal rates and the reversal rates.

Proposition 2. *(a) The partial effects of r, α and K are as follows (where D (resp. P) denotes the rate corresponding to the defendants (resp. plaintiffs)):*

	Trial win rates	Appeal Rates	Reversal Rates
$r \uparrow$	$D \uparrow$	D (see part (b))	$D \uparrow$
	$P \downarrow$	$P \downarrow$	$P \downarrow$
$\alpha \uparrow$	$D \downarrow$	$D \downarrow$	D *(no effect)*
	$P \uparrow$	P *(no effect)*	P *(no effect)*
$K \uparrow$	D *(no effect)*	$D \uparrow$	D (see part (c))
	P *(no effect)*	P *(no effect)*	P *(no effect)*

(b) If $t_{10}F(EG_D^0 + \delta_0 K) > t_{11}F(EG_D^1 + \delta_1 K)$ then an increase in r increases the appeal rate of the defendants. If $t_{10}F(EG_D^0 + \delta_0 K) < t_{11}F(EG_D^1 + \delta_1 K)$ then an increase in r increases the defendant appeal rate if and only if

$$\frac{\alpha}{1-\alpha} < \frac{t_{10}t_{11}[F(EG_D^0 + \delta_0 K) - F(EG_D^1 + \delta_1 K)]}{t_{11}F(EG_D^1 + \delta_1 K) - t_{10}F(EG_D^0 + \delta_0 K)}. \tag{12}$$

(c) As K increases, the defendant reversal rate increases if and only if

$$\delta_0 \frac{f(EG_D^0 + \delta_0 K)}{F(EG_D^0 + \delta_0 K)} > \delta_1 \frac{f(EG_D^1 + \delta_1 K)}{F(EG_D^1 + \delta_1 K)} \tag{13}$$

Before we discuss the implications of the above proposition, note that the effects outlined in the proposition are the partial effects. Hence, when we say that K has no effect on the defendants' and plaintiffs' trial win rates, what we are saying is that it has no effect once we hold r constant. Of course, if a change in K changes r, then that will change the trial win rates as noted in the table above.

Proposition 2 says that an increase in r decreases plaintiffs' trial win rate but an increase in α increases it. This suggests that if α is sufficiently large then the plaintiffs' trial win rate will be larger than that of the defendants, even when defendants' stakes are higher and cause the case-mix at the trial level to be relatively pro-defendant. Hence, unlike as suggested by Eisenberg and Heise (2009), we cannot rule out a pro-defendant case-mix at the trial level merely based on a higher plaintiff trial win rate.

With respect to the appeal rates, Proposition 2 says that, holding other things constant, as r increases the plaintiff appeal rate decreases. Factors α and K have no effect on the plaintiff appeal rate. In the case of defendants, the recorded appeal rate decreases with α, increases with K, and also increases as r increases if α is not very large. This suggests a possibility that as r, α and K increase (relative to the benchmark case), then the defendant appeal rate rises or, if it falls, it does not fall as much as the plaintiff appeal rate, resulting in the pattern in the appeal rates that is seen in the data (see Table 2 in Appendix B).

Finally, let us consider the reversal rates. The proposition says that an increase in r decreases the plaintiff reversal rate and increases the defendant reversal rate. The size of α has no effect on the reversal rates. An increase in K can either decrease or increase their reversal rates. This suggests that, for any given K, if r is sufficiently large then that can result in the defendants' reversal rate being higher than the plaintiffs' reversal rate, even if the selection hypothesis does not operate at the appellate level.

In the foregoing discussion we suggest, separately for trial win rates, appeal rates and the reversal rates, how a pro-defendant case-mix, a presence of "damages" cases and higher stakes of defendants can result in the relative magnitudes of these rates (across defendants and plaintiffs) that are observed in the data. However, the conditions on the parameters (r, α and K) that are favorable for achieving the observed relative magnitudes (across defendants and plaintiffs) in the case of one rate, may seem to contradict the conditions that are favorable for achieving the observed relative magnitudes in the case of another rate.[19] This may lead one to

[19]For example, in the case of the trial win rates, we suggest that for a given $r > 0.5$ if α is sufficiently large then the recorded plaintiff trial win rate can be higher than the recorded defendant win rate.

wonder whether there are values of r, α and K such that the observed patterns in all the three rates are simultaneously achieved. The answer is in the affirmative. We first illustrate this with a numerical example. In Table 1 the trial win, appeal and reversal rates for one specific combination of r, α and K values are presented. Following that, in Figure 1 we present the entire range of r and α values that result in a higher plaintiff win rate, a higher defendant appeal rate and a higher defendant reversal rate, when the values of other parameters are the same as those used in the example in Table 1.

