EFFECTS OF SELECTION PROCESSES AND PRIVATE INFORMATION ON CRIMINAL ANTITRUST CASE OUTCOMES

By

David Buford Sapper

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

May, 2006

Nashville, Tennessee

Approved:

John J. Siegfried

Mark A. Cohen

William W. Damon

Robert A. Driskill

George H. Sweeney

To my graciously supportive wife, Lisa,

and

To my beloved son, Sergei,

ACKNOWLEDGEMENTS

I am grateful to John Siegfried. As a mentor, co-author, supervisor, and committee chairman, he taught me – usually by example – how to thoughtfully and responsibly use economics. I thank Mark Cohen, whose research largely motivated this study, for providing most of the data this study employs. I thank J.S. Butler for fortifying my resolve and econometrics training.

With gratitude, I recognize my favorite bosses' support. Chris Klein's suggestions, as well as Lynn Greer's and Robert Garvin's constant encouragement were invaluable. Significantly, they allowed me to apply and expand my training through policy-making and independent research, including this study.

For my wonderful family members, who, with love, never quit asking about my progress, I am very thankful. I am especially indebted to my parents, Neil and Jo Ann, as well as my sister, Erica, for loving reassurance that continue to boost my pursuits. I thank Lisa and Sergei, my incredible wife and son, for providing more than enough love, inspiration, and joy to motivate me to propose and, more importantly, to complete this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Previous studies of the determinants of federal criminal antitrust case outcomes generally focus on a few observable and measurable factors that could influence one or a few of the observed outcomes at various stages of case disposition. For example, the studies examine whether the industry represented by the defendant or the corporate rank of the defendant influence the government's case-bringing activity or the outcomes of plea bargaining, trial, or sentencing. These studies acknowledge in various degrees that the case selection and plea bargaining processes could influence observed case outcomes. None of these studies recognizes that (or examines how) case outcomes could be affected by the existence of private information about the defendant's chances at trial that one of the litigants knows during plea bargaining.

This study contributes to the economics literature a comprehensive theoretical framework that: specifies the mechanics and predicted effects of the process by which the prosecutor decides to initiate a case; the process by which the defendant chooses a plea; the process by which the judge (or jury) decides to convict a defendant at trial; and the existence of private information about the defendant's chances at trial that one of the litigants knows during plea bargaining. This study also supplies new empirical tests of these theoretical predictions.

This study examines the implications of the decision process by which a federal prosecutor initiates criminal antitrust cases. The prosecutor's case-bringing decision determines the mix of case characteristics (e.g., the litigants' respective trial costs, the amount and strength of the prosecutor's evidence) observed in cases that the prosecutor initiates. Controlling for the prosecutor's case selection process is important because it is not only the value of a particular case characteristic, but also the values of other case characteristics determined by the prosecutor's case selection process, that influence expected trial rates, trial conviction rates, and sentencing decisions.

This study tests – for the first time – how the case selection process determines the observed characteristics of initiated cases. Specifically, this study tests the extent to which prosecutors initiated observed cases as predicted. If the process by which prosecutors choose cases to initiate is non-random, then the determinants of case outcomes include the case selection process.

The case selection model developed in this study generally specifies the prosecutor's case-bringing decision in terms of the prosecutor's *ex ante* expected value of initiating a potential case. In the case selection model, the prosecutor decides to initiate a case if the *ex ante* expected value of bringing the case is non-negative. Existing empirical studies of federal antitrust enforcement inform the case selection model's specification of the prosecutor's *ex ante* expected value of initiating a potential case. In particular, the empirical findings and expert opinions of antitrust scholars and practitioners, including Posner (1970), Siegfried (1975), Weaver (1977), and Baker (1978), strongly suggest that obtaining convictions is a Department of Justice (DOJ) Antitrust Division prosecutor's primary motivation for initiating a criminal antitrust case. Accordingly, in this study's case selection model, the prosecutor's *ex ante* expected value of initiating a case increases in the probabilities of conviction by trial or pleas of guilty or *nolo contendere*.

The prosecutor's case selection process is conceptually analogous to the trade-offs required by binding budget constraints. Imagine a child in a candy store with a dollar in her pocket, a love for chocolate bars that cost 75 cents apiece, and a lesser preference for bubble gum that costs 25 cents apiece. She cannot afford to buy more than one chocolate bar and one piece of bubble gum. If the child walks out of the candy store with no money left in her pocket, then it is reasonable to expect that she purchased a chocolate bar. If, for whatever reason, she spends her dollar but walks out of the store with two pieces of bubble gum, then we know that the other candy she purchased did not include one of her beloved chocolate bars.

Higher expected penalties elevate the prosecutor's *ex ante* expected value of bringing a case. Suppose the prosecutor initiates a case with a relatively high (or higher) expected sentence upon trial conviction, where the higher expected value is from an exogenous shock like the enactment of the Antitrust Policies and Procedures Act of 1974 (APPA).[1] Then it is likely that the *initiated* case will feature characteristics such as relatively low conspiracy sales or other relatively weak evidence with values (relative to the values expected in initiated cases with lower

---

[1] This law raised violations of Section 1 of the Sherman Act (e.g., price-fixing and bid-rigging) from misdemeanor to felony status and increased the statutory maximum penalties for these offenses.

expected penalties) that, *ceteris paribus*, reduce the prosecutor's *ex ante* expected value of bringing a case.

That is because the relatively high expected penalty due to the provisions of the APPA, *ceteris paribus*, increases the prosecutor's *ex ante* expected value of initiating the case. With that "surplus" in her *ex ante* expected value of bringing the case, the prosecutor can better "afford" to initiate the case even if it is further characterized by relatively weak evidence.[2]

Notwithstanding the APPA's tendency to raise expected sentences, weaker evidence in initiated cases due to the prosecutor's case selection process could translate, for example, into lower observed trial conviction rates. Snyder (1989, 1990) does not consider the potential effects of the case selection process, but his findings suggest that cases in which the provisions of the APPA applied involved relatively lower dollar amounts of conspiracy sales than cases in which the APPA did not apply. Snyder also finds that the observed trial conviction rate was lower in cases in which the APPA applied than in cases in which the APPA did not apply.

The theoretical framework developed below combines several existing models of the case disposition process in order to provide a more complete model of case disposition – from pre-indictment through sentencing. For example, this study uses the litigation model developed by Katz (1988) to specify the *ex ante* expected (optimal) values of the litigants' chosen levels of trial expenditure should the case go to trial. The litigants' respective expected trial expenditures determine the value of the endogenous variables that represents the litigants' common *ex ante* expected probability of trial conviction, as well as their common *ex ante* expected sentences following a trial conviction. These expected trial expenditures also influence the values of the endogenous variables that represent the *ex ante* expected probability of the defendant choosing to go to trial and the *ex ante* expected value of a sentence following a guilty or *nolo contendere* plea. All of these *ex ante* expected values affect the prosecutor's *ex ante* expected value of bringing a case, which is the fundamental equation of the case selection model. In this way this study uses the litigation model developed by Katz (1988) to improve Eisenberg and Farber's (2003) case selection model.

Moreover, specifying the litigants' trial expenditures as choice variables and the *ex ante* expected probability of trial conviction as an endogenous variable improves upon the plea bargaining models generally used in the law and economics literature. The existing models treat

---

2 Of course, the prosecutor generally would prefer initiating cases strengthened by sound evidence.

the *ex ante* probability of trial conviction as an exogenous variable and, at least implicitly, do not treat litigants' trial expenditures as choice variables. Such specifications inadequately capture the endogenous nature of the litigants' trial expenditure decisions and the *ex ante* probability of trial conviction.

This study also originates empirical tests for the existence of private information during plea bargaining, as well as the identity of the privately informed litigant, if one exists, during plea bargaining in federal criminal antitrust cases. The identity of the litigant that possesses private information in plea bargaining has important policy implications. In general, plea bargaining is expected to be socially efficient when the defendant is privately or better informed about his chances at trial, while plea bargaining is expected to be socially inefficient when the prosecutor is privately or better informed.

The defendant could have private or superior information about his chances at trial because he might be the only one who truly knows whether he committed the alleged crime. As noted by Reinganum (1988), however, the prosecutor could have better information about the defendant's chances at trial. In particular, the prosecutor could have better information on aspects of the case which do not necessarily depend on the actual actions of the defendant than the defendant has. For example, the prosecutor may have appeared more frequently or more recently before the judge(s) than the defendant's attorney(s) have. This could give the prosecutor more intimate familiarity with the judge's (current) legal philosophy or other judicial tendencies. This knowledge, in addition to the prosecutor's familiarity with the details of her case against the defendant, could give the prosecutor better information about the defendant's chances at trial.

When the defendant possesses private or better information about his chances at trial (including knowledge about the defendant's "true culpability"), plea bargaining facilitates gathering accurate (and socially useful) information about the defendant. Plea bargaining induces defendants with a high probability of conviction (which is safely assumed to be correlated with "true guilt") to plead guilty. The defendant's decision to accept or reject a plea deal can thus reveal his culpability.

Further, when defendants are privately informed about their chances at trial (or true level of culpability), plea bargained penalties that are lower than expected penalties upon trial conviction are justified in socially efficiency terms. That is because the discounted penalties can induce guilty defendants to admit their guilt and avoid adding to courts' or government prosecutors' workloads.[3]

In contrast, if the prosecutor is privately informed about the defendant's chances at trial, then the prosecutor can offer a sufficiently enticing plea deal that puts great pressure on defendants who may not be guilty (in truth or after trial) to plead guilty. In this scenario, wrongful convictions are more likely. As a result, society receives mixed or improper signals about legal behavior.[4]

Using theoretical and empirical analyses, this study provides answers to the following questions:

- How do the observable characteristics of cases influence i) the federal prosecutor's decision to initiate a criminal antitrust case, ii) the defendant's plea choice, iii) the judge's or jury's decision to convict a defendant at trial, and iv) the judge's sentencing decision after a trial conviction?

- How do the selection processes by which i) the federal prosecutor decides to initiate a case, ii) the defendant chooses a plea, and iii) the judge (or jury) decides to convict a defendant at trial influence the observable characteristics of criminal antitrust cases that proceed to subsequent stages of case disposition?

- How do these processes influence the predicted effects of case characteristics on observed federal criminal antitrust case outcomes?

- Did the observed characteristics of federal criminal antitrust cases influence, as theoretically predicted, the defendants' decisions to go to trial, the judges' (or juries') decisions to convict defendants at trial, and the judges' sentencing decisions after trial convictions in cases initiated from 1956 through 1979?

---

3 Grossman and Katz (1983); Froeb (1993).
4 Reinganum (1988); Froeb (1993).

- Did selection processes influence, as theoretically predicted, the observed characteristics of federal criminal antitrust cases that proceeded to subsequent stages of case disposition in cases initiated from 1956 through 1979?

- How does private information about the defendant's chances at trial held by one of the litigants during plea bargaining influence observed federal criminal antitrust case outcomes?

- Did defendants, prosecutors, or neither hold private information regarding their chances at trial during plea bargaining in federal criminal antitrust cases initiated from 1956 through 1979?

Chapter II of this study provides a review of relevant studies from the law and economics literature, including the antitrust literature. Previous studies, including Siegfried's (1975) analysis of DOJ Antitrust Division case-bringing activity, strongly suggest that the prosecutor's selection of cases to pursue – the prosecutor's "case selection process" – is not a random process in federal criminal antitrust enforcement. Snyder (1989, 1990) demonstrates that defendants in criminal antitrust cases do not randomly make plea choices. Meanwhile, Hylton's (1993) empirical findings from data on civil antitrust cases suggest that it is reasonable to expect defendants to possess private information during plea bargaining in federal criminal antitrust cases.

Chapters III-V fully develop components of the prosecutor's *ex ante* expected value of bringing a case. Chapter III defines the prosecutor's *ex ante* expected value of bringing a case. This expected value depends on the prosecutor's *ex ante* expected probabilities of possible plea bargaining outcomes and the prosecutor's *ex ante* expected values of sentences across modes of conviction. The prosecutor's *ex ante* expected value of initiating a case is central to the theoretical model of the prosecutor's case-bringing decision rule. According to this decision rule, the prosecutor will initiate a case if her *ex ante* expected value of initiating the case is non-negative.

Chapter IV develops a model of plea bargaining similar to Eisenberg and Farber's (2003) model, in which the defendant, instead of the prosecutor, has private information about his chances at trial. This model relies on a Nash solution in the value of a plea agreement. This model develops more fully the definition of the prosecutor's *ex ante* expected value of a negotiated plea agreement and her *ex ante* expected probability of trial.

Chapter V develops a model of litigation similar to the Katz (1988) model in which the prosecutor and defendant choose their levels of trial expenditure, given their unit costs of trial and other case characteristics, in order to maximize their expected values of litigation. The litigation model relies on Nash solutions in the litigants' trial expenditures and predicts how changes in the values of case characteristics will affect the litigants' optimal levels of trial expenditure. The litigants' optimal levels of trial expenditure directly affect the prosecutor's *ex ante* probability of trial and the prosecutor's *ex ante* expected probability of trial conviction. These *ex ante* expected probabilities are central components of the prosecutor's *ex ante* expected value of bringing a case. Accordingly, the litigation model predictions are integral to several testable implications of the theoretical framework.

Chapter VI defines and explains the case selection effect, which requires determining how marginal changes in individual case characteristics affect the prosecutor's expected value of bringing a case. Chapter VI also defines the *ex post* probability of trial, which is conditioned by the case selection process. In part, changes in the values of case characteristics are expected to influence the *ex post* probability of trial as they are expected to influence the *ex ante* expected probability of trial. Changes in the values of case characteristics also indirectly influence the *ex post* expected probability of trial through the "case selection effects" created by the case selection process that conditions the *ex post* expected probability of trial.

Chapter VII defines the *ex post* probability of trial conviction, which is conditioned by the case selection process and the trial selection process. Changes in the values of case characteristics influence the *ex post* probability of trial conviction, directly and, through case selection effects and trial selection effects, indirectly.

Chapter VIII defines the *ex post* expected sentence following a trial conviction, which is conditioned by the case selection process, the trial selection process, and the trial conviction selection process. All of these processes influence expected sentences imposed after trial convictions.

Chapter IX explains the normative and positive implications of one of the parties possessing private or better information during plea bargaining regarding the defendant's chances at trial. Chapter IX also explains how the case selection process can distort statistics popularly used in tests for the existence and possessor of private information in plea bargaining and other settlement negotiations. This chapter explores alternative approaches to testing the information structure of plea bargaining in federal criminal antitrust cases initiated from 1956 through 1980.

Chapter X describes the data used for this study's empirical tests and reports the test results. The results suggest that the selection processes involved in case disposition do not sort defendants (and the associated cases) randomly, and that they sort defendants according to the theoretical predictions developed in this study. The results further suggest that this study's theoretical specifications of the prosecutor's case selection process, the defendant's plea decision, and the judge's (or jury's) trial conviction decision process are applicable to different types of criminal cases.

This study conducts six tests of the information structure of plea bargaining in the sample of cases used in this study. Two test results suggest that the defendants held private or better information during plea bargaining, one result suggests that the prosecutors were privately or better informed, and three results suggest that neither party possessed private information during plea bargaining. These mixed results suggest that the informational structure of plea bargaining in federal criminal antitrust cases remains an open question. In normative terms, these mixed results leave open the question of whether plea bargaining was socially efficient or not in federal criminal antitrust cases initiated from 1956 through 1979.

Chapter XI summarizes this study's contributions to theoretical and empirical analyses of federal criminal antitrust enforcement. This chapter concludes with a discussion of unexplored research topics in federal criminal antitrust enforcement.

CHAPTER II

LITERATURE REVIEW

The following literature review summarizes the relevant contributions of primarily two branches of the law and economics literature. The first set of studies examines federal criminal antitrust enforcement. The second set explores the nature of sample selection biases created by the prosecutor's (or plaintiff's in the civil litigation literature) decision to initiate a case and the defendant's decision to go to trial. This literature review also identifies areas for investigation unexplored by the previous studies.

Prosecutors' Motivations for Initiating Cases

Posner (1970) began a criminal antitrust literature that provides alternative descriptions of and explanations for observed federal criminal antitrust enforcement activity and case outcomes. In particular, the criminal antitrust literature has focused on prosecutors' case-bringing activity, defendants' plea choices, trial conviction rates, and sentencing.

Posner (1970) suggests that the government's high winning percentage in antitrust cases could reflect an excessively cautious enforcement policy, from a legal perspective more than an economic efficiency perspective.[5] Posner's description of Department of Justice (DOJ) enforcement strongly suggests that prosecutors in the DOJ Antitrust Division place the highest priority on convictions. According to Posner,

> [I]t would appear that both legal doctrine and the enforcement machinery are geared more to the apprehension of unsuccessful attempts to fix prices than to the apprehension of successful price fixing. In general, the fact of an agreement or conspiracy to fix prices is all that the government need prove in a price-fixing case and all that it attempts to prove.[6]

Consistent with Posner's (1970) view of federal antitrust enforcement, Baker (1978) explains that the DOJ Antitrust Division generally files a companion civil case with any criminal

---

5 Posner also notes that the DOJ's high success rate in Supreme Court cases could reflect careful screening by the Solicitor General, as suggested by the weaker Supreme Court record of private claimants.

6 Posner (1970), 410

indictment when the subject conduct has not ended. According to Baker (1978), however, if the DOJ Antitrust Division does not think it can meet the "proof beyond a reasonable doubt" standard for blatantly illegal conduct, the Division simply does not initiate the case.

Siegfried (1975), meanwhile, reasons that if the DOJ Antitrust Division's objective is to maximize something akin to social welfare, then the Division will allocate resources to activities that create benefits that exceed the costs of successful prosecution. Potential social welfare benefits result from efficiency gains and income redistribution. Siegfried's empirical tests expand and improve upon the tests conducted by Long *et al*. (1973). Neither study, however, finds evidence that measures of the social welfare benefits from federal antitrust enforcement play significant roles in explaining DOJ antitrust case-bringing.

According to Siegfried (1975), these findings are not surprising given that the reward structure for prosecutors in the Antitrust Division of the DOJ is likely to favor winning cases rather than reducing social efficiency losses or social inequities. Nonetheless, Siegfried notes that pursuing cases with a high probability of successful prosecution could be consistent with some sort of economic welfare maximization. For instance, the DOJ may seek to maintain a credible threat of successful prosecution in order to dispose of cases by plea bargaining or deter potential violations.

Similarly, Asch (1975) concludes that the ambiguities of the DOJ's case selection process places limits on the inferences drawn from relations between antitrust enforcement activity and associated industry economic characteristics, such as the numbers of firms in the industries and average firm size. He notes, for example, that prosecutors may expect legal precedents created in certain cases to affect unrelated industries. Although Asch does not discuss this point, it is worth noting that trial convictions (more than guilty or *nolo contendere* plea agreements) set precedent favorable and useful to the DOJ. Thus, prosecutors use litigation for precedent-setting, which suggests both their interest and commitment to win at trial.

Weaver (1977) studies the factors that influence federal prosecutors' decisions in antitrust cases and finds that prosecutors are less likely to bring cases based on their economic significance and more likely to bring cases when the alleged activity is clearly illegal or when the cases otherwise can be won easily. Such empirical evidence of a typical DOJ Antitrust Division prosecutors' interest in convictions supports the assumption in this study that the prosecutor is conviction-motivated in order to specify theoretical predictions regarding the indirect effects of the prosecutor's case-bringing decision on the observed trial rate, trial conviction rate, and sentencing outcomes.

Per Se Illegality of Certain Types of Sherman Act Violations

Section 1 of the Sherman Act, which prohibits price-fixing and bid-rigging,[7] is particularly important in antitrust enforcement. According to Baker (1978), the *Trenton Potteries* case (1927) clearly established the *per se* illegality of price-fixing. The Supreme Court's decision in *Appalachian Coal* (1933) created some confusion regarding the *Trenton Potteries* decision. Then the *Saucony-Vacuum Oil* (1940) decision firmly reestablished the *per se* illegality of price-fixing.

Baker (1978) explains that traditional enforcement involves two types of cases: 1) cases that involve *per se* illegal activity regardless of other circumstances, and 2) cases that involve the application of the rule of reason that require consideration of circumstances beyond the business activity in question. Meanwhile, federal courts effectively recognize three categories of antitrust cases: 1) "hard core" *per se* cases involving price fixing and mark*et al*location;[8] 2) "soft core" per se cases in which courts consider facts surrounding conduct that is technically subject to per se rules, such as various vertical arrangements and joint ventures;[9] and 3) cases in which the rule of reason applies and all facts are relevant.[10]

The U.S. Supreme Court, however, has consistently rejected the creation of the middle category of "soft core" offenses that are *per se* illegal, even though the Supreme Court has changed its opinion about certain types of conduct. As Baker notes, for example, in *United States v. Arnold, Schwinn & Co.* (1967), the Supreme Court ruled that territorial restrictions are

---

7 Currently, Section 1 of the Sherman Antitrust Act, 15 U.S.C. §§ 1-7, which contains the provisions regarding "trusts, etc., in restraint of trade illegal; penalty," states:

> Every contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations, is declared to be illegal. Every person who shall make any contract or engage in any combination or conspiracy hereby declared to be illegal shall be deemed guilty of a felony, and, on conviction thereof, shall be punished by fine not exceeding $10,000,000 if a corporation, or, if any other person, $350,000, or by imprisonment not exceeding three years, or by both said punishments, in the discretion of the court.

8 See e.g., *United States v. Socony-Vacuum Oil Co*. 310 U.S. 150 (1940).

9 See e.g., *Fortner Enterprises v. United States Steel Corp*., 394 U.S. 495 (1969). See also *Worthern Bank & Trust Co. v. National Bank-Americard Inc*., 485 F.2d 119 (8th Cir. 1973).

10 See e.g., *Continental T.V., Inc. v. GTE Sylvania Inc*., 433 U.S. 36 (1977).

per se illegal. Then in *Continental T.V., Inc. v. GTE Sylvania Inc.* (1977), the Supreme Court reversed (or refined) its position about which types of territorial restraints are *per se* illegal.[11]

Justice Thurgood Marshall captured a popular justification for the judicial rules that Sherman Act Section 1 violations are per se illegal when he wrote:

> They are justified on the assumption that the gains from imposition of the rule will far outweigh the losses and that significant administrative advantages will result.[12]

Elzinga and Wood (1988) and Wood (1993) take issue with such cost-benefit justifications of *per se* rules against price-fixing and horizontal market division. They argue that the cost savings and the benefits of such *per se* rules may have been overstated and, because of case selection or litigation strategies, *per se* cases may be more expensive and produce fewer convictions than rule of reason cases.[13]

Notwithstanding the various views of the DOJ's methods of deciding which antitrust cases to initiate as well as the merits of *per se* rules against horizontal restraints of trade like price-fixing and bid-rigging, Gallo *et al*. (2000) provide useful statistics on observed DOJ antitrust case-bringing activity. They report that from 1955 through 1997, 80 percent of the total number of cases recorded by the Commerce Clearing House ("CCH") involved horizontal *per se* violations ("HPSVs"). Of the cases involving HPSVs, 80 percent were criminal cases. Further, criminal prosecution of HPSVs increased after 1980, such that the ratio of criminal to civil prosecutions of HPSVs went from about one-to-one in the 1960s to about seventeen-to-one in the 1980s.[14] Moreover, the DOJ's focus on HPSV cases during the 1980s is consistent with a

---

11 Echoing Posner's (1970) concerns, Baker (1978) also observes that the Supreme Court has greatly expanded the types of business conduct that is per se illegal, even conduct that enhances competition.

12 *U.S. v. Container Corp. of America*, 393 U.S. 333, 341 (1969).

13 Wood (1993) theoretically demonstrates how *per se* rules can increase litigation costs and lower the observed probability of conviction. As Wood explains intuitively, in cases involving *per se* rules, both parties may change their strategies. The defendant would increase legal efforts, which could increase the plaintiff's costs to the point of lowering the probability of conviction. Meanwhile, the plaintiff might respond to *per se* rules by filing weak cases that require more legal effort, thereby increasing litigation costs and lowering the probability of conviction. In his sample of private antitrust cases, Wood finds that civil price-fixing cases involved greater expenditure of legal resources than other types of civil antitrust cases under a rule of reason, as measured by the number of docket entries, duration, number of trial days, depositions, and judges' rulings. He also finds empirical evidence of lower conviction rates in *per se* cases compared to rule of reason cases.

14 Former Antitrust Division chief Donald I. Baker (1978) argues that the DOJ Antitrust Division's decision to bring a civil or criminal case follows extensive analysis of the facts and issues by trial staff and senior officials. Ultimately, the Assistant Attorney General decides whether to bring a civil or criminal case based upon articulated

prosecutorial objective of convicting defendants. From 1955 to 1997, the DOJ won about 92 percent of the total number of cases involving HPSVs.

Judicial Discretion in Sentencing

Another important aspect of federal criminal antitrust enforcement is the influence of idiosyncratic judicial discretion in sentencing prior to the introduction of federal sentencing guidelines.[15] Existing studies demonstrate, by the nature of their hypotheses or their empirical findings, the importance of judicial discretion in sentencing in federal criminal antitrust cases.

For example, Posner (1970) emphasizes that a change in statutory maximum penalties is an indirect method of changing imposed penalties and the magnitude of the changes in imposed penalties is uncertain. Posner appears to consider the effects of judicial discretion in sentencing the source of this uncertainty.

Cohen (1989, 1992) examines sentences imposed on criminal antitrust offenders from 1955 to 1980. Judges can impose fines on corporate defendants and fines and prison on individual defendants. Overall, Cohen's (1992) regression analysis of sentencing decisions shows that a judge's promotion potential positively affects the level of fines imposed on corporate defendants in federal HPSV cases. More specifically, Cohen (1992) also finds that judges appointed by Presidents Nixon and Carter imposed corporate fines that were statistically significantly lower than judges appointed by President Johnson imposed in federal HPSV antitrust cases. Thus, while the nature of the influence of political ideology on judges' discretionary (i.e., pre-sentencing guidelines) sentencing decisions, for example, may not be obvious, there is evidence for its existence.

Gallo *et al*. (2000) report that following the enactment of each penalty-enhancing law since 1955, there has been a significant increase in average real fines for both firm and individual defendants, but less than the proportionate increases in the statutory maximum penalties. The fines imposed on firms have increased more than fines imposed on individuals. According to Gallo *et al*. (2000), since 1955, the imprisonment of convicted individuals increased in

---

principles and the facts of the particular case, but also based partly upon intuition and experience. According to Baker, criminal cases usually involve conduct that gives the prosecutor no reasonable choice other than criminal prosecution.

[15] Only cases involving criminal antitrust offenses committed after November 1, 1987 are subject to the Sentencing Guidelines.

frequency, especially after 1974. Gallo *et al*. (2000) posit that the DOJ and the federal courts pragmatically fit the punishment to the crime instead of relying solely on fines or imprisonment for deterrence. Such speculation lends further support for the importance of judicial (and prosecutorial) discretion in sentences imposed for Sherman Act violations.

For further empirical support of the influence of judicial discretion in sentencing, consider the following empirical findings regarding the effect of the passage of the Antitrust Policies and Procedures Act ("APPA") of 1974. The APPA elevated violations of Section 1 of the Sherman Act from misdemeanor to felony status. The APPA also dramatically increased the statutory maximum penalties for violations of Section 1 of the Sherman Act. Statutory maximum fines increased from $50,000 to $1,000,000 per count.

The statutory maximum prison sentences rose from one year to three years per count. Gallo, *et al.* (1994) report nearly a six-fold increase in the average nominal corporate fine and nearly a threefold increase in the nominal fine per individual for cases following enactment of the APPA. In dollar amounts, however, these increases were small relative to the dollar increases in the statutory maximums.

Additionally, Cohen's (1992) regression analysis demonstrates that corporate fines were statistically significantly lower by more than $60,000 in cases brought under the misdemeanor regime compared to the felony regime. Further, the percentage of individuals receiving the maximum fine rose by just over one percent after the enactment of the APPA. In addition, Snyder (1990) reports that the change to the felony status brought significant increases in average fines. Meanwhile, Gallo, *et al*. (2000) report that the corresponding percentage of firms receiving the maximum fine decreased by roughly five percent.

With respect to incarceration, the APPA seems to have increased sanctions. According to Gallo, *et al*. (1994), the number of cases after 1974 in which judges imposed imprisonment doubled the corresponding number of cases for the previous fourteen-year period, yet the number of individuals sentenced to prison remained nearly constant. Moreover, Gallo, *et al*. find that the APPA brought a significant increase in the imposition of the maximum prison term.

Cohen (1992) reports a statistically significant lower probability of imprisonment for convictions occurring before the increase in statutory penalties for antitrust offenses. Cohen does not, however, find a statistically significant difference in the magnitudes of prison sentences before and after the APPA. Meanwhile, Snyder (1990) reports that felony cases, compared to misdemeanor cases, were associated with a higher frequency of prison sentences for individual defendants, and a higher average prison term per defendant.

Determinants of Observed Trial and Trial Conviction Rates

Existing studies in the criminal antitrust literature also examine the indirect effects of the shift from misdemeanor to felony penalties on non-sentencing aspects of case disposition. Snyder (1989, 1990) discusses and tests empirically various hypotheses regarding the determinants of observed trial rates and trial conviction rates in bid-rigging and price-fixing cases. Snyder's data cover the period 1970-1985. In particular, his hypotheses consider how defendants, judges, and, to a lesser extent, prosecutors would react to the enhanced penalties allowed under the APPA.

According to Snyder (1989, 1990) the possibility of harsher felony penalties reduces the frequency and severity of criminal activity. Snyder (1990) assumes that the probability of conviction and the expected penalty likely increase with the severity of the offense. Snyder expects that higher penalties will deter "marginal offenders" for whom the expected returns from illegal business conduct are similar to the expected returns from legal activity. Meanwhile, he predicts that potential defendants not deterred by higher penalties – the "intramarginal offenders" – will tend to reduce the frequency or severity of potentially illegal conduct.

The higher stakes associated with felony penalties also are likely to increase the defendants' willingness to go to trial and otherwise increase their defensive efforts.[16] According to Snyder, such reactions by potential and actual antitrust offenders to the shift to felony penalties from misemeanor penalties help explain the decline in economically meaningful criminal antitrust cases and the increase in trial rates.

In addition, Snyder (1989, 1990) proposes that social efficiency considerations cause the courts to be more concerned about legal error and thus exercise more care in deciding guilt in felony cases than when misdemeanor penalties applied. Since he finds no evidence that the courts were moving towards harsher penalties for criminal antitrust violations before the passage

16 Snyder bases these arguments on the assumption that risk-neutral potential defendants jointly decide whether to commit a crime, the severity of the crime, and their level of defensive efforts according to the maximization of income net of the expected costs from illegal activity. His assumptions are consistent with the foundations of this study's case selection and litigation models.

of the APPA, Snyder expects judges to raise the (implied) burden of proof for prosecutors under a felony penalty regime compared to a misdemeanor penalty regime. Thus, Snyder (1989, 1990) suggests that the courts' reluctance to apply harsher sanctions for antitrust offenses caused the trial conviction rate to decline.

Snyder (1990) also stresses the importance of the sequential nature of case disposition in determining the effects on trial outcomes from the shift to felony penalties. The cases that go to trial are selected according to the defendants' plea decisions. Because the set of cases that go to trial are not randomly drawn from the set of initiated cases, observed changes in trial conviction rates may be caused by changes in the types of cases that go to trial. Accordingly, notwithstanding Snyder's hypothesis that judges would raise their conviction standards in cases brought under the APPA, Snyder (1990) recognizes that the trial conviction rate could increase because the additional cases that go to trial (because of higher stakes of conviction) are likely to be drawn from cases involving defendants with greater probability of trial conviction.[17]

Meanwhile, Snyder (1990) notes that the government's case selection process could produce effects that mitigate the expected positive effect on trial rates that he expects the shift to felony penalties to produce. Snyder recognizes the potential for political influence on the enforcement decisions of prosecutors in the Antitrust Division of the DOJ, but concludes that the DOJ consistently vigorously enforces Section 1 of the Sherman Act.[18]

Instead, Snyder (1990) hypothesizes that prosecutors may raise their standards for the likelihood of conviction in cases they bring under a felony penalty regime. He notes that if prosecutors raise their standards for the likelihood of conviction in cases they bring under a felony penalty regime, they may bring fewer cases. Moreover, the defendants whom the

---

[17] Snyder (1990) does not mention any assumptions regarding the information structure of plea bargaining with respect to his expectation that additional cases that go to trial will be associated with higher probabilities of trial conviction. As explained later, this expectation is consistent with defendants possessing private information during plea bargaining about their chances at trial.

[18] Gallo *et al*. (2000) find that contrary to common assumptions, the presidential administration does not have much influence on the level of antitrust enforcement activity. Gallo *et al*. (2000) report that for the period 1953-1997, relative to the number of years spent controlling the White House, Democrat administrations brought slightly disproportionately more cases than Republican administrations. In addition, Gallo *et al*. (2000) find that Democrat administrations are more likely than Republican administrations to explore uncharted areas of antitrust law.

Likewise, Snyder (1989) rules out explanations based on the DOJ Antitrust Division's enforcement policy. Even though the Reagan administration sought to relax antitrust prohibitions on vertical restraints and mergers, it did not advocate reducing efforts to prosecute price-fixing cases, which was consistent with previous enforcement regimes. Snyder also rules out explanations based on budget cuts during the Reagan administration and the 1977 *Illinois Brick* decision. The *Illinois Brick* decision limits eligibility to sue for treble damages to direct customers of the cartel but nonetheless assigned damage rights, thereby encouraging appropriate civil claims while deterring price-fixing activity.

prosecutors indict (with higher standards of conviction) may decide to avoid trial.[19] Snyder also suggests that this increase in the prosecutors' standards of conviction could increase trial conviction rates.

Snyder's (1989, 1990) empirical findings regarding the effects of the shift to felony penalties support his predictions as well as my case selection model's predictions under certain assumptions. In his earlier study, Snyder (1989) uses multinomial logit regression techniques to estimate the determinants of the relative probabilities of plea outcomes. In his later study, Snyder (1990) uses bivariate probit analysis to estimate the determinants of the probability of trial (i.e., not guilty pleas versus *nolo contendere* and guilty pleas) and the probability of trial conviction.

In particular, Snyder (1989) observes that after 1979,[20] the total number of HPSV cases filed increased.[21] With respect to trial rates, Snyder (1989) reports that for individual and corporate defendants, the statutory application of felony penalties reduces the probability of *nolo contendere* pleas in favor of not guilty pleas.[22] For individual defendants, however, Snyder (1989) finds that the shift to felony penalties had no significant effect on the probability of guilty pleas relative to not guilty pleas. In contrast, for corporate defendants, he finds that felony penalties increase the probability of guilty pleas relative to not guilty pleas. For HPSV cases brought during the period 1970-1985, Snyder (1990) finds that the applicability of felony penalties increases the observed trial rate for corporate and individual defendants.

Based on their inspection of the data they compile regarding federal antitrust enforcement, Gallo *et al*. (2000) observe changes in plea outcomes in criminal antitrust cases after 1979, the first year cases were brought under the APPA. They disagree with Snyder (1989,

---

19 This reason for lower trial rates in cases brought under the APPA implicitly assumes or recognizes that relative to the statutory maximum penalties allowed under the APPA, judges did not significantly increase the levels of the sentences they imposed under the APPA. It is possible, therefore, that the parties' expected sentences under the APPA did not constitute sufficiently large trial stakes to induce defendants to trial despite the prosecutors' elevated conviction standard.

20 Felony penalties under the APPA applied for violations that occurred after 1974. Thus, in legal terms, the provisions of the APPA apply in cases initiated in 1979 and later.

21 The data reported by Gallo *et al*. (2000) supports Snyder's observation.

22 Gallo, *et al*. (2000) provide an institutional explanation for the timing of the observed decrease in the frequency of *nolo contendere* pleas. As previously mentioned, the DOJ has a long-standing policy of opposing a *nolo contendere* plea if the violation was blatant, if the defendant would not be exposed to the same sentence as she would by pleading guilty, or if a conviction would aid plaintiffs in treble damage suits against the defendant. The DOJ reinforced this policy in 1979, however, by directing prosecutors to oppose *nolo contendere* pleas unless they receive oversight approval to do otherwise.

1990), however, about whether or not the trial rate increased after 1979. They also disagree slightly with Snyder about whether the shift to felony penalties under the APPA or a shift in DOJ Antitrust Division plea bargaining policy caused the changes in plea outcomes.

Gallo, *et al*. (2000) do not find an increase in the trial rate after 1979. They find changes in the observed annual rates of *nolo contendere* and guilty pleas, however. According to their calculations, the percentage of total HPSV cases ending in conviction that were disposed by *nolo contendere* pleas fell from 88 percent during 1975 to 1979 to 24 percent during 1980 to 1984. Meanwhile, the same statistic for guilty pleas was 11 percent for the period 1975 to 1979 and 75 percent for the period 1980 to 1984.

According to Snyder (1990), "[p]rice-fixing cases are the core of Antitrust Division enforcement activity, are a key input into private enforcement, and their deterrent effects may account for the greatest welfare gains from antitrust enforcement." In contrast, Snyder finds that bid-rigging cases usually involve less economic harm. In cases initiated between 1970 and 1985, Snyder finds that the conspiracy sales involved in price-fixing cases were many times larger than the amount of commerce involved in bid-rigging cases.

Snyder (1990) also reports a sharp decrease in average sales involved in price-fixing cases after 1979.[23] Snyder (1990) finds, however, that the number of bid-rigging cases steadily increased while the number of price-fixing cases fell after the shift to felony penalties.

Snyder (1989) also finds an increase in the number of cases disposed by plea agreement from 1970 to 1985, most of which were bid-rigging cases.

Snyder also analyzes the effects of higher conspiracy sales on the disposition of price-fixing and bid-rigging cases. Snyder (1989) finds that the higher the sales affected by the alleged

---

[23] Snyder's finding applies across case outcomes and comports with Joyce's (1989) finding that corporate defendants in price-fixing cases during the 1980s tended to have few assets. Gallo *et al*. (2000) also find that after 1980, criminal convictions have almost exclusively involved relatively small firms. In addition, Snyder (1990), like Marvel, Netter and Robinson (1988), points out that the decline in the number of significant price-fixing cases following the enactment of the APPA may explain the decline in the number of follow-on private suits.

Marvel, *et al.* argue that federal prosecutors were able to win many cases in the 1970s that involved ineffective cartels because the defendants in those cases did not strongly contest the charges, even if they did not actually attempt to fix prices. The reason for weak defense efforts in the 1970s, according to Marvel *et al*., is that the expected penalties following criminal convictions did not outweigh the defendants' expected litigation costs. After matching a sample of federal criminal price fixing cases from 1972-1979 with subsequent civil (treble) damage filings against the same defendants, Marvel *et al*. find that nearly half of the price-fixing cases in their sample did not result in any civil follow-on civil suits for damages. Based on these findings, they argue that the cartels that were prosecuted were likely to have been ineffective in elevating prices above competitive levels and earning monopoly rents. Moreover, their findings and conclusions are consistent with those of Posner (1970), Long *et al*. (1973), Siegfried (1975) and Weaver (1977) regarding federal antitrust enforcement in general.

conspiracy, the higher the probability that corporate defendants will plead not guilty relative to the other pleas. Snyder (1990) finds that a 10 percent increase in sales over its mean increases the probability that the defendant goes to trial by just over one percent (and decreases the probability of trial conviction by less than one percent).

Snyder (1990) argues that, taken together, his regression estimates of the determinants of plea choice demonstrate that the shift to felony penalties produced a broad range of effects. In fact, Snyder (1990) suggests that his findings of a rise in the trial rate and a reduction in the trial conviction rate after 1979 provides evidence of a combination of effects on case disposition from the shift to felony penalties. In particular, he argues that the rise in the trial rate after 1979 clearly indicates that defendants (whom felony penalties did not deter from potentially illegal activity) increased their defensive efforts in response to the higher-stakes felony penalty regime. Moreover, Snyder argues that these effects outweighed the mitigating effects of other (e.g., prosecutors' and judges') reactions to the APPA.

This study follows Snyder's (1990) suggestion to research further the effects of harsher potential sanctions on individual behavior and the enforcement process. This study does not simply restate others' research in terms of the case selection model developed later. In fact, this study's case selection model relies on factors not considered by Snyder.

As previously mentioned, Snyder's explanation for the higher trial rate when (alleged) conspiracy sales are higher is that a defendant is willing to go to trial – even when the evidence is stacked against him – because he expects the sympathetic judge to raise the conviction standard (at least in that particular case).[24] Notwithstanding the accuracy of Snyder's depiction of federal judges, this study demonstrates that the process by which federal prosecutors initiate certain cases instead of other potential cases could at least partly drive Snyder's empirical findings.

For example, the case selection model demonstrates how the prosecutor's case selection process, indpendent of the potential deterrent effect highlighted by Snyder, increases the likelihood that the prosecutor will initiate weaker cases (from the prosecutor's perspective) when felony penalties apply than when misdemeanor penalties applied. The case selection model generally specifies the prosecutor's case-bringing decision in terms of the prosecutor's *ex ante*

---

[24] Cohen and Scheffman (1989) use this logic to predict the effects of sentencing guidelines on federal criminal antitrust case disposition.

expected value of initiating a potential case. If that expected value is non-negative, the prosecutor is "duty bound" to bring the case.

In particular, as shown below, higher expected penalties elevate the prosecutor's *ex ante* expected value of bringing a case. Suppose the prosecutor initiates a case with a relatively high (or higher) expected sentence upon trial conviction (where the higher expected value is from an exogenous shock like the enactment of the APPA). Then it is just as likely that the *initiated* case will feature case characteristics such as relatively low conspiracy sales or other relatively weak evidence with values (relative to the values expected in initiated cases with lower expected penalties) that, *ceteris paribus*, tend to reduce the prosecutor's *ex ante* expected value of bringing a case, as it is likely that such an initiated case will feature relatively strong evidence that, *ceteris paribus*, tend to increase the prosecutor's *ex ante* expected value of bringing a case. Weaker evidence in initiated cases could translate into higher observed trial rates, lower observed trial conviction rates, and lower observed sentences.

In this way, the shift to felony penalties can affect the mix of case characteristics in cases that the prosecutor decides to initiate. The mix of case characteristics in initiated cases that results from the prosecutor's case-bringing decision (i.e., case selection) process, in turn, is expected to influence several observed case outcomes, including the defendant's plea decision, the judge's trial conviction decision, and the judge's sentencing decision after conviction by trial or plea.

To various but limited extents, the criminal antitrust literature discusses and empirically tests for selection effects (i.e., selection bias) on observed case outcomes, where the defendant's choice of type (or degree) of illegal business activity and his plea choice, as well as the judge's trial conviction decision create these selection effects.[25] The authors pay little attention,

---

[25] For example, in developing his empirical test of the determinants of judges' prison sentencing decisions, Cohen observes that the determination of a jail sentence is a sequential decision problem, spanning the conviction decision through the judge's determination of the length of the jail term. This description of the determination of a jail sentence focuses mainly, if not exclusively, on judges' decisions at various stages of criminal antitrust case disposition. The prosecutor, defendant, and judge all may make decisions that determine a jail sentence. For example, in the sequence of decisions that determine a prison sentence, the prosecutor decides to bring the case, the defendant decides which plea to enter, (the prosecutor could accept that plea or object to that plea, as Cohen notes), and the judge decides whether to accept or reject that plea (possibly over the prosecutor's objection).

Cohen (1992) does not find any statistically significant difference in corporate fines imposed across means of conviction – i.e., trial conviction, non-negotiated guilty plea, negotiated guilty plea, negotiated *nolo contendere* plea, and negotiated *nolo contendere* plea accepted by the judge over the government's objection. In contrast, Cohen finds that the type of plea is a significant determinant of prison sentencing. Specifically, Cohen finds a statistically significantly higher probability of a jail sentence and a statistically significantly longer jail sentence in however, to the case selection process by which prosecutors decide which cases to initiate. The implications of the prosecutor's case selection process deserve more attention.

Information Structure of Plea Bargaining

The criminal antitrust literature does not explicitly address the issue of asymmetric or private information in plea bargaining and how private information during plea bargaining could alter the trial selection process and create (additional) bias in observed case outcomes. This is a serious omission given the policy implications of private information during plea bargaining. Privately informed prosecutors harm social welfare because they can fashion a settlement offer that induces (truly) innocent defendants to plead guilty or *nolo contendere*. Such pleas send false signals about socially efficient business conduct.[26] In contrast, privately informed defendants during plea bargaining benefit social welfare because the (truly) guilty ones are more likely to plead guilty and send the correct signals about appropriate business conduct. This dissertation attempts to contribute analyses of case selection and private information during plea bargaining in price-fixing and bid-rigging cases that are lacking in the criminal antitrust literature.

The asymmetric/private information theory developed by Bebchuck (1984) and others in the trial selection bias literature receives increasing attention. According to this theory, one party knows the probability that the plaintiff will win at trial, while the other knows only the distribution of plaintiff victory probabilities. When the defendant is better or privately informed, he will accept the (uninformed) plaintiff's settlement offer if he is relatively pessimistic about his chances. Meanwhile, the better or privately informed defendant will go to trial if he (correctly) expects to win at trial. Under this theory, the selection of cases for trial is one-sided, and the *ex*

---

cases in which the defendant pleaded not guilty. Cohen also finds a statistically significantly lower probability of a jail sentence in cases in which the defendant entered a *nolo contendere* plea over the prosecutor's objection. (Cohen uses the finding of the longer jail sentences for not guilty pleas to support the notion that judges "penalize" guilty defendants for using court time for trial.) Elsewhere, however, Cohen notes that he found no evidence of selection bias when estimating sample selection models using Heckman's (1976) estimation technique, with its well-known "Heckman's lambda" term.

[26] Cohen and Scheffman (1989) point out the potential for socially inefficient plea bargaining with respect to the bargaining power transferred to prosecutors by sentencing guidelines, but they do not discuss the role of private information *per se*.

post probability of a plaintiff trial victory is systematically below the fraction of successful plaintiffs in the pool of legal disputes.[27]

Empirical tests for the existence of private information (and which litigant possesses private information) in settlement negotiations also focus mainly on the correlation between trial rates and plaintiff victories at trial. In general, this literature, which includes a wide variety of theoretical modeling approaches, posits that a positive (negative) correlation between trial rates and plaintiff victories at trial is evidence of privately informed defendants (plaintiffs) during settlement negotiations.

Waldfogel (1998) uses three variables to control for the likely informational structure of settlement negotiations in civil cases. First, he suspects the fraction of *pro se* plaintiffs may be related to the quality of plaintiffs' information because attorneys are better informed about likely case outcomes than are non-attorneys. Second, he hypothesizes that parties who are repeat players in civil litigation are better informed than are non-repeat players. Third, he expects institutions to be better informed than individuals. The last two approaches to controlling for private information could be used for federal criminal antitrust case data. For example, repeat defendants could be relatively better informed than are non-repeat defendants (compared to prosecutors). Moreover, corporate defendants could be relatively better informed than are individual defendants (compared to prosecutors).

Hylton (1993) expects that defendants in civil antitrust cases are likely to have good information about their chances at trial. Hylton also claims that if defendants possess more or better information about their chances at trial than plaintiffs possess, then as legal doctrine

---

27 Priest and Klein (1984) and others that followed develop the divergent expectations theory, in which each party estimates, with error, the likelihood of a plaintiff victory at trial, and legal disputes proceed to trial when the plaintiff is (randomly) sufficiently more optimistic than the defendant. Accordingly, legal disputes in which the parties' valuation of the quality of the plaintiff's claim(s) is far above or below the parties' expected decision standard are more likely to settle. Conversely, legal disputes that proceed to trial are likely those in which the difference between the parties' perceived quality of the plaintiff's claim(s) and the parties' expected decision standard is smaller. The *ex post* probability of a plaintiff trial victory is expected to be close to 50 percent regardless of the fraction of plaintiff winners in the pool of legal disputes.

The private/asymmetric information and divergent expectations theories compete in the trial selection bias literature to predict and explain largely the same types of data and sometimes the same data. The data used in these studies almost exclusively measure only observed trial rates and, in some studies, observed trial outcomes across different types of cases (e.g., product liability or employment discrimination). In particular, the empirical tests of divergent expectations theories usually consist of comparisons of trial outcomes (i.e., the rate of plaintiff trial victory) across types of cases or comparisons of observed trial outcomes to the fifty percent plaintiff victory rate. Fifty percent is the central tendency for plaintiff trial victory rates under the divergent expectations theory. Shavell (1996), however, demonstrates that it is possible for civil cases that go to trial to result in plaintiff victory with any probability.

reduces the likelihood of judicial error over time, the plaintiffs' win rate will fall over time. Hylton notes that Salop and White (1988) find that the annual average plaintiff win rate in civil antitrust cases has fallen over time.

Hylton's (1993) hypothesis and Salop and White's (1988) finding invite examination of the information structure of plea bargaining in federal criminal antitrust cases. The development of legal doctrine is pronounced in federal criminal antitrust cases (as well as in civil antitrust cases). Using Hylton's hypothesis, measures of prosecutors' trial win rates over time or with the development of antitrust legal doctrine could test for private information among litigants in federal criminal antitrust cases.

The trial selection bias literature focuses almost entirely on civil litigation. This literature *completely ignores* criminal antitrust data. Moreover, Froeb (1993) appears to be one of a few studies, along with Grossman and Katz (1984), to theoretically model plea bargaining with asymmetric information in criminal cases. Froeb (1993) also appears to be the *only* study to empirically test for the trial selection effects of private information during plea bargaining in criminal cases (and he does not use criminal antitrust case data).

The studies that develop or consider models of asymmetric information during settlement negotiations seem to arbitrarily treat the probability of a plaintiff's or prosecutor's victory at trial as an exogenous endowment about which one of the litigants has better information. The true probability of a plaintiff's or prosecutor's victory at trial (which one party knows) is assumed to be revealed at trial, which is consistent with Nash solutions to games of settlement negotiations or plea bargaining according to Froeb (1993). (These studies do treat the probability of trial and the value of the settlement as endogenous variables, however.) The theoretical framework in this study also treats the probability of trial conviction as an endogenous variable.

In departures from the trial selection bias literature, Eisenberg and Farber (1997, 2003) were some of the first, if not the only, studies to analyze the potential for the selection process by which a plaintiff chooses to file a suit to create a "case selection effect" which can influence (bias) observed case outcomes. When investigating or interpreting observed case outcomes, the case selection effect is just as legitimate as the trial selection effect described in the trial selection bias literature. Eisenberg and Farber (1997, 2003) use models of settlement negotiation that assume, without much discussion at all, that the defendant has private information about his chances at trial. They do not consider the potential for the plaintiff to have private information about the defendant's chances at trial. Further, like the studies found in the trial selection literature, Eisenberg and Farber (1997, 2003) treat the probability of a plaintiff trial victory (and expected damage award) as exogenous, such that case characteristics do not directly influence the probability of plaintiff trial victory. Instead, in the Eisenberg and Farber (2003) model, only Nature and the (indirect) case selection effect determine the observed rate of plaintiff victory at trial.

It is not clear that Eisenberg and Farber (2003) actually test for case selection bias in observed trial and trial conviction rates. They focus on changes in the values of case characteristics that produce (additive) direct and indirect case selection effects with the same sign. In addition, the prediction that forms the basis of one of their tests depends on the assumption that the direct effect outweighs the indirect case selection effect.

Department of Justice Antitrust Division Budget Levels and the (Opportunity) Costs of Prosecutors' Trial Expenditures

A few studies in the criminal antitrust literature briefly discuss the implications of the DOJ Antitrust Division budget levels on enforcement activity. Posner (1970) finds that between 1956 and 1967, the DOJ Antitrust Division's funding more than doubled, yet the number of Antitrust Division personnel rose by less than one-third. This increase in personnel, albeit less than proportional to the budget increase, was material. Posner does not find a corresponding significant increase in the number of cases brought, however.

According to Posner (1970), these findings support his hypothesis that the price of resources used for antitrust enforcement could increase faster than prices in general. In addition, these findings could suggest a high (private sector) market demand for antitrust attorneys, whose fees increase faster than inflation. If the average salary for a private antitrust attorney is higher than the average salary for a government prosecutor, then it is reasonable to assume that the defendant's unit cost of litigation efforts is relatively higher than the prosecutor's unit cost.

Snyder (1990) discusses DOJ Antitrust Division budget cuts as an explanation for a decline in federal antitrust enforcement from the mid-1970s through the 1980s. He notes that during the 1980s, budgets remained roughly constant in nominal dollars and the attorney staff was cut by about one-third. He argues that while enforcement must have declined in some areas, enforcement of Section 1 of the Sherman Act did not decline. Snyder (1989) notes that the Reagan administration was never accused of not pursuing alleged price-fixing and bid-rigging activity.

Despite his discussion of budget effects on DOJ Antitrust Division enforcement, Snyder (1989, 1990) does not include DOJ Antitrust Division budget information as explanatory variables in his regressions to explain trial and trial conviction rates. Similarly, Long *et al*. (1973), Siegfried (1975), and Cohen (1992) do not make use of DOJ Antitrust Division budget information in their empirical tests of the determinants of case-bringing activity and sentencing decisions. Moreover, Gallo *et al*. (2000) does not track such budget information. In fact, Posner (1970) does not present raw budget data, either.

This factor suggests a significant omission in the empirical studies found in the criminal antitrust literature. To fill this research void, this study investigates such questions as whether DOJ Antitrust Division budgets influence trial rates.

CHAPTER III

CASE SELECTION MODEL

The "case selection model" of case disposition that is developed in this study focuses on five stages of case disposition: the pre-indictment stage when evidence is gathered, the case selection (indictment) stage , the plea decision stage (with simultaneous plea bargaining), the trial litigation stage (if the case goes to trial) and the sentencing stage at the end of the paths to conviction – pleas of guilty or *nolo contendere* and conviction at trial. The fundamental decision rule that drives the case selection model of case outcomes is the federal prosecutor's decision to bring a case, based on the information available to her before she indicts an individual or corporate defendant for violating the Sherman Act.

The Prosecutor's Decision to Initiate a Case

In the pre-indictment phase (which could include a grand jury investigation, for example), the federal prosecutor is endowed with knowledge about a potential Sherman Act violation such as price-fixing or bid-rigging that she knows she should prosecute criminally.[28] At this stage she also knows the identity of the alleged criminal(s). For simplicity, this study assumes the prosecutor is randomly tipped off by some credible informant. Thus, if the prosecutor's case selection process does not create a biased sample of cases relative to the universe of potential cases, then cases are randomly chosen in the model. This study does not consider, for example, the deterrent effect of antitrust enforcement.

Based on the evidence she gathers before she indicts and her pre-indictment (i.e., *ex ante*) expected values of certain case characteristics, the prosecutor evaluates her *ex ante* expected probability of trial, Θ , as well as her *ex ante* expected value of her payoff from the case if it goes to trial, V^T , and if it is disposed by plea agreement, V^B . The prosecutor then puts all of the information together in order to evaluate her *ex ante* expected value of bringing that particular potential case $E(V)$, which can be expressed generally as

[28] In general, Baker (1978) asserts that the area of overlap of the Sherman Act's criminal and civil provisions is small and criminal cases usually involve conduct that gives the prosecutor no reasonable choice other than criminal prosecution.

$$E(V) = \Theta V^T + (1-\Theta)V^B . \qquad [1]$$

Assume the prosecutor is "duty bound" to indict a defendant and thus initiate a case if $E(V) \geq 0$. Accordingly, the predictions of the case selection model of case disposition depend on the influence of various factors on $E(V)$. Different values of these factors produce different values of $E(V)$ and thus determine the types or characteristics of cases the prosecutor initiates pursuant to the Sherman Act. .

A brief description of the prosecutor's case selection process is useful in describing the key insight provided by the case selection model. Assume the prosecutor faces a potential case with a relatively extreme-valued case characteristic that has a positive (negative) effect on $E(V)$. With the "cushion" ("deficit") created by that extreme-valued case characteristic, the prosecutor is willing to initiate the case even if (only if) other case characteristics – even (especially) those with relatively extreme values – have negative (positive) effects on $E(V)$ as long as (so that) the condition $E(V) \geq 0$ holds.

Variables Used to Define the Prosecutor's *Ex Ante* Expected Value of Bringing a Case

In order to use the case selection model to generate specific predictions that can be tested with a rich set of case-level data, this chapter expands the components of the prosecutor's *ex ante* expected value of bringing a case $E(V)$. In the pre-indictment phase, the federal prosecutor is endowed (i.e., provided by her informant) with a given quantity and/or quality of evidence, represented by the vector e. The prosecutor's evidence, e, determines the merits of the case. She also knows the values of the exogenous case characteristics, which include the prosecutor's unit cost of trial expenditure, C, and the defendant's unit cost of trial expenditure, K.

The litigation model developed below provides the optimal (Nash equilibrium) values of the prosecutor's and the defendant's respective choices of levels of trial expenditure, $X = X^*$ and $Y = Y^*$. At the pre-indictment stage, the prosecutor evaluates all of her *ex ante* expected values with $X = X^*$ and $Y = Y^*$. (As shown below, both X^* and Y^* are implicit functions of the exogenous case characteristics and other factors as specified by the litigation model.)

An endogenous variable that the prosecutor evaluates in the pre-indictment stage (using a plea-bargaining model developed in a later chapter) is her *ex ante* expected probability of a trial (i.e., not guilty plea), Θ. The value of Θ depends on the exogenous case characteristics, the litigants' *ex ante* expected chosen levels of trial expenditure, $X = X^*$ and $Y = Y^*$, as well as the parties' common *ex ante* expected trial sentence, S^T, which is endogenous and depends on the parties' levels of trial expenditure, X and Y and the prosecutor's evidence, e.

In the pre-indictment stage, the prosecutor expects the parties to have potentially divergent expectations about the probability of a trial conviction. Specifically, the prosecutor's *ex ante* expected probability of trial conviction, Φ, is defined as

$$\Phi = \hat{\Pi} + \rho, \qquad [2]$$

and the defendant's expected probability of trial conviction (from the prosecutor's perspective) is analogously defined as

$$\Gamma = \hat{\Pi} + \delta. \qquad [3]$$

These expected probabilities of trial conviction share a common component, $\hat{\Pi}$, while ρ and δ are idiosyncratic components, the values of which may diverge.

Assume ρ and δ have means of zero and variances of one. Thus, ρ and δ can be of either sign and can be positively or negatively correlated with the *ex ante* probability of trial conviction, $\hat{\Pi}$. For example, the prosecutor would be "optimistic" if ρ is positively correlated with $\hat{\Pi}$, and "pessimistic" if δ is negatively correlated with $\hat{\Pi}$. Both parties can be optimistic or pessimistic about their chances at trial to the same or differing degrees. For example, both parties could be optimistic, or one party can be optimistic while the other is pessimistic.

As discussed later, the parties commonly *ex ante* expected probability of trial conviction, $\hat{\Pi}$, is an endogenous variable that is explicitly a function of the parties' levels of trial expenditure, X and Y; the prosecutor's evidence, e; as well as the parties' common *ex ante* expected level of variability of the trial conviction standard, λ.

The prosecutor is assumed to be risk neutral and expects *ex ante* the following net value[29] of going to trial:

$$V^T = \Phi S^T - CX \ . \qquad [4]$$

The prosecutor's *ex ante* expected value of accepting a *nolo contendere* or guilty plea is

$$V^B = S^B , \qquad [5]$$

The prosecutor's *ex ante* expectation of the defendant's net value of going to trial is the following:

$$U^T = -\Gamma S^T - KY \ . \qquad [6]$$

From the prosecutor's *ex ante* perspective, the defendant's value of entering a *nolo contendere* or guilty plea agreement is

$$U^B = -S^B . \qquad [7]$$

For a list of the key variables used in the theoretical frame work developed above and below, see Table 1. Table 1 defines the variables and identifies whether the variables are exogenous, endogenous, or choice variables in the theoretical framework.

[29] For convenience in developing my case and trial selection model, assume the costs and benefits to the prosecutor from her pre-indictment activity are sunk or otherwise irrelevant here. As shown later, the litigation model considers the prosecutor's pre-indictment activity.

Table 1. Definitions of variables used in theoretical framework

Theoretical Framework Variable	Type of Variable	Theoretical Framework Definition
C	Exogenous	the unit cost of trial expenditure for the prosecutor
K	Exogenous	the unit cost of trial expenditure for the defendant
$\mathbf{e}$	Exogenous	the strength and amount of the prosecutor's evidence
λ	Exogenous	the litigants' common expected variability of the judge's trial conviction standard
ρ	Exogenous	the idiosyncratic component of the prosecutor's *ex ante* expected probability of trial conviction
M	Exogenous	the statutory maximum penalty
m	Exogenous	the statutory minimum penalty (which is zero until the implementation of the federal sentencing guidelines in 1987)
X	Choice	the prosecutor's level of trial expenditure
Y	Choice	the defendant's level of trial expenditure
X^*	optimal value of a choice variable and implicit function of other variables	the prosecutor's optimal level of trial expenditure
Y^*	optimal value of a choice variable and implicit function of other variables	the defendant's optimal level of trial expenditure
Θ	endogenous, $\Theta = \Psi\left(\rho - \frac{CX + KY}{S^T}\right)$	the *ex ante* expected probability of trial
$\hat{\Pi}$	endogenous, $\hat{\Pi} = \Pi([\mathbf{e} + P(X) - D(Y)]/\lambda)$	the litigants' common *ex ante* expected probability of trial conviction
S^T	endogenous, $S^T = m + \mu^T(\mathbf{e}, P(X), D(Y))(M - m)$	the litigants' common *ex ante* expected penalty following a trial conviction
S^B	endogenous, $S^B = (\hat{\Pi} + \rho)S^T + \frac{KY - CX}{2}$	the litigants' common *ex ante* expected penalty following a conviction by plea of guilty or *nolo contendere*

CHAPTER IV

PLEA BARGAINING MODEL

The previous chapter introduced several *ex ante* expected values of possible case outcomes. This chapter more fully defines the prosecutor's *ex ante* expected probability of trial, Θ, and her *ex ante* expected value of a plea agreement, $V^B = S^B$.

The Prosecutor's *Ex Ante* Expected Probability of a Not Guilty Plea

In order to derive a mathematical definition of the prosecutor's *ex ante* expected probability of trial, Θ, first consider the necessary conditions for the prosecutor to accept a negotiated plea with an *ex ante* value to the prosecutor of V^B.[30] The *ex ante* value of a negotiated settlement, V^B, must fall within the applicable "settlement zone". A settlement zone for a particular type of plea is the range of potential plea bargaining outcomes involving the particular type of plea that both the prosecutor and defendant would prefer over going to trial – all from the prosecutor's perspective.

That is, the prosecutor will accept a guilty or *nolo contendere* plea if the prosecutor values that plea at least as much as going to trial and the defendant will make an (acceptable) offer only if he values that plea at least as much as going to trial. Thus, the following condition must hold:

$$V^T \le V^B \le -U^T, \qquad [8]$$

which can also be expressed as

$$\Phi S^T - CX \le S^B \le \Gamma S^T + KY, \qquad [8']$$

or

[30] According to Sullivan (1977), the DOJ Antitrust Division does not initiate negotiations but will participate in negotiations initiated by defendants.

$$S^B \leq \Gamma S^T - \Phi S^T + KY + CX \ . \qquad [8'']$$

In terms of my model's parameters, this implies the following settlement zone:

$$Z = (\Gamma - \Phi)S^T + CX + KY \ . \qquad [9]$$

For a case to be disposed by trial, the settlement zone must be non-zero. Thus, for a case to be disposed by trial, the settlement zone must be negative and the following constraint must hold:

$$\delta < \rho - \frac{(CX + KY)}{S^T}, \qquad [10]$$

Importantly, assume here that at the time the prosecutor decides to bring a case based on the valuation described by Equation [1], she does not know the idiosyncratic component of the defendant's expected probability of a trial conviction, δ. Instead, assume that at the time she selects a case, she knows that δ is drawn from a distribution with the cumulative distribution function $\Psi(\cdot)$.[31] Thus, the prosecutor's *ex ante* probability of trial, Θ, can be defined as follows:

$$0 \leq \Theta = \Psi\left(\rho - \frac{CX + KY}{S^T}\right) \leq 1 \, . \qquad [11]$$

The *ex ante* probability of trial from the prosecutor's perspective, Θ, is higher when the idiosyncratic component of the prosecutor's expected probability of trial conviction is higher. Conversely, Θ is lower when either parties' trial costs are higher. Lastly, note that the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, does not appear on the right side of Equation [20] and thus does not directly affect Θ.

---

[31] Later, this study examines the implications of assuming instead that the prosecutor has an informational advantage over the defendant.

Nash Equilibrium Definition of the Prosecutor's (Conditional)
Ex Ante Expected Value of a Plea Agreement

For convenience, assume that from the prosecutor's perspective, the defendant's expected net gain from going to trial depends on the prosecutor's *ex ante* expected probability of trial conviction, Φ. That is, in the model of the plea agreement, substitute the prosecutor's *ex ante* expected probability of trial conviction, $\Phi = \hat{\Pi} + \rho$, for the defendant's *ex ante* expected probability of trial conviction, $\Gamma = \hat{\Pi} + \delta$. In other words, assume that the idiosyncratic difference between the prosecutor's and defendant's *ex ante* expected probabilities of trial does not exist (i.e., $\rho = \delta$).[32] With that assumption, the grounds for negotiation stem from the differences in the prosecutor's and the defendant's respective total trial expenditures, CX and KY.

With these simplifying assumptions, the prosecutor's *ex ante* expected value of a plea agreement, $V^B = S^B$, is a Nash solution that maximizes Ω, which is the product of the net gains the prosecutor expects to accrue to each party from a plea agreement:[33]

$$\begin{aligned} \Omega &= \left(V^B - V^T\right) \bullet \left[-V^B - U^T\right] \\ &= \left[S^B - \left(\Phi S^T - CX\right)\right] \bullet \left[-S^B - \left(-\Phi S^T - KY\right)\right] \end{aligned}, \qquad [12]$$

[32] As Eisenburg and Farber (2003) point out, this assumption implies that the prosecutor ignores any information there might be in the fact that the parties cannot negotiate a plea agreement and assumes the defendant expects the same probability of trial conviction as the prosecutor. In the alternative and without changing the basic character of the theoretical results of my model, it is reasonable to assume that the prosecutor is somehow informed by the fact that plea negotiations failed, or to assume that the prosecutor uses some other expected value of Γ based on $\Theta = \Psi(\cdot)$, the *ex ante* probability that the case will go to trial. These alternatives do not alter the basic character of the case selection model's predictions.

[33] Depending on the informational structure of the game, the parties may not know with certainty different case characteristics at various stages in games that could possibly underlie my selection model. In these situations, assume the parameters of prior distributions of case characteristics inform the parties' expectations.

Recall that $V^B = S^B$ is the prosecutor's *ex ante* expected value of a case that ends in a plea of guilty or *nolo contendere*, $V^T = \Phi S^T - CX$ is the prosecutor's *ex ante* expected value of a case that goes to trial, and $U^T = -\Gamma S^T - KY$ $(= -\Phi S^T - KY$, by assumption).

The first order condition from maximizing Ω with respect to V^B produces the prosecutor's expected value of a negotiated settlement, which is given by

$$\begin{aligned} V^B &= S^B \\ &= \Phi S^T + \frac{KY - CX}{2} \quad . \qquad [13] \\ &= \left(\hat{\Pi} + \rho\right)S^T + \frac{KY - CX}{2} \end{aligned}$$

It is interesting to note that the prosecutor's expected value of a negotiated settlement, V^B, decreases only with the prosecutor's expected unit cost of trial expenditure, C, and her expected level of trial expenditure, X. Meanwhile, V^B increases with every other variable on the right side of Equation [13].

Putting the separate components together, the prosecutor's *ex ante* expected value of indicting the (potential) defendant, $E(V)$, can be further expanded as follows:

$$\begin{aligned} E(V) &= \Theta V^T + (1 - \Theta)V^B \\ &= \Theta\left(\Phi S^T - CX\right) + (1 - \Theta)\left(\Phi S^T + \frac{KY - CX}{2}\right) \quad , \qquad [14] \\ &= \left(\hat{\Pi} + \rho\right)S^T + \frac{KY - CX}{2} - \Theta\frac{KY + CX}{2} \end{aligned}$$

where

$$V^T = \Phi S^T - CX \;, \tag{4}$$

$$\Theta = \Psi\left(\rho - \frac{CX + KY}{S^T}\right), \text{ and} \tag{11}$$

$$V^B = S^B = \Phi S^T + \frac{KY - CX}{2} = \left(\hat{\Pi} + \rho\right) S^T + \frac{KY - CX}{2}, \tag{13}$$

The litigation model developed in the next chapter specifies how changes in the values of case characteristics affect the prosecutor's *ex ante* expected values of X and Y. The litigation model also shows how changes in the values of X and Y affect the litigants' (from the prosecutor's perspective) common *ex ante* expected (Nash equilibrium) probability of trial conviction, $\hat{\Pi}$. Thus, given that $\hat{\Pi}$ affects $E(V)$, as shown in Equation [14], the litigation model helps to specify how changes in the values of X and Y (that are created by changes in the values of case characteristics) affect $E(V)$ through $\hat{\Pi}$. Understanding these dynamics is important for understanding how changes in the values of case characteristics ultimately affect $E(V)$.

As shown in Equations [11] and [14], changes in the values of X and Y also affect $E(V)$ through the prosecutor's *ex ante* expected probability of trial, Θ. In addition, inspection of Equations [4] and [14] shows that changes in the value of X affects $E(V)$ through the prosecutor's *ex ante* expected payoff from trial, V^T; and, as shown by Equations [13] and [14], through the prosecutor's *ex ante* expected payoff from a plea agreement, V^B. Equations [13] and [14] also show that changes in the value of Y affect $E(V)$ through V^B.

CHAPTER V

LITIGATION MODEL

This chapter develops a (sub-game) model of (trial) litigation with an endogenous trial sentence that will be used to predict the effects of changes in the values of certain variables on the *ex ante* expected values of the prosecutor's and defendant's (Nash equilibrium) optimal levels of trial expenditure, X and Y, respectively. The directions of the predicted changes in the values of X and Y inform the predicted subsequent (net) direct (first- and second-order) effects of changes in the values of X and Y on the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, the parties' common *ex ante* expected trial sentence, S^T, as well as the *ex ante* probabilty of trial, Θ. In later chapters, these predictions determine the expected directions of the effects of variations in the values of case characteristics on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, as well as the *ex post* (observed) probability of trial, the *ex post* probability of trial conviction, and the *ex post* expected trial sentence. The litigation model thus expands my case selection model's predictions regarding the effects on $E(V)$ of changing values of case characteristics.

The Litigants' Common *Ex Ante* Expected Probability of Trial Conviction and *Ex Ante* Expected Value of the Penalty Imposed upon Trial Conviction

As previously mentioned, in the pre-indictment stage, the prosecutor knows her endowment of evidence, e, which determines the merits of the case and includes the type(s) of alleged violation(s), the duration of the violation(s), the number of firms involved in the charge(s), the offender's previous antitrust convictions, the defendant's culpability and the economic impact associated with the charges, such as the dollar value of affected commerce. The prosecutor's evidence, e, also could include other information, such as the year and region of the country in which the alleged offense(s) happened, or other industry-specific information. Assume that the prosecutor's evidence does not change after the pre-indictment stage and that the prosecutor never drops a case after indictment.

Although the litigation model (and, hence, my case selection model) does not include an explicit discovery process, the pre-indictment stage in my models could serve discovery purposes. Nonetheless, this study assumes that a defendant who has exculpatory information cannot establish it during the pre-indictment stage.[34] If the defendant could, Shavell (1989) shows that failure to do so would itself signal guilt. Based on the federal rules of disclosure, it is reasonable to assume that all of the prosecutor's evidence is provided in the indictment.

At the trial stage, the prosecutor and defendant choose their levels of effort at trial, X and Y, respectively. It is useful to note that X and Y are the only choice variables in my model. Meanwhile, $P(X)$ and $D(Y)$ are scalars denoting the persuasive value of testimony and legal arguments regarding the defendant's culpability and liability made at trial by the prosecutor and defendant, respectively.

