The Adverse Selection of Cases for Trial

LUKE FROEB

Economic Analysis Group, United States Department of Justice, Washington, DC 20001, USA

I. Introduction

In their review of the economics of legal disputes, Cooter and Rubinfeld (1989) state that empirical research lags "woefully" behind theoretical advances. Part of the problem is the fragility of bargaining models to changes in the information structure or in the strategic play of the game. It is difficult to draw inferences about bargaining games when the predictions of the models are sensitive to changes in mostly unobservable factors. The challenge is to develop reduced forms, consistent with different classes of models, that permit inference about the game being played.

This paper develops a simple reduced form that allows inference about the existence of private information in a plea bargaining game. The reduced form is consistent with a wide class of extensive games and uses incentive compatibility constraints (e.g., Myerson, 1979) to derive testable implications. Incentive compatibility implies that defendants who possess private information about their chances at trial will not voluntarily disclose the information without an incentive to do so. This puts restrictions on the form that plea bargains can take. In particular, it implies that individuals with relatively good chances at trial are more likely to go to trial. This outcome is analogous to adverse selection in an insurance market. Just as relatively "bad" insurance risks are more likely to purchase insurance, so are defendants with relatively bad chances at trial more likely to reach plea settlements.

This kind of adverse selection in plea bargaining is shown to imply a positive correlation between trial rates and conviction rates. Evidence of such a positive correlation is found using aggregate federal criminal data.

II. Adverse Selection in Plea Bargaining

If defendants possess private information about their chances at trial, plea bargaining can facilitate the process of gathering information about defendants because a defendant's guilt or innocence can be revealed through his choice to accept or reject
an offered plea. Plea bargaining thus contributes to the accuracy of the legal process by inducing defendants with a high probability of conviction to identify themselves as guilty. In this scenario, plea discounts off of expected trial penalties can be justified as payments to relatively guilty defendants to reveal that they are indeed guilty.

Of course, this view of plea bargaining makes no sense unless defendants possess private information about their chances at trial. If prosecutors, rather than or in addition to defendants, possess private information about a defendant’s chances at trial, then the prosecutor can make plea offers where “. . . the greatest pressures to plead guilty are brought to bear on defendants who may be innocent. The universal rule is that the sentence differential between guilty-plea and trial defendants increases in direct proportion to the likelihood of acquittal.” In this scenario, plea discounts induce defendants with relatively good chances at trial to plead guilty. The desirability of this type of plea bargaining is questionable because it can lead to false convictions. Thus, knowing whether plea bargaining exhibits adverse selection is important for determining how plea bargaining affects the legal process.

Adverse selection in plea bargaining also has implications for forecasting changes in case disposition based on changes in legal policy. For example, during prohibition, when the United States legal system was flooded with defendants charged with liquor crimes, plea agreement became the dominant method of conviction in both the state and federal courts. At the same time that trial rates fell, conviction rates also fell for most federal crimes (Padgett, 1987). If we had known that plea bargaining was subject to adverse selection, then such a change would have been predictable.

Finally, a finding that plea bargaining exhibits adverse selection can be used as a guide to theoretical model building. Plea bargaining games with equilibria inconsistent with adverse selection can be rejected as inconsistent with the data.

III. Incentive Compatibility Constraints and Private Information

In this section, I show how private information can affect the selection of cases for trial by deriving incentive compatibility constraints for a plea bargaining game in which defendants possess private information about their chances of prevailing at trial. The constraints characterize the outcome of any extensive game with the postulated information structure and payoff functions, which are taken as the primitives of our model. The extensive structure can be quite general, like a take-it-or-leave-it or alternating offers game.

The information of the defendant is \( \pi \), which represents the defendant’s chances of being convicted at trial. The defendant, but not the prosecutor, knows the value of \( \pi \). If convicted, the defendant faces a sentence of length \( S \), and the defendant has a utility function over sentence length, \( U(\cdot) \). The utility of a zero sentence length, \( U(0) \), is normalized to zero so that the expected utility of the trial lottery is \( \pi U(S) \).