Numerical Example: Of the cases adjudicated by the trial court, let $\alpha = 0.15$. Also, suppose that the higher stakes of the defendants results in a better selection of cases by them at the trial level, resulting in $r = 0.55$. The values of other relevant parameters are the following: $t_{00} = 0.9, a_{00} = 0.95, q_D^0 = 0.8, q_D^1 = 0.65, H_1 = 120, H_0 = 10, K = 20$. Also, suppose that the distribution of the litigant's cost of appeal is given by the two-sided power distribution, $TSP(0, 40, 90, 3)$, where zero is the minimum value of the distribution, 40 is the mode, 90 is the maximum value and 3 is the shape parameter.[20] When we plug these numbers in our model, we get the following:

As the figures in Table 1 show, the observed higher plaintiff win rate at the trial level and higher defendant appeal and reversal rates at the appellate stage are consistent with higher defendant stakes once we take into account that there are likely to be cases in which the issue is that of damages and not liability.

Keeping the values of all the parameters, except α and r, the same as those in the above example, it can be shown, that for any (α, r) combination in the hatched area in Figure 1, the plaintiff win rate is higher than the defendant win rate, and that defendant appeal and reversal rates exceed those of the plaintiff. The Equal T line represents combinations for which the defendants and the plaintiffs have the same trial win rate. The Equal A line represents the combinations for which the two have

For the appeal rates, we suggest that if α is not very large then a pro-defendant case-mix ($r > 0.5$), along with higher stakes of the defendants, can result in a higher appeal rate of the defendants.

[20] A random variable X with probability density function $f(\cdot)$ is said to follow a TSP(a, m, b, n) distribution if

$$f(x|a,m,b,n) = \begin{cases} \frac{n}{b-a}\left(\frac{x-a}{m-a}\right)^{n-1} & a < x \leq m, \\ \frac{n}{b-a}\left(\frac{b-x}{b-m}\right)^{n-1} & m \leq x < b. \end{cases}$$

where, a (resp., b) is the minimum (resp. maximum) value that X can take and for any $n > 1$, m is the mode of the distribution. See van Dorp and Kotz (2002) for further information on the TSP distribution.

Table 1: Trial Win, Appeal and Reversal Rates (Numerical Example)

Pr(Defendant win at trial level)	43%
Pr(Plaintiff win at trial level)	57%
Pr(D appeal$\|$D loss)	14%
Pr(P appeal$\|$P loss)	10%
Pr(Reversal$\|$D appeal)	37%
Pr(Reversal$\|$P appeal)	30%

the same appeal rate. Finally, the Equal R line represents combinations for which both have the same reversal rate.

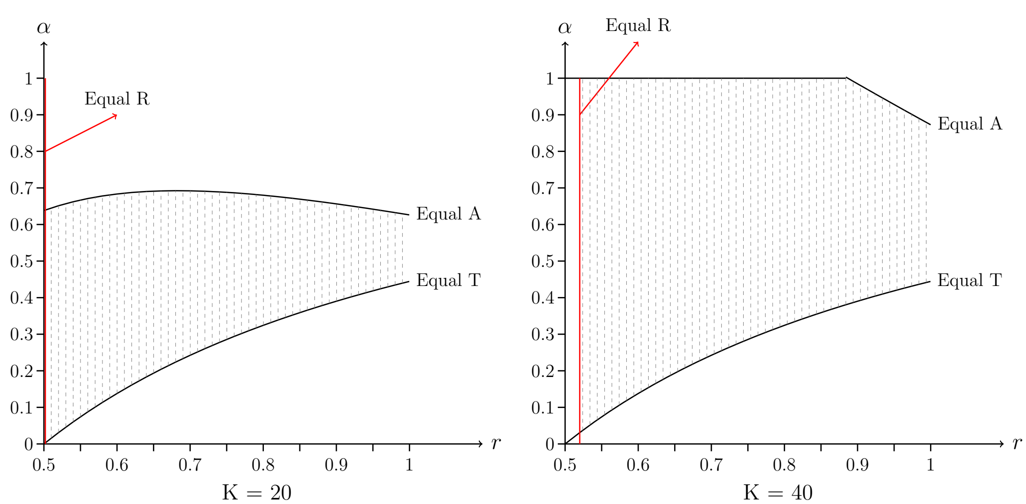

Figure 1 The dashed area indicates the combinations of (r, α) values that yield the same pattern as observed in the data: higher plaintiff trial win rate; higher defendant appeal and reversal rates.