Assume that the probability of trial conviction, Π, (which is the same as the litigants' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, which was discussed in the previous chapter) increases in e and $P(X)$ and decreases in $D(Y)$. Further, $P_X > 0$, $D_Y > 0$, $P_{XX} < 0$, and $D_{YY} < 0$. Also assume that $P(X)$ and $D(Y)$ are independent so that $P_Y = D_X = P_{XY} = D_{YX} = 0$. That is, a litigant's ability to form her arguments, given the charges and evidence, does not depend on her opponent's arguments. Nevertheless, by construction, any effect $P(X)$ has on Π may be offset by $D(Y)$, and *vice versa*. Thus, the persuasive value of a litigant's arguments, but not the litigant's ability to present her case, depend on the other litigant's arguments.

It is plausible that the federal prosecutor's trial costs are, on average, lower than those of a private defendant. As Eisenberg and Farber (2003) point out and completely rely upon for interpreting the results of their empirical tests, the federal government (or, for my purposes, the Antitrust Division of the U.S. Department of Justice in particular), unlike many private defendants, has a standing corps of attorneys willing and able to litigate cases. The Department of Justice need not retain counsel on an hourly basis to litigate a case, although this does occur. Because it is reasonable to assume that unit costs from the parties' efforts across plea outcomes are such that the prosecutor's unit costs are lower than the defendant's, assume $C < K$ for all X

[34] This assumption is also consistent with assuming a Nash equilibrium for the plea-bargaining game underlying my model.

and Y.[35] For simplicity in developing the litigation model, however, the cost per unit of the defendant's trial expenditure, K, is normalized to one, so that C reflects the prosecutor's unit cost of trial expenditure relative to (as a fraction of) the defendant's.

At trial, the parties expect the trial court (i.e., jury and/or judge) to be presented with e, $P(X)$ and $D(Y)$, which the trial court uses to convict or exonerate the defendant. The outcome of a trial is also affected by several case-specific random factors, however. These factors include the idiosyncratic biases of judges and juries, as well as random elements from the argument production functions and the burden of proof function. For example, as argued by Snyder (1989, 1990) as well as Cohen and Scheffman (1989), the discretionary ability of judges to adjust the conviction rule or standard according to their preferences, even for *per se* violations, is an important determinant of the defendant's probability of conviction. These random factors associated with the judge's conviction standard are represented by v, which is distributed according to the cumulative density function $\Pi(v)$ and the probability density function $\pi(v)$. The litigants' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, which was previously introduced, is analogous to the cumulative density function $\Pi(v)$.

The distribution of v is assumed to be commonly known. Several assumptions characterize the probability density function $\pi(v)$. First, assume that all deviations have some positive probability. Second, assume that the mean of v is zero. Third, assume that $\pi(v)$ has a single mode and that its mode is equal to its median. Thus, larger deviations from the median are less common than smaller deviations.

Whatever a judge's motives,[36] the leeway granted judges in making discretionary decisions at trial is unavoidably limited by institutional constraints of the legal system. For

---

[35] Posner finds that between 1956 and 1967, for example, the DOJ Antitrust Division's appropriations more than doubled, yet the number of Antitrust Division personnel rose by less than one-third. This increase in personnel, albeit less than proportional to the budget increase, was material. Posner does not find a corresponding significant increase in the number of cases brought.

According to Posner (1970), these findings support his hypothesis that the price of resources used for antitrust enforcement could increase faster than prices in general. In addition, these findings could suggest a high (private sector) market demand for antitrust attorneys, whose fees increase faster than inflation. If the average salary for a private antitrust attorney is higher than the average salary for a government prosecutor, then it is reasonable to assume that the defendant's unit cost of litigation efforts is relatively higher than the prosecutor's unit cost.

[36] The economics literature contains several papers postulating a variety of factors that motivate judicial behavior. According to Posner (1993), federal judges maximize "the same thing everybody else does," their utility, which depends on both pecuniary and non-pecuniary benefits like leisure, prestige, and power. Others have more narrowly focused on potential influences on judicial behavior, including a judge's political ideology (Ashenfelter,

example, important procedural or evidentiary decisions of district court judges may be overturned by higher courts on appeal. Thus, *v* is multiplied by a positive spread factor, λ, which is independent of *v*. The spread factor λ represents the parties common expectation regarding the latitude the legal system allows judges in making discretionary decisions at trial.

The parties assume the judicial conviction rule is to find the defendant guilty if and only if $[\mathbf{e} + P(X) - D(Y)] > \lambda v$. Thus, the common component of the parties' *ex ante* expected probabilities of trial conviction is defined as

$$\hat{\Pi} = \Pi([\mathbf{e} + P(X) - D(Y)]/\lambda). \qquad [15]$$

Next, assume that the judge determines sentences simultaneously with conviction by plea or by trial. That is, there is no separate sentencing stage in the case disposition process (although the conviction path could influence sentencing decisions, as discussed later).

In the event that the defendant chooses to go to trial and is convicted at trial, from the prosecutor's perspective, the defendant receives a sentence represented by S^T. As described previously, S^T is comprised of the applicable statutory minimum sentence, *m*, where $m \geq 0$, added to some expected fraction, μ^T, of the difference between the applicable statutory maximum sentence, *M*, and *m*. In addition, assume $0 \leq \mu^T \leq 1$.

Because the indictment and the parties' testimony and legal arguments at trial regard the defendant's culpability and penalty exposure, they are determinants of μ^T. That is, the parties expect μ^T to increase as evidence, e, and the prosecutor's trial arguments, *P*(*X*), increase. Conversely, the parties expect μ^T to decrease in *D*(*Y*). Further assume that the prosecutor always argues for higher penalties than the defendant. The specific effects of e, *P*(*X*), and *D*(*Y*) on μ^T may depend on the parties' common expectation of judicial objectives or motives in

---

Eisenberg and Schwab, 1995) or his reputation among peers or litigants, (Higgins and Rubin, 1980; Miceli and Cosgel, 1994; Rasmusen, 1994).

Snyder (1990) posits that judges' standards for conviction may depend on the existing statutory sentencing parameters. Thus, depending on the discretionary behavior of judges, many of the factors that affect the expected trial penalty may also affect the probability of conviction. Snyder's thesis is important as a general proposition because it recognizes the mutual dependence among factors affecting the probability of trial conviction and factors that influence the sentencing decision.

sentencing.[37] The parties' common *ex ante* expected sentence following a trial conviction can be expressed as follows

$$S^T = m + \mu^T(\mathbf{e}, P(X), D(Y))(M - m). \qquad [16]$$

Nash Equilibrium Definitions of the Parties' Optimal Levels of Trial Expenditure

Assume that the outcome of the litigation game is a Nash equilibrium in the prosecutor's and defendant's trial expenditures, *X* and *Y*, respectively. A Nash equilibrium implies that the parties' choices of *X* and *Y* are optimal responses taking the opponent's (expected) choices as given. As Katz points out, the Nash equilibrium does not require communiction among the parties and judge. Assume also that agency problems involving attorneys do not exist.

The optimal levels of *X* and *Y* are solutions to a maximization problem. The prosecutor and defendant solve the following simultaneous maximization problems:

$$\underset{X}{Max} \quad V^T = \Pi(\mathbf{e}, X, Y)S^T(\mathbf{e}, X, Y, M, m) - CX \text{ , and} \qquad [17]$$

$$\underset{Y}{Max} \quad U^T = -\Pi(\mathbf{e}, X, Y)S^T(\mathbf{e}, X, Y, M, m) - Y. \qquad [18]$$

Differentiating V^T and U^T with respect to *X* and *Y*, respectively, produces the first order conditions for an optimum. The first-order conditions are:

$$V_X^T = \Pi_X S^T + \Pi S_X^T - C = 0 \text{, and} \qquad [19]$$

[37] Empirical evidence from antitrust enforcement suggests that the number, and possibly the severity of charges of conviction affect sentencing. For example, Gallo, *et al*. (1994) suggest that the increases in the severity of sentences imposed since the late 1970s reflect, in part, the prosecutorial strategy of bundling charges of antitrust violations with charges of non-antitrust violations, which became increasingly popular around 1977. They find that since 1955, the average real fine net of suspensions was about three times higher for both individuals and corporate defendants convicted of bundled offenses relative to those convicted of price-fixing violations only.

A similar finding holds for prison sentences since 1955. The average prison term net of suspensions imposed on individuals convicted of pure antitrust offenses was just over two months compared to over four months for those convicted of both antitrust and non-antitrust offenses. Likewise, Cohen finds that cases involving more counts of conviction bring statistically significantly higher fines for corporate defendants and longer prison sentences for individual defendants.

$$U_Y^T = -\left(\Pi_Y S^T + \Pi S_Y^T + 1\right) = 0\,, \qquad [20]$$

where $\Pi_X = \pi P_X$ and $\Pi_Y = \pi \ \ D_Y$.

The marginal value of trial expenditure (for example, $\pi P_X S^T + \Pi S_X^T - C$ for the prosecutor) is proportional to the probability that an additional argument will be decisive in the trial conviction decision, where "decisiveness" relates to the probability that $\mathbf{e} + P(X) - D(Y)$ exceeds λv. The probability that additional trial expenditure (and the associated additional argument at trial) will be decisive is given by the probability density function $\pi(\mathbf{e}, X, Y)$.

Recall the assumptions regarding the characteristics of $\Pi(\mathbf{e}, X, Y)$ and $\pi(\mathbf{e}, X, Y)$. When the prosecutor includes an additional charge in the indictment or when a party makes an additional argument at trial, the "conviction threshold value" of the trial conviction rule, $\mathbf{e} + P(X) - D(Y)$, moves relative to λv. This movement in the threshold value changes the probability of trial conviction, Π, which affects π and thus the marginal value of trial expenditure. Threshold value movement in close cases offsets small[38] random influences on the probability of trial conviction; in cases that are less close, movement in the threshold value offsets the effects of larger random shocks.

---

[38] The random shocks would have to be small or else the case would not be close.

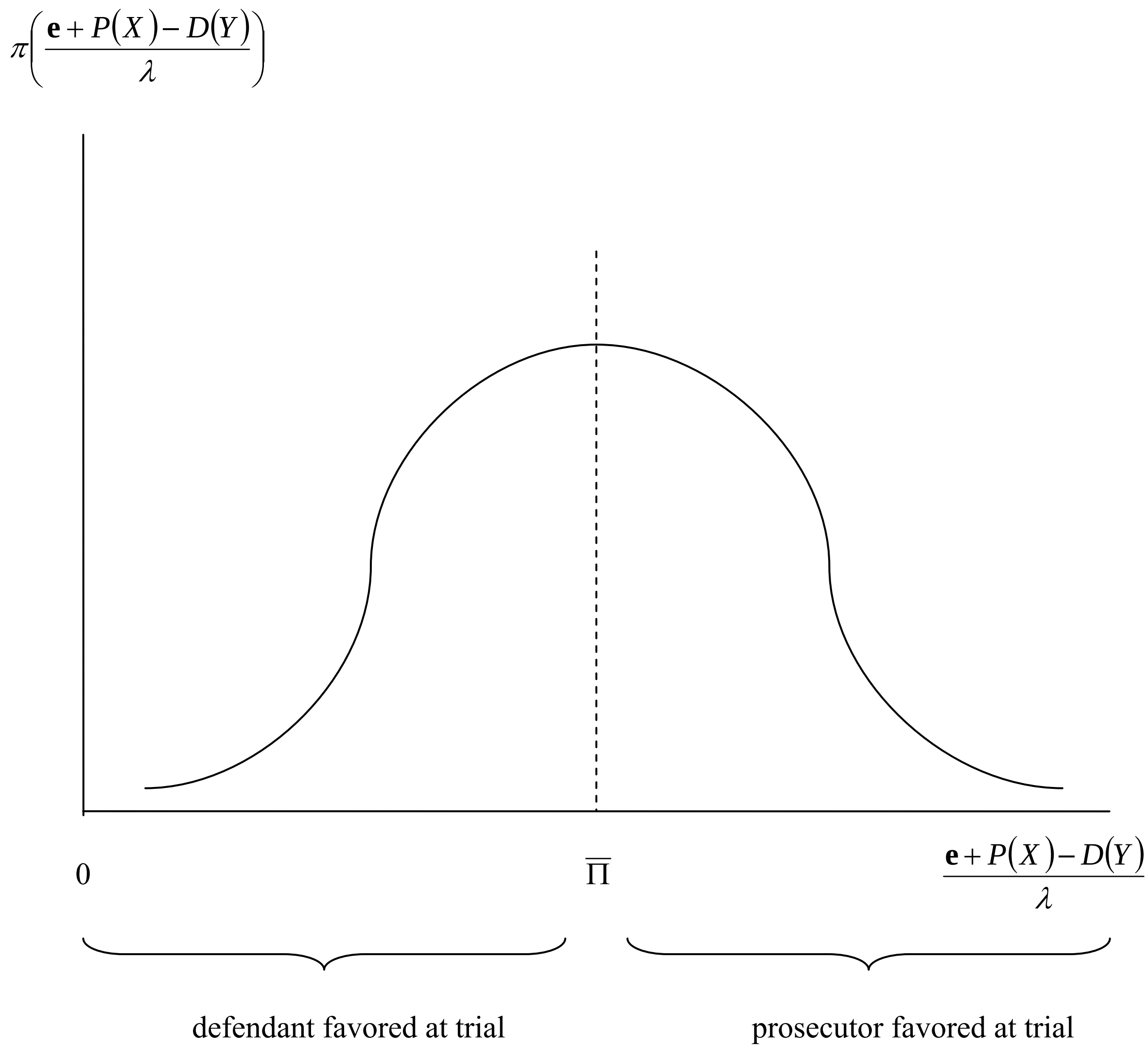

Figure 1. A graphical representation of the probability distribution function π

Because the density, π, is higher around the median of the distribution, the marginal value of trial effort rises as the probability of trial conviction approaches the median trial conviction threshold and the parties are more evenly matched as trial opponents. Assuming smaller random disturbances at trial are more likely than larger ones (i.e., small probability of "*Perry Mason* endings" at trial), additional arguments in the indictment or at trial have greater influences, positive and negative, on the probability of trial conviction in closer cases.

The predicted effects of additional trial arguments depend on three factors: 1) the pre-indictment value of the trial conviction threshold relative to the median trial conviction threshold value; 2) the pre-indictment level of π relative to the median value of π; and 3) the predicted directions of the effects of additional arguments on the trial conviction threshold value (or the probability of trial conviction). These are the central insights underlying the predictions of the litigation model.

Like Katz (1988), this study denotes the "favorite" at trial as the party with a chance of winning that exceeds some threshold value like 50 percent, where 50 percent is chosen as an intuitively appealing example.[39] Meanwhile, the "underdog" is the party with less than, for example, a 50 percent chance of winning. It is not necessarily the case that the favored party has better trial skills. Assume the favorite/underdog distinction is commonly known before the trial commences.

The assumption that the prosecutor is the "favorite" has implications regarding trial conviction rates, trial rates and the prosecutors' case-bringing decisions. Specifically, as the favorite makes additional arguments (by increasing his level of expenditures at trial), the conviction threshold value moves away (positively or negatively) from the mode of π, the probability density function of Π. Such movement along the density function π decreases the favorite's marginal value of trial expenditure (in terms of the probability of trial conviction when ignoring sentencing effects). In contrast, arguments made by the underdog move the trial conviction threshold value toward the mode of π, which raises the marginal value of trial expenditure for the underdog.

The second-order conditions for an optimum in trial expenditure are:

$$V_{XX}^{T} = \Pi_{XX} S^{T} + 2\Pi_{X} S_{X}^{T} + \Pi S_{XX}^{T} < 0 \text{, and} \qquad [21]$$

$$-U_{YY}^{T} = -\left(\Pi_{YY} S^{T} + 2\Pi_{Y} S_{Y}^{T} + \Pi S_{YY}^{T}\right) < 0 \text{.} \qquad [22]$$

The determinants of the parties' choices of trial expenditures, *X* and *Y*, are implicitly specified by the first- and second-order conditions for an optimum. Further assume that all first-

[39] The intuitively appealing assumption that the mean of $\hat{\Pi}$ is 50 percent is consistent with assuming that *v* is distributed according to the logistic distribution function, for example. It is important to note that this "assumption" is more of a normalization. A nonzero mean for *v* would imply a mean of $\hat{\Pi}$ that is different from 50 percent.

and second-order necessary conditions are satisfied to ensure that each party's optimal trial expenditure varies continuously with the opponent's optimal trial expenditure.

Under these conditions, it is convenient to refer to the implied function that describes the relationship between a party's optimal trial expenditure and his opponent's optimal trial expenditure as the party's "trial expenditure reaction function." The implied prosecutor's trial expenditure reaction function is $X(Y)$, and the implied defendant's trial expenditure reaction function is $Y(X)$. The prosecutor's trial expenditure reaction function is implied by Equation [19]; similarly, the defendant's trial expenditure reaction function is implied by Equation [20]. The second-order conditions assist in signing first derivatives in the analysis below.

The slope of the prosecutor's trial expenditure reaction function is derived by totally differentiating Equation [19] with respect to X and Y and is defined as $dY/dX = -\left(V_{XX}^{T}\right)/V_{XY}^{T}$. The slope of the defendant's trial expenditure reaction function is found by totally differentiating Equation [20] with respect to X and Y and is defined as $dY/dX = -U_{YX}^{T}/U_{YY}^{T}$. From the second-order optimization conditions, $V_{XX}^{T} < 0$ and $-U_{YY}^{T} < 0$ (or $U_{YY}^{T} > 0$). At a Nash equilibrium under these conditions and provided both parties' trial expenditures affect the trial outcome and have diminishing returns, the slopes of the two trial expenditure reaction functions must be of opposite signs when they intersect. Further, stability of the Nash equilibrium requires the slope of the prosecutor's trial expenditure reaction function to be steeper than the slope of the defendant's trial expenditure reaction function when compared as oriented on common axes.

Differently signed slopes for the two trial expenditure reaction functions (when using the same axes to graph the reaction functions) implies only one party's additional trial expenditure leads to an increase in the opponent's trial expenditure. Differently signed slopes also implies that an increase in only one party's trial expenditure leads to a decrease in the other party's trial expenditure. In the former case, following the nomenclature of Katz, a marginal increase in a party's trial expenditure is "provocative" if it leads to a marginal increase in the other party's trial expenditure; conversely, in the latter case, a marginal increase in a party's trial expenditure is "deterring" if it leads to a marginal decrease in the other party's trial expenditure.[40]

The sign of each party's reaction function can be determined by expanding my model's specification of the slope of the reaction functions and making use of the favorite/underdog

[40] The provocative/deterring distinction is analogous to the distinction between a "strategic complement" and a "strategic substitute" made by Bulow *et al*. (1985).

assumption. For example, the slope of the defendant's trial expenditure reaction function can be expanded as follows:

$$\frac{dY}{dX} = \frac{U_{YX}^{T}}{-U_{YY}^{T}} = \frac{\left(\Pi_{YX} S^{T} + \Pi_{Y} S_{X}^{T} + \Pi_{X} S_{Y}^{T} + \Pi S_{YX}^{T}\right)}{\Pi_{YY} S^{T} + 2\Pi_{Y} S_{Y}^{T} + \Pi \ S_{YY}^{T}}. \qquad [23]$$

It is reasonable to assume that S_{YX}^{T} is negative like the second and third terms of the numerator of the right-hand side of equation [23]. For convenience, assume that $\Pi_{YX} S^{T}$ is the dominant (largest absolute value) term in the numerator. Also, according to Equation [22], the denominator of Equation [23] is positive.

Because *D*(*Y*) is independent of *X* and $D_{YX} = 0$, the sign of the slope of the defendant's trial expenditure reaction function is $\operatorname{sgn}\left[\Pi_{YX}\right] = \operatorname{sgn}\left[\pi' D_{Y} P_{X}\right] = \operatorname{sgn}\left[\pi'\right]$. Making use of the shape of the probability density function π as it relates to values of $\Pi = \overline{\Pi}$, the slope, π', of the probability distribution function is positive when the defendant is favored and negative when he is the underdog. Accordingly, if $\Pi_{YX} S^{T}$ is assumed to be the dominant (largest absolute value) term in the numerator of the defendant's trial expenditure reaction function, then the slope of that reaction function is positive when the defendant is the trial favorite and negative when he is the underdog.[41]

A key insight above is that trial expenditure by the favorite makes the case less close and trial expenditure by the underdog makes the case closer. Consideration of this insight and the (primary) determinant of the signs of the slopes of the trial expenditure reaction functions reveals another insight: trial expenditure by the favorite is deterring, and trial expenditure by the underdog is provocative.

Figure 2 graphically depicts the Nash equilibrium of the litigation game and the (necessary) characteristics of the trial expenditure reaction functions. Note in Figure 2 that the defendant is favored in all equilibria above the line defined by $\hat{\Pi} = \overline{\Pi}$ and the prosecutor is favored in all equilibria below. Also, $X(Y)$ is "steeper" or more positive than $Y(X)$ at a Nash equilibrium, (X^{*}, Y^{*}).

[41] Without this simplifying assumption, the slope of the defendant's trial expenditure reaction function is less negative when the defendant is the trial favorite and more negative when he is the underdog.

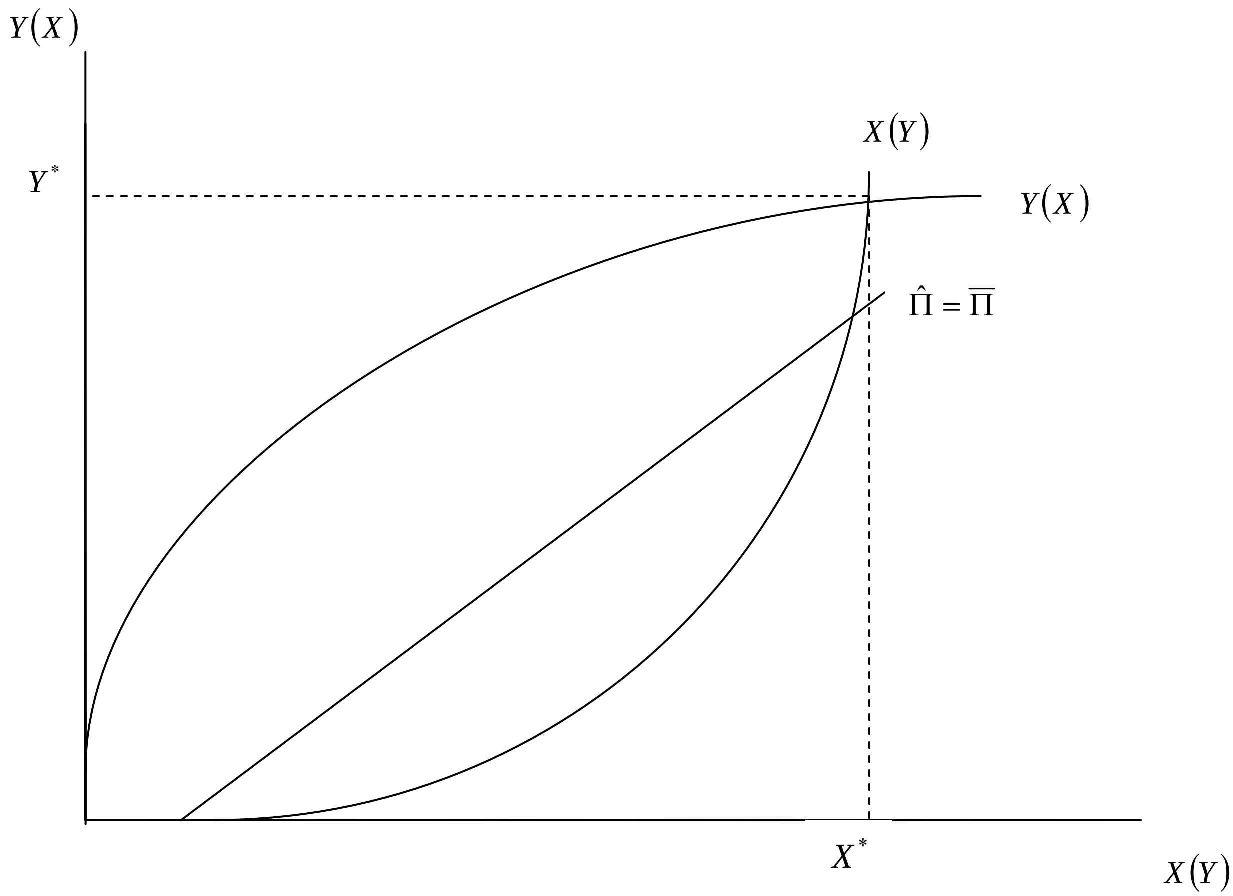

Figure 2. Trial expenditure reaction functions, $X(Y)$ and $Y(X)$

It is important to note that along a party's trial expenditure reaction function, the party's chances at trial always worsen as the opponent spends more. Along the defendant's trial expenditure reaction function, for example,

$$\frac{d\Pi(Y^*)}{dX} = \Pi_X + \Pi_Y \frac{dY^*}{dX}. \qquad [24]$$

Equation [24] shows the direct effect of the prosecutor's trial expenditure is to decrease the defendant's chances at trial ($\Pi_X > 0$). If the prosecutor is the favorite, then her trial expenditure is deterring ($dY^*/dX < 0$) and the direct and indirect effects have the same signs. Then the prosecutor's trial expenditure obviously worsens the defendant's chances at trial and Equation [24] is positive.

In contrast, if the prosecutor is the underdog, then, at the margin, her trial expenditure is provocative ($dY^*/dX > 0$). In this case, the indirect effect of a marginal increase in *X* is negative and mitigates the positive direct effect of an increase in *X*. The indirect effect must be weaker than the direct effect, however. If the indirect effect could overtake the direct effect, the marginal value of trial expenditure would not rise, which implies that there should have been no increase in *X* in the first place. In my model, therefore, the indirect effect of the parties' adjustments (through their trial expenditure) can alter the magnitude but not the direction of the total effect. Thus, Equation [24] is positive regardless of whether the prosecutor or the defendant is the trial favorite.

Equation [24] is positive regardless of whether the prosecutor is the favorite or underdog at trial. Thus, my model generally predicts that trial expenditure by either party, at the margin, improves that party's chances at trial.

The effects of small changes in case characteristics are found by first totally differentiating the first-order maximization conditions as follows:

$$\begin{bmatrix} V_{XX}^T & V_{XY}^T \\ -U_{YX}^T & -U_{YY}^T \end{bmatrix} \begin{bmatrix} dX \\ dY \end{bmatrix} = [L] \begin{bmatrix} dX \\ dY \end{bmatrix}. \qquad [25]$$

Cramer's Rule can be employed to show in mathematical terms the effects of changes in the values of case characteristics on the parties' Nash equilibrium choices of trial expenditure.

Assuming a locally strictly stable equilibrium implies that the following conditions hold: $-\left(V_{XX}^T U_{YY}^T\right) > -\left(V_{XY}^T U_{YX}^T\right)$ and $[L] < 0$. This stable equilibrium assumption also implies the previously mentioned stability requirement – that the prosecutor's trial expenditure reaction function be "steeper" than the defendant's. In turn, these conditions allow me to specify the litigation model's predictions regarding the effects of changes in the values of case characteristics on the parties' choices of *X* and *Y*. Those predictions can then be used for other predictions regarding the marginal effects of changes in case characteristics on Nash equilibrium trial outcomes.

Table 2 reports the litigation model's predicted signs of the total marginal effects on the values of 1) the prosecutor's optimal level of trial expenditure, X^*, 2) the defendant's optimal level of trial expenditure, Y^*, 3) the litigants' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$ (evaluated at $X = X^*$ and $Y = Y^*$), and 4) the litigants' common *ex ante* expected penalty following a trial conviction, S^T (evaluated at $X = X^*$ and $Y = Y^*$), from marginal changes in the values of i) the prosecutor's level of trial expenditure, X, ii) the defendant's level of trial expenditure, Y, iii) the strength and amount of the prosecutor's evidence, $\mathbf{e}$, iv) the unit cost of trial expenditure for the prosecutor, C, v) the unit cost of trial expenditure for the defendant, K, vi) the litigants' common *ex ante* expected penalty following a trial conviction, S^T, and vii) the litigants' common expected variability of the judge's trial conviction standard, λ.

In other words, Table 2 summarizes the litigation model's predictions (some of which are not discussed above) regarding the effects of changes in the values of certain variables on my litigation model's Nash equilibrium expected values of the prosecutor's and defendant's optimal levels of trial expenditure, *X* and *Y*, respectively. The directions of the predicted changes in the values of *X* and *Y* inform the predicted subsequent (net) direct (first- and second-order) effects of changes in the values of *X* and *Y* on the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, the parties' common *ex ante* expected trial sentence, S^T, as well as the *ex ante* probabilty of trial, Θ. In later chapters, these predictions determine the expected directions of the effects of variations in the values of case characteristics on the prosecutor's *ex ante* expected value of bringing a case, *E*(*V*), as well as the *ex post* (observed) probability of trial, the *ex post* probability of trial conviction, and the *ex post* expected trial sentence.

Table 2. The litigation model's predicted signs of the total marginal effects on a) the litigants' optimal levels of trial expenditure, X^* and Y^*; b) the litigants' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$; and c) the litigants' common *ex ante* expected penalty following a trial conviction, S^T, from marginal changes in the values of case characteristics

Mathematical Expression(s)	Litigation Model Prediction(s)
$\frac{dX^*}{dY}, \frac{dY^*}{dX}$	Trial expenditure by the favorite is deterring (i.e., reduces trial expenditure by the opponent), and trial expenditure by the underdog is provocative (i.e., increases trial expenditure by the opponent).
$\frac{d\hat{\Pi}}{dX}, \frac{d\hat{\Pi}}{dY}$	Trial expenditure by either party improves that party's expected chances at trial.
$\frac{dS^T}{dX}, \frac{dS^T}{dY}$	Trial expenditure by either party improves that party's value of the sentencing decision.
$\frac{dX^*}{d\mathbf{e}}, \frac{dY^*}{d\mathbf{e}}$	Changes in the merits of the case through increases in the amount or strength of evidence that benefit the underdog will cause both parties to increase trial expenditure. Conversely, changes in the amount or strength of evidence that benefit the favorite will cause both parties to reduce trial expenditure.
$\frac{d\hat{\Pi}}{d\mathbf{e}}$	The party that benefits from a change in e will have better chances at trial.
$\frac{dS^T}{d\mathbf{e}}$	Changes in the value of e have ambiguous effects on the *ex ante* expected trial sentence. That is because of the changes in *both* parties' trial expenditure in response to changes in the value of e.
$\frac{dX^*}{dC}, \frac{dY^*}{dC}, \frac{dX^*}{dK}, \frac{dY^*}{dK}$	An increase (a decrease) in a single party's cost per unit of trial expenditure, C or K, will lead that party to reduce (increase) trial expenditure. The opponent will reduce (increase) trial expenditure if and only if the opponent is the trial favorite.
$\frac{d\hat{\Pi}}{dC}, \frac{d\hat{\Pi}}{dK}$	An increase (a decrease) in a single party's trial costs will reduce (improve) that party's probability of winning at trial.
$\frac{dS^T}{dC}, \frac{dS^T}{dK}$	If the prosecutor (defendant) is the trial favorite, then an increase in C (K) will cause X (Y) to fall and Y (X) to rise, which will reduce (increase) S^T. Meanwhile, if the prosecutor (defendant) is the trial favorite, then an increase in K (C) will cause X and Y to fall, which has ambiguous (net) effects on S^T.

Table 2, continued

Mathematical Expression(s)	Litigation Model Prediction(s)
$\frac{dX^*}{dS^T}, \frac{dY^*}{dS^T}$	An (exogenous) increase in the trial stakes for both parties, S^T, or a decrease in both parties' trial costs will induce both parties to increase their trial expenditure. Meanwhile, a decrease in common trial stakes or an increase in the trial costs of both parties will lead them to reduce their trial expenditures.
$\frac{d\hat{\Pi}}{dS^T}$	The chances at trial of the party with relatively weaker diminishing returns in trial expenditure will improve.
$\frac{dX^*}{d\lambda}, \frac{dY^*}{d\lambda}$	An increase (a decrease) in the variability of the trial court's conviction decision, λ, will cause parties in relatively close cases to reduce (increase) trial expenditure, and will cause parties in relatively one-sided cases to increase (decrease) trial expenditure.
$\frac{d\hat{\Pi}}{d\lambda}$	As long as the parties' rates of diminishing returns from trial expenditure are sufficiently close, an increase (a decrease) in λ will increase (decrease) the underdog's chances at trial.
$\frac{dS^T}{d\lambda}$	Changes in the value of λ have ambiguous effects on the *ex ante* expected trial sentence, S^T. That is because of the theoretical specification of S^T and the changes in *both* parties' trial expenditure in response to changes in the value of λ.

CHAPTER VI

DETERMINANTS OF THE OBSERVED TRIAL RATE

This chapter defines the *ex post* probability of trial (i.e., observed trial rate), $\widetilde{\Theta}$. Among the determinants of observed trial rates in federal criminal antitrust cases, it is important to consider both the direct effects of changes in the values of case characteristics and the indirect effects created by the prosecutor's case selection process. The prosecutor's case selection process conditions the sample of initiated cases that proceed to plea bargaining, which influences the observed trial rate.

These considerations are necessary for meaningful empirical tests of the determinants of trial rates in federal criminal antitrust cases. A brief description of the prosecutor's case selection process specified in the case selection model is useful to describe generally the key insight provided by the case selection model. Assume the prosecutor is faced with a potential case that has a certain case characteristic with a relatively extreme value that has a positive (negative) effect on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$. With that "cushion" ("deficit") created by that extreme-valued case characteristic, the prosecutor would be willing to initiate the case even if (only if) other case characteristics – even (especially) with relatively extreme values – have negative (positive) effects on $E(V)$ as long as (so that) the condition $E(V) \geq 0$ holds.

Empirically analyzing the sample selection bias created by the prosecutor's case selection process is generally difficult, however, because of the lack of information regarding the potential cases that prosecutors do *not* pursue. Thus, for example, controlling for sample selection bias created by the prosecutors' case-bringing decisions in regression models of defendants' plea choices is not possible using the usual approaches (e.g., the approach introduced by Heckman).

Moreover, as explained in this section, the second-order direct effects referred to as "trial expenditure effects" and "stakes effects," as well as the indirect "case selection effects" on the *ex post* expected probability of trial from changes in the values of case characteristics generally render ambiguous all predicted net effects on the *ex post* expected probability of trial from changes in the values of case characteristics. Thus, the following theoretical analysis generally provides ambiguous predictions for estimated coefficients on empirical variables that represent case characteristics in regression models of defendants' plea choices.

Despite the ambiguous expected indirect case selection effects on the observed trial rate, the theoretical implications of the prosecutor's case selection process, as specified in the case selection model, are nonetheless empirically testable. That is because the theoretical model of the prosecutor's case-bringing decision implies particular expected mixes of case characteristics, given the relative value of at least one case characteristic, in cases that the prosecutor decides to initiate. These predictions suggest firm relationships among the relative values of different pairs of case characteristics in cases that are initiated (i.e., along the *same* game tree branch stemming from the case-bringing decision mode). These relationships can be tested meaningfully using sample means tests using the sample of initiated (observed) cases.

Direct Effects on the Observed Trial Rate

The conditional *ex post* probability of trial is defined as follows:

$$\begin{aligned}\tilde{\Theta} &= \Pr\left[(\rho-\delta)S^{T} - CX - KY > 0 \,\middle|\, E(V) \geq 0\right] \\ &= E\left(\Theta \middle| E(V) \geq 0\right)\end{aligned} \qquad [26]$$

As shown in Equation [26], the *ex post* expected probability of trial, $\tilde{\Theta}$, is a conditional expected value of the *ex ante* expected probability of trial, Θ. The weak inequality, $E(V) \geq 0$, which conditions the *ex post* probability of trial, $\tilde{\Theta}$, was introduced in Equation [1]. The condition $E(V) \geq 0$ simply states that the prosecutor's *ex ante* expected value of bringing the case was non-negative, which is why the prosecutor initiated the case.

Both $E(V)$ and Θ, which is a component of $E(V)$, are functions of several variables that have already been introduced. The litigation model provides the optimal (Nash equilibrium) values of the prosecutor's and the defendant's respective choices of levels of trial expenditure, so that all of the *ex ante* expected values are evaluated with $X = X^{*}$ and $Y = Y^{*}$. Both X^{*} and Y^{*}

are implicit functions of C, K, S^T, $\mathbf{e}$, λ, as well as each other as specified by the litigation model.

The exogenous variables C and K represent the prosecutor's and defendant's respective unit costs of trial expenditure. The prosecutor's evidence that forms the merits of the prosecutor's case is represented by the vector of exogenous variables, e. The exogenous variable λ represents the parties' common expected value of the variability of the judge's trial conviction decision standard.

Recall that the exogenous variables ρ and δ are the idiosyncratic elements of the prosecutor's and defendant's *ex ante* expected probability of trial conviction, $\Phi = \hat{\Pi} + \rho$ and $\Gamma = \hat{\Pi} + \delta$, respecively, where $\hat{\Pi}$ is an endogenous variable that represents the parties' common *ex ante* expected probability of trial conviction. For completeness, recall that $\hat{\Pi}$ is a component of $E(V)$ and is a function of X, Y, e, and λ. Specifically,

$$\hat{\Pi} = \Pi\left(\left[\mathbf{e} + P(X) - D(Y)\right]/\lambda\right). \qquad [15]$$

Recall further that the endogenous variable S^T is the parties' common *ex ante* expected trial sentence, which is a function of X, Y, e, as well as the (exogenous) statutory maximum and minimum penalties for Sherman Act violations, M and m, respectively. Specifically,

$$S^T = m + \mu^T\left(\mathbf{e}, P(X), D(Y)\right)(M - m). \qquad [16]$$

This section focuses on the effects of the case characteristics C, K, S^T, ρ, e, and λ on the first element of the right-hand side of Equation [26], the expected value of the *ex ante* expected probability of trial, $E(\Theta)$ (which equals Θ). A full discussion of the signs of the total marginal effects of the case characteristics (as well as X and Y) on Θ is found in Appendix A, as part of the examination of the effects of different factors on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, as discussed later.

It is important to recognize that because the *ex ante* probability of trial, Θ, is a function of endogenous variables it is necessary to examine the first- and second-order effects of factors affecting Θ. The second-order effects on Θ arise from the effects of changes in the values of these factors on the endogenous variables X, Y, and S^T. The second-order effects that involve *X* and *Y* are the "trial expenditure effects" on Θ and the second-order effects that involve S^T are the "stakes effects" on Θ. The expected signs of the the total (net) marginal effects of C, K, and S^T (but not ρ) on Θ are ambiguous because of the ambiguity created by the trial expenditure effects and the stakes effects.

For example, according to the litigation model, an increase in a single party's cost per unit of trial expenditure, C or *K*, will lead that party to reduce trial expenditure. The opponent will reduce trial expenditure if and only if the opponent is the trial favorite. It follows that if the prosecutor is the trial favorite (underdog), an increase in her unit trial cost, C, will reduce X^* and increase (reduce) Y^*. Thus, as discussed in Appendix A, the trial expenditure effect on $\widetilde{\Theta}$ through X^* and Y^* from an increase in C (K) is ambiguous when the prosecutor (defendant) is the trial favorite. In contrast, when the prosecutor (defendant) is the trial underdog, the trial expenditure effect on the *ex ante* probability of trial, Θ, through X^* and Y^* from an increase in C (K) is positive, which counters the negative expected direct effect on Θ from an increase in the value of C (K).

Meanwhile, an exogenous increase in the trial stakes for both parties, including the parties' common expected sentence following a trial conviction, S^T, will induce the prosecutor and defendant to increase their (Nash) optimal choices of levels of trial expenditures, X^* and Y^*, respectively. Thus, the trial expenditure effect on the *ex ante* probability of trial, Θ, through both X^* and Y^* from exogenous increases in S^T is negative, which counters the positive expected direct effect on Θ from an increase in the value of S^T.

It is also important to note that there are other exogenous variables that affect *X*, *Y*, and S^T that do not directly influence the value of Θ but nonetheless create purely second-order trial expenditure effects or stakes effects on Θ. For example, in this study's theoretical framework (i.e., in the case selection model and the litigation model embedded in the case selection model), the amount or strength of evidence, e, produced in the pre-indictment stage is taken by the prosecutor as *given* at the case selection stage in which she evaluates her *ex ante* expected value of bringing a case, $E(V)$. The evidence represents the merits of the case. For example, higher values of the sales or profits attributed to an alleged conspiracy or more counts of conspiracy in an indictment provide stronger evidence of the existence (and success) of a conspiracy than lower returns from an alleged conspiracy. In terms of the litigation model, higher conspiracy sales or more counts in the indictment imply higher values of e.

When evaluating the direct determinants of the *ex ante* probability of trial, Θ, (as part of the evaluation of the *ex post* probability of trial, $\widetilde{\Theta}$) it is important to recognize the influence of more evidence on Θ through the effects of evidence on the *ex ante* expected trial sentence, S^T, which are the "stakes" effects of evidence on Θ. (Evidence influences S^T directly and through X^* and Y^*, thus creating two types of stakes effects.) As discussed in Appendix A, the marginal direct effect of more or better evidence, e, on the *ex ante* probability of trial, Θ, is ambiguous because of the ambiguous signs and magnitudes of the trial expenditure effects and the stakes effects of evidence on Θ.

As explained in Appendix A, the litigation model provides predictions regarding the effects on the prosecutor's and the defendant's respective levels of trial expenditure, *X* and *Y*, from changes in the value of the parties' common expected value of the variability of the judge's trial conviction decision standard, λ. Those predicted changes in *X* and *Y* create the trial expenditure effects that determine the total marginal effect on the *ex ante* probability of trial, Θ. Those predictions, however, require the assumption that the litigants' respective diminishing returns from trial expenditure are "close." Without any reasonable justification for that assumption, the expected sign of the net trial expenditure effect on the *ex ante* probability of trial, Θ, from increases in λ is ambigous.

In (stark) contrast, the endogenous variables *X*, *Y*, and S^T are not functions of ρ. Thus, changes in the value of ρ do not create second-order trial expenditure effects or stakes effects and the expected sign of the marginal direct effect of ρ on Θ is unambiguously positive.

Indirect Case Selection Effects on the Observed Trial Rate

Specifying Equation [26] to include the condition $E(V) \geq 0$ in the definition of the (conditional) *ex post* probability of trial, $\widetilde{\Theta}$, highlights the theoretical and empirical need to view $\widetilde{\Theta}$ from the perspectives associated with direct effects as well as the indirect case selection effects of changes in the values of case characteristics. That is, with respect to the direct effects of changes in the values of case characteristics on $\tilde{\Theta}$, it is important and convenient to view $\tilde{\Theta}$ as the probability that a case that was initiated went to trial. With respect to the indirect selection effects of changes in the values of case characteristics on $\tilde{\Theta}$, however, $\tilde{\Theta}$ also should be viewed as the probability that a case that was going to trial was initiated by the prosecutor in the first place (and observed).

If the bias created by the case selection process is not taken into consideration, the direct effect of changes in case characteristics on the *ex post* probability of trial, $\tilde{\Theta}$, may be over- or under-stated – whether theoretically predicted or empirically estimated. That is, while Equation [26] can be used to predict the direct effects of changes in values of case characteristics on $\tilde{\Theta}$, these changes in the values of case characteristics also have indirect effects on $\tilde{\Theta}$ through the prosecutor's case selection process (i.e., initiate a case if $E(V) \geq 0$).

A brief description of a stylized version of the prosecutor's case selection process is useful to describe the indirect case selection effects from changes in the values of case characteristics. Assume the prosecutor is faced with a potential case that has a certain case characteristic with a relatively extreme value that has a positive (negative) effect on $E(V)$. With that "cushion" ("deficit") created by that extreme-valued case characteristic, the prosecutor would be willing to initiate the case even if (only if) other case characteristics – even (especially) with relatively extreme values – have negative (positive) effects on $E(V)$ as long as (so that) the condition $E(V) \geq 0$ holds.

Appendix A provides the mathematical analysis of the expected total marginal effects of changes in the values of particular case characteristics on the prosecutor's *ex ante* expected value of initiating a case, $E(V)$. The case characteristics that are included in this analysis are the following: the parties' commonly expected trial sentence, S^T; the prosecutor's and defendant's unit costs of trial expenditure, C and K, respectively; the prosecutor's evidence, e; the variability of the judge's conviction decision standard, λ; as well as the idiosyncratic component, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, $\Phi = \hat{\Pi} + \rho$.

The idiosyncratic component, δ, of the defendant's *ex ante* expected probability of trial conviction, $\Gamma = \hat{\Pi} + \delta$, is not expected to create a case selection effect on the *ex post* probability of trial, $\tilde{\Theta}$. That is because of the innocuous simplifying assumption that when the prosecutor evaluates $E(V)$ at the case selection stage, she expects the defendant to adopt the prosecutor's *ex ante* expected probability of trial conviction, $\Phi = \hat{\Pi} + \rho$, in the defendant's expected payoff from going to trial. Intuitively, it is reasonable to expect that when the prosecutor evaluates $E(V)$ that she would rely on her idiosyncratic *ex ante* expected probability of trial conviction instead of her expectation of the defendant's idiosyncratic *ex ante* expected probability of trial conviction. (Moreover, it is changes in the difference between ρ and δ that are expected to affect the observed trial rate, so it is reasonable to focus on changes in the value of ρ, holding other factors constant.)

Now recall the definition of $E(V)$:

$$
\begin{aligned}
E(V) &= \Theta V^T + (1-\Theta)V^B \\
&= \Theta\left(\Phi S^T - CX\right) + (1-\Theta)\left(\Phi S^T + \frac{KY - CX}{2}\right) \quad, \qquad [14] \\
&= \left(\hat{\Pi} + \rho\right)S^T + \frac{KY - CX}{2} - \Theta\frac{KY + CX}{2}
\end{aligned}
$$

where

$$V^T = \Phi S^T - CX \;, \qquad [4]$$

$$V^B = S^B = \Phi S^T + \frac{KY - CX}{2} = \left(\hat{\Pi} + \rho\right)S^T + \frac{KY - CX}{2}, \qquad [13]$$

$$\Theta = \Psi\left(\rho - \frac{CX + KY}{S^T}\right), \qquad [11]$$

$$\hat{\Pi} = \Pi\left(\left[\mathbf{e} + P(X) - D(Y)\right]/\lambda\right).\text{and} \qquad [15]$$

$$S^T = m + \mu^T(\mathbf{e}, P(X), D(Y))(M - m) \qquad [16]$$

In addition, recall that the litigation model provides the optimal (Nash equilibrium) values of the prosecutor's and the defendant's respective choices of levels of trial expenditure, so that all of the *ex ante* expected values are evaluated with $X = X^*$ and $Y = Y^*$. Both X^* and Y^* are implicit functions of C, K, S^T, $\mathbf{e}$, λ, as well as each other as specified by the litigation model.

The following equations are discussed in Appendix A. They show the expected total marginal effects of increases in the values of case characteristics on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$.

$$\begin{aligned}\frac{dE(V)}{dC} &= \frac{\partial E(V)}{\partial C}\frac{dC}{dC} + \frac{\partial E(V)}{\partial X}\frac{dX^*}{dC} + \frac{\partial E(V)}{\partial Y}\frac{dY^*}{dC} + \frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{dC} + \frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{dC} + \frac{\partial E(V)}{\partial S^T}\frac{dS^T}{dC} \\ &= -\frac{(1+\Theta)X^*}{2} - \frac{[(1+\Theta)C]}{2}\frac{dX^*}{\partial C} + \frac{[(1-\Theta)K]}{2}\frac{dY^*}{dC} - \frac{(KY^* + CX^*)}{2}\frac{d\Theta}{dC} \\ &\quad + S^T\frac{d\hat{\Pi}}{dC} + (\hat{\Pi} + \rho)\frac{dS^T}{dC}\end{aligned} \qquad [30]$$

$$\begin{aligned}\frac{dE(V)}{dK} &= \frac{\partial E(V)}{\partial K}\frac{dK}{dK} + \frac{\partial E(V)}{\partial X}\frac{dX^*}{dK} + \frac{\partial E(V)}{\partial Y}\frac{dY^*}{dK} + \frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{dK} + \frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{dK} + \frac{\partial E(V)}{\partial S^T}\frac{dS^T}{dK} \\ &= \frac{(1-\Theta)Y^*}{2} - \frac{[(1+\Theta)C]}{2}\frac{dX^*}{dK} + \frac{[(1-\Theta)K]}{2}\frac{dY^*}{dK} - \frac{(KY^* + CX^*)}{2}\frac{d\Theta}{dK} \\ &\quad + S^T\frac{d\hat{\Pi}}{dK} + (\hat{\Pi} + \rho)\frac{dS^T}{dK}\end{aligned} \qquad [31]$$

$$\begin{aligned}\frac{dE(V)}{dS^T} &= \frac{\partial E(V)}{\partial S^T}\frac{dS^T}{dS^T} + \frac{\partial E(V)}{\partial X}\frac{dX^*}{dS^T} + \frac{\partial E(V)}{\partial Y}\frac{dY^*}{dS^T} + \frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{dS^T} + \frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{dS^T} \\ &= (\hat{\Pi} + \rho) - \frac{[(1+\Theta)C^T]}{2}\frac{dX^*}{dS^T} + \frac{[(1-\Theta)K]}{2}\frac{dY^*}{dS^T} - \frac{(KY^* + CX^*)}{2}\frac{d\Theta}{dS^T} + S^T\frac{d\hat{\Pi}}{dS^T}\end{aligned}. \qquad [32]$$

$$\frac{dE(V)}{d\mathbf{e}} = \frac{\partial E(V)}{\partial \mathbf{e}}\frac{d\,\mathbf{e}}{d\,\mathbf{e}} + \frac{\partial E(V)}{\partial X}\frac{dX^*}{d\,\mathbf{e}} + \frac{\partial E(V)}{\partial Y}\frac{dY^*}{d\,\mathbf{e}} + \frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{d\,\mathbf{e}} + \frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{d\,\mathbf{e}} + \frac{\partial E(V)}{\partial S^T}\frac{dS^T}{d\,\mathbf{e}}$$

$$= 0 - \frac{[(1+\Theta)C]}{2}\frac{dX^*}{d\mathbf{e}} + \frac{[(1-\Theta)K]}{2}\frac{dY^*}{d\mathbf{e}} + \frac{\left(KY^* + CX^*\right)}{2}\frac{d\Theta}{d\,\mathbf{e}} \qquad , \quad [33]$$

$$+ S^T\frac{d\hat{\Pi}}{d\mathbf{e}} + (\hat{\Pi} + \rho)\frac{dS^T}{d\mathbf{e}}$$

$$\frac{dE(V)}{d\lambda} = \frac{\partial E(V)}{\partial \lambda}\frac{d\,\lambda}{d\,\lambda} + \frac{\partial E(V)}{\partial X}\frac{dX^*}{d\,\lambda} + \frac{\partial E(V)}{\partial Y}\frac{dY^*}{d\,\lambda} + \frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{d\,\lambda} + \frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{d\,\lambda} + \frac{\partial E(V)}{\partial S^T}\frac{dS^T}{d\lambda}$$

$$= 0 - \frac{[(1+\Theta)C]}{2}\frac{dX^*}{d\lambda} + \frac{[(1-\Theta)K]}{2}\frac{dY^*}{d\lambda} + \frac{\left(KY^* + CX^*\right)}{2}\frac{d\Theta}{d\,\lambda} \qquad , \quad [34]$$

$$+ S^T\frac{d\hat{\Pi}}{d\lambda} + \left(\hat{\Pi} + \rho\right)\frac{dS^T}{d\lambda}$$

and

$$\frac{dE(V)}{d\,\rho} = \frac{\partial E(V)}{\partial \rho}\frac{d\,\rho}{d\,\rho} + \frac{\partial E(V)}{\partial X}\frac{dX^*}{d\,\rho} + \frac{\partial E(V)}{\partial Y}\frac{dY^*}{d\,\rho} + \frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{d\,\rho} + \frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{d\,\rho} + \frac{\partial E(V)}{\partial S^T}\frac{dS^T}{d\,\rho}$$

$$= S^T + 0 + 0 - \frac{\left(KY^* + CX^*\right)}{2}\frac{d\Theta}{d\,\rho} + 0 + 0 \qquad . \quad [35]$$

These equations feature several common elements. They all feature a direct effect on $E(V)$ and four potential indirect effects on $E(V)$ from increases in the values of the case characteristics. It is also important to note that this study's case selection model realistically allows the prosecutor to focus (i.e., "orient" herself, in the terminology introduced in Appendix A) on (or be movtivated by the outcomes of) particular aspects of case disposition when evaluating whether or not to initiate a potential federal criminal antitrust case. The particular aspects of case disposition on which the prosecutor may focus correspond to different endogenous elements of the mathematical definition of $E(V)$. These aspects are : i) her own expected trial effort or trial expenditure, $X = X^*$; ii) the defendant's expected trial effort or expenditure, $Y = Y^*$, iii) the *ex ante* probability of trial, Θ; iv) the *ex ante* probability of trial conviction, $\hat{\Pi}$; or v) the *ex ante* expected trial sentence, S^T. These aspects of case disposition are associated with the five types of predicted indirect (second-order) effects on $E(V)$ from increases in the values of case characteristics.

In particular, as discussed below and again in Appendix A, it is reasonable to give relatively more weight to the predicted marginal effect of $\frac{\partial E(V)}{\partial \hat{\Pi}} > 0$. This assumption also implies that the case selection model's predictions should place less weight on the predicted marginal effects of $\frac{\partial E(V)}{\partial X} < 0$, $\frac{\partial E(V)}{\partial Y} > 0$, $\frac{\partial E(V)}{\partial \Theta} < 0$, and $\frac{\partial E(V)}{\partial S^T} > 0$.

Studies by Posner (1970), Siegfried (1975), Weaver (1977) and Baker (1978) strongly suggest that government prosecutors are primarily motivated by (total) conviction rates, regardless of how convictions are achieved (i.e., either through trial or plea agreement). In particular, Posner (1970) is sharply critical of the DOJ Antitrust Division. According to Posner,

> [I]t would appear that both legal doctrine and the enforcement machinery are geared more to the apprehension of unsuccessful attempts to fix prices than to the apprehension of successful price fixing. In general, the fact of an agreement or conspiracy to fix prices is all that the government need prove in a price-fixing case and all that it attempts to prove.[42]

Like the study by Long *et al*. (1973) upon which his empirical analysis builds , Siegfried's empirical tests fail to find much evidence that measures of the social welfare benefits from federal antitrust enforcement play significant roles in explaining DOJ antitrust case-bringing activity. According to Siegfried,

> Perhaps this is not too surprising if we consider the reward structure confronting decision makers in the Antitrust Division. It is probably more important to win cases than to reduce economic losses or inequities in order to move up the success ladder in the Justice Department.

[42] Posner (1970), 410.

> Unfortunately, it has not been possible to identify a variable to measure the expected probability of winning various alternative cases.[43]

(Incidentally, this study seeks to find what Siegfried was seeking – a variable or variables that measure the expected probability of winning various alternative cases.)

Consistent with Posner's and Siegfried's empirical findings, Weaver's (1977) empirical analysis suggests that prosecutors are less likely to bring cases based on their economic significance and more likely to bring cases when the alleged activity is clearly illegal or when the cases otherwise can be won easily. In addition, former U.S. Assistant Attorney General for Antitrust, Donald I. Baker (1978) claims that if the government does not think it can meet the "proof beyond a reasonable doubt" standard for a suspected price-fixing offense, it simply does not bring the case. In a more recent study, Snyder (1989) reports that the Antitrust Division tends to secure more convictions in cases involving lower sales. If federal prosecutors are willing to target cases involving low sales, it is reasonable to suspect that they are primarily targeting cases that are easy to win.

Henceforth, prosecutors who are primarily motivated by (total) conviction rates are characterized as "conviction-motivated". The assumption that federal prosecutors are "conviction-motivated" in criminal antitrust cases is intuitively appealing and has solid and abundant support from empirical tests and other studies conducted by experts in the area of criminal antitrust enforcement – including practitioners such as a notable Seventh Circuit judge and a former U.S. Assistant Attorney General for Antitrust. It is a reasonable assumption.

Again, in terms of the case selection model (i.e., $E(V) \geq 0$), assuming the prosecutor is conviction-motivated implies that relatively more weight should be placed on the predicted marginal effect of $\frac{\partial E(V)}{\partial \hat{\Pi}} > 0$. This assumption also implies that the case selection model's predictions should place less weight on the predicted marginal effects of $\frac{\partial E(V)}{\partial X} < 0$, $\frac{\partial E(V)}{\partial Y} > 0$, $\frac{\partial E(V)}{\partial \Theta} < 0$, and $\frac{\partial E(V)}{\partial S^T} > 0$.

Although the assumption that the prosecutor is conviction-motivated involves the litigants' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, the assumption does not

43 Siegfried (1975), 573.

mean that the prosecutor is motivated only by the prospects of a trial conviction. Consider the structure of the equation $E(V) = \Theta V^T + (1 - \Theta)V^B$. To the extent that a marginal increase in (or a higher value of) $\hat{\Pi}$ in a potential case increases her *ex ante* expected value of bringing that case, *E*(*V*), the prosecutor is motivated by the prospects of a conviction, even if the case goes to trial, since $\frac{\partial V^T}{\partial \hat{\Pi}} > 0$. The prosecutor is also (just as) motivated to bring a case by the other (second-order) positive effects on *E*(*V*) that are created by a marginal increase in (or a higher value of) $\hat{\Pi}$ in cases that she considers pursuing in the pre-indictment stage. In particular, the prosecutor's decision to bring a case is also motivated by a marginal increase in (higher value of) $\hat{\Pi}$ because a marginal increase in $\hat{\Pi}$ creates a marginal increase in V^B, which, in turn, creates a marginal increase in *E*(*V*).

The predicted total (net) marginal effects on E(V) from increases in the values of case characteristics are summarized in Tables 21 through 26 at the end of Appendix A. These tables show the implications of different assumptions about the prosecutor's focus on different aspects of case disposition.

Two points of additional background will facilitate the following examination of the indirect case selection effects on observed case outcomes, including the observed trial rate. First, consider how the values of case characteristics are determined. Some case characteristics' values may be determined by external (institutional) shocks and some case characteristics' values may be inherently associated with certain types of cases or certain types of defendants, while the remaining case characteristics' values could be determined through the prosecutor's case selection process. The external shocks and the inherent differences across types of cases (i.e., bid-rigging versus price-fixing) or types of defendants provide opportunities for empirical tests of the implications of the case selection model (including the plea bargaining model and the litigation model).

Second, according to the specifications of the model of the prosecutor's case selection process, after the prosecutor has committed to prosecuting a case, neither the values of the case characteristics nor her (non-negative) *ex ante* expected value of bringing that case, $E(V)$ can change. These values are "frozen" once the prosecutor decides to initiate a case.

An example is useful for explaining the implications of this "frozen" case characteristic value concept. Suppose that an *initiated* case is observed to have a particular case characteristic with a relative value that is expected to increase $E(V)$ – for example, a relatively low value of the prosecutor's unit cost of trial expenditure, C. In this situation, *because that case was initiated*, we know that the relatively low value of C did not create sufficiently large negative second-order effects on $E(V)$ that would render $E(V)$ to be negative.[44] This concept of the "frozen" values of the case characteristics in the sample of cases observed after a particular decision-stage of the case disposition process is important, especially because this concept facilitates analysis in subsequent chapters of this study (i.e., Chapters VII , VIII, and IX).[45]

Case selection effects on observed trial rates from the amount or strength of the prosecutor's evidence

As previously mentioned, at the case selection stage, the prosecutor expects that the given amount and strength of her evidence, e, and other exogenous case characteristics (e.g., the prosecutor's unit cost of trial expenditure, C) will affect the prosecutor's and the defendant's behavior at trial. From the prosecutor's perspective at the case selection stage, the prosecutor's and defendant's expected behavior at trial, if the case should go to trial, is captured by their expected optimal choices of levels of trial expenditure, X^* and Y^*, respectively. At the case selection stage, the prosecutor expects that the litigants' optimal choices of trial expenditure will affect the litigants' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$. Working backwards in game theoretical fashion, the prosecutor's expectations of the parties' choices of

44 On the other hand, it is still possible that the relatively low value of *C* created positive second-order effects on *E*(*V*) that made *E*(*V*) higher than it would have been if the value of *C* were not relatively low. This possibility is of no consequence, however.

Considerations of the expected second-order effects of case characteristics on which *E*(*V*) does not directly depend (i.e., the prosecutor's evidence, **e**, the idiosyncratic element, ρ , of the prosecutor's *ex ante* expected probability of trial, and the variability of the judge's conviction decision standard, λ) are important for evaluating whether relatively low or high values of these case characteristics are likely to produce relatively high or low values of *E*(*V*). As explained later, understanding these relationships between the relative values of case characteristics and *E*(*V*) is an integral part of the analysis of expected indirect case selection effects on the *ex post* probability of trial, the *ex post* probability of trial conviction, as well as expected sentences following conviction by trial or plea agreement.

45 The concept of "frozen" values of the case characteristics is not as important in the analysis of indirect case selection effects in the rest of this section, because it is not possible for the second-order direct effects on *E*(*V*) from changes in the values of case characteristics to confound the predicted first-order effects when the prosecutor is assumed to be "conviction-motivated."

trial expenditure will affect her *ex ante* expected probability of trial, Θ, and, finally, the prosecutor's *ex ante* expected value of bringing the case, $E(V)$. Thus, because the given level (in terms of quantity and quality) of evidence, e, affects $E(V)$, the level of evidence creates indirect case selection effects on the *ex post* probability of trial, $\tilde{\Theta}$.

As discussed in Appendix A and shown in Table 24. b., an increase in the prosecutor's evidence, e, increases $E(V)$ when the prosecutor is reasonably assumed to be "conviction-motivated." Thus, if a conviction-motivated prosecutor decides to initiate a case characterized by a relatively *low* level of evidence, e, the relative values of the other case characteristics in the initiated (observed) case are expected to *increase* $E(V)$, *ceteris paribus*.

Conversely, if a conviction-motivated prosecutor decides to initiate a case characterized by a relatively *high* level of evidence, e, it is possible, as with the previous scenario (i.e., low level of evidence), that the initiated (observed) case will have other case characteristics that, *ceteris paribus*, are expected to *increase* $E(V)$. But it is also more likely in this scenario (i.e., high level of evidence) than the previous scenario (i.e., low level of evidence) that the initiated (observed) case will have other case characteristics that, *ceteris paribus*, are expected to *decrease* $E(V)$.[46] It is worth noting that Eisenberg and Farber (2003) mention the latter possibility and ignore the former possibility.

For example, *ceteris paribus*, compared to an initiated case in which the prosecutor has weaker or less evidence, an initiated case in which the prosecutor has relatively stronger or more evidence is just as likely to to be characterized by a relatively high value of the prosecutor's unit cost of trial expenditure, C, as such an initiated case is likely to be characterized by a relatively low value of C. The same predictions apply for the defendant's unit cost of trial expenditure, K, the parties' common *ex ante* expected trial sentence, S^T, the parties' common *ex ante*

[46] While this expected case selection "adjustment" in the mix of case characteristics may seem counter-intuitive in several respects, it is perfectly consistent with the case selection rule. First, it is important to recognize that there are more factors than the amount of evidence that determine the parties' common *ex ante* expected probability of trial conviction. Second, it is important to recognize that the assumption that the prosecutor is conviction-motivated deals only with the weight that the prosecutor's *generally* places on various (endogenous) aspects of case disposition for all types of federal criminal antitrust cases.

Understanding the previous two paragraphs in the body is crucial for understanding the predicted case selection effects from changes in the values of case characteristics. Understanding the previous paragraphs conceptually is also crucial for understanding the other predicted sample selection effects (e.g., trial selection effect) discussed later.

expected probability of trial conviction, $\hat{\Pi}$,[47] and the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial.

In addition, when compared to an initiated case with weaker or less evidence, an initiated case with a relatively higher level of evidence may be characterized by relatively high or low values of the variability of the judge's conviction decision standard, λ. This prediction is based on the ambigous effect of λ on $E(V)$, as discussed in Appendix A, and not due to the mechanics of the modeled case selection process.

In sum, the total (net) indirect case selection effect of more or stronger evidence, e, on the observed trial rate, $\tilde{\Theta}$, is ambiguous. Nonetheless, this ambiguity does not preclude meaningful empirical tests of the implications of the case selection process and initiated cases with (systematically) more or stronger – or less or weaker – evidence. In general, the case selection process followed by a conviction-motivated prosecutor, as modeled, suggests different mixes of case characteristics in initiated cases with relatively extreme values of certain case characteristics, like e. The conviction-motivated prosecutor's expected case selection adjustments, and the associated expected mixes of case characteristics in initiated cases that are implied by the case selection model (i.e., initiate if $E(V) \geq 0$) suggest meaningful empirical tests, as discussed later.

Case selection effects of allegations of bid-rigging

Based on the insight into the prosecutor's case-bringing decision process provided by former U.S. Assistant Attorney General for Antitrust, Donald I. Baker (1978), it is reasonable to assume that a high probability of conviction is a criterion for federal criminal antitrust cases that the DOJ Antitrust Division initiates. As several researchers and, in particular, Snyder (1989, 1990), have discussed, prosecutors can generally prove bid-rigging conspiracies more easily than they can prove price-fixing conspiracies. Victims of bid-rigging are government agencies who gather the conspirators' bids, so probative evidence is generally easier for prosecutors to discover and gather in bid-rigging cases compared to price-fixing cases. In addition, bid-rigging cases

[47] For completeness, note that $\hat{\Pi}$ does not directly influence $\tilde{\Theta}$.

usually feature direct testimony about the conspiracy (Snyder, 1989) given by a conspirator (Snyder, 1990).

In terms of the case selection model, it follows that the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, is higher in bid-rigging cases than non-bid-rigging cases like cases involving alleged exclusionary practices or price-fixing. According to the prosecutor's case selection process, if the prosecutor initiates a bid-rigging case, compared to a federal criminal antitrust case that does not involve alleged bid-rigging, the initiated case is just as likely to be characterized by a relatively low value of the defendant's unit cost of trial expenditure, K, which would decrease $E(V)$, as the initiated case is likely to be characterized by a relatively high value of K, which would increase $E(V)$.[48]

The same predictions apply for the parties' common *ex ante* expected trial sentence, S^T, the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, the expected relative value(s) of the level of the prosecutor's evidence, e,.[49] as well as the prosecutor's unit cost of trial expenditure, C. Meanwhile, as before, an initiated case with a relatively high value of $\hat{\Pi}$ may be characterized by relatively high or low values of the variability of the judge's conviction decision standard, λ. In sum, the total (net) indirect case selection effect on the observed trial rate, $\tilde{\Theta}$, from a (systematically) relatively high value of $\hat{\Pi}$ in certain types of initiated cases is ambiguous.

Case selection effects of higher (opportunity) trial costs for the prosecutor

48 The previous explanation of the case selection effects of relatively low and high values of the level of evidence in initiated cases was meant to provide the reader with a complete understanding of the implications of the case selection process. Because this study uses Chapter VI to analyze the implications of marginal *increases* in the values of case characteristics, which are relevant for regression analysis and for comparison to the analogous predictions developed by Eisenberg and Farber (2003), this study does not discuss the implications of marginal decreases in $\hat{\Pi}$ (or a relatively low value of $\hat{\Pi}$ observed in an initiated case).

Discussing both scenarios would add unnecessarily to the volume of this already voluminous study. Thus, the reader should keep in mind the implications of the case selection process when the value of the subject case characteristic is relatively high or low in initiated cases. Both scenarios will be relevant for the sample means tests reported in Chapter X.

49 Recall that a vector, **e**, represents evidence so that it can be measured in terms of quality and quantity. Thus, a prosecutor may initiate a bid-rigging case with good but little evidence. That is, sometimes quality matters more than quantity.

Researchers like Posner (1970) and Gallo, *et al*. (2000) have commented on the DOJ Antitrust Division's traditionally low funding levels. Lower funding or budget levels imply higher per unit (opportunity) costs of trial expenditure, C. Higher values of C are expected to negatively directly affect the *ex post* probability of trial, $\widetilde{\Theta}$, as shown in Equation [26]. Higher values of C in the cases that the prosecutor initiates, are also expected to create indirect case selection effects (and a net case selection effect) on $\widetilde{\Theta}$.

Importantly, in contrast to the above analyses of case selection effects created by more evidence, e, or a relatively high value of $\hat{\Pi}$, the model of the prosecutor's case selection decision implies particular relative values of case characteristics other than C in *initiated* cases that are characterized by a relatively high value of C. That is because, *ceteris paribus*, a relatively high value of C lowers the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, and creates a "deficit" in $E(V)$. In order for a case to be initiated, this "deficit" must be at least balanced (if not outweighed) by relative values of other case characteristics that increase $E(V)$. Thus, if an initiated case is characterized by a relatively high value of C, the model of the prosecutor's case-bringing decision suggests that such an initiated case must be further characterized by relative values of other case characteristics that increase $E(V)$. Hence the specificity of the following predictions.

Under the reasonable assumption that the prosecutor is (purely) "conviction-motivated," if the prosecutor considers initiating a case characterized by a relatively high value of C, the model of the case selection process implies that the prosecutor would be likely to initiate that case if it is also characterized by relatively (compared to initiated cases with lower values of C) high values of the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$; the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial conviction; the prosecutor's evidence, e; the defendant's unit trial cost, K; and the *ex ante* expected trial sentence, S^T.[50] Meanwhile, an initiated case with a relatively high value of C may be

[50] These expectations for the relative values of $\hat{\Pi}$ (and ρ) in initiated cases characterized by higher values of *C* suggest that if the DOJ Antitrust Division's budgeted resources are limited, antitrust prosecutors will be more inclined to bring cases that are relatively easier to successfully prosecute. Some argue that the DOJ's rising conviction rate after the late 1970s reflects the effects of prosecutors' limited budgets. As Snyder (1989) notes, when the DOJ's antitrust enforcement budgets diminished during this period, the cases brought usually involved unsophisticated and otherwise easily prosecuted conspirators. Thus, if a particular case is brought that requires a

characterized by relatively high or low values of the variability of the judge's conviction decision standard, λ.

As discussed in more detail in Appendix A, in initiated cases with a (systematically) relatively high values of C, the expected relatively (compared to cases not characterized by relatively high values of C) higher values of S^T, e, and K have ambiguous direct marginal effects on the *ex ante* expected probability of trial, Θ, and thus, the observed trial rate, $\tilde{\Theta}$. The expected relatively higher value of ρ in initiated cases with relatively high values of C is expected to have a positive effect on Θ and thus $\tilde{\Theta}$. (Note that $\hat{\Pi}$ does not directly affect $\tilde{\Theta}$.) In sum, the total (net) indirect case selection effect on the observed trial rate, $\tilde{\Theta}$, from a (systematically) relatively high value of C in initiated cases is ambiguous.

Case selection effects of higher (opportunity) trial costs for the defendant

The defendant's unit cost of litigation, K, is expected to vary by type of defendant. As the law and economics literature (e.g., Eisenberg and Farber (2003)) commonly assumes, it is reasonable to expect corporate defendants to face lower unit costs of trial expenditure than individual defendants face, (and for higher ranking corporate officials to have lower unit costs of trial expenditure than lower ranking defendants).

The composition of types of defendants in federal antitrust cases and, in particular, the types of criminal defendants, has varied over the years. Gallo *et al*. (2000) report that from 1955-1997 about 88 percent of the total number of CCH cases (criminal and civil) involved firms as defendants and about 40 percent involved individuals as defendants. After 1980, the total number of defendants in all CCH cases (criminal and civil) decreased, but the percentage of cases with individuals as defendants increased relative to the percentage of cases with firms as defendants. According to Gallo *et al*. (2000), the majority of individual defendants are high-ranking corporate officials.

---

significant amount of trial expenditure (or effort) by the prosecutor, then, given that the case has been brought, the case is one in which the prosecutor anticipates a trial conviction. Snyder's arguments are consistent with this study's case selection model's predicted case selection effects on the observed trial conviction rate from higher trial costs for the prosecutor.