An outcome of the plea bargaining game is a pair of functions, \( (p(\pi), s(\pi)) \), that map the information of the defendant into the probability of going to trial, \( p(\pi) \), and the settlement payoff, \( s(\pi) \). The term \( s(\pi) \) is the sentence length that the defendant accepts to avoid trial.

Incentive compatibility constraints are derived from a revelation game where defendants announce their private information, \( \pi \), and then receive the payoff associated with the outcome \( (p(\pi), s(\pi)) \). In a Nash equilibrium, it is not possible for a defendant to misrepresent his private information to gain a more favorable settlement. In particular, it is not possible for a defendant with information \( \pi \) to pretend to have information \( \pi^* \) to gain a better settlement. By pretending to have information \( \pi^* \), the defendant gains the outcome \( (p(\pi^*), s(\pi^*)) \). In equilibrium, the expected utility of this outcome must be less than the expected utility of the truth telling outcome, \( (p(\pi), s(\pi)) \):

\[
U_d(\pi) \geq (1-p(\pi))U(s(\pi)) + p(\pi)U(S) \quad \text{for all } \pi, \pi^*.
\]

(1)

\( U_d(\pi) \) is the expected utility function of the defendant who truthfully reports his private information, \( \pi \). The right-hand side of the inequality is the expected utility of a defendant with information \( \pi^* \) who pretends to have information \( \pi \).

It is possible to rewrite the incentive compatibility constraint as follows:

\[
U_d(\pi) \geq U_d(\pi^*) + U(S)p(\pi)(\pi - \pi^*).
\]

(2)

Add the constraint for a defendant of type \( \pi \), to the constraint for a defendant of type \( \pi^* \) to derive inequality (3):

\[
\Rightarrow p(\pi) \geq p(\pi^*) \quad \text{for all } \pi_1 < \pi_2.
\]

(3)

In this case, incentive compatibility implies that the trial probability function is non-increasing in the defendant’s type, that is, the higher a defendant’s probability of conviction at trial, the lower the probability that he will go to trial.

If the inequality in Equation (3) is strict, at least for some defendant types, then we say that there is adverse selection. Defendants with relatively good chances of prevailing at trial are more likely to go to trial. Those with higher probabilities of conviction are more likely to accept plea bargains. If \( p(\pi_1) = p(\pi_2) \) for all \( \pi \) and \( \pi^* \), then defendants behave identically (pool) and there is no adverse selection.

So far I have limited attention to a class of games characterized by private information possessed by only the defendant. In general, it is possible that the prosecutor, instead of or in addition to the defendant, possesses private information. If the prosecutor has private information about a defendant’s chances at trial, then it is possible to proceed exactly as above to derive the incentive compatibility constraint of the prosecutor. If he or she is maximizing an expected utility function of sentence length, as in Landes (1971), then the incentive compatibility restriction is the opposite of that implied by the game in which the defendant has private information. In this context, incentive compatibility implies that defendants with relatively small chances of prevailing at trial are more likely to go to trial. In this case we say that plea bargaining exhibits reverse adverse selection.

If both the prosecutor and the defendant have private information, then it is not possible to say much about the outcome of the game. In this case, the outcome is a function of two variables, the information of the prosecutor and that of the defendant, and there are two sets of incentive compatibility constraints, one for the prosecutor and one for the defendant. The incentive compatibility constraints are similar to the constraints with one-sided uncertainty, but are expressed in terms of expected value so that they depend on the joint distribution of the information sets of both.

