As a threshold matter, high market shares are considered informative of potential underlying competitive dynamics, especially issues of market power and dominance. However, as is well known, high shares may not tell the entire story. Beyond issues related to market definition, entry, and efficiencies, this article notes an additional reason why caution may be warranted before drawing conclusions about dominance or market power from share evidence. Specifically, such conclusions can be statistically unsupportable in the presence of small sample issues. In markets with relatively few transactions (that is, “thinly” traded markets), the implications of observed high market shares are much less clear than in more thickly traded markets. It is entirely possible that situations that appear to implicate a dominant firm actually reflect pure random chance in a competitive process involving equally matched or nondominant firms. This article discusses theories and methods for distinguishing between outcomes that exhibit strong statistical evidence of dominance as opposed to those that merely reflect the random distribution of winnings among nondominant firms.

* Economist, Encino, CA


AUTHORS’ NOTE: This article reflects our views only and does not necessarily represent the views of the Federal Trade Commission or any individual commissioner. Mark Glick, Dave Schmidt, Mike Vita, Nicholas Franczyk, Malcolm Coate, and John Yun provided helpful comments and suggestions. Remaining errors are ours.
KEY WORDS: antitrust, market structure, monopolization, dominance, small samples, statistical significance
I. INTRODUCTION

It is generally maintained in competition matters that a high market share, such as seventy percent, provides prima facie evidence of monopoly power or dominance.¹ Such a

¹ Under U.S. competition law, evidence of market power or monopoly power is a requisite condition for a monopolization or attempted monopolization offense. See, e.g., United States v. Grinnell Corp., 384 U.S. 563, 570–71 (1966). The courts have defined market power and monopoly power in related but different manners. The Supreme Court has stated that “market power exists whenever prices can be raised above the levels that would be charged in a competitive market.” Jefferson Parish Hosp. Dist. No. 2 v. Hyde, 466 U.S. 2, 27 n.46 (1984). Monopoly power has been defined as "the power to control prices or exclude competition.” United States v. E. I. du Pont de Nemours & Co. (Cellophane), 351 U.S. 377, 391 (1956). EU competition law generally defines a firm as dominant if it behaves "to an appreciable extent independently of its competitors, customers and ultimately of its consumer." Case 27/76, United Brands Cont’l BV v. Comm’n, 1978 E.C.R. 207. These legal definitions of market power, monopoly power and dominance may not fully comport with economic notions. For example, Carlton & Perloff state:

It is common practice to say that whenever a firm can profitably set its price above its marginal cost without making a loss, it has monopoly power or market power. One might usefully distinguish between the terms by using monopoly power to describe a firm that makes a profit if it sets its price optimally above its marginal cost, and market power to describe a firm that earns only the competitive profit when it sets its price optimally above its marginal cost. However, people do not always make this
conclusion is consistent with the way U.S. courts as well as U.S. and international antitrust agencies evaluate general structural evidence. Of course, such structural evidence is a
distinction, and generally use the two terms interchangeably, sometimes creating confusion.

DENNIS CARLTON & JEFFREY PERLOFF, MODERN INDUSTRIAL ORGANIZATION 93 (2005)

2 Discussions of the requisite market share required for monopoly power or dominance commonly begin with Judge Hand's statement in United States v. Aluminum Co. of America that while a market share of ninety percent "is enough to constitute a monopoly[,] it is doubtful whether sixty or sixty-four percent would be enough; and certainly thirty-three per cent is not." 148 F.2d 416, 424 (2d Cir. 1945). Following Alcoa and American Tobacco, U.S. courts typically have required a dominant market share before inferring the existence of monopoly power. The Fifth Circuit observed, "monopolization is rarely found when the defendant's share of the relevant market is below 70%." Exxon Corp. v. Berwick Bay Real Estate Partners, 748 F.2d 937, 940 (5th Cir. 1984) (per curiam). Similarly, the Tenth Circuit noted that to establish "monopoly power, lower courts generally require a minimum market share of between 70% and 80%." Colo. Interstate Gas Co. v. Natural Gas Pipeline Co. of Am., 885 F.2d 683, 694 n.18 (10th Cir. 1989) (citation omitted). Likewise, the Third Circuit stated that "a share significantly larger than 55% has been required to establish prima facie market power" and held that a market share between seventy-five percent and eighty percent of sales is "more than adequate to establish a prima facie case of power." United States v. Dentsply Int'l, Inc., 399 F.3d 181, 187 (3d Cir. 2005).