With respect to criminal CCH cases in particular, Gallo *et al*. (2000) find that since 1955 about 80 percent of the criminal cases involved firms as defendants and about 27 percent of the criminal cases involved individuals as defendants. The number of criminal cases with individuals as defendants generally increased since 1955, especially after 1979, but the total number of individual defendants decreased in the 1980s. Further, nearly every criminal case from 1955-1972 involved a firm as a defendant, but that is not true for the period 1972-1997. Moreover, Snyder (1989, 1990) finds that most of the corporate defendants in the 1980s were small firms. In general, these trends suggest an increase in the defendant's unit cost of trial expenditure in federal criminal antitrust cases from the 1950s to at least the 1990s.

According to Equation [26], a higher value of the defendant's unit cost of trial expenditure, K, negatively affects the *ex post* probability of trial, $\widetilde{\Theta}$. Meanwhile, because increases in K positively affect $E(V)$, if the prosecutor is conviction-motivated and is willing to bring a case against an individual defendant (i.e., with a relatively high values of K compared to a case involving a corporate defendant), the initiated case is just as likely to be characterized by a relatively low value of prosecutor's unit cost of trial expenditure, C, which would increase $E(V)$, as the initiated case is likely to be characterized by a relatively high value of C, which would decrease $E(V)$.[51]

The same predictions apply for the parties' common *ex ante* expected trial sentence, S^T, the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, and the expected relative value(s) of the level of the prosecutor's evidence, e. In contrast, as before, an initiated case with a relatively high value of K may be characterized by relatively high or low values of the variability of the judge's conviction decision standard, λ. Thus, the total (net) indirect case

---

[51] The previous explanation of the case selection effects of relatively low and high values of the level of evidence in initiated cases was meant to provide the reader with a complete understanding of the implications of the case selection process. Because this study uses Chapter VI to analyze the implications of marginal *increases* in the values of case characteristics, which are relevant for regression analysis and for comparison to the analogous predictions developed by Eisenberg and Farber (2003), this study does not discuss the implications of marginal decreases in $\hat{\Pi}$ (or a relatively low value of $\hat{\Pi}$ observed in an initiated case).

Discussing both scenarios would add unnecessarily to the volume of this study. Thus, the reader should keep in mind the implications of the case selection process when the value of the subject case characteristic is relatively high or low in initiated cases. Both scenarios will be relevant for the sample means tests reported in Chapter X.

selection effect on the observed trial rate, $\widetilde{\Theta}$, from a (systematically) relatively high value of K in initiated cases is ambiguous.

Case selection effects of the change from misdemeanor to felony status of certain offenses

On several occasions Congress has increased statutory penalties for federal criminal antitrust offenses. Notably, the Antitrust Procedures and Penalties Act of 1974 ("APPA") elevated Sherman Act violations from misdemeanor to felony status. This increased the maximum fine to $100,000 for individuals and to $1 million for corporations, thereby doubling the maximum fine for individuals and increasing maximum fines for corporate defendants by a factor of twenty. For individual offenders, the shift to felony penalties also increased the maximum prison term from one to three years per count.[52]

As discussed in Chapter II, Cohen (1992), Gallo, *et al.* (1994, 2000), and Snyder (1990) find significant increases in fines imposed by judges following the enactment of the APPA. In terms of the case selection model, the elevation of maximum penalties for federal criminal antitrust offenses translates into a higher value of the prosecutor's *ex ante* expected sanctions following a trial conviction, S^T.[53] This exogenous institutional increase in S^T positively directly affects the *ex post* probability of trial, $\widetilde{\Theta}$.

In order to specify the case selection model's predicted indirect effects on case outcomes from the shift to felony penalties, continue to assume that the prosecutor is (purely) "conviction motivated" and the prosecutor has weaker diminishing returns in trial expenditure. Given these

[52] In 1977, the DOJ adopted a new policy regarding sentencing recommendations for Sherman Act convictions (Gallo, *et al*. 2000). The aim of the DOJ policy change was to increase deterrence through stiffer antitrust penalties. It mandated that prosecutors seek a base corporate fine of ten percent of the convicted firm's total sales affected by the violation. The directive also suggested upward adjustments for evidence of excessively high markups and for recidivism; but it allowed downward adjustments if the defendant cooperated with the government or if the suggested base fine jeopardized the defendant firm's competitive viability.

[53] The provisions of the APPA only apply to violations that occurred after November 1, 1974. It is still possible that judges, in order to enhance their promotion potential (Cohen 1992), would be motivated, post-APPA, to increase the sanctions they impose even when the provisions of the APPA do not apply. Imposing harsher penalties for criminal antitrust violations could elevate their standings in the eyes of politicians or prosecutors who supported the spirit of the APPA to enhance penalties for criminal antitrust violations. As Cohen (1992) explains, the White House and key U.S. Department of Justice officials are main actors in appeals court appointments.

Gallo *et al*. (2000) report that following the enactment of each penalty-enhancing law since 1955, there has been a significant increase in average real fines for both firm and individual defendants, but less than the proportionate increases in the statutory maximum penalties. The fines imposed on firms have increased more than fines imposed on individuals.

assumptions, the prosecutor's case selection process implies that a federal criminal antitrust case that the prosecutor initiates under the APPA (an exogenous shock that increases the value of S^T) is likely to be characterized by a relatively (compared to pre-APPA cases) high or low values of the prosecutor's unit cost of trial expenditure, C, the defendant's unit trial cost, K; the prosecutor's evidence, e; the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$; and the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial conviction. On the contrary, an initiated case with a relatively high value of S^T may be characterized by relatively high or low values of the variability of the judge's conviction decision standard, λ.

As discussed above with respect to the expected direct effects of case characteristics on the observed trial rate, and in more detail in Appendix A, in initiated cases with a (systematically) relatively high value of S^T (e.g., post-APPA) compared to other types of cases, the expected relatively high value of C and the expected relatively low values of e and K have ambiguous direct marginal effects on the *ex ante* expected probability of trial, Θ, and thus, on the observed trial rate, $\tilde{\Theta}$. In contrast, the expected relatively low value of ρ is expected to decrease the *ex ante* expected probability of trial, Θ, and thus, the observed trial rate, $\tilde{\Theta}$. (Note that $\hat{\Pi}$ does not directly influence $\tilde{\Theta}$.) In sum, the total (net) indirect case selection effect on the observed trial rate, $\tilde{\Theta}$, from (systematically) relatively high values of S^T (caused by an exogenous shock to S^T) in certain types of initiated cases (e.g., post-APPA) is ambiguous.

Case selection effects from the alleged amount of economic harm from the antitrust violation

Higher alleged conspiracy sales implies greater alleged economic harm from the conspiracy, which implies a higher value of the parties' common *ex ante* expected sentence following a trial conviction, S^T. The analysis and predictions regarding the case selection effects on the *ex post* probability of trial, $\tilde{\Theta}$, are identical to those regarding the enactment of the APPA. The various determinants of S^T provide several potential empirical tests.

Case selection effects of a more confident prosecutor

Suppose that based on previous trial outcomes in similar or recent cases, for example, the prosecutor is relatively more confident about the chances of a trial conviction in a particular case than she is in other cases. The prosecutor's confidence can be expressed by a relatively high value of the idiosyncratic component, ρ, of the prosecutor's expected probability of trial conviction. An increase in ρ (and/or a decrease in the idiosyncratic component of the defendant's expected probability of trial conviction, δ) is expected to increase the observed trial rate. As discussed in Appendix A and shown in Table 26. b., an increase in ρ is expected to increase a conviction-motivated prosecutor's *ex ante* expected value of initiating a case, $E(V)$.

According to the case selection process, when compared to cases in which the (conviction-motivated) prosecutor is less confident, initiated cases in which the prosecutor is relatively more confident are just as likely to be characterized by relatively high or low values of the prosecutor's unit cost of trial expenditure, C, the defendant's unit trial cost, K; the prosecutor's evidence, e; the parties' common expected trial sentence, S^T, and even the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$. (Note that $\hat{\Pi}$ does not directly influence $\tilde{\Theta}$.) As before, an initiated case with a relatively high value of ρ may be characterized by relatively high or low values of the variability of the judge's conviction decision standard, λ. Thus, as with case selection effects associated with other case characteristics, the total (net) indirect case selection effect on the observed trial rate, $\tilde{\Theta}$, from a (systematically) relatively high value of ρ in certain types of initiated cases is ambiguous.

Empirically Testable Implications of Indictment as a Selection Process

Table 3 provides the theoretical implications of the prosecutor's case selection process and the defendant's plea decision (i.e., trial selection) process with respect to the influence of marginal increases in the values of case characteristics on the observed trial rate.[54] The theoretical implications summarized in Table 3 are relevant for regression analysis used to

[54] The predictions in Table 3 assume the prosecutor is "conviction-motivated" and has weaker diminishing returns in trial expenditure than the defendant.

identify the determinants of the observed trial rate. Inspection of Table 3 demonstrates that the second-order direct "stakes effects" and "litigation expenditure effects," as well as the indirect "case selection effects" on the *ex post* expected probability of trial from changes in the values of case characteristics generally render ambiguous all predicted net effects on the *ex post* expected probability of trial from marginal increases in the values of case characteristics.

These regression-related predictions are relevant at least for purposes of comparisons to the predictions in Eisenberg and Farber (2003), who developed the case selection model upon which the case selection model is built. Because Eisenberg and Farber did not treat the litigants' trial expenditures as choice variables or the expected penalty as an endogenous variable, they did not consider the direct second-order "stakes effects" or "litigation expenditure effects" that create ambiguity in the predictions of this study. Moreover, Eisenberg and Farber employed a "majority rules" signing convention to determine a net effect of indirect case selection effects. This questionable signing convention allowed them to report unambiguous indirect case selection effects on the *ex post* probability of trial.

These regression-related predictions are also relevant for comparisons to the predictions in Snyder (1989, 1990). Snyder recognized, through limited discussion, the potential effects of the prosecutor's case-bringing decision on the observed trial rate. He did not, however, systematically consider case selection effects on the observed trial rate as this study does.

Table 3. Signs of expected direct and indirect effects on the observed trial rate, $\tilde{\Theta}$

Case Characteristic	Expected First-Order Direct Effect on $\tilde{\Theta}$	Expected Net Second-Order Direct "Trial Expenditure" Effect on $\tilde{\Theta}$ *		Expected Net Second-Order Direct "Stakes" Effect on $\tilde{\Theta}$ *		Expected Indirect Case Selection Effect on $\tilde{\Theta}$	Expected Total (Net) Effect on $\tilde{\Theta}$
		When $\hat{\Pi} \geq \overline{\Pi}$	When $\hat{\Pi} < \overline{\Pi}$	When $\hat{\Pi} \geq \overline{\Pi}$	When $\hat{\Pi} < \overline{\Pi}$		
$\hat{\Pi}$ **	0	0		0		?	?
C	-	?	+	-	?	?	?
K	-	+	?	?	+	?	?
S^T **	+	-		0		?	?
e	0	+	-	?		?	?
λ	0	?		?		?	?
ρ	+	0		0		?	?

* As discussed above and, in detail, in Appendix A,.these second-order direct effects involve movements in both litigants' optimal levels of trial expenditure, sometimes in the same direction but sometimes in opposite directions. Hence the "net" reference.

** The marginal increases in $\hat{\Pi}$ and S^T are from exogenous shocks.

Despite the ambiguous expected indirect case selection effects on the observed trial rate shown in Table 3, the theoretical implications of the prosecutor's case selection process, as specified in the case selection model, are empirically testable. That is because the theoretical model of the prosecutor's case-bringing decision implies particular expected mixes of case characteristics, given the relative value of at least one case characteristic, in cases that the prosecutor does initiate. These predictions suggest firm relationships among the relative values of different pairs of case characteristics in cases that are initiated (i.e., along the *same* game tree branch stemming from the case-bringing decision node).[55] These relationships can be tested meaningfully using sample means tests using the sample of initiated (observed) cases.

[55] If potential cases that prosecutors chose not to initiate were observable, then it would be possible to test other predictions provided by the model of the case selection process related to the more straightforward question of whether relative values for case characteristics that are associated with a higher (lower) value of $E(V)$ are found in cases that the prosecutor does (does not) initiate. That is, if cases that are not initiated were observable, then another meaningful test would be to look for differences in the relative values of the same case characteristic in the sample of cases that are initiated versus the sample of cases that are not initiated (i.e., across game tree branches stemming from the case-bringing decision node).

Table 4 summarizes the predicted relationships among the relative values of different pairs of case characteristics in cases that are initiated. To understand Table 4 it is useful to first understand what relationships among the relative values of different pairs of case characteristics in initiated cases that the model of the prosecutor's case selection process does *not* predict.

For example, suppose an initiated case is characterized by a relatively *high* value of the prosecutor's unit cost of trial expenditure, C. *Ceteris paribus*, a relatively *high* value of C is expected to be associated with a relatively *low* value of the prosecutor's *ex ante* expected value of initiating a case, $E(V)$. In this situation, the model of the prosecutor's case selection process does *not* predict that the *initiated* case also would be characterized by values of other case characteristics that are expected to *further reduce* $E(V)$. In this situation, relative values of other case characteristics *not* predicted by the model of the prosecutor's case selection process include relatively *low* values of the defendant's unit trial cost, K; the level of the prosecutor's evidence, e; the parties' common expected trial sentence, S^T; the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$; or the idiosyncratic component, ρ, of the prosecutor's expected probability of trial conviction.

Continuing the preceding example, if an initiated case is characterized by a relatively *high* value of C, then observed relatively *high* values of K, e, S^T, $\hat{\Pi}$, or ρ in such an initiated case would support the predictions of the model of the prosecutor's case selection process. Meanwhile, if an initiated case is characterized instead by a relatively *low* value of C, then observed relatively *high or low* values of K, e, S^T, $\hat{\Pi}$, or ρ in such an initiated case also would be consistent with the predictions of the model of the prosecutor's case selection process.

Table 4 provides the combinations of relative values of pairwise combinations of case characteristics, in cases that are initiated, that are consistent and inconsistent with the predictions implied by the model of the prosecutor's case selection process.[56] Table 4 serves as the basis of sample means tests reported in Chapter X.

[56] The predictions in Table 4 assume the prosecutor is "conviction-motivated" and has weaker diminishing returns in trial expenditure than the defendant.

Table 4. Pairs of relative values of case characteristics in initiated cases and their consistency with the predictions of the case selection model

	High $\hat{\Pi}$	Low $\hat{\Pi}$	High C	Low C	High K	Low K	High S^T	Low S^T	High **e**	Low **e**	High ρ	Low ρ
High $\hat{\Pi}$	C*											
Low $\hat{\Pi}$	C*	I*										
High C	C	I	I*									
Low C	C	C	C*	C*								
High K	C	C	C	C	C*							
Low K	C	I	I	C	C*	I*						
High S^T	C	C	C	C	C	C	C*					
Low S^T	C	I	I	C	C	I	C*	I*				
High **e**	C	C	C	C	C	C	C	C	C*			
Low **e**	C	I	I	C	C	I	C	I	C*	I*		
High ρ	C	C	C	C	C	C	C	C	C	C	C*	
Low ρ	C	I	I	C	C	I	C	I	C	I	C*	I*

* As explained in Chapter X, different empirical variables can represent the same theoretical variable. Thus, combinations of relatively high and low values of the same theoretical variable represented by more than one empirical variable is possible for a given empirical observation. Pairs that are consistent with case selection model are denoted with "C" and pairs that are inconsistent are denoted with "I". Horizontal axis provides the possible relative values of first case characteristic in a possible pair. Vertical axis provides the possible relative values of second case characteristic in a possible pair.

CHAPTER VII

DETERRMINANTS OF THE OBSERVED TRIAL CONVICTION RATE

This chapter defines the conditional *ex post* probability of trial conviction, $\widetilde{\Pi}$, which is the observed trial conviction rate. In order to observe a case disposed by trial conviction, two things must happen: the prosecutor must initiate the case and the case must go to trial. The prosecutor's case selection and the defendant's plea decision (i.e., trial selection) processes thus condition the sample of initiated cases that go to trial and thus create case and trial selection effects, respectively, which influence the observed trial conviction rate. This study's theoretical framework thus diverges from and improves upon Snyder's (1989, 1990) as well as Eisenberg and Farber's (2003) models by recognizing that the case selection process, in addition to the trial selection process, affects the mix of case characteristics in the cases that go to trial, thus influencing the observed trial conviction rate, $\widetilde{\Pi}$.

These considerations are necessary for meaningful empirical tests of the determinants of trial conviction rates in federal criminal antitrust cases. The implications of the prosecutor's case selection process were explained in the previous chapter. The implications of the defendant's plea decision (i.e., trial selection) process specified above are explained in this chapter. They are conceptually analogous to the implications of the case selection process, however. Assume the defendant is faced with a case against him that has a certain case characteristic with a relatively extreme value that has a positive (negative) effect on the *ex ante* probability of trial, Θ, (and hence the *ex post* probability of trial, $\widetilde{\Theta}$). With that "cushion" ("deficit") created by that extreme-valued case characteristic, the defendant would be willing to go to trial even if (only if) other case characteristics – even (especially) with relatively extreme values – have negative (positive) effects on the probability of trial as long as (so that) the condition $\Theta > 0$ holds.

Empirically analyzing the sample selection bias created by the defendant's trial selection process is possible because cases that go to trial and cases that do not go to trial are observable. One way is to control for sample selection bias created by the defendant's plea decision in regression models of trial outcomes using the usual approaches (e.g., the approach introduced by Heckman).

As explained below, theory implies unambigous predicted direct effects on the observed trial conviction rate from marginal increases in the values of case characteristics. Except for a couple of instances, however, the predicted indirect case selection effects and trial selection effects are ambiguous. Thus, the predicted total (net) marginal effects on the observed trial rate from marginal increases in the values of case characteristics are ambigous.

Despite these ambiguities, other meaningful empirical tests of the implications of the model of the defendant's plea decision are possible. The theoretical model of the defendant's plea decision implies firm relationships among the relative values of the same case characteristic in the sample of cases that go to trial versus the sample of cases that do not go to trial (i.e., across *different* game tree branches stemming from the plea decision mode). These relationships can be tested meaningfully with sample means tests using the samples of (initiated) cases that do and do not go to trial.

In addition, the predictions of the model of the defendant's plea decision suggest firm relationships among the relative values of different pairs of case characteristics either in the sample of cases that go to trial or in the sample of cases that do not go to trial (i.e., along the *same* game tree branch stemming from the plea decision mode). These relationships can be tested meaningfully using sample means tests using either the sample of cases that go to trial or the sample of cases that do not go to trial (or both samples).

Direct Effects on the Observed Trial Conviction Rate

The observed trial conviction rate is the *ex post* expected trial conviction rate conditional on the case being filed *and* the case going to trial. This implies the following expanded definition of the observed trial conviction rate:

$$E(\tilde{\Pi}) = E\left[\hat{\Pi} \,\middle|\, E(V) \geq 0, (\rho - \delta)S^T - CX - KY > 0\right], \qquad [27]$$

where $\hat{\Pi} = \Pi([\mathbf{e} + P(X) - D(Y)]/\lambda)$.

The definitions of the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, the prosecutor's arguments at trial, $P(X)$, and the defendant's trial arguments, $D(Y)$, in Equation [27] are specified by the litigation model. Meanwhile, $X = X^*$ and $Y = Y^*$ are the optimal (Nash equilibrium) values of the prosecutor's and the defendant's respective choices of levels of trial expenditure. Both X^* and Y^* are implicit functions of C, K, S^T, $\mathbf{e}$, λ, as well as each other as specified by the litigation model.

The expected marginal direct effects on the (Nash equilibrium) value of $\hat{\Pi}$ from changes in the values of case characteristics are also the expected marginal effects from changes in the values of case characteristics on the value of $\tilde{\Pi}$. First consider the marginal effects of case characteristics on the *ex ante* probability of trial conviction, $\hat{\Pi}$. The litigation model predicts that $\hat{\Pi}$ is positively related to the prosecutor's evidence, e. The litigation models also predicts that an increase in the defendant's unit cost of trial expenditure, K, will increase $\hat{\Pi}$.[57] An increase in C lowers $\hat{\Pi}$, however. Given the assumption that the prosecutor has weaker diminishing returns in trial expenditure, an increase in the trial stakes for both parties, S^T, will increase $\hat{\Pi}$. Meanwhile, the prosecutor's and the defendant's respective idiosyncratic components, ρ and δ, of their *ex ante* expected probabilities of trial conviction, do not directly affect $\hat{\Pi}$.

Finally, the litigation model predicts that as long as the parties' rates of diminishing returns from trial expenditure are sufficiently close, an increase (a decrease) in λ will increase (decrease) the underdog's chances at trial. Because assumptions regarding the "closeness" of the parties' trial productivities cannot be easily justified, the expected effect of λ on $\hat{\Pi}$ is ambiguous. In sum, these are the expected "direct effects" of case characteristics on the *ex ante* probability of trial conviction, $\hat{\Pi}$ (even if they affect $\hat{\Pi}$ indirectly through $X = X^*$ and $Y = Y^*$).[58]

---

[57] Gallo *et al*. (2000) report that for the period 1955-1997, the overall conviction rate in criminal CCH cases does not vary significantly by type of defendant or by the corporate rank of individual defendants. For civil CCH cases, however, the DOJ convicted high-ranking individuals less frequently than the DOJ convicted lower raking individuals.

[58] Since these expected "direct" effects are based on the predictions of the litigation model, they account for the effects on the *ex ante* expected probability of trial conviction, $\hat{\Pi}$, from changes in the litigants' optimal levels of trial expenditure caused by changes in the values of case characteristics. The predicted directions of the changes in

Case Selection Effects on the Observed Trial Conviction Rate

This study's theoretical framework diverges from and improves upon Snyder's (1989, 1990) as well as Eisenberg and Farber's (2003) models by recognizing that the case selection process, in addition to the trial selection process, affects the mix of case characteristics in the cases that go to trial, thus influencing the observed trial conviction rate, $\tilde{\Pi}$. Note that the first conditioning event expressed in Equation [27] is that the prosecutor brought the case. This conditioning event, represented by $E(V) \geq 0$, produces indirect "case selection effects" on the observed trial conviction rate that are analogous to the indirect case selection effects discussed in Chapter VI with regard to the observed trial rate. Specifically, if a conviction-motivated prosecutor considers initiating a case with relatively more or stronger evidence, e, then because her *ex ante* expected value of bringing a case, $E(V)$, unambiguously increases in e when she is conviction-motivated, she would be willing to tolerate case characteristics that otherwise reduce $E(V)$. Of course, she also would be interested (even more interested) in initiating such a case if it was also characterized by case characteristics that raise $E(V)$.

Thus, an *initiated* case with more or stronger evidence is just as likely to be further characterized by a relatively low value of the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, (compared to cases not characterized by more or stronger evidence) as it is likely to be further characterized by a relatively high value of $\hat{\Pi}$. (Recall that $E(V)$ increases in $\hat{\Pi}$.) Thus, the expected case selection effect of more or stronger evidence on the observed trial conviction rate, $\tilde{\Pi}$, is ambiguous.[59]

In addition, if the prosecutor initiates a case with relatively more or stronger evidence, e, then the initiated case is just as likely to feature relatively higher or lower values of the

---

the litigants' optimal levels of trial expenditure that are used to determine the direct effect on $\hat{\Pi}$ from changes in the values of case characteristics are the same predictions that determine the expected indirect "trial expenditure effects" on the *ex post* probability of trial, $\tilde{\Theta}$, from changes in the values of case characteristics, as discussed in Chapter VI.

[59] As discussed in Chapter VI, this prediction contradicts the prediction that would follow from the analytical approach of Eisenberg and Farber (2003). It is still useful and important to recognize that if the prosecutor initiates a case with less or weaker evidence, then, *ceteris paribus*, the initiated case is likely to be characterized by a relatively high value of $\hat{\Pi}$. Otherwise, *ceteris paribus*, the condition $E(V) \geq 0$ is not likely to hold and the prosecutor would not initiate the case in the first place.

prosecutor's trial cost, C, the defendant's trial cost, K, the *ex ante* expected trial sentence, S^T, and the idiosyncratic component, ρ, of the prosecutor's expected probability of trial conviction. Importantly, the prosecutor is likely to initiate a case with strong or abundant evidence and further characterized by these (ambiguous) relative values of C, K, and S^T without regard for the effects of C, K, and S^T on the endogenous $\hat{\Pi}$.

On the other hand, if the prosecutor initiates a case in which she has a relatively high unit trial cost, C, which reduces $E(V)$, then the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, is likely to be higher, since $E(V)$ increases in $\hat{\Pi}$. Thus, the indirect case selection effect on the observed trial conviction rate, $\tilde{\Pi}$, from *higher* unit costs of trial expenditure for the prosecutor is positive. Without changing the expected sign of the case selection effect on $\tilde{\Pi}$ from initiated cases with a relatively high value of C, the case selection process also implies that an initiated case with a relatively high value of C is also likely to be characterized by a relatively high value of K, and relatively high values of S^T, e, and ρ (in addition to the higher value of $\hat{\Pi}$).

Conversely, if the prosecutor initiates a case characterized by a relatively high unit cost per defendant's trial expenditure, K, then the initiated case is just as likely to feature relatively higher or lower values of the *ex ante* expected probability of trial conviction $\hat{\Pi}$, the level of evidence, e, the prosecutor's trial cost, C, the *ex ante* expected trial sentence, S^T, and the idiosyncratic component, ρ, of the prosecutor's expected probability of trial conviction. Thus, the expected indirect case selection effect on $\tilde{\Pi}$ from initiated cases with a relatively high value of K is ambiguous.

Using the same logic as the situations involving initiated cases with relatively high values of e and K, the predicted case selection effects on $\tilde{\Pi}$ from higher values of S^T and ρ are ambiguous. Meanwhile, the marginal effect on $E(V)$ of the variability of the judge's trial conviction standard, λ, is ambiguous. Thus, the expected sign of the case selection effect on $\tilde{\Pi}$ from an increase in λ is also ambiguous.

Trial Selection Effects on the Observed Trial Conviction Rate

The trial selection process indirectly affects $\tilde{\Pi}$ in a manner that is similar to the way that the case selection process indirectly affects the *ex post* probability of trial.[60] Despite the conceptual similarity, the condition for a case to go to trial (i.e., the defendant's trial decision rule) differs mathematically from the condition for a case to be initiated. That is, the second conditioning event on the *ex post* probability of trial conviction, $\tilde{\Pi}$, is that the case went to trial, which is represented by $(\rho - \delta)S^T - CX - KY > 0$. This conditioning event does not depend directly on the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$ (like the case selection rule does); however, it does depend on variables that determine $\hat{\Pi}$. The defendant's decision to go to trial shapes the mix of case characteristics in cases that go to trial, which, in turn, affects the observed trial conviction rate.

The total (net) marginal effects of changes in the values of case characteristics have ambiguous effects on the *ex ante* expected probability of trial, Θ. This has no bearing on the following evaluation of the indirect trial selection effects of case characteristics on the *ex post* expected probability of trial conviction, $\tilde{\Pi}$, however. That is because the evaluation of the marginal effects of changes in the values of case characteristics on Θ is done from an *ex ante* perspective (e.g., when evaluating the effects of case characteristics on Θ as they relate to $E(V)$ or $\tilde{\Theta}$, as done in Appendix A and Chapter VI, respectively).[61] In contrast, the evaluation of the trial selection effect on the observed trial rate from changes in the values of case characteristics is performed from an *ex post* perspective.

When evaluating *ex post* the trial selection effects, all that matters are the relative values of case characteristics that, when taken together or at least in pairs, *ceteris paribus*, are expected to satisfy the condition for a case to go to trial, $(\rho - \delta)S^T - CX - KY > 0$. Once a case has gone to trial, the values of the case characteristics upon which the defendant relied in making his plea decision (and upon which the prosecutor relied in her case-bringing decision) are "frozen." In particular, the endogenous variables $X = X^*$ and $Y = Y^*$, as well as S^T, which cause the

60 In this way, this study's theoretical framework diverges from and improves upon Eisenberg and Farber's (2003) model, since Eisenberg and Farber consider only the indirect effects of the case selection process on the observed plaintiff trial win rate.

61 See Appendix A for discussions of the indirect "trial expenditure effects" on Θ.

second-order direct "trial expenditure effects" and "stakes effects" on Θ and thus $\tilde{\Theta}$, are "frozen." That is, once a case goes to trial, the second-order effects on the *ex ante* probability of trial, Θ, that the defendant may consider when making his plea decision are not relevant. As demonstrated below, all that matters is the relative value of a given case characteristic and the fact that the defendant *has chosen* to go to trial (or to avoid trial). With that background, the evaluation of the expected indirect trial selection effects of changes in the values of case characteristics on the observed trial conviction rate, $\tilde{\Pi}$, is relatively straightforward.

Suppose a case that *goes to trial* is characterized by a relatively high *ex ante* commonly expected trial sentence, S^T, the first-order direct effect of which on the *ex post* probability of trial, $\tilde{\Theta}$, is positive. In that situation, it is just as likely that the case is characterized further by values of case characteristics that, *ceteris paribus*, are expected to decrease $\tilde{\Theta}$, as it is likely that the case is characterized further by values of case characteristics that, *ceteris paribus*, are expected to increase $\tilde{\Theta}$.[62]

For example, if a case that goes to trial is characterized by a relatively high value of S^T, then that case could be characterized by relatively high or low values of the litigants' unit costs of trial expenditure, *C* and *K*. The expected *ex post* probability of trial, $\tilde{\Pi}$, could rise or fall with such a range of possible values of *C* and *K*. Thus, the expected (net) trial selection effects on the observed trial conviction rate, $\tilde{\Pi}$, from a marginal increase in the value of S^T are ambiguous.

Using the same logic, the expected (net) trial selection effect on $\tilde{\Pi}$ from a marginal increase in the values of the idiosyncratic component, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, is ambiguous. The expected trial selection effect on $\tilde{\Pi}$ from a marginal increase in *C* is also ambiguous. As explained below, meaningful empirical tests of the implications of the defendant's plea choice (i.e., trial selection) and the implications of cases going to trial with higher values of S^T, ρ, and *C* are nonetheless possible.[63]

---

62 In contrast, if a case that *goes to trial* is characterized by a relatively low value of S^T, which is expected to *decrease* the *ex post* probability of trial, $\tilde{\Theta}$, then that case is likely to be further characterized by case characteristics that, *ceteris paribus*, are expected to increase $\tilde{\Theta}$.

63 Recall that $\hat{\Pi}$ and thus $\tilde{\Pi}$ do not depend directly on the idiosyncratic component, ρ, of the prosecutor's *ex ante* expected probability of trial conviction.

In contrast, if a case that goes to trial is characterized by a relatively high value of the defendant's unit cost of trial expenditure, *K*, then that case is expected to be further characterized by a relatively low value of the prosecutor's unit costs of trial expenditure, *C*, a relatively high value of S^T, as well as relatively low values of the litigants (pre-trial) optimal levels of trial expenditure, X^* and Y^*. Assuming the prosecutor has weaker diminishing returns in trial expenditure, these expected relative values of *C*, S^T, X^* and Y^* imply that the expected indirect trial selection effect on the *ex post* probability of trial conviction, $\widetilde{\Pi}$, from a marginal increase in *K* is positive.

Empirically Testable Implications of the Defendant's Decision to Go to Trial as a Selection Process

Table 5 summarizes the theoretical implications of the judge's trial conviction decision (i.e., trial conviction selection) process with respect to the influence of case characteristics on the observed trial conviction rate.[64] The theoretical implications summarized in Table 5 apply to regression analysis used to identify the determinants of the observed trial rate. Such regression analysis must control for sample selection bias from the defendant's plea decision (i.e. trial selection) process. Traditional regression techniques do not allow controlling for the sample selection bias created by the prosecutor's case selection process, however.

These regression-related predictions are relevant for comparisons to the predictions in Snyder (1989, 1990). Snyder recognized, through limited discussion, the potential effects of the plea decision process and the prosecutor's case-bringing decision on the observed trial conviction rate. He did not, however, systematically consider those selection effects on the observed trial conviction rate as this study does.

[64] The predictions in Table 5 assume the prosecutor is "conviction-motivated" and has weaker diminishing returns in trial expenditure than the defendant.

Table 5. Signs of expected direct and indirect effects on the observed trial conviction rate, $\tilde{\Pi}$

Case Characteristic	Expected Direct Effect on $\tilde{\Pi}$ *	Expected Indirect Case Selection Effect on $\tilde{\Pi}$	Expected Indirect Trial Selection Effect on $\tilde{\Pi}$	Expected Total (Net) Effect on $\tilde{\Pi}$
$\hat{\Pi}$ **	+	?	0	?
C	-	+	?	?
K	+	?	+	?
S^T **	+	?	?	?
e	+	?	?	?
λ	?	?	0	?
ρ	0	?	?	?

* The expected direct effects include first- and second-order effects, as explained in Appendix A.

** The marginal increase in $\hat{\Pi}$ and S^T are from exogenous shocks.

Despite the ambiguous expected indirect trial selection effects on the observed trial rate shown in Table 5, the theoretical implications of the defendant's trial selection process, as specified in the model of the defendant's plea decision, are empirically testable. Table 6 summarizes the predicted relationships among the relative values of different pairs of case characteristics in cases that go to trial and in cases that end in a plea of guilty or *nolo contendere*. To understand Table 6, first consider the relationships among the relative values of different pairs of case characteristics in cases that go to trial that are *not* predicted by the model of the defendant's plea decision (i.e., trial selection) process. For example, suppose a case that goes to trial is characterized by a relatively *high* value of the prosecutor's unit cost of trial expenditure, C. *Ceteris paribus*, a case that goes to trial is *not* expected to have a relatively *high* value of C.

In that situation, the model of the defendant's plea decision (i.e., trial selection) process does *not* predict that the case that *goes to trial* also would be characterized by values of other case characteristics that are expected to *further reduce* $\tilde{\Theta}$. In this situation, relative values of other case characteristics *not* predicted by the model of the defendant's trial selection process include relatively *high* values of the defendant's unit trial cost, K; or relatively *low* values of the parties' common expected trial sentence, S^T; or the idiosyncratic component, ρ, of the prosecutor's expected probability of trial conviction.

Continuing the preceding example, if a case that goes to trial is characterized by a relatively *low* value of C, then observed relatively *high* values of S^T or ρ, or relatively *low* values of K in such an initiated case would support the predictions of the model of the defendant's trial selection process. Meanwhile, if a case that goes to trial is characterized instead by a relatively *low* value of C, then observed relatively *high or low* values of K, S^T, or ρ in such a case that goes to trial also would be consistent with the predictions of the model of the defendant's trial selection process.

These predictions of the model of the defendant's plea decision suggest firm relationships among the relative values of different pairs of case characteristics either in the sample of cases that go to trial or in the sample of cases that do not go to trial (i.e., along the *same* game tree branch stemming from the plea decision node). These relationships can be tested meaningfully with sample means tests using either the sample of cases that go to trial or the sample of cases that do not go to trial (or both).

Table 6 provides the combinations of relative values of pairwise combinations of case characteristics in the sample of cases that *go to trial* that are consistent and inconsistent with the predictions implied by the model of the defendant's trial selection process.[65] Importantly, if the combinations that are inconsistent with the predictions of the model of the defendant's trial selection process for the sample of cases that *go to trial* are observed in the sample of cases that do not go to trial, then those observed combinations in the sample of cases that *do not go to trial* provide support for the predictions of the trial selection model. That is, if a combination of case characteristic values *is not* conducive for a defendant choosing to go to trial, then that combination *is* conducive for a defendant choosing *not* to go to trial. Thus, all (statistically significant) observed combinations of case characteristic values in the sample of initiated cases can be used to test the implications of the trial selection model. Table 6 serves as the basis of sample means tests reported in Chapter X.

[65] The predictions in Table 6 assume the prosecutor is "conviction-motivated" and has weaker diminishing returns in trial expenditure than the defendant.

Table 6. Pairs of relative values of case characteristics in cases that go to trial and their consistency with the predictions of the trial selection model

	High C	Low C	High K	Low K	High S^T	Low S^T	High ρ	Low ρ
High C	I*							
Low C	C*	C*						
High K	I	C	I*					
Low K	C	C	C*	C*				
High S^T	C	C	C	C	C*			
Low S^T	I	C	I	C	C*	I*		
High ρ	C	C	C	C	C	C	C*	
Low ρ	I	C	I	C	C	I	C*	I*

* As explained in Chapter X, different empirical variables can represent the same theoretical variable. Thus, combinations of relatively high and low values of the same theoretical variable represented by more than one empirical variable is possible for a given empirical observation. Pairs that are consistent with plea bargaining model are denoted with "C" and pairs that are inconsistent are denoted with "I". Horizontal axis provides the possible relative values of first case characteristic in a possible pair. Vertical axis provides the possible relative values of second case characteristic in a possible pair.

Other empirical tests of the implications of the model of the defendant's plea decision are possible. The theoretical model of the defendant's plea decision implies firm relationships among the relative values of the same case characteristic in the sample of cases that go to trial versus the sample of cases that do not go to trial (i.e., across *different* game tree branches stemming from the plea decision mode). These relationships can be tested meaningfully with sample means tests using the samples of (initiated) cases that do and do not go to trial.

Specifically, cases that go to trial (do not go to trial) are expected to be characterized by a relatively low (high) value of the prosecutor's unit cost of trial expenditure, C, and the defendant's unit trial cost, K. Meanwhile, cases that go to trial (do not go to trial) are expected to be characterized by a relatively high (low) value of the parties' common expected trial sentence, S^T; and the idiosyncratic component, ρ, of the prosecutor's expected probability of trial conviction.

CHAPTER VIII

DETERMINANTS OF OBSERVED TRIAL SENTENCES

This chapter defines the conditional *ex post* expected sentence imposed following a trial conviction, $\widetilde{S}^T$, which is the observed trial sentence. Observing a case in which a trial sentence is imposed requires the prosecutor to bring the case, the case to go to trial, and the trial to end in a conviction. The processes that determine the prosecutor's case selection, the defendant's plea decision (i.e., trial selection), and the judge's (or jury's) trial conviction decision thus condition the sample of initiated cases that go to trial and end in conviction at trial. Accordingly, these processes create case selection, trial selection, and trial conviction selection effects, respectively, which influence the observed trial sentence.

Consideration of these indirect selection effects is necessary and useful for empirical tests of the determinants of observed trial sentences in federal criminal antitrust cases. Consistent with the analyses in the previous chapters of this study, the predicted indirect case selection effects, trial selection effects, and trial conviction selection effects are generally ambiguous and render ambiguous the predicted total (net) marginal effects on the observed trial sentences from marginal increases in the values of case characteristics. Meaningful empirical tests similar to those described above, however, are possible in this case as well.

Direct Effects on Observed Trial Sentences

It is important to recognize that a trial sentence is observed only if the case is initiated, the case goes to trial, and the defendant is convicted at trial. The expected value of the observed trial sentence can be expressed as the following conditional *ex post* expected value:

$$E\left(\widetilde{S}^T\right) = E\left[S^T \left| \begin{array}{l} E(V) \geq 0, \\ (\rho - \delta)S^T - (CX + KY) > 0, \\ \mathbf{e} + P(X) + D(Y) > \lambda v \end{array} \right. \right], \qquad [28]$$

where, according to Equation [16],

$$S^T = m + \mu^T(\mathbf{e}, P(X), D(Y))(M - m). \qquad [16]$$

and,

$$E(V) = \Theta V^T + (1-\Theta)V^B$$

$$= \Theta(\Phi S^T - CX) + (1-\Theta)\left(\Phi S^T + \frac{KY - CX}{2}\right) \quad , \qquad [14]$$

$$= (\hat{\Pi} + \rho)S^T + \frac{KY - CX}{2} - \Theta\frac{KY + CX}{2}$$

where

$$V^T = \Phi S^T - CX \ , \qquad [4]$$

$$V^B = S^B = \Phi S^T + \frac{KY - CX}{2} = (\hat{\Pi} + \rho)S^T + \frac{KY - CX}{2}, \qquad [13]$$

$$\Theta = \Psi\left(\rho - \frac{CX + KY}{S^T}\right), \text{ and} \qquad [11]$$

$$\hat{\Pi} = \Pi([\mathbf{e} + P(X) - D(Y)]/\lambda). \qquad [15]$$

In addition, recall that the litigation model provides the optimal (Nash equilibrium) values of the prosecutor's and the defendant's respective choices of levels of trial expenditure, so that all of the *ex ante* expected values are evaluated with $X = X^*$ and $Y = Y^*$. Both X^* and Y^* are implicit functions of C, K, S^T, $\mathbf{e}$, λ, as well as each other as specified by the litigation model. As shown in Equation [28] the definition of the conditional *ex post* expected trial sentence draws from all of the elements of the case selection model of federal criminal antitrust case disposition, including the model of plea bargaining and the litigation model (with endogenous sentencing).

This section focuses on the effects of the case characteristics C, K, ρ, e, and λ on the first element of the right-hand side of Equation [28], the expected value of the *ex ante* expected probability of trial, $E\left(S^T\right)$, which equals S^T. A full discussion of the signs of the total marginal effects of the case characteristics (as well as *X* and *Y*) on S^T is found in Appendix A, as part of the examination of the effects of different factors on the prosecutor's *ex ante* expected value of bringing a case, *E*(*V*).

Briefly, the expected sign of the marginal direct effect on S^T and thus $\widetilde{S}^T$ from increases in the prosecutor's unit cost of trial expenditure, C, is negative (amibiguous) when the prosecutor (defendant) is the trial favorite. Further, the expected sign of the marginal direct effect on S^T and thus $\widetilde{S}^T$ from increases in the defendant's unit cost of trial expenditure, K, is positive (ambiguous) when the prosecutor (defendant) is the trial favorite.

An increase in the amount and/or strength of evidence against the defendant, e, causes a unidirectional change in the prosecutor's and defendant's optimal levels of trial expenditure, $X = X^*$ and $Y = Y^*$, respectively, which are implicit functions of e. This type of change in $X = X^*$ and $Y = Y^*$ has ambiguous marginal effects on S^T. Thus, the marginal direct effect on $\widetilde{S}^T$ from an increase in e is ambiguous.

The litigation model's predictions about the effects of the variability of the trial court's conviction standard, λ, on the parties' trial expenditures, *X* and *Y*, rely on the "closeness" of the parties' diminishing returns in trial expenditure. Without any reasonable means of accurately predicting (or measuring) this "closeness," the expected sign of the marginal direct effect on $\widetilde{S}^T$ from an increase in λ is ambiguous.

Neither the *ex ante* expected trial sentence, S^T, nor the parties' trial expenditures, *X* and *Y*, are functions of the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probabilty of trial conviction, $\hat{\Pi}$. Thus, the expected value of the marginal direct effect on $\widetilde{S}^T$ from an increase in ρ is zero.

Case Selection Effects on Observed Trial Sentences

Note that the first conditioning event expressed in Equation [28] is that the prosecutor brought the case. This conditioning event, represented by $E(V) \geq 0$, produces indirect "case selection effects" on the observed trial sentences that are analogous, but not identical, to the indirect case selection effects discussed in Chapters V and VI with regard to the observed trial rate and the observed trial conviction rate, respectively.

This analysis takes advantage of the fact that *E*(*V*) is a function of S^T, which simplifies determining the case selection effects on S^T from marginal increases in the values of case characteristics. Specifically, if the prosecutor initiates a case in which she has a relatively high unit trial cost, *C*, then the parties' common *ex ante* expected probability of trial conviction, S^T, in the initiated case is likely to be higher. Thus, the indirect cases selection effect on the observed trial conviction rate, $\widetilde{S}^T$, from higher unit costs of trial expenditure for the prosecutor is positive. That is because the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, unambiguously decreases in C when she is conviction-motivated and, in order to initiate such a case, other case characteristics such as S^T would need to raise $E(V)$.

On the other hand, if a conviction-motivated prosecutor initiates a case with relatively more or stronger evidence, e, then the initiated case is just as likely to be characterized by a relatively low value of S^T as it is likely to be characterized by a relatively high value of S^T. Thus, the expected case selection effect of more or stronger evidence on the observed trial sentence, $\widetilde{S}^T$, is ambiguous.

Similarly, since $E(V)$ increases in marginal increases in the values of the defendant's unit cost of trial expenditure, K, the parties commonly expected *ex ante* expected probability of trial conviction, $\hat{\Pi}$, and the idiosyncratic component, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, the expected indirect case selection effects on $\widetilde{S}^T$ from marginal increases in K, $\hat{\Pi}$, and ρ are ambiguous. These ambiguities do not preclude meaningful empirical tests of the model of the case selection process, however.

Meanwhile, the marginal effect on $E(V)$ of the variability of the judge's trial conviction standard, λ, is ambiguous. Thus, the expected sign of the indirect case selection effect on $\widetilde{S}^T$ from an increase in λ is also ambiguous.

Trial Selection Effects on Observed Trial Sentences

The second conditioning event on the *ex post* expected trial sentence, $\widetilde{S}^T$, is that the case went to trial, which is represented by $(\rho - \delta)S^T - CX - KY > 0$. This conditioning event depends directly on the *ex ante* expected trial sentence, S^T. The defendant's decision to go to trial shapes the mix of case characteristics in cases that go to trial, including S^T, which, in turn, affects the *ex post* expected trial sentence, $\widetilde{S}^T$.[66] The following abbreviated description of the trial selection effects on the *ex post* expected trial sentence, $\widetilde{S}^T$, from marginal increases in the values of case characteristics, takes advantage of the fact that the *ex ante* expected probability of trial, Θ, is a function of S^T.[67]

Suppose a case that is characterized by either a relatively high value of the unit cost of trial expenditure for the prosecutor, *C*, or a relatively high value of the unit cost of the

[66] As previously mentioned, in the sequence of selection processes and decisions that determine a prison sentence, the prosecutor decides to bring the case, the defendant decides which plea to enter, the prosecutor could accept that plea or object to that plea, as Cohen (1992) notes, and the judge decides whether to accept or reject that plea (possibly over the prosecutor's objection). Cohen (1992) does not find any statistically significant difference in *corporate fines* imposed across means of conviction – i.e., trial conviction, non-negotiated guilty plea, negotiated guilty plea, negotiated *nolo contendere* plea, and negotiated *nolo contendere* plea accepted by the judge over the government's objection. In contrast, Cohen finds that the type of plea is a significant determinant of prison sentencing. Elsewhere, however, Cohen notes that he found no evidence of selection bias when estimating sample selection models of sentencing decisions using the estimation technique introduced by Heckman (1976).

[67] It is worth mentioning here that the fact that the total (net) marginal effects of changes in the values of case characteristics have ambiguous effects on the *ex ante* expected probability of trial, Θ, has no bearing on the following evaluation of the indirect trial selection effects of case characteristics on $\widetilde{S}^T$. That is because the evaluation of the marginal effects of changes in the values of case characteristics on Θ is done from an *ex ante* perspective (e.g., when evaluating the effects of case characteristics on Θ as they relate to $E(V)$ or $\widetilde{\Theta}$, as done in Appendix A and Chapter VI, respectively). In contrast, the evaluation of the trial selection effect on the observed trial sentence from changes in the values of case characteristics is performed from an *ex post* perspective.

In an *ex post* evaluation of the trial selection effects, all that matters are the relative values of case characteristics that, when taken together or at least in pairs, *ceteris paribus*, are expected to satisfy the condition for a case to go to trial, $(\rho - \delta)S^T - CX - KY > 0$. Once a case has gone to trial, the values of the case characteristics (in particular, the endogenous variables $X = X^*$ and $Y = Y^*$ that cause the "trial expenditure effects" on Θ) upon which the defendant relied in making his plea decision (and upon which the prosecutor relied in her case-bringing decision) are "frozen."

defendant's trial expenditure, *K* – or by relatively high values of *C* and *K* – goes to trial. The model of the defendant's plea decision predicts that such a case would be further characterized by a relatively high value of S^T. Thus, the indirect trial selection effects on the *ex post* expected trial sentence, $\tilde{S}^T$, from marginal increases in *C* or *K* are positive.

Now suppose a case that goes to trial is characterized by a relatively high value of the idiosyncratic component, ρ, of the prosecutor's *ex ante* expected probability of trial conviction. In that circumstance, the case that goes to trial is just as likely to be characterized by a relatively low value of the *ex ante* commonly expected trial sentence, S^T, as it is likely to be characterized by a relatively high value of S^T. Thus, the predicted indirect trial selection effect on the *ex post* expected trial sentence, $\tilde{S}^T$, from a marginal increase in ρ is ambiguous. Using the same logic (because the *ex post* expected probability of trial, $\tilde{\Theta}$, increases in S^T), the predicted indirect trial selection effect on the *ex post* expected trial sentence, $\tilde{S}^T$, from a marginal increase in the value of the *ex ante* expected trial sentence, S^T, is ambiguous.

Trial Conviction Selection Effects on Observed Trial Sentences

The judge's (or jury's) trial conviction decision censors the sample of cases that proceed to sentencing following a trial conviction. Thus, the trial conviction decision creates selection (bias) effects associated with increases in the values of case characteristics on the observed sentences following conviction at trial. The mechanics of the trial conviction selection effects from changes in the values of case characteristics are analogous to the mechanics of the case selection effects and trial selection effects that have already been discussed.

For example, the litigation model predicts that a case characterized by a relatively high value of the unit cost of trial expenditure for the prosecutor, *C*, is likely to have a lower *ex ante* expected probability of trial conviction, $\hat{\Pi}$, and thus a lower *ex post* expected probability of trial conviction, $\tilde{\Pi}$. Accordingly, if a case characterized by a high value of *C* results in a trial conviction, then, *ceteris paribus*, that case is also likely to be characterized by a high (low) value of the *ex ante* expected trial sentence, S^T, when the prosecutor (defendant) has weaker diminishing returns from trial expenditure. That is because such relative values of S^T under such circumstances are expected to increase the value of $\hat{\Pi}$ (and $\tilde{\Pi}$) according to the litigation model. Thus, the expected indirect trial conviction selection effect on the *ex post* expected trial sentence, $\tilde{S}^T$, from an increase in the value of *C* is positive (negative) when the prosecutor (defendant) has weaker diminishing returns in trial expenditure.

On the other hand, if a case characterized by a relatively high value of *K* results in a trial conviction, then that case is just as likely to be characterized by a relatively low value of the *ex ante* expected trial sentence, S^T, as it is likely to be characterized by a relatively high value of S^T. That is because $\hat{\Pi}$ (and $\tilde{\Pi}$) increases in *K*. Similarly, if a case characterized by a high value of e results in a trial conviction, then, *ceteris paribus*, that case is also just as likely to be characterized by a relatively low value of the *ex ante* expected trial sentence, S^T, as it is likely to be characterized by a relatively high value of S^T.

In contrast, the marginal effect on the *ex ante* expected probability of trial conviction from an increase in the variability of the judge's trial conviction decision standard, λ, is ambiguous. Thus, the expected sign of the trial conviction selection effect on $\tilde{S}^T$ from an increase in λ is ambiguous.

Finally, the idiosyncratic component, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, influences neither the parties' levels of trial expenditure nor the *ex ante* expected probability of trial conviction. Thus, increases in ρ do not create any trial conviction selection effects on $\tilde{S}^T$.

Empirically Testable Implications of Trial Conviction as a Selection Process

Table 7 provides the theoretical implications of the judge's sentencing decision following a trial conviction with respect to the influence of case characteristics on observed trial sentences.[68] The theoretical implications summarized in Table 7 apply to empirical tests consisting of regression analysis to identify the determinants of observed trial sentences. Such regression analysis would need to control for sample selection bias from the plea decision (i.e.

---

68 The predictions in Table 7 assume the prosecutor is "conviction-motivated" and has weaker diminishing returns in trial expenditure than the defendant.

trial selection) and the trial conviction processes. Traditional regression techniques do not allow controlling for the sample selection bias created by the prosecutor's case selection process.

The predictions shown in Table 7 are useful for comparisons to Cohen's (1992) analysis of the determinants of sentences imposed in federal criminal antitrust cases. As previously mentioned, Cohen recognizes that the defendant's plea decision and the trial conviction decision could influence sentencing outcomes. He does not discuss the potential effects on observed sentences from the prosecutor's decision to initiate particular cases, however.

Table 7. Signs of expected direct and indirect effects on the observed trial sentencing decision, $\tilde{S}^T$

Case Characteristic	Expected Direct Effect on $\tilde{S}^T$ *		Expected Indirect Case Selection Effect on $\tilde{S}^T$	Expected Indirect Trial Selection Effect on $\tilde{S}^T$	Expected Indirect Trial Conviction Selection Effect on $\tilde{S}^T$	Expected Total (Net) Effect on $\tilde{S}^T$
	When $\hat{\Pi} \geq \overline{\Pi}$	When $\hat{\Pi} < \overline{\Pi}$				
$\hat{\Pi}$ **	0		?	0	?	?
C	-	?	+	+	+	?
K	+	?	?	+	?	?
S^T **	+		?	?	?	?
e	?		?	?	?	?
λ	?		?	0	?	?
ρ	0		?	?	0	?

* The expected direct effects include first- and second-order effects, as discussed in Appendix A.

** The marginal increase in $\hat{\Pi}$ and S^T are from exogenous shocks.

Despite the ambiguous expected indirect trial conviction selection effects on observed trial sentences shown in Table 7, the theoretical implications of the trial conviction selection process, as specified in the model of the rule by which the judge (or jury) decides to convict a defendant after trial, are empirically testable. Table 8 summarizes the predicted relationships among the relative values of different pairs of case characteristics in cases that end in trial conviction. To understand Table 8, first consider the relationships among the relative values of different pairs of case characteristics for cases ending in trial conviction that are *not* predicted by the model.

For example, suppose a case that is disposed by a trial conviction is characterized by a relatively *high* value of the prosecutor's unit cost of trial expenditure, C. *Ceteris paribus*, a litigated case that ends in trial conviction is not expected to be associated with a relatively *high* value of C, since C is expected to *negatively* directly affect the *ex post* probability of trial conviction, $\tilde{\Pi}$.

In that situation, the model of the judge's trial conviction decision rule (i.e., trial conviction selection process) does *not* predict that the case that *ends in a trial conviction* also is characterized by values of other case characteristics that are expected to *further reduce* $\tilde{\Pi}$. In this situation, relative values of other case characteristics *not* predicted by the model of the trial conviction selection process include relatively *low* values of the defendant's unit trial cost, K; or relatively *low* values of the parties' common expected trial sentence, S^T; (assuming the prosecutor has weaker diminishing returns in trial expenditure than the defendant), or the amount or strength of the prosecutor's evidence, e.

Continuing the preceding example, if a case that is disposed by trial conviction is characterized by a relatively *high* value of C, then observed relatively *high* values of S^T, e, or K in such a case would be consistent with the predictions of the model of the judge's trial conviction decision rule. Meanwhile, if a case that ends in trial conviction is characterized instead by a relatively *low* value of C, then observed relatively *high or low* values of K, S^T, or e in such a case also would be consistent with the predictions of the model of the trial conviction decision rule.

These predictions of the model of the defendant's plea decision suggest firm relationships among the relative values of different pairs of case characteristics either in the sample of cases that end in trial conviction or in the sample of cases that go to trial but do not end in trial conviction (i.e., along the *same* game tree branch stemming from the trial conviction decision node). These relationships can be tested meaningfully using sample means tests using either the sample of cases that end in trial conviction or the sample of cases that go to trial but do not end in trial conviction (or both).

Table 8 provides the combinations of relative values of pairwise combinations of case characteristics for the sample of cases *ending in trial conviction* that are consistent and inconsistent with the predictions implied by the model of the trial conviction selection process.[69] Note well that if the combinations that are inconsistent with the predictions of the model of the judicial trial conviction decision for the sample of cases that *are disposed by trial conviction* are observed in the sample of cases that *go to trial but do not end in trial conviction*, then those observed combinations in the sample of cases that *go to trial but do not end in trial conviction* provide support for the predictions of the model of the judicial trial conviction decision. That is, if a combination of case characteristic values *is not* conducive for a judge *convicting* a defendant at trial, then that combination *is* conducive for a judge *acquitting* that defendant at trial (or dismissing the case). Thus, all (statistically significant) observed combinations of case characteristic values in the sample of litigated cases can be used to test the implications of the trial conviction selection model. Table 8 serves as the basis of sample means tests reported in Chapter X.

[69] The predictions in Table 8 assume the prosecutor is "conviction-motivated" and has weaker diminishing returns in trial expenditure than the defendant.

Table 8. Pairs of relative values of case characteristics in cases that end in trial conviction and their consistency with the predictions of the trial conviction model

	High C	Low C	High K	Low K	High S^T	Low S^T	High **e**	Low **e**
High C	I*							
Low C	C*	C*						
High K	C	C	C*					
Low K	I	C	C*	I*				
High S^T	C	C	C	C	C*			
Low S^T	I	C	C	I	C*	I*		
High **e**	C	C	C	C	C	C	C*	
Low **e**	I	C	C	I	C	I	C*	I*

* As explained in Chapter X, different empirical variables can represent the same theoretical variable. Thus, combinations of relatively high and low values of the same theoretical variable represented by more than one empirical variable is possible for a given empirical observation. Pairs that are consistent with plea bargaining model are denoted with "C" and pairs that are inconsistent are denoted with "I". Horizontal axis provides the possible relative values of first case characteristic in a possible pair. Vertical axis provides the possible relative values of second case characteristic in a possible pair.

Other empirical tests of the implications of the model of the trial conviction decision are possible. The theoretical model of the trial conviction decision implies firm relationships among the relative values of the same case characteristic in the sample of cases ending in trial conviction versus the sample of cases that go to trial but do not end in conviction (i.e., across *different* game tree branches stemming from the trial decision node). These relationships can be tested meaningfully using sample means tests using the samples of (initiated) cases that end in trial conviction and (versus) cases that go to trial but do not end in conviction.

Specifically, assuming the prosecutor has weaker diminishing returns in trial expenditure than the defendant, cases that end in trial conviction (go to trial but end in acquittal, dismissal, or directed verdict in favor of the defendant) are expected to be characterized by a relatively low (high) value of the prosecutor's unit cost of trial expenditure, C. Meanwhile, cases that end in trial conviction (go to trial but do not end in conviction) are expected to be characterized by a relatively high (low) value of the defendant's unit trial cost, K, the parties' common expected trial sentence, S^T; and the amount or strength of the prosecutor's evidence, e.

CHAPTER IX

INFORMATION STRUCTURE OF PLEA BARGAINING

As highlighted in the literature review in Chapter II, the criminal antitrust literature ignores (or does not explicitly address) the issue of how private information in plea bargaining could alter the trial selection process and create (additional) bias in observed case outcomes. This is a significant omission given the policy implications of private information during plea bargaining.

Privately informed prosecutors harm social welfare because they can fashion a settlement offer that induces (truly) innocent defendants to plead guilty or *nolo contendere*. Such pleas send false signals about legal and socially efficient business conduct. In contrast, privately informed defendants during plea bargaining benefit social welfare because the (truly) guilty ones are more likely to plead guilty and send the correct signals about appropriate business conduct.

This section explains the implications of potential indirect case selection effects on plea decisions and trial outcomes, as well as potential indirect trial selection effects on trial outcomes, in a commonly used empirical test for the existence and type of private information during plea bargaining (or settlement negotiations in civil litigation). As explained below, the potential for indirect case selection effects and trial selection effects could confound and render meaningless the correlation between trial rates and trial conviction rates, which is the empirical test commonly used to identify the information structure of plea bargaining (settlement negotiations).

This chapter concludes by describing five alternative empirical testing approaches that avoid such problems. Four of the alternative approaches are suggested by Hylton (1993) and one is contributed by this study. It is important to empirically assess the underlying information structure of plea bargaining in the observed federal criminal antitrust cases used to test the theoretical predictions developed in this study. That is because the models of the defendant's plea decision, as well as the prosecutor's *ex ante* expected value of the sentence following a plea of guilty or *nolo contendere*, assume that the defendant is privately informed about his chances at trial.

Information Structure of Plea Bargaining in This Study's Theoretical Framework

As discussed in Chapter IV, the models of the defendant's plea decision, as well as the prosecutor's *ex ante* expected value of the sentence following a plea of guilty or *nolo contendere*, assume that the defendant is privately informed about his chances at trial. This is a fundamental assumption. Thus, it is important to test empirically the validity of the assumption.

If it were not appropriate to assume that during plea bargaining the defendant is privately informed about his chances at trial, then the theoretical framework used in this study would need to be changed significantly. For example, an incentive compatibility constraint that defines the probability of trial would still be relevant, but it likely would involve the prosecutor's objective(s) (e.g., conviction rate maximization, trial cost minimization). Accordingly, a privately informed prosecutor could set the value of the plea negotiated sentence just low enough to induce a guilty or *nolo contendere* plea. In this study's theoretical framework, the value of the negotiated penalty would not be the Nash equilibrium solution to a non-cooperative plea bargaining game. Instead, for example, the Nash equilibrium solution could be the choice-variable solution to the prosecutor's maximization of some objective.

Existing Empirical Tests of Information Structure of Plea Bargaining

According to the asymmetric/private information theory developed by Bebchuck (1984) and others, including Froeb (1993), one party knows the probability that the prosecutor will win at trial, while the other knows only the distribution of plaintiff victory probabilities. When the defendant is better or privately informed, he will accept the (uninformed) prosecutor's settlement offer if he is relatively pessimistic about his chances. Meanwhile, the better or privately informed defendant will go to trial if he (correctly) expects to win at trial. Under this theory, the selection of cases for trial is one-sided, and the *ex post* probability of a prosecutor's trial victory is systematically below the fraction of trial convictions in the pool of initiated cases.

Previous empirical tests for the existence of private information (and which litigant possesses private information) in plea bargaining, including the test used by Froeb (1993), focus mainly on the correlation between trial rates and plaintiff victories at trial. In general, existing studies posit that a positive (negative) correlation between trial rates and prosecutors' victories at trial is evidence of privately informed defendants (prosecutors) during settlement negotiations.

With respect to testing for trial selection bias from plea bargaining, Froeb (1993) observes that determining whether only the defendant or only the prosecutor is privately informed during plea bargaining is a simple matter of testing whether the trial conviction rate of cases that go to trial were lower than the trial conviction rate of cases that do not go to trial. The latter statistic does not exist, however, rendering this approach infeasible.

Froeb devises an alternative empirical testing strategy to determine what type of trial selection process underlies observed case outcomes. He illustrates his approach by referring to a specific plea bargaining game in which a group of identically charged defendants receives the same take-it-or-leave-it plea offer from the prosecutor. If these defendants are privately informed about their private information, those with higher probabilities of trial conviction will accept the offer, while those with lower probabilities of trial conviction will reject it.

Now suppose the prosecutor offers a harsher sentence in exchange for a guilty plea. Some defendants with poor chances at trial who had previously accepted the more lenient offer will reject the harsher offer and go to trial. This increases the mean probability of trial conviction among the sample of cases that go to trial, which creates positive correlation between observed trial rates and observed trial conviction rates. Conversely, according to Froeb's testing strategy, if plea bargaining exhibits reverse adverse selection, then observed trial and trial conviction rates are negatively correlated. Pooling (e.g., both parties privately informed), meanwhile, would imply that there is no correlation.

Froeb further argues that this testing strategy does not depend on the specification of plea bargaining according to a take-it-or-leave-it model. According to Froeb, this testing strategy only requires that changes in trial rates are exogenous and sufficiently large. He points to evidence of adverse trial selection from plea bargaining suggested by increases in trial rates and associated increases in trial conviction rates in Alaska and Arizona when plea bargaining was banned. Froeb notes that these examples provide useful natural experiments because the changes in trial rates were probably exogenous, since the plea bargaining bans were likely the result of a change in the local political climates.

In addition, using federal court data for various crimes, including white-collar crimes other than antitrust offenses, Froeb estimates cross-district and cross-crime regression models and finds a positive correlation between trial and trial conviction rates. That is, these regressions suggest adverse selection in plea bargaining. Tax fraud cases are the only sample for which Froeb's time-series regressions provide evidence of adverse selection in plea bargaining. For all other types of cases, his time-series regressions suggest pooling (i.e., no correlation between trial and trial conviction rates).

Existing studies of federal criminal antitrust case outcomes find that following an increase in the trial rate, the trial conviction rate fell. Specifically, Snyder (1990) finds, for HPSV cases brought during the period 1970-1985, the applicability of felony penalties increases the (*ex post*) probability of trial for corporate and individual defendants. Meanwhile, Snyder (1990) finds that giving HPSV offenses felony status lowered trial conviction rates.

In comparison, the available empirical evidence suggests that the Antitrust Guideline had some impact on the defendants' choices between *nolo contendere* and guilty pleas. Gallo, *et al*. (1996) report that the percentage of convicted defendants who plead *nolo contendere* fell from 10 percent during 1985 to 1989 to 2 percent during 1990 to 1994. Meanwhile, cases disposed by guilty pleas rose from 89 percent of all convictions during 1985-89 to 98 percent during 1990-94.[70] Gallo *et al*. (1996) do not report trial rates separately. Nonetheless, the changes in the relative percentages of cases disposed by *nolo contendere* or guilty pleas that are reported by Gallo *et al*. (1996) suggest a decrease in trial rates following the Antitrust Guideline.[71] Meanwhile, Gallo, *et al*. (1996) reports a dramatic decrease in the overall conviction rate from 97 percent of all criminal antitrust cases in 1991 to 2 percent in 1992. Moreover, DOJ prosecutors were winless in 1993 and won only 12 percent of all criminal antitrust cases in 1994.[72] From this evidence, it is reasonable to infer that after the introduction of the Antitrust Guideline, both the observed trial rate and trial conviction rate fell.

Based on the (inferred) trends of trial and trial conviction rates in federal criminal antitrust cases following the passage of the APPA, a direct application of Froeb's test would suggest that plea bargaining in HPSV cases creates reverse adverse trial selection (i.e.,

---

[70] As the authors note, these observations may reflect the government's increasing opposition to the *nolo contendere* plea.

[71] Gallo *et al*. (2000) provide other support for the notion that trial rates declined following the introduction of the Antitrust Guideline. They report that the supporting data for their tables show that the percentage of not guilty verdicts to total verdicts fell from slightly more than 1 percent during 1980-1984 to less than 0.13 percent during 1985-1997.

[72] See Gallo, *et al*. (1996), Table XV. This table, which presents the annual "win rates" in criminal antitrust cases, apparently defines a win as any type of conviction at the case-level.

prosecutor's possess private information during plea bargaining). In contrast, the inferred increase in the plea rate and decrease in the overall conviction rate in federal criminal antitrust cases following the introduction of the Antitrust Guideline suggests adverse selection in plea bargaining (i.e., defendants possess private information during plea bargaining).[73]

The Prosecutor's Case Selection Process and Empirical Tests of the Information Structure of Plea Bargaining

The above discussion of trial selection bias from private information during plea bargaining in federal criminal antitrust cases is not necessarily persuasive. Although Froeb's reduced form model of trial selection is (Nash) consistent with this study's expanded form case selection model, Froeb's reduced form model does not account for indirect case selection effects on the defendant's plea decision or the judge's (or jury's) decision to convict a defendant at trial. The persuasiveness of the anecdotal and empirical evidence cited by Froeb (1993) therefore depends on an (implicit) assumption that the prosecutor randomly chooses cases to initiate.

It is preferable to avoid assuming (effectively) that prosecutors do not control which federal criminal antitrust cases are initiated. In fact, it is reasonable to expect that DOJ prosecutors are very deliberate in their decisions to initiate these cases, usually after gaining familiarity with many characteristics of potential cases.[74] Thus, case seleciton effects (in addition to trial selection effects) should be considered in analyses of the determinants of federal criminal antitrust case outcomes, including tests for trial selection bias in plea bargaining.

Froeb recognizes that empirically testing for trial selection bias is complicated because potentially unobserved factors such as prosecutors' case-bringing decisions could induce spurious covariation between trial and trial conviction rates. Citing Flanders (1976), Froeb notes that DOJ prosecutors initiate various types of cases at various rates depending on their views of

---

[73] The expected effects of the Antitrust Guideline on case disposition are beyond the scope of this study. Nonetheless, it is worth noting that the academic concerns regarding mandatory guideline sentencing (the guidelines are no longer mandatory), as implemented by the United States Sentencing Commission, focused on the asymmetric bargaining power that such guidelines provided prosecutors. The particular concern was that the guidelines, as implemented, were based on the charges in the indictment instead of the counts of conviction. Since the prosecutor better knows what charges she will file against the defendant, the sentencing guidelines, as implemented, provided the prosecutor with private information during plea bargaining. Thus, the finding inferred from the statistics reported by Gallo *et al.*, that the trial rate and trial conviction rates were positively correlated after the introduction of the Antitrust Guideline contradicts the general academic prediction.

[74] Likewise, it is reasonable to expect that defendants in federal criminal antitrust cases – especially corporate defendants – make non-random decisions about whether or not to go to trial based on case characteristics.

enforcement needs, resources, possible problems in state law enforcement, or DOJ policy and other organizational pressures. Froeb also notes that federal prosecutors' case-bringing decisions depend on the rates at which various federal crimes are committed. To the extent that crime rates depend on deterrence efforts, they, too, may not be randomly determined.

As Froeb explains, if the determinants of case-bringing decisions or other potentially unobserved factors induce spurious correlation between trial rates and trial conviction rates, then the identification of adverse selection or reverse adverse selection from plea bargaining would require a policy instrument that shifts the probability of trial. For example, Landes's (1971) theory of prosecutorial behavior suggests that a (binding) budget constraint could serve as such a policy instrument. [75]

Froeb was unable to find a policy instrument that shifted the trial rate in his samples of cases. Thus he warns that his regression results should be interpreted with care because he cannot rule out spurious correlations between trial and trial conviction rates in his data. With respect to federal criminal antitrust cases for the period 1955-1980, several policy instruments applied by Congress (e.g., DOJ budget appropriations, statutory sentencing reform), federal judges (e.g., the development of the doctrine of the *per se* illegality of horizontal price-fixing), and federal prosecutors (e.g., limits on plea agreements involving *nolo contendere* pleas) could shift the trial rate over time.

The theoretical framework developed in this study demonstrates how policy or institutional shifts in federal criminal antitrust enforcement are expected to influence the relative values of certain characteristics of the cases that prosecutors consider initiating. In turn, changes in the values of case characteristics are expected to directly impact observed plea and trial outcomes. Changes in the values of case characteristics also are expected to produce indirect case selection effects (and trial selection effects) on plea and trial outcomes that could mute or even reverse the signs of the direct effects. In addition, as explained above, the predicted indirect case selection effects on observed trial rates and trial conviction rates, as well as the predicted indirect trial selection effects on observed trial conviction rates, are generally ambiguous.

---

[75] For his sample of cases, Froeb (1993) asserts that prosecutorial budgets change slowly, perhaps with a sufficiently long lag that trial rates can be treated as exogenous. While this assertion may be true for the DOJ as an entire law enforcement agency, it may not be true for particular divisions of the DOJ like the Antitrust Division. The descriptive statistics reported in Appendix B suggest that Froeb's assertion may not apply to the DOJ Antitrust Division.

If policy instruments or other institutional factors do not create the expected exogenous shifts in the trial rate – that is, if the trial rate shifts for unexplained reasons – then, consistent with Froeb's concerns, correlations between observed trial rates and trial conviction rates could be spurious. Thus, sole reliance on correlations between observed trial rates and trial conviction rates is not sufficient for identifying private information in plea bargaining if case selection is non-random and/or correlated with observed trial rates or trial conviction rates. For example, the mere observation that trial and trial conviction rates are negatively correlated following the passage of the APPA does not necessarily lend support for the existence of reverse adverse trial selection in federal criminal antitrust cases.

Alternative Approaches to Empirical Tests of the Information Structure of Plea Bargaining

This study uses alternatives to the empirical testing methodology employed by Froeb (1993) in the criminal litigation context and others in the civil litigation context. The methods of Froeb and others focus on the correlation between trial rates and trial conviction rates for identifying the existence and type of private information during plea bargaining. Alternative testing approaches are desirable because of the potentially confounding effect of the (non-random) case selection and trial selection processes, as explained above.

Hylton (1993) suggests several alternative approaches to testing for private information in plea bargaining that do not rely on the correlation between the trial rate and trial conviction rate. First, Hylton explains that the distance between the expected liability (expected penalty and trial cost) of an innocent defendant and the expected liability of a guilty defendant increases as the probabilities of type I and type II legal errors[76] regarding the defendant's compliance with a legal standard fall. This *increases* the zone of acceptable plea offers (i.e., penalties that defendants would accept by pleading guilty) for guilty defendants and *decreases* the zone of acceptable plea offers for innocent defendants. Under these circumstances, it follows that the proportion of innocent defendants who litigate likely increases.