5. CONCLUSION

This paper uses a formal model of the adjudication process to show that the pattern in the trial win, appeal and reversal rates of plaintiffs and defendants, that is observed in the Supplemental Survey of Civil Appeals data, can be explained by higher defendant stakes, or any other factor that causes a pro-defendant case-mix at the trial level. This is true even if, (i) defendants are *not* more selective about which cases they push on appeal (and, hence, agree with Eisenberg and Heise (2009) that the selection hypothesis does not apply at the appellate stage); (ii) neither the trial court nor the appellate court is biased; (iii) both the defendants and the plaintiffs have the same cost of appeal; and finally, (iv) both the defendants and plaintiffs are equally biased in their beliefs about the correct outcome being in their favor. By showing that the result holds even in the absence of any bias on the part of either trial or appellate court adjudicators, we show that the conclusion of appellate court "plaintiphobia" drawn by Eisenberg and Heise (2009) does not necessarily follow from the data.

In deriving this result, we use the fact that the variable in the data that measures plaintiff wins is constructed such that any trial in which the defendant is deemed liable and has to pay damages, is recorded as a plaintiff win. This, as we explain in the paper, can skew the recorded trial win rate in favor of the plaintiffs, even if the case-mix at the trial level is pro-defendant. This is a crucial point that needs to be taken into account when using these data and that, to the best our knowledge, no other studies have acknowledged.

APPENDIX A: PROOFS

Proof of Proposition 1

Substituting $r = 0.5$ and $\alpha = 0$ in expressions (6) and (7), we have $\Pr(\text{D win at trial}) = 0.5$ and $\Pr(\text{P win at trial}) = 0.5$.

Similarly, substituting $r = 0.5$, $\alpha = 0$ and $K = 0$, in (8), we have $\Pr(\text{D appeal}|\text{D loss}) = t_{10}F(EG_D^0) + t_{11}F(EG_D^1)$. Now from equations (1) and (4),

$$EG_D^0 = \left(\frac{q_D^0 t_{10}}{q_D^0 t_{10} + (1 - q_D^0) t_{11}} a_{00} + \frac{(1 - q_D^0) t_{11}}{q_D^0 t_{10} + (1 - q_D^0) t_{11}} a_{01} \right) \Delta H. \quad (14)$$

Using assumptions (A.2), (A.4) and (A.6), EG_D^0 can be written as

$$EG_D^0 = \left(\frac{(1-q_P^1)t_{01}}{(1-q_P^1)t_{01} + q_P^1 t_{00}} a_{11} + \frac{q_P^1 t_{00}}{(1-q_P^1)t_{01} + q_P^1 t_{00}} a_{10} \right) \Delta H. \quad (15)$$

This expression along with (2) and (3) suggests that $EG_D^0 = EG_P^1$. Similarly, it can be shown that $EG_D^1 = EG_P^0$ when $r = 0.5, \alpha = 0$ and $K = 0$. Using these equalities and assumption (A.2), we can write $\Pr(\text{D appeal}|\text{D loss}) = t_{10}F(EG_D^0) + t_{11}F(EG_D^1) = t_{01}F(EG_P^1) + t_{00}F(EG_P^0)$. From (9), this last term equals $\Pr(\text{P appeal}|\text{P loss})$.

Using $r = 0.5$, $\alpha = 0$, $K = 0$, $EG_D^0 = EG_P^1$ and $EG_D^1 = EG_P^0$, we can write the defendants reversal rate as,

$$\frac{t_{10}F(EG_D^0)a_{00} + t_{11}F(EG_D^1)a_{01}}{t_{10}F(EG_D^0) + t_{11}F(EG_D^1)} = \frac{t_{10}F(EG_P^1)a_{00} + t_{11}F(EG_P^0)a_{01}}{t_{10}F(EG_P^1) + t_{11}F(EG_P^0)} \quad (16)$$

Using the equalities in assumptions (A.2) and (A.4), one can see that the right hand side expression in (16) is the same as the plaintiff's reversal rate given in expression (11). ■

Proof of Proposition 2

Trial Win Rates

Differentiating the defendants' trial win rate (expression (6)), with respect to r we have $d[\Pr(\text{D win at trial})]/dr = (1-\alpha)(t_{00} - t_{01}) > 0$. Similarly, differentiating the plaintiffs' trial win rate, (7), with respect to r, we have $d[\Pr(\text{P win at trial})]/dr = (1-\alpha)(t_{10} - t_{11}) < 0$.