On the other hand, the European Commission's Tenth Report on Competition indicated a dominant position can generally be said to exist once a market share on the order of forty to
rebuttable presumption. Conclusions about monopoly power or dominance can be made only after evaluating possible offsetting considerations and justifications, such as entry, and transaction or practice-specific efficiencies, such as innovation, among others.

This article discusses an additional reason that caution may be warranted before drawing conclusions about dominance or market power from share evidence alone. It is argued that in markets with relatively few transactions (that is, “thinly” traded markets), the implications of observed high market shares are much less clear than in more thickly traded markets.\(^3\) Simply forty-five percent is reached. It also emphasized the general unevenness in shares in reaching this conclusion:

Although this share does not in itself automatically give control of the market, if there are large gaps between the position of the firm concerned and those of its closest competitors and also other factors likely to place it at an advantage as regards competition, a dominant position may well exist.


\(^3\) There are numerous industries that are of antitrust interest that are thinly traded, including defense contracting, aerospace, school milk contracting, public works contracting, and various commodities (agricultural as well as chemicals, metals and other industrial supplies), although this may be affected by how markets are defined. One example of a recent case where issues of market thinness may have been relevant is *Race Tires Am., Inc. v. Hoosier Racing Tire Corp.*, 614 F.3d 57 (3d Cir. 2010). This case involved exclusive equipment requirements adopted by motor sports sanctioning bodies. The market involved “tires for dirt oval track racing in the
put, it is entirely possible that situations that appear to implicate a dominant firm actually reflect
pure random chance in a competitive process involving equally matched or nondominant firms. In thinly traded markets with firms that are equally matched in every way, it would not be unusual to observe one firm with a market share exceeding generally accepted thresholds for monopoly power or dominance, even though these observed market shares are a result of a stochastic outcome.\(^4\) The goal of this article is to discuss methods and theory to distinguish

\begin{quote}
United States and Canada.” The sanctioning bodies at issue in this market required that racers use a specific tire type and brand in all of their races for the season. There were three major manufacturers and three major “customers” which were the sanctioning bodies. The plaintiff alleged that between 2003 and 2007 the sanctioning companies had requested competitive bids only seven times. The defendant, Hoosier Tire, had a seventy-nine percent share of the market for dirt track racing tires. The court found that notwithstanding the defendant’s high market share, the sanctioning bodies’ rule “creates more exciting races, ensures equal access to a uniform product, tends to increase safety, and lowers the cost of tires” and thus was procompetitive. \textit{Id.} at 81. We would note, that despite the court’s ruling that the exclusive dealing arrangement was procompetitive on other bases, the thinness of the market at issue suggests the court would have had reason to be cautious in drawing the conclusion that the defendant was a “dominant firm” in this matter.
\end{quote}

\(^4\) Our caveats are similar but not identical to those raised in the literature on Gibrat’s Law See ROBERT GIBRAT, \textit{LES INEGALITES ECONOMIQUES} (1931). Under Gibrat-like processes, the market consists of identical firms, and annual growth for each firm is determined by a random sampling from an identical distribution of growth rates. Despite the fact that firms in such situations would all appear to be equally situated, such markets can nonetheless evolve toward
between outcomes that show strong statistical evidence of dominance versus those that merely reflect the random distribution of winnings among nondominant firms.

The remainder of this article is organized as follows. Section II develops the above contention in the context of a two-firm model with “equally matched” firms. Probability theory is used to demonstrate the likelihood that a firm in such a context can have an observed market share consistent with monopoly power or dominance. The analysis also discusses how this conclusion is directly affected by the “thickness” of the alleged market. Next, we address statistical techniques that can be used to determine when a high share may actually be a sign of dominance. The appropriate tests vary depending on the nature of exogenous information regarding the firm observed to have a “large” share and the relative weight placed on false positives and false negatives. Section III extends the analysis to three- and multifirm cases. Again, the appropriate test depends on the nature of exogenous information regarding the firm observed to have a “large” share and the relative weight placed on false positives and false negatives. Under certain strong assumptions, a simple “one-firm” test can be used. If there are not such strong priors, there are literally an infinite number of tests to measure dominance in these cases. However, it is found that two different assumptions bound the likelihood of a finding of nondominance. Which assumption is most appropriate depends on the policymaker or other stakeholder’s relative concern with (tolerance for) false positives and false negatives. Section IV concludes by summarizing the results and briefly noting the importance of small highly concentrated structures that persist over long periods of time. Thus, what appear to be highly concentrated markets with stable market leaders are nothing more than a reflection of “luck.” See John Sutton, Gibrat’s Legacy, 35 J. ECON. LITERATURE 40 (1997), for a survey of this literature.
sample questions for other issues in industrial organization and, more specifically, antitrust and competition law.