Thus, assuming the probabilities of type I and type II legal errors regarding defendants' compliance with legal standards are low (e.g., less than 50 percent, according to Hylton) and

[76] Type I legal errors occur when truly guilty defendants are found not guilty and type II errors occur when truly innocent defendants are found guilty.

roughly equal, if defendants have private information regarding their chances at trial, then the trial conviction rate will be below 50 percent. That is, when legal errors are relatively unlikely, a trial conviction rate below 50 percent is evidence that defendants are privately informed during plea bargaining about their chances at trial. Conversely, a trial conviction rate above 50 percent is evidence that prosecutors possess private information during plea bargaining. A trial conviction rate of 50 percent is evidence that neither party holds private information. According to Hylton, for a given sample of defendants, comparing the value of the trial conviction rate to the 50 percent mark is the best test of the informational structure of plea bargaining.

Hylton also explains hypotheses regarding changes in the trial conviction rate as legal doctrine develops or as time passes[77] that can be used to test for the information structure of plea bargaining. He posits that if the development of legal doctrine over time reduces the likelihood of (type I and type II) judicial error, then the trial rate and, separately, the prosecutors' win rate at trial will fall over time if plea bargaining (i.e., trial selection) operates as expected when defendants are privately informed during plea bargaining about their chances at trial.[78]

Since its enactment in 1890, the Sherman Act's definitions of illegal business conduct have not changed. Still, legal doctrine regarding applications of the Sherman Act developed over time – especially with Supreme Court decisions that confirm, refine, contradict, or overturn precedent.

For example, as discussed previously in Chapter II, antitrust case law has come to view certain cartel activities like price-fixing and bid-rigging as horizontal *per se* violations ("HPSVs") of Section 1 of the Sherman Act. According to former Assistant Attorney General for Antitrust Donald I. Baker (1978), the *Trenton Potteries* case in 1927 clearly established the *per se* illegality of price-fixing. The Supreme Court's decision in the *Appalachian Coal* case in 1933 created some confusion regarding the *Trenton Potteries* decision. Then the *Socony-Vacuum Oil* case in 1940 firmly reestablished the *per se* illegality of price-fixing.

Also according to Baker (1978), federal courts effectively recognize three categories of antitrust cases. The first category includes "hard core" *per se* cases involving price fixing and

[77] Hylton (1993) uses the passage of time as a proxy for the development of legal doctrine. In Chapter X, this study uses separate measures of the passage of time and the incremental development of legal doctrine to apply Hylton's tests.

[78] Hylton (1993) expects defendants in civil antitrust cases to possess better information about their chances at trial than plaintiffs possess. In support of this notion, Hylton notes that Salop and White (1988) find that the annual average plaintiff win rate in civil antitrust cases has fallen over time.

market allocation.[79] The second category includes “soft core” *per se* cases in which courts consider facts surrounding conduct that is technically subject to per se rules, such as various vertical arrangements and joint ventures.[80] The third category includes cases in which the “rule of reason” applies and all facts are relevant.[81]

In contrast to the lower courts, however, Baker (1978) notes that the Supreme Court has consistently rejected the creation of the middle category of “soft core” offenses that are *per se* illegal, even though the Supreme Court has changed its opinion about certain types of conduct. For example, in United States v. Arnold, Schwinn & Co., 388 U.S. 365 (1967), the Supreme Court ruled that territorial restrictions are *per se* illegal. Then in Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36 (1977), the Supreme Court refined its position and held that some territorial restraints are not *per se* illegal.[82]

This legal history suggests that the development of legal doctrine over time with respect to application of Section 1 of the Sherman Act has clarified what business conduct is legal and illegal and has reduced judicial error in HPSV cases. Thus, if defendants (alternatively, prosecutors) possess private information during plea bargaining in HPSV cases, the trial rate should decrease (increase), over time or as antitrust legal doctrine develops incrementally. In addition, the trial conviction rate should decrease (increase), over time or as antitrust legal doctine develops, if defendants (prosecutors) possess private information during plea bargaining.

Gallo *et al*. (2000) identify the years in which the U.S. Supreme Court made decisions that various legal scholars categorize, though not unanimously, as “landmark” decisions. This study uses these identified decisions to create an empirical variable that controls for the incremental development of federal criminal antitrust legal doctrine. Observed plea decisions and trial outcomes are regressed on explanatory variables representing time and the issuance of landmark antitrust court decisions. Consistent with Hylton’s hypotheses, negative estimated coefficients on the variables representing time or the issuance of landmark antitrust court decisions would imply that, during plea bargaining, defendants are privately informed about their chances at trial.

---

79 See e.g., United States v. Socony-Vacuum Oil Co. 310 U.S. 150 (1940).

80 See e.g., Fortner Enterprises v. United States Steel Corp., 394 U.S. 495 (1969). See also Worthern Bank & Trust Co. v. National Bank-Americard Inc., 485 F.2d 119 (8th Cir. 1973).

81 See e.g., Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36 (1977).

82 Echoing Posner’s (1970) concerns, Baker (1978) also observes that the Supreme Court has greatly expanded the types of business conduct that is *per se* illegal, even conduct that arguably enhances competition.

Using the (statistically signficant) estimated regression coefficients on variables representing the passage of time and the incremental development of antitrust legal doctrine to test for the existence and type of private information during plea bargaining avoids the potentially confounding effects (for purposes of empirical testing) of the case selection and trial selection processes. *Ceteris paribus*, these factors are not expected *a priori* to effect negatively or positively the prosecutor's *ex ante* expected value of initiating a case or the defendant's *ex ante* expected value of going to trial.

That is, the reduction of legal error (or the variability of the universe of potential judges' or juries' applications of law) simply "clears the air" and cannot be relied upon by prosecutors in their *ex ante* evaluation of the benefits of initiating a case, as represented by the model of the prosecutor's case-bringing decision, or by defendants in their *ex ante* evaluation of the benefits of going to trial, as represented by the model of the defendant's plea decision. The "clarified air" created by the development of legal doctrine over time simply allows a potentially privately informed litigant to employ more effectively (or more distinctly, for purposes of empirical testing) his or her assessment of the defendant's chances at trial, given the case characteristics represented in the theoretical framework.[83]

In addition to his hypotheses regarding the overall trial conviction rate in a sample of cases and changes in the trial conviction rate as legal doctrine develops as time passes, Hylton provides a related hypothesis: The trial rate will fall as legal doctrine develops over time (thus reducing type I and type II legal errors) when defendants are privately informed during plea negotiations about their chances at trial. Hylton does not explain this particular hypothesis. Nonetheless, this hypothesis follows from the reasonable assumption that, because prosecutors gather evidence against defendants before indicting defendants, prosecutors do not indict defendants that the prosecutors believe to be innocent. Given that assumption, because the reduction of legal error with the development of legal doctrine over time increases the zone of acceptable plea offers to guilty defendants, fewer indicted defendants (whom the prosecutors think are guilty based on evidence) will reject plea offers and go to trial.

[83] This reasoning might not hold in the impossible event that one of the litigants in a particular case (successfully) influenced the development of legal doctrine according to that litigant's prospective interests in that particular case. That is, unless one of the litigants is or could be omniscient, omnipresent, and omnipotent, this reasoning is firm.

In addition to the empirical tests suggested by Hylton, this study contributes another approach to empirically testing for the existence and type of private information during plea bargaining. This approach also avoids the problems of potentially confounding indirect selection effects that reduce the usefulness of the correlation of the trial rate and trial conviction rate. That approach makes use of the (statistically significant) estimated coefficient of the inverse Mills ratio used as an explanatory variable in a regression model of trial outcomes that controls for sample selection bias in the sample of cases that are chosen for trial.

The inverse Mills ratio used as a regressor in the trial outcome regression model derives from estimates of the probability of trial that are created by the regression model of the probability of trial. The regression model of the probability of trial estimates the probability of trial for each observation used in the probability of trial regression model, whether or not the observation involves a defendant that goes to trial.

When included as an explanatory variable in a (second-stage) regression model of trial outcomes, the inverse Mills ratio controls for (total or net) sample selection bias created by the defendant's decision to go to trial. As Snyder (1990) notes, estimating separate models of plea decisions and trial outcomes ignores the sequential nature of the case disposition process. Plea decisions determine the sample of cases that go to trial, which implies that plea decisions may not randomly determine the set of defendants that go to trial. This implies that the set of defendants that go to trial is a biased representation of the set of indicted defendants. Thus, observed changes in the trial conviction rate can be due to the change in the distribution of defendants who go to trial from the distribution of indicted defendants.

Ignore, as Froeb (1993) does, the potential for legal error at trial and assume that trials perfectly identify innocent and guilty defendants. If the prosecutor holds private information during plea bargaining and offers harsher penalties for a guilty plea than she previously offered, then truly innocent defendants (whom the prosecutor was previously offering lighter penalties in plea deals) are likely to be the defendants that reject the harsher plea offers and go to trial. In this situation, a defendant who is more likely to go to trial is less likely to be convicted at trial than a defendant who is less likely to go to trial.

In contrast, if defendants are privately informed and the prosecutor offers harsher penalties than she previously offered, then the truly guilty defendants are likely to go to trial (because truly innocent defendants were already going to trial). In this situation, a defendant who is more likely to go to trial is more likely to be convicted at trial than a defendant who is less likely to go to trial.

The sign of the estimated coefficient for the inverse Mills ratio that is estimated by the probability of trial regression model and included in a regression model of trial outcomes – where both regression models use defendant-level data – indicates whether defendants that are estimated to be more likely to go to trial are more or less likely to be convicted at trial (if they go to trial), holding other factors constant. A regression model of trial outcomes can control for the reduction of legal error at trial (which Froeb assumes away) by including time and occurances of landmark decisions as regressors.

In such a regression model of trial outcomes, if the sign of the estimated coefficient for the inverse Mills ratio suggests that defendants that are more likely to go to trial are *more* likely to be convicted at trial, then such a regression result suggests that (truly) *guilty* defendants are more likely to go to trial and that (truly) *innocent* defendants are more likely to accept plea bargains. Such a regression result would suggest that prosecutors are privately informed during plea bargaining about the defendants' chances at trial and are able to fashion plea arrangements that induce (truly) innocent defendants to plead guilty or *nolo contendere*.

Conversely, in a regression model that controls for the reduction in legal error at trial, if the sign of the estimated coefficient for the inverse Mills ratio suggests that defendants that are more likely to go to trial are *less* likely to be convicted at trial, then such a regression result suggests that (truly) *innocent* defendants are more likely to go to trial and that (truly) *guilty* defendants are more likely to accept plea bargains. Such a regression result would suggest that defendants are privately informed during plea bargaining about the their chances at trial and prosecutors are unable to fashion plea arrangements that induce such (truly) innocent defendants to plead guilty or *nolo contendere*.

Importantly, the regression models of the plea decision and trial outcome effectively control for the relevant case characteristics that create the selection effects explained in this study. Thus, using the sign of the coefficient of the inverse Mills ratio included in the trial outcome regression model to test for the information structure of plea bargaining avoids the previously discussed potential for indirect trial selection effects (as defined in this study) to confound the simple correlation of the observed trial rate and trial conviction rate that Froeb (1993) and other studies use to test for the information structure of plea bargaining.

CHAPTER X

EMPIRICAL ANALYSES

This study's empirical tests use over 3,000 observations of defendant-level data describing federal criminal antitrust cases initiated from 1956 through 1980. The empirical tests include two-sample means tests and regression analysis controlling for sample selection bias. The purposes of the empirical tests are to test the implications of the case selection model, trial selection model, and trial conviction model, as well as to test the implications of private information held by one of the litigants during plea bargaining

Using two different types of means tests and three regression models, this study examines the implications of the selection models. The means tests inspect differences in the five distinct observable samples of cases created by the case selection, trial selection, and trial conviction selection processes. The means tests suggest that the observed prosecutors did not randomly choose cases to initiate, the observed defendants did not randomly choose to go to trial, and the observed judges (or juries) did not randomly convict defendants at trial. The two-sample means tests also lend support to many predictions of the case selection, trial selection, and trial conviction models. The selection models thus appear ripe for general application to different types of criminal cases.

The first type of means test confirm that the means of variables representing case characteristics in the sample of cases that go to trial differ from the means of the same variables in the sample of cases that do *not* go to trial. These means differ in ways implied by the trial selection model. Similar means tests support the trial conviction model. This type of means test is not appropriate for testing the implications of the case selection model, because the sample of *unitiated* cases is unobserved.

The second type of means test supports the case selection model. Those means tests employ data from the sample of initiated cases (i.e., the full dataset) to produce (statistically significant) combinations of the relative extreme values of pairs of empirical variables. The identified combinations are consistent with the case selection model's predictions. Similar means tests – using the samples of cases that went to trial, cases that did not go to trial, cases that ended in trial conviction, and cases that went to trial but did not end in conviction – support the trial selection and trial conviction models.

In various ways, the regression analysis complements the previous means tests in assessing the implications of the trial selection model and trial conviction model. In general, the regression analysis supports the relevance of trial selection and trial conviction selection as determinants of federal criminal antitrust case outcomes.

Meanwhile, this study depends wholly on the regression models of the probabilities of trial and trial conviction to conduct tests for the existence and possessor of private (or better) information during plea bargaining in federal criminal antitrust cases. As highlighted in the literature review in Chapter II, the criminal antitrust literature ignores (or does not explicitly address) the issue of how private information in plea bargaining could alter the trial selection process and create (additional) bias in observed case outcomes. This is a significant omission given the policy implications of private information during plea bargaining.

Privately informed prosecutors harm social welfare because they can fashion a settlement offer that induces (truly) innocent defendants to plead guilty or *nolo contendere*. Such pleas send false signals about legal and socially efficient business conduct. In contrast, privately informed defendants during plea bargaining benefit social welfare because the (truly) guilty ones are more likely to plead guilty and send the correct signals about appropriate business conduct.

Regression models of the probabilities of trial and trial conviction suggest the existence of private information during plea bargaining. What is not clear from the regression results, however, is whether the defendants or prosecutors in the observed cases possess private or better information.

With these findings, this study's empirical findings contribute to the antitrust literature as well as the law and economics literature. The test results of this study are comparable to the empirical findings of other studies, but the empirical approaches differ from those of existing studies. The tests conducted in this study differ in terms of statistical approach and data from previously described empirical analyses found in the federal (criminal) antitrust enforcement literature (e.g., Cohen 1992, Gallo *et al*. 2000, Snyder 1989, 1990). For example, instead of using regression or other statistical tests, Gallo *et al*. (2000) describe trends in federal antitrust inforcement using various descriptive statistics by year, in the style of Posner (1970). This study's regression models of the *ex post* probabilities of trial and trial conviction in federal criminal antitrust cases differ from Snyder's (1989, 1990) in terms of model specifications, timeframe of data (Snyder uses 1970-1985), and approach to controlling for heteroskedasticity.

This study's regression model of sentencing decisions covers roughly the same period of observed sentencing decisions, 1956-1980, but it differs from Cohen's (1992) regression models in a number of ways. Cohen estimates general models of penalties imposed after any means of conviction – trial or pleas of guilty or *nolo contendere*. In contrast, this study focuses on the sentences imposed following conviction by trial. The model specifications differ in terms of explanatory variables, controlling for sample selection bias, and addressing heteroskedasticity.

The tests also differ from the tests conducted in studies that focus on the implications and determinants of case selection (e.g., Eisenberg and Farber 1997, 2003) and trial selection, including the information structure of plea bargaining/settlement negotiations (e.g., Froeb 1993, Hylton 1993, and Waldfogel 1998),. None of the existing studies uses data from federal criminal antitrust cases. Eisenberg and Farber (1997, 2003) test the implications of indirect effects of the plaintiff's case selection process on case outcomes by using regression models of observed trial rates and trial outcomes in federal civil cases. In contrast, this study uses two-sample means tests.

Froeb (1993) tests for the existence of private information and the litigant who holds (more or better) private information about the defendant's chances at trial during plea bargaining by estimating the correlation of trial rates and trial conviction rates in a variety of types of federal criminal cases. Similarly, Waldfogel (1998) examines the correlation of trial rates and plaintiff win rates at trial in various types of cases in which Waldfogel expects *a priori* whether plaintiffs or defendants possess private (or better) information about the defendant's chances at trial. Hylton (1993) tests the implications of informational advantages of litigants for trial rates and plaintiffs' success rates at trial by examining plaintiff win rates in various types of civil cases in which Hylton expects *a priori* whether plaintiffs or defendants possess private (or better) information about the defendant's chances at trial.

In contrast to Froeb (1993), Waldfogal (1998), and Hylton (1993), this study tests for the existence of private information and the identity of the privately informed defendant by comparing the predicted and actual signs of the estimated coefficients of various explanatory variables in regression models of defendants' decisions to go to trial and trial outcomes. As previously mentioned, the *K*ey explanatory variables are those that represent the passage of time, the development of legal doctrine, as well as the "Heckman's lambda" term that controls for sample selection bias created by the defendant's plea decision.

Data

Professor Mark Cohen graciously provided most of the data that describe defendants, the charges they face, and the judge in federal criminal antitrust cases initiated from 1956 through 1980. To those data were added data about the levels of profit and employee compensation in the industries represented by the defendants described by Cohen's data.[84] Data also were added regarding the annual appropriations used to fund the Antitrust Division of the U.S. Department of Justice, as well as the party of the U.S. President in the indictment year and sentencing year. The new data came from various sources – Gallo *et al*. (2000), U.S. Department of Justice ("DOJ") Antitrust Division, U.S. Bureau of Labor Statistics ("BLS"). Additional variables were created from the collected data to suit this study's testing needs. Table 9 describes the empirical variables.

Beginning with 3,907 defendant-level observations of cases that Professor Cohen provided, observations in which the defendant was a trade association or a union were deleted. These observations were deleted because the descriptions of the cases provided by the Commerce Clearing House ("CCH") give the Standard Industrial Code ("SIC") for trade associations or unions as "industries" in and of themselves, instead of the more relevant SIC for the industries represented by the defendant trade associations or unions.

If a hypothetical observed case involves a trade association of yo-yo manufacturers, the CCH description of the case would include the SIC for trade associations instead of the SIC for yo-yos (or toys). As shown in Table 9, the SIC provided by the CCH (as found in the Cohen dataset) was used to create empirical variables representing the annual levels of profits and employee compensation of industries represented by the observed defendants. These variables represent the defendants' unit (opportunity) cost of trial expenditure. In the hypothetical, this study would take interest in the annual level of profit and employee compensation in the yo-yo (or toy) industry, but would not be able to compile that information with the SIC provided by the CCH.

---

[84] Cohen's data includes the Standard Industrial Code, at least at the two-digit level, for the industry associated with each observed defendant.

Also deleted were observations that regard industries for which the BLS does not provide information regarding profit or employee compensation. In addition, observations in which variables related to the variance of the previous year's trial (conviction) outcomes are undefined because of division by zero, and observations in which variables related to the previous year's trial conviction rate are undefined because of a lack of a previous observed year were deleted. In order to avoid this problem, an alternative approach might be to use a three- or four-year average or otherwise interpolate a trial conviction rate for a previous period. This study uses the previous year's trial conviction rate based on the notion that a trial lawyer is only as good as her last case. Further, the DOJ Antitrust Division likely serves as a "revolving door" for attorneys seeking long-term careers in the private sector. Turnover among staff attorneys at the DOJ Antitrust Division implies that the previous observed year could be the most relevant in terms of capturing the Antitrust Division prosecutors' expectations of their abilities to convict defendants at trial in the observed indictment year.

Lastly, observations in which the number of counts of conviction exceeded the number of counts of indictment were deleted, since that appears to be a nonsensical situation. This situation could be the result of consolidation of charges in the indictment, but the benefits of trying to resolve this problem by going back to the original source of the data are not likely to exceed the costs.

The full dataset used for testing has 3,093 observations.[85] Table 1 defines the variables used in the theoretical analysis discussed above, and is presented here for ease of reference. Table 9 defines the empirical variables.

---

[85] This study used the statistics computer program *Stata Intercooled Version 9.0* to perform the means tests and estimate the regression models.

Table 1. Definitions of variables used in theoretical framework

Theoretical Framework Variable	Type of Variable	Theoretical Framework Definition
C	Exogenous	the unit cost of trial expenditure for the prosecutor
K	Exogenous	the unit cost of trial expenditure for the defendant
$\mathbf{e}$	Exogenous	the strength and amount of the prosecutor's evidence
λ	Exogenous	the litigants' common expected variability of the judge's trial conviction standard
ρ	Exogenous	the idiosyncratic component of the prosecutor's *ex ante* expected probability of trial conviction
M	Exogenous	the statutory maximum penalty
m	Exogenous	the statutory minimum penalty (which is zero until the implementation of the federal sentencing guidelines in 1987)
X	Choice	the prosecutor's level of trial expenditure
Y	Choice	the defendant's level of trial expenditure
X^*	optimal value of a choice variable and implicit function of other variables	the prosecutor's optimal level of trial expenditure
Y^*	optimal value of a choice variable and implicit function of other variables	the defendant's optimal level of trial expenditure
Θ	Endogenous, $\Theta = \Psi\left(\rho - \frac{CX + KY}{S^T}\right)$	the *ex ante* expected probability of trial
$\hat{\Pi}$	Endogenous, $\hat{\Pi} = \Pi([\mathbf{e} + P(X) - D(Y)]/\lambda)$	the litigants' common *ex ante* expected probability of trial conviction
S^T	Endogenous, $S^T = m + \mu^T(\mathbf{e}, P(X), D(Y))(M - m)$	the litigants' common *ex ante* expected penalty following a trial conviction
S^B	Endogenous, $S^B = (\hat{\Pi} + \rho)S^T + \frac{KY - CX}{2}$	the litigants' common *ex ante* expected penalty following a conviction by plea of guilty or *nolo contendere*

Table 9. Definitions of empirical variables

Dependent Variables	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
TRIAL	= 1 if defendant pleads *not* guilty; = 0 otherwise	Θ, $\tilde{\Theta}$	Plea
TRIAL CONVICTION	= 1 if defendant convicted at trial; = 0 otherwise	$\hat{\Pi}$, $\tilde{\Pi}$	Trial
FINE	= nominal dollars fined / annual CPI (1967 = 100) for sentencing year, in thousands ($000) (source of nominal amount of fine and sentencing year: Cohen dataset; source of CPI data: Bureau of Labor Statistics website)	S^T, $\tilde{S}^T$	Sentencing

Explanatory Variables Representing Prosecutor Characteristics	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
DOJBUDGET	= nominal dollars of annual appropriations to the DOJ Antitrust Division / annual CPI (1967 = 100) for indictment year, in thousands ($000) (source of appropriation data: DOJ website)	C	Plea; Trial; Sentencing

Table 9, continued

Explanatory Variables Representing Prosecutor Characteristics (cont.)	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
HIGH DOJBUDGET	= 1 if real (1967 = 100) annual appropriations to the DOJ Antitrust Division exceeds $8,915,747, the mean for the full sample (n = 3093); = 0 otherwise	C	Plea; trial; sentencing
%CRIMINAL	Percentage of opened Commerce Clearing House antitrust cases that are criminal per year (source: Gallo, *et al*. 2000)	C	Plea; trial; sentencing
ANNUAL FORTUNE500	Annual number of CCH federal criminal antitrust cases involving Fortune 500 companies (source: Gallo, *et al*. 2000)	C	Plea; Trial; sentencing
HIGH ANNUAL FORTUNE500	= 1 if the annual number of CCH federal criminal antitrust cases involving Fortune 500 companies exceeds 23 , the mean for the full sample (n = 3093); = 0 otherwise.	C	Plea; trial; sentencing
REPUB PRES INDICT YEAR	= 1 if President is Republican; = 0 if President is Democrat	Determinant of Θ, $\hat{\Pi}$, and S^T	Plea; Trial; Sentencing

Table 9, continued

Explanatory Variables Representing Defendant Characteristics	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
#CODEFEND	Number of other defendants in same case	K	Plea; Trial; Sentencing
#CORP CODEFEND	Number of other firms that are defendants in same case	K	Plea; Trial; Sentencing
ONLY CORP DEFEND	Case involves only firms as defendants	K	Plea; Trial; Sentencing
ONLY INDIV DEFEND	Case involves only individuals as defendants	K	Plea; Trial; Sentencing
ONLY HIGH RANK	= 1 if individual defendants in case are all of high corporate rank (e.g., president); = 0 otherwise	K	Plea; Trial; Sentencing
ONLY LOW RANK	= 1 if individual defendants in case are all of low corporate rank (e.g., below president); = 0 otherwise	K	Plea; Trial; Sentencing
ANY LOW RANK	= 1 if any individual defendants in case is of low corporate rank (e.g., below president); = 0 otherwise	K	Plea; Trial; Sentencing
PROFIT	= Nominal annual corporate profit by 2-digit SIC code / annual CPI (1967 = 100) for the indictment year, in millions ($000,000) (source: U.S. Bureau of Labor Standards)	$\mathbf{e}$, K	Plea; Trial; Sentencing

Table 9, continued

Explanatory Variables Representing Defendant Characteristics (cont.)	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
HIGH PROFIT	= 1 if the value of *PROFIT* exceeds $2,806.854 million, the mean for the full sample (n = 3093); = 0 otherwise	$\mathbf{e}$, K	Plea; Trial; Sentencing
COMPENSATION	= Nominal employee compensation by 2-digit SIC code / annual CPI (1967 = 100) for the indictment year, in millions ($000,000) (source: U.S. Bureau of Labor Standards)	$\mathbf{e}$, K	Plea; Trial; Sentencing
HIGH COMPENSATION	= 1 if the value of *COMPENSATION* exceeds $15,782.14 million, the mean for the full sample (n = 3093); = 0 otherwise	$\mathbf{e}$, K	Plea; Trial; Sentencing
#OTHER INDICT	Number of other indictments for the defendant in the same year	$\mathbf{e}$, K	Plea; Trial; Sentencing
ZERO OTHER INDICT	= 1 if the number of other indictments for the defendant in the same year equals zero; = 0 otherwise	$\mathbf{e}$, K	Plea; Trial; Sentencing

Table 9, continued

Explanatory Variables Representing Defendant Characteristics (cont.)	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
#PREV CONVICT	Number of previous antitrust convictions for the defendant	**e**	Plea; Trial; Sentencing
ZERO PREV CONVICT	= 1 if the number of previous antitrust convictions for the defendant equals zero; = 0 otherwise	**e**	Plea; Trial; Sentencing

Explanatory Variables Representing Violation Characteristics	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
COMMERCE	= Nominal dollar amount of commerce involved in case / annual CPI (1967 = 100) for the indictment year, in thousands ($000)	**e**	Plea; Trial; Sentencing
HIGH COMMERCE	= 1 if the value of *HIGH COMMERCE* exceeds $64,396.21 thousand, the mean for the full sample (n = 3093); = 0 otherwise	**e**	Plea; Trial; Sentencing
#COUNTS CONVICT	Number of counts of conviction	**e**	Sentencing
MULTI COUNTS CONVICT	= 1 if the number of counts of conviction exceeds one; = 0 otherwise	**e**	Sentencing

Table 9, continued

Explanatory Variables Representing Violation Characteristics (cont.)	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
#COUNTS INDICT	Number of counts of indictment	**e**	Plea; Trial; Sentencing
MULTI COUNTS INDICT	= 1 if the number of counts of indictment exceeds one; = 0 otherwise	**e**	Plea; Trial; Sentencing
COUNTS INDICT-CONVICT	Number of counts of indictment minus number of counts of conviction	**e**	Plea; Trial; Sentencing
DURATION	Duration of alleged conspiracy in years	**e**	Plea; Trial; Sentencing
HIGH DURATION	= 1 if the duration of the alleged conspiracy exceeds 5 years, the mean for the full sample (n = 3093); = 0 otherwise	**e**	Plea; Trial; Sentencing
ANY BID RIG	= 1 if any count of indictment alleges bid rigging; = 0 otherwise	**e**, ρ, and $\hat{\Pi}$	Plea; Trial; Sentencing
PRICE FIX NOT BID RIG	= 1 if the defendant is charged with price fixing but not bid rigging; = 0 otherwise	**e**, ρ, and $\hat{\Pi}$	Plea; Trial; Sentencing
BID RIG NOT PRICE FIX	= 1 if the defendant is charged with bid rigging but not price fixing; = 0 otherwise	**e**, ρ, and $\hat{\Pi}$	Plea; Trial; Sentencing
CONSUMER VIC	= 1 if the alleged victim is a consumer; = 0 otherwise	**e**	Plea; Trial; Sentencing

Table 9, continued

Explanatory Variables Representing Sentencing Determinants	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
SENTENCE YEAR	year of sentencing decision	Determinant of S^T	Sentencing
CASE TIME SPAN	Year of sentencing minus year of indictment	Determinant of S^T (X or Y)	Sentencing
POST-APPA SENTENCE YEAR	= 1 if the sentencing year is 1975 or later; = 0 otherwise	Determinant of S^T	Sentencing
POST-APPA INDICT YEAR	= 1 if the indictment year is 1975 or later; = 0 otherwise	Determinant of S^T	Plea; Trial

Explanatory Variables Representing (Sentencing) Judge Characteristics	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
REPUB JUDGE	= 1 if judge is a Republican in the sentencing year (according to Cohen's source); = 0 otherwise	Determinant of $\hat{\Pi}$, S^T	Trial; Sentencing
REPUB PRES SENTENCE YEAR	= 1 if President is a Republican in the sentencing year; = 0 otherwise	Determinant of S^T	Sentencing
PRES & JUDGE SAME PARTY SENTENCE YEAR	= 1 if President and judge are in the same political party in the sentencing year; = 0 otherwise	Determinant of S^T	Sentencing

Table 9, continued

Explanatory Variables Representing (Sentencing) Judge Characteristics (cont.)	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
PRIOR PROSECUTOR	= 1 if the judge previously served as a government prosecutor; = 0 otherwise	Determinant of $\hat{\Pi}$, S^T	Trial; Sentencing
TENURE	Number of years the judge has served on the federal district court	Determinant of $\hat{\Pi}$, S^T	Trial; Sentencing
#FILINGS	Number of cases filed in district per judge	Determinant of $\hat{\Pi}$, S^T	Trial; Sentencing

Other Explanatory Variables	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
INDICT YEAR	Year of indictment	Determinant of Θ and $\hat{\Pi}$	Plea; Trial
LANDMARK	Average number of "landmark" antitrust cases identified in leading antitrust law treatises. (source: Gallo *et al*. 2000)	Determinant of Θ, $\hat{\Pi}$, and S^T	Plea; Trial; Sentencing
DOJ WIN PREV YEAR BY VIOL#1	Trial conviction rate for previous observed year by type of first violation in the indictment	ρ	Plea

Table 9, continued

Other Explanatory Variables (cont.)	Definition	Associated Theoretical Model Variable(s)	Case Disposition Stage of Empirical Model
HIGH DOJ WIN PREV YEAR BY VIOL#1	= 1 if the trial conviction rate for the previous observed year by type of first violation in the indictment exceeds 28.70472, which is the mean value for the full sample (n = 3093); = 0 otherwise	ρ	Plea
DOJ WIN PREV YEAR	Trial conviction rate for previous observed year	ρ	Plea
VAR DOJ WIN PREV YEAR BY VIOL#1	Variance of trial outcomes for previous observed year by type of first violation in the indictment	λ	Trial
VAR DOJ WIN PREV YEAR	Variance of trial outcomes for previous observed year	λ	Trial
MILLS TRIAL	Inverse Mills ratio that controls for probability of trial	Determinant of $\hat{\Pi}$, S^T	Trial; Sentencing
MILLS TRIAL CONVICT	Inverse Mills ratio that controls for probability of trial conviction	Determinant of S^T	Sentencing

Cross-Branch Means Tests of Trial Selection and Trial Conviction Selection Processes

This section discusses the results of two-sample means tests used to test predictions of how, given the relative value of one case characteristic, the trial selection process and trial conviction process shape the mix of the other characteristics of cases that proceed to subsequent stages of case disposition. Cursory reviews of Tables B-1 through B-5 in Appendix B[86] reveals that the distributions of the variables differ across sub-samples associated with various stages of case disposition (i.e., various branches stemming from various nodes of the game tree of case disposition). Differences in these distributions support the notion that the defendant's plea decision, and the judge's or jury's decision to convict at trial are sorting processes that alter the mixes of the characteristics of cases as they proceed through the legal process.

For purposes of testing the selection models' predictions, correlation and regression analyses are generally not suitable because the direction of the marginal changes in the values of, or the relative values of, the case characteristics under consideration matters. In part, that is because the theoretical predictions differ according to whether a marginal *increase* or *decrease* in the value of a case characteristic under consideration occurred, or whether a relatively *high* or *low* level of a case characteristic under consideration is observed.

As discussed in Chapter VI, potential cases that prosecutors do not initiate are unobserved. Thus, it is not possible to test the predictions of the model of the prosecutor's case selection process by comparing the means of case characteristics across samples of cases that are and are not initiated (i.e., across the two game tree branches stemming from the node of the prosecutor's case selection decision). Such tests are possible for the model of the defendant's plea decision and the judge's or jury's trial conviction decision, however.

Overall, the results only moderately support the predictions of the trial selection model. This is not surprising given that corporate defendants and white-collar individual defendants, as well as their (presumably) expert attorneys likely base decisions (or advice) on entered pleas on considerations not easily captured by the empirical variables created from available data. In

---

[86] Appendix B provides summary statistics of the data, by samples of cases at various stages of case disposition. Table B-1 summarizes the full set of data; Table B-2, the sample of observations in which the defendants entered pleas of guilty of *nolo contendere*; Table B-3, the sample of observations in which the defendants entered not guilty pleas; Table B-4, the sample of observations in which the defendants were convicted at trial; and Table B-5, the sample of observations in which the defendants were not convicted at trial (i.e., the defendant was acquitted, the case was dismissed, or the case ended in a directed verdict in the defendant's favor).

comparison to the results of the means tests of the trial selection model, the means test results provide more support for the model of the judicial trial conviction decision. Some findings provide somewhat ambiguous support for the model's predictions that requires alternative explanations, however. In addition, several results of both sets of means tests are not statistically significant.

Tests of the trial selection process

As discussed in Chapter VII, the theoretical model of the defendant's plea decision implies firm relationships among the relative values of the same case characteristic in the sample of defendants in cases that go to trial versus the sample of defendants in cases that do not go to trial (i.e., across *different* game tree branches stemming from the plea decision mode). These relationships can be tested meaningfully using sample means tests using the samples of defendants in (initiated) cases that go to trial and defendants in cases that do not go to trial. Figure 3 provides a game tree diagram that shows the two samples used in the following tests of the trial selection model – the sample of defendants who choose to go to trial and the sample of defendants who choose not to go to trial – produced by the defendant's plea decision.

Specifically, cases that go to trial (do not go to trial) are expected to be characterized by a relatively low (high) value of the prosecutor's unit cost of trial expenditure, C, and the defendant's unit trial cost, K. Moreover, cases that go to trial (do not go to trial) are expected to be characterized by a relatively high (low) value of the parties' common expected trial sentence, S^T and a relatively high value of the idiosyncratic component, ρ, of the prosecutor's expected probability of trial conviction.

Overall, the results of the following tests of the difference in the values of the means of empirical variables representing case characteristics in the samples of defendants who do (n = 601) and do not go to trial (n = 2492) provide mixed support for the model of the defendant's plea decision (i.e., the trial selection process). Appendix C reports the means test results. Table 10 summarizes the results.

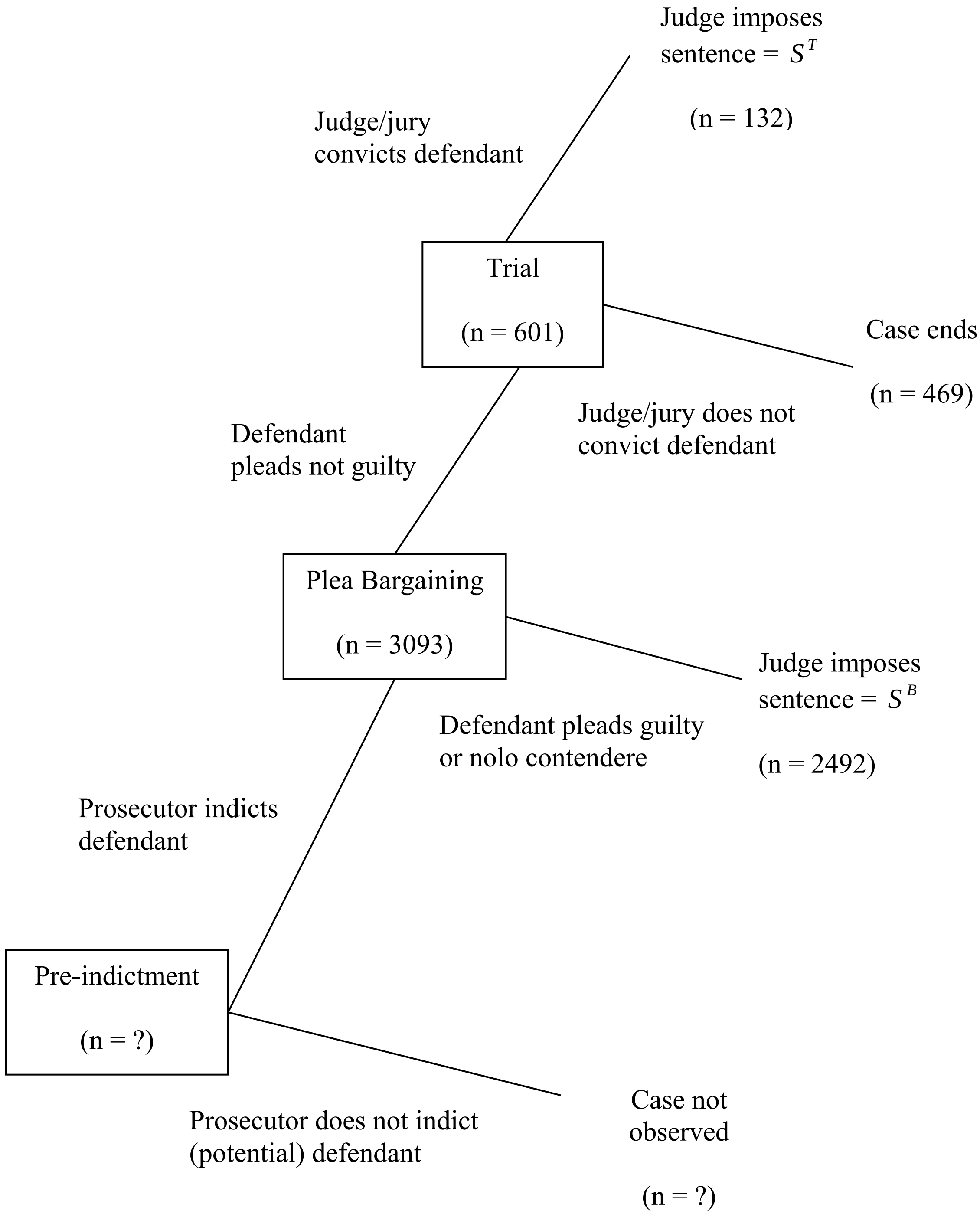

Figure 3. Game tree representation of the case disposition process and relevant sample sizes

Table 10.a. Results of cross-branch means tests of the defendant's plea decision as a selection process

Empirical Variable	Associated Case Characteristic(s)	Mean Value in Sample of Litigated Cases (n = 601)*	Mean Value in Sample of Non-Litigated Cases (n = 2492)*
DOJBUDGET	C	8139.42	9102.98
%CRIMINAL	C	0.4874542	0.4719141
ONLY INDIV DEFEND	K	0.0449251	0.0232745
ONLY CORP DEFEND	K	0.5990017	0.6508828
PROFIT	K	3013.09	2757.12
COMPENSATION	K	14552.8	16078.62
#CORP CODEFEND	K	7.90183	6.831461
#OTHER INDICT	$\mathbf{e}$, K	0.1647255	0.4357945
COUNTS INDICT-CONVICT	$\mathbf{e}$, S^T, ρ	1.073211	1.26565
#COUNTS INDICT	$\mathbf{e}$, S^T	1.291181	1.369583
DURATION	$\mathbf{e}$, S^T	4.650582	5.672151
COMMERCE	$\mathbf{e}$, S^T	45534.5	68945.12
BID RIG NOT PRICE FIX	ρ	0.1397671	0.1175762
ANY BID RIG	ρ	0.2312812	0.2624398
DOJ WIN PREV YEAR	ρ	25.43097	27.02271
DOJ WIN PREV YEAR BY VIOL#1	ρ	26.82892	29.15711
CASE TIME SPAN	None	1.329451	0.8896469
POST-APPA INDICT YEAR	S^T	Statistically Insignificant	
#PREV CONVICT	$\mathbf{e}$, S^T	Statistically Insignificant	
ANNUAL FORTUNE500	C	Statistically Insignificant	
ONLY HIGH RANK	K	Statistically Insignificant	
ONLY LOW RANK	K	Statistically Insignificant	
COMPENSATION	K	Statistically Insignificant	

* Sample mean values are statistically significant unless otherwise noted.

Table 10.b. Comparison of expected and observed means test results

Empirical Variable	Associated Case Characteristic(s)	Sample in Which a Larger Value is Expected	Sample in Which a Larger Value is Observed
DOJBUDGET	C	Litigated	Non-litigated
%CRIMINAL	C	?	Litigated
ONLY INDIV DEFEND	K	Non-litigated	Litigated
ONLY CORP DEFEND	K	Litigated	Non-Litigated
PROFIT	K	Litigated	Litigated
COMPENSATION	K	Litigated	Non-litigated
#CORP CODEFEND	K	Litigated	Litigated
#OTHER INDICT	$\mathbf{e}$, K	Non-litigated for K; Litigated for e	Non-litigated
COUNTS INDICT-CONVICT	$\mathbf{e}$, S^T, ρ	Non-litigated	Non-litigated
#COUNTS INDICT	$\mathbf{e}$, S^T	Litigated	Non-litigated
DURATION	$\mathbf{e}$, S^T	Litigated	Non-litigated
COMMERCE	$\mathbf{e}$, S^T	Litigated	Non-litigated
BID RIG NOT PRICE FIX	ρ	Litigated	Litigated
ANY BID RIG	ρ	Litigated	Non-litigated
DOJ WIN PREV YEAR	ρ	Litigated	Non-litigated
DOJ WIN PREV YEAR BY VIOL#1	ρ	Litigated	Non-litigated
CASE TIME SPAN	None	Litigated	Litigated
POST-APPA INDICT YEAR	S^T	Litigated	Neither
#PREV CONVICT	$\mathbf{e}$, S^T	Litigated	Neither
ANNUAL FORTUNE500	C	Non-litigated	Neither
ONLY HIGH RANK	K	Litigated	Neither
ONLY LOW RANK	K	Non-litigated	Neither
COMPENSATION	K	Litigated	Neither

The means test results suggest that the trial selection process is complicated and that the empirical variables do not capture all of the subtle considerations that go into the defendants' plea decisions. Alternatively, the number of statistically insignificant results could suggest that the defendant's plea decision is a random process.

The latter interpretation is not satisfying intuitively in the context of federal criminal antitrust cases, however. These cases involve white-collar individual and corporate defendants represented by skilled attorneys. These defendants are not likely prone to whimsical, random decision-making in the face of generally high stakes in terms of personal and corporate finances and reputations.

In order to interpret the means test results in relation to the theoretical predictions, recall the condition that, according to the theoretical framework, must hold in order for a defendant to go to trial: $(\rho - \delta)S^T - CX - KY > 0$. The condition for a defendant to be observed choosing to go to trial suggests that cases that go to trial are more likely to be characterized by a relatively low value of the defendant's unit cost of trial expenditure, K, compared to value of K observed in cases that do not go to trial. The means test results regarding *PROFIT*, *#OTHER INDICT*, and *#CORP CODEFEND* support the theoretical predictions if, as expected, the defendants' (opportunity) cost of trial expenditure: falls with the subject defendant industry's profits, rises with the number of indictments the subject defendant faces in the same year, and falls with the number of corporate defendants in the subject case.

On the other hand, the results of two means tests related to the defendant's trial costs contradict the theoretical predictions. As discussed above and by Eisenberg and Farber (2003), the defendant's unit cost of trial expenditure, K, is expected to be higher for individual defendants than corporate defendants. Accordingly, *ceteris paribus*, cases in which all of the defendants are individuals are expected to end in a plea of guilty or *nolo contendere*, while cases involving all corporate defendants are expected to go to trial. Thus, the findings that the mean values or the empirical variables *ONLY INDIV DEFEND* and *ONLY CORP DEFEND* are higher and lower, respectively, in cases that proceeded to trial versus cases that did not proceed to trial run counter to the predictions of the model of the defendant's plea decision.

Similarly, the finding that the value of *COMPENSATION* is higher in the sample of non-litigated cases contradicts the trial selection model's prediction that defendants with lower unit (opportunity) costs of trial expenditure are more likely to go to trial. As the defendant's level of compensation rises, his opportunity cost from trial falls.

Assuming that the prosecutor is more confident about her chances of winning at trial in a bid-rigging case than a price-fixing case, then the idiosyncratic component, ρ, of the prosecutor's expected probability of trial conviction is higher in a bid-rigging case than in a price-fixing case. Thus, the finding that cases that involved bid-rigging charges but not price-fixing charges, as represented by the dummy variable *BID RIG NOT PRICE FIX*, were more likely to go to trial than to end in a plea of guilty or *nolo contendere* supports the theoretical predictions of the model of the defendant's plea decision.

In contrast, the values of the empirical variables that are supposed to more directly represent ρ -- *DOJ WIN PREV YEAR* and *DOJ WIN PREV YEAR BY VIOL#1* – have lower mean values in cases that go to trial than in cases that do not go to trial. These findings contradict the theoretical predictions.

A smaller difference in the number of charges in the indictment versus charges for which the defendant is convicted suggests stronger evidence, e. Further, the *ex ante* expected trial sentence, S^T, is expected to increase in e. Thus, the finding that the mean value of *COUNTS INDICT-CONVICT* is lower in cases that go to trial than in cases that do not go to trial supports the theoretical prediction of the model of the defendant's plea choice.[87]

On the other hand, the findings that the mean values of *COMMERCE* and *#COUNTS INDICT* are lower in the sample of cases that go to trial than in the sample of cases that end in pleas of guilty or *nolo contendere* does not support the theoretical predictions of the model of the defendant's plea decision. That is because the *ex ante* expected trial sentence, S^T, and the probability of trial are expected to increase in the amount of commerce involved in an alleged conspiracy and, more generally the amount of evidence (notwithstanding the strength of the evidence).

Meanwhile, the finding that the mean value of *DURATION* is lower in the sample of cases that go to trial than in the sample of cases that do not go to trial does not necessarily support or contradict the theoretical predictions of the model of the defendant's plea choice. To the extent that the *ex ante* expected sentence, and hence the probability of trial, increases in the

87 Meanwhile the finding that the value of *CASE TIME SPAN* is higher in cases that go to trial than in cases that do not go to trial simply supports the common sense prediction that cases that go to trial last longer.

duration of the alleged conspiracy, this finding contradicts the theoretical predictions. The *ex ante* expected sentence does not necessarily increase in the duration of the conspiracy, however, if the conspiracy was unsuccessful and did not cause much economic harm.

Similarly, the finding that the mean value of *%CRIMINAL* is higher in the sample of cases that go to trial than in the sample of cases ending in a plea agreement does not necessarily contradict or support the theory developed in this study. If the prosecutor's unit (opportunity) cost of trial expenditure, C, increases with the percentage of criminal cases in which the DOJ Antitrust Division is involved in a given year, then the finding does not support the model's predictions. It is possible, however, that there are some economies of scale or economies of focus from bringing a larger percentage of criminal cases in a given year. If so, this means test result supports the theory.

Taken together, some of these means test results confirm the theoretical predictions of the model of the defendant's plea choice. Some of the findings undermine the model's predictions, however. Some findings provide ambiguous support for the model's predictions. Several means test results are not statistically significant.[88]

Such mixed support for the model of the defendant's plea choice are not surprising given that corporate defendants and white-collar individual defendants, as well as their (presumably) expert attorneys likely base decisions (or advice) on entered pleas on considerations not easily captured by the empirical variables created from available data. For example, the empirical variables do not control for the defendant's pre-trial assessment of the strength of his legal representation or the credibility of his best defense if he chooses to go to trial.

Tests of the trial conviction selection process

The following tests compare the means of empirical variables representing case characteristics in the sample of defendants in cases that proceed to trial (i.e., "litigated cases") and end in trial conviction (n = 132) with the sample of defendants in litigated cases that do not

[88] Specifically, the tests of the difference of means values across the samples of cases that did and did not go to trial, using the following variables, did not produce statistically significant results: *POST-APPA INDICT YEAR* (which is positively related to S^T), *#PREV CONVICT* (which is positively related to S^T), *ANNUAL FORTUNE500* (which is positively related to *C*), *ONLY HIGH RANK* (which is negatively related to *K*), and *ONLY LOW RANK* (which is positively related to *K*).)

end in trial conviction (n = 469). The results generally support the predictions of the theoretical model of the rule by which the judge or jury decides to convict the defendant at trial (i.e., the trial conviction process).

The theoretical model of the trial conviction decision implies certain relationships among the relative values of the same case characteristic in the sample of defendants in litigated cases that end in trial conviction and the sample of defendants in litigated cases that do not end in conviction (i.e., across *different* game tree branches stemming from the trial decision node). These relationships can be tested meaningfully using sample means tests using the samples of defendants in (initiated) cases that end in trial conviction and samples of defendants in litigated cases that do not end in conviction.

Specifically, the trial conviction model predicts that when the prosecutor has weaker diminishing returns in trial expenditure than the defendant, a case characterized by a relatively low value of the prosecutor's unit cost of trial expenditure, C is more likely to end in trial conviction. Meanwhile, a case characterized by a relatively high value of C is more likely to end in acquittal, dismissal, or directed verdict in favor of the defendant. Meanwhile, a case that ends in trial conviction is expected to be characterized by relatively high values of the defendant's unit trial cost, K, the parties' common expected trial sentence, S^T; and the amount or strength of the prosecutor's evidence, e. Out of all cases that go to trial, cases that do not end in trial conviction are expected to be characterized by relatively low values of K, S^T, e.

With two exceptions, the means tests discussed below use the same empirical variables used in the tests related to the trial selection process. The two exceptions are the omission of empirical variables representing the idiosyncratic component, ρ, of the prosecutor's expected probability of trial conviction, and the inclusion of empirical variables representing λ, the parties common expectation regarding the latitude that the legal system allows judges in making discretionary decisions at trial,

Most of the statistically significant test results are consistent with the theoretical predictions discussed in Chapter VIII. None of the statistically significant results is completely inconsistent with the theoretical predictions, although several of the test results are statistically insignificant. Appendix D reports the means test results. Table 11 summarizes the statistically significant results.

Table 11.a. Results of cross-branch means tests of the trial conviction decision as a selection process

Empirical Variable	Associated Case Characteristic(s)	Mean Value in Sample of Litigated Cases That End in Conviction (n = 132)*	Mean Value in Sample of Litigated Cases That Do Not End in Conviction (n = 469)*
ANNUAL FORTUNE500	C	19.06818	25.18763
%CRIMINAL	C	0.515303	0.4796162
#CODEFEND	C, K	9.560606	12.55437
ONLY INDIV DEFEND	K	0.0833333	0.0341151
ONLY HIGH RANK	$\mathbf{e}$, K	0.8939394	0.7654584
ONLY LOW RANK	$\mathbf{e}$, K	0.0909091	0.2089552
PROFIT	K	2579.31	3135.18
DURATION	$\mathbf{e}$	5.545455	4.398721
#COUNTS INDICT	$\mathbf{e}$	1.181818	1.321962
COMMERCE	$\mathbf{e}$, S^T	61510.3	41038.11
PRIOR PROSECUTOR	$\hat{\Pi}$, S^T	0.5909091	0.4626866
PRICE FIX NOT BID RIG	$\hat{\Pi}$	0.6590909	0.7377399
BID RIG NOT PRICE FIX	$\hat{\Pi}$	0.2045455	0.1215352
VAR DOJ WIN PREV YEAR BY VIOL#1	λ	11.00835	12.96607
DOJBUDGET	C	Statistically Insignificant	
ONLY CORP DEFEND	K	Statistically Insignificant	
COMPENSATION	K	Statistically Insignificant	
#OTHER INDICT	$\mathbf{e}$, K	Statistically Insignificant	
#PREV CONVICT	$\mathbf{e}$	Statistically Insignificant	
COUNTS INDICT-CONVICT	$\mathbf{e}$	Statistically Insignificant	
POST-APPA INDICT YEAR	S^T	Statistically Insignificant	
VAR DOJ WIN PREV YEAR	λ	Statistically Insignificant	
CASE TIME SPAN	None	Statistically Insignificant	

* Sample mean values are statistically significant unless otherwise noted.

Table 11.b. Comparison of expected and observed means test results

Empirical Variable	Associated Case Characteristic(s)	Sample in Which a Larger Value is Expected	Sample in Which a Larger Value is Observed
ANNUAL FORTUNE500	C	Not conviction	Not conviction
%CRIMINAL	C	?	Conviction
#CODEFEND	C, K	?	Not conviction
ONLY INDIV DEFEND	K	Conviction	Conviction
ONLY HIGH RANK	$\mathbf{e}$, K	Conviction for $\mathbf{e}$; Not conviction for K	Conviction
ONLY LOW RANK	$\mathbf{e}$, K	Not conviction for $\mathbf{e}$; Conviction for K	Not conviction
PROFIT	K	Not conviction	Not conviction
DURATION	$\mathbf{e}$	Conviction	Conviction
#COUNTS INDICT	$\mathbf{e}$	?	Not conviction
COMMERCE	$\mathbf{e}$, S^T	Conviction	Conviction
PRIOR PROSECUTOR	$\hat{\Pi}$, S^T	Conviction	Conviction
PRICE FIX NOT BID RIG	$\hat{\Pi}$	Not conviction	Not conviction
BID RIG NOT PRICE FIX	$\hat{\Pi}$	Conviction	Conviction
VAR DOJ WIN PREV YEAR BY VIOL#1	λ	?	Not conviction
DOJBUDGET	C	Conviction	Neither
ONLY CORP DEFEND	K	Not conviction	Neither
COMPENSATION	K	Not conviction	Neither
#OTHER INDICT	$\mathbf{e}$, K	Conviction	Neither
#PREV CONVICT	$\mathbf{e}$	Conviction	Neither
COUNTS INDICT-CONVICT	$\mathbf{e}$	Not conviction	Neither
POST-APPA INDICT YEAR	S^T	Conviction	Neither
VAR DOJ WIN PREV YEAR	λ	?	Neither
CASE TIME SPAN	None	?	Neither

To see how the results relate to the theoretical predictions, recall from Chapter V the theoretical condition that the judicial trial conviction rule is to find the defendant guilty if and only if $[\mathbf{e} + P(X) - D(Y)] > \lambda v$. Recall that e is the quality or quantity of the prosecutor's evidence against the defendant. *P(X)* and *D(Y)* are scalars that denote the persuasive value of testimony and legal arguments regarding the defendant's culpability and liability made at trial by the prosecutor and defendant, respectively. *X* and *Y* represent the prosecutor's and defendant's levels of trial expenditure, respectively; *v* represents the random factors associated with the judge's conviction standard; and the spread factor λ represents the parties' common expectation regarding the latitude the legal system allows judges in making discretionary decisions at trial.[89]

The theory developed earlier predicts that the probability of trial conviction increases in the *ex ante* expected trial sentence, S^T, if the prosecutor has weaker diminishing returns in trial expenditure than the defendant. Also, the probability of trial conviction increases in the defendant's unit (opportunity) cost of trial expenditure, K, and decreases in the prosecutor's unit (opportunity) cost of trial expenditure, C.

As shown in Table 11, several means test results are not statistically significant.[90] Nonetheless, the means test results reported in Table 11 provide more support for the model of the trial conviction decision than the results of the tests regarding the trial selection model reported in Table 10. It is possible that the empirical variables employed in the means tests capture the factors relevant in trial conviction decisions better than they capture the factors that influence plea decisions. That is, the empirical variables likely reflect entries into evidenciary records on which judges and juries are legally bound to base their trial conviction decisions. In contrast, it is less likely that the empirical variables capture more speculative or nuanced considerations, based on the information contained in the empirical variables or other information, that influence defendants' plea decisions.

---

89 Since the expected value of *v* is zero, the judicial trial conviction rule can be rewritten as $\mathbf{e} + P(X) > D(Y)$, which simplifies the instant analysis.

90 In particular, the tests of the difference of means values across the samples of litigated cases that did and did not end in trial conviction, using the following variables, did not produce statistically significant results: *DOJBUDGET* (which is negatively related to *C*); *ONLY CORP DEFEND*, *COMPENSATION* (which is negatively related to *K*); *#OTHER INDICT* (which is positively related to *K*); *#PREV CONVICT* and *COUNTS INDICT-CONVICT* (which are positively and negatively related to **e**, respectively); *POST-APPA INDICT YEAR* (which is positively related to S^T); as well as *VAR DOJ WIN PREV YEAR* (which is positively related to λ). In addition, the means test involving *CASE TIME SPAN* produced statistically insignificant results.

The mean value of the empirical variable *ANNUAL FORTUNE500* is lower in the sample of cases that go to trial (henceforth, for brevity, "litigated cases") and end in conviction than in the sample of litigated cases that do not end in conviction. It is reasonable to expect the prosecutor's unit (opportunity) cost of trial expenditure, C, to rise as the DOJ Antitrust Division brings more cases against Fortune 500 companies. Thus, this finding supports the prediction that cases characterized by a relatively low value of C will end in conviction if they go to trial.

The mean of *%CRIMINAL* is higher in the sample of litigated cases that end in conviction than in the sample of litigated cases that do not. If C increases in the percentage of criminal cases the DOJ Antitrust Division initiates in a given year, then this finding contradicts the predictions of the model of the trial conviction decision. On the other hand, if there are economies of scale or focus from bringing a higher percentage of criminal antitrust cases, then a relatively higher value of *%CRIMINAL* implies a relatively lower value of C, and this finding supports the theoretical prediction of the model of the trial conviction decision.

The mean of *#CODEFEND* is higher in the sample of litigated cases that do *not* end in conviction than in the sample of cases that end in trial conviction. Defendants could cooperate or coordinate their litigation efforts in such a way to reduce their combined trial costs – for example, by sharing pre-trial information. If so, then the level of the defendant's unit cost of trial expenditure, K, is higher when the value of *#CODEFEND* is larger. The prosecutor's unit cost of trial expenditure, C, also could increase with the number of defendants if, for example, the prosecutor's preparations for cross-examining witnesses require more time or other resources as the number of defendants increases. If so, then C increases in *#CODEFEND*. The means test result for *#CODEFEND* is consistent with these interpretations of how *#CODEFEND* represents C and K, since the probability of trial conviction, as modeled, decreases in C and increases in K.[91]

The probability of trial conviction, as modeled, increases in the defendant's unit cost of trial expenditure, K.It is reasonable to expect individual defendants to have higher (opportunity) costs of trial than corporate defendants. Thus, the finding that the mean value of *ONLY INDIV*

---

[91] On the other hand, it is plausible that defendants cannot or choose not to cooperate or coordinate, which implies that K increases in *#CODEFEND*. The prosecutor could leverage such non-cooperation among defendants into a lower unit cost of trial expenditure, C, if, for example, at least some of the defendants are willing to cooperate with the prosecutor to some extent. The more defendants, the more potentially cooperative witnesses exist. The finding that the mean of *#CODEFEND* is higher in the sample of non-trial conviction cases does not square with this interpretation of the relationship between *#CODEFEND* and C or K.

DEFEND is higher in the sample of litigated cases that end in trial conviction than in the sample of litigated cases that do not end in conviction supports the model.

The higher the defendant's industry's level of profits, the lower the defendant's opportunity cost from trial expenditure, K. Thus, the finding that the mean value of *PROFIT* is higher in the sample of litigated cases also supports the prediction of the trial conviction model.

The finding that the mean value of *ONLY LOW RANK* is lower in the sample of litigated cases that end in conviction suggests that litigated cases in which all of the individual defendants have low corporate ranks (i.e., lower than president of the company) and, accordingly, *higher* unit costs of trial expenditure, are *less* likely to end in conviction. Similarly, the finding that the mean value of *ONLY HIGH RANK* is higher in the sample of litigated cases that end in conviction suggests that a case is *more* likely to end in conviction if all of the individual defendants are high-ranking and have *lower* unit costs of trial expenditure. If the variables *ONLY LOW RANK* and *ONLY HIGH RANK* represent only the defendant's unit cost of trial expenditure, *K*, then one could reasonably interpret these findings as counter to the predictions of the model of the trial conviction (selection) process.

It is also reasonable, however, to suspect that cases involving only low-ranking or high-ranking individual defendants also feature relatively weak or strong evidence, respectively.[92] If *ONLY LOW RANK* and *ONLY HIGH RANK* represent the amount or strength of evidence, $\mathbf{e}$, then the findings regarding *ONLY LOW RANK* and *ONLY HIGH RANK* support the prediction of the model of the conviction decision process.

The mean values of *DURATION* and *BID RIG NOT PRICE FIX* are higher in the sample of litigated cases that end in trial conviction, confirming the trial conviction model's prediction that more or stronger evidence increases the probability of trial conviction. The latter finding supports Snyder's (1990) assertion that bid-rigging cases are easier to prosecute successfully because the prosecutor can more easily gather convincing evidence in those cases, such as the testimony of co-conspirators.

Snyder does not mention the likelihood that rigging bids for providing or supplying government services produces more or stronger evidence than fixing prices for goods or services

[92] Otherwise, why would prosecuting the chief executive officers in the Enron and WorldCom fraud cases be so important to government prosecutors? In the sample of litigated cases, the highest estimated correlation coefficient between *ONLY LOW RANK* and empirical variables representing the quantity or quality of evidence was 0.13 (between *ONLY LOW RANK* and *MULTI HPSV COUNTS*). The empirical variables created from the available data do not necessarily fully capture the strength of the evidence in any given case, however.

sold to the private sector, *ceteris paribus*. Government procurement processes are more likely to capture and maintain relevant information that is more readily available to prosecutors than information captured and maintained by private sector individuals or organizations. Thus, the finding that the mean of *PRICE FIX NOT BID RIG* is higher in the sample of litigated cases that did not end in conviction than in the sample of cases disposed by trial conviction adds support for the notion that the probability of trial conviction is higher in bid-rigging cases than price-fixing cases.

To the extent that fewer counts of indictment implies weaker evidence, the finding that the mean value of *#COUNTS INDICT* is lower in litigated cases that end in conviction runs counter to the model's prediction that the probability of trial conviction rises with the strength of the prosecutor's evidence. If, on the other hand, prosecutors are likely to bundle charges when they lack any single piece of strong evidence, this finding could be interpreted to support the model's prediction.

If the prosecutor has weaker diminishing returns in trial expenditure than the defendant, then the trial conviction model predicts that the probability of trial conviction increases in the *ex ante* expected trial sentence, S^T. In addition, as Cohen (1992) and others hypothesize, S^T rises in the dollar value of commerce involved in the conspiracy. Thus, the finding that the mean value of *COMMERCE* is higher in the sample of litigated cases that end in conviction supports the trial conviction model's predictions.

The finding that the mean value of *VAR DOJ WIN PREV YEAR BY VIOL#1*, which represents the mean-preserving spread factor λ, is lower in the sample of litigated cases that end in conviction is interesting. In terms of the litigation model, this finding suggests that the observed litigated cases were relatively "close." That is, the trial conviction model predicts that the probability of trial conviction increases with λ only under the assumption that, when choosing their optimal levels of trial expenditure, both litigants think that their respective chances of winning are sufficiently close (e.g., the common *ex ante* expected probability of trial is numerically close to 50 percent). "Sufficiently close" is determined by the shape of the probability distribution function for trial outcomes. Starting from a probability of trial conviction close to 50 percent, an increase in the *mean-preserving* spread factor, λ, decreases the probability of trial conviction and increases the probability of trial conviction when the starting point is sufficiently near the tail of the probability distribution.[93]

Judicial discretion at trial may play a role in trial outcomes. Consider the finding that the mean of *PRIOR PROSECUTOR* is higher in the sample of cases disposed by trial conviction than in the sample of litigated cases that do not end in conviction. This finding suggests that judges who previously worked as government prosecutors tend to favor prosecutors' trial arguments over defendants' arguments and use their discretionary influence over a trial to benefit the prosecutors.

Same-Branch Means Tests of the Case Selection, Trial Selection, and Trial Conviction Selection Processes

This section reports additional types of two-sample means tests that test with specificity (and for the first time) the theoretical implications of the model of the prosecutor's case selection decision, and with more specificity, the defendant's plea decision and the judge's or jury's decision to convict (or acquit) at trial. Unlike the previously reported means tests, these tests focus on the predicted relationships among the values of case characteristics in five samples of observed cases created by the prosecutor's case selection decision, the defendant's plea decision, and the judge's (or jury's) decision to convict (or not) at trial. The selection processes represented by those decisions imply different mixes of the relative values of various pairs of case characteristics, as discussed in Chapters VI, VII, and VIII.

The following five sets of means tests are designed to capture selection "adjustments" observed in the mix of the relative values of various pairs of characteristics in the five observed samples of defendants created as cases are disposed.. Those samples represent defendants in cases that are i) initiated by the prosecutor (n = 3093); ii) initiated and proceed to trial (n = 601); iii) initiated and *do not* proceed to trial (n = 2492); iv) initiated, go to trial, and end in conviction (n = 132); or v) initiated, go to trial, and *do not* end in conviction (n = 469). The game tree diagram in Figure 3 shows which selection process – the prosecutor's case-bringing decision, the

---

93 Meanwhile, an increase in the *mean-preserving* spread factor, λ, increases the probability of trial conviction when the starting point is sufficiently near either tail of the probability distribution of trial conviction outcomes.

defendant's plea choice, or the judge's or jury's trial conviction decision – produces the five samples used in these means tests.

The following means tests involve fourteen variables that represent the case characteristics featured in the theoretical framework. While the means tests identify statistical significance, the tests employ binary variables that mitigate, if not eliminate, the potential for misleading results caused by observations with outlier values for certain variables.[94] The following list groups these fourteen variables by the associated case characteristic.

<u>Prosecutor's trial cost (C)</u>

HIGH DOJBUDGET

HIGH ANNUAL FORTUNE500

<u>Defendant's trial cost (K)</u>

ONLY LOW RANK

HIGH PROFIT

HIGH COMPENSATION

ZERO OTHER INDICT (also related to the amount or strength of evidence, **e**)

<u>Amount or strength of evidence (**e**)</u>

ZERO PREV CONVICT

HIGH DURATION

HIGH COMMERCE (also related to the expected trial sentence, S^T)

MULTI COUNTS CONVICT

MULTI COUNTS INDICT

<u>Trial sentence (S^T)</u>

POST-APPA INDICT YEAR

---

[94] The newly created binary variables were not used in the previous means tests because there was already a clear-cut method of dividing the samples for the previous means tests using the defendant's decision to go to trial (or not) and the judge's or jury's decision to convict at trial (or not). In these tests, dummy variables had to be created in order to create two samples based on the relative values of a continuous variable representing a case characteristic.

Probability of trial conviction ($\hat{\Pi}$)

BID RIG NOT PRICE FIX (also related to the idiosyncratic component of the prosecutor's *ex ante* expected probability of trial, ρ)

Idiosyncratic component of the prosecutor's *ex ante* expected probability of trial (ρ)

HIGH DOJ WIN PREV YEAR BY VIOL#1

In order to avoid redundancy, the means tests use each possible pair of relevant variables only once. All fourteen of the above variables are relevant as factors that directly influence indictment and plea decisions, thus creating ninety-one (91) individual means tests (i.e., 13 pairs + 12 pairs + 11 pairs $\cdots$ + 2 pairs + 1 pair = 91 possible pairs of the fourteen variables used for 91 means tests). The means tests associated with the trial stage of case disposition do not use the variable *HIGH DOJ WIN PREV YEAR BY VIOL#1* because ρ does not directly influence the probability of trial conviction in the specification of the litigation model. Thus, only thirteen of the variables listed above are relevant as factors that directly influence trial conviction decisions, which creates seventy-eight (78) means tests (i.e., 12 pairs + 11 pairs $\cdots$ + 2 pairs + 1 pair = 78 possible pairs of the thirteen variables used for 78 means tests).

The results of the individual means tests fall into three categories: i) statistically significant and consistent with theoretical predictions; ii) statistically significant and inconsistent with the theoretical expectations; and iii) statistically insignificant. In order to be categorized as consistent with the predictions of the theoretical models of the case selection, trial selection, and trial conviction processes, the results of the following means test must be consistent with the predicted relationships among pairs of variables representing case characteristics summarized previously in Tables 4, 6, and 8, respectively. For convenience, those tables are shown below.

Reporting the results of up to ninety-one (91) two-sample means tests is undesirable and unnecessary. It is instructive, however, to examine samples of the additional means tests conducted to test the theoretical predictions of the models of the case selection, trial selection, and trial conviction processes. The results of some illustrative means tests are reported later.

Tests of the case selection process

Data for potential federal criminal antitrust cases that prosecutors decided *not* to pursue are unavailable. Thus, it is not possible to test the difference in the means of the value of the same case characteristic in the sample of defendants in initiated cases versus the sample of potential defendants in uninitiated cases (as done in the first type of means tests of the implications of the plea and trial conviction decisions reported previously).

Despite the lack of data for *uninitiated* cases, another approach to testing the implications of the case selection model is possible. The alternative approach uses data exclusively from the sample of 3093 defendants in initiated cases and tests the implications for the relationships of values of case characteristics in cases the prosecutor decides to initiate. Specifically, the alternative approach tests the difference in the means of one empirical variable representing one case characteristic across samples created by the binary values of a different empirical variable representing either the same case characteristic[95] or a different case characteristic.

For example, the case selection model predicts the following combinations of case characteristics, in terms of the case characteristics' relative values, in *initiated* cases: (relatively high value of $\mathbf{e}$, relatively low value of *C*), (relatively high value of $\mathbf{e}$, relatively high value of *C*), and (relatively low value of $\mathbf{e}$, relatively low value of *C*). One way to state the rationale for those expected combinations of $\mathbf{e}$ and *C* is that, *ceteris paribus*, the prosecutor will *not* initiate a case with relatively weak evidence, $\mathbf{e}$, when her unit (opportunity) cost of trial expenditure, *C*, is relatively high (because of a low DOJ budget).

Accordingly, one means test examines the difference in the mean of *HIGH COMMERCE* in observed initiated cases in which *HIGH DOJBUDGET* = 1 and the mean of *HIGH COMMERCE* in observed initiated cases in which *HIGH DOJBUDGET* = 0. That is, the means test examines whether the prosecutor's evidence, $\mathbf{e}$, is relatively *strong* (e.g., *HIGH COMMERCE* = 1) or *weak* (e.g., *HIGH COMMERCE* = 0) in *initiated* cases in which the prosecutor's unit (opportunity) cost of trial expenditure, *C*, is relatively *high* (e.g., *HIGH DOJBUDGET* = 0) or *low* (e.g., *HIGH DOJBUDGET* = 1).

---

[95] Recall that within the set of empirical variables used in this study, sometimes different empirical variables represent the same case characteristic. For example, *ONLY LOW RANK, HIGH PROFIT, HIGH COMPENSATION,* and *ZERO OTHER INDICT* represent the defendant's unit cost of trial expenditure, *K*.

Table 4 summarizes the predicted relationships among the relative values of different pairs of case characteristics in the sample of indicted defendants (i.e., defendants in cases that prosecutors initiate). (Table 4 first appeared in Chapter VI.)