---

3Models of plea bargaining games that result in this type of separating equilibrium have been studied by Grossman and Katz (1980), Kobayashi and Lott (1989), and Reinganum (1988).
the prosecutor and defendant. To derive refutable implications from games of two-sided information, the extensive form of the game must be considered. 3

IV. Empirical Implications of Incentive Compatibility Constraints

Thus far, I have discussed an equilibrium relationship between the probability of trial and the probability of conviction for a class of plea bargaining games under three different scenarios: In the case of private information possessed by only the defendant, defendants with relatively good chances at trial go to trial, and in the case of private information possessed by only the prosecutor, defendants with relatively bad chances at trial go to trial. For the case of private information possessed by both parties, there is no simple relationship. Thus, neglecting the case of two-sided private information, determining whether defendants or prosecutors possess private information would be a simple matter of testing whether conviction rates of the trial group were lower than conviction rates of the plea group. This strategy is not feasible, however, because conviction rates for the plea group are not observed.

An alternate testing strategy can be illustrated by reference to a specific extensive game in which a group of defendants with private information receives the same take-it-or-leave-it plea offer from a prosecutor. If the defendants are identical, except for their private information, the offer induces adverse selection among the defendants: Those defendants with relatively good chances at trial reject the offer, while those with relatively bad chances at trial accept it. 4

Now consider what happens when the prosecutor offers a longer sentence in exchange for a guilty plea. Some defendants with a relatively high probability of conviction who had previously accepted the offer now reject it and opt for trial. This increases the mean probability of conviction for the trial group, inducing a positive correlation between trial rates and conviction rates. Thus, if there is adverse selection in plea bargaining, trial rates are positively related to conviction rates at trial. Conversely, if there is reverse adverse selection, then trial rates are negatively related to conviction rates. Pooling would imply that there is no correlation.

Although this hypothesis-testing strategy is illustrated with a take-it-or-leave-it model, the model is not necessary for the strategy to work. As long as changes in trial rates are exogenous and large enough, the incentive compatibility constraints imply that some defendants, with a relatively high probability of conviction, will be moved out of the settlement group and into the trial group. This change induces a positive correlation between trial rates and conviction rates. I have found two good natural experiments, one in Alaska and one in El Paso, where just such changes occurred, and in both cases trial rates increased at the same time that the conviction rates increased. 5 These are good natural experiments because the changes in the trial rates were probably exogenous, brought on by changes in the local political climate.

These historical experiments, though certainly consistent with the existence of adverse selection in plea bargaining, are difficult to generalize. For most types of non-

---

3Drawing intuition from an exchange bargaining model of two-sided uncertainties (Osborne and Rubinstein, 1990), all that can be said about the outcome is that when the gains from trade are larger (when the costs of a trial are larger) then trade (settlements) is more likely.

4See Belochuk (1984) or Froeb (1989) for examples of this game.


---

experimental data, testing for adverse selection is not so easy because changes in the conviction rate cannot be considered exogenous. In particular, there may be unobserved factors, inducing spurious covariation between trial and conviction rates, so that identification of adverse selection in plea bargaining would require an instrument that shifts the probability of trial.

Appropriate instruments are suggested by theories of prosecutorial behavior under a budget constraint, like that of Landes (1971). In his model, the prosecutor maximizes expected sentences subject to a resource constraint. In equilibrium, the constraint is binding so that loosening the constraint leads to more trials. Unfortunately, I was unable to find a policy instrument that shifted the trial rate. Consequently, the regression results presented below should be interpreted with care because the possibility of a spurious correlation cannot be ruled out. On the other hand, it is clear that prosecutorial budgets change slowly, perhaps with a long enough lag so that trial rates can be considered exogenous.

V. Data and Results


The basic equation to be estimated is a regression of conviction rates on trial rates. In a second specification, I use dismissal rates as a control for the pool of criminal types that are aggregated. The conviction rates are aggregates of individual convictions so that the variance of each aggregate is dependent on the number of defendants in the aggregate and upon the conviction rate itself. The chosen weights are proportional to $\pi(1 - \pi)/n$, where $\pi$ is the estimated probability of conviction and $n$ is the number of defendants in each aggregate. 6 Results for the district data are presented in Table 1. Results for the crime data are presented in Table 2, and results from time series regressions for five different crimes are presented in Table 3.