II. TWO-FIRM TEST

A. Testing for dominance

To set ideas, first consider the following stylized duopoly. Assume there are two firms in a well-defined market for widgets: CoggCo and SprocketCo. Suppose that CoggCo has a seventy percent share and SprocketCo has a thirty percent share. A threshold question for competition analysis is whether CoggCo’s seventy percent share is prima facie evidence of monopoly power or dominance. Suppose that in all respects, CoggCo and SprocketCo and their respective products are perfect substitutes, their cost structures are identical, and there are no other economic or technical substitute products available to market participants. The firms and products are equally matched in the sense that customers are universally indifferent between the firms and their respective products, products are undifferentiated, and there are no cost factors motivating variation in prices. Under these assumptions, a customer’s decision to choose a firm is effectively equivalent to a coin toss. Each firm has a fifty percent probability of winning any given bid or contest. However, even with perfectly matched firms, observed wins and losses may, by pure random chance, be quite unequal.

Suppose the firms compete through a simple bidding market mechanism. Further, assume that only ten bidding contests between the evenly matched CoggCo and SprocketCo have taken place. Under this construct, there is a nontrivial probability of observing one firm winning 5

5 We later relax this assumption and allow for win probabilities other than fifty percent.

6 Although we assume the mechanism of competition is through a simple bidding market, our results do not depend on competition between the two companies being a bidding situation.
a substantial majority of the contests. In fact, even though, by design, the two firms are evenly matched (analogous to heads and tails on a fair coin), there is more than a thirty-four percent chance of observing one firm winning at least seventy percent of the contests (just as there is a thirty-four percent chance that at least seven out of ten flips of a coin will be all heads or tails).  

Under these simplifying assumptions, in ten contests, the perfectly evenly matched firms have a substantial probability of a highly uneven (that is, asymmetric) market outcome. Even though by construct the firms are evenly matched, there is a substantial likelihood that the observed seventy

\[ \text{The probability of observing one firm with a seventy percent market share can be calculated using the binomial distribution (each contest, that is, trial or coin flip, is independent and occurs with replacement). Suppose } A \text{ and } B \text{ are the observed shares of firms } A \text{ and } B, \text{ respectively. Each firm has a fifty percent chance of winning a given contest. The probability that at least one firm obtains a seventy percent share in ten such contests is the probability that } \text{either firm } A \text{ or firm } B \text{ obtains a seventy percent share: } P(A \geq 0.7 \cup B \geq 0.7) = P(A \geq 0.7) + P(B \geq 0.7), \]

where the equality follows because of mutual exclusivity (that is, firms } A \text{ and } B \text{ cannot both obtain more than a seventy percent share). Because there are ten contests, this probability can be rewritten as

\[ P(N_A \geq 7) + P(N_B \geq 7) = P(N_A = 7) + P(N_A = 8) + P(N_A = 9) + P(N_A = 10) + P(N_B = 7) + P(N_B = 8) + P(N_B = 9) + P(N_B = 10), \]

where } N_A \text{ and } N_B \text{ are the number of wins of Firms } A \text{ and } B, \text{ respectively. These probabilities can be obtained from the binomial distribution, under the assumption that the probability of success equals } 50\%: P(N_A \geq 7) + P(N_B \geq 7) = 2(0.117188 + 0.043945 + 0.009766 + 0.00097656).

For a more complete exposition of the binomial distribution, see, for example, DENNIS D. WACKERLY, WILLIAM MENDENHALL III & RICHARD L. SCHEAFFER, MATHEMATICAL STATISTICS AND APPLICATIONS 97–109 (2001).
percent market share is the result of a purely random outcome.\textsuperscript{8} As such, in this situation, caution would be warranted in interpreting a high market share as prima facie evidence of dominance. Market shares may not wholly reveal critically important underlying competitive dynamics, even before various offsetting issues, such as entry and efficiencies, are considered.