Table 4. Pairs of relative values of case characteristics in initiated cases and their consistency with the predictions of the case selection model

	High $\hat{\Pi}$	Low $\hat{\Pi}$	High C	Low C	High K	Low K	High S^T	Low S^T	High **e**	Low **e**	High ρ	Low ρ
High $\hat{\Pi}$	C*											
Low $\hat{\Pi}$	C*	I*										
High C	C	I	I*									
Low C	C	C	C*	C*								
High K	C	C	C	C	C*							
Low K	C	I	I	C	C*	I*						
High S^T	C	C	C	C	C	C	C*					
Low S^T	C	I	I	C	C	I	C*	I*				
High **e**	C	C	C	C	C	C	C	C	C*			
Low **e**	C	I	I	C	C	I	C	I	C*	I*		
High ρ	C	C	C	C	C	C	C	C	C	C	C*	
Low ρ	C	I	I	C	C	I	C	I	C	I	C*	I*

* As explained in Chapter X, different empirical variables can represent the same theoretical variable. Thus, combinations of relatively high and low values of the same theoretical variable represented by more than one empirical variable is possible for a given empirical observation. Pairs that are consistent with case selection model are denoted with "C" and pairs that are inconsistent are denoted with "I". Horizontal axis provides the possible relative values of first case characteristic in a possible pair. Vertical axis provides the possible relative values of second case characteristic in a possible pair.

To understand Table 4, as it relates to the following means tests, it is useful first to understand what relationships among the relative values of different pairs of case characteristics in an initiated case the model of the prosecutor's case selection process does *not* predict. For example, suppose an *initiated* case is characterized by relatively *high* values of the prosecutor's unit cost of trial expenditure, C, which is represented empirically by a value of zero (0) for the dummy variable *HIGH DOJBUDGET*.[96] This empirical variable represents the level of the DOJ Antitrust Division's annual budget relative to the total sample (n = 3093) mean value of the variable *DOJBUDGET*.[97] *Ceteris paribus*, a relatively *high* value of C is expected to be associated with a relatively *low* value of the prosecutor's *ex ante* expected value of inititating a case, $E(V)$. Thus, *ceteris paribus*, a relatively high value of C is not expected to characterize an initiated case *unless* the initiated case is further characterized, for example, by a relatively high value of the defendant's unit cost of trial expenditure, *K*, and/or a relatively high value of the amount or strength of the prosecutor's evidence, e, and/or a relatively high value of the *ex ante* expected trial sentence, S^T.

In this situation (i.e., *HIGH DOJBUDGET* = 0), the model of the prosecutor's case selection process does *not* predict that this *initiated* case also would be characterized by values of other case characteristics that are expected to *further reduce* $E(V)$. In this situation, relative values of other case characteristics *not* predicted by the model of the prosecutor's case selection process include relatively *low* values of the defendant's unit trial cost, K (e.g., *HIGH PROFIT* = 1); the level of the prosecutor's evidence, e (e.g., *HIGH DURATION* = 0); the parties' common expected trial sentence, S^T (e.g., *POST-APPA INDICT YEAR* = 0); the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$ (e.g., *BID RIG NOT PRICE FIX* = 0); or the idiosyncratic component, ρ, of the prosecutor's expected probability of trial conviction (e.g., *HIGH DOJ WIN PREV YEAR BY VIOL#1* = 0).

Continuing the preceding example, if an initiated case is characterized by relatively *high* values of C (e.g., *HIGH DOJBUDGET* = 0), then observed relatively *high* values of K (e.g.,

---

[96] The DOJ Antitrust Division's budgets are determined prosectively by Congressional appropriations. Lower budgets imply higher opportunity costs from going to trial. Thus, the value of *C* is inversely related to *DOJBUDGET* and *HIGH DOJBUDGET*. Similarly, the defendant's unit trial cost, *K*, is inversely related to the level of profits and employee compensation for the industry affiliated with the defendant, represented by *PROFIT* and *COMPENSATION*, respectively. Other empirical variables that represent *C* and *K* (e.g., *ANNUAL FORTUNE500* and *ONLY LOW*) are positively related to *C* and *K*. After considering the trade-offs in terms of ease of understanding, the empirical variables were coded according to the values of the empirical variables and not their associated theoretical variables.

[97] All of the dummy variables that were created especially for these additional means tests in order to capture relatively high (and low) values of case characteristics use the overall sample (n =3093) mean value as the benchmark.

HIGH PROFIT = 0), e (e.g., *HIGH DURATION* = 1), S^T (e.g., *POST-APPA INDICT YEAR* = 1), $\hat{\Pi}$ (e.g., *BID RIG NOT PRICE FIX* = 1), or ρ (e.g., *HIGH DOJ WIN PREV YEAR BY VIOL#1* = 1) in such an initiated case would support the predictions of the model of the prosecutor's case selection process. Meanwhile, if an initiated case is characterized instead by relatively *low* values of C (e.g., *HIGH DOJBUDGET* = 1), then observed relatively *high or low* values of K (e.g., *HIGH PROFIT* = 0 or 1), e (e.g., *HIGH DURATION* = 0 or 1), S^T (e.g., *POST-APPA INDICT YEAR* = 0 or 1), $\hat{\Pi}$ (e.g., *BID RIG NOT PRICE FIX* = 0 or 1), or ρ (e.g., *HIGH DOJ WIN PREV YEAR BY VIOL#1* = 0 or 1) in such an initiated cases also would support the predictions of the model of the prosecutor's case selection process.

In general, the combinations found in the sample of initiated cases that are consistent with the case selection model can be stated in terms of the prosecutor's value of bringing a case, E(V): (low value of E(V), high value of E(V)), (high value of E(V), low value of E(V)), and (high value of E(V), high value of E(V)). The combination (low value of E(V), low value of E(V)) is inconsistent with the case selection model. Table 12 reports the statistically significant results of illustrative two-sample means tests of the implications of the model of the prosecutor's case selection process summarized in Table 4.

Table 12. Sample same-branch means tests of indictment as a selection process

Variables Used in Two-Sample Means Tests	Evidence of Expected Case Selection "Adjustments" Predicted by the Model of the Prosecutor's Case Selection Process			Not Expected According to Case Selection Model
	High *C*, Δ E(V) > 0	Low *C*, Δ E(V) < 0	Low *C*, Δ E(V) > 0	High *C*, Δ E(V) < 0
ttest *HIGH ANNUAL FORTUNE500*, by(*HIGH DOJBUDGET*)*			1	1
ttest *ONLY LOW RANK*, by(*HIGH DOJBUDGET*)	0**	0	0	0
ttest *HIGH PROFIT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *HIGH COMPENSATION*, by(*HIGH DOJBUDGET*)	1	1		
ttest *ZERO OTHER INDICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *ZERO PREV CONVICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *HIGH DURATION*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH COMMERCE*, by(*HIGH DOJBUDGET*)			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *MULTI COUNTS INDICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH DOJBUDGET*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH DOJBUDGET*)			1	1
TOTALS	6	6	6	6

* The variable *HIGH DOJBUDGET*, which represents the prosecutor's unit cost of trial expenditure, *C*, pairs with other variables in the tests reported here.

** "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results.

A two-sample means test will always suggest the combination (high value of *C,* low value of *K*) or, more generally, (low value of E(V), low value of E(V)) when it suggests the combination (low value of *C,* high value of *K*) or, more generally, (high value of E(V), high value of E(V)). Thus, in the sample of defendants in initiated cases, the percentage of combinations that are consistent with the case selection model's predictions cannot fall below 50 percent and cannot exceed 100 percent of the combinations of pairs of variables the *statistically significant* means test results support.

As shown in Table 12, a full 75 percent of the combinations that the statistically significant means test results support are consistent with the predictions of the case selection model. For example, in the sample of 3093 defendants in initiated cases, when *HIGH DOJBUDGET* = 1, *HIGH PROFIT* = 1. This combination implies that in cases that are initiated, when the defendant's unit (opportunity) cost of trial expendicture, K, is low, the prosecutor's (opportunity) unit cost of trial expenditure, C , is low.

Although 50 percent marks the minimum of the range of the percentage of potential statistically significant combinations, 50 percent does not imply zero support for the case selection model. For example, the prosecutor could tolerate the combination (high value of *C,* low value of *K*) and initiate a case if the prosecutor sufficiently highly values another combination (or other combinations) of relative values of characteristics of the case. In that situation, 50 percent could imply support for model.

The case selection model's implications apply to both the sample of initiated cases and the sample of uninitiated cases. *Ideally*, this study would apply the means tests to both samples. This is not possible because data are not available for the sample of uninitiated cases. Of course, if the sample of uninitiated cases were available and the means tests were applied to that (hypothetical) sample of uninitiated cases, the results of the means tests applied to the (hypothetical) sample of uninitiated cases might add to or detract from the support provided by the means tests applied to the available sample of initiated cases.

It is reasonable for the unavailability of uninitiated case data to condition reliance on the results of the means tests of the implications of the case selection model. On the other hand, it is reasonable to expect that prosecutors non-randomly select cases to initiate. Thus, data limitations should not necessarily prevent attempts to use available information to assess observed case selection effects that are potentially important determinants of observed case outcomes – especially when no other study directly tests the implications of prosecutors' case selection decisions.

As reported in Appendix E, the results of the rest of the ninety-one (91) two-sample means tests are comparable to the results reported in Table 12. Specifically, 105 combinations corresponding to statistically significant results support the case selection model, while 37 statistically significant combinations are inconsistent with the predictions of the case selection model. Statistically insignificant results produce the remaining 40 combinations. Of the combinations corresponding to statistically significant results, 74 percent support the case selection model.

These ninety-one (91) two-sample means test results suggest that theoretical or empirical analyses of the determinants of federal criminal antitrust case outcomes (i.e., plea decisions, trial verdicts, and sentencing decisions) should not ignore the prosecutor's case selection process or treat it as a random selection process. These results also suggest that the case selection model developed in this study reasonably captures the process by which DOJ Antitrust Division prosecutors decided to initiate federal criminal antitrust cases against 3093 defendants from 1956 through 1980.

Tests of the trial selection process

The previous type of two-sample means tests of the implications of the model of the defendant's plea decision tested the difference in the mean values of the *same* empirical variable representing a particular case characteristic across *two* samples of cases: the sample of cases that go to trial (i.e., "litigated cases") and the sample of cases that do not go to trial (i.e. "non-litigated cases"). That is, the means tests reported previously focused on the relative values of the *same* empirical variable representing a particular case characteristic *across* game tree branches stemming from the node at which the defendant chooses whether to go to trial or to avoid trial by entering a plea of guilty or *nolo contendere*. The mixed results of those tests provided limited support for the model of the defendant's plea decision (i.e., trial selection process).

The following tests represent a different approach to testing the implications of the trial selection model. The tests examine the respective relative values of *pairs* of empirical variables, representing either the same or different case characteristics, in the *same* sample of cases: *first* in the sample of litigated cases, and *next* in the sample of non-litigated cases. That is, there is one distinct set of tests for each sample.

In other words, the tests focus on the respective relative values of *pairs* of empirical variables *along the same branch* emanating from the node at which the defendant chooses whether to go to trial or to avoid trial by entering a plea of guilty or *nolo contendere*. The tests are *not* the same as comparing, in a single test, values of empirical variables *across* game tree branches stemming from the node at which the defendant chooses whether to go to trial or to avoid trial, as was done previously.

Specifically, the tests use the sample of 601 defendants in litigated cases as well as the sample of 2492 defendants in non-litigated cases. The tests compare the difference in the mean values of one empirical variable representing one case characteristic across samples created by the binary values of another variable representing either the same case characteristic or a different case characteristic.

For example, the trial selection model predicts the following combinations of case characteristics, in terms of the case characteristics' relative values, in *litigated* cases: (relatively high value of S^T, relatively low value of K), (relatively high value of S^T, relatively high value of K), and (relatively low value of S^T, relatively low value of K). One way to state the rationale for those expected combinations of S^T and K is that, *ceteris paribus*, the defendant will *not* go to trial when the stakes of the case, as measured by the expected trial sentence, S^T is low, if his unit (opportunity) cost of trial expenditure, K, is relatively high.

Thus, one means test uses the sample of *litigated* cases and examines the difference in the mean of *HIGH COMMERCE* when *HIGH PROFIT* = 1 and the mean of *HIGH COMMERCE* when *HIGH PROFIT* = 0. That is, the means test examines whether the expected trial sentence, S^T, is relatively *high* (e.g., *HIGH COMMERCE* = 1) or *low* (e.g., *HIGH COMMERCE* = 0) in *litigated* cases in which the (corporate) defendant's unit (opportunity) cost of trial expenditure, K, is relatively *high* (e.g., *HIGH PROFIT* = 0) or *low* (e.g., *HIGH PROFIT* = 1).

Table 6 was introduced and explained in Chapter VII. Table 6 summarizes the predicted relationships among the relative values of different pairs of case characteristics in the sample of defendants in cases that go to trial as well as the sample of defendants in cases that end in a plea of guilty or *nolo contendere*.

Table 6. Pairs of relative values of case characteristics in cases that go to trial and their consistency with the predictions of the trial selection model

	High C	Low C	High K	Low K	High S^T	Low S^T	High ρ	Low ρ
High C	I*							
Low C	C*	C*						
High K	I	C	I*					
Low K	C	C	C*	C*				
High S^T	C	C	C	C	C*			
Low S^T	I	C	I	C	C*	I*		
High ρ	C	C	C	C	C	C	C*	
Low ρ	I	C	I	C	C	I	C*	I*

* As explained in Chapter X, different empirical variables can represent the same theoretical variable. Thus, combinations of relatively high and low values of the same theoretical variable represented by more than one empirical variable is possible for a given empirical observation. Pairs that are consistent with plea bargaining model are denoted with "C" and pairs that are inconsistent are denoted with "I". Horizontal axis provides the possible relative values of first case characteristic in a possible pair. Vertical axis provides the possible relative values of second case characteristic in a possible pair.

To understand Table 6, as it relates to the following means tests, it is useful to first understand what relationships among the relative values of different pairs of case characteristics in a case that goes to trial are *not* predicted by the model of the defendant's plea decision (i.e., trial selection) process. For example, suppose a case that goes to trial is characterized by a relatively *high* value of the prosecutor's unit cost of trial expenditure, C, which is represented empirically by a value of zero (0) for the dummy variable *HIGH DOJBUDGET*. *Ceteris paribus*, a case that goes to trial is *not* expected to have a relatively *high* value of C.

In this situation (i.e., *HIGH DOJBUDGET* = 0), the model of the defendant's plea decision (i.e., trial selection) process does *not* predict that the case that *goes to trial* also would be characterized by values of other case characteristics that are expected to *further reduce* $\widetilde{\Theta}$. In this situation, relative values of other case characteristics *not* predicted by the model of the defendant's trial selection process include relatively *high* values of the defendant's unit trial cost, K (e.g., *HIGH PROFIT* = 0); or relatively *low* values of the parties' common expected trial sentence, S^T (e.g., *POST-APPA INDICT YEAR* = 0); or the idiosyncratic component, ρ, of the prosecutor's expected probability of trial conviction (e.g., *HIGH DOJ WIN PREV YEAR BY VIOL#1* = 0).

Continuing the preceding example, if a case that goes to trial is characterized by a relatively *low* value of C (e.g., *HIGH DOJBUDGET* = 1), then observed relatively *high* values of S^T (e.g., *POST-APPA INDICT YEAR* = 1) or ρ (e.g., *HIGH DOJ WIN PREV YEAR BY VIOL#1* = 1), or relatively *low* values of K (e.g., *HIGH PROFIT* = 1) in such a case would support the predictions of the model of the defendant's trial selection process. Meanwhile, if a case that goes to trial is characterized instead by a relatively *low* value of C (e.g., *HIGH DOJBUDGET* = 1), then observed relatively *high or low* values of K (e.g., *HIGH PROFIT* = 0 or 1), e (e.g., *HIGH DURATION* = 0 or 1), S^T (e.g., *POST-APPA INDICT YEAR* = 0 or 1), or ρ (e.g., *HIGH DOJ WIN PREV YEAR BY VIOL#1* = 0 or 1) in such a case that goes to trial also would be consistent with the predictions of the model of the defendant's trial selection process.

In general, the combinations found in the sample of defendants who choose to go to trial that are consistent with the trial selection model can be stated in terms of the relative value of the *ex post* probability of trial, $\tilde{\Theta}$. The combinations consistent with the trial selection model's predictions are:: (low value of $\tilde{\Theta}$, high value of $\tilde{\Theta}$), (high value of $\tilde{\Theta}$, low value of $\tilde{\Theta}$), and (high value of $\tilde{\Theta}$, high value of $\tilde{\Theta}$). The combination (low value of $\tilde{\Theta}$, low value of $\tilde{\Theta}$) is inconsistent with the trial selection model. Table 13 reports the statistically significant results of illustrative two-sample means tests of the implications of the model of the defendant's trial selection process summarized in Table 6.

Table 13. Sample same-branch means tests of the defendant's plea decision as a selection process

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model			Not Expected According to Trial Selection Model
	High *C*, $\Delta\,\tilde{\Theta} > 0$	Low *C*, $\Delta\,\tilde{\Theta} < 0$	Low *C*, $\Delta\,\tilde{\Theta} > 0$	High *C*, $\Delta\,\tilde{\Theta} < 0$
ttest *HIGH ANNUAL FORTUNE500*, by(*HIGH DOJBUDGET*)*			1	1
ttest *ONLY LOW RANK*, by(*HIGH DOJBUDGET*)	0**	0	0	0
ttest *HIGH PROFIT*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH COMPENSATION*, by(*HIGH DOJBUDGET*)			1	1
ttest *ZERO OTHER INDICT*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *ZERO PREV CONVICT*, by(*HIGH DOJBUDGET*)***			1 (p<0.15)	1 (p<0.15)
ttest *HIGH DURATION*, by(*HIGH DOJBUDGET*)***			1	1
ttest *HIGH COMMERCE*, by(*HIGH DOJBUDGET*)***			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH DOJBUDGET*)***	1	1		
ttest *MULTI COUNTS INDICT*, by(*HIGH DOJBUDGET*)***	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH DOJBUDGET*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DOJBUDGET*)			1	1
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH DOJBUDGET*)			1	1
TOTALS	2	2	9	9

* The variable *HIGH DOJBUDGET*, which represents the prosecutor's unit cost of trial expenditure, *C*, pairs with other variables in the tests reported here.

** "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results.

*** Assume that, *ceteris paribus*, more or better evidence, e, increases ρ and/or S^T and $\tilde{\Theta}$.

As Table 13 shows, 13 of the 22 combinations corresponding to statistically significant means test results are consistent with the predictions of the trial selection model. As reported in Appendix F, the results of the rest of the ninety-one (91) two-sample means tests using data associated with 601 defendants in litigated cases lend somewhat more support to the trial selection model than the results reported in Table 13.

Specifically, 84 (or 68 percent) of the 124 combinations suggested by the statistically significant test results support the trial selection model, while 40 statistically significant combinations are inconsistent with the predictions of the case selection model. Statistically insignificant results produce the remaining 58 combinations.

Appendix G reports the results for ninety-one (91) means tests conducted using data associated with 2492 defendants in *non-litigated* cases. Recall from Chapter VII that if the combinations that are inconsistent with the predictions of the model of the defendant's trial selection process for the sample of cases that *go to trial* are observed in the sample of cases that do not go to trial, then those observed combinations in the sample of cases that *do not go to trial* provide support for the predictions of the trial selection model. That is, if a combination of case characteristic values *is not* conducive for a defendant choosing to go to trial, then that combination *is* conducive for a defendant choosing *not* to go to trial. Thus, all (statistically significant) observed combinations of case characteristic values in the sample of initiated cases can be used to test the implications of the trial selection model.

The results of the means tests using the sample of non-litigated cases provide more support for the trial selection model than the tests using the sample of litigated cases. Of the 146 combinations corresponding to statistically significant test results, 45 (or 31 percent) support the trial selection model, while 101combinations contradict the trial selection model. The remaining 36 combinations correspond to statistically insignificant test results.

Determining the total support for the model from these means tests requires some explanation and calculations. A means test always will suggest (high value of *C,* high value of *K*) or, more generally, (low value of $\widetilde{\Theta}$, low value of $\widetilde{\Theta}$) when it suggests (low value of *C,* low value of *K*) or, more generally, (high value of $\widetilde{\Theta}$, high value of $\widetilde{\Theta}$). Thus, in the sample of defendants *who choose to go to trial*, the percentage of statistically significant combinations consistent with the trial selection model cannot fall below 50 percent and cannot exceed 100 percent of the *statistically significant* combinations of pairs of variables that the means tests produce.

Similarly, a means test always will suggest (low value of *C,* low value of *K*) or, more generally, (high value of $\widetilde{\Theta}$, high value of $\widetilde{\Theta}$) when it suggests (high value of *C,* high value of *K*) or, more generally, (low value of $\widetilde{\Theta}$, low value of $\widetilde{\Theta}$). Thus, using the sample of defendants who choose *not* to go to trial, the percentage of statistically significant combinations consistent with the trial selection model cannot fall below zero percent and cannot exceed 50 percent of the *statistically significant* combinations of pairs of variables that the means tests produce.

Given that the first test produced 124 statistically significant combinations and the second test produced 146, the results of the first and second sets of means tests should be weighted by 0.46 (= 124/270) and 0.54, respectively. Thus, 50 percent (= 46 * (68-50)/50 + 54 * (31/50)) of the combinations corresponding to statistically significant results are consistent with the trial selection model's predictions.

On balance, these two-sample means test results suggest that theoretical or empirical analyses of the determinants of trial verdicts and sentencing outcomes in federal criminal antitrust cases should not treat the trial selection process as a random process. These results also suggest that the trial selection model reasonably captures the process by which 3093 defendants chose to go to trial or to avoid trial in federal criminal antitrust cases initiated from 1956 through 1980.

At the same time, the divergence of support provided by tests using the sample of defendants in litigated cases versus the sample of defendants in non-litigated cases merits further consideration. The larger (relevant) percentage of statistically significant results supporting the theory using the sample of defendants in non-litigated cases is striking.

These results suggest at least two possibilities. First, the determinants of the defendant's decision to go to trial might differ systematically from the determinants of his decision to avoid trial by pleading guilty or *nolo contendere*.[98] Second, but related to the first possibility, it is

[98] One reason for this suspicion is the differences in the implications of pleading not guilty, guilty, or *nolo contendere*. As the statistics reported by Gallo *et al*. (2000) make clear, *nolo contendere* pleas are quite popular in federal criminal antitrust cases. The popularity of the *nolo contendere* plea stems from the defendant's expected financial exposure to follow-on civil suits for treble damages following a *nolo contendere* plea. While a plea of *nolo contendere* is a conviction, it does not provide *prima facia* evidence of a criminal violation in a follow-on civil suit and it forces plaintiffs to prove the violation. Moreover, proving a violation in follow-on civil suits can be costly for potential plaintiffs. For example, between the time of the disposition of the criminal case and the beginning of the

quite plausible that corporate defendants and white-collar individual defendants, as well as their (presumably) expert attorneys likely base decisions or advice on entered pleas on considerations not easily captured by the empirical variables created from available data. For example, the defendant's risk cost from going to trial (possibly based in part on his attorney's assessment) could be a case characteristic that drives his plea decision. Since the defendant's risk cost from going to trial includes potential damage to his professional or personal reputation, it is a very difficult case characteristic to measure empirically.

Tests of the trial conviction selection process

The results of the first type of two-sample means tests of the implications of the model of the judge's (or jury's) decision to convict the defendant at trial were presented earlier. Those means tests tested the difference in the mean values of the same empirical variable representing a particular case characteristic *across two samples*: the sample of litigated cases that end in conviction and the sample of litigated cases that do not end in conviction. In other words, the previous means tests focused on the relative values of the same empirical variable across game tree branches stemming from the trial node at which the judge or jury decides whether to convict the defendant or not convict the defendant. Those tests' mixed results provided reasonable support for the model of the judicial trial conviction process.

The following tests represent a different approach to testing the implications of the trial conviction model. The tests examine the respective relative values of *pairs* of empirical variables, representing either the same or different case characteristics, in the *same* sample of cases: *first* in the sample of litigated cases that end in conviction, and *next* in the sample of litigated cases that do *not* end in conviction. That is, there is one distinct set of tests for each sample.

In other words, the tests focus on the respective relative values of *pairs* of empirical variables *along the same branch* emanating from the node at which the judge or jury decides whether to convict the defendant at trial. These tests are *not* the same as comparing values of empirical variables *across* game tree branches in a single test, as was done previously.

follow-on civil case, the defendant could bolster his defense strategy or introduce new facts or facts that he otherwise did not reveal during the criminal case. Thus, nolo contedere pleas can deter follow-on civil suits or at least reduce the probability of follow-on civil suits in which plaintiffs win.

Specifically, the tests use the sample of 132 defendants in litigated cases that end in conviction as well as the sample of 469 defendants in litigated cases that do not end in conviction. The tests compare the difference in the mean values of one empirical variable representing one case characteristic across samples created by the binary values of another variable representing the same case characteristic or a different case characteristic.

For example, the trial conviction model predicts the following combinations of case characteristics, in terms of the case characteristics' relative values, in litigated cases *that end in conviction*: (relatively low value of *C*, relatively high value of *K*), (relatively low value of *C*, relatively low value of *K*), and (relatively high value of *C*, relatively high value of *K*). One way to state the rationale for those expected combinations of *C* and *K* is that, *ceteris paribus*, the judge will *not* convict the defendant when the defendant's unit cost of trial expenditure, *K*, is relatively low, if the prosecutor's unit cost of trial expenditure, *C*, is relatively high.

Thus, one means test uses the sample of litigated cases *that end in conviction* and examines the difference in the mean of *HIGH DOJBUDGET* when *HIGH PROFIT* = 1 and the mean of *HIGH DOJBUDGET* when *HIGH PROFIT* = 0. That is, the means test examines whether the prosecutor's unit (opportunity) cost of trial expenditure, *C*, is relatively *high* (e.g., *HIGH DOJBUDGET* = 0) or *low* (e.g., *HIGH DOJBUDGET* = 1) in litigated cases *that end in conviction* in which the (corporate) defendant's unit (opportunity) cost of trial expenditure, *K*, is relatively *high* (e.g., *HIGH PROFIT* = 0) or *low* (e.g., *HIGH PROFIT* = 1).

Table 8 was introduced and explained in Chapter VIII. Table 8 summarizes the predicted relationships among the relative values of different pairs of case characteristics in the sample of defendants in cases that end in trial conviction as well as the sample of defendants in cases that end in acquittal, dismissal, or directed verdict.

Table 8. Pairs of relative values of case characteristics in cases that end in trial conviction and their consistency with the predictions of the trial conviction model

	High C	Low C	High K	Low K	High S^T	Low S^T	High **e**	Low **e**
High C	I*							
Low C	C*	C*						
High K	C	C	C*					
Low K	I	C	C*	I*				
High S^T	C	C	C	C	C*			
Low S^T	I	C	C	I	C*	I*		
High **e**	C	C	C	C	C	C	C*	
Low **e**	I	C	C	I	C	I	C*	I*

* As explained in Chapter X, different empirical variables can represent the same theoretical variable. Thus, combinations of relatively high and low values of the same theoretical variable represented by more than one empirical variable is possible for a given empirical observation. Pairs that are consistent with plea bargaining model are denoted with "C" and pairs that are inconsistent are denoted with "I". Horizontal axis provides the possible relative values of first case characteristic in a possible pair. Vertical axis provides the possible relative values of second case characteristic in a possible pair.

Table 8 summarizes the predicted relationships among the relative values of different pairs of case characteristics in cases that go to trial. To understand Table 8 it is useful to first understand what relationships among the relative values of different pairs of case characteristics in a litigated case that ends in trial conviction are *not* predicted by the trial conviction model.

For example, suppose a litigated case that ends in trial conviction is characterized by a relatively *high* value of the prosecutor's unit cost of trial expenditure, C, which is represented empirically by a value of zero (0) for the dummy variable *HIGH DOJBUDGET*. *Ceteris paribus*, a litigated case that ends in trial conviction is not expected to be associated with a relatively *high* value of C, since C is expected to *negatively* directly affect the *ex post* probability of trial conviction, $\widetilde{\Pi}$.

In that situation, the model of the judge's trial conviction decision rule (i.e., trial conviction process) does *not* predict that the case that *ends in a trial conviction* also would be characterized by values of other case characteristics that are expected to *further reduce* $\widetilde{\Pi}$. In this situation, relative values of other case characteristics *not* predicted by the model of the trial conviction process include relatively *low* values of the defendant's unit trial cost, K (e.g., *HIGH PROFIT* = 1); the parties' common expected trial sentence, S^T; (assuming the prosecutor has weaker diminishing returns in trial expenditure than the defendant) (e.g., *POST-APPA INDICT YEAR* = 0), or the amount or strength of the prosecutor's evidence, e (e.g., *HIGH DURATION* = 0).

Continuing the preceding example, if a case that is disposed by trial conviction is characterized by a relatively *high* value of C, then observed relatively *high* values of S^T (e.g., *POST-APPA INDICT YEAR* = 1), e (e.g., *HIGH DURATION* = 1), or K (e.g., *HIGH PROFIT* = 0) in such a case would be consistent with the predictions of the model of the judge's trial conviction decision rule. Meanwhile, if a case that ends in trial conviction is characterized instead by a relatively *low* value of C (e.g., *HIGH DOJBUDGET* = 1), then observed relatively *high or low* values of K (e.g., *HIGH PROFIT* = 0 or 1), S^T (e.g., *POST-APPA INDICT YEAR* = 0 or 1), or e (e.g., *HIGH DURATION* = 0 or 1) in such a case would be consistent with the predictions of the model of the trial conviction decision rule.

In general, the combinations found in the sample of defendants who *are convicted at trial* that are consistent with the trial conviction model's predictions can be stated in terms of the *ex post* probability of trial conviction, $\tilde{\Pi}$, as follows: (low value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$), (high value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$), and (high value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$). The combination (low value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$) is inconsistent with the trial conviction model's predictions. Table 14 reports the statistically significant results of illustrative two-sample means tests of the implications of the model of the judicial trial conviction rule summarized in Table 8.

Table 14. Sample same-branch means tests of the trial conviction decision as a selection process

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule			Not Expected According to Trial Conviction Model
	High *C*, $\Delta\,\tilde{\Pi} > 0$	Low *C*, $\Delta\,\tilde{\Pi} < 0$	Low *C*, $\Delta\,\tilde{\Pi} > 0$	High *C*, $\Delta\,\tilde{\Pi} < 0$
ttest *HIGH ANNUAL FORTUNE500*, by(*HIGH DOJBUDGET*)*			1	1
ttest *ONLY LOW RANK*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH PROFIT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *HIGH COMPENSATION*, by(*HIGH DOJBUDGET*)	1	1		
ttest *ZERO OTHER INDICT*, by(*HIGH DOJBUDGET*)	0**	0	0	0
ttest *ZERO PREV CONVICT*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *HIGH DURATION*,*** by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH COMMERCE*, *** by(*HIGH DOJBUDGET*)			1	1
ttest *MULTI COUNTS CONVICT*, *** by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, *** by(*HIGH DOJBUDGET*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH DOJBUDGET*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DOJBUDGET*)			1	1
TOTALS	3	3	6	6

* The variable *HIGH DOJBUDGET*, which represents the prosecutor's unit cost of trial expenditure, *C*, pairs with other variables in the tests reported here.

** "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results.

*** Assume that, *ceteris paribus*, more evidence, e, increases the value of $\tilde{\Pi}$.

Of the 18 statistically significant combinations associated with the results reported in Table 14, 12 (or 67 percent) support the trial conviction model. As reported in Appendix H, the results of the rest of the 78 tests using data associated with 132 defendants in litigated cases disposed by conviction are comparable to the results reported in Table 14.

In fact, 53 (or 72 percent) of the 74 combinations corresponding to statistically significant results support the trial conviction model, while 21 combinations corresponding to statistically significant results are inconsistent with the predictions of the trial conviction model. A large number, 82, of test results using data from the set of defendants convicted at trial are statistically insignificant, however.

Appendix I reports the results for an additional 78 means tests conducted using data associated with 469 defendants in litigated cases not disposed by conviction. As explained in Chapter VIII, if the combinations that are inconsistent with the predictions of the model of the judicial trial conviction decision for the sample of cases that *are disposed by trial conviction* are observed in the sample of cases that *go to trial but do not end in trial conviction*, then those observed combinations in the sample of cases that *go to trial but do not end in trial conviction* provide support for the predictions of the model of the judicial trial conviction decision.

That is, if a combination of case characteristic values *is not* conducive for a judge *convicting* a defendant at trial, then that combination *is* conducive for a judge *acquitting* that defendant at trial (or dismissing the case or issuing a directed verdict). Thus, all (statistically significant) observed combinations of case characteristic values in the sample of litigated cases can be used to test the implications of the trial conviction model.

In general, the combination found in the sample of defendants who *go to trial but are not convicted* that is consistent with the trial conviction model's predictions can be stated in terms of the *ex post* probability of trial conviction, $\tilde{\Pi}$, as follows: (low value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$). The combinations (low value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$), (high value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$), and (high value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$) are inconsistent with the trial conviction model's predictions.

The results of the means tests using the sample of defendants in litigated cases that are not disposed by trial conviction provide more empirical support for the trial conviction model. Of the 100 combinations suggested by statistically significant results, 28 support the trial conviction model. Statistically insignificant results suggested 56 combinations.

Determining the total support for the model from these means tests requires some explanation and calculations. A means test always will suggest (high value of *C,* low value of *K*) or, more generally, (low value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$), when it suggests (low value of *C,* high value of *K*) or, more generally, (high value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$). Thus, using the sample of defendants *convicted at trial*, the percentage of statistically significant combinations that are consistent with the trial conviction model cannot fall below 50 percent and cannot exceed 100 percent of the *statistically significant* combinations of pairs of variables that the means tests report. The reported results should be judged accordingly.

Similarly, a means test always will suggests the combination (low value of *C,* high value of *K*) or, more generally, (high value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$) when it suggests (high value of *C,* low value of *K*) or, more generally, (low value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$). Thus, using the sample of defendants *who went to trial but were not convicted*, the percentage of combinations consistent with the trial conviction model's predictions cannot fall below zero percent and cannot exceed 50 percent of the statistically significant combinations of pairs of variables that the means tests produce. The reported results should be judged accordingly.

Given that the first test of the trial conviction model produced 74 statistically significant combinations and the second test produced 100, the results of the first and second sets of means tests should be weighted by 0.43 (= 74/174) and 0.57, respectively. Thus, over half – about 51 percent (= 43 * (72-50)/50 + 57 * (28/50)) – of the combinations associated with statistically significant results support the trial conviction model.

Notwithstanding the number of statistically insignificant means test results, these test results suggest that theoretical or empirical analyses of the determinants of federal criminal antitrust trial sentencing outcomes should not treat the trial selection process as a random process. These results also suggest that this study's model of the judicial trial conviction decision reasonably captures the process by which judges or juries delivered trial verdicts in federal criminal antitrust cases initiated from 1956 through 1980.

The theoretical model of the judicial trial conviction decision makes reasonable use of the range of factors that determine a trial verdict. These tests employ a range of empirical variables related to the theoretical factors. Thus, the test results suggest that the trial conviction model stands as a reasonable characterization of the non-random process by which a judge or jury decides to convict a defendant at trial.

Still, some intuitively appealing factors are difficult to represent theoretically or empirically. Such factors include the credibility of the witnesses or their performances on the witness stand, since credibility does not ensure persuasiveness in favor of the desired litigant. Such factors also include the comparative or absolute strengths of the attorneys' preparations for trial or performances at trial; as well as the political, social, or moral persuasions of the jury and/or judge. Given the potential influence of such factors, neither the number of statistically significant results nor the limited amount of model support they provide is surprising.

Regression Analysis

As explained in Chapters VI, VII, and VIII, the theoretical models of the *ex post* probabilities of trial and trial conviction, as well as sentences imposed after trial conviction produce ambiguous predictions regarding the total marginal effects of marginal increases in the values of case characteristics. That is because of the predicted influence of indirect selection effects produced by the prosecutor's case selection process, the defendant's trial selection process, and the process by which the judge or jury decides to convict the defendant at trial. Tables 3, 5, and 7, which appear below, summarize the predicted confounding influence of the indirect selection effects.

This study estimates regression models of the probability of trial, the probability of trial conviction, as well as fines imposed after trial conviction. Like Snyder (1990) and Cohen (1992), this study recognizes that trial selection effects are determinants of plea bargaining, trial, and sentencing outcomes. As Snyder (1990) and Cohen (1992) note, when estimating models of case outcomes, it is important to recognize the sequential nature of the case disposition process.

For example, plea decisions determine the sample of cases that go to trial, which implies that plea decisions may not randomly determine the set of defendants that go to trial. This implies that the set of defendants that go to trial is a biased representation of the set of indicted defendants. Thus, observed changes in the trial conviction rate can be due to the change in the distribution of defendants who go to trial from the distribution of indicted defendants. In this way, trial selection bias can confound the interpretations of estimates of the effects of other factors (e.g., behavior of litigants or judges, institutional changes in antitrust enforcement).[99]

Following Heckman's (1979) method, this study uses the estimated model of the probability of trial to estimate an inverse Mills ratio for all observations used in the regression models of the probability of trial conviction and fines imposed following trial convictions. An inverse Mills ratio associated with the probability of trial is an explanatory variable in the regression models of trial and sentencing outcomes. Likewise, the specification of the regression model of fines imposed after trial convictions includes an inverse Mills ratio associated with the probability of trial conviction.

Because data regarding uninitiated cases are unavailable, this study does not estimate a regression model of prosecutors' decisions to initiate cases. In other words, the unavailability of uninitiated case data renders this study's regression models unable to identify case selection effects. Thus, the estimated regression coefficients for the explanatory variables representing case characteristics are estimates of the combined (first- and second-order) direct and indirect case selection effects on the *ex post* probabilities of trial and trial conviction from marginal increases in the values of case characteristics.[100]

Such data limitations do not prevent the regression analysis to examine the implications of trial selection effects and trial conviction selection effects. For example, calculations involving estimated regression coefficients provide the estimated trial selection effects on the *ex post* probability of trial conviction from marginal increases in the values of case characteristics.

With these estimated trial selection effects and other regression estimates, the regression analysis complements the previous means tests in assessing the implications of the trial selection model and trial conviction model. In general, the regression analysis supports the importance of

---

[99] With respect to another consideration in regression analysis, Cohen (1992) and Snyder (1989, 1990) note the potential for heteroskedasticity because defendants in the same case likely do not constitute independent observations. They control for heteroskedasticity by clustering observations with similar types of defendants (individual or corporate) who make similar plea choices in the same case. This aggregation approach sacrifices a significant amount of information.

The regression models in this study control for heteroskedasticity using the "cluster" option in the statistics computer program *Stata*. Observations used in the regressions are "clustered" by the Commerce Clearing House case number. This approach yields robust standard errors, as suggested by White (1980, 1982). This approach is similar to the approach for controlling for heteroskedasticity used in Siegfried and Sapper (1994).

[100] In other words, the estimated coefficients in this study's regression models represent the (first- and second-order) direct effects of marginal increases in the values of case characteristics conditioned by prosecutors' decisions to indict the observed defendants (i.e., initiate the observed cases).

trial selection and trial conviction selection as determinants of federal criminal antitrust case outcomes.

Meanwhile, this study depends wholly on the regression models of the probabilities of trial and trial conviction to conduct tests for the existence and possessor of private (or better) information during plea bargaining in federal criminal antitrust cases. As highlighted in the literature review in Chapter II, the criminal antitrust literature ignores (or does not explicitly address) the issue of how private information in plea bargaining could alter the trial selection process and create (additional) bias in observed case outcomes. This is a significant omission given the policy implications of private information during plea bargaining.

Privately informed prosecutors harm social welfare because they can fashion a settlement offer that induces (truly) innocent defendants to plead guilty or *nolo contendere*. Such pleas send false signals about legal and socially efficient business conduct. In contrast, privately informed defendants during plea bargaining benefit social welfare because the (truly) guilty ones are more likely to plead guilty and send the correct signals about appropriate business conduct.

This study's regression-based tests for private information during plea negotiations use the signs of the estimated coefficients for explanatory variables that do not represent the case characeristics used in the theoretical framework. Those explanatory variables represent the passage of time, the occurance of landmark antitrust court decisions, and the inverse Mills ratio (or Heckman's lambda) that controls for the estimated probability that the observed defendant would go to trial. [101]

The use of the estimated coefficients of those explanatory variables for testing the information structure of plea bargaining are described in Chapter IX. These tests produce mixed evidence regarding the information structure of plea bargaining. Some of the tests suggest that during plea bargaining the observed defendants possessed private (or better) information regarding their chances at trial, some of the tests suggest that the prosecutors held private information, and some tests suggest that neither litigant was privately informed.

[101] The overall trial conviction rate for the sample of defendants that go to trial also will be used to test the information structure of plea bargaining. This test is also discussed in Chapter IX.

Probit model of the observed trial rate

As explained in Chapter VI (and by reference, Appendix A), the theoretical model of the *ex post* probability of trial produces ambiguous predictions for the total marginal effects on the *ex post* probability of trial from marginal increases in the values of case characteristics. That occurs because of the predicted confounding influence of second-order direct effects and indirect selection effects produced by the prosecutor's case selection process, as shown in Table 3 from Chapter VI.

Table 3. Signs of expected direct and indirect effects on the observed trial rate, $\tilde{\Theta}$

Case Characteristic	Expected First-Order Direct Effect on $\tilde{\Theta}$	Expected Net Second-Order Direct "Trial Expenditure" Effect on $\tilde{\Theta}$ *		Expected Net Second-Order Direct "Stakes" Effect on $\tilde{\Theta}$ *		Expected Indirect Case Selection Effect on $\tilde{\Theta}$	*Expected Total (Net) Effect on* $\tilde{\Theta}$
		When $\hat{\Pi} \geq \overline{\Pi}$	When $\hat{\Pi} < \overline{\Pi}$	When $\hat{\Pi} \geq \overline{\Pi}$	When $\hat{\Pi} < \overline{\Pi}$		
$\hat{\Pi}$ **	0	0		0		?	?
C	-	?	+	-	?	?	?
K	-	+	?	?	+	?	?
S^T **	+	-		0		?	?
e	0	+	-	?		?	?
λ	0	?		?		?	?
ρ	+	0		0		?	?

* As discussed above and, in detail, in Appendix A,.these second-order direct effects involve movements in both litigants' optimal levels of trial expenditure, sometimes in the same direction but sometimes in opposite directions. Hence the "net" reference.

** The marginal increases in $\hat{\Pi}$ and S^T are from exogenous shocks.

Table 15 provides the coefficients – not the marginal effects – estimated by a probit regression model of the defendant's decision to enter either a plea of not guilty (*TRIAL* = 1) and go to trial, or a plea of guilty or *nolo contendere* (*TRIAL* = 0). This probit model uses data regarding 3093 defendants associated with 411 federal criminal antitrust cases initiated from 1956 through 1980.

Table 15. Probit model of the *ex post* probability of trial

Dependent variable: *TRIAL*

Explanatory Variables	Sign of the Expected First-Order Direct Effect	Estimated Coefficient	Robust Standard Error†	Z	P > |z|
Constant	+	2.478745	2.428912	1.02	0.307
INDICT YEAR	?	-.0496947	.0486459	-1.02	0.307
LANDMARK	?	-.1627714**	.081754	-1.99	0.046
DOJBUDGET	+	-.0000156	.0001117	-0.14	0.889
%CRIMINAL	?	-.5892453	.9331985	-0.63	0.528
ANNUAL FORTUNE500	-	.0098087	.0089952	1.09	0.276
REPUB PRES INDICT YEAR	?	.2753818	.2035389	1.35	0.176
#CODEFEND	?	-.0197135	.0141426	-1.39	0.163
#CORP CODEFEND	+	.0500579	.0311249	1.61	0.108
ONLY INDIV DEFEND	-	.4634297 *	.2486836	1.86	0.062
ANY LOW RANK	-	.1200737	.0932485	1.29	0.198
#OTHER INDICT	-	-.1131335 **	.0470123	-2.41	0.016
#PREV CONVICT	+	.0375383	.0410079	0.92	0.360
PROFIT	+	.0001367 ***	.0000509	2.69	0.007
COMPENSATION	+	-.0000277 ***	9.61e-06	-2.88	0.004
DURATION	+	.0014363	.015017	0.10	0.924
COMMERCE	+	-6.24e-07	5.94e-07	-1.05	0.294
# COUNTS INDICT	+	-.02146	.1698846	-0.13	0.899
ANY BID RIG	+	.0210531	.4535056	0.05	0.963
BID RIG NOT PRICE FIX	+	.7357755 **	.3613693	2.04	0.042
PRICE FIX NOT BID RIG	?	.27898	.4153695	0.67	0.502
CONSUMER VIC	?	.2437632	.1971788	1.24	0.216

† Standard error adjusted for 411 clusters of Commerce Clearinghouse-defined cases.
* 0.10 level of signficance.
** 0.05 level of significance.
*** 0.01 level of significance.

Table 15, continued

Dependent variable: *TRIAL*

Explanatory Variables	Sign of the Expected First-Order Direct Effect	Estimated Coefficient	Robust Standard Error†	Z	P > |z|
DOJ WIN PREV YEAR BY VIOL#1	+	-.0160364	.0112504	-1.43	0.154
(DOJ WIN PREV YEAR BY VIOL#1)^2	?	.0002253 *	.0001268	1.78	0.076
POST-APPA INDICT YEAR	+	.7429855 *	.4026015	1.85	0.065
Regression Statistics					
N	3093				
Wald χ^2 (25)	58.88				
Prob > χ^2	0.0001				
Log pseudolikelihood	-1334.8585				
Pseudo R^2	0.1236				

† Standard error adjusted for 411 clusters of Commerce Clearinghouse-defined cases.
* 0.10 level of signficance.
** 0.05 level of significance.
*** 0.01 level of significance.

As shown in Table 15, the estimated coefficients for several explanatory variables representing theoretical model variables are statistically significant. Table 16 provides the estimated marginal effects on the *ex post* probability of trial from marginal increases in the values of the explanatory variables that had statistically significant estimated coefficients in the probit regression model reported in Table 15.

Because we cannot observe the set of potential cases that the prosecutor chooses not to pursue, the regression model summarized in Table 15 also does not identify indirect case selection effects on the *ex post* probability of trial from marginal increases in the explanatory variables representing case characteristics. Thus, the marginal effects reported in Table 16 are conditional on prosecutors initiating the observed cases. That is, the marginal effects reported in Table 16 include direct and case selection effects.

Table 16. Estimated marginal direct effects on the *ex post* probability of trial, conditional on prosecutors initiating the observed cases

Dependent Variable (y): Pr(*TRIAL*)

Mean Value of Pr(*TRIAL*) = 0.16219994

Explanatory Variable	Sign of the Expected First-Order Direct Effect	Estimated Marginal Effect (dy/dx)	Standard Error	Z	P>|z|	Mean Value of Explanatory Variable (x)
LANDMARK	?	-.0399587	.02042	-1.96	0.050	3.28449
*ONLY INDIV DEFEND**	-	.137314	.0851	1.61	0.107	.027481
#OTHER INDICT	-	-.0277731	.01162	-2.39	0.017	.383123
PROFIT	+	.0000336	.00001	2.66	0.008	2806.85
COMPENSATION	+	-6.80e-06	.00000	-2.88	0.004	15782.1
*BID RIG NOT PRICE FIX**	+	.2260117	.12944	1.75	0.081	.121888
(DOJ WIN PREV YEAR BY VIOL#1)^2	?	.0000553	.00003	1.79	0.074	1710.16
*POST-APPA INDICT YEAR**	+	.2092021	.12634	1.66	0.098	.28516

* Estimated marginal effect (dy/dx) is for discrete change of dummy variable from 0 to 1.

According to the estimates reported in Table 16, the probability of trial fell by 0.040 percent with increases in the average number of Supreme Court decisions in the presidential administration prior to the observed year that three law treatises cited by Gallo *et al*. (2000) consider "landmark" decisions (*LANDMARK*). The implications of this finding and others regarding the information structure of plea bargaining in the observed cases are discussed later.

If the defendants in a case are only individual defendants (*ONLY INDIV DEFEND*), then the probability of trial increases by 0.137 percent, *ceteris paribus*. This is contrary to the predicted negative first-order direct effect of an increase in the defendant's unit cost of trial expenditure (*K*).

Ceteris paribus, the probability of trial in the observed case is 0.028 percent lower for every other indictment in another case that the observed defendant faces in the year of the observed case (*#OTHER INDICT*). It is reasonable to expect the defendant's unit cost of trial expenditure (K) to increase with the number of legal disputes in which he is involved. This finding is therefore consistent with the predicted negative first-order direct effect on the probability of trial from an increase in the defendant's unit cost of trial expenditure.

Ceteris paribus, every additional thousand dollars (in real dollars, 1967 = 100) in the profit level (*PROFIT*) and total employee compensation level (*COMPENSATION*) of the industry with which the defendant is associated raised and lowered by trace amounts the probability of trial, respectively. The defendant's unit (opportunity) cost of trial expenditure is expected to fall with increases in these variables. Thus, the former finding is consistent, while the latter finding is inconsistent, with the predicted positive first-order direct effects on the probability of trial from a decrease in the defendant's unit cost of trial expenditure.

The probability that defendants charged with big-rigging but not price-fixing (*BID RIG NOT PRICE FIX*) would plead not guilty is 0.226 percent higher than it is for defendants not charged with bid-rigging, *ceteris paribus*. As previously discussed, Snyder (1989) explains that bid-rigging cases are relatively easy to prosecute successfully. Thus, in cases that involve bid-rigging but not price-fixing charges, the idiosyncratic component, ρ, of the prosecutor's *ex ante* expected probability of trial conviction is expected to be relatively high. This finding is consistent with the predicted positive first-order direct effect on the probability of trial from an increase in ρ.

The trial conviction rate in the previous observed year for cases involving the same charge as the first charge in the indictment of the observed defendant (*DOJ WIN PREV YEAR BY VIOL#1*), which is positively related to ρ, had a statistically insignificant effect on the probability that the observed defendant would plead not guilty. The square of that variable, however, had a positive but de minimus impact on the probability of trial, however. This finding regarding the squared term could be construed as consistent with the predicted positive first-order direct effect on the probability of trial from an increase in ρ. The theory does not specify whether the marginal effect on the observed trial rate from a change in the value of ρ – whether the change in ρ is positive or negative – is expected to be increasing or decreasing.

Finally, the probability that defendants indicted in years following the passage of the Antitrust Policy and Procedures Act of 1974 (*POST-APPA INDICT YEAR*) would go to trial is 0.209 percent higher than the probability of trial is for defendants indicted before the passage of the Act. To the extent that the passage of the Act induced defendants to expect judges to impose harsher penalties, increases in this explanatory variable capture increases in the value of the *ex ante* expected penalty imposed after trial conviction, S^T. This finding is thus consistent with the predicted positive first-order direct effect on the probability of trial from increases in S^T.

These estimated marginal (total) direct effects of increases in the values of case characteristics on the probability of trial have only purely descriptive value in this study. As previously explained, the estimated regression coefficients for the explanatory variables representing case characteristics are not readily useful, if at all, to test the predicted total marginal effects of changes in the values of case characteristics on the *ex post* probability of trial.

That is because, as previously mentioned, the estimated regression coefficients do not identify the second-order direct trial expenditure or stakes effects. Nor do the estimated coefficients identify indirect case selection effects on the *ex post* probability of trial from marginal increases in the explanatory variables representing case characteristics. Further discussion of the estimated marginal effects for variables representing case characteristics is beyond the scope of this study.

Probit model of the observed trial conviction rate

As explained in Chapter VII, the theoretical model of the *ex post* probability of trial conviction produces ambiguous predictions for the total marginal effects on the *ex post* probability of trial from marginal increases in the values of case characteristics. That is because of the predicted confounding influence of indirect selection effects produced by the prosecutor's case selection process and the defendant's decision whether to go to trial or avoid trial, as shown in Table 5 from Chapter VII.

Table 5. Signs of expected direct and indirect effects on the observed trial conviction rate, $\tilde{\Pi}$

Case Characteristic	Expected Direct Effect on $\tilde{\Pi}$ *	Expected Indirect Case Selection Effect on $\tilde{\Pi}$	Expected Indirect Trial Selection Effect on $\tilde{\Pi}$	*Expected Total (Net) Effect on* $\tilde{\Pi}$
$\hat{\Pi}$ **	+	?	0	?
C	-	+	?	?
K	+	?	+	?
S^T **	+	?	?	?
e	+	?	?	?
λ	?	?	0	?
ρ	0	?	?	?

* The expected direct effects include first- and second-order effects, as explained in Appendix A.

** The marginal increase in $\hat{\Pi}$ and S^T are from exogenous shocks.

Table 17 reports the coefficients – not the marginal effects – estimated by probit regression models of the judge's or jury's decision to convict the defendant at trial (*TRIAL CONVICTION* = 1), or not to convict at trial (*TRIAL CONVICTION* = 0) through an acquittal, directed verdict, or dismissal. This probit model uses data regarding 128 federal criminal antitrust cases in which 601 defendants chose to go to trial.

The specification of the *ex post* probability of trial conviction tracks closely to the specification of the model of the *ex post* probability of trial, but the models differ in a few ways. The probit model of the probability of trial conviction controls for trial selection by including the explanatory variable *MILLS TRIAL*. This variable represents the inverse Mills ratio associated with the probability of trial, as estimated by the plea model reported in Table 15. That is, this variable represents the "Heckman's lambda" associated with the trial selection process that produces the sample of defendants included in the trial model specification.

For identification, the model of the probability of trial conviction includes two other explanatory variables not included in the probability of trial model and omits two variables that were included in the latter model. The other new variables are *VAR DOJ WIN PREV YEAR BY VIOL#1* and the squared term, (*VAR DOJ WIN PREV YEAR BY VIOL#1*)^2. These variables represent λ, the variance of the previous observed year's trial conviction rate in cases involving the same violation in the first charge of the indictment.

The variable *DOJ WIN PREV YEAR BY VIOL#1* and its square, (*DOJ WIN PREV YEAR BY VIOL#1*)^2, which were included in the model of the *ex post* probability of trial are omitted in the specification of the model of the *ex post* probability of trial conviction. These variables relate to the idiosyncratic element, ρ, of the prosecutor's *ex ante* probability of trial conviction. According to the theory, ρ, is not expected to directly affect the *ex post* probability of trial conviction.

Further, identification of the model of the observed probability of trial conviction is on firmer ground by excluding *DOJ WIN PREV YEAR BY VIOL#1* and (*DOJ WIN PREV YEAR BY VIOL#1*)^2. The information provided by those variables is not totally lost, however, since those variables influence the estimate of the inverse Mills ratio controlling for the probability of trial. The statistical significance of the estimated coefficient for (*DOJ WIN PREV YEAR BY VIOL#1*)^2 in the probit model of the probability of trial suggests successful identification of the selection model of the probability of trial conviction reported in Table 17.

Table 17. Probit model of the *ex post* probability of trial conviction

Dependent variable: *TRIAL CONVICTION*

Explanatory Variables	Sign of the Expected Direct Effect	Estimated Coefficient	Robust Standard Error†	z	P > |z|
Constant	?	.6889495	4.870969	0.14	0.888
INDICT YEAR	?	-.039796	.1283927	-0.31	0.757
LANDMARK	?	-.0941648	.2088911	-0.45	0.652
DOJBUDGET	+	-.0001638	.0003552	-0.46	0.645
%CRIMINAL	?	.3679875	1.604738	0.23	0.819
ANNUAL FORTUNE500	-	-.0182584	.026341	-0.69	0.488
REPUB PRES INDICT YEAR	?	.389279	.764732	0.51	0.611
#CODEFEND	?	.0184174	.0299802	0.61	0.539
#CORP CODEFEND	-	-.1015646	.0672765	-1.51	0.131
ONLY INDIV DEFEND	+	1.098368 *	.6260658	1.75	0.079
ANY LOW RANK	+	-.6627685 **	.2926731	-2.26	0.024
#OTHER INDICT	+	-.2327006	.1895925	-1.23	0.220
#PREV CONVICT	+	-.107974	.2188513	-0.49	0.622
PROFIT	-	.0000701	.0001354	0.52	0.605
COMPENSATION	-	-.0000153	.0000268	-0.57	0.569
DURATION	+	.0419429	.0283366	1.48	0.139
COMMERCE	+	2.59e-07	1.32e-06	0.20	0.845
#COUNTS INDICT	+	-.1576443	.3277478	-0.48	0.631
ANY BID RIG	+	-1.879359 **	.8865568	-2.12	0.034
BID RIG NOT PRICE FIX	+	3.376874 ***	1.072895	3.15	0.002
PRICE FIX NOT BID RIG	?	-.3311841	.5816041	-0.57	0.569
CONSUMER VIC	?	1.545526 ***	.3993938	3.87	0.000
VAR DOJ WIN PREV YEAR BY VIOL#1	?	-.0188339	.060144	-0.31	0.754
(VAR DOJ WIN PREV YEAR BY VIOL#1)^2	?	.0006211	.0019613	0.32	0.751
POST-APPA INDICT YEAR	+	1.082489	1.165289	0.93	0.353

† Standard error adjusted for 128 clusters of Commerce Clearinghouse-defined cases.

* 0.10 level of signficance.

** 0.05 level of significance.

*** 0.01 level of significance.

Table 17, continued

Dependent variable: *TRIAL CONVICTION*

Explanatory Variables	Sign of the Expected Direct Effect	Estimated Coefficient	Robust Standard Error[†]	z	P > |z|
REPUB JUDGE	?	.3612863	.2733146	1.32	0.186
PRIOR PROSECUTOR	?	.3080878	.2973616	1.04	0.300
TENURE	?	-.013631	.025368	-0.54	0.591
MILLS TRIAL	?	1.948936 *	1.144289	1.70	0.089
Regression Statistics					
N	601				
Wald χ^2 (26)	56.68				
Prob > χ^2	0.0011				
Log pseudolikelihood	-219.62286				
Pseudo R^2	0.3059				

[†] Standard error adjusted for 128 clusters of Commerce Clearinghouse-defined cases.

* 0.10 level of signficance.

** 0.05 level of significance.

*** 0.01 level of significance.

Table 17 shows that several explanatory variables representing theoretical model variables have statistically significant estimated regression coefficients. Table 18 provides the estimated marginal effects on the *ex post* probability of trial conviction from marginal increases in the values of explanatory variables that had statistically significant estimated coefficients in the probit regression model reported in Table 17. The estimates reported in Table 18 are conditional on prosecutors initiating the observed cases.

Table 18. Estimated marginal direct effects on the *ex post* probability of trial conviction, conditional on prosecutors initiating the observed cases

Dependent Variable (y): Pr(*TRIAL CONVICTION*)

Mean Value of Pr(*TRIAL CONVICTION*) = 0.12493958

Explanatory Variable	Sign of Expected Direct Effect	Estimated Marginal Effect (dy/dx)	Standard Error	Z	P>|z|	Mean Value of Explanatory Variable (x)
*ONLY INDIV DEFEND**	+	.3444573	.24135	1.43	0.154	.044925
*ANY LOW RANK**	+	-.1085118	.04095	-2.65	0.008	.206323
*ANY BID RIG**	+	-.2322762	.08083	-2.87	0.004	.231281
*BID RIG NOT PRICE FIX**	+	.9079712	.09602	9.46	0.000	.139767
*CONSUMER VIC**	?	.4120851	.12126	3.40	0.001	.31614
MILLS TRIAL	?	.4010598	.23651	1.70	0.090	1.23755

* Estimated marginal effect (dy/dx) is for discrete change of dummy variable from 0 to 1.

According to the estimates reported in Table 18, when the defendants in an observed case are all individuals (*ONLY INDIV DEFEND*), the probability of trial conviction is 0.344 percent higher than when not all the defendants are individuals, *ceteris paribus*. This is finding is consistent with the predicted positive direct effect on the *ex post* probability of trial from an increase in the value of *K*.

On the other hand, when any of the defendants in an observed case have low corporate rank (i.e., below the rank of president) and *ANY LOW RANK* = 1, *ceteris paribus*, the probability of trial conviction is 0.109 percent lower than when none of the defendants are low ranking. The expected direct effect on the *ex post* probability of trial conviction from an increase in *K* is positive. Since lower-ranking defendants' unit costs of trial expenditure are presumably higher than high-ranking defendants' unit costs of trial expenditure, this finding contradicts the theoretical prediction.[102]

[102] This finding might suggest that prosecutors are motivated to convict higher ranking defendants and put forth better efforts at trial when prosecuting higher ranking defendants.

When the observed defendant faces any bid-rigging charge (*ANY BID RIG*), the observed trial conviction rate is 0.232 percent lower than when the defendant does not face any such charge. Prosecutors are more likely to prosecute bid-rigging cases with more success than other types of cases, as suggested by Snyder (1989), because bid-rigging cases produce more or stronger evidence than other types of cases. This finding is inconsistent with direct effect expectations.

When the observed defendant is charged with big-rigging but not price-fixing (*BID RIG NOT PRICE FIX*), however, the observed trial conviction rate is 0.908 percent higher than it is when the observed defendant is not charged with bid-rigging, *ceteris paribus*. This finding is consistent with the expected direct effect of more or stronger evidence on the observed trial conviction rate.

In addition, when the victim of the antitrust crime is a consumer or group of consumers (*CONSUMER VIC*), *ceteris paribus*, the probability of trial conviction is 0.412 percent higher than when the victim is a firm, a competitor, or the government. For purposes of positive analysis of observed applications of law, this is an interesting empirical finding. Section 1 of the Sherman Act does not distinguish between the antitrust crimes committed against different types of victims. That suggests that the probability of trial conviction should not necessarily depend on the type of victim. Hypotheses regarding reasons for this finding or normative considerations regarding this finding are beyond the scope of this study.

As mentioned previously, the estimated regression coefficients and estimated marginal effects derived from the regression results do not identify indirect case selection effects from increases in the values of the explanatory variables that represent case characteristics. While the reported results are useful for positive analysis of federal criminal antitrust enforcement, those involving explanatory variables that represent case characteristics are not useful for testing all of the implications of the model of the *ex post* probability of trial conviction.

Notwithstanding that caveat, with every percentage increase in the estimated probability that the observed defendant would go to trial (*MILLS TRIAL*), the *ex post* probability of trial conviction is 0.475 percent higher, *ceteris paribus*. This finding has implications regarding the information structure of plea bargaining, as discussed later.

Here it is worth noting that the statistical significance of the estimated coefficient for *MILLS TRIAL* supports the notion that increases in the marginal values of case characteristics (that are also included in the probability of trial model) create indirect trial selection effects on the *ex post* probability of trial conviction. In fact, the trial selection effects shape the Heckman's lambda (or inverse Mills ratio), *MILLS TRIAL*. Separating the estimated magnitudes of individual trial selection effects of specific case characteristics requires additional calculations, however.

Along with the means test results presented earlier, this finding provides another reason why analysis of trial outcomes in federal criminal antitrust cases should not ignore potential indirect trial selection effects. In fact, the statistical significance of the estimated coefficient for *MILLS TRIAL* may help explain the relatively weaker support the first type of means tests provided for the trial selection model[103] compared to the support the first type of means tests provided for the trial conviction model.[104]

As previously mentioned, the relatively weak support for the model of the defendant's plea choice from the first type of means tests reported in Table 10 suggests that unobserved or unobservable factors not captured by this study's empirical variables could be primary determinants of defendants' plea choices. Such unobserved or unobservable factors could incluce the defendant's pre-trial assessment of the strength of his legal representation or the credibility of his best defense if he chooses to go to trial.

It is therefore important to note that Heckman's lambda associated with the probability of trial (*MILLS TRIAL*) controls for trial selection effects on the *ex post* probability of trial conviction created by observable factors (e.g. the probability of trial probit regressors) *as well as* unobservable factors that influence the error term of the probit model of the probability of trial. Thus, the combination of the weak support for the trial selection model provided by the first type of means tests reported in Table 10 and the statistical significance of the coefficient for *MILLS TRIAL* support the notion that unobserved or unobservable factors largely determine defendants' decisions to go to trial.

---

103 See Table 10.
104 See Table 10.

Estimates of indirect trial selection effects on the observed trial conviction rate

The estimated coefficient for *MILLS TRIAL* identifies the effect of Heckman's lambda, which is the expected value of the disturbance in the selection equation (here, going to trial) on the outcome (here, the trial outcome). It is possible to calculate, from several inputs, the signs and estimated magnitudes of trial selection effects, which form Heckman's lambda, on the observed trial conviction rate from marginal increases in the values of specific case characteristics. The inputs are the estimated coefficient on *MILLS TRIAL*, the coefficients in the probit model of the observed trial rate, and the relationship between the estimated promit model index and the variable *MILLS TRIAL*. All explanatory variables that appear in the probit model of the probability of trial have indirect effects on the trial outcome, including their effects on the observed probability of trial.[105]

One input in the calculations, $\partial(MILLS\ TRIAL)/\partial E(\tilde{\Theta})$, where $\tilde{\Theta}$ is the *ex post* probability of trial, requires explanation. Recall that *MILLS TRIAL* is Heckman's lambda, conceptually the expected value of the disturbance in the probit model of the probability of trial given that the observed defendant actually goes to trial. The mathematical definition of this variable is $MILLS\ TRIAL = \phi(E(\tilde{\Theta}))/1 - \Phi(E(\tilde{\Theta}))$, where ϕ is the standard normal probability distribution function (pdf) and Φ is the cumulative normal distribution function (cdf) evaluated at the mean of the data. Accordingly, *MILLS TRIAL* is a single index function – that is, a function of a linear index of explanatory variables – and $\partial(MILLS\ TRIAL)/\partial E(\tilde{\Theta})$ is a constant proportionality factor.

As previously mentioned, 601 of the 3093 observed indicted defendants, or about 19.4 percent of the indicted defendants, chose to go to trial. When a probit model of defendants' decision to go to trial is used, an observed trial rate of 19.4 percent implies $\Phi(E(\tilde{\Theta}))$ equals 0.806.[106] This, in turn, implies an average z-value in the data of 0.86124. That is, the average z-value (or the average index value) of 0.86124 is the value in the standard normal at which the

105 The variables need not have statistically significant coefficients in the probability of trial conviction model.
106 Selection is based on the high end of the index (i.e., *TRIAL* = 1), so the 19.4 percent is at the high end.

80.6 percentile is found. Thus, $\phi\left(E\left(\tilde{\Theta}\right)\right)$ equals 0.275. That is, 0.275 is the pdf value associated with a standard normal value (z) of 0.86124 and a cdf value of 0.806.[107]

These three numbers (z = 0.86124, $\phi\left(E\left(\tilde{\Theta}\right)\right)$ = 0.275, and $\Phi\left(E\left(\tilde{\Theta}\right)\right)$ = 0.806) are used to compute the numerical estimate of the factor of proportionality linking the probit index, also Known as a propensity score, to Heckman's lambda. The first step of the computation is calculating the derivative of Heckman's lambda with respect to the average z-value. This derivative, which is a function of $\Phi\left(E\left(\tilde{\Theta}\right)\right)$ and $\phi\left(E\left(\tilde{\Theta}\right)\right)$, is defined as

$$\frac{\partial\ Heckman's\ lambda}{\partial\ z} = \frac{z\phi}{1-\Phi} + \left(\frac{\phi}{1-\Phi}\right)^2 .$$

The second step is to insert the values z = 0.86124, $\phi\left(E\left(\tilde{\Theta}\right)\right)$ = 0.275, and $\Phi\left(E\left(\tilde{\Theta}\right)\right)$ = 0.806. The resulting numerical estimate of the factor of proportionality is 0.784.

In sum, the indirect effect on the probability of trial conviction from a marginal increase in the value of an explanatory variable in the probit model of the probability of trial is the following: the explanatory variable's estimated coefficient in the probit model of the probability of trial times the value of the derivative of Heckman's lambda with respect to the average z-value (here, 0.784) times the estimated coefficient on Heckman's lambda (here, the value of the estimated coefficient on *MILLS TRIAL*, 0,401). The last two parts of the calculation are the same for all explanatory variables in the probit model of the probability of trial. The first part of the equation, the estimated coefficient from the probit model of the probability of trial, differs for each explanatory variable.[108]

---

[107] Note the calculation is in reverse: the cdf is observed, the z-value is inferred, then the pdf is calculated. Here, the z-value identifies the number of standard deviations above the mean the 80.6 percentile is located.

[108] The question of statistical significance arises here because more than just the probit model of the probability of trial is involved in calculating the estimated indirect trial selection effects on the observed probability of trial conviction from marginal changes in the values of case characteristics. The calculation involves a non-linear transformation to Heckman's lambda, which is represented by *MILLS TRIAL*, and the estimated coefficient on *MILLS TRIAL* in the probit model of the observed probability of trial conviction. It follows that the estimated indirect trial selection effect of an explanatory variable can be statistically insignificant even if the estimated coefficient in the probit model of the probability of trial is statistically significant.

That concern reasonably applies only when the statistical significance of the relevant explanatory variables in the probit model of the probability of trial is marginal. That is because the values of Heckman's lambda (*MILLS TRIAL*) and thus the estimated coefficient on *MILLS TRIAL* in the probit model of the observed probability of trial conviction are based on more than 3,000 observations and apply equally to all explanatory variables in the probit

Table 19 shows the derivation of the estimated trial selection effects on the *ex post* probability of trial, $\widetilde{\Pi}$, from marginal increases in the values of several case characteristics: the defendant's unit cost of trial expenditure, K, as measured by *ONLY INDIV DEFEND*, *#OTHER INDICT*, *PROFIT*, and *COMPENSATION*; the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, as measured by *BID RIG NOT PRICE FIX*; and the parties' common *ex ante* expected trial sentence, S^T, as measured by *POST-APPA INDICT YEAR*.

Table 19. Estimated trial selection effects on the *ex post* probability of trial conviction, conditional on prosecutors initiating the observed cases

	(A)	(B)	(C)
Explanatory Variable in the Probit Model of the Observed Probability of Trial	Estimated Coefficients in the Probit Model of the Observed Probability of Trial*	Calculated Factor of Proportionality Linking the Index of the Probit Model of the Probability of Trial to Heckman's lambda associated with Probability of Trial (*MILLS TRIAL*) **	Estimated Coefficient for *MILLS TRIAL* in the Probit Model of the Observed Probability of Trial Conviction***
ONLY INDIV DEFEND	0.137314	0.784	0.4010598
#OTHER INDICT	-0.0277731	0.784	0.4010598
PROFIT	0.0000336	0.784	0.4010598
COMPENSATION	-6.80e-06	0.784	0.4010598
BID RIG NOT PRICE FIX	0.2260117	0.784	0.4010598
POST-APPA INDICT YEAR	0.2092021	0.784	0.4010598

* Estimated marginal (total direct) effects on the *ex post* probability of trial reported in Table 16.

** Value obtained by inserting z = 0.86124, $\phi\left(E\left(\widetilde{\Theta}\right)\right) = 0.275$, and $\Phi\left(E\left(\widetilde{\Theta}\right)\right) = 0.806$, which are implied by the fraction of indicted defendants that chose to go to trial (601/3093), into the formula for the derivative of Heckman's lambda with respect to the z-value of the data.

*** Value reported in Table 18.

model of the probability of trial. Of course, if the estimated coefficient on *MILLS TRIAL* in the probit model of the observed probability of trial conviction is statistically insignificant, then indirect trial selection effects is a moot issue.

Table 19, continued

	(D) = (A) x (B) x (C)	
Explanatory Variable in the Probit Model of the Ex Post Probability of Trial	Estimated Indirect Trial Selection Effect on the Ex Post Probability of Trial Conviction from a Marginal Increase in the Explanatory Variable in the Probit Model of the Ex Post Probability of Trial	Expected Sign of the Indirect Trial Selection Effect on the Ex Post Probabiity of Trial Conviction from a Marginal Increase in the Explanatory Variable in the Probit Model of the Ex Post Probability of Trial ****
ONLY INDIV DEFEND	0.043176	+
#OTHER INDICT	-0.00873	? for **e** , + for *K*
PROFIT	1.06E-05	-
COMPENSATION	-2.1E-06	-
BID RIG NOT PRICE FIX	0.071065	?
POST-APPA INDICT YEAR	0.06578	?

**** Predicted signs reported in Table 5.

Table 19 shows that the (trial) sample-selection probit model of the *ex post* probability of trial conviction provides mixed support for the predictions of the trial conviction model, nothwithstanding the ambiguity of the predicted signs for some indirect trial selection effects. These results, as well as the results of the other tests of the implications of the trial selection model nonetheless suggest that trial selection effects should not be ignored as determinants of observed federal criminal antitrust case outcomes.