In Table 1, results from the U.S. Attorney data are presented. Ten districts were dropped that had either a 100% or 0% conviction rate, which implies a zero variance and an infinite weight. These districts were relatively small (they all had less than 10 trials) so dropping them probably does not affect the results. Based on their unusual set of local crimes or laws, I also excluded the District of Columbia, Puerto Rico, the Virgin Islands, Guam, and Hawaii. The final number of districts in the data set is 77. I estimated the pooled regressions by pooling data for 1984 and 1985 after a Chow test for stability of slope coefficients across the 2 years proved insignificant. The

6Different United States Attorneys prosecute various kinds of cases at different rates depending on their views concerning current needs, the relationship of federal to state resources, their perceptions of possible problems in the state law enforcement mechanism, their interpretations of Justice Department policy, pressures from enforcement agencies, and, of course, differences in the rates at which various federal crimes are actually committed in the district. (Flanders, 1976, p. 58)
Table 1. Pooled and within-district weighted regressions: Conviction rates on dismissal and trial rates

<table>
<thead>
<tr>
<th></th>
<th>Dismissals</th>
<th>Trials</th>
<th>R-squared</th>
<th>Deg freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pooled district</td>
<td>-0.248</td>
<td>0.159**</td>
<td>9.1%*</td>
<td>145</td>
</tr>
<tr>
<td>Model 1</td>
<td>(2.87)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>(2.61)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within district</td>
<td>.030</td>
<td>0.214</td>
<td>0.1%</td>
<td>71</td>
</tr>
<tr>
<td>Model 1</td>
<td>(0.11)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>(1.08)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1Data are defendants terminated by the U.S. Attorney's office in calendar years 1984 and 1985, aggregated by district. Seventy-seven districts were used. Regression coefficients are above 1-statistics in parentheses. *Significance at the 5% level; **significance at 1% level. Model 1 measures trials as a percentage of trials and pleas, and dismissals as a percentage of trials, pleas, and dismissals. Model 2 measures trials as a percentage of trials, pleas, and dismissals. Pooled-district regressions estimated by a seemingly unrelated weighted regression with cross-equation restrictions: within-district estimation by weighted regression on the difference data.

pooled regressions show that trial rates are significantly and positively related to conviction rates. The within-district estimation results are also presented in Table 1. The within-district estimator is constructed by differencing the data for the 2 years in each district. The within-district results in Table 1 show a positive relation between trial rates and conviction rates, but the coefficients lack statistical significance. There is probably not enough within-district variation to estimate the coefficients precisely.

In Table 2, results from the crime data are presented. I estimated the pooled district crime regressions by pooling data for 1983, 1984, and 1985 after a Chow test for stability of slope coefficients across the 3 years proved insignificant. The pooled crime regressions show that trial rates are significantly and positively related to conviction rates. The within-crime estimator is constructed by adding crime dummies to the regression. The within-crime results show a marginally significant positive relation between trial rates and conviction rates. The similarity of the slope coefficients from the within- and pooled-crime regressions do not suggest the existence of crime-specific confounding effects. This suggests that shifts in trial rates across crimes can be considered exogenous (Hausman, 1978), at least with respect to crime-specific effects.

The time-series regressions by crime are presented in Table 3. The relation between trial rates and conviction rates is insignificant for all but the crime of income tax fraud. For tax fraud, trial rates are positively and significantly related to conviction rates. Except for tax fraud, the time-series results seem to imply pooling. I can only guess why these results appear different from the cross-sectional results in Tables 1 and 2. One referee has suggested that the increasing use of pleas over the sample period could induce a spurious negative correlation between trial rates and conviction rates that would mask an underlying positive correlation.

To summarize the results, the cross-district and cross-crime regressions provide evidence of a positive correlation between conviction and trial rates. Except for the crime of tax fraud, the time-series regressions suggest no correlation.