Differentiating expression (6) with respect to α we have $d[\Pr(\text{D win at trial})]/d\alpha = -(rt_{00} + (1-r)t_{01}) < 0$; and differentiating (7) with respect to α, we have $d[\Pr(\text{P win at trial})]/d\alpha = 1 - (rt_{10} + (1-r)t_{11}) > 0$.

Since K does not enter either of the trial rate expressions, it has no direct effect on the trial rates.

Appeal Rates

Differentiating the defendants' appeal rate (expression (8)) with respect to r, we have

$$\begin{aligned} d[\Pr(\text{D appeal}|\text{D loss})]/dr \propto \\ &\alpha[t_{10}F(EG_D^0 + \delta_0 K) - t_{11}F(EG_D^1 + \delta_1 K)] \\ &+ (1-\alpha)(t_{10}t_{11})[F(EG_D^0 + \delta_0 K) - F(EG_D^1 + \delta_1 K)]. \quad (17) \end{aligned}$$

Since $EG_D^0 + \delta_0 K > EG_D^1 + \delta_1 K$, the second term is clearly positive. But since $t_{10} < t_{11}$, the sign of the first term is not obvious. If $t_{10}F(EG_D^0 + \delta_0 K) > t_{11}F(EG_D^1 + \delta_1 K)$, then the first term is also positive and the derivative is positive. If, however, $t_{10}F(EG_D^0 + \delta_0 K) < t_{11}F(EG_D^1 + \delta_1 K)$, then the derivative is positive only if α is sufficiently small. That is, with $t_{10}F(EG_D^0 + \delta_0 K) < t_{11}F(EG_D^1 + \delta_1 K)$, $\alpha[t_{10}F(EG_D^0 + \delta_0 K) - t_{11}F(EG_D^1 + \delta_1 K)] + (1-\alpha)(t_{10}t_{11})[F(EG_D^0 + \delta_0 K) - F(EG_D^1 + \delta_1 K)] > 0$ if and only if

$$\frac{\alpha}{1-\alpha} < \frac{t_{10}t_{11}[F(EG_D^0 + \delta_0 K) - F(EG_D^1 + \delta_1 K)]}{t_{11}F(EG_D^1 + \delta_1 K) - t_{10}F(EG_D^0 + \delta_0 K)}.$$

The effect of α on appeal rates is straightforward. From (8), the fact that α appears only in the denominator and that $rt_{10} + (1-r)t_{11} < 1$, it is clear that an increase in α decreases the defendant appeal rate. And, from (9) it is clear that α does not affect the plaintiff appeal rate. Similarly, it is easy to see that an increase in K increases the defendant appeal rate, and does not affect plaintiff appeal rate.

Differentiating expression (9) with respect to r, we have $d[\Pr(\text{P appeal}|\text{P loss})]/dr \propto t_{00}t_{01}[F(EG_P^0) - F(EG_P^1)] < 0$.

Reversal Rates

Differentiating (10) with respect to r, we have $d[\Pr(\text{Reversal}|\text{D loss})]/dr \propto (a_{00} - a_{01})F(EG_D^1 + \delta_1 K)F(EG_D^0 + \delta_0 K)t_{11}t_{10} > 0$. Similarly, differentiating (11) with respect to r, we have $d[\Pr(\text{Reversal}|\text{P loss})]/dr \propto (a_{10} - a_{11})F(EG_P^1)F(EG_P^0)t_{01}t_{00} < 0$. Hence, a pro-defendant case-mix increases defendant reversal rates and decreases plaintiff reversal rate.

Since α does not enter either of the reversal rates, it is clear that it does not affect the reversal rates.