Ordinary least squares regression model of observed trial sentences

As explained in Chapter VIII, the theoretical model of the *ex post* expected sentence following a trial conviction produces ambiguous predictions for the total marginal effects on the *ex post* expected trial sentence from marginal increases in the values of case characteristics. That is because of the predicted confounding influence of indirect selection effects produced by the prosecutor's case selection process, the defendant's decision whether to go to trial or avoid trial, and the judge's or jury's decision to convict the defendant at trial. These direct and indirect effects are shown in Table 7 taken from Chapter VIII.

Table 7. Signs of expected direct and indirect effects on the observed trial sentencing decision, $\tilde{S}^T$

Case Characteristic	Expected Direct Effect on $\tilde{S}^T$ *		Expected Indirect Case Selection Effect on $\tilde{S}^T$	Expected Indirect Trial Selection Effect on $\tilde{S}^T$	Expected Indirect Trial Conviction Selection Effect on $\tilde{S}^T$
	When $\hat{\Pi} \geq \overline{\Pi}$	When $\hat{\Pi} < \overline{\Pi}$			
$\hat{\Pi}$ **	0		?	0	?
C	-	?	+	+	+
K	+	?	?	+	?
S^T **	+		?	?	?
e	?		?	?	?
λ	?		?	0	?
ρ	0		?	?	0

* The expected direct effects include first- and second-order effects, as discussed in Appendix A.

** The marginal increase in $\hat{\Pi}$ and S^T are from exogenous shocks.

Table 20 provides the estimated coefficients, which are also the marginal effects, estimated by OLS regression, of the determinants of the fine imposed by judges on defendants convicted at trial. This regression uses data regarding 132 defendants convicted at trial in 43 cases initiated from 1956 through 1980.

The fine model specifications differ from the plea and trial model specifications in several ways. The fine model specification replaces *INDICT YEAR*, which controls for the year of the indictment, with *SENTENCE YEAR*, which controls for the year of the observed sentence. The fine model also includes new variables that represent characteristics of sentencing judges, such as whether the judge and the President belonged to the same political party (*PRES & JUDGE SAME PARTY SENTENCE YEAR*), the judge's tenure on the federal bench (*TENURE*), and the judge's workload in terms of the annual number of filings in his court (*#FILINGS*). Similarly, this model includes a variable that controls for the political party of the President during the sentencing year, *REPUB PRES SENTENCE YEAR*. These variables control for judicial discretion in sentences, which could be politically motivated (especially as it relates judges' motives to be promoted to a higher court).

The fine model includes the variable *MILLS TRIAL*, which represents the inverse Mills ratio associated with the probability of trial as estimated using the plea model specification reported in Table 15. This model also includes the variable *MILLS TRIAL CONVICT*, which represents the inverse Mills ratio associated with the probability of trial conviction. The values of this variable are estimates calculated with the trial conviction model specification reported in Table 17. These variables ensure proper identification and consistent estimates given the potential for sample selection bias created by the plea decision and the trial outcome.

The model of fines imposed after trial conviction also omits the variable *DOJ WIN PREV YEAR BY VIOL#1*, which represents the previous observed year's trial conviction rate in cases involving the same violation in the first charge of the indictment, *VAR DOJ WIN PREV YEAR BY VIOL#1*, which represents the variance of the previous observed year's trial conviction rate in cases involving the same violation in the first charge of the indictment, and their squared terms. Identification of the model of the fines imposed following trial conviction is on firmer ground by excluding these terms. The information provided by these omitted variables is not totally lost, however, since those variables influence the estimate of the inverse Mills ratios controlling for the probabilities of trial and trial conviction.

Because the estimated coefficients for neither *VAR DOJ WIN PREV YEAR BY VIOL#1* nor (*VAR DOJ WIN PREV YEAR BY VIOL#1*)^2 are statistically significant in the probit model of the observed trial conviction rate, the OLS selection model of fines imposed after trial conviction is not identified by omission of explanatory variables. Nonetheless, the Heckman's lambdas associated with the probabilities of trial and trial conviction identify the model non-linearly.

If the non-linear idenfication is weak because of collinearity of explanatory variables, then large standard errors and no statistically significant estimated effects would result. This is not the case here. The OLS selection model of fines imposed after trial conviction contains several statistically significant predictors, which confirms successful non-linear identification.

Table 20. Ordinary least squares model of the fine imposed after trial conviction

Dependent Variable = *FINE*

Explanatory Variables	Sign of the Expected First-Order Direct Effect	Estimated Coefficient	Robust Standard Error†	T	P > |t|
Constant	?	1267.441	1992.148	0.64	0.528
SENTENCE YEAR	?	-22.70772	54.15703	-0.42	0.677
LANDMARK	?	-131.7301 **	63.3963	-2.08	0.044
DOJBUDGET	?	-.12315	.1245212	-0.99	0.328
%CRIMINAL	?	-1139.185 **	436.8057	-2.61	0.013
ANNUAL FORTUNE500	?	-43.76564 **	16.20524	-2.70	0.010
REPUB PRES INDICT YEAR	?	417.3306*	227.5169	1.83	0.074
#CODEFEND	?	48.67149 **	19.88925	2.45	0.019
#CORP CODEFEND	?	-150.1411 ***	53.65624	-2.80	0.008
ONLY INDIV DEFEND	?	948.2954 ***	293.3048	3.23	0.002
ANY LOW RANK	?	-530.9433 **	214.9137	-2.47	0.018
#OTHER INDICT	+	-44.9275	100.285	-0.45	0.656
#PREV CONVICT	+	142.1528	143.0881	0.99	0.326
PROFIT	?	.0846083 **	.0254776	3.32	0.002
COMPENSATION	?	-.0083601 *	.0046261	-1.81	0.078
DURATION	+	20.04302	12.54574	1.60	0.118
COMMERCE	+	-.0002604	.0005453	-0.48	0.635

† Standard error adjusted for 43 clusters of Commerce Clearinghouse-defined cases.

* 0.10 level of signficance.

** 0.05 level of significance.

*** 0.01 level of significance.

Table 20, continued

Dependent Variable = *FINE*

Explanatory Variables	Sign of the Expected First-Order Direct Effect	Estimated Coefficient	Robust Standard Error†	T	P > |t|
#COUNTS INDICT	+	-207.6047	124.5319	-1.67	0.103
#COUNTS CONVICT	+	183.3675	150.3582	1.22	0.229
ANY BID RIG	?	-2324.146 ***	705.1223	-3.30	0.002
BID RIG NOT PRICE FIX	?	3150.505 ***	906.3422	3.48	0.001
PRICE FIX NOT BID RIG	?	-766.4474 ***	226.189	-3.39	0.002
CONSUMER VIC	?	1177.95 ***	404.0791	2.92	0.006
CASE TIME SPAN	?	51.12032	63.23025	0.81	0.423
POST-APPA SENTENCE YEAR	+	-146.532	215.7079	-0.68	0.501
REPUB JUDGE	?	272.7474 **	129.2187	2.11	0.041
REPUB PRES SENTENCE YEAR	?	-697.4412 ***	246.4497	-2.83	0.007
PRES & JUDGE SAME PARTY SENTENCE YEAR	?	-91.86725	91.80664	-1.00	0.323
PRIOR PROSECUTOR	?	151.2172	98.07022	1.54	0.131
TENURE	?	4.763831	7.008494	0.68	0.500
#FILINGS	?	.8391433 *	.4242685	1.98	0.055
MILLS TRIAL	?	1639.205 ***	461.33	3.55	0.001
MILLS TRIAL CONVICT	?	1119.554 ***	363.9936	3.08	0.004

† Standard error adjusted for 43 clusters of Commerce Clearinghouse-defined cases.

* 0.10 level of signficance.

** 0.05 level of significance.

*** 0.01 level of significance.

Table 20, continued

Dependent Variable = *FINE*

Regression Statistics		
N	132	
F(32, 42)	31.34	
Prob > F	0.0000	
R-squared	0.5351	
Root MSE	187.57	

As explained in Chapter VIII, the theoretical model of the *ex post* expected trial sentence, $\tilde{S}^T$, depends largely on the theoretical model of the *ex ante* expected trial sentence, S^T. The theoretical model of S^T, which is defined in Chapter V, is the least developed element of the theoretical framework of this study. That is largely because judicial discretion is an important determinant of sentencing outcomes, especially during the pre-sentencing guidelines period covered by the dataset of federal criminal antitrust cases that this study employs. The influence of case characteristics – other than S^T – on sentencing decisions made largely at the discretion of judges is extremely difficult to model theoretically. This study does not attempt to model theoretically judicial discretion in judges' sentencing decisions.

While this study does not attempt to model judicial discretion in sentencing, this study does treat the *ex ante* expected trial sentence, S^T, as an endogenous variable that depends on factors that also influence the litigants' chosen levels of trial expenditure and the *ex ante* probability of trial conviction, $\hat{\Pi}$. As summarized in Table 2 in Chapter V, the theoretical predictions regarding expected changes in S^T from changes in the values of case characteristics depend on the relative magnitudes of changes in the levels of litigants' trial expenditures from changes in the values of case characteristics. Those predictions also depend on whether the prosecutor or the defendant is the trial favorite and the degree that one litigant's chances at trial exceed the other litigant's chances. *A priori* assumptions regarding these conditions are not easily defendable.

Thus, even ignoring the role of judicial discretion in sentencing, predicting the total, direct, and indirect marginal effects on observed sentencing decisions from changes in the values of case characteristics is extremely challenging. Accordingly, as shown in Table 20, the theorectical predictions regarding the signs of the direct effects on the *ex post* expected trial sentence from marginal increases in the values of case characteristics – except those positively related to S^T – are ambiguous. In light of these challenges and ambiguities, interpretations of some but not all of the estimated coefficients in the regression model of sentencing decisions following trial convictions are beyond the scope of this study.

Notwithstanding the challenges in predicting and interpreting the regression results reported in Table 20, several regression results are worth noting. For example, the estimated coefficient for *POST-APPA SENTENCE YEAR* is statistically insignificant. This finding differs from Cohen's (1992) finding of a statistically significant positive influence of felony penalties (instead of misdemeanor penalties) on fines imposed on convicted corporations. Differences in regression model specifications and data[109] likely contribute to the differences in regression results. The statistical insignificance of the estimated coefficient for *POST-APPA SENTENCE YEAR* suggests the importance of judicial discretion in (pre-sentencing guidelines) sentencing decisions in federal criminal antitrust cases.

The estimated coefficients for judge-specific explanatory variables in this study's regression model of fines imposed after trial conviction further support the importance of judicial discretion. For example, Republican judges (*REPUB JUDGE* = 1) impose real fines after trial convictions that are more than $270,000 higher than Democrat judges, *ceteris paribus*. In addition, when the president in the year that the fine is imposed is a Republican (*REPUB PRES SENTENCE YEAR = 1*), judges impose real fines that are nearly $700,000 lower than fines imposed when the president during the sentencing year is a Democrat. Fines imposed after trial conviction do not change when the judge and president belong to the same political party, however, other factors held constant.

The estimated coefficient for the explanatory variable representing the annual number of cases filed in the judge's court (*#FILINGS*) suggests that, *ceteris paribus*, judges impose fines after trial convictions that rise by $1,000 (in real terms) for each filing in the judge's court. This

[109] Cohen's (1992) OLS regression model of imposed fines used different explanatory variables and employed data for corporate defendants convicted by pleas of guilty and *nolo contendere* as well as those convicted at trial.

supports Cohen's (1992) hypothesis that judges "penalize" convicted defendants for using scarce court resources.

Other regression results reported in Table 20 relate directly to a central focus of this study – the influence of trial selection effects and trial conviction selection effects on sentences imposed after trial conviction. The estimated coefficients for the explanatory variables *MILLS TRIAL* and *MILLS TRIAL CONVICT* are statistically significant and positive. These estimated coefficients suggest that, *ceteris paribus*, the higher the probabilities that the observed defendant would go to trial and would be convicted at trial, the larger the fine the defendant will receive upon conviction at trial. In terms of the indirect selection effects considered in this study, these estimated coefficients support the notion that the processes by which defendants choose to go to trial and judges or juries decide to convict defendants at trial contribute to higher fines imposed after trial convictions.

The estimated coefficients for *MILLS TRIAL* and *MILLS TRIAL CONVICT* do not identify the signs or magnitudes of indirect trial selection effects and trial conviction selection effects created by marginal increases in the values of case characteristics. Nonetheless, they are consistent with the notion that indirect trial selection effects and trial conviction selection effects from changes in the values of case characteristics are determinants of the observed fines imposed on the sample of 132 defendants included in the regression model presented in Table 20. Along with the means tests results presented earlier, these findings support the need to recognize potential indirect trial selection effects and trial conviction selection effects in analyses of the determinants of federal criminal antitrust case outcomes.

Regression-based tests of the information structure of plea bargaining

Privately informed prosecutors harm social welfare because they can fashion a settlement offer that induces (truly) innocent defendants to plead guilty or *nolo contendere*. Such pleas send false signals about legal and socially efficient business conduct. In contrast, privately informed defendants during plea bargaining benefit social welfare because the (truly) guilty ones are more likely to plead guilty and send the correct signals about appropriate business conduct.

Chapter IX presents several approaches to testing the information structure of plea bargaining that are alternatives to the approach taken by Froeb (1993) and others. The same intuition motivates the testing approaches used in this study and by Froeb. When the prosecutor is privately informed about the defendant's chances at trial, the prosecutor can fashion plea offers that are likely to induce even a truly innocent defendant to plead guilty (or *nolo contendere*) in lieu of going to trial. In contrast, when defendants possess private information about their chances at trial, truly guilty defendants are likely to accept plea offers but truly innocent defendants are likely to maintain their innocence and go to trial. From that common starting point, this study's empirical tests of the information structure of plea bargaining diverge from Froeb's testing approach.

Recall from Chapter IX that Froeb assumes away the potential for legal error at trial and assumes that trials perfectly identify innocent and guilty defendants. Under this assumption, if the prosecutor holds private information during plea bargaining and offers harsher penalties for a guilty plea than she previously offered, then truly innocent defendants (whom the prosecutor was previously offering lighter penalties in plea deals) are likely to be the defendants that reject the harsher plea offers and go to trial. In this situation, as the trial rate increases, the trial conviction rate is likely to fall.

In contrast, if defendants are privately informed and the prosecutor offers harsher penalties than she previously offered, then the truly guilty defendants are likely to go to trial (because truly innocent defendants were already going to trial). In this situation, as the trial rate increases, the trial conviction rate is likely to increase.

Froeb uses the correlation between the observed trial rate and trial conviction rate to test the information structure of plea bargaining. Under this approach, a positive correlation between trial and trial conviction rates suggests that in the sample of cases used, defendants were privately informed during plea bargaining. A negative correlation suggests that prosecutors held private information.

As discussed in Chapter IX, Froeb recognizes that the case selection process – and not necessarily the information structure of plea bargaining – could produce defendants who are more likely to go to trial and either more or less likely to be convicted at trial. To this potential confounding influence of the case selection process, this study points out the potential for a similarly confounding influence of the trial selection process whereby a mix of case characteristics that are likely to lead defendants to choose to go to trial could influence the probability of trial conviction.

This study's empirical tests of the information structure of plea bargaining do not rely on correlations between the trial rate and trial conviction rate for a given sample of defendants. Instead, this study uses five tests of the information structure of plea bargaining that stem from hypotheses developed by Hylton (1993) and introduces one test.

As discussed in Chapter IX, Hylton explains that the distance between the expected liability (expected penalty and trial cost) of an innocent defendant and the expected liability of a guilty defendant increases as the probabilities of (type I and type II) legal errors regarding the defendant's compliance with a legal standard fall. This *increases* the zone of acceptable plea offers (i.e., penalties that defendants would accept by pleading guilty) for guilty defendants and *decreases* the zone of acceptable plea offers for innocent defendants. Under these circumstances, it follows that the proportion of innocent defendants who litigate likely increases.

According to Hylton, the strongest evidence regarding the information structure of plea bargaining is the overall trial conviction rate for a given sample of defendants. As explained in Chapter IX, assuming the probabilities of type I and type II legal errors regarding defendants' compliance with legal standards are low[110] and roughly equal, if defendants have private information regarding their chances at trial, then the trial conviction rate will be below 50 percent. That is, when legal errors are relatively unlikely, a trial conviction rate below 50 percent is evidence that defendants are privately informed during plea bargaining about their chances at trial. Conversely, a trial conviction rate above 50 percent is evidence that prosecutors possess private information during plea bargaining. A trial conviction rate of 50 percent is evidence that neither party holds private information. According to Hylton, for a given sample of defendants, comparing the value of the trial conviction rate to the 50 percent mark is the best test of the information structure of plea bargaining.

It is reasonable to assume that the legal error at trial is sufficiently low[111] and that the probabilities of type I and type II errors are roughly equal in this study's sample of case initiated from 1956 through 1980. In the sample of defendants used in this study, out of the 3,093 indicted defendants, (only) 601 defendants chose to go to trial. Of these 601 defendants, judges and juries convicted only 132. Thus, for the sample of defendants used in this study, the trial conviction rate is 132/601*100 = 21.96 percent. A trial conviction rate of less than 22 percent

110 Hylton claims that a probability of legal error at trial below 50 percent is sufficiently low.

111 Notwithstanding the landmark Supreme Court decisions over the past 50 years, by 1956, the Sherman Act had been law for over 60 years.

suggests that, during plea negotiations, the observed defendants in federal criminal antitrust cases from 1956 through 1980 held private information about their chances at trial.

The other tests suggested by Hylton involve movements in the trial rate and trial conviction rate as legal doctrine develops over time. As legal doctrine develops over time, the likelihood of (type I and type II) legal error at trial falls. Hylton explains that the distance between the expected liability (expected penalty and trial cost) of an innocent defendant and the expected liability of a guilty defendant increases as the probability of legal error at trial falls. This *increases* the zone of acceptable plea offers (i.e., penalties that defendants would accept by pleading guilty) for guilty defendants and *decreases* the zone of acceptable plea offers for innocent defendants.

Accordingly, as the probability of legal error falls, then the trial conviction rate will fall over time if plea bargaining (i.e., trial selection) operates as expected when defendants are privately informed (or better informed than prosecutors) about their chances at trial during plea bargaining. Based on Salop and White's (1988) finding that the annual average plaintiff win rate in civil antitrust trials fell over time, Hylton expects defendants in civil antitrust cases to possess better information about their chances at trial than plaintiffs possess.

Hylton uses the passage of time as a proxy for the development of legal doctrine. In contrast, the explanatory variables used in this study's regression models of the *ex post* probabilities of trial and trial conviction include separate measures of the passage of time (*INDICT YEAR*) and the incremental development of legal doctrine (*LANDMARK*) in order to apply Hylton's tests.[112]

As reported previously in Table 17, the estimated coefficients for *INDICT YEAR* and *LANDMARK* in the regression model of the *ex post* probability of trial conviction are statistically insignificant (but negative). These findings suggest that neither litigant held private information during plea bargaining.

Hylton also hypothesizes that the trial rate will fall as legal doctrine develops over time (thus reducing type I and type II legal errors) when, during plea negotiations, defendants are

---

[112] One advantage of departing from Hylton's approach and using regression models is that the regression models control for several factors that could influence the observed trial rate and trial conviction rate. This is important because some factors, such as the DOJ Antitrust Division's budget, for example, could be correlated with time (although symptoms of mulicolinearity do not exist). This study's application of Hylton's hypotheses thus better isolate the correlation of the trial rate and trial conviction rate with the cumulative (*INDICT YEAR*) and incremental (*LANDMARK*) development of legal doctrine.

privately informed (or have better information) about their chances at trial. As reported previously in Table 15, in the regression model of the observed trial rate, the estimated coefficient for *LANDMARK* is negative and statistically significant. Like the test that used the overall trial conviction rate, this finding suggests that defendants were privately informed during plea bargaining.

Meanwhile, the estimated coefficient for *INDICT YEAR* in the regression model of the *ex post* probability of trial is statistically insignificant (but negative). The explanatory variable *INDICT YEAR* controls for the passage of time as well as the cumulative development of antitrust legal doctrine. The statistical insignificance of the marginal effect of the passage of time on the observed trial rate suggests that, during plea bargaining, neither defendants nor prosecutors held private or better information regarding the defendants' chances at trial.

The test of the information structure of plea bargaining introduced in this study uses the estimated coefficient of the explanatory variable *MILLS TRIAL* that is included in the regression model of the observed trial conviction rate. The explanatory variable *MILLS TRIAL* controls for the estimated probability that the observed defendant would go to trial.

As explained in Chapter IX, ignoring the potential for legal error at trial, if the prosecutor holds private information during plea bargaining and offers harsher penalties for a guilty plea than she previously offered, then truly innocent defendants (whom the prosecutor was previously offering lighter penalties in plea deals) are likely to be the defendants that reject the harsher plea offers and go to trial. In this situation, a defendant who is more likely to go to trial is less likely to be convicted at trial than a defendant who is less likely to go to trial.

In contrast, if defendants are privately informed and the prosecutor offers harsher penalties than she previously offered, then the truly guilty defendants are likely to go to trial (because truly innocent defendants were already going to trial). In this situation, a defendant who is more likely to go to trial is more likely to be convicted at trial than a defendant who is less likely to go to trial.

Accordingly, a positive sign of the estimated coefficient for *MILLS TRIAL* in a regression model of trial outcomes suggests that defendants that are more likely to go to trial are *more* likely to be convicted at trial, then such a regression result suggests that (truly) *guilty* defendants are more likely to go to trial and that (truly) *innocent* defendants are more likely to accept plea bargains. Such a regression result would suggest that prosecutors are privately informed during plea bargaining about the defendants' chances at trial and are able to fashion plea arrangements that induce (truly) innocent defendants to plead guilty or *nolo contendere*.

Conversely, if the sign of the estimated coefficient for *MILLS TRIAL* in a regression model of trial outcomes suggests that defendants that are more likely to go to trial are *less* likely to be convicted at trial, then such a regression result suggests that (truly) *innocent* defendants are more likely to go to trial and that (truly) *guilty* defendants are more likely to accept plea bargains. Such a regression result would suggest that defendants are privately informed during plea bargaining about the their chances at trial and prosecutors are unable to fashion plea arrangements that induce such (truly) innocent defendants to plead guilty or *nolo contendere*.

As reported in Table 17, the estimated coefficient for *MILLS TRIAL* in this study's regression model of trial outcomes is positive and statistically signficiant. This finding suggests that, during plea bargaining, the prosecutors possessed private or better information about the defendants' chances at trial.

In sum, this study conducts six tests of the information structure of plea bargaining in the sample of cases used in this study. Two test results suggest that the defendants held private or better information during plea bargaining, one result suggests that the prosecutors were privately or better informed, and three results suggest that neither party possessed private information during plea bargaining. These mixed results suggest that the information structure of plea bargaining in federal criminal antitrust cases remains an open question. In normative terms, these mixed results leave open the question of whether plea bargaining was socially efficient or not in federal criminal antitrust cases initiated from 1956 through 1980.

CHAPTER XI

CONCLUSION

This study explores two categories of determinants of federal criminal antitrust case outcomes: selection processes by which cases are disposed, and the information structure of plea bargaining. While this study's examination of these factors was thorough, several areas remain ripe for expanded future research.

This study models, as functions of case characteristics, i) the federal prosecutor's decision to initiate a criminal antitrust case, ii) the defendant's decision to go to trial or not, iii) the judge's or jury's decision to convict a defendant at trial, and iv) the judge's sentencing decision after a trial conviction. This study also explores how selection processes by which i) the federal prosecutor decides to initiate a case, ii) the defendant chooses a plea, and iii) the judge (or jury) decides to convict a defendant at trial influence the characteristics of criminal antitrust cases that proceed to subsequent stages of case disposition.

By influencing the mix of characteristics in cases that pass through these selection processes, the selection processes can influence the case outcomes that follow. That is, the case selection process can influence plea decisions, trial outcomes, and sentencing decisions. Similarly and separately, the trial selection process can affect trial outcomes and sentencing outcomes. Likewise, the trial conviction process can shape sentencing outcomes.

Snyder's (1989, 1990) study of plea bargaining and trial outcomes in federal criminal antitrust cases, as well as Cohen's (1992) study of sentencing outcomes, recognize the potential selection effects of the defendant's plea choice and the trial process on subsequent stages of case disposition. This study is the first to examine the potential influence of the prosecutor's selection of cases to initiate.

If these selection processes randomly sort defendants, then ignoring them is reasonable. Using defendant-level data for cases initiated from 1956 through 1980, this study's empirical tests indicate that prosecutors did not randomly decide which cases to initiate, defendants did not randomly decide whether or not to plead not guilty and go to trial, and judges and juries did not randomly convict defendants at trial. Further, this study's empirical tests suggest that the models developed in this study reasonably depict the processes by which prosecutors chose cases to pursue, defendants decided to plead not guilty instead of pleading guilty or nolo contendere, and judges and juries decided whether to convict or acquit (or otherwise not convict) defendants at trial.

This study also examines how the information structure of plea bargaining can influence observed federal criminal antitrust case outcomes. The information structure of plea bargaining has implications for normative and positive analysis of law enforcement.

When defendants hold private or better information about their chances at trial, plea bargaining is socially efficient. Guilty defendants plead guilty and save court resources. Meanwhile, innocent defendants maintain their innocence and (ignoring legal error) prevail at trial, thereby sending proper signals about the legality of the alleged conduct.

In contrast, when prosecutors possess better information about defendants' chances at trial, they are likely to use that informational advantage to fashion plea offers that truly innocent defendants accept. Such plea agreements are socially inefficient because they do not accurately signal the legality of alleged conduct. Moreover, for purposes of positive analysis, defendants with lower probabilities of trial conviction are less likely to go to trial when prosecutors hold private information than when defendants are privately informed.

This study conducts six tests of the information structure of plea bargaining in the sample of cases used in this study. Five of the tests stem from hypotheses posed by Hylton (1993). The results of two of those tests suggest that the defendants held private or better information during plea bargaining, but the results of the other three tests derived from Hylton's hypotheses imply that neither litigant possessed private information during plea bargaining.

This study contributes the sixth test, which employs a regression model of trial outcomes that controls for sample selection bias in order to determine the relationship, if any, between observed defendants' estimated probabilities of trial and trial conviction. This study finds that the estimated probability of trial conviction rises with observed defendants' estimated probability of going to trial. In contrast to the results of tests suggested by Hylton (1993), this finding suggests that the prosecutors held private or better information during plea bargaining.

Based on these mixed results, the informational structure of plea bargaining in federal criminal antitrust cases remains an open question. In normative terms, these mixed results leave open the question of whether plea bargaining was socially efficient or inefficient in federal criminal antitrust cases initiated from 1956 through 1980.

This study's focus on effects of selection processes and private information on federal criminal antitrust case outcomes, as well as the historical nature of the data used for empirical testing leave several topics for future research. In several ways, respectively, this study's theoretical and empirical analyses merit further consideration and expansion. In addition, developments in federal criminal antitrust enforcement that occurred after 1980 suggest future research.

In particular, extending Cohen's (1992) study to explore further the influence of judicial discretion in sentencing would be useful. Extensions of theoretical and empirical models of sentencing could and should examine the implications of changes in the latitude the legal system gives judges to use their discretion in sentencing. Such model extensions could exploit the natural experiments created by the implementation of mandatory and, later, discretionary guideline sentencing over the past two decades.

The details of the implementation and refinement of guideline sentencing raises important normative questions regarding federal criminal antitrust enforcement. For example, does (mandatory and/or discretionary) *charge-based* guideline sentencing give the prosecutor an informational advantage during plea bargaining? It seems plausible that the prosecutors could control the expected penalty imposed upon trial conviction by controlling the charges of indictment. This would give the prosecutor an informational advantage during plea bargaining.

Based on the regression results of this study, information about expected penalties may be the most relevant information during plea bargaining. In this regard, the statistically significant negative estimated regression coefficients for *LANDMARK* in the plea outcome and trial sentencing decision equations are noteworthy. Using the logic of Hylton's (1993) hypotheses, the signs of these coefficients suggest that the development of legal doctrine allows privately informed defendants to choose to go to trial when they expect judges to impose lower penalties upon trial conviction.

Another post-1980 development in federal criminal antitrust enforcement suggests a useful natural experiment for testing the implications of this study's theoretical models of the *ex ante* and *ex post* probabilities of trial conviction. Snyder (1989) and Marvel *et al*. (1988) document the tendency of prosecutors in the Reagan administration to bring criminal price-fixing cases against relatively small firms involved in alleged conspiracies involving small amounts of commerce. In terms of this study's theoretical models of the *ex ante* and *ex post* probabilities of trial conviction, such cases imply a relatively small value of the variable representing the *ex ante* expected trial sentence, S^T. Further, the high conviction rate in these cases implies a relatively high value of the variable representing the *ex ante* expected probability of trial conviction, $\hat{\Pi}$. Moreover, the lop-sided nature of these cases (in favor of the prosecutor) could be used to test the implications of the direct and indirect selection effects on the *ex post* probability of trial conviction from marginal changes in the value of λ, which represents the variability of the trial conviction decision.

APPENDIX A

PREDICTED EFFECTS ON THE PROSECUTOR'S *EX ANTE* EXPECTED VALUE OF BRINGING A CASE FROM CHANGES IN THE VALUES OF CASE CHARACTERISTICS

Recall that the prosecutor brings a case (i.e., a case is observed) if and only if $E(V) \geq 0$. Because the condition $E(V) \geq 0$ must be maintained in order for the prosecutor to initiate a case, it is important to evaluate the total marginal effects of changes in the values of case characteristics on $E(V)$. Now recall the definition of $E(V)$:

$$E(V) = \Theta V^T + (1-\Theta)V^B$$

$$= \Theta\left(\Phi S^T - CX\right) + (1-\Theta)\left(\Phi S^T + \frac{KY - CX}{2}\right) \quad , \qquad [14]$$

$$= \left(\hat{\Pi} + \rho\right)S^T + \frac{KY - CX}{2} - \Theta\frac{KY + CX}{2}$$

where

$$V^T = \Phi S^T - CX \ , \qquad [4]$$

$$V^B = S^B = \Phi S^T + \frac{KY - CX}{2} = \left(\hat{\Pi} + \rho\right)S^T + \frac{KY - CX}{2}, \qquad [13]$$

$$\Theta = \Psi\left(\rho - \frac{CX + KY}{S^T}\right), \qquad [11]$$

$$\hat{\Pi} = \Pi\left(\left[\mathbf{e} + P(X) - D(Y)\right]/\lambda\right).\text{and} \qquad [15]$$

$$S^T = m + \mu^T\left(\mathbf{e}, P(X), D(Y)\right)(M - m) \qquad [16]$$

In addition, recall that the litigation model provides the optimal (Nash equilibrium) values of the prosecutor's and the defendant's respective choices of levels of trial expenditure, so that all of the *ex ante* expected values are evaluated with $X = X^*$ and $Y = Y^*$. Both X^* and Y^* are implicit functions of C, K, S^T, $\mathbf{e}$, λ, as well as each other as specified by the litigation model.

In addition, recall that at the case selection stage (i.e., after the pre-indictment stage), the prosecutor takes as given an amount and/or quality of evidence, e. At the case selection stage, the prosecutor also takes as given the (exogenous) values of the other case characteristics (e.g., the prosecutor's unit cost of trial expenditure, C).

Intuitively, from the prosecutor's perspective at the case selection stage, the prosecutor's and defendant's expected behavior at trial, if the case should go to trial, is captured by their expected optimal choices of levels of trial expenditure, X^* and Y^*, respectively. At the case selection stage, the prosecutor expects that the litigants' (expected) optimal choices of trial expenditure will inform the prosecutor's *ex ante* expected value of going to trial, V^T, including the litigants' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, and their common *ex ante* expected sentence following a trial conviction, S^T. The prosecutor also uses X^* and Y^* to form her *ex ante* expected value of a plea agreement, V^B, her *ex ante* expected probability of trial, Θ. Putting it all together, the prosecutor uses X^* and Y^* to evaluate *ex ante* her expected value of bringing the case, $E(V)$.

I can now show mathematically the total (i.e. first-order and second-order) effects on $E(V)$ of marginal changes in individual case characteristics that the prosecutor is likely to consider.

$$\frac{dE(V)}{dC} = \frac{\partial E(V)}{\partial C}\frac{dC}{dC} + \frac{\partial E(V)}{\partial X}\frac{dX^*}{dC} + \frac{\partial E(V)}{\partial Y}\frac{dY^*}{dC} + \frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{dC} + \frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{dC} + \frac{\partial E(V)}{\partial S^T}\frac{dS^T}{dC}$$

$$= -\frac{(1+\Theta)X^*}{2} - \frac{[(1+\Theta)C]}{2}\frac{dX^*}{\partial C} + \frac{[(1-\Theta)K]}{2}\frac{dY^*}{dC} - \frac{(KY^* + CX^*)}{2}\frac{d\Theta}{dC} + S^T\frac{d\hat{\Pi}}{dC} + (\hat{\Pi} + \rho)\frac{dS^T}{dC} \quad [30]$$

$$\frac{dE(V)}{dK}=\frac{\partial E(V)}{\partial K}\frac{d\,K}{d\,K}+\frac{\partial E(V)}{\partial X}\frac{dX^*}{d\,K}+\frac{\partial E(V)}{\partial Y}\frac{dY^*}{d\,K}+\frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{d\,K}+\frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{d\,K}+\frac{\partial E(V)}{\partial S^T}\frac{dS^T}{d\,K}$$

$$=\frac{(1-\Theta)Y^*}{2}-\frac{[(1+\Theta)C]}{2}\frac{dX^*}{dK}+\frac{[(1-\Theta)K]}{2}\frac{dY^*}{dK}-\frac{(KY^*+CX^*)}{2}\frac{d\Theta}{dK}+S^T\frac{d\hat{\Pi}}{dK}+(\hat{\Pi}+\rho)\frac{dS^T}{d\,K} \qquad [31]$$

$$\frac{dE(V)}{dS^T}=\frac{\partial E(V)}{\partial S^T}\frac{dS^T}{d\,S^T}+\frac{\partial E(V)}{\partial X}\frac{dX^*}{d\,S^T}+\frac{\partial E(V)}{\partial Y}\frac{dY^*}{d\,S^T}+\frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{d\,S^T}+\frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{d\,S^T}$$

$$=(\hat{\Pi}+\rho)-\frac{[(1+\Theta)C^T]}{2}\frac{dX^*}{dS^T}+\frac{[(1-\Theta)K]}{2}\frac{dY^*}{dS^T}-\frac{(KY^*+CX^*)}{2}\frac{d\Theta}{dS^T}+S^T\frac{d\hat{\Pi}}{dS^T}. \qquad [32]$$

$$\frac{dE(V)}{d\mathbf{e}}=\frac{\partial E(V)}{\partial \mathbf{e}}\frac{d\,\mathbf{e}}{d\,\mathbf{e}}+\frac{\partial E(V)}{\partial X}\frac{dX^*}{d\,\mathbf{e}}+\frac{\partial E(V)}{\partial Y}\frac{dY^*}{d\,\mathbf{e}}+\frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{d\,\mathbf{e}}+\frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{d\,\mathbf{e}}+\frac{\partial E(V)}{\partial S^T}\frac{dS^T}{d\,\mathbf{e}}$$

$$=0-\frac{[(1+\Theta)C]}{2}\frac{dX^*}{d\mathbf{e}}+\frac{[(1-\Theta)K]}{2}\frac{dY^*}{d\mathbf{e}}+\frac{(KY^*+CX^*)}{2}\frac{d\Theta}{d\,\mathbf{e}}+S^T\frac{d\hat{\Pi}}{d\mathbf{e}}+(\hat{\Pi}+\rho)\frac{dS^T}{d\mathbf{e}}, \qquad [33]$$

$$\frac{dE(V)}{d\lambda}=\frac{\partial E(V)}{\partial \lambda}\frac{d\,\lambda}{d\,\lambda}+\frac{\partial E(V)}{\partial X}\frac{dX^*}{d\,\lambda}+\frac{\partial E(V)}{\partial Y}\frac{dY^*}{d\,\lambda}+\frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{d\,\lambda}+\frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{d\,\lambda}+\frac{\partial E(V)}{\partial S^T}\frac{dS^T}{d\lambda}$$

$$=0-\frac{[(1+\Theta)C]}{2}\frac{dX^*}{d\lambda}+\frac{[(1-\Theta)K]}{2}\frac{dY^*}{d\lambda}+\frac{(KY^*+CX^*)}{2}\frac{d\Theta}{d\,\lambda}+S^T\frac{d\hat{\Pi}}{d\lambda}+(\hat{\Pi}+\rho)\frac{dS^T}{d\lambda}, \qquad [34]$$

and

$$\frac{dE(V)}{d\rho} = \frac{\partial E(V)}{\partial \rho}\frac{d\rho}{d\rho} + \frac{\partial E(V)}{\partial X}\frac{dX^*}{d\rho} + \frac{\partial E(V)}{\partial Y}\frac{dY^*}{d\rho} + \frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{d\rho} + \frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{d\rho} + \frac{\partial E(V)}{\partial S^T}\frac{dS^T}{d\rho}$$
$$= S^T + 0 + 0 - \frac{(KY^* + CX^*)}{2}\frac{d\Theta}{d\rho} + 0 + 0 \quad . \quad [35]$$

Determining the signs of the right-hand sides of equations [30]-[35] requires some (admittedly tedious) discussion. This discussion is helpful, however, for understanding the determinants of expected case selection bias in observed case outcomes. A better understanding of this bias is important for understanding the implications of policy or institutional decisions regarding antitrust enforcement such as funding levels for the DOJ Antitrust Division, for example.

The first terms on the right-hand side of equations [30]-[32] and [35] are the direct partial (or "first-order") effects on $E(V)$ from increases in the values of the case characteristics C, K, S^T, and ρ, respectively. Inspection of Equation [14] confirms that the partial effects of increases in the values of K, S^T, and ρ on $E(V)$ are positive, while increases in C have a negative first-order effect on $E(V)$. Meanwhile, inspection of Equation [14] and the fact that $0 \le \Theta \le 1$ demonstrate that $\frac{\partial E(V)}{\partial X} < 0$, $\frac{\partial E(V)}{\partial Y} > 0$, $\frac{\partial E(V)}{\partial \Theta} < 0$, $\frac{\partial E(V)}{\partial \hat{\Pi}} > 0$, and $\frac{\partial E(V)}{\partial S^T} > 0$. Note that the prosecutor's evidence, e, and the variability of the trial court's conviction standard, λ, do not directly affect $E(V)$. Thus, the first term on the right-hand side of Equations [33] and [34] are zero.

With respect to the second and third terms on the right-hand sides of equations [30]-[34], I use predictions about litigation behavior from my litigation model to determine the signs. For example, according to the litigation model, an increase in a single party's cost per unit of trial expenditure, C or K, will lead that party to reduce trial expenditure. Further, the opponent will reduce trial expenditure if (and only if) the opponent is the trial "favorite."

Thus, when the prosecutor is the trial favorite (i.e., $\hat{\Pi} \ge \overline{\Pi}$), $\frac{dX^*}{\partial C} < 0$ and $\frac{dY^*}{dC} > 0$, while $\frac{dX^*}{dK} < 0$ and $\frac{dY^*}{dK} < 0$. In contrast, when the defendant is the trial favorite (i.e., $\hat{\Pi} < \overline{\Pi}$),

$\frac{dX^*}{dC} < 0$, and $\frac{dY^*}{dC} < 0$, while $\frac{dX^*}{dK} > 0$ and $\frac{dY^*}{dK} < 0$. This implies that the second term of Equation [30] is always positive, and the third term is positive when the prosecutor is the trial favorite and negative when she is the underdog at trial.

Similarly, with respect to right-hand side of Equation [31], the second term is positive when the prosecutor is the trial favorite and negative when she is the underdog at trial. The third term on the right-hand side of Equation [31], meanwhile, is always negative (just as the second term on the right-hand side of Equation [30] is always positive).

The litigation model predicts that an increase in S^T will lead both parties to increase trial expenditures. Thus, $\frac{dX^*}{dS^T} > 0$ and $\frac{dY^*}{dS^T} > 0$, which implies that the second and third terms in Equation [32] are negative and positive, respectively.

The litigation model predicts that if the prosecutor is the trial underdog (favorite), then an increase in the prosecutor's evidence, e, which increases the merits of the case, will increase (decrease) both parties' trial expenditures. Thus, if the prosecutor is the trial underdog (favorite), $\frac{dX^*}{d\mathbf{e}} > 0$ and $\frac{dY^*}{d\mathbf{e}} > 0$ ($\frac{dX^*}{d\mathbf{e}} < 0$ and $\frac{dY^*}{d\mathbf{e}} < 0$), and the second and third terms on the right-hand side of Equation [33] are positive (negative).

The litigation model's predictions regarding the affect on the parties trial expenditures from increases in the variability of the trial court's conviction standard, λ, depend on assumptions regarding the "closeness" of the parties' relative diminishing returns in trial expenditure. Such assumptions are not easily defended, so the signs of the second and third terms on the right-hand side of Equation [34] are ambiguous.

Meanwhile, the litigation model's specification that does not include the idiosycratic component, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, $\hat{\Pi}$, as an implicit determinant of the litigants' optimal levels of trial expenditure, X^* and Y^*. Thus, the second and third terms on the right-hand side of Equation [35] are zero.

Now, consider the signs of the fourth term on the right-hand side of equation [30]. Recall that Equation [11] defines the *ex ante* probability of trial from the prosecutor's perspective, and forms the basis of my trial selection model. An increase in the prosecutor's unit cost of trial expenditure, C, has the following total marginal effect on the prosecutor's *ex ante* probability of trial, Θ, given the implicit functions represented by the Nash equilibrium values of the parties' levels of trial expenditure, $X = X^*$ and $Y = Y^*$:

$$\frac{d\Theta}{dC} = \frac{\partial\Theta}{\partial C}\frac{dC}{dC} + \left(\frac{\partial\Theta}{\partial X} + \frac{\partial\Theta}{\partial S^T}\frac{\partial S^T}{\partial P}\frac{\partial P}{\partial X}\right)\frac{dX^*}{dC} + \left(\frac{\partial\Theta}{\partial Y} + \frac{\partial\Theta}{\partial S^T}\frac{\partial S^T}{\partial D}\frac{\partial D}{\partial Y}\right)\frac{dY^*}{dC}.$$

The sign of $\frac{d\Theta}{dC}$ depends on the relative signs and magnitudes of the unambiguously negative marginal direct effect of C on Θ (i.e., $\frac{\partial\Theta}{\partial C} = \psi\left(\rho - \frac{CX^* - KY^*}{S^T}\right)\left[-\left(\frac{X^*}{S^T}\right)\right] < 0$), the marginal "trial expenditure effects" of C on Θ (i.e., $\frac{\partial\Theta}{\partial X}\frac{dX^*}{dC} + \frac{\partial\Theta}{\partial Y}\frac{dY^*}{dC}$), and the marginal "stakes effects" of C on Θ (i.e., $\frac{\partial\Theta}{\partial S^T}\frac{\partial S^T}{\partial P}\frac{\partial P}{\partial X}\frac{dX^*}{dC} + \frac{\partial\Theta}{\partial S^T}\frac{\partial S^T}{\partial D}\frac{\partial D}{\partial Y}\frac{dY^*}{dC}$). As noted above, the litigation model's predictions imply that $\frac{dX^*}{dC}$ is always negative and that $\frac{dY^*}{dC}$ is negative if the prosecutor is the trial underdog. Meanwhile, the sign of the partial marginal direct effect on the *ex ante* probability of trial from an increase in the prosecutor's level of trial expenditure,

$$\frac{\partial\Theta}{\partial X} = \psi\left(\rho - \frac{CX^* + KY^*}{S^T}\right)\left[-\left(\frac{C}{S^T}\right)\right],$$

is negative. The sign of the marginal partial direct effect on the *ex ante* probability of trial from an increase in the defendant's level of trial expenditure,

$$\frac{\partial\Theta}{\partial Y} = \psi\left(\rho - \frac{CX^* + KY^*}{S^T}\right)\left[-\left(\frac{K}{S^T}\right)\right]$$

is also negative. In contrast, the sign of the marginal partial direct effect on the *ex ante* expected probability of trial from an increase in the *ex ante* expected trial sentence,

$$\frac{\partial \Theta}{\partial S^T} = \psi\left(\rho - \frac{CX^* + KY^*}{S^T}\right)\left[\frac{CX^* + KY^*}{\left(S^T\right)^2}\right],$$

is positive.

In sum, the sign of the fourth term on the right-hand side of Equation [30] is ambiguous because the sign of $\frac{d\Theta}{dC}$ is ambiguous without any defensible simplifications regarding the relative weights of the partial direct effect, the partial "trial expenditure effects," and the partial "stakes effects." Similarly, the fourth term on the right-hand side of Equation [31] is ambiguous because the sign of $\frac{d\Theta}{dK}$ is ambiguous.

The fourth term in equation [32] requires some discussion. This term is the marginal effect on the prosecutor's *ex ante* expected value of bringing a case, *E*(*V*), from a change in the *ex ante* probabiilty of trial, Θ, that is created by an increase in the value of the parties' common *ex ante* expected trial sentence, S^T (where the change in S^T is caused by an exogenous shock). The sign of the fourth term on the right-hand side of equation [32] depends on the sign of $\frac{d\Theta}{dS^T}$, which can be expressed as follows, when evaluated at equilibrium levels of trial expenditure, $X = X^*$ and $Y = Y^*$:

$$\frac{d\Theta}{dS^T} = \frac{\partial \Theta}{\partial S^T}\frac{dS^T}{dS^T} + \frac{\partial \Theta}{\partial X}\frac{dX^*}{dS^T} + \frac{\partial \Theta}{\partial Y}\frac{dY^*}{dS^T}.$$

By inspection, the sign of $\frac{d\Theta}{dS^T}$ depends on the relative magnitudes of the positive partial marginal effect on the *ex ante* expected probability of trial from an increase in the *ex ante* expected trial sentence, (i.e., $\frac{\partial \Theta}{\partial S^T} = \psi\left(\rho - \frac{CX^* + KY^*}{S^T}\right)\left[\frac{CX^* + KY^*}{\left(S^T\right)^2}\right] > 0$, as previously discussed) and the negative trial expenditure effects (i.e., $\frac{\partial\Theta}{\partial X}\frac{dX^*}{dS^T}+\frac{\partial\Theta}{\partial Y}\frac{dY^*}{dS^T}<0$).[113] Without more information regarding these relative magnitudes, the sign of the fourth term on the right-hand side of Equation [32] is ambiguous.

Because the *ex ante* expected probability of trial, Θ, does not directly depend on the prosecutor's evidence, e, the total marginal effect of e on Θ comes entirely from the partial trial expenditure effects and (the two types of) stakes effects on Θ from increases in e. These partial effects can be expressed mathematically as follows, when evaluated at equilibrium levels of trial expenditure, $X = X^*$ and $Y = Y^*$:

$$\frac{d\Theta}{d\mathbf{e}}=\left(\frac{\partial\Theta}{\partial X}+\frac{\partial\Theta}{\partial S^T}\frac{\partial S^T}{\partial P}\frac{\partial P}{\partial X}\right)\frac{dX^*}{d\mathbf{e}}+\left(\frac{\partial\Theta}{\partial Y}+\frac{\partial\Theta}{\partial S^T}\frac{\partial S^T}{\partial D}\frac{\partial D}{\partial Y}\right)\frac{dY^*}{d\mathbf{e}}+\frac{\partial\Theta}{\partial S^T}\frac{dS^T}{d\mathbf{e}}$$

As previously discussed, the litigation model predicts that if the prosecutor is the trial underdog (favorite), then $\frac{dX^*}{d\mathbf{e}}>0$ and $\frac{dY^*}{d\mathbf{e}}>0$ ($\frac{dX^*}{d\mathbf{e}}<0$ and $\frac{dY^*}{d\mathbf{e}}<0$). Nonetheless, the ambiguous relative magnitudes of the trial expenditure effects and the two types of partial stakes effects (i.e., the stakes effects from increases in S^T due to increases in *X* and *Y*, and separately from increases in S^T due to increases in e), as well as the ambiguous sign of the net stakes effect renders the sign of $\frac{d\Theta}{d\mathbf{e}}$ ambiguous. Accordingly, the sign of the fourth term on the right-hand side of Equation [33] is ambiguous.

Similarly, the variability of the trial court's conviction standard, λ, only influences the *ex ante* expected probability of trial, Θ, through the trial expenditure effects and stakes effects. This is evident by inspection of the following evaluation of $\frac{d\Theta}{d\lambda}$ at equilibrium levels of trial expenditure, $X = X^*$ and $Y = Y^*$:

---

113 Recall the litigation model predicts $\frac{dX^*}{dS^T}>0$ and $\frac{dY^*}{dS^T}>0$. As previously discussed, $\frac{\partial\Theta}{\partial X}<0$ and $\frac{\partial\Theta}{\partial Y}<0$.

$$\frac{d\Theta}{d\lambda} = \left(\frac{\partial\Theta}{\partial X} + \frac{\partial\Theta}{\partial S^T} \frac{\partial S^T}{\partial P} \frac{\partial P}{\partial X} \right) \frac{dX^*}{d\lambda} + \left(\frac{\partial\Theta}{\partial Y} + \frac{\partial\Theta}{\partial S^T} \frac{\partial S^T}{\partial D} \frac{\partial D}{\partial Y} \right) \frac{dY^*}{d\lambda} .$$

Because of the previously discussed ambiguity of the relative magnitudes of sign of $\frac{\partial\Theta}{\partial X} < 0$ and $\frac{\partial\Theta}{\partial S^T} \frac{\partial S^T}{\partial P} \frac{\partial P}{\partial X} > 0$ and because the litigation model's predictions regarding $\frac{dX^*}{d\lambda}$ and $\frac{dY^*}{d\lambda}$ are weak, the sign of the fourth term on the right-hand side of Equation [34] is ambiguous.

Due to the specification of the litigation model (i.e., X^* and Y^* do not depend on ρ), ρ does not create any confounding "trial expenditure effects" on Θ. Thus, it is only necessary to recognize that Θ increases in ρ and $E(V)$ decreases in Θ. Thus, the fourth term on the right-hand side of Equation [35] is negative.

With respect to the fifth term on the right-hand sides of equations [30]-[35], I will use the litigation model to predict the second-order effects of increases in the values of case characteristics on $E(V)$ through $\hat{\Pi}$. For instance, according to the litigation model, an increase in a single party's cost per unit of trial expenditure, C or K, will decrease that party's probability of winning at trial. Thus, with respect to Equation [30], the litigation model predicts that

$$\frac{d\hat{\Pi}}{dC} = \frac{\partial\hat{\Pi}}{\partial P} \frac{\partial P}{\partial X} \frac{dX^*}{dC} + \frac{\partial\hat{\Pi}}{\partial D} \frac{\partial D}{\partial Y} \frac{dY^*}{dC} < 0 .$$ [114]

Accordingly, the fifth term of equation [30] is negative. In Equation [31], however, the litigation model predicts that $\frac{d\hat{\Pi}}{dK} > 0$, which implies that the fifth term is positive.

With respect to the fifth term of Equation [32], my litigation model predicts that an increase in the trial stakes for both parties through an increase in S^T, for example, will induce

---

114 While it is possible for $\frac{dY^*}{dC}$ to be negative, it would be irrational for the defendant to decrease his trial expenditure to the point of letting the *ex ante* probability of trial conviction increase.

both parties to increase their trial expenditure. This will improve the chances at trial of the party with relatively weaker diminishing returns in trial expenditure.[115] Assumptions regarding the parties' relative rate of diminishing returns in their own trial expenditures require some discussion.

It is reasonable to assume that the prosecutor is the party with relatively weaker diminishing returns in trial expenditure (i.e., $D_{YY} / D_Y^2 < P_{XX} / P_X^2 < 0$). In my litigation model and, arguably, in practice, the defendant must react to the case, $A(e)$, that the prosecutor develops in the pre-indictment stage. Other factors held constant, the defendant has less "control" over the case than the prosecutor. Along those lines, it is well established in the game theoretical industrial organization literature that by having more strategic "tools" than the defendant, the prosecutor should be more productive at trial than the defendant. For example, in the litigation model, the prosecutor has evidence, e, and trial expenditure, X, as her tools at trial, while defendant only has his trial expenditure, Y.

In addition, as Cohen (1992) argues to underpin his principal hypothesis regarding the promotion potential of judges as a primary motivator of their sentencing decisions, the prosecutor generally has a more advantageous relationship with the typical federal district judge than the defendant has for purposes of winning at trial. Further, defendants in criminal antitrust trials could be confined in how far they can develop arguments if doing so could (inadvertently) work against other arguments or otherwise increase their liability.

For completeness of discussion, however, the assumption that the prosecutor has stronger diminishing returns from trial expendture (i.e., $P_{XX} / P_X^2 < D_{YY} / D_Y^2 < 0$) has some appeal. Defense attorneys often specialize in certain types of cases. In fact, anecdotally, antitrust defense and litigation support have blossomed as cottage industries. This suggests high levels of trial productivity for federal criminal antitrust defendants' litigation teams. The criminal

---

115 Meanwhile, as discussed in Chapter V, an increase in the trial costs of both parties, or a decrease in common trial stakes, will lead them to reduce their trial expenditure. Again, this will improve the chances at trial and at sentencing for the party with relatively weaker diminishing returns in trial expenditure.

The intuition behind these predictions is that the party with weaker diminishing returns in trial efforts will react to the same increase in trial stakes with a larger increase in trial arguments because that party finds it "cheaper" to increase trial expenditure, and *vice versa*. That is, the party with weaker diminishing returns gets more "bang per unit" from increasing trial effort and expenditure, so that party increases trial expenditure more than the opponent does. Using the same logic in the opposite direction, the party with stronger diminishing returns in trial efforts will react to a common decrease in trial stakes with a smaller decrease in trial arguments.

antitrust defense industry has not always been as well developed as it may be today, however, and my empirical tests (will) cover years going back to 1955.

Further, anecdotally, prosecutors in the U.S. Department of Justice Antitrust Division are possibly more singly focused on prosecuting antitrust offenses than private defense attorneys who often have several types of clients at any given time. To the extent this is an empirical question, the DOJ's historically high conviction rate, particularly in price-fixing and bid-rigging cases, suggests that federal prosecutors in criminal antitrust cases are, in aggregate or on average, more productive than criminal antitrust defense teams.

Thus, for the purposes of this study, it is reasonably safe to assume that the prosecutor has weaker diminishing returns in trial expenditure than the defendant has. When the prosecutor has weaker diminishing returns from trial expenditure, the litigation model predicts that the parties' common *ex ante* expected probability of trial conviction increases in the parties' common trial stakes. That is, when evaluated at equilibrium levels of trial expenditure, $X = X^*$ and $Y = Y^*$,

$$\frac{d\hat{\Pi}}{dS^T} = \frac{\partial\hat{\Pi}}{\partial P}\frac{\partial P}{\partial X}\frac{dX^*}{dS^T} + \frac{\partial\hat{\Pi}}{\partial D}\frac{\partial D}{\partial Y}\frac{dY^*}{dS^T}$$

is positive.[116] Thus, when the prosecutor (defendant) has weaker diminishing returns from trial expenditure, the fifth term in Equation [32] is positive (negative).

The litigation model predicts that the party that benefits from a change in the merits of the prosecutor's case will have better chances at trial. Because the merits of a case increase in the prosecutor's evidence, e, the fifth term on the right-hand side of Equation [33] is positive.

Meanwhile, the litigation model's predictions regarding the affect on the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, from increases in the variability of the trial court's conviction standard, λ, depend on assumptions regarding the "closeness" of the parties' diminishing returns in trial expenditure. As with the second and third terms of Equation [34], because such assumptions are not easily defended, the sign of the fifth term on the right-

116 According to my litigation model, when the prosecutor has weaker diminishing returns in trial expenditure, $\frac{d\hat{\Pi}}{dS^T} > 0$ for *decreases* in S^T as well.

hand side of Equation [34] is ambiguous. Moreover, the fifth term on the right-hand side of Equation [35] is zero because neither the parties' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, nor the parties' levels of trial expenditure, *X* and *Y*, depend on the idiosyncratic component, ρ, of her (personal) *ex ante* expected probability of trial conviction, Φ.

The litigation model predicts that if the prosecutor (defendant) is the trial favorite, then an increase in the prosecutor's unit cost of trial expenditure, C, (the defendant's unit cost of trial expenditure, *K*) will cause the prosecutor (defendant) to decrease her (his) level of trial expenditure, *X* (*Y*), and the defendant (prosecutor) to increase his (her) level of trial expenditure, *Y* (*X*), which will reduce (increase) S^T. Meanwhile, if the prosecutor (defendant) is the trial favorite, then an increase in K (*C*) will cause *X* and *Y* to fall, which has ambiguous net effects on S^T. Thus, the sign of the sixth term on the right hand side of Equation [30] is negative (amibiguous) when the prosecutor (defendant) is the trial favorite. Further, the sign of the sixth term on the right-hand side of Equation [31] is positive (ambiguous) when the prosecutor (defendant) is the trial favorite. (There is no sixth term on the right hand side of equation [32].)

The sixth term of the right-hand side of Equation [33] is the marginal effect of more evidence on the prosecutor's *ex ante* expected value of bringing a case through the effect of more evidence on the value of the *ex ante* expected trial sentence. The litigation model predicts that changes in the merits of the case through increases in the amount or strength of evidence, e, that benefit the underdog will cause both parties to increase trial expenditure. Conversely, changes in e that benefit the favorite will cause both parties to reduce trial expenditure. Thus, the unidirectional changes in the parties' trial expenditure from changes in the value of e have ambiguous effects on the *ex ante* expected trial sentence, S^T. Accordingly, the sign of the sixth term on the right-hand side of Equation [33] is ambiguous.

The litigation model's predictions about the effects of the variability of the trial court's conviction standard, λ, on the parties' trial expenditures, *X* and *Y*, rely on the "closeness" of the parties' diminishing returns in trial expenditure. Without any reasonable means of accurately measuring this "closeness," the sixth term on the right-hand side of Equation [34] is ambiguous.

Neither the *ex ante* expected trial sentence, S^T, nor the parties' trial expenditures, *X* and *Y*, are functions of the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probabilty of trial conviction, $\hat{\Pi}$. Thus, the value of the sixth term on the right-hand side of Equation [35] is zero.

I now turn to the expected sign of the net effect of changes in the values of case characteristics on the prosecutor's *ex ante* expected value of bringing a case, *E*(*V*). Intuitively appealing and otherwise defendable assumptions about the prosecutor's relative preferences for different aspects of case disposition clarify the predicted signs of the marginal effects of changes in the values of case characteristics on *E*(*V*).

At the margin, it is reasonable to suppose that a variety of possible instiutional or individual preferences regarding case disposition affect case-related decisions of prosecutors (or their supervisors) and judges. These possible case-disposition preferences imply different "analytical weights" that should be applied to different types of partial case selection effects on *E*(*V*) created by changes (or just increases) in the values of case characteristics.

For example, the prosecutor may be adverse to trial effort or expenditure, as it represents difficult work. Attorneys often complain about the time consumption or stressful conditions of work related to trial litigation. I label such prosecutors "own workload-averse." In terms of my case selection model, assuming the prosecutor is "workload-oriented" implies that more predictive weight should be placed on $\frac{\partial E(V)}{\partial X} < 0$ relative to $\frac{\partial E(V)}{\partial Y} > 0$, $\frac{\partial E(V)}{\partial \Theta} < 0$, $\frac{\partial E(V)}{\partial \hat{\Pi}} > 0$, and $\frac{\partial E(V)}{\partial S^T} > 0$.

Similarly, it is possible that in her decision to bring a case the prosecutor could give primary consideration to the amount of work a case would create for the defendant or the defendant's attorney. I call that type of prosecutor "defendant's workload-motivated", which implies that relatively more predictive weight should be placed on $\frac{\partial E(V)}{\partial Y} > 0$ relative to $\frac{\partial E(V)}{\partial X} > 0$, $\frac{\partial E(V)}{\partial \Theta} < 0$, $\frac{\partial E(V)}{\partial \hat{\Pi}} > 0$, and $\frac{\partial E(V)}{\partial S^T} > 0$.[117]

---

[117] There may be practical, if not normative, justifications for ignoring the case selection effects of *Y* (i.e., letting $\frac{\partial E(V)}{\partial Y} = 0$). By not making that assumption, my case selection model allows the prosecutor to base her case-bringing decision primarily or solely based on how much effort, *Y*, a case will create for the defendant or the defendant's attorney(s). For the defendant's attorney(s) or their law firm, more expected trial effort could translate into higher salaries or profits. In the extreme, questions of professional competence or ethics might be raised if a

For some time, long and growing backlogs of court dockets have been a problem in federal courts. This implies that judges, as well as government prosecutors who repeatedly appear before them, would prefer federal criminal antitrust cases, which are notoriously associated with long and complex trials, to be disposed by plea agreement (while ensuring, in the judges' minds at least, that justice is served). Under these circumstances, it would be reasonable to expect the prosecutor to be "trial-averse," which implies that $\frac{\partial E(V)}{\partial \Theta} < 0$ should be given more weight in my case selection model's predictions relative to $\frac{\partial E(V)}{\partial X} < 0$, $\frac{\partial E(V)}{\partial Y} > 0$, $\frac{\partial E(V)}{\partial \hat{\Pi}} > 0$, and $\frac{\partial E(V)}{\partial S^T} > 0$.

It is also possible that the prosecutor is "penalty-motivated" and, in deciding which cases to pursue, derives great value from the parties' common *ex ante* expected trial sentence, possibly for deterrence reasons. A penalty-motivated prosecutor implies that $\frac{\partial E(V)}{\partial S^T} > 0$ should be given more weight relative to $\frac{\partial E(V)}{\partial X} < 0$, $\frac{\partial E(V)}{\partial Y} > 0$, $\frac{\partial E(V)}{\partial \Theta} < 0$, and $\frac{\partial E(V)}{\partial \hat{\Pi}} > 0$.

Meanwhile, studies by Posner (1970), Siegfried (1975), Weaver (1977) and Baker (1978) strongly suggest that government prosecutors are primarily motivated by (total) conviction rates, regardless of how convictions are achieved (i.e., either through trial or plea agreement). In particular, Posner (1970) is sharply critical of the DOJ Antitrust Division. According to Posner,

> [I]t would appear that both legal doctrine and the enforcement machinery are geared more to the apprehension of unsuccessful attempts to fix prices than to the apprehension of successful price fixing. In general, the fact of an agreement or conspiracy to fix prices is all that the government need prove in a price-fixing case and all that it attempts to prove. (p. 410)

Siegfried (1975) reasons that expected net social benefits from federal antitrust enforcement should determine enforcement activity if the DOJ Antitrust Division's objective is to maximize something akin to social welfare. Like the study by Long *et al*. (1973) upon which Siegfried's empirical analysis builds using less aggregated data, however, Siegfried's empirical

---

prosecutor selected cases based (only) on considerations of how much trial effort the case is likely to require of the defendant or the defendant's attorney(s). Further, it is reasonable to assume that K and $\mathbf{e}$ capture the prosecutor's expectations regarding how costly trial will be for the defendant.

tests fail to find much evidence that measures of the social welfare benefits from federal antitrust enforcement play significant roles in explaining DOJ antitrust case-bringing activity. According to Siegfried,

> Perhaps this is not too surprising if we consider the reward structure confronting decision makers in the Antitrust Division. It is probably more important to win cases than to reduce economic losses or inequities in order to move up the success ladder in the Justice Department. Unfortunately, it has not been possible to identify a variable to measure the expected probability of winning various alternative cases. (p. 573)

(Incidentally, this study seeks to find what Siegfried was seeking – a variable or variables that measure the expected probability of winning various alternative cases.)

Consistent with Posner's and Siegfried's empirical findings, Weaver's (1977) empirical analysis suggests that prosecutors are less likely to bring cases based on their economic significance and more likely to bring cases when the alleged activity is clearly illegal or when the cases otherwise can be won easily. In addition, former U.S. Assistant Attorney General for Antitrust, Donald I. Baker (1978) claims that if the government does not think it can meet the "proof beyond a reasonable doubt" standard for a suspected price-fixing offense, it simply does not bring the case. In a more recent study, Snyder (1989) reports that the Antitrust Division tends secure more convictions in cases involving lower sales. If federal prosecutors are willing to target cases involving low sales, it is reasonable to suspect that they are primarily targeting cases that are easy to win.

Henceforth, I refer to prosecutors who are primarily motivated by (total) conviction rates as "conviction-motivated". The assumption that federal prosecutors are "conviction-motivated" in criminal antitrust cases is intuitively appealing and has solid and abundant support from empirical tests and other studies conducted by experts in the area of criminal antitrust enforcement – including practitioners like a notable Severth Circuit judge and a former U.S. Assistant Attorney General for Antitrust. It is a reasonable assumption.

In terms of my case selection model (i.e., $E(V) > 0$), assuming the prosecutor is conviction-motivated implies that relatively more weight should be placed on the previously predicted marginal effect of $\frac{\partial E(V)}{\partial \hat{\Pi}} > 0$. This assumption also implies that my case selection model's predictions should place less weight on the previously predicted marginal effects of $\frac{\partial E(V)}{\partial X} < 0$, $\frac{\partial E(V)}{\partial Y} > 0$, $\frac{\partial E(V)}{\partial \Theta} < 0$, and $\frac{\partial E(V)}{\partial S^T} > 0$.

Although the assumption that the prosecutor is conviction-motivated involves the litigants' common *ex ante* expected probability of trial conviction, $\hat{\Pi}$, the assumption does not mean that the prosecutor is motivated only by the prospects of a trial conviction. Consider the structure of the equation $E(V) = \Theta V^T + (1 - \Theta)V^B$. To the extent that a marginal increase in (or a higher value of) $\hat{\Pi}$ in a potential case increases her *ex ante* expected value of bringing that case, *E*(*V*), the prosecutor is motivated by the prospects of a conviction, even if the case goes to trial, since $\frac{\partial V^T}{\partial \hat{\Pi}} > 0$. The prosecutor is also (just as) motivated to bring a case by the other (second-order) positive effects on *E*(*V*) that are created by a marginal increase in (or a higher value of) $\hat{\Pi}$ in cases that she considers pursuing in the pre-indictment stage. In particular, the prosecutor's decision to bring a case is also motivated by a marginal increase in (higher value of) $\hat{\Pi}$ because a marginal increase in $\hat{\Pi}$ creates a marginal increase in V^B, which, in turn, creates a marginal increase in *E*(*V*).

In summary, it is important to account for the rational self-interested motivations of prosecutors (and judges) in the disposition of federal criminal antitrust cases. By accounting for different types of prosecutorial preferences, my case selection model can provide a meaningful and useful theoretical framework for identifying and analyzing potential sources and types of case selection bias in observed federal criminal antitrust case outcomes.[118]

It is worth noting an alternative or additional rationale for giving more analytical weight to the probability of trial conviction element of the prosecutor's *ex ante* expected value of bringing a case. Given the general degree of ambiguity regarding the prosecutor's *ex ante*

---

[118] It would be ideal to develop approaches that obviate the need for this signing convention. Given diseconomies of more "applied mathematics," looking for simplifying assumptions (that could be difficult to reflect empirically) is inappropriate. It is also possible to simplify the analysis by making certain endogenous variables exogenous instead. This would not allow for inspection of possibilities that previous researchers ignored by making certain variables exogenous.

Eisenberg and Farber (2003), for example, make simplifying assumptions to ease determining signs with no more or less justification than asserting reasonableness. In contrast, relying on a stylized "conviction-motivated" prosecutor in my model is justified by several academic studies and associated empirical findings. Researchers should make use of findings in previous studies – especially consensus empirical findings – when the subject of those findings (i.e., prosecutorial motivation) has not likely changed over time.

(theoretical) expectations of the effects of different case characteristics on her *ex ante* expected value of bringing a case, $E(V)$, it is reasonable to expect the prosecutor to focus on the single element of $E(V)$ associated with the least amount of ambiguity. That is, given the relative clarity of the expected effects of changes in the values of case characteristics on the *ex ante* expected probability of trial conviction, $\hat{\Pi}$, it is reasonable to expect the prosecutor to focus on those (partial) effects in her *ex ante* evaluation of $E(V)$ as part of her case bringing decision process.

Tables A-1.a. through A-6.b. summarize the predicted signs of the partial and total effects on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from increases (and decreases) in different case characteristics, as shown in Equations [30]-[35] and discussed above. In these tables, I make various assumptions regarding the whether the prosecutor or defendant is the trial favorite and the parties' relative diminishing returns in trial expenditure. These tables also show how predictions regarding the net effects of changes in the values of case characteristics on $E(V)$ are influenced by different assumptions regarding the prosecutor's relative preferences for different aspects of federal criminal antitrust case disposition. The predictions summarized in these tables are integral to predictions regarding case selection effects on the (conditional) *ex post* probabilities of trial and trial conviction, as well as *ex post* expected sentences.

Table 21. a. Predicted signs of the partial effects on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the prosecutor's cost per unit of trial expenditure, C

Predicted Partial Effects on $E(V)$ from Increases (or Decreases) In C	Prosecutor is Trial Favorite ($\hat{\Pi} \geq \overline{\Pi}$)	Prosecutor is Trial Underdog ($\hat{\Pi} < \overline{\Pi}$)
$\frac{\partial E(V)}{\partial C}$	< 0 (> 0)	
$\frac{\partial E(V)}{\partial X}\frac{dX^*}{dC}$	> 0 (< 0)	
$\frac{\partial E(V)}{\partial Y}\frac{dY^*}{dC}$	> 0 (< 0)	< 0 (> 0)
$\frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{dC}$	? (?)	
$\frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{dC}$	< 0 (> 0)	
$\frac{\partial E(V)}{\partial S^T}\frac{dS^T}{dC}$	<0 (>0)	? (?)

Table 21. b. Predicted signs of the total (net) effect on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the prosecutor's cost per unit of trial expenditure, C

Predicted Total Effect on $E(V)$ from Increases (or Decreases) in C – i.e., $\frac{dE(V)}{dC}$ – for Different Prosecutorial Motivations	Prosecutor is Trial Favorite ($\hat{\Pi} \geq \overline{\Pi}$)	Prosecutor is Trial Underdog ($\hat{\Pi} < \overline{\Pi}$)
Own Workload-Averse	? (?)	
Defendant's Workload-Motivated	? (?)	< 0 (> 0)
Trial-Averse	? (?)	
Conviction-Motivated	< 0 (> 0)	
Penalty-Motivated	< 0 (> 0)	? (?)

Table 22. a. Predicted signs of the partial effects on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the defendant's cost per unit of trial expenditure, K

Predicted Partial Effects on $E(V)$ from Increases (or Decreases) in K	Prosecutor is Trial Favorite ($\hat{\Pi} \geq \overline{\Pi}$)	Prosecutor is Trial Underdog ($\hat{\Pi} < \overline{\Pi}$)
$\frac{\partial E(V)}{\partial K}$	> 0 (< 0)	
$\frac{\partial E(V)}{\partial X} \frac{dX^*}{dK}$	> 0 (< 0)	< 0 (> 0)
$\frac{\partial E(V)}{\partial Y} \frac{dY^*}{dK}$	< 0 (> 0)	
$\frac{\partial E(V)}{\partial \Theta} \frac{d\Theta}{dK}$	? (?)	
$\frac{\partial E(V)}{\partial \hat{\Pi}} \frac{d\hat{\Pi}}{dK}$	> 0 (< 0)	
$\frac{\partial E(V)}{\partial S^T} \frac{dS^T}{dK}$	> 0 (< 0)	? (?)

Table 22. b. Predicted signs of the total (net) effect on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the defendant's cost per unit of trial expenditure, K

Predicted Total Effect on $E(V)$ from Increases (or Decreases) in K – i.e., $\frac{dE(V)}{dK}$ – for Different Prosecutorial Motivations	Prosecutor is Trial Favorite ($\hat{\Pi} \geq \overline{\Pi}$)	Prosecutor is Trial Underdog ($\hat{\Pi} < \overline{\Pi}$)
Own Workload-Averse	> 0 (< 0)	? (?)
Defendant's Workload-Motivated	? (?)	
Trial-Averse	? (?)	
Conviction-Motivated	> 0 (< 0)	
Penalty-Motivated	> 0 (< 0)	? (?)