Table 2. Pooled and within-crime weighted regressions: Conviction rates on dismissal and trial rates

<table>
<thead>
<tr>
<th></th>
<th>Dismissals</th>
<th>Trials</th>
<th>R-squared</th>
<th>Deg freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pooled crime</td>
<td>-0.81**</td>
<td>0.564**</td>
<td>33.2%**</td>
<td>45</td>
</tr>
<tr>
<td>Model 1</td>
<td>(-3.47)</td>
<td>(4.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>(3.31)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within crime</td>
<td>-0.059</td>
<td>0.384*</td>
<td>11.8%</td>
<td>30</td>
</tr>
<tr>
<td>Model 1</td>
<td>(-0.15)</td>
<td>(1.96)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>(1.95)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1Data are defendants terminated in the calendar years 1983, 1984, and 1985 in the U.S. district courts, aggregated by crime. Sixteen different crimes are used. Regression coefficients are above t-statistics in parentheses. *Significance at the 5% level; **significance at the 1% level. Model 1 measures trials as a percentage of trials and pleas, and dismissals as a percentage of trials, pleas, and dismissals. Model 2 measures trials as a percentage of trials, pleas, and dismissals.

VI. Conclusions

The positive correlation between trial rates and conviction rates can be interpreted as preliminary evidence of private information possessed by federal criminal
defendants. The practice of plea bargaining seems to be contributing to the accuracy of the legal process by inducing relatively guilty defendants to identify themselves as guilty through their acceptance of offered pleas. The evidence is preliminary because the reduced form relationship between trial rates and conviction rates, though consistent with other specific historical experiments, could be spurious.

The finding of adverse selection in plea bargaining can be used to forecast changes in case dispositions based on changes in legal policy or as a guide to theoretical model building. Bargaining games with equilibria where defendants sort themselves according to the information that they possess (e.g., Grossman and Katz, 1983) may be considered more accurate characterizations of the practice of plea bargaining than games with other types of equilibria (e.g., Reinganum, 1988).

References

Rubinstein, Michael L., Stevens Clarke, and Teresa White, Alaska Land Plea Bargaining, United States Department of Justice, National Institutes of Justice (July, 1980).

Limited Liability and Incentives when Firms Can Inflict Damages Greater than Net Worth

Lisa Lipowski Posev

Pennsylvania State University, University Park, PA 16802 USA

It is well known that, under certain assumptions, strict liability causes a firm to internalize the social cost of accidents, so it chooses the socially optimal activity and preventive care levels. A familiar modification of this result is made in models that allow bankruptcy. When potential damages to third parties exceed the firm's net worth, part of the risk is externalized. Consequently, the firm externalizes part of the benefit of restraining its activity level and increasing the level of preventive care. This argument, which has been emphasized in the literature, implies that limited liability causes the firm's activity level to be too high and the level of preventive care to be too low.

This note emphasizes another consideration that points in the opposite direction. Expenditures on preventive care reduce the firm's wealth, thus reducing the size of a court judgement that it can pay. The reduction in the size of the judgement that it can pay increases the cost of bankruptcy borne by potential creditors, including tort victims. Thus limited liability shifts part of the expected cost of preventive care to the victims of tort.

A similar argument applies to the firm's activity level. A lower activity level often implies a reduction in the firm's stock of capital. A reduction in the firm's stock of capital reduces the firm's wealth, which increases the cost of bankruptcy borne by potential creditors. Thus limited liability shifts part of the expected cost of lowering activity levels to the victims of tort.

The general conclusion can be summarized as follows. Limited liability has two effects. First, it externalizes the firm's expected cost of accidents, which causes a higher activity level and a lower level of preventive care. Second, it externalizes part of the firm's expected cost of care and undercapitalization, which causes a higher level of care and a lower level of activity. In general, one effect is not necessarily larger than the other. This note highlights this argument with an example and then proves it mathematically.

An Example: Generators and Disposers of Toxic Waste

Firms that generate or dispose of toxic waste often have the potential to inflict damages that are greater than their ability to compensate victims. Environmental law has