Differentiating (10) with respect to K, we have

$$\begin{aligned} &d[\Pr(\text{Reversal}|\text{D loss})]/dr \propto \\ &\qquad r(1-r)t_{10}t_{11}(a_{00} - a_{01})\Big[\delta_0 F(EG_D^1 + \delta_1 K)f(EG_D^0 + \delta_0 K) \\ &\qquad\qquad - \delta_1 F(EG_D^0 + \delta_0 K)f(EG_D^1 + \delta_1 K)\Big]. \end{aligned} \tag{18}$$

This expression is positive if and only if

$$\delta_0 \frac{f(EG_D^0 + \delta_0 K)}{F(EG_D^0 + \delta_0 K)} > \delta_1 \frac{f(EG_D^1 + \delta_1 K)}{F(EG_D^1 + \delta_1 K)}.$$

■

APPENDIX B: DATA

The patterns in the data that we refer to in the main body are based on the Supplemental Survey of Civil Appeals, 2001. In Table 2 we show the composition of civil cases at the trial level. Further, in Table 3 we show the summary statistics pertaining to the trial win rates, appeal rates and reversal rates of the defendants and the plaintiffs from this data set. These statistics are the same as those used by Eisenberg and Heise (2009) to motivate their hypothesis that appellate courts exhibit "plaintiphobia".

Table 2: Types of Civil Cases at the Trial Level

Case Type	Frequency
Tort cases	5451
Automobile tort	2819
Premises Liability	924
Product Liability	117
Intentional tort	251
Medical malpractice	850
Professional malpractice	68
Slander/libel	61
Other tort	361
Contract	2411
Fraud	428
Seller Plaintiff	772
Buyer Plaintiff	531
Mortgage Foreclosure	13
Employment Discrimination	106
Employment Dispute: Other	175
Rental/Lease agreement	188
Other contract	198
Property	176
Eminent domain	33
Other Property	143
All civil cases	8038

Of the initial 8,038, we eliminated 165 civil trial cases in which both litigants appealed. We eliminated 10 cases for which the trial outcome was missing. This gives a total of 7,863 cases. The trial win rate for the defendants is defined as (Total # of Trials Won by the Defendant)/(Total # of Trials). The trial win rate of the plaintiffs is similarly defined. Of the 7,863 trials, 991 (12.6%) were appealed. As Table 3 shows, the appeal rate for the defendants is higher than that of the plaintiffs. The appeal rate for defendants is defined as

$$\text{Defendant Appeal Rate} = \frac{\text{Total \# of Defendant Appeals}}{\text{Total \# of Cases in which Defendants Lose}} \tag{19}$$

The appeal rate for the plaintiffs is similarly defined. The defendants brought 573 (57.8%) of the 991 cases appealed.

We categorize a case as being reversed if it was "reversed in whole" or "reversed in part" or "remanded" or "affirmed in part". The reversal rate for the defendants is defined as

$$\text{Defendant Reversal Rate} = \frac{\text{Total \# of Defendant Appeals Reversed}}{\text{Total \# of Defendant Appeals}} \tag{20}$$

The reversal rate for the plaintiffs is similarly defined.

Table 3: Trial Win, Appeal and Reversal Rates of Defendants and Plaintiffs

Party	Trial Win Rate	Appeal Rate	Reversal Rate
Defendants	46.3	13.6	34.3
Plaintiffs	53.7	11.5	23.7
p-value[a]	0.000	0.000	0.000

[a] p-value is the probability value for the t-test of the difference between the defendant and the plaintiff rates.

REFERENCES

Bhole, B. and B. Gleeson Hanna (2009). An analytical framework for interpreting appellate court data. *Economics Bulletin 29*(2), 1163–1174.

Daughety, A. and J. Reinganum (2000a). Appealing judgments. *Rand Journal of Economics 31*, 502–525.

Daughety, A. and J. Reinganum (2000b). On the economics of trials: Adversarial process, evidence, and equilibrium bias. *Journal of Law, Economics and Organization 16*, 365–394.

Eisenberg, T. and M. Heise (2009). Plaintiphobia in state courts? an empirical study of state court trials on appeal. *The Journal of Legal Studies 38*.

Lempert, R. (1998). Why do juries get a bum rap? reflections on the work of valerie hans. *DePaul Law Review 48*, 453–62.

Priest, G. and B. Klein (1984). The selection of disputes for litigation. *The Journal of Legal Studies 13*, 1–55.

Saks, M. J. (1998). Public opinion about the civil jury: Can reality be found in illusions? *DePaul L. Review 48*, 221–43.

Shavell, S. (1995). The appeals process as a means of error correction. *Journal of Legal Studies XXIV*, 379–426.

van Dorp, J. R. and S. Kotz (2002). A novel extension of the triangular distribution and its parameter estimation. *The Statistician 51*(1), 63–79.

Vidmar, N. (1998). The performance of the american civil jury: An empirical perspective. *Arizona Law Review 40*, 849–899.

Waldfogel, J. (1995). The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory. *The Journal of Political Economy 103*, 229–260.