Table 23. a. Predicted signs of the partial effects on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the commonly expected sentence following trial conviction, S^T

Predicted Partial Effects on $E(V)$ from Increases (or Decreases) in S^T*	Prosecutor is Trial Favorite or Underdog ($0 \leq \hat{\Pi} \leq 1$)
$\frac{\partial E(V)}{\partial S^T}$	> 0 (< 0)
$\frac{\partial E(V)}{\partial X}\frac{dX^*}{dS^T}$	< 0 (> 0)
$\frac{\partial E(V)}{\partial Y}\frac{dY^*}{dS^T}$	> 0 (< 0)
$\frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{dS^T}$	? (?)
$\frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{dS^T}$	> 0 (> 0)**

* Assume the prosecutor has weaker diminishing returns in trial expenditure than the defendant has.

** Note that the direction of the predicted partial change in $E(V)$ is the same whether S^T increases or decreases.

Table 23. b. Predicted signs of the total (net) effect on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the commonly expected sentence following trial conviction, S^T

Predicted Total Effect on $E(V)$ from Increases (or Decreases) in S^T – i.e., $\frac{dE(V)}{dS^T}$ – for Different Prosecutorial "Orientations"*	Prosecutor is Trial Favorite or Underdog ($0 \leq \hat{\Pi} \leq 1$)
Own Workload-Averse	? (?)
Defendant's Workload-Motivated	> 0 (< 0)
Trial-Averse	? (?)
Conviction-Motivated	> 0 (?)
Penalty-Motivated	> 0 (< 0)

* Assume the prosecutor has weaker diminishing returns in trial expenditure than the defendant has.

Table 24. a. Predicted signs of the partial effects on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the prosecutor's evidence, **e**

Predicted Partial Effects on $E(V)$ from Increases (or Decreases) in e	Prosecutor is Trial Favorite ($\hat{\Pi} \geq \overline{\Pi}$)	Prosecutor is Trial Underdog ($\hat{\Pi} < \overline{\Pi}$)
$\frac{\partial E(V)}{\partial \mathbf{e}}$	0 (0)	
$\frac{\partial E(V)}{\partial X}\frac{dX^*}{d\mathbf{e}}$	< 0 (< 0)	> 0 (> 0)
$\frac{\partial E(V)}{\partial Y}\frac{dY^*}{d\mathbf{e}}$	< 0 (> 0)	> 0 (> 0)
$\frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{d\mathbf{e}}$	? (?)	
$\frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{d\mathbf{e}}$	> 0 (< 0)	
$\frac{\partial E(V)}{\partial S^T}\frac{dS^T}{d\mathbf{e}}$	? (?)	

Table 24. b. Predicted signs of the total (net) effect on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the prosecutor's evidence, **e**

Predicted Total Effect on $E(V)$ from Increases (or Decreases) in e – i.e., $\frac{dE(V)}{d\mathbf{e}}$ – for Different Prosecutorial Motivations	Prosecutor is Trial Favorite ($\hat{\Pi} \geq \overline{\Pi}$)	Prosecutor is Trial Underdog ($\hat{\Pi} < \overline{\Pi}$)
Own Workload-Averse	< 0 (< 0)	> 0 (> 0)
Defendant's Workload-Motivated	< 0 (> 0)	> 0 (> 0)
Trial-Averse	? (?)	
Conviction-Motivated	> 0 (< 0)	
Penalty-Motivated	? (?)	

Table 25. a. Predicted signs of the partial effects on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the variability of the trial court's conviction standard, λ

Predicted Partial Effects on $E(V)$ from Increases (or Decreases) in λ *	Prosecutor is Trial Favorite or Underdog ($0 \leq \hat{\Pi} \leq 1$)
$\frac{\partial E(V)}{\partial \lambda}$	0 (0)
$\frac{\partial E(V)}{\partial X}\frac{dX^*}{d\lambda}$	? (?)
$\frac{\partial E(V)}{\partial Y}\frac{dY^*}{d\lambda}$	? (?)
$\frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{d\lambda}$	? (?)
$\frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{d\lambda}$	? (?)
$\frac{\partial E(V)}{\partial S^T}\frac{dS^T}{d\lambda}$	? (?)

* Assume the prosecutor has weaker diminishing returns in trial expenditure than the defendant has.

Table 25. b. Predicted signs of the total (net) effect on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the variability of the trial court's conviction standard, λ

Predicted Total Effect on $E(V)$ from Increases (or Decreases) in λ – i.e., $\frac{dE(V)}{d\lambda}$ – for Different Prosecutorial Motivations*	Prosecutor is Trial Favorite or Underdog ($0 \leq \hat{\Pi} \leq 1$)
Own Workload-Averse	? (?)
Defendant's Workload-Motivated	? (?)
Trial-Averse	? (?)
Conviction-Motivated	? (?)
Penalty-Motivated	? (?)

* Assume the prosecutor has weaker diminishing returns in trial expenditure than the defendant has.

Table 26. a. Predicted signs of the partial effects on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the idiosyncratic component, ρ, of the prosecutor's *ex ante* expected probability of a trial conviction.

Predicted Partial Effects on $E(V)$ from Increases (or Decreases) in ρ	Prosecutor is Trial Favorite or Underdog ($0 \leq \hat{\Pi} \leq 1$)
$\frac{\partial E(V)}{\partial \rho}$	> 0 (< 0)
$\frac{\partial E(V)}{\partial X}\frac{dX^*}{d\rho}$	0 (0)
$\frac{\partial E(V)}{\partial Y}\frac{dY^*}{d\rho}$	0 (0)
$\frac{\partial E(V)}{\partial \Theta}\frac{d\Theta}{d\rho}$	< 0 (> 0)
$\frac{\partial E(V)}{\partial \hat{\Pi}}\frac{d\hat{\Pi}}{d\rho}$	0 (0)
$\frac{\partial E(V)}{\partial S^T}\frac{dS^T}{d\rho}$	0 (0)

Table 26. b. Predicted signs of the total (net) effect on the prosecutor's *ex ante* expected value of bringing a case, $E(V)$, from changes in the idiosyncratic component, ρ, of the prosecutor's *ex ante* expected probability of a trial conviction.

Predicted Total Effect on $E(V)$ from Increases (or Decreases) in ρ – i.e., $\frac{dE(V)}{d\rho}$ – for Different Prosecutorial "Orientations"	Prosecutor is Trial Favorite or Underdog ($0 \leq \hat{\Pi} \leq 1$)
Own Workload-Averse	> 0 (< 0)
Defendant's Workload-Motivated	> 0 (< 0)
Trial-Averse	? (?)
Conviction-Motivated	> 0 (< 0)
Penalty-Motivated	> 0 (< 0)

APPENDIX B

SUMMARY STATISTICS OF EMPIRICAL VARIABLES

The following tables summarize the data used for empirical testing, by samples of cases at various stages of case disposition. Table 27 summarizes the full set of data; Table 28, the sample of observations in which the defendants entered pleas of guilty of nolo contendere; Table 29, the sample of observations in which the defendants entered not guilty pleas; Table 30, the sample of observations in which the defendants were convicted at trial; and Table 31, the sample of observations in which the defendants were not convicted at trial (i.e., the defendant was acquitted, the case was dismissed, or the case ended in a directed verdict). Cursory reviews of Tables 27 through 31 reveal that the distributions of the variables differ across sub-samples associated with various stages and branches of case disposition.

Table 27 Summary statistics of the sample of observed defendants who were indicted

Variable	N	Mean	Std. Dev.	Min.	Max.
TRIAL	3093	.1943097	.3957323	0	1
TRIAL CONVICTION	3093	.042677	.2021606	0	1
FINE	3093	132.4474	334.9416	0	5120.323
INDICT YEAR	3093	67.52861	7.657799	56	79
LANDMARK	3093	3.284491	1.832706	1	8
DOJBUDGET	3093	8915.747	4096.589	4255.528	17229.21
%CRIMINAL	3093	.4749337	.1356336	.02	.72
ANNUAL FORTUNE500	3093	23.12997	13.63491	8	52
HIGH ANNUAL FORTUNE500	3093	.3301002	.470325	0	1
REPUB PRES INDICT YEAR	3093	.5189137	.4997229	0	1
#CODEFEND	3093	12.15745	11.5066	0	72
#CORP CODEFEND	3093	7.039444	5.490059	0	28

Table 27, continued

Variable	N	Mean	Std. Dev.	Min.	Max.
ONLY CORP DEFEND	3093	.6408018	.479843	0	1
ONLY INDIV DEFEND	3093	.0274814	.1635079	0	1
ONLY HIGH RANK	3093	.7982541	.4013684	0	1
ONLY LOW RANK	3093	.1732945	.3785629	0	1
ANY LOW RANK	3093	.2017459	.4013684	0	1
#OTHER INDICT	3093	.3831232	1.656947	0	20
ZERO OTHER INDICT	3093	.8677659	.3387998	0	1
#PREV CONVICT	3093	.1357905	.820376	0	19
ZERO PREV CONVICT	3093	.938571	.2401544	0	1
PROFIT	3093	2806.854	2371.634	-42.49872	12251.26
HIGH PROFIT	3093	.3912059	.4880993	0	1
COMPENSATION	3093	15782.14	13783.53	775.6696	66998.46
HIGH COMPENSATION	3093	.2938894	.4556156	0	1
DURATION	3093	5.47365	5.073748	1	40
HIGH DURATION	3093	.3624313	.4807802	0	1
COMMERCE	3093	64396.21	166428.9	0	2116402
HIGH COMMERCE	3093	.2033624	.402565	0	1
#COUNTS CONVICT	3093	.1260912	.3689188	0	2
MULTI COUNTS CONVICT	3093	.0129324	.1130014	0	1
MULTI COUNTS INDICT	3093	.3032654	.4597432	0	1
COUNTS INDICT-CONVICT	3093	1.228257	.6221573	0	3
PRICE FIX NOT BID RIG	3093	.6970579	.4596047	0	1
BID RIG NOT PRICE FIX	3093	.1218881	.3272095	0	1
ANY BID RIG	3093	.2563854	.4367077	0	1

Table 27, continued

Variable	N	Mean	Std. Dev.	Min.	Max.
CONSUMER VIC	3093	.2169415	.4122291	0	1
CASE TIME SPAN	3093	.9751051	1.05187	0	7
DOJ WIN PREV YEAR BY VIOL#1	3093	88.03136	11.5329	53	100
DOJ WIN PREV YEAR	3093	86.30391	12.10904	53	100
HIGH DOJ WIN PREV YEAR BY VIOL#1	3093	.4028451	.4905494	0	1
VAR DOJ WIN PREV YEAR BY VIOL#1	3093	13.46567	11.88678	0	33.33333
VAR DOJ WIN PREV YEAR	3093	15.5146	9.477293	0	33.33333
SENTENCE YEAR	3093	68.50372	7.496597	56	81
POST-APPA SENTENCE YEAR	3093	.3359198	.4723874	0	1
POST-APPA INDICT YEAR	3093	.28516	.4515636	0	1
REPUB JUDGE	3093	.3750404	.4842116	0	1
REPUB PRES SENTENCE YEAR	3093	.4390559	.4963522	0	1
PRES & JUDGE SAME PARTY SENTENCE YEAR	3093	.5706434	.4950644	0	1
PRIOR PROSECUTOR	3093	.5851924	.4927685	0	1
TENURE	3093	9.842548	7.422962	0	35
#FILINGS	3093	343.1261	114.7727	97	812

Table 28 Summary statistics of the sample of observed defendants who entered pleas of guilty or *nolo contendere*

Variable	N	Mean	Std. Dev.	Min.	Max.
TRIAL	2492	0	0	0	0
TRIAL CONVICTION	2492	0	0	0	0
FINE	2492	154.213	363.7022	0	5120.323
INDICT YEAR	2492	68.02729	7.432235	56	79
LANDMARK	2492	3.387685	1.935138	1	8
DOJBUDGET	2492	9102.976	4032.501	4255.528	17229.21
%CRIMINAL	2492	.4719141	.1391065	.17	.72
ANNUAL FORTUNE500	2492	22.95787	13.71904	8	52
HIGH ANNUAL FORTUNE500	2492	.3326645	.4712621	0	1
REPUB PRES INDICT YEAR	2492	.502809	.5000925	0	1
#CODEFEND	2492	12.2203	12.12626	0	72
#CORP CODEFEND	2492	6.831461	5.08148	0	23
ONLY CORP DEFEND	2492	.6508828	.4767867	0	1
ONLY INDIV DEFEND	2492	.0232745	.1508042	0	1
ONLY HIGH RANK	2492	.7993579	.4005611	0	1
ONLY LOW RANK	2492	.170947	.3765382	0	1
ANY LOW RANK	2492	.2006421	.4005611	0	1
#OTHER INDICT	2492	.4357945	1.785657	0	19
ZERO OTHER INDICT	2492	.8623596	.3445914	0	1
#PREV CONVICT	2492	.144061	.8494176	0	19
ZERO PREV CONVICT	2492	.9357945	.2451678	0	1
PROFIT	2492	2757.115	2274.806	-42.49872	12251.26
HIGH PROFIT	2492	.3972713	.4894312	0	1
COMPENSATION	2492	16078.62	13754.51	775.6696	66998.46
HIGH COMPENSATION	2492	.2973515	.4571843	0	1

Table 28, continued

Variable	N	Mean	Std. Dev.	Min.	Max.
DURATION	2492	5.672151	5.148191	1	40
HIGH DURATION	2492	.3760032	.4844781	0	1
COMMERCE	2492	68945.12	179346.9	0	2116402
HIGH COMMERCE	2492	.209069	.4067254	0	1
#COUNTS CONVICT	2492	.1039326	.3317057	0	2
MULTI COUNTS CONVICT	2492	.008427	.0914292	0	1
MULTI COUNTS INDICT	2492	.3138042	.464131	0	1
COUNTS INDICT-CONVICT	2492	1.26565	.6240656	0	3
PRICE FIX NOT BID RIG	2492	.6914125	.4620031	0	1
BID RIG NOT PRICE FIX	2492	.1175762	.3221703	0	1
ANY BID RIG	2492	.2624398	.4400487	0	1
CONSUMER VIC	2492	.1930177	.394746	0	1
CASE TIME SPAN	2492	.8896469	1.003137	0	7
DOJ WIN PREV YEAR BY VIOL#1	2492	88.53772	11.12419	53	100
DOJ WIN PREV YEAR	2492	87.15449	11.72352	53	100
HIGH DOJ WIN PREV YEAR BY VIOL#1	2492	.4097111	.491879	0	1
VAR DOJ WIN PREV YEAR BY VIOL#1	2492	13.68986	11.96106	0	33.33333
VAR DOJ WIN PREV YEAR	2492	15.50195	9.48503	0	33.33333

Table 28, continued

SENTENCE YEAR	2492	68.91693	7.292804	56	80
POST-APPA SENTENCE YEAR	2492	.3374799	.472945	0	1
POST-APPA INDICT YEAR	2492	.2905297	.4540979	0	1
REPUB JUDGE	2492	.3579454	.4794922	0	1
REPUB PRES SENTENCE YEAR	2492	.4357945	.49596	0	1
PRES & JUDGE SAME PARTY SENTENCE YEAR	2492	.5778491	.4940015	0	1
PRIOR PROSECUTOR	2492	.6079454	.4883067	0	1
TENURE	2492	10.15048	7.562308	0	35
#FILINGS	2492	342.75	113.7765	97	812

Table 29 Summary statistics of the sample of observed defendants who entered not guilty pleas

Variable	N	Mean	Std. Dev.	Min.	Max.
TRIAL	601	1	0	1	1
TRIAL CONVICTION	601	.2196339	.4143435	0	1
FINE	601	42.19807	137.2063	0	1619.433
INDICT YEAR	601	65.4609	8.219218	56	79
LANDMARK	601	2.856606	1.238733	1	8
DOJBUDGET	601	8139.416	4268.916	4255.528	17229.21
%CRIMINAL	601	.4874542	.1194613	.02	.72
ANNUAL FORTUNE500	601	23.84359	13.26771	8	52
HIGH ANNUAL FORTUNE500	601	.3194676	.4666587	0	1
REPUB PRES INDICT YEAR	601	.5856905	.4930127	0	1
#CODEFEND	601	11.89684	8.469908	0	72
#CORP CODEFEND	601	7.90183	6.87013	0	28
ONLY CORP DEFEND	601	.5990017	.4905089	0	1
ONLY INDIV DEFEND	601	.0449251	.2073123	0	1
ONLY HIGH RANK	601	.7936772	.405002	0	1
ONLY LOW RANK	601	.1830283	.3870118	0	1
ANY LOW RANK	601	.2063228	.405002	0	1
#OTHER INDICT	601	.1647255	.9225799	0	20
ZERO OTHER INDICT	601	.890183	.3129219	0	1
#PREV CONVICT	601	.1014975	.6865477	0	10
ZERO PREV CONVICT	601	.9500832	.2179545	0	1
PROFIT	601	3013.089	2729.285	3.229279	12251.26
HIGH PROFIT	601	.3660566	.4821265	0	1
COMPENSATION	601	14552.8	13847.1	775.6696	45984.84
HIGH COMPENSATION	601	.2795341	.4491441	0	1

Table 29, continued

Variable	N	Mean	Std. Dev.	Min.	Max.
DURATION	601	4.650582	4.667373	1	25
HIGH DURATION	601	.3061564	.4612794	0	1
COMMERCE	601	45534.5	93584.29	0	807319.7
HIGH COMMERCE	601	.1797005	.3842576	0	1
#COUNTS CONVICT	601	.21797	.4838146	0	2
MULTI COUNTS CONVICT	601	.031614	.1751158	0	1
#COUNTS INDICT	601	1.291181	.5196846	1	3
COMMERCE	601	45534.5	93584.29	0	807319.7
HIGH COMMERCE	601	.1797005	.3842576	0	1
MULTI COUNTS INDICT	601	.2595674	.4387624	0	1
COUNTS INDICT-CONVICT	601	1.073211	.5898852	0	3
PRICE FIX NOT BID RIG	601	.7204659	.4491441	0	1
BID RIG NOT PRICE FIX	601	.1397671	.347034	0	1
ANY BID RIG	601	.2312812	.422003	0	1
CONSUMER VIC	601	.3161398	.4653555	0	1
CASE TIME SPAN	601	1.329451	1.169593	0	6
DOJ WIN PREV YEAR BY VIOL#1	601	85.93178	12.89161	53	100
DOJ WIN PREV YEAR	601	82.77704	13.02127	53	100
HIGH DOJ WIN PREV YEAR BY VIOL#1	601	.374376	.4843645	0	1
VAR DOJ WIN PREV YEAR BY VIOL#1	601	12.53608	11.537	0	33.33333
VAR DOJ WIN PREV YEAR	601	15.56709	9.452843	0	33.33333

Table 29, continued

Variable	N	Mean	Std. Dev.	Min.	Max.
SENTENCE YEAR	601	66.79035	8.071925	56	81
POST-APPA SENTENCE YEAR	601	.3294509	.4704054	0	1
POST-APPA INDICT YEAR	601	.2628952	.4405727	0	1
REPUB JUDGE	601	.4459235	.4974812	0	1
REPUB PRES SENTENCE YEAR	601	.452579	.4981608	0	1
PRES & JUDGE SAME PARTY SENTENCE YEAR	601	.5407654	.4987505	0	1
PRIOR PROSECUTOR	601	.4908486	.5003327	0	1
TENURE	601	8.565724	6.670539	0	27
#FILINGS	601	344.6855	118.899	123	812

Table 30 Summary statistics of the sample of observed defendants convicted at trial

Variable	N	Mean	Std. Dev.	Min.	Max.
TRIAL	132	1	0	1	1
TRIAL CONVICTION	132	1	0	1	1
FINE	132	192.1291	239.1611	0	1619.433
INDICT YEAR	132	66.53788	8.903559	56	79
LANDMARK	132	2.774242	1.28473	1	8
DOJBUDGET	132	8525.326	4445.345	4255.528	17229.21
%CRIMINAL	132	.515303	.1228755	.02	.72
ANNUAL FORTUNE500	132	19.06818	8.712712	8	52
HIGH ANNUAL FORTUNE500	132	.1136364	.3185781	0	1
REPUB PRES INDICT YEAR	132	.7424242	.438965	0	1
#CODEFEND	132	9.560606	7.855817	0	72
#CORP CODEFEND	132	5.94697	4.221563	0	23
ONLY CORP DEFEND	132	.6287879	.4849696	0	1
ONLY INDIV DEFEND	132	.0833333	.2774383	0	1
ONLY HIGH RANK	132	.8939394	.3090882	0	1
ONLY LOW RANK	132	.0909091	.288575	0	1
ANY LOW RANK	132	.1060606	.3090882	0	1
#OTHER INDICT	132	.1136364	.3185781	0	1
ZERO OTHER INDICT	132	.8863636	.3185781	0	1
#PREV CONVICT	132	.0606061	.2695083	0	2
ZERO PREV CONVICT	132	.9469697	.2249476	0	1
PROFIT	132	2579.307	2633.807	136.4176	12251.26
HIGH PROFIT	132	.280303	.4508583	0	1
COMPENSATION	132	15201.25	14570.15	2070.64	45984.84
HIGH COMPENSATION	132	.3333333	.4732004	0	1

Table 30, continued

Variable	N	Mean	Std. Dev.	Min.	Max.
DURATION	132	5.545455	5.315057	1	20
HIGH DURATION	132	.3712121	.4849696	0	1
COMMERCE	132	61510.3	113788.4	0	807319.7
HIGH COMMERCE	132	.2121212	.4103676	0	1
#COUNTS CONVICT	132	.1515152	.3805349	0	2
MULTI COUNTS CONVICT	132	.0075758	.0870388	0	1
#COUNTS INDICT	132	1.181818	.4247708	1	3
MULTI COUNTS INDICT	132	.1666667	.3740977	0	1
COUNTS INDICT-CONVICT	132	1.030303	.5654073	0	3
PRICE FIX NOT BID RIG	132	.6590909	.4758206	0	1
BID RIG NOT PRICE FIX	132	.2045455	.4049057	0	1
ANY BID RIG	132	.2272727	.4206667	0	1
CONSUMER VIC	132	.4166667	.4948848	0	1
CASE TIME SPAN	132	1.219697	1.100412	0	5
DOJ WIN PREV YEAR BY VIOL#1	132	83.15909	15.18579	53	100
DOJ WIN PREV YEAR	132	81.62879	14.64981	53	100
HIGH DOJ WIN PREV YEAR BY VIOL#1	132	.4166667	.4948848	0	1
VAR DOJ WIN PREV YEAR BY VIOL#1	132	11.00835	11.59196	0	33.33333
VAR DOJ WIN PREV YEAR	132	16.40768	8.399533	0	33.33333

Table 30, continued

Variable	N	Mean	Std. Dev.	Min.	Max.
SENTENCE YEAR	132	67.75758	8.80243	56	81
POST-APPA SENTENCE YEAR	132	.4393939	.498204	0	1
POST-APPA INDICT YEAR	132	.2878788	.4544992	0	1
REPUB JUDGE	132	.6590909	.4758206	0	1
REPUB PRES SENTENCE YEAR	132	.5606061	.498204	0	1
PRES & JUDGE SAME PARTY SENTENCE YEAR	132	.4772727	.501386	0	1
PRIOR PROSECUTOR	132	.5909091	.4935391	0	1
TENURE	132	7.590909	6.049357	0	25
#FILINGS	132	352.2197	122.2653	194	728

Table 31 Summary statistics of the sample of observations in which the defendants went to trial and were not convicted at trial

Variable	N	Mean	Std. Dev.	Min.	Max.
TRIAL	469	1	0	1	1
TRIAL CONVICTION	469	0	0	0	0
FINE	469	0	0	0	0
INDICT YEAR	469	65.15778	8.000043	56	79
LANDMARK	469	2.879787	1.225886	1	8
DOJBUDGET	469	8030.802	4216.444	4255.528	17229.21
%CRIMINAL	469	.4796162	.1174291	.02	.72
ANNUAL FORTUNE500	469	25.18763	14.00683	8	52
HIGH ANNUAL FORTUNE500	469	.3773987	.4852535	0	1
REPUB PRES INDICT YEAR	469	.5415778	.4988003	0	1
#CODEFEND	469	12.55437	8.527966	0	72
#CORP CODEFEND	469	8.452026	7.358083	0	28
ONLY CORP DEFEND	469	.5906183	.4922448	0	1
ONLY INDIV DEFEND	469	.0341151	.1817188	0	1
ONLY HIGH RANK	469	.7654584	.4241644	0	1
ONLY LOW RANK	469	.2089552	.4069965	0	1
ANY LOW RANK	469	.2345416	.4241644	0	1
#OTHER INDICT	469	.1791045	1.030471	0	20
ZERO OTHER INDICT	469	.891258	.3116477	0	1
#PREV CONVICT	469	.1130064	.7637775	0	10
ZERO PREV CONVICT	469	.9509595	.2161832	0	1
PROFIT	469	3135.177	2745.949	3.229279	12251.26
HIGH PROFIT	469	.3901919	.488314	0	1
COMPENSATION	469	14370.29	13647.29	775.6696	45984.84
HIGH COMPENSATION	469	.2643923	.44148	0	1

Table 31, continued

Variable	N	Mean	Std. Dev.	Min.	Max.
DURATION	469	4.398721	4.442039	1	25
HIGH DURATION	469	.2878465	.4532426	0	1
COMMERCE	469	41038.11	86670.1	0	586510.3
HIGH COMMERCE	469	.1705757	.3765394	0	1
#COUNTS CONVICT	469	.2366738	.5079059	0	2
MULTI COUNTS CONVICT	469	.0383795	.1923159	0	1
#COUNTS INDICT	469	1.321962	.5398318	1	3
MULTI COUNTS INDICT	469	.2857143	.4522363	0	1
COUNTS INDICT-CONVICT	469	1.085288	.5966224	0	3
PRICE FIX NOT BID RIG	469	.7377399	.440333	0	1
BID RIG NOT PRICE FIX	469	.1215352	.3270971	0	1
ANY BID RIG	469	.2324094	.4228196	0	1
CONSUMER VIC	469	.2878465	.4532426	0	1
CASE TIME SPAN	469	1.360341	1.187634	0	6
DOJ WIN PREV YEAR BY VIOL#1	469	86.71215	12.07227	53	100
DOJ WIN PREV YEAR	469	83.10021	12.523	53	100
HIGH DOJ WIN PREV YEAR BY VIOL#1	469	.3624733	.4812278	0	1
VAR DOJ WIN PREV YEAR BY VIOL#1	469	12.96607	11.49727	0	33.33333
VAR DOJ WIN PREV YEAR	469	15.33051	9.723957	0	33.33333

Table 31, continued

Variable	N	Mean	Std. Dev.	Min.	Max.
SENTENCE YEAR	469	66.51812	7.842615	56	80
POST-APPA SENTENCE YEAR	469	.2985075	.4580919	0	1
POST-APPA INDICT YEAR	469	.2558635	.4368114	0	1
REPUB JUDGE	469	.3859275	.4873334	0	1
REPUB PRES SENTENCE YEAR	469	.4221748	.4944335	0	1
PRES & JUDGE SAME PARTY SENTENCE YEAR	469	.5586354	.4970802	0	1
PRIOR PROSECUTOR	469	.4626866	.4991382	0	1
TENURE	469	8.840085	6.816114	0	27
#FILINGS	469	342.565	117.9805	123	812

APPENDIX C

CROSS-BRANCH MEANS TESTS OF THE TRIAL SELECTION MODEL

This appendix reports the results of the means tests used to test the predictions of the trial selection model. Table 10 in Chapter X summarizes the results reported here. These means tests test the difference in the values of the means of the values of empirical variables that represent case characteristics in the sample of defendants who go to trial versus the sample of defendants who avoid trial.

```
ttest COUNTS INDICT-CONVICT, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492     1.26565    .0125013    .6240656    1.241136    1.290164
       1 |     601    1.073211    .0240619    .5898852    1.025955    1.120467
---------+--------------------------------------------------------------------
combined |    3093    1.228257    .0111869    .6221573    1.206323    1.250192
---------+--------------------------------------------------------------------
    diff |            .1924388    .0280654                .1374101    .2474675
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   6.8568
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000

.

ttest BID RIG NOT PRICE FIX, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .1175762    .0064537    .3221703     .104921    .1302315
       1 |     601    .1397671    .0141558     .347034    .1119661     .167568
---------+--------------------------------------------------------------------
combined |    3093    .1218881    .0058835    .3272095    .1103522    .1334241
---------+--------------------------------------------------------------------
    diff |           -.0221908    .0148668               -.0513407    .0069591
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -1.4926
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0678         Pr(|T| > |t|) = 0.1356          Pr(T > t) = 0.9322

.
```

ttest *PROFIT*,by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    2757.115    45.56909    2274.806    2667.758    2846.473
       1 |     601    3013.089    111.3299    2729.285    2794.446    3231.733
---------+--------------------------------------------------------------------
combined |    3093    2806.854    42.64398    2371.634     2723.24    2890.467
---------+--------------------------------------------------------------------
    diff |           -255.9737    107.6961               -467.137   -44.81049
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -2.3768
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0088         Pr(|T| > |t|) = 0.0175          Pr(T > t) = 0.9912
```

ttest *COMPENSATION*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    16078.62    275.5313    13754.51    15538.33    16618.92
       1 |     601     14552.8    564.8349     13847.1    13443.51    15662.09
---------+--------------------------------------------------------------------
combined |    3093    15782.14    247.8395    13783.53    15296.19    16268.09
---------+--------------------------------------------------------------------
    diff |            1525.824    625.8819                298.6372     2753.01
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   2.4379
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9926         Pr(|T| > |t|) = 0.0148          Pr(T > t) = 0.0074
```

ttest *#OTHER INDICT*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .4357945    .0357704    1.785657    .3656517    .5059373
       1 |     601    .1647255    .0376328    .9225799    .0908174    .2386335
---------+--------------------------------------------------------------------
combined |    3093    .3831232    .0297933    1.656947    .3247065    .4415398
---------+--------------------------------------------------------------------
    diff |            .2710691    .0751528                .1237145    .4184236
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   3.6069
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9998         Pr(|T| > |t|) = 0.0003          Pr(T > t) = 0.0002
```

```
ttest  #CORP CODEFEND, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    6.831461    .1017926     5.08148    6.631854    7.031068
       1 |     601     7.90183    .2802384     6.87013    7.351463    8.452198
---------+--------------------------------------------------------------------
combined |    3093    7.039444    .0987159    5.490059    6.845889    7.232999
---------+--------------------------------------------------------------------
    diff |            -1.07037     .248788               -1.558176    -.582563
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -4.3023
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000

.
ttest CASE TIME SPAN, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .8896469    .0200949    1.003137    .8502424    .9290513
       1 |     601    1.329451    .0477087    1.169593    1.235755    1.423147
---------+--------------------------------------------------------------------
combined |    3093    .9751051    .0189135     1.05187    .9380208    1.012189
---------+--------------------------------------------------------------------
    diff |            -.439804    .0471502               -.5322529   -.3473552
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -9.3277
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000


ttest  LANDMARK, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    3.387685    .0387648    1.935138     3.31167    3.463699
       1 |     601    2.856606     .050529    1.238733    2.757371    2.955841
---------+--------------------------------------------------------------------
combined |    3093    3.284491    .0329536    1.832706    3.219878    3.349104
---------+--------------------------------------------------------------------
    diff |            .5310789    .0827498                .3688287    .6933292
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   6.4179
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

Ttest *PRIOR PROSECUTOR*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .6079454    .0097818    .4883067    .5887641    .6271267
       1 |     601    .4908486     .020409    .5003327    .4507668    .5309303
---------+--------------------------------------------------------------------
combined |    3093    .5851924    .0088604    .4927685    .5678195    .6025652
---------+--------------------------------------------------------------------
    diff |            .1170968    .0222979                .0733767    .1608169
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   5.2515
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

ttest *CONSUMER VIC*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .1930177    .0079076     .394746    .1775115    .2085238
       1 |     601    .3161398    .0189822    .4653555      .27886    .3534195
---------+--------------------------------------------------------------------
combined |    3093    .2169415    .0074122    .4122291    .2024081    .2314749
---------+--------------------------------------------------------------------
    diff |           -.1231221    .0186051               -.1596018   -.0866424
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -6.6176
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

ttest *COMPETITOR VIC*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .0276886    .0032875     .164112    .0212421    .0341351
       1 |     601    .0415973    .0081514    .1998335    .0255886     .057606
---------+--------------------------------------------------------------------
combined |    3093    .0303912    .0030871     .171689    .0243382    .0364442
---------+--------------------------------------------------------------------
    diff |           -.0139087    .0077995               -.0292015     .001384
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -1.7833
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0373         Pr(|T| > |t|) = 0.0746          Pr(T > t) = 0.9627
.
```

ttest *GOV VIC*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .2207063    .0083094     .414806    .2044122    .2370004
       1 |     601    .1780366    .0156173    .3828622    .1473654    .2087078
---------+--------------------------------------------------------------------
combined |    3093    .2124151    .0073557    .4090832    .1979927    .2268376
---------+--------------------------------------------------------------------
    diff |            .0426697    .0185776                .0062439    .0790954
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   2.2968
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9892         Pr(|T| > |t|) = 0.0217          Pr(T > t) = 0.0108
.
```

ttest *FIRM VIC*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .5585875     .009949    .4966553    .5390782    .5780967
       1 |     601    .4642263    .0203601     .499134    .4242406     .504212
---------+--------------------------------------------------------------------
combined |    3093    .5402522    .0089627    .4984577    .5226788    .5578256
---------+--------------------------------------------------------------------
    diff |            .0943612     .022592                .0500643    .1386581
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   4.1767
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

ttest *#COUNTS INDICT*_hpsv, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    1.098716    .0085304    .4258394    1.081988    1.115443
       1 |     601    1.043261      .01516    .3716522    1.013488    1.073034
---------+--------------------------------------------------------------------
combined |    3093    1.087941     .007487    .4163851    1.073261     1.10262
---------+--------------------------------------------------------------------
    diff |            .0554547    .0188991                .0183987    .0925106
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   2.9343
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9983         Pr(|T| > |t|) = 0.0034          Pr(T > t) = 0.0017
```

ttest *ANY BID RIG*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .2624398    .0088151    .4400487    .2451542    .2797255
       1 |     601    .2312812    .0172139     .422003    .1974745    .2650879
---------+--------------------------------------------------------------------
combined |    3093    .2563854    .0078524    .4367077     .240989    .2717818
---------+--------------------------------------------------------------------
    diff |            .0311586    .0198411               -.0077445    .0700618
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   1.5704
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9418         Pr(|T| > |t|) = 0.1164          Pr(T > t) = 0.0582
```

.

.

ttest *MULTI COUNTS INDICT*_hpsv, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .1448636     .007052    .3520339    .1310352    .1586919
       1 |     601    .0915141    .0117714    .2885791     .068396    .1146323
---------+--------------------------------------------------------------------
combined |    3093    .1344973    .0061358     .341241    .1224666    .1465279
---------+--------------------------------------------------------------------
    diff |            .0533494    .0154802                .0229968     .083702
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   3.4463
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9997         Pr(|T| > |t|) = 0.0006          Pr(T > t) = 0.0003
```

.

ttest *MULTI COUNTS INDICT*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .3138042    .0092975     .464131    .2955725    .3320358
       1 |     601    .2595674    .0178975    .4387624     .224418    .2947167
---------+--------------------------------------------------------------------
combined |    3093    .3032654    .0082666    .4597432    .2870569     .319474
---------+--------------------------------------------------------------------
    diff |            .0542368    .0208733                .0133099    .0951637
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   2.5984
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9953         Pr(|T| > |t|) = 0.0094          Pr(T > t) = 0.0047
```

.

```
ttest #COUNTS INDICT, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    1.369583    .0117608    .5871002    1.346521    1.392645
       1 |     601    1.291181    .0211984    .5196846    1.249549    1.332813
---------+--------------------------------------------------------------------
combined |    3093    1.354349    .0103458    .5753772    1.334063    1.374634
---------+--------------------------------------------------------------------
    diff |            .0784013    .0261138                .0271992    .1296034
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   3.0023
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9986         Pr(|T| > |t|) = 0.0027          Pr(T > t) = 0.0014
```

```
ttest  %CRIMINAL, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .4719141    .0027866    .1391065    .4664499    .4773784
       1 |     601    .4874542    .0048729    .1194613    .4778842    .4970243
---------+--------------------------------------------------------------------
combined |    3093    .4749337    .0024388    .1356336    .4701519    .4797156
---------+--------------------------------------------------------------------
    diff |           -.0155401    .0061584               -.0276151   -.0034651
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -2.5234
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0058         Pr(|T| > |t|) = 0.0117          Pr(T > t) = 0.9942
```

```
ttest DOJ TRIAL WIN PREV YEAR, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    27.02271    .4675967     23.3424    26.10579    27.93963
       1 |     601    25.43097    .8643365    21.18947    23.73347    27.12846
---------+--------------------------------------------------------------------
combined |    3093    26.71342    .4125747    22.94524    25.90447    27.52237
---------+--------------------------------------------------------------------
    diff |            1.591745    1.042504               -.4523262    3.635816
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   1.5268
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9365         Pr(|T| > |t|) = 0.1269          Pr(T > t) = 0.0635

.
```

ttest *DURATION*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    5.672151     .103129    5.148191    5.469924    5.874378
       1 |     601    4.650582    .1903861    4.667373    4.276678    5.024487
---------+--------------------------------------------------------------------
combined |    3093     5.47365    .0912303    5.073748    5.294772    5.652528
---------+--------------------------------------------------------------------
    diff |            1.021569    .2298766                .5708421    1.472295
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   4.4440
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

ttest *ONLY INDIV DEFEND*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .0232745    .0030209    .1508042    .0173507    .0291983
       1 |     601    .0449251    .0084564    .2073123    .0283173     .061533
---------+--------------------------------------------------------------------
combined |    3093    .0274814      .00294    .1635079    .0217168     .033246
---------+--------------------------------------------------------------------
    diff |           -.0216506    .0074215               -.0362022   -.0070991
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -2.9173
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0018         Pr(|T| > |t|) = 0.0036          Pr(T > t) = 0.9982
```

ttest *ONLY CORP DEFEND*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .6508828     .009551    .4767867    .6321541    .6696116
       1 |     601    .5990017    .0200083    .4905089    .5597069    .6382964
---------+--------------------------------------------------------------------
combined |    3093    .6408018     .008628     .479843    .6238847     .657719
---------+--------------------------------------------------------------------
    diff |            .0518812    .0217896                .0091575    .0946048
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   2.3810
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9913         Pr(|T| > |t|) = 0.0173          Pr(T > t) = 0.0087
```

```
ttest  DOJBUDGET, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    9102.976    80.77937    4032.501    8944.575    9261.378
       1 |     601    8139.416    174.1327    4268.916    7797.432      8481.4
---------+--------------------------------------------------------------------
combined |    3093    8915.747    73.66012    4096.589     8771.32    9060.175
---------+--------------------------------------------------------------------
    diff |            963.5605    185.3881                600.0642    1327.057
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   5.1975
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

```
ttest COMMERCE, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    68945.12    3592.691    179346.9    61900.16    75990.09
       1 |     601     45534.5    3817.383    93584.29    38037.44    53031.55
---------+--------------------------------------------------------------------
combined |    3093    64396.21    2992.532    166428.9    58528.66    70263.76
---------+--------------------------------------------------------------------
    diff |            23410.63    7552.726                8601.756     38219.5
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   3.0996
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9990         Pr(|T| > |t|) = 0.0020          Pr(T > t) = 0.0010
```

```
ttest DOJ TRIAL WIN PREV YEAR BY VIOL#1, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    29.15711    .5979812    29.85118    27.98452     30.3297
       1 |     601    26.82892    1.199323    29.40176    24.47354     29.1843
---------+--------------------------------------------------------------------
combined |    3093    28.70472    .5353602    29.77392    27.65502    29.75442
---------+--------------------------------------------------------------------
    diff |            2.328188    1.352624               -.3239435     4.98032
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   1.7212
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9573         Pr(|T| > |t|) = 0.0853          Pr(T > t) = 0.0427
```

```
ttest COMMERCE PER YEAR, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    18921.37     806.337    40252.29    17340.21    20502.53
       1 |     601    14748.56    1592.712    39045.82    11620.59    17876.53
---------+--------------------------------------------------------------------
combined |    3093    18110.55    720.1064    40048.53    16698.62    19522.49
---------+--------------------------------------------------------------------
    diff |            4172.815    1818.721                 606.791    7738.839
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   2.2944
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9891         Pr(|T| > |t|) = 0.0218          Pr(T > t) = 0.0109
```

```
ttest PRICE FIX NOT BID RIG, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .6914125    .0092549    .4620031    .6732645    .7095606
       1 |     601    .7204659     .018321    .4491441    .6844849    .7564469
---------+--------------------------------------------------------------------
combined |    3093    .6970579    .0082641    .4596047    .6808542    .7132615
---------+--------------------------------------------------------------------
    diff |           -.0290534    .0208832               -.0699998     .011893
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -1.3912
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0821         Pr(|T| > |t|) = 0.1643          Pr(T > t) = 0.9179
```

```
ttest ANY HPSV, by(TRIAL)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .9538523    .0042037    .2098469    .9456093    .9620954
       1 |     601    .9517471    .0087488    .2144787    .9345652     .968929
---------+--------------------------------------------------------------------
combined |    3093    .9534433     .003789    .2107215    .9460141    .9608724
---------+--------------------------------------------------------------------
    diff |            .0021052    .0095776               -.0166738    .0208842
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   0.2198
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.5870         Pr(|T| > |t|) = 0.8260          Pr(T > t) = 0.4130

.
```

ttest *POST-APPA INDICT YEAR*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .2905297    .0090965    .4540979    .2726922    .3083672
       1 |     601    .2628952    .0179713    .4405727    .2276008    .2981895
---------+--------------------------------------------------------------------
combined |    3093      .28516    .0081195    .4515636    .2692399    .3010802
---------+--------------------------------------------------------------------
    diff |            .0276345    .0205183               -.0125963    .0678653
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   1.3468
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9109         Pr(|T| > |t|) = 0.1781          Pr(T > t) = 0.0891
```

.

ttest *ANNUAL FORTUNE500*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    22.95787     .274821    13.71904    22.41896    23.49677
       1 |     601    23.84359    .5412012    13.26771    22.78072    24.90647
---------+--------------------------------------------------------------------
combined |    3093    23.12997    .2451671    13.63491    22.64926    23.61068
---------+--------------------------------------------------------------------
    diff |           -.8857288    .6195232               -2.100448    .3289899
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -1.4297
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0765         Pr(|T| > |t|) = 0.1529          Pr(T > t) = 0.9235
```

ttest *#CODEFEND*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492     12.2203    .2429142    12.12626    11.74397    12.69664
       1 |     601    11.89684    .3454948    8.469908    11.21831    12.57536
---------+--------------------------------------------------------------------
combined |    3093    12.15745    .2068983     11.5066    11.75178    12.56312
---------+--------------------------------------------------------------------
    diff |            .3234664    .5229606               -.7019191    1.348852
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   0.6185
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.7319         Pr(|T| > |t|) = 0.5363          Pr(T > t) = 0.2681
```

ttest *ONLY HIGH RANK*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .7993579    .0080241    .4005611    .7836234    .8150925
       1 |     601    .7936772    .0165204     .405002    .7612324     .826122
---------+--------------------------------------------------------------------
combined |    3093    .7982541    .0072169    .4013684    .7841036    .8124046
---------+--------------------------------------------------------------------
    diff |            .0056807    .0182425               -.030088    .0414495
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   0.3114
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.6222         Pr(|T| > |t|) = 0.7555          Pr(T > t) = 0.3778
```

ttest *ONLY LOW RANK*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492     .170947    .0075428    .3765382    .1561561    .1857379
       1 |     601    .1830283    .0157865    .3870118    .1520247    .2140319
---------+--------------------------------------------------------------------
combined |    3093    .1732945    .0068069    .3785629    .1599481     .186641
---------+--------------------------------------------------------------------
    diff |           -.0120813    .0172049              -.0458155     .021653
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -0.7022
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.2413         Pr(|T| > |t|) = 0.4826          Pr(T > t) = 0.7587

.
.
```

ttest *ANY LOW RANK*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492    .2006421    .0080241    .4005611    .1849075    .2163766
       1 |     601    .2063228    .0165204     .405002     .173878    .2387676
---------+--------------------------------------------------------------------
combined |    3093    .2017459    .0072169    .4013684    .1875954    .2158964
---------+--------------------------------------------------------------------
    diff |           -.0056807    .0182425              -.0414495     .030088
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -0.3114
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.3778         Pr(|T| > |t|) = 0.7555          Pr(T > t) = 0.6222
```

ttest *#PREV CONVICT*, by(*TRIAL*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |    2492     .144061    .0170156    .8494176    .1106948    .1774272
       1 |     601    .1014975    .0280049    .6865477     .046498     .156497
---------+--------------------------------------------------------------------
combined |    3093    .1357905    .0147511     .820376    .1068676    .1647133
---------+--------------------------------------------------------------------
    diff |            .0425635    .0372795               -.0305317    .1156586
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   1.1417
Ho: diff = 0                                     degrees of freedom =     3091

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.8732         Pr(|T| > |t|) = 0.2537          Pr(T > t) = 0.1268
```

APPENDIX D

CROSS-BRANCH MEANS TESTS OF THE TRIAL CONVICTION MODEL

This appendix reports the results of the means tests used to test the predictions of the trial conviction model. Table 11 in Chapter X summarizes these results. These means tests test the difference in the values of the means of the values of empirical variables that represent case characteristics in the sample of defendants convicted at trial versus the sample of defendants who go to trial but avoid trial conviction.

```
ttest PROFIT, by(TRIAL CONVICTION)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    3135.177    126.7962    2745.949    2886.017    3384.337
       1 |     132    2579.307    229.2435    2633.807    2125.809    3032.806
---------+--------------------------------------------------------------------
combined |     601    3013.089    111.3299    2729.285    2794.446    3231.733
---------+--------------------------------------------------------------------
    diff |            555.8696    268.1779                 29.1863    1082.553
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   2.0728
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9807         Pr(|T| > |t|) = 0.0386          Pr(T > t) = 0.0193
```

```
ttest MULTI COUNTS INDICT, by(TRIAL CONVICTION)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .2857143    .0208823    .4522363    .2446795     .326749
       1 |     132    .1666667     .032561    .3740977    .1022532    .2310801
---------+--------------------------------------------------------------------
combined |     601    .2595674    .0178975    .4387624     .224418    .2947167
---------+--------------------------------------------------------------------
    diff |            .1190476    .0429926                .0346131    .2034821
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   2.7690
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9971         Pr(|T| > |t|) = 0.0058          Pr(T > t) = 0.0029
```

ttest *#HPSV COUNTS INDICT*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    1.081023    .0169312    .3666699    1.047753    1.114294
       1 |     132    .9090909    .0312707    .3592729      .84723    .9709518
---------+--------------------------------------------------------------------
combined |     601    1.043261      .01516    .3716522    1.013488    1.073034
---------+--------------------------------------------------------------------
    diff |            .1719325    .0359695                .1012909    .2425741
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   4.7800
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

ttest *MULTI HPSV*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .1108742    .0145136    .3143116    .0823544     .139394
       1 |     132    .0227273     .013021    .1496004   -.0030315     .048486
---------+--------------------------------------------------------------------
combined |     601    .0915141    .0117714    .2885791     .068396    .1146323
---------+--------------------------------------------------------------------
    diff |            .0881469    .0282283                .0327085    .1435854
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   3.1226
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9991         Pr(|T| > |t|) = 0.0019          Pr(T > t) = 0.0009

.
```

ttest *ANY HPSV*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .9701493    .0078664    .1703572    .9546915     .985607
       1 |     132    .8863636    .0277287    .3185781    .8315097    .9412176
---------+--------------------------------------------------------------------
combined |     601    .9517471    .0087488    .2144787    .9345652     .968929
---------+--------------------------------------------------------------------
    diff |            .0837856    .0208711                .0427962     .124775
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   4.0144
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0001          Pr(T > t) = 0.0000

.
.
```

ttest *CONSUMER VIC*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .2878465    .0209288    .4532426    .2467204    .3289725
       1 |     132    .4166667    .0430742    .4948848    .3314556    .5018777
---------+--------------------------------------------------------------------
combined |     601    .3161398    .0189822    .4653555      .27886    .3534195
---------+--------------------------------------------------------------------
    diff |           -.1288202    .0455864               -.2183488   -.0392916
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -2.8258
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0024         Pr(|T| > |t|) = 0.0049          Pr(T > t) = 0.9976
```

ttest *FIRM VIC*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .4840085    .0231007    .5002778    .4386146    .5294024
       1 |     132    .3939394    .0426911    .4904831    .3094863    .4783925
---------+--------------------------------------------------------------------
combined |     601    .4642263    .0203601     .499134    .4242406     .504212
---------+--------------------------------------------------------------------
    diff |            .0900691    .0490824               -.0063254    .1864637
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   1.8351
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9665         Pr(|T| > |t|) = 0.0670          Pr(T > t) = 0.0335
```

ttest *#COUNTS INDICT*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    1.321962    .0249271    .5398318    1.272979    1.370945
       1 |     132    1.181818    .0369716    .4247708     1.10868    1.254957
---------+--------------------------------------------------------------------
combined |     601    1.291181    .0211984    .5196846    1.249549    1.332813
---------+--------------------------------------------------------------------
    diff |            .1401434    .0509258                .0401286    .2401582
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   2.7519
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9969         Pr(|T| > |t|) = 0.0061          Pr(T > t) = 0.0031
```

ttest *VAR DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    12.96607    .5308947    11.49727    11.92283     14.0093
       1 |     132    11.00835    1.008951    11.59196    9.012399    13.00429
---------+--------------------------------------------------------------------
combined |     601    12.53608    .4706039      11.537    11.61185    13.46032
---------+--------------------------------------------------------------------
    diff |            1.957721    1.134861               -.2710692    4.186512
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   1.7251
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9575         Pr(|T| > |t|) = 0.0850          Pr(T > t) = 0.0425
```

ttest *%CRIMINAL*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .4796162    .0054224    .1174291     .468961    .4902714
       1 |     132     .515303    .0106949    .1228755    .4941459    .5364602
---------+--------------------------------------------------------------------
combined |     601    .4874542    .0048729    .1194613    .4778842    .4970243
---------+--------------------------------------------------------------------
    diff |           -.0356868    .0116896               -.0586445   -.0127292
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -3.0529
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0012         Pr(|T| > |t|) = 0.0024          Pr(T > t) = 0.9988
```

ttest *#CODEFEND*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    12.55437    .3937849    8.527966    11.78057    13.32818
       1 |     132    9.560606    .6837611    7.855817    8.207964    10.91325
---------+--------------------------------------------------------------------
combined |     601    11.89684    .3454948    8.469908    11.21831    12.57536
---------+--------------------------------------------------------------------
    diff |            2.993765    .8262217                1.371122    4.616408
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   3.6234
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9998         Pr(|T| > |t|) = 0.0003          Pr(T > t) = 0.0002
```

ttest *PRIOR PROSECUTOR*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .4626866    .0230481    .4991382    .4173961    .5079771
       1 |     132    .5909091    .0429571    .4935391    .5059298    .6758884
---------+--------------------------------------------------------------------
combined |     601    .4908486     .020409    .5003327    .4507668    .5309303
---------+--------------------------------------------------------------------
    diff |           -.1282225    .0490594               -.2245719   -.0318731
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -2.6136
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0046         Pr(|T| > |t|) = 0.0092          Pr(T > t) = 0.9954
```

ttest *ANNUAL FORTUNE500*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    25.18763    .6467752    14.00683    23.91669    26.45858
       1 |     132    19.06818    .7583442    8.712712      17.568    20.56837
---------+--------------------------------------------------------------------
combined |     601    23.84359    .5412012    13.26771    22.78072    24.90647
---------+--------------------------------------------------------------------
    diff |            6.119451    1.284231                 3.59731    8.641593
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   4.7651
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

ttest *ONLY INDIV DEFEND*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .0341151     .008391    .1817188    .0176265    .0506038
       1 |     132    .0833333    .0241479    .2774383     .035563    .1311036
---------+--------------------------------------------------------------------
combined |     601    .0449251    .0084564    .2073123    .0283173     .061533
---------+--------------------------------------------------------------------
    diff |           -.0492182    .0203442               -.0891727   -.0092637
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -2.4193
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0079         Pr(|T| > |t|) = 0.0158          Pr(T > t) = 0.9921
```

ttest *DURATION*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    4.398721    .2051143    4.442039    3.995662     4.80178
       1 |     132    5.545455    .4626163    5.315057    4.630289     6.46062
---------+--------------------------------------------------------------------
combined |     601    4.650582    .1903861    4.667373    4.276678    5.024487
---------+--------------------------------------------------------------------
    diff |           -1.146734    .4578639               -2.045948   -.2475201
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -2.5045
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0063         Pr(|T| > |t|) = 0.0125          Pr(T > t) = 0.9937
```

ttest *COMMERCE*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    41038.11    4002.054     86670.1    33173.89    48902.33
       1 |     132     61510.3    9904.006    113788.4    41917.82    81102.79
---------+--------------------------------------------------------------------
combined |     601     45534.5    3817.383    93584.29    38037.44    53031.55
---------+--------------------------------------------------------------------
    diff |           -20472.19    9190.468               -38521.65   -2422.738
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -2.2275
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0131         Pr(|T| > |t|) = 0.0263          Pr(T > t) = 0.9869
```

.

.
ttest *PRICE FIX NOT BID RIG*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .7377399    .0203327     .440333    .6977852    .7776945
       1 |     132    .6590909    .0414149    .4758206    .5771624    .7410194
---------+--------------------------------------------------------------------
combined |     601    .7204659     .018321    .4491441    .6844849    .7564469
---------+--------------------------------------------------------------------
    diff |             .078649    .0441739               -.0081056    .1654035
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   1.7804
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9622         Pr(|T| > |t|) = 0.0755          Pr(T > t) = 0.0378
```

ttest *BID RIG NOT PRICE FIX*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .1215352    .0151039    .3270971    .0918552    .1512151
       1 |     132    .2045455    .0352425    .4049057    .1348273    .2742636
---------+--------------------------------------------------------------------
combined |     601    .1397671    .0141558     .347034    .1119661     .167568
---------+--------------------------------------------------------------------
    diff |           -.0830103    .0340529               -.1498879   -.0161326
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -2.4377
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0075         Pr(|T| > |t|) = 0.0151          Pr(T > t) = 0.9925
```

ttest *ONLY HIGH RANK*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .7654584    .0195861    .4241644    .7269709     .803946
       1 |     132    .8939394    .0269027    .3090882    .8407195    .9471593
---------+--------------------------------------------------------------------
combined |     601    .7936772    .0165204     .405002    .7612324     .826122
---------+--------------------------------------------------------------------
    diff |            -.128481    .0395912               -.2062354   -.0507265
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -3.2452
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0006         Pr(|T| > |t|) = 0.0012          Pr(T > t) = 0.9994
```

ttest *ONLY LOW RANK*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .2089552    .0187934    .4069965    .1720254     .245885
       1 |     132    .0909091    .0251172     .288575    .0412212     .140597
---------+--------------------------------------------------------------------
combined |     601    .1830283    .0157865    .3870118    .1520247    .2140319
---------+--------------------------------------------------------------------
    diff |            .1180461    .0378577                .0436962     .192396
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   3.1182
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9990         Pr(|T| > |t|) = 0.0019          Pr(T > t) = 0.0010
```

ttest *ANY LOW RANK*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .2345416    .0195861    .4241644     .196054    .2730291
       1 |     132    .1060606    .0269027    .3090882    .0528407    .1592805
---------+--------------------------------------------------------------------
combined |     601    .2063228    .0165204     .405002     .173878    .2387676
---------+--------------------------------------------------------------------
    diff |             .128481    .0395912                .0507265    .2062354
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   3.2452
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9994         Pr(|T| > |t|) = 0.0012          Pr(T > t) = 0.0006
```

ttest *LANDMARK*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    2.879787    .0566061    1.225886    2.768553     2.99102
       1 |     132    2.774242    .1118214     1.28473    2.553033    2.995452
---------+--------------------------------------------------------------------
combined |     601    2.856606     .050529    1.238733    2.757371    2.955841
---------+--------------------------------------------------------------------
    diff |            .1055444    .1220767               -.1342061    .3452948
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   0.8646
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.8062         Pr(|T| > |t|) = 0.3876          Pr(T > t) = 0.1938

.
.
```

. ttest *DOJBUDGET*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    8030.802    194.6973    4216.444    7648.212    8413.391
       1 |     132    8525.326    386.9176    4445.345    7759.911    9290.742
---------+--------------------------------------------------------------------
combined |     601    8139.416    174.1327    4268.916    7797.432      8481.4
---------+--------------------------------------------------------------------
    diff |           -494.5249    420.4775               -1320.314    331.2645
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -1.1761
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.1200         Pr(|T| > |t|) = 0.2400          Pr(T > t) = 0.8800
```

```
. ttest ONLY CORP DEFEND, by(TRIAL CONVICTION)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .5906183    .0227298    .4922448    .5459533    .6352833
       1 |     132    .6287879    .0422112    .4849696    .5452841    .7122917
---------+--------------------------------------------------------------------
combined |     601    .5990017    .0200083    .4905089    .5597069    .6382964
---------+--------------------------------------------------------------------
    diff |           -.0381695    .0483445               -.1331149    .0567758
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -0.7895
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.2151         Pr(|T| > |t|) = 0.4301          Pr(T > t) = 0.7849


. ttest #OTHER INDICT, by(TRIAL CONVICTION)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .1791045    .0475827    1.030471    .0856022    .2726067
       1 |     132    .1136364    .0277287    .3185781    .0587824    .1684903
---------+--------------------------------------------------------------------
combined |     601    .1647255    .0376328    .9225799    .0908174    .2386335
---------+--------------------------------------------------------------------
    diff |            .0654681    .0909373               -.1131267    .2440629
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   0.7199
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.7641         Pr(|T| > |t|) = 0.4719          Pr(T > t) = 0.2359


. ttest #PREV CONVICT, by(TRIAL CONVICTION)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .1130064     .035268    .7637775    .0437032    .1823096
       1 |     132    .0606061    .0234577    .2695083    .0142012     .107011
---------+--------------------------------------------------------------------
combined |     601    .1014975    .0280049    .6865477     .046498     .156497
---------+--------------------------------------------------------------------
    diff |            .0524003    .0676674               -.0804939    .1852945
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   0.7744
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.7805         Pr(|T| > |t|) = 0.4390          Pr(T > t) = 0.2195

.
.
```

. ttest *COMPENSATION*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    14370.29    630.1731    13647.29    13131.97    15608.61
       1 |     132    15201.25    1268.169    14570.15    12692.51    17709.99
---------+--------------------------------------------------------------------
combined |     601     14552.8    564.8349     13847.1    13443.51    15662.09
---------+--------------------------------------------------------------------
    diff |           -830.959    1365.056               -3511.837    1849.919
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -0.6087
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.2715         Pr(|T| > |t|) = 0.5429          Pr(T > t) = 0.7285
```

. ttest *COMMERCE PER YEAR*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    14874.69     1824.53    39512.76     11289.4    18459.97
       1 |     132    14300.41      3262.4    37482.12    7846.605    20754.21
---------+--------------------------------------------------------------------
combined |     601    14748.56    1592.712    39045.82    11620.59    17876.53
---------+--------------------------------------------------------------------
    diff |           574.2793    3850.282               -6987.415    8135.973
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   0.1492
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.5593         Pr(|T| > |t|) = 0.8815          Pr(T > t) = 0.4407
```

ttest *COUNTS INDICT-CONVICT*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    1.085288    .0275495    .5966224    1.031152    1.139424
       1 |     132    1.030303    .0492124    .5654073    .9329492    1.127657
---------+--------------------------------------------------------------------
combined |     601    1.073211    .0240619    .5898852    1.025955    1.120467
---------+--------------------------------------------------------------------
    diff |           .0549848    .0581259               -.0591704    .1691401
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   0.9460
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.8277         Pr(|T| > |t|) = 0.3445          Pr(T > t) = 0.1723

.
.
.
```

. ttest *ANY BID RIG*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .2324094     .019524    .4228196    .1940438    .2707749
       1 |     132    .2272727    .0366143    .4206667    .1548408    .2997046
---------+--------------------------------------------------------------------
combined |     601    .2312812    .0172139     .422003    .1974745    .2650879
---------+--------------------------------------------------------------------
    diff |            .0051367    .0416137               -.0765898    .0868631
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   0.1234
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.5491         Pr(|T| > |t|) = 0.9018          Pr(T > t) = 0.4509
```

. ttest *COMPETITOR VIC*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .0405117    .0091135    .1973666    .0226032    .0584203
       1 |     132    .0454545    .0181992    .2090924    .0094523    .0814568
---------+--------------------------------------------------------------------
combined |     601    .0415973    .0081514    .1998335    .0255886     .057606
---------+--------------------------------------------------------------------
    diff |           -.0049428    .0197048               -.0436417     .033756
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -0.2508
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.4010         Pr(|T| > |t|) = 0.8020          Pr(T > t) = 0.5990
```

ttest *GOV VIC*, by(*TRIAL CONVICTION*)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .1876333    .0180471    .3908359    .1521698    .2230967
       1 |     132    .1439394    .0306695    .3523655    .0832678    .2046109
---------+--------------------------------------------------------------------
combined |     601    .1780366    .0156173    .3828622    .1473654    .2087078
---------+--------------------------------------------------------------------
    diff |            .0436939    .0377123               -.0303704    .1177582
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   1.1586
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.8765         Pr(|T| > |t|) = 0.2471          Pr(T > t) = 0.1235
```

```
. ttest CASE TIME SPAN, by(TRIAL CONVICTION)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    1.360341    .0548399    1.187634    1.252578    1.468104
       1 |     132    1.219697    .0957786    1.100412    1.030224     1.40917
---------+--------------------------------------------------------------------
combined |     601    1.329451    .0477087    1.169593    1.235755    1.423147
---------+--------------------------------------------------------------------
    diff |            .1406442    .1151917               -.0855845    .3668728
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   1.2210
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.8887         Pr(|T| > |t|) = 0.2226          Pr(T > t) = 0.1113
```

```
. ttest VAR DOJ TRIAL WIN PREV YEAR, by(TRIAL CONVICTION)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    15.33051    .4490106    9.723957    14.44818    16.21283
       1 |     132    16.40768    .7310855    8.399533    14.96141    17.85394
---------+--------------------------------------------------------------------
combined |     601    15.56709    .3855895    9.452843    14.80982    16.32436
---------+--------------------------------------------------------------------
    diff |           -1.077168    .9311162               -2.905817    .7514814
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -1.1569
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.1239         Pr(|T| > |t|) = 0.2478          Pr(T > t) = 0.8761
```

```
ttest POST-APPA INDICT YEAR, by(TRIAL CONVICTION)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     469    .2558635    .0201701    .4368114    .2162284    .2954987
       1 |     132    .2878788    .0395591    .4544992    .2096215    .3661361
---------+--------------------------------------------------------------------
combined |     601    .2628952    .0179713    .4405727    .2276008    .2981895
---------+--------------------------------------------------------------------
    diff |           -.0320152    .0434257               -.1173003    .0532699
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -0.7372
Ho: diff = 0                                     degrees of freedom =      599

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.2306         Pr(|T| > |t|) = 0.4613          Pr(T > t) = 0.7694
```

APPENDIX E

SAME-BRANCH MEANS TESTS THE CASE SELECTION MODEL USING SAMPLE OF INDICTED DEFENDANTS

This appendix reports the results of two-sample means tests that determine the relationships between variables representing the characteristics of initiated cases. All of the possible pairs of 14 variables allowed 91 separate means tests. Tables 33-36 report the means test results, by the pairs of variables used in the means tests. Each means test suggests two combinations (out of a possible four combinations) of the subject pair of variables in terms of their relative values. The means tests can suggest 182 different possible combinations of pairs of 14 variables in terms of the variables' relative values.

Statistical insignificance suggests that prosecutors either used a random process for selecting federal criminal antitrust cases to initiate from 1956 through 1979, or that the case selection process otherwise produced a random mix of case characteristics in initiated cases. In the tables, "0" denotes statistically insignificant results. Statistically significant results, denoted by "1," that are inconsistent with the case selection model's predictions suggest that prosecutors used a non-random case selection process that differs from the case selection model specification of that process. All other results will lend empirical support for the case selection model's predictive value.

The means tests reported in Tables 33-36 use the sample of defendants chosen for indictment (n = 3093), so the case selection model's predictions, which are summarized in Table 4, apply. Suppose a variable representing the prosecutor's unit cost of trial expenditure, *C,* is paired with a variable representing the defendant's unit cost of trial expenditure, *K*, in a means test. Then the means test will suggest either of the following combinations of C and *K*: i) (high value of *C,* high value of *K*) and (low value of *C,* low value of *K*); or ii) (high value of *C,* low value of *K*) and (low value of *C,* high value of *K*). Three combinations of pairs of the relative values of the prosecutor's unit cost of trial expenditure, *C,* and the defendant's, *K*, in the sample of initiated cases, are consistent with the case selection model's predictions: (high value of *C,* high value of *K*), (low value of *C,* low value of *K*) and (low value of *C,* high value of *K*). The combination of (high value of *C,* low value of *K*) in the sample of initiated cases is inconsistent with the case selection model's predictions.

In general, the combinations found in the sample of initiated cases that are consistent with the case selection model can be stated in terms of the prosecutor's value of bringing a case, *E(V)*: (low value of *E(V)*, high value of *E(V)*), (high value of *E(V)*, low value of *E(V)*), and (high value of *E(V)*, high value of *E(V)*). The combination (low value of *E(V)*, low value of *E(V)*) is inconsistent with the case selection model.

A two-sample means test will always suggest (high value of *C,* low value of *K*) or, more generally, (low value of *E(V)*, low value of *E(V)*) when it suggests (low value of *C,* high value of *K*) or, more generally, (high value of *E(V)*, high value of *E(V)*). Thus, in the sample of defendants in initiated cases, the percentage of combinations that are consistent with the case selection model's predictions cannot fall below 50 percent and cannot exceed 100 percent of the *statistically significant* combinations of pairs of variables that the means tests produce. The reported means tests results should be judged accordingly. Table 32 summarizes the results of the individual means tests covered by Tables 33-36, which follow.

Table 32 Results of means tests of implications of the case selection process, using the sample of indicted defendants (n = 3093)

Appendix E Table Reporting Results	Statistically Insignificant Results	Statistically Significant Results			
		Number of Combinations Consistent with Case Selection Model's Predictions			Number of Combinations Inconsistent with Case Selection Model's Predictions
Table 33	2	14	14	10	10
Table 34	16	13	13	17	17
Table 35	22	7	7	7	7
Table 36	0	0	0	3	3
GRAND TOTALS	40	34	34	37	37
	40	105			37

Table 33 Descriptions of results of means tests that include empirical variables representing the prosecutor's unit (opportunity) cost of trial expenditure, *C*, and other variables from the sample of indicted defendants (n = 3093)

Variables Used in Two-Sample Means Tests	Evidence of Expected Case Selection "Adjustments" Predicted by the Model of the Prosecutor's Case Selection Process			Not Expected According to Case Selection Model
	High *C*, $\Delta E(V) > 0$	Low *C*, $\Delta E(V) < 0$	Low *C*, $\Delta E(V) > 0$	High *C*, $\Delta E(V) < 0$
ttest *HIGH ANNUAL FORTUNE500*, by(*HIGH DOJBUDGET*)			1	1
ttest *ONLY LOW RANK*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *HIGH PROFIT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *HIGH COMPENSATION*, by(*HIGH DOJBUDGET*)	1	1		
ttest *ZERO OTHER INDICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *ZERO PREV CONVICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *HIGH DURATION*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH COMMERCE*, by(*HIGH DOJBUDGET*)			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *MULTI COUNTS INDICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH DOJBUDGET*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DOJBUDGET*)			1	1
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH DOJBUDGET*)			1	1
SUB-TOTALS	6	6	6	6

Table 33, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Case Selection "Adjustments" Predicted by the Model of the Prosecutor's Case Selection Process			Not Expected According to Case Selection Model
	High *C*, $\Delta E(V) > 0$	Low *C*, $\Delta E(V) < 0$	Low *C*, $\Delta E(V) > 0$	High *C*, $\Delta E(V) < 0$
ttest *ONLY LOW RANK*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *HIGH PROFIT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *HIGH COMPENSATION*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *ZERO OTHER INDICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *ZERO PREV CONVICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *HIGH DURATION*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *HIGH COMMERCE*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *MULTI COUNTS CONVICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *MULTI COUNTS INDICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH ANNUAL FORTUNE500*)			1	1
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH ANNUAL FORTUNE500*)			1	1
SUB-TOTALS	8	8	4	4
TOTALS	14	14	10	10

Note: The variables *HIGH DOJBUDGET* and *HIGH ANNUAL FORTUNE500*, which represent the prosecutor's unit cost of trial expenditure, *C*, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results.

Table 34 Descriptions of results of means tests that include empirical variables representing the defendant's unit (opportunity) cost of trial expenditure, *K*, and other variables from the sample of indicted defendants (n = 3093)

Variables Used in Two-Sample Means Tests	Evidence of Expected Case Selection "Adjustments" Predicted by the Model of the Prosecutor's Case Selection Process			Not Expected According to Case Selection Model
	Low *K*, $\Delta E(V) > 0$	High *K*, $\Delta E(V) < 0$	High *K*, $\Delta E(V) > 0$	Low *K*, $\Delta E(V) < 0$
ttest *HIGH PROFIT*, by(*ONLY LOW RANK*)			1	1 (p<0.10)
ttest *HIGH COMPENSATION*, by(*ONLY LOW RANK*)			1	1
ttest *ZERO OTHER INDICT*, by(*ONLY LOW RANK*)	1	1		
ttest *ZERO PREV CONVICT*, by(*ONLY LOW RANK*)	1	1		
ttest *HIGH DURATION*, by(*ONLY LOW RANK*)			1	1
ttest *HIGH COMMERCE*, by(*ONLY LOW RANK*)			1	1
ttest *MULTI COUNTS CONVICT*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ONLY LOW RANK*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *BID RIG NOT PRICE FIX*, by(*ONLY LOW RANK*)	1	1		
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*ONLY LOW RANK*)	1	1		
SUB-TOTALS	4	4	5	5

Table 34, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Case Selection "Adjustments" Predicted by the Model of the Prosecutor's Case Selection Process			Not Expected According to Case Selection Model
	Low *K*, $\Delta E(V) > 0$	High *K*, $\Delta E(V) < 0$	High *K*, $\Delta E(V) > 0$	Low *K*, $\Delta E(V) < 0$
ttest *HIGH COMPENSATION*, by(*HIGH PROFIT*)			1	1
ttest *ZERO OTHER INDICT*, by(*HIGH PROFIT*)			1	1
ttest *ZERO PREV CONVICT*, by(*HIGH PROFIT*)	0	0	0	0
ttest *HIGH DURATION*, by(*HIGH PROFIT*)	1	1		
ttest *HIGH COMMERCE*, by(*HIGH PROFIT*)			1 (p<0.15)	1 (p<0.15)
ttest *MULTI COUNTS CONVICT*, by(*HIGH PROFIT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH PROFIT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*HIGH PROFIT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*HIGH PROFIT*)	0	0	0	0
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH PROFIT*)	1	1		
SUB-TOTALS	3	3	4	4

Table 34, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Case Selection "Adjustments" Predicted by the Model of the Prosecutor's Case Selection Process			Not Expected According to Case Selection Model
	Low *K*, $\Delta E(V) > 0$	High *K*, $\Delta E(V) < 0$	High *K*, $\Delta E(V) > 0$	Low *K*, $\Delta E(V) < 0$
ttest *ZERO OTHER INDICT*, by(*HIGH COMPENSATION*)			1	1
ttest *ZERO PREV CONVICT*, by(*HIGH COMPENSATION*)			1	1
ttest *HIGH DURATION*, by(*HIGH COMPENSATION*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*HIGH COMPENSATION*)			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH COMPENSATION*)			1	1
ttest *MULTI COUNTS INDICT*, by(*HIGH COMPENSATION*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*HIGH COMPENSATION*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*HIGH COMPENSATION*)	1	1		
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH COMPENSATION*)	1	1		
SUB-TOTALS	3	3	5	5

Table 34, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Case Selection "Adjustments" Predicted by the Model of the Prosecutor's Case Selection Process			Not Expected According to Case Selection Model
	Low *K* or e, $\Delta E(V) > 0$	High *K* or e, $\Delta E(V) < 0$	High *K* or e, $\Delta E(V) > 0$	Low *K* or e, $\Delta E(V) < 0$
ttest *ZERO PREV CONVICT*, by(*ZERO OTHER INDICT*)			1	1
ttest *HIGH DURATION*, by(*ZERO OTHER INDICT*)	1	1		
Ttest *HIGH COMMERCE*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*ZERO OTHER INDICT*)			1 (p<0.10)	1
ttest *MULTI COUNTS INDICT*, by(*ZERO OTHER INDICT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*ZERO OTHER INDICT*)	1	1		
Ttest *BID RIG NOT PRICE FIX*, by(*ZERO OTHER INDICT*)	0	0	0	0
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*ZERO OTHER INDICT*)	1	1		
SUB-TOTALS	3	3	3	3
TOTALS	13	13	17	17

Note: The variables *ONLY LOW RANK*, *HIGH PROFIT*, *HIGH COMPENSATION*, and *ZERO OTHER INDICT*, which represent the defendant's unit cost of trial expenditure, *K*, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results.

Table 35 Descriptions of results of means tests that include empirical variables representing the amount or strength of the prosecutor's evidence, **e**, and other variables from the sample of indicted defendants (n = 3093)

Variables Used in Two-Sample Means Tests	Evidence of Expected Case Selection "Adjustments" Predicted by the Model of the Prosecutor's Case Selection Process			Not Expected According to Case Selection Model
	Low e, $\Delta E(V) > 0$	High e, $\Delta E(V) < 0$	High e, $\Delta E(V) > 0$	Low e, $\Delta E(V) < 0$
ttest *HIGH DURATION*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *POST-APPA INDICT YEAR*, by(*ZERO PREV CONVICT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*ZERO PREV CONVICT*)	0	0	0	0
SUB-TOTALS	1	1	0	0

Table 35, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Case Selection "Adjustments" Predicted by the Model of the Prosecutor's Case Selection Process			Not Expected According to Case Selection Model
	Low e, $\Delta E(V) > 0$	High e, $\Delta E(V) < 0$	High e, $\Delta E(V) > 0$	Low e, $\Delta E(V) < 0$
ttest *HIGH COMMERCE*, by(*HIGH DURATION*)			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH DURATION*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH DURATION*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*HIGH DURATION*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DURATION*)			1	1
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH DURATION*)			1	1
SUB-TOTALS	0	0	5	5
ttest *MULTI COUNTS CONVICT*, by(*HIGH COMMERCE*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH COMMERCE*)	0	0	0	0
ttest *POST-APPA INDICT YEAR*, by(*HIGH COMMERCE*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH COMMERCE*)	1	1		
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH COMMERCE*)	1	1		
SUB-TOTALS	2	2	1	1

Table 35, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Case Selection "Adjustments" Predicted by the Model of the Prosecutor's Case Selection Process			Not Expected According to Case Selection Model
	Low e, $\Delta E(V) > 0$	High e, $\Delta E(V) < 0$	High e, $\Delta E(V) > 0$	Low e, $\Delta E(V) < 0$
ttest *MULTI COUNTS INDICT*, by(*MULTI COUNTS CONVICT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*MULTI COUNTS CONVICT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*MULTI COUNTS CONVICT*)	1	1		
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*MULTI COUNTS CONVICT*)	0	0	0	0
SUB-TOTALS	2	2	1	1
ttest *POST-APPA INDICT YEAR*, by(*MULTI COUNTS INDICT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*MULTI COUNTS INDICT*)	0	0	0	0
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*MULTI COUNTS INDICT*)	1	1		
SUB-TOTALS	2	2	0	0
TOTALS	7	7	7	7

Note: The variables *ZERO PREV CONVICT*, *HIGH DURATION*, *HIGH COMMERCE*, *MULTI COUNTS CONVICT*, and *MULTI COUNTS INDICT*, which represent the level of the prosecutor's evidence, **e**, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results.

Table 36 Descriptions of results of means tests that include empirical variables representing the litigants' *ex ante* expected trial sentence, S^T, the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, and other variables from the sample of indicted defendants (n = 3093)

Variables Used in Two-Sample Means Tests	Evidence of Expected Case Selection "Adjustments" Predicted by the Model of the Prosecutor's Case Selection Process			Not Expected According to Case Selection Model
	Low S^T, $\Delta E(V) > 0$	High S^T, $\Delta E(V) < 0$	High S^T, $\Delta E(V) > 0$	Low S^T, $\Delta E(V) < 0$
ttest *BID RIG NOT PRICE FIX*, by(*POST-APPA INDICT YEAR*)			1	1
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*POST-APPA INDICT YEAR*)			1	1
SUB-TOTALS	0	0	2	2
	Low ρ, $\Delta E(V) > 0$	High ρ, $\Delta E(V) < 0$	High ρ, $\Delta E(V) > 0$	Low ρ, $\Delta E(V) < 0$
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*BID RIG NOT PRICE FIX*)			1	1
SUB-TOTALS	0	0	1	1
TOTALS	0	0	3	3

Note: The variable *POST-APPA INDICT YEAR*, which represents the litigants' *ex ante* expected trial sentence, S^T, and *BID RIG NOT PRICE FIX*, which represents the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results.

APPENDIX F

SAME-BRANCH MEANS TESTS OF THE TRIAL SELECTION MODEL USING SAMPLE OF DEFENDANTS THAT GO TO TRIAL

This appendix reports the results of two-sample means tests that determine the relationships between variables representing the characteristics of cases that went to trial. All of the possible pairs of 14 variables allowed 91 separate means tests. Tables 38-41 report the means test results, by the pairs of variables used in the means tests. Each means test suggests two combinations (out of a possible four combinations) of the subject pair of variables in terms of their relative values. The means tests can suggest 182 different possible combinations of pairs of 14 variables in terms of the variables' relative values.

Statistical insignificance suggests that defendants either used a random process for deciding to go to trial in federal criminal antitrust cases to initiate from 1956 through 1979, or used a trial selection process that otherwise produced a random mix of case characteristics in cases that went to trial. "0" denotes statistically insignificant results. Statistically significant results, denoted by "1," that are inconsistent with the trial selection model's predictions suggest that defendants in the observed cases used a non-random plea decision process that differs from the trial selection model specification of that decision. All other results will lend empirical support for the trial selection model's predictive value.

The means tests reported in Tables 38-41 use the sample of defendants who go to trial (n = 601), so the trial selection model's predictions, which are summarized in Table 6, apply. Suppose a variable representing the prosecutor's unit cost of trial expenditure, *C,* is paired with a variable representing the defendant's unit cost of trial expenditure, *K,* in a means test. Then the means test will suggest either of the following combinations of C and *K*: i) (high value of *C,* high value of *K*) and (low value of *C,* low value of *K*); or ii) (high value of *C,* low value of *K*) and (low value of *C,* high value of K).

Three combinations of pairs of the relative values of the prosecutor's unit cost of trial expenditure, *C,* and the defendant's, *K,* in the sample of litigated cases, are consistent with the trial selection model's predictions: (low value of *C,* low value of *K*), (high value of *C,* low value of *K*), and (low value of *C,* high value of *K*). The combination of (high value of *C,* high value of *K*) is inconsistent with the trial selection model's predictions.

In general, the combinations found in the sample of defendants who choose to go to trial that are consistent with the trial selection model can be stated in terms of the relative value of the ex post probability of trial, $\tilde{\Theta}$. The combinations consistent with the trial selection model's predictions are:: (low value of $\tilde{\Theta}$, high value of $\tilde{\Theta}$), (high value of $\tilde{\Theta}$, low value of $\tilde{\Theta}$), and (high value of $\tilde{\Theta}$, high value of $\tilde{\Theta}$). The combination (low value of $\tilde{\Theta}$, low value of $\tilde{\Theta}$) is inconsistent with the trial selection model.

A means test always will suggest (high value of *C,* high value of K) or, more generally, (low value of $\tilde{\Theta}$, low value of $\tilde{\Theta}$) when it suggests (low value of *C,* low value of K) or, more generally, (high value of $\tilde{\Theta}$, high value of $\tilde{\Theta}$). Thus, in the sample of defendants *who choose to go to trial*, the percentage of statistically significant combinations consistent with the trial selection model cannot fall below 50 percent and cannot exceed 100 percent of the *statistically signficant* combinations of pairs of variables that the means tests produce. The results should be judged accordingly. Table 37 summarizes the results of the individual means tests covered by Tables 38-41, which follow.

Table 37 Results of means tests of implications of the trial selection process, using the sample of defendants that went to trial (n = 601)

Appendix F Table Reporting Results	Statistically Insignificant Results	Statistically Significant Results			
		Number of Combinations Consistent with Trial Selection Model's Predictions			Number of Combinations Inconsistent with Trial Selection Model's Predictions
Table 38	10	7	7	13	13
Table 39	30	8	8	15	15
Table 40	18	7	7	9	9
Table 41	0	0	0	3	3
GRAND TOTALS	58	22	22	40	40
	58	84			40

Table 38 Descriptions of results of means tests that include empirical variables representing the prosecutor's unit (opportunity) cost of trial expenditure, *C*, and other variables from the sample of defendants that went to trial (n = 601)

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model			Not Expected According to Trial Selection Model
	High *C*, $\Delta \tilde{\Theta} > 0$	Low *C*, $\Delta \tilde{\Theta} < 0$	Low *C*, $\Delta \tilde{\Theta} > 0$	High *C*, $\Delta \tilde{\Theta} < 0$
ttest *HIGH ANNUAL FORTUNE500*, by(*HIGH DOJBUDGET*)			1	1
ttest *ONLY LOW RANK*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *HIGH PROFIT*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH COMPENSATION*, by(*HIGH DOJBUDGET*)			1	1
ttest *ZERO OTHER INDICT*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *ZERO PREV CONVICT*, by(*HIGH DOJBUDGET*)*			1 (p<0.15)	1 (p<0.15)
ttest *HIGH DURATION*, by(*HIGH DOJBUDGET*)*			1	1
ttest *HIGH COMMERCE*, by(*HIGH DOJBUDGET*)*			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH DOJBUDGET*)*	1	1		
ttest *MULTI COUNTS INDICT*, by(*HIGH DOJBUDGET*)*	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH DOJBUDGET*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DOJBUDGET*)			1	1
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH DOJBUDGET*)			1	1
SUB-TOTALS	2	2	9	9

*

Table 38, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model			Not Expected According to Trial Selection Model
	High *C*, $\Delta\,\tilde{\Theta} > 0$	Low *C*, $\Delta\,\tilde{\Theta} < 0$	Low *C*, $\Delta\,\tilde{\Theta} > 0$	High *C*, $\Delta\,\tilde{\Theta} < 0$
ttest *ONLY LOW RANK*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *HIGH PROFIT*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *HIGH COMPENSATION*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *ZERO OTHER INDICT*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *ZERO PREV CONVICT*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *HIGH DURATION*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *HIGH COMMERCE*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *MULTI COUNTS CONVICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *MULTI COUNTS INDICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH ANNUAL FORTUNE500*)			1	1
SUB-TOTALS	5	5	4	4
TOTALS	7	7	13	13

Note: The variables *HIGH DOJBUDGET* and *HIGH ANNUAL FORTUNE500*, which represent the prosecutor's unit cost of trial expenditure, *C*, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. These results assume that, *ceteris paribus*, more or better evidence, **e,** increases ρ and/or S^T and $\tilde{\Theta}$.

Table 39 Descriptions of results of means tests that include empirical variables representing the defendant's unit (opportunity) cost of trial expenditure, *K*, and other variables from the sample of defendants that went to trial (n = 601)

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model			Not Expected According to Trial Selection Model
	High *K*, $\Delta \tilde{\Theta} > 0$	Low *K*, $\Delta \tilde{\Theta} < 0$	Low *K*, $\Delta \tilde{\Theta} > 0$	High *K*, $\Delta \tilde{\Theta} < 0$
ttest *HIGH PROFIT*, by(*ONLY LOW RANK*)			1	1
ttest *HIGH COMPENSATION*, by(*ONLY LOW RANK*)			1	1
ttest *ZERO OTHER INDICT*, by(*ONLY LOW RANK*)			1 (p<0.10)	1 (p<0.10)
ttest *ZERO PREV CONVICT*, by(*ONLY LOW RANK*)			1 (p<0.10)	1 (p<0.10)
ttest *HIGH DURATION*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ONLY LOW RANK*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *BID RIG NOT PRICE FIX*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*ONLY LOW RANK*)	0	0	0	0
SUB-TOTALS	1	1	4	4

Table 39, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model			Not Expected According to Trial Selection Model
	High *K*, $\Delta\tilde{\Theta} > 0$	Low *K*, $\Delta\tilde{\Theta} < 0$	Low *K*, $\Delta\tilde{\Theta} > 0$	High *K*, $\Delta\tilde{\Theta} < 0$
ttest *HIGH COMPENSATION*, by(*HIGH PROFIT*)			1	1
ttest *ZERO OTHER INDICT*, by(*HIGH PROFIT*)			1	1
ttest *ZERO PREV CONVICT*, by(*HIGH PROFIT*)	0	0	0	0
ttest *HIGH DURATION*, by(*HIGH PROFIT*)			1	1
ttest *HIGH COMMERCE*, by(*HIGH PROFIT*			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH PROFIT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH PROFIT*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH PROFIT*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH PROFIT*)			1	1
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH PROFIT*)			1	1
SUB-TOTALS	1	1	7	7

Table 39, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Selection “Adjustments” Predicted by Trial Selection Model			Not Expected According to Trial Selection Model
	High *K*, $\Delta\,\tilde{\Theta} > 0$	Low *K*, $\Delta\,\tilde{\Theta} < 0$	Low *K*, $\Delta\,\tilde{\Theta} > 0$	High *K*, $\Delta\,\tilde{\Theta} < 0$
ttest *ZERO OTHER INDICT*, by(*HIGH COMPENSATION*)	1	1		
ttest *ZERO PREV CONVICT*, by(*HIGH COMPENSATION*)	0	0	0	0
ttest *HIGH DURATION*, by(*HIGH COMPENSATION*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*HIGH COMPENSATION*)			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH COMPENSATION*)	1	1		
ttest *MULTI COUNTS INDICT*, by(*HIGH COMPENSATION*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH COMPENSATION*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH COMPENSATION*)			1	1
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH COMPENSATION*)	0	0	0	0
SUB-TOTALS	3	3	3	3

Table 39, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model			Not Expected According to Trial Selection Model
	Low e or *K*, $\Delta\,\tilde{\Theta} > 0$*	High e or *K*, $\Delta\,\tilde{\Theta} < 0$*	High e or *K*, $\Delta\,\tilde{\Theta} > 0$*	Low e or *K*, $\Delta\,\tilde{\Theta} < 0$*
ttest *ZERO PREV CONVICT*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *HIGH DURATION*, by(*HIGH* by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*ZERO OTHER INDICT*)	1 (p<0.10)	1 (p<0.10)		
ttest *MULTI COUNTS CONVICT*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ZERO OTHER INDICT*)			1 (p<0.10)	1 (p<0.10)
ttest *POST-APPA INDICT YEAR*, by(*ZERO OTHER INDICT*)	1 (p<0.10)	1 (p<0.10)		
ttest *BID RIG NOT PRICE FIX*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*ZERO OTHER INDICT*)	1	1		
SUB-TOTALS	3	3	1	1
TOTALS	8	8	15	15

Note: The variables *ONLY LOW RANK*, *HIGH PROFIT*, *HIGH COMPENSATION*, and *ZERO OTHER INDICT*, which represent the defendant's unit cost of trial expenditure, *K*, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. These results assume that, *ceteris paribus*, more or better evidence, **e,** increases ρ and/or S^T and $\tilde{\Theta}$.

Table 40 Descriptions of results of means tests that include empirical variables representing the amount or strength of the prosecutor's evidence, **e**, and other variables from the sample of defendants that went to trial (n = 601)

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model			Not Expected According to Trial Selection Model
	Low e, $\Delta \tilde{\Theta} > 0$*	High e, $\Delta \tilde{\Theta} < 0$*	High e, $\Delta \tilde{\Theta} > 0$*	Low e, $\Delta \tilde{\Theta} < 0$*
ttest *HIGH DURATION*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ZERO PREV CONVICT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *BID RIG NOT PRICE FIX*, by(*ZERO PREV CONVICT*)	0	0	0	0
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*ZERO PREV CONVICT*)	1	1		
SUB-TOTALS	1	1	1	1

Table 40, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model			Not Expected According to Trial Selection Model
	Low e, $\Delta\,\tilde{\Theta} > 0$*	High e, $\Delta\,\tilde{\Theta} < 0$*	High e, $\Delta\,\tilde{\Theta} > 0$*	Low e, $\Delta\,\tilde{\Theta} < 0$*
ttest *HIGH COMMERCE*, by(*HIGH DURATION*)			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH DURATION*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH DURATION*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*HIGH DURATION*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DURATION*)			1	1
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH DURATION*)			1	1
SUB-TOTALS	0	0	5	5
ttest *MULTI COUNTS CONVICT*, by(*HIGH COMMERCE*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH COMMERCE*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*HIGH COMMERCE*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH COMMERCE*)	1	1		
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH COMMERCE*)	1	1		
SUB-TOTALS	2	2	2	2

Table 40, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model			Not Expected According to Trial Selection Model
	Low e, $\Delta\, \tilde{\Theta} > 0$*	High e, $\Delta\, \tilde{\Theta} < 0$*	High e, $\Delta\, \tilde{\Theta} > 0$*	Low e, $\Delta\, \tilde{\Theta} < 0$*
ttest *MULTI COUNTS INDICT*, by(*MULTI COUNTS CONVICT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*MULTI COUNTS CONVICT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*MULTI COUNTS CONVICT*)	1 ($p<0.10$)	1 ($p<0.10$)		
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*MULTI COUNTS CONVICT*)	0	0	0	0
SUB-TOTALS	2	2	1	1
ttest *POST-APPA INDICT YEAR*, by(*MULTI COUNTS INDICT*)	1 ($p<0.10$)	1 ($p<0.10$)		
ttest *BID RIG NOT PRICE FIX*, by(*MULTI COUNTS INDICT*)	1 ($p<0.12$)	1 ($p<0.12$)		
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*MULTI COUNTS INDICT*)	0	0	0	0
SUB-TOTALS	2	2	0	0
TOTALS	7	7	9	9

Note: The variables *ZERO PREV CONVICT*, *HIGH DURATION*, *HIGH COMMERCE*, *MULTI COUNTS CONVICT*, and *MULTI COUNTS INDICT*, which represent the level of the prosecutor's evidence, **e,** pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. Assume that, *ceteris paribus*, more or better evidence, **e,** increases ρ and/or S^T and $\tilde{\Theta}$.

Table 41 Descriptions of results of means tests that include empirical variables representing the litigants' *ex ante* expected trial sentence, S^T, the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, and other variables from the sample of defendants that went to trial (n = 601)

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model			Not Expected According to Trial Selection Model
	Low S^T, $\Delta\,\tilde{\Theta} > 0$	High S^T, $\Delta\,\tilde{\Theta} < 0$	High S^T, $\Delta\,\tilde{\Theta} > 0$	Low S^T, $\Delta\,\tilde{\Theta} < 0$
ttest *BID RIG NOT PRICE FIX*, by(*POST-APPA INDICT YEAR*)			1	1
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*POST-APPA INDICT YEAR*)			1	1
SUB-TOTALS	0	0	2	2
	Low ρ, $\Delta\,\tilde{\Theta} > 0$	High ρ, $\Delta\,\tilde{\Theta} < 0$	High ρ, $\Delta\,\tilde{\Theta} > 0$	Low ρ, $\Delta\,\tilde{\Theta} < 0$
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*BID RIG NOT PRICE FIX*)			1	1
SUB-TOTALS	0	0	1	1
TOTALS	0	0	3	3

Note: The variable *POST-APPA INDICT YEAR*, which represents the litigants' *ex ante* expected trial sentence, S^T, and *BID RIG NOT PRICE FIX*, which represents the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. pairs with other variables in the tests reported here.

APPENDIX G

SAME-BRANCH MEANS TESTS OF THE TRIAL SELECTION MODEL USING SAMPLE OF DEFENDANTS THAT AVOID TRIAL

Tables 43-46 report the results of means tests that use the sample of defendants who choose *not* to go to trial (n = 2492). The means tests reported here use 14 variables are are similar to those reported in Appendix F. The trial selection model's predictions, which Table 6 summarizes, apply with one caveat: combinations labeled "inconsistent" in Table 6 are labeled "consistent" here. For example, in terms of the prosecutor's unit cost of trial expenditure, *C,* and the defendant's, *K,* the combination (high value of *C,* high value of K) is the only combination of C and *K,* in terms of the relative values of the variables, that is consistent with the trial selection model's predictions. Meanwhile, the combinations (low value of *C,* low value of K), (high value of *C,* low value of K), and (low value of *C,* high value of K) are inconsistent with the trial selection model's predictions.

In general, the combination found in the sample of defendants who *do not* choose to go to trial that is consistent with the trial selection model's predictions can be stated in terms of the relative value of the ex post probability of trial, $\tilde{\Theta}$, as follows: (low value of $\tilde{\Theta}$, low value of $\tilde{\Theta}$). The combinations (low value of $\tilde{\Theta}$, high value of $\tilde{\Theta}$), (high value of $\tilde{\Theta}$, low value of $\tilde{\Theta}$), and (high value of $\tilde{\Theta}$, high value of $\tilde{\Theta}$) are inconsistent with the trial selection model's predictions.

A means test always will suggest (low vaue of *C,* low value of K) or, more generally, (high value of $\tilde{\Theta}$, high value of $\tilde{\Theta}$) when it suggests (high value of *C,* high value of K) or, more generally, (low value of $\tilde{\Theta}$, low value of $\tilde{\Theta}$). Thus, using the sample of defendants who choose *not* to go to trial, the percentage of statistically significant combinations consistent with the trial selection model cannot fall below zero percent and cannot exceed 50 percent of the *statistically significant* combinations of pairs of variables that the means tests produce. The reported results should be judged accordingly. Table 42 summarizes the results of the individual means tests covered by Tables 43-46, which follow.

Table 42 Results of means tests of implications of the trial selection process, using the sample of defendants that did not go to trial (n = 2492)

Appendix G Table Reporting Results	Statistically Insignificant Results	Statistically Significant Results			
		Number of Combinations Inconsistent with Trial Selection Model's Predictions			Number of Combinations Consistent with Trial Selection Model's Predictions
Table 43	4	7	7	16	16
Table 44	12	12	12	19	19
Table 45	18	9	9	7	7
Table 46	0	0	0	3	3
GRAND TOTALS	36	28	28	45	45
	36	101			45

Table 43 Descriptions of results of means tests that include empirical variables representing the prosecutor's unit (opportunity) cost of trial expenditure, *C*, and other variables from the sample of defendants that did not go to trial (n = 2492)

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Selection Model			Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model
	Low *C*, $\Delta\tilde{\Theta} < 0$	High *C*, $\Delta\tilde{\Theta} > 0$	High *C*, $\Delta\tilde{\Theta} < 0$	Low *C*, $\Delta\tilde{\Theta} > 0$
ttest *HIGH ANNUAL FORTUNE500*, by(*HIGH DOJBUDGET*)			1	1
ttest *ONLY LOW RANK*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *HIGH PROFIT*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH COMPENSATION*, by(*HIGH DOJBUDGET*)			1	1
ttest *ZERO OTHER INDICT*, by(*HIGH DOJBUDGET*)			1	1
ttest *ZERO PREV CONVICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *HIGH DURATION*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH COMMERCE*, by(*HIGH DOJBUDGET*)			1 ($p<0.12$)	1 ($p<0.12$)
ttest *MULTI COUNTS CONVICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *MULTI COUNTS INDICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH DOJBUDGET*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH DOJBUDGET*)			1	1
SUB-TOTALS	3	3	9	9

Table 43, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Selection Model			Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model
	Low *C*, $\Delta\,\widetilde{\Theta} < 0$	High *C*, $\Delta\,\widetilde{\Theta} > 0$	High *C*, $\Delta\,\widetilde{\Theta} < 0$	Low *C*, $\Delta\,\widetilde{\Theta} > 0$
ttest *ONLY LOW RANK*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *HIGH PROFIT*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *HIGH COMPENSATION*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *ZERO OTHER INDICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *ZERO PREV CONVICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *HIGH DURATION*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *HIGH COMMERCE*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *MULTI COUNTS CONVICT*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH ANNUAL FORTUNE500*)			1	1
SUB-TOTALS	4	4	7	7
TOTALS				

Note: The variables *HIGH DOJBUDGET* and *HIGH ANNUAL FORTUNE500*, which represent the prosecutor's unit cost of trial expenditure, *C*, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. Assume that, *ceteris paribus*, more or better evidence, **e,** increases ρ and/or S^T and $\widetilde{\Theta}$.

Table 44 Descriptions of results of means tests that include empirical variables representing the defendant's unit (opportunity) cost of trial expenditure, *K*, and other variables from the sample of defendants that did not go to trial (n = 2492)

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Selection Model			Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model
	Low *K*, $\Delta \tilde{\Theta} < 0$	High *K*, $\Delta \tilde{\Theta} > 0$	High *K*, $\Delta \tilde{\Theta} < 0$	Low *K*, $\Delta \tilde{\Theta} > 0$
ttest *HIGH PROFIT*, by(*ONLY LOW RANK*)			1	1
ttest *HIGH COMPENSATION*, by(*ONLY LOW RANK*)			1	1
ttest *ZERO OTHER INDICT*, by(*ONLY LOW RANK*)			1	1
ttest *ZERO PREV CONVICT*, by(*ONLY LOW RANK*)			1	1
ttest *HIGH DURATION*, by(*ONLY LOW RANK*)	1	1		
ttest *HIGH COMMERCE*, by(*ONLY LOW RANK*)	1	1		
ttest *MULTI COUNTS CONVICT*, by(*ONLY LOW RANK*)				1
ttest *MULTI COUNTS INDICT*, by(*ONLY LOW RANK*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *BID RIG NOT PRICE FIX*, by(*ONLY LOW RANK*)			1	1
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*ONLY LOW RANK*)			1	1
SUB-TOTALS	3	3	6	6

Table 44, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Selection Model			Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model
	Low *K*, $\Delta\tilde{\Theta} < 0$	High *K*, $\Delta\tilde{\Theta} > 0$	High *K*, $\Delta\tilde{\Theta} < 0$	Low *K*, $\Delta\tilde{\Theta} > 0$
ttest *HIGH COMPENSATION*, by(*HIGH PROFIT*)			1	1
ttest *ZERO OTHER INDICT*, by(*HIGH PROFIT*)			1	1
ttest *ZERO PREV CONVICT*, by(*HIGH PROFIT*)	0	0	0	0
ttest *HIGH DURATION*, by(*HIGH PROFIT*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*HIGH PROFIT*)	1	1		
ttest *MULTI COUNTS CONVICT*, by(*HIGH PROFIT*)			1	1
ttest *MULTI COUNTS INDICT*, by(*HIGH PROFIT*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH PROFIT*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH PROFIT*)	0	0	0	0
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH PROFIT*)			1	1
SUB-TOTALS	2	2	5	5

Table 44, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Selection Model			Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model
	Low *K*, $\Delta \tilde{\Theta} < 0$	High *K*, $\Delta \tilde{\Theta} > 0$	High *K*, $\Delta \tilde{\Theta} < 0$	Low *K*, $\Delta \tilde{\Theta} > 0$
ttest *ZERO OTHER INDICT*, by(*HIGH COMPENSATION*)			1	1
ttest *ZERO PREV CONVICT*, by(*HIGH COMPENSATION*)	1	1		
ttest *HIGH DURATION*, by(*HIGH COMPENSATION*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*HIGH COMPENSATION*)	1	1		
ttest *MULTI COUNTS CONVICT*, by(*HIGH COMPENSATION*)	1	1		
ttest *MULTI COUNTS INDICT*, by(*HIGH COMPENSATION*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH COMPENSATION*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH COMPENSATION*)			1	1
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH COMPENSATION*)			1	1
SUB-TOTALS	4	4	4	4

Table 44, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Selection Model			Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model
	High *K* or e, $\Delta\,\tilde{\Theta} < 0$	Low *K* or e, $\Delta\,\tilde{\Theta} > 0$	Low *K* or e, $\Delta\,\tilde{\Theta} < 0$	High *K* or e, $\Delta\,\tilde{\Theta} > 0$
ttest *ZERO PREV CONVICT*, by(*ZERO OTHER INDICT*)			1	1
ttest *HIGH DURATION*, by(*ZERO OTHER INDICT*)	1	1		
ttest *HIGH COMMERCE*, by(*ZERO OTHER INDICT*)			1 ($p<0.10$)	1 ($p<0.10$)
ttest *MULTI COUNTS CONVICT*, by(*ZERO OTHER INDICT*)			1	1
ttest *MULTI COUNTS INDICT*, by(*ZERO OTHER INDICT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*ZERO OTHER INDICT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*ZERO OTHER INDICT*)	0	0	0	0
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*ZERO OTHER INDICT*)	1	1		
SUB-TOTALS	3	3	4	4
TOTALS				

Note: The variables *ONLY LOW RANK*, *HIGH PROFIT*, *HIGH COMPENSATION*, and *ZERO OTHER INDICT*, which represent the defendant's unit cost of trial expenditure, *K*, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. Assume that, *ceteris paribus*, more or better evidence, **e,** increases ρ and/or S^T and $\tilde{\Theta}$.

Table 45 Descriptions of results of means tests that include empirical variables representing the amount or strength of the prosecutor's evidence, **e**, and other variables from the sample of defendants that did not go to trial (n = 2492)

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Selection Model			Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model
	High e, $\Delta\tilde{\Theta} < 0$	Low e, $\Delta\tilde{\Theta} > 0$	Low e, $\Delta\tilde{\Theta} < 0$	High e, $\Delta\tilde{\Theta} > 0$
ttest *HIGH DURATION*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *POST-APPA INDICT YEAR*, by(*ZERO PREV CONVICT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*ZERO PREV CONVICT*)	1	1		
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*ZERO PREV CONVICT*)	0	0	0	0
SUB-TOTALS	2	2	0	0

Table 45, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Selection Model			Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model
	High e, $\Delta \tilde{\Theta} < 0$	Low e, $\Delta \tilde{\Theta} > 0$	Low e, $\Delta \tilde{\Theta} < 0$	High e, $\Delta \tilde{\Theta} > 0$
ttest *HIGH COMMERCE*, by(*HIGH DURATION*)			1 (p<0.11)	1 (p<0.11)
ttest *MULTI COUNTS CONVICT*, by(*HIGH DURATION*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH DURATION*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*HIGH DURATION*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DURATION*)			1	1
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH DURATION*)			1	1
SUB-TOTALS	0	0	5	5
ttest *MULTI COUNTS CONVICT*, by(*HIGH COMMERCE*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH COMMERCE*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH COMMERCE*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH COMMERCE*)	1	1		
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*HIGH COMMERCE*)	1	1		
SUB-TOTALS	3	3	1	1

Table 45, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Selection Model			Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model
	High e, $\Delta\,\tilde{\Theta} < 0$	Low e, $\Delta\,\tilde{\Theta} > 0$	Low e, $\Delta\,\tilde{\Theta} < 0$	High e, $\Delta\,\tilde{\Theta} > 0$
ttest *MULTI COUNTS INDICT*, by(*MULTI COUNTS CONVICT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*MULTI COUNTS CONVICT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*MULTI COUNTS CONVICT*)	1	1		
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*MULTI COUNTS CONVICT*)	0	0	0	0
SUB-TOTALS	2	2	1	1
ttest *POST-APPA INDICT YEAR*, by(*MULTI COUNTS INDICT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*MULTI COUNTS INDICT*)	0	0	0	0
Ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*MULTI COUNTS INDICT*)	1	1		
SUB-TOTALS	2	2	0	0
TOTALS				

Note: The variables *ZERO PREV CONVICT*, *HIGH DURATION*, *HIGH COMMERCE*, *MULTI COUNTS CONVICT*, and *MULTI COUNTS INDICT*, which represent the level of the prosecutor's evidence, **e**, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. Assume that, *ceteris paribus*, more or better evidence, **e,** increases ρ and/or S^T and $\tilde{\Theta}$.

Table 46 Descriptions of results of means tests that include empirical variables representing the litigants' *ex ante* expected trial sentence, S^T, the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, and other variables from the sample of defendants that did not go to trial (n = 2492)

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Selection Model			Evidence of Expected Trial Selection "Adjustments" Predicted by Trial Selection Model
	High S^T, $\Delta\tilde{\Theta} < 0$	Low S^T, $\Delta\tilde{\Theta} > 0$	Low S^T, $\Delta\tilde{\Theta} < 0$	High S^T, $\Delta\tilde{\Theta} > 0$
ttest *BID RIG NOT PRICE FIX*, by(*POST-APPA INDICT YEAR*)			1	1
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*POST-APPA INDICT YEAR*)			1	1
SUB-TOTALS	0	0	2	2
	High ρ, $\Delta\tilde{\Theta} < 0$	Low ρ, $\Delta\tilde{\Theta} > 0$	Low ρ, $\Delta\tilde{\Theta} < 0$	High ρ, $\Delta\tilde{\Theta} > 0$
ttest *HIGH DOJ TRIAL WIN PREV YEAR BY VIOL#1*, by(*BID RIG NOT PRICE FIX*)			1	1
SUB-TOTALS	0	0	1	1
TOTALS				

Note: The variable *POST-APPA INDICT YEAR*, which represents the litigants' *ex ante* expected trial sentence, S^T, and *BID RIG NOT PRICE FIX*, which represents the idiosyncratic element, ρ, of the prosecutor's *ex ante* expected probability of trial conviction, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results.

APPENDIX H

SAME-BRANCH MEANS TESTS OF THE TRIAL CONVICTION MODEL USING SAMPLE OF DEFENDANTS CONVICTED AT TRIAL

This appendix reports the results of two-sample means tests conducted to compare the relationship between variables representing case characteristics in the sample of defendants convicted at trial in order to test the predictions of the trial conviction model. All of the possible pairs of 13 variables allowed 78 separate means tests. Tables 48-51 report the means test results, by the pairs of variables used in the means tests. Each means test suggests two combinations (out of a possible four combinations) of the subject pair of variables in terms of their relative values. The means tests can suggest 156 different possible combinations of pairs of the relative values of the 13 variables.

Statistically insignificant results either suggest that judges and juries randomly convicted defendants at trial in the observed cases, or the results suggest that the judges and juries used a conviction decision process that otherwise produced a random mix of case characteristics in the cases that end in trial conviction. "0" denotes statistically insignificant results. Statistically significant results that are inconsistent with the trial conviction model's predictions sugges that the judges and juries non-randomly convicted defendants using a process that differs from the process specified by the trial conviction model. All other results lend empirical support for the trial conviction model's predictions.

The tables use the sample of defendants who were convicted at trial (n = 132), so the trial conviction model predictions, which Table 8 summarizes, apply. For example, if a variable representing the prosecutor's unit cost of trial expenditure, *C,* is paired with a variable representing the defendant's unit cost of trial expenditure, *K,* in a means test, then the means test will suggest either of the following combinations: i) (high value of *C,* high value of K) and (low value of *C,* low value of K); or ii) (high value of *C,* low value of K) and (low value of *C,* high value of K). The following combinations of pairs of variables representing C and *K,* in terms of the relative values of the variables, are consistent with the trial conviction model's predictions: (high value of *C,* high value of K), (low value of *C,* low value of K) and (low value of *C,* high value of K). The combination of (high value of *C,* low value of K) is inconsistent with the trial conviction model's predictions.

In general, the combinations found in the sample of defendants who *are convicted at trial* that are consistent with the trial conviction model's predictions can be stated in terms of the ex post probability of trial conviction, $\tilde{\Pi}$, as follows: (low value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$), (high value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$), and (high value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$). The combination (low value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$) is inconsistent with the trial conviction model's predictions.

A means test always will suggest (high value of *C,* low value of K) or, more generally, (low value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$), when it suggests (low value of *C,* high value of K) or, more generally, (high value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$). Thus, using the sample of defendants *convicted at trial*, the percentage of statistically significant combinations that are consistent with the trial conviction model cannot fall below 50 percent and cannot exceed 100 percent of the *statistically significant* combinations of pairs of variables that the means tests report. The reported results should be judged accordingly. Table 47 summarizes the results of the individual means tests covered by Tables 48-51, which follow.

Table 47 Results of means tests of implications of the trial conviction selection process, using the sample of defendants convicted at trial (n = 132)

Appendix H Table Reporting Results	Statistically Insignificant Results	Statistically Significant Results			
		Number of Combinations Consistent with Trial Conviction Model's Predictions			Number of Combinations Inconsistent with Trial Conviction Model's Predictions
Table 48	14	7	7	9	9
Table 49	44	7	7	5	5
Table 50	22	2	2	7	7
Table 51	2	0	0	0	0
GRAND TOTALS	82	16	16	21	21
	82	53			21

Table 48 Descriptions of results of means tests that include empirical variables representing the prosecutor's unit (opportunity) cost of trial expenditure, *C*, and other variables from the sample of defendants convicted at trial (n = 132)

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule			Not Expected According to Trial Conviction Model
	High *C*, $\Delta\,\tilde{\Pi} > 0$	Low *C*, $\Delta\,\tilde{\Pi} < 0$	Low *C*, $\Delta\,\tilde{\Pi} > 0$	High *C*, $\Delta\,\tilde{\Pi} < 0$
ttest *HIGH ANNUAL FORTUNE500*, by(*HIGH DOJBUDGET*)			1	1
ttest *ONLY LOW RANK*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH PROFIT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *HIGH COMPENSATION*, by(*HIGH DOJBUDGET*)	1	1		
ttest *ZERO OTHER INDICT*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *ZERO PREV CONVICT*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *HIGH DURATION*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH COMMERCE*, by(*HIGH DOJBUDGET*)			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH DOJBUDGET*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DOJBUDGET*)			1	1
SUB-TOTALS	3	3	6	6

Table 48, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule			Not Expected According to Trial Conviction Model
	High *C*, $\Delta\,\tilde{\Pi} > 0$	Low *C*, $\Delta\,\tilde{\Pi} < 0$	Low *C*, $\Delta\,\tilde{\Pi} > 0$	High *C*, $\Delta\,\tilde{\Pi} < 0$
ttest *ONLY LOW RANK*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *HIGH PROFIT*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *HIGH COMPENSATION*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *ZERO OTHER INDICT*, by(*HIGH ANNUAL FORTUNE500*)	1 (p<0.15)	1 (p<0.15)		
ttest *ZERO PREV CONVICT*, by(*HIGH ANNUAL FORTUNE500*)	1 (p<0.15)	1 (p<0.15)		
ttest *HIGH DURATION*, by(*HIGH ANNUAL FORTUNE500*)			1 (p<0.15)	1 (p<0.15)
ttest *HIGH COMMERCE*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *MULTI COUNTS INDICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
SUB-TOTALS	4	4	3	3
TOTALS				

Note: The variables *HIGH DOJBUDGET* and *HIGH ANNUAL FORTUNE500*, which represent the prosecutor's unit cost of trial expenditure, *C*, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. Assume that, *ceteris paribus*, more evidence, **e,** increases the value of $\tilde{\Pi}$.

Table 49 Descriptions of results of means tests that include empirical variables representing the defendant's unit (opportunity) cost of trial expenditure, *K*, and other variables from the sample of defendants convicted at trial (n = 132)

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule			Not Expected According to Trial Conviction Model
	Low *K*, $\Delta\,\tilde{\Pi} > 0$	High *K*, $\Delta\,\tilde{\Pi} < 0$	High *K*, $\Delta\,\tilde{\Pi} > 0$	Low *K*, $\Delta\,\tilde{\Pi} < 0$
ttest *HIGH PROFIT*, by(*ONLY LOW RANK*)			1 (p<0.12)	1 (p<0.12)
ttest *HIGH COMPENSATION*, by(*ONLY LOW RANK*)			1	1
ttest *ZERO OTHER INDICT*, by(*ONLY LOW RANK*)			1	1
ttest *ZERO PREV CONVICT*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *HIGH DURATION*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*ONLY LOW RANK*)	1	1		
ttest *MULTI COUNTS CONVICT*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *POST-APPA INDICT YEAR*, by(*ONLY LOW RANK*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*ONLY LOW RANK*)	0	0	0	0
SUB-TOTALS	2	2	3	3

Table 49, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule			Not Expected According to Trial Conviction Model
	Low *K*, $\Delta\,\tilde{\Pi} > 0$	High *K*, $\Delta\,\tilde{\Pi} < 0$	High *K*, $\Delta\,\tilde{\Pi} > 0$	Low *K*, $\Delta\,\tilde{\Pi} < 0$
ttest *HIGH COMPENSATION*, by(*HIGH PROFIT*)			1	1
ttest *ZERO OTHER INDICT*, by(*HIGH PROFIT*)			1	1
ttest *ZERO PREV CONVICT*, by(*HIGH PROFIT*)	0	0	0	0
ttest *HIGH DURATION*, by(*HIGH PROFIT*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*HIGH PROFIT*)	1	1		
ttest *MULTI COUNTS CONVICT*, by(*HIGH PROFIT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH PROFIT*)	0	0	0	0
ttest *POST-APPA INDICT YEAR*, by(*HIGH PROFIT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*HIGH PROFIT*)	0	0	0	0
SUB-TOTALS	2	2	2	2

Table 49, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Conviction “Adjustment” Predicted by the Model of the Judge’s or Jury’s Trial Conviction Decision Rule			Not Expected According to Trial Conviction Model
	Low *K*, $\Delta\,\tilde{\Pi} > 0$	High *K*, $\Delta\,\tilde{\Pi} < 0$	High *K*, $\Delta\,\tilde{\Pi} > 0$	Low *K*, $\Delta\,\tilde{\Pi} < 0$
ttest *ZERO OTHER INDICT*, by(*HIGH COMPENSATION*)	0	0	0	0
ttest *ZERO PREV CONVICT*, by(*HIGH COMPENSATION*)	0	0	0	0
ttest *HIGH DURATION*, by(*HIGH COMPENSATION*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*HIGH COMPENSATION*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*HIGH COMPENSATION*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH COMPENSATION*)	0	0	0	0
ttest *POST-APPA INDICT YEAR*, by(*HIGH COMPENSATION*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*HIGH COMPENSATION*)	1	1		
SUB-TOTALS	2	2	0	0

Table 49, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule			Not Expected According to Trial Conviction Model
	Low *K* or e, $\Delta \tilde{\Pi} > 0$	High *K* or e, $\Delta \tilde{\Pi} < 0$	High *K* or e, $\Delta \tilde{\Pi} > 0$	Low *K* or e, $\Delta \tilde{\Pi} < 0$
ttest *ZERO PREV CONVICT*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *HIGH DURATION*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *POST-APPA INDICT YEAR*, by(*ZERO OTHER INDICT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*ZERO OTHER INDICT*)	0	0	0	0
SUB-TOTALS	1	1	0	0

Note: The variables *ONLY LOW RANK*, *HIGH PROFIT*, *HIGH COMPENSATION*, and *ZERO OTHER INDICT*, which represent the defendant's unit cost of trial expenditure, *K*, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. Assume that, *ceteris paribus*, more evidence, **e,** increases the value of $\tilde{\Pi}$.

Table 50 Descriptions of results of means tests that include empirical variables representing the amount or strength of the prosecutor's evidence, **e**, and other variables from the sample of defendants convicted at trial (n = 132)

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule			Not Expected According to Trial Conviction Model
	Low e, $\Delta\,\tilde{\Pi} > 0$	High e, $\Delta\,\tilde{\Pi} < 0$	High e, $\Delta\,\tilde{\Pi} > 0$	Low e, $\Delta\,\tilde{\Pi} < 0$
ttest *HIGH DURATION*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ZERO PREV CONVICT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *BID RIG NOT PRICE FIX*, by(*ZERO PREV CONVICT*)			1 (p<0.15)	1 (p<0.15)
SUB-TOTALS	0	0	2	2
ttest *HIGH COMMERCE*, by(*HIGH DURATION*)			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH DURATION*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH DURATION*)	0	0	0	0
ttest *POST-APPA INDICT YEAR*, by(*HIGH DURATION*)	0	0	0	0
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DURATION*)			1	1
SUB-TOTALS	0	0	2	2

Table 50, continued

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule			Not Expected According to Trial Conviction Model
	Low e, $\Delta \tilde{\Pi} > 0$	High e, $\Delta \tilde{\Pi} < 0$	High e, $\Delta \tilde{\Pi} > 0$	Low e, $\Delta \tilde{\Pi} < 0$
ttest *MULTI COUNTS CONVICT*, by(*HIGH COMMERCE*)			1	1
ttest *MULTI COUNTS INDICT*, by(*HIGH COMMERCE*)	0	0	0	0
ttest *POST-APPA INDICT YEAR*, by(*HIGH COMMERCE*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH COMMERCE*)	1	1		
SUB-TOTALS	1	1	2	2
ttest *MULTI COUNTS INDICT*, by(*MULTI COUNTS CONVICT*)	0	0	0	0
ttest *POST-APPA INDICT YEAR*, by(*MULTI COUNTS CONVICT*)	0	0	0	0
ttest *BID RIG NOT PRICE FIX*, by(*MULTI COUNTS CONVICT*)	0	0	0	0
SUB-TOTALS	0	0	0	0
ttest *POST-APPA INDICT YEAR*, by(*MULTI COUNTS INDICT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*MULTI COUNTS INDICT*)			1	1
SUB-TOTALS	1	1	1	1
TOTALS				

Note: The variables *ZERO PREV CONVICT*, *HIGH DURATION*, *HIGH COMMERCE*, *MULTI COUNTS CONVICT*, and *MULTI COUNTS INDICT*, which represent the level of the prosecutor's evidence, **e**, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. Assume that, *ceteris paribus*, more evidence, **e,** increases the value of $\tilde{\Pi}$.

Table 51 Descriptions of results of means tests that include an empirical variable representing the litigants' *ex ante* expected trial sentence, S^T, and other variables from the sample of defendants convicted at trial (n = 132)

Variables Used in Two-Sample Means Tests	Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule			Not Expected According to Trial Conviction Model
	Low S^T, $\Delta\tilde{\Pi} > 0$	High S^T, $\Delta\tilde{\Pi} < 0$	High S^T, $\Delta\tilde{\Pi} > 0$	Low S^T, $\Delta\tilde{\Pi} < 0$
ttest *BID RIG NOT PRICE FIX*, by(*POST-APPA INDICT YEAR*)	0	0	0	0
TOTALS	0	0	0	0

Note: The variable *POST-APPA INDICT YEAR*, which represents the litigants' *ex ante* expected trial sentence, S^T, pairs with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results.

APPENDIX I

SAME-BRANCH MEANS TESTS OF THE TRIAL CONVICTION MODEL
USING SAMPLE OF DEFENDANTS NOT CONVICTED AT TRIAL

Tables 53-55 report the results of means tests that use the sample of defendants who went to trial but were *not* convicted (n = 469). The means tests reported here use 13 variables and are similar to the means tests reported in Appendix H. The trial conviction model predictions summarized in Table 8 apply, with one caveat: the combinations labeled "inconsistent" in Table 8 are labeled "consistent" here. Suppose a variable representing the prosecutor's unit cost of trial expenditure, *C,* is paired with a variable representing the defendant's unit cost of trial expenditure, *K,* in a means test. The following combination of the relative values of C and *K* is consistent with the trial conviction model's predictions: (low value of *C,* high value of K). The combinations of (high value of *C,* high value of K), (low value of *C,* low value of K), and (high value of *C,* low value of K) are inconsistent with the trial conviction model's predictions for the sample of defendants who went to trial but were *not* convicted (n = 469).

In general, the combination found in the sample of defendants who *go to trial but are not convicted* that is consistent with the trial conviction model's predictions can be stated in terms of the ex post probability of trial conviction, $\tilde{\Pi}$, as follows: (low value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$). The combinations (low value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$), (high value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$), and (high value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$) are inconsistent with the trial conviction model's predictions.

A means test always will suggests the combination (low value of *C,* high value of K) or, more generally, (high value of $\tilde{\Pi}$, high value of $\tilde{\Pi}$) when it suggests (high value of *C,* low value of K) or, more generally, (low value of $\tilde{\Pi}$, low value of $\tilde{\Pi}$). Thus, using the sample of defendants *who went to trial but were not convicted*, the percentage of combinations consistent with the trial conviction model's predictions cannot fall below zero percent and cannot exceed 50 percent of the statistically significant combinations of pairs of variables that the means tests produce. The reported results should be judged accordingly. Table 52 summarizes the results of the individual means tests covered by Tables 53-55, which follow.

Table 52 Results of means tests of implications of the trial conviction selection process, using the sample of defendants convicted at trial (n = 132)

Appendix I Table Reporting Results	Statistically Insignificant Results	Statistically Significant Results			
		Number of Combinations Inconsistent with Trial Conviction Model's Predictions			Number of Combinations Consistent with Trial Conviction Model's Predictions
Table 53	16	6	6	9	9
Table 54	26	11	11	10	10
Table 55	14	5	5	8	8
Table 56	0	0	0	1	1
GRAND TOTALS	56	22	22	28	28
	56	72			28

Table 53 Descriptions of results of means tests that include empirical variables representing the prosecutor's unit (opportunity) cost of trial expenditure, *C*, and other variables from the sample of defendants that went to trial and were not convicted (n = 469)

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Conviction Model			Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule
	Low *C*, $\Delta\,\tilde{\Pi} < 0$	High *C*, $\Delta\,\tilde{\Pi} > 0$	High *C*, $\Delta\,\tilde{\Pi} < 0$	Low *C*, $\Delta\,\tilde{\Pi} > 0$
ttest *HIGH ANNUAL FORTUNE500*, by(*HIGH DOJBUDGET*)			1	1
ttest *ONLY LOW RANK*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *HIGH PROFIT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *HIGH COMPENSATION*, by(*HIGH DOJBUDGET*)	1	1		
ttest *ZERO OTHER INDICT*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *ZERO PREV CONVICT*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH DURATION*, by(*HIGH DOJBUDGET*)			1	1
ttest *HIGH COMMERCE*, by(*HIGH DOJBUDGET*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *MULTI COUNTS INDICT*, by(*HIGH DOJBUDGET*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH DOJBUDGET*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DOJBUDGET*)			1	1
SUB-TOTALS	4	4	5	5

Table 53, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Conviction Model			Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule
	Low *C*, $\Delta\,\widetilde{\Pi} < 0$	High *C*, $\Delta\,\widetilde{\Pi} > 0$	High *C*, $\Delta\,\widetilde{\Pi} < 0$	Low *C*, $\Delta\,\widetilde{\Pi} > 0$
ttest *ONLY LOW RANK*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *HIGH PROFIT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *HIGH COMPENSATION*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *ZERO OTHER INDICT*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *ZERO PREV CONVICT*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *HIGH DURATION*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH ANNUAL FORTUNE500*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH ANNUAL FORTUNE500*)	1	1		
ttest *POST-APPA INDICT YEAR*, by(*HIGH ANNUAL FORTUNE500*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH ANNUAL FORTUNE500*)			1	1
SUB-TOTALS	2	2	4	4
TOTALS	6	6	9	9

Note: The variables *HIGH DOJBUDGET* and *HIGH ANNUAL FORTUNE500*, which represent the prosecutor's unit cost of trial expenditure, *C*, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. Assume that, *ceteris paribus*, more evidence, **e,** increases the value of $\widetilde{\Pi}$.

Table 54 Descriptions of results of means tests that include empirical variables representing the defendant's unit (opportunity) cost of trial expenditure, *K*, and other variables from the sample of defendants that went to trial and were not convicted (n = 469)

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Conviction Model			Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule
	High *K*, $\Delta\,\tilde{\Pi} < 0$	Low *K*, $\Delta\,\tilde{\Pi} > 0$	Low *K*, $\Delta\,\tilde{\Pi} < 0$	High *K*, $\Delta\,\tilde{\Pi} > 0$
ttest *HIGH PROFIT*, by(*ONLY LOW RANK*)			1	1
ttest *HIGH COMPENSATION*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *ZERO OTHER INDICT*, by(*ONLY LOW RANK*)	1 (p<0.10)	1 (p<0.10)		
ttest *ZERO PREV CONVICT*, by(*ONLY LOW RANK*)	1 (p<0.15)	1 (p<0.15)		
ttest *HIGH DURATION*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*ONLY LOW RANK*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ONLY LOW RANK*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*ONLY LOW RANK*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*ONLY LOW RANK*)	0	0	0	0
SUB-TOTALS	2	2	3	3

Table 54, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Conviction Model			Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule
	High *K*, $\Delta\,\tilde{\Pi} < 0$	Low *K*, $\Delta\,\tilde{\Pi} > 0$	Low *K*, $\Delta\,\tilde{\Pi} < 0$	High *K*, $\Delta\,\tilde{\Pi} > 0$
ttest *HIGH COMPENSATION*, by(*HIGH PROFIT*)			1	1
ttest *ZERO OTHER INDICT*, by(*HIGH PROFIT*)	0	0	0	0
ttest *ZERO PREV CONVICT*, by(*HIGH PROFIT*)	0	0	0	0
ttest *HIGH DURATION*, by(*HIGH PROFIT*)	1	1		
ttest *HIGH COMMERCE*, by(*HIGH PROFIT*)	1	1		
ttest *MULTI COUNTS CONVICT*, by(*HIGH PROFIT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH PROFIT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*HIGH PROFIT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*HIGH PROFIT*)	1	1		
SUB-TOTALS	4	4	2	2

Table 54, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Conviction Model			Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule
	High *K*, $\Delta \tilde{\Pi} < 0$	Low *K*, $\Delta \tilde{\Pi} > 0$	Low *K*, $\Delta \tilde{\Pi} < 0$	High *K*, $\Delta \tilde{\Pi} > 0$
ttest *ZERO OTHER INDICT*, by(*HIGH COMPENSATION*)			1	1
ttest *ZERO PREV CONVICT*, by(*HIGH COMPENSATION*)			1 (p<0.14)	1 (p<0.14)
ttest *HIGH DURATION*, by(*HIGH COMPENSATION*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*HIGH COMPENSATION*)	1	1		
ttest *MULTI COUNTS CONVICT*, by(*HIGH COMPENSATION*)			1 (p<0.14)	1 (p<0.14)
ttest *MULTI COUNTS INDICT*, by(*HIGH COMPENSATION*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*HIGH COMPENSATION*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*HIGH COMPENSATION*)	1	1		
SUB-TOTALS	3	3	4	4

Table 54, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Conviction Model			Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule
	High *K* or e, $\Delta\,\tilde{\Pi} < 0$	Low *K* or e, $\Delta\,\tilde{\Pi} > 0$	Low *K* or e, $\Delta\,\tilde{\Pi} < 0$	High *K* or e, $\Delta\,\tilde{\Pi} > 0$
ttest *ZERO PREV CONVICT*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *HIGH DURATION*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*ZERO OTHER INDICT*)	1 ($p<0.10$)	1 ($p<0.10$)		
ttest *MULTI COUNTS CONVICT*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ZERO OTHER INDICT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*ZERO OTHER INDICT*)	0	0	0	0
ttest *BID RIG NOT PRICE FIX*, by(*ZERO OTHER INDICT*)	1	1		
SUB-TOTALS	2	2	1	1
TOTALS	11	11	10	10

Note: The variables *ONLY LOW RANK*, *HIGH PROFIT*, *HIGH COMPENSATION*, and *ZERO OTHER INDICT*, which represent the defendant's unit cost of trial expenditure, *K*, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. Assume that, *ceteris paribus*, more evidence, **e,** increases the value of $\tilde{\Pi}$.

Table 55 Descriptions of results of means tests that include empirical variables representing the amount or strength of the prosecutor's evidence, **e**, and other variables from the sample of defendants that went to trial and were not convicted (n = 469)

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Conviction Model			Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule
	High e, $\Delta\,\tilde{\Pi} < 0$	Low e, $\Delta\,\tilde{\Pi} > 0$	Low e, $\Delta\,\tilde{\Pi} < 0$	High e, $\Delta\,\tilde{\Pi} > 0$
ttest *HIGH DURATION*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *HIGH COMMERCE*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *MULTI COUNTS CONVICT*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*ZERO PREV CONVICT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*ZERO PREV CONVICT*)	0	0	0	0
ttest *BID RIG NOT PRICE FIX*, by(*ZERO PREV CONVICT*)	1	1		
SUB-TOTALS	1	1	1	1

Table 55, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Conviction Model			Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule
	High e, $\Delta\tilde{\Pi} < 0$	Low e, $\Delta\tilde{\Pi} > 0$	Low e, $\Delta\tilde{\Pi} < 0$	High e, $\Delta\tilde{\Pi} > 0$
ttest *HIGH COMMERCE*, by(*HIGH DURATION*)			1	1
ttest *MULTI COUNTS CONVICT*, by(*HIGH DURATION*)			1 (p<0.15)	1 (p<0.15)
ttest *MULTI COUNTS INDICT*, by(*HIGH DURATION*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*HIGH DURATION*)			1	1
ttest *BID RIG NOT PRICE FIX*, by(*HIGH DURATION*)			1	1
SUB-TOTALS	0	0	5	5
ttest *MULTI COUNTS CONVICT*, by(*HIGH COMMERCE*)	0	0	0	0
ttest *MULTI COUNTS INDICT*, by(*HIGH COMMERCE*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*HIGH COMMERCE*)	0	0	0	0
ttest *BID RIG NOT PRICE FIX*, by(*HIGH COMMERCE*)	1	1		
SUB-TOTALS	1	1	1	1

Table 55, continued

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Conviction Model			Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule
	High e, $\Delta\, \tilde{\Pi} < 0$	Low e, $\Delta\, \tilde{\Pi} > 0$	Low e, $\Delta\, \tilde{\Pi} < 0$	High e, $\Delta\, \tilde{\Pi} > 0$
ttest *MULTI COUNTS INDICT*, by(*MULTI COUNTS CONVICT*)			1	1
ttest *POST-APPA INDICT YEAR*, by(*MULTI COUNTS CONVICT*)	1	1		
ttest *BID RIG NOT PRICE FIX*, by(*MULTI COUNTS CONVICT*)	1 (p<0.15)	1 (p<0.15)		
SUB-TOTALS	2	2	1	1
ttest *POST-APPA INDICT YEAR*, by(*MULTI COUNTS INDICT*)	0	0	0	0
ttest *BID RIG NOT PRICE FIX*, by(*MULTI COUNTS INDICT*)	1	1		
SUB-TOTALS	1	1	0	0
TOTALS	5	5	8	8

Note: The variables *ZERO PREV CONVICT*, *HIGH DURATION*, *HIGH COMMERCE*, *MULTI COUNTS CONVICT*, and *MULTI COUNTS INDICT*, which represent the level of the prosecutor's evidence, **e**, pair with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results. Assume that, *ceteris paribus*, more evidence, **e,** increases the value of $\tilde{\Pi}$.

Table 56 Descriptions of results of means tests that include empirical variables representing the litigants' *ex ante* expected trial sentence, S^T, and other variables from the sample of defendants that went to trial and were not convicted (n = 469)

Variables Used in Two-Sample Means Tests	Not Expected According to Trial Conviction Model			Evidence of Expected Trial Conviction "Adjustment" Predicted by the Model of the Judge's or Jury's Trial Conviction Decision Rule
	High S^T, $\Delta\tilde{\Pi} < 0$	Low S^T, $\Delta\tilde{\Pi} > 0$	Low S^T, $\Delta\tilde{\Pi} < 0$	High S^T, $\Delta\tilde{\Pi} > 0$
ttest *BID RIG NOT PRICE FIX*, by(*POST-APPA INDICT YEAR*)			1	1
TOTALS	0	0	1	1

Note: The variable *POST-APPA INDICT YEAR*, which represents the litigants' *ex ante* expected trial sentence, S^T, pairs with other variables in the tests reported here. "0" denotes statistically insignificant results, so that the reported totals only reflect statistically significant results.

BIBLIOGRAPHY

Alexander, Cindy R., Jennifer Arlen and Mark A. Cohen, "Regulating Corporate Criminal Sanctions: Evidence on the Effect of the U.S. Sentencing Guidelines," XLII *Journal of Law and Economics* (1999).

Alschuler, Albert W., "The Failure of Sentencing Guidelines: A Plea for Less Aggregation," 58 *University of Chicago Law Review* (1991).

Alschuler, Albert W., "The Prosecutor's Role in Plea Bargaining," 36 *University of Chicago Law Review* (1968).

Anderson, James M., Jeffrey R. Kling and Kate Stith, "Measuring Inter-Judge Sentencing Disparity Before and After the Federal Sentencing Guidelines," XLII *Journal of Law and Economics* (1999).

Andreoni, James, "Reasonable Doubt and the Optimal Magnitude of Fines: Should the Penalty Fit the Crime?" 22 *RAND Journal of Economics* (1991).

Asch, Peter, "The Determinants and Effects of Antitrust Activity," 18 *Journal of Law and Economics* (1975).

Baker, Donald I., "To Indict or Not to Indict: Prosecutorial Discretion in Sherman Act Enforcement," 63 *Cornell Law Review* (1978).

Bebchuk, Lucian Arye and Louis Kaplow, "Optimal Sanctions and Differences in Individuals' Likelihood of Avoiding Detection," 13 *International Review of Law and Economics* (1993).

Block, Michael K. and William M. Rhodes, "Forecasting the Impact of the Federal Sentencing Guidelines," 7 *Behavioral Sciences & the Law* (1989).

Briggs, Hugh C., Kathleen D. Huryn, and Mark E. McBride, "Treble Damages and the Incentive to Sue and Settle," 27 *RAND Journal of Economics* (1996).

Bryant, Peter G. and E. Woodrow Eckard, "Price Fixing: The Probability of Getting Caught," 73 *Review of Economics and Statistics* (1991).

Cartwright, Phillip A. and David R. Kamerschen, "Variations in Antitrust Enforcement Activity," 3 *Review of Industrial Organization* (1986).

Claybault, James M. and Michael K. Block, Sherman Act Indictments, 1955-1980, New York: Federal Legal Publications (1981).

Cohen, Mark A. and David T. Scheffman, "The Antitrust Sentencing Guideline: Is the Punishment Worth the Costs?" 27 *American Criminal Law Review* (1989).

Cohen, Mark A., "Explaining Judicial Behavior or What's 'Unconstitutional' About the Sentencing Commission," 7 *Journal of Law, Economics and Organization* (1991).

Cohen, Mark A., "The Motives of Judges: Empirical Evidence from Antitrust Sentencing," 12 *International Review of Law and Economics* (1992).

Cohen, Mark A., "The Role of Criminal Sanctions in Antitrust Enforcement," 7 *Contemporary Policy Issues* (1989).

Cooter, Robert D. and Daniel L. Rubinfeld, "Economic Analysis of Legal Disputes and Their Resolution," 27 *Journal of Economic Literature* (1989).

Cutter, Stephanie A., "Organizational Sentencing," 32 *American Criminal Law Review* (1995).

Easterbrook, Frank H., "Criminal Process as a Market System," 12 *Journal of Legal Studies* (1983).

Eisenberg, Theodore, "Testing the Selection Effect: A New Theoretical Framework with Empirical Tests," 19 *Journal of Legal Studies* (1990).

Eisenberg, Theodore and Henry S. Farber, "The Litigious Plaintiff Hypothesis: Case Selection and Resolution," 28 *RAND Journal of Economics* (1997).

Eisenberg, Theodore and Henry Farber, "The Government as Litigant: Further Tests of the Case Selection Model," 5 *American Law and Economics Review* (2003).

Elder, Harold W., "Trials and Settlements in the Criminal Courts: An Empirical Analysis of Dispositions and Sentencing," 18 *Journal of Legal Studies* (1989).

Finkelstein, Michael O., "A Statistical Analysis of Guilty Plea Practices in the Federal Courts," 89 *Harvard Law Review* (1975).

Forst, Brian and Kathleen B. Brosi, "A Theoretical and Empirical Analysis of the Prosecutor," 6 *Journal of Legal Studies* (1977).

Fournier, Gary M. and Thomas W. Zuehlke, "Litigation and Settlement: An Empirical Approach," 71 *Review of Economics and Statistics* (1989).

Fraas, Arthur G., and Douglas F. Greer, "Market Structure and Price Collusion: An Empirical Analysis," 26 *Journal of Industrial Economics* (1977).

Froeb, Luke, "Criminal Plea Bargaining with Uncertainty: An Empirical Model of Sentencing," Working Paper, The Law School, The University of Chicago (1989).

Froeb, Luke, “The Adverse Selection of Cases for Trial,” 13 *International Review of Law and Economics* (1993).

Gallo, Joseph C., Joseph L. Craycraft and Steven C. Bush, “Guess Who Came to Dinner: An Empirical Study of Federal Antitrust Enforcement for the Period 1963-1984,” 2 *Review of Industrial Organization* (1985).

Gallo, Joseph C., Kenneth Dau-Schmidt, Joseph L. Craycraft, and Charles J. Parker, “Department of Justice Antitrust Enforcement, 1955-1997: An Empirical Study,” 17 *Review of Industrial Organization* (2000).

Gallo, Joseph C., Kenneth G. Dau-Schmidt, Joseph L. Craycraft and Charles J. Parker, “Criminal Penalties Under the Sherman Act: A Study of Law and Economics,” 16 *Research in Law and Economics* (1994).

Gay, Gerald D., Martin F. Grace, Jayant R. Kale and Thomas H. Noe, “Noisy Juries and the Choice of Trial Mode in a Sequential Signaling Game: Theory and Evidence,” 20 *Rand Journal of Economics* (1989).

Greene, William H., *Econometric Analysis*, *Second Edition*, New York: Macmillan Publishing Company (1993).

Grossman, Gene M. and Michael L. Katz, “Plea Bargaining and Social Welfare,” 73 *American Economic Review* (1983).

Hay, George and Daniel Kelley, “An Empirical Study of Price Fixing Conspiracies,” 17 *Journal of Law and Economics* (1974).

Heckman, James J., “Sample Selection Bias as a Specification Error,” 47 *Econometrica* (1979).

Higgins, Richard S. and Paul H. Rubin, “Judicial Discretion,” 9 *Journal of Legal Studies* (1980).

Hylton, Keith N., “Asymmetric Information and the Selection of Disputes for Litigation,” 22 *Journal of Legal Studies* (1993).

Joyce, Jon M., “Effect of Firm Organizational Structure on Incentives to Engage in Price Fixing,” 7 *Contemporary Policy Issues* (1989).

Kaplow, Louis, “The Optimal Probability and Magnitude of Fines for Acts That Definitely are Undesirable,” 12 *International Review of Law and Economics* (1992).

Katz, Avery, “Judicial Decisionmaking and Litigation Expenditure,” 8 *International Review of Law and Economics* (1988).

Klein, Christopher C., “Predation in the Courts: Legal Versus Economic Analysis in Sham Litigation,” 10 *International Review of Law and Economics* (1990).

Kobayashi, Bruce H., “Deterrence with Multiple Defendants: An Explanation for ‘unfair’ plea bargains,” 23 *RAND Journal of Economics* (1992).

Kobayashi, Bruce H., John R. Lott, Jr., “In Defense of Criminal Defense Expenditures and Plea Bargaining,” 16 *International Review of Law and Economics* (1996).

Kobayashi, Bruce H., John R. Lott, Jr., “Low Probability-High-Penalty Enforcement Strategies and the Efficient Operation of the Plea-Bargaining System,” 12 *International Review of Law and Economics* (1992).

LaCasse, Chantele and A. Abigail Payne, “Federal Sentencing Guidelines and Mandatory Minimum Sentences: Do Defendants Bargain in the Shadow of the Judge?” XLII *Journal of Law and Economics* (1999).

Lachman, Judith A., “Daring the Courts: Trial and Bargaining Consequences of Minimum Penalties,” 90 *Yale Law Journal* (1981).

Landes, William M., “An Economic Analysis of the Courts,” 14 *Journal of Law and Economics* (1971).

Landes, William M., “Optimal Sanctions for Antitrust Violations,” 50 *University of Chicago Law Review* (1983).

Long, William F., Richard Schramm and Robert Tollison, “The Economic Determinants of Antitrust Activity,” 16 *Journal of Law and Economics* (1973).

Maddala, G.S., *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press (1983).

Marvel, Howard P., Jeffry M. Netter and Anthony M. Robinson, “Price Fixing and Civil Damages: An Economic Analysis,” 40 *Stanford Law Review* (1988).

Miceli, Thomas J. and Metin M. Cosgel,”Reputation and Judicial Decision-making,” 23 *Journal of Economic Behavior and Organization* (1994).

Nalebluff, Barry, “Credible Pretrial Negotiation,” 18 *RAND Journal of Economics* (1987).

Note, “*Nolo* Pleas in Antitrust Cases,” 79 *Harvard Law Review* (1966).

P’ng, I. P. L., “Strategic Behavior in Suit, Settlement, and Trial,” *The Bell Journal of Economics* (1983).

Payne, A. Abigail, "Does Inter-Judge Disparity Really Matter? An Analysis of the Effects of Sentencing Reforms in Three Federal District Courts," 17 *International Review of Law and Economics* (1997).

Posner, Richard A., "A Statistical Study of Antitrust Enforcement," 13 *Journal of Law and Economics* (1970).

Priest, George L. and Benjamin Klein, "The Selection of Disputes for Litigation," 13 *Journal of Legal Studies* (1984).

Priest, George L., "Reexamining the Selection Hypothesis: Learning from Wittman's Mistakes," 14 *Journal of Legal Studies* (1985).

Priest, George L., "Selective Characteristics of Litigation," 9 *Journal of Legal Studies* (1980).

Rasmusen, Eric, "Predictable and Unpredictable Error in Tort Awards: The Effect of Plaintiff Self-Selection and Signaling," 15 *International Review of Law and Economics* (1995).

Reinganum, Jennifer F., "Plea Bargaining and Prosecutorial Discretion," 78 *American Economic Review* (1988).

Reinganum, Jennifer F., "Sentencing Guidelines, Judicial Discretion and Plea Bargaining," 31 *Rand Journal of Economics* (2000).

Rhodes, William M., "The Economics of Criminal Courts: A Theoretical and Empirical Investigation," 5 *Journal of Legal Studies* (1976).

Shavell, Steven, "Any Frequency of Plaintiff Victory at Trial is Possible," 25 *Journal of Legal Studies* (1996).

Siegfried, John J., "The Determinants of Antitrust Activity," 18 Journal of Law and Economics (1975).

Siegfried, John J. and David B. Sapper, "A Cloud of Doubt About the Shadow Price of Sunshine: We're Still in the Dark," 49 *Public Finance* (1994).

Snyder, Edward A., "New Insights into the Decline of Antitrust Enforcement," 7 *Contemporary Policy Issues* (1989).

Snyder, Edward A., "The Effect of Higher Criminal Penalties on Antitrust Enforcement," 33 *Journal of Law and Economics* (1990).

Waldfogel, Joel, "Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation," 41 *Journal of Law and Economics* (1998).

Waldfogel, Joel, “The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory,” 103 *Journal of Political Economy* (1995).

Wheeler, Stanton, Kenneth Mann and Austin Sarat, *Sitting in Judgment: The Sentencing of White-Collar Criminals*, New Haven: Yale University Press (1988).

Wittman, Donald, “Dispute Resolution, Bargaining, and the Selection of Cases for Trial: A Study of the Generation of Biased and Unbiased Data” 17 *Journal of Legal Studies* (1988).

Wittman, Donald, “Is the Selection of Cases for Trial Biased?” 14 *Journal of Legal Studies* (1985).

Wood, William C., “Costs and Benefits of Per Se Rules in Antitrust Enforcement,” *Antitrust Bulletin* (Winter 1993).