Such an asymmetric outcome is largely the result of the “small sample” nature of the competition between CoggCo and SprocketCo (that is, the limited number of contests in which the two firms competed). As highlighted below, the problem of random asymmetric outcomes increases as there are fewer contests and decreases as the number of contests increases.\textsuperscript{9}

Given this simple result, a next step is to pursue a more scientific, statistical means to determine whether a firm’s share is indicative of dominance or merely an artifact of random chance. To address this question, begin with the following definitions in a two-firm construct:

\textsuperscript{8} For intuition, consider how likely it is to observe an unequal number of heads and tails in a series of coin flips. With three flips, an even outcome (that is, an equal number of heads and tails) is impossible; the only possible outcome of three tosses is unequal. In a series of four flips, it would surprise few to see either three heads or three tails. With ten flips, though it may seem surprising, there is still more than a one-third chance of seeing seven heads or tails.

\textsuperscript{9} This point follows from the Law of Large Numbers and the Central Limit Theorem. The observed share is technically a sample mean and, as a consequence of the Law of Large Numbers and the Central Limit Theorem, as the number of observations increases (here, the number of contests), the sample mean converges to the true population mean (here, the true win probability, which is fifty percent in this example). See, e.g., ORLEY ASHENFELTER, PHILLIP LEVINE & DAVID ZIMMERMAN, STATISTICS AND ECONOMETRICS: METHODS AND APPLICATIONS 109–110 (2002), and WACKERLY, MENDENHALL & SCHEAFFER, supra note 7, at 346–54 & 423.
Definition 1—The true win probability is the chance that a firm has of winning any given contest. For example, if customers are indifferent between the two firms, the true win probability of each is fifty percent.

Definition 2—A firm is weakly dominant if it has a marginal advantage, that is, its true win probability is epsilon more than fifty percent (where epsilon is defined as some “very small number”).

This is not to contend that a firm that has slightly more than a fifty percent chance of winning is dominant in a legal sense. Rather, the definition serves as the weakest possible illustration of dominance (with two firms). The definition of weak dominance is used to ground the baseline analytical framework. Concepts that address stronger definitions of dominance are considered below and parallel the fifty percent threshold.

Under the proposed weak dominance definition, the relevant competition question can be rephrased: When can statistical methods based on observed shares be considered reliable evidence that one firm has more than a fifty percent chance of winning? What is an appropriate statistical test for weak dominance?

Hypothesis testing provides a scientifically accepted method. Hypothesis testing involves setting up a null hypothesis and seeking evidence of whether, based on the observed data, the proposed null hypothesis can be rejected with a particular degree of statistical certainty. If the null is rejected, it is said to be rejected in favor of an alternative hypothesis.

---

With the definition of weak dominance developed here, the null is that neither firm is
dominant, that is, they are “equal competitors.” The alternative is that at least one firm has an
edge. Heuristically, the null and alternative hypothesis can be written as follows:

\[ H_0: \text{The two firms are equal competitors} \]
\[ H_a: \text{The two firms are not equal competitors} \]

where \( H_0 \) and \( H_a \) are the notations that indicate the null and alternative hypotheses, respectively.
The null is assumed true unless there is statistical evidence to the contrary; this is akin to an
“innocent until proven guilty” standard.

Given our definition of weak dominance, the null is that each firm has a fifty percent
chance of winning:

\[ H_0: p_A = p_B = 0.5 \]

where \( p_A \) and \( p_B \) represent the probability that firms A and B, respectively, win any given
contract. In the proposed example, if both firms have a fifty percent chance of winning, neither
would be considered dominant. The firms would be equal competitors. The alternative
hypothesis is that the firms are not equal competitors:

\[ H_a: p_A \neq p_B \neq 0.5 \]

Under the alternative hypothesis (and using the definition of weak dominance proposed above),
at least one firm has an edge in winning a contract, that is, a greater than fifty percent chance of
winning.\(^{11}\)

Choosing between \( H_0 \) and \( H_a \) depends, in part, on the \( p \)-value (also known as the attained
level of significance). The intuition for the \( p \)-value is as follows. Consider the CoggCo and

\(^{11}\) This particular test is agnostic as to whether it is firm \( A \) or firm \( B \) that is dominant under \( H_a \).

This issue is further discussed below.
SprocketCo example above. It was assumed that CoggCo has a seventy percent observed market share. In this example, the \( p \)-value is the probability of observing a seventy percent or greater share in ten contests if the null hypothesis were true (that is, \textit{if each firm actually had a fifty percent chance of winning any given contract}). In other words, even if customers are completely indifferent between the two firms, as in the ten flips of a coin example, there would be some calculable chance (probability) that one firm would win seventy percent of the contests purely at random. This probability is the \( p \)-value. In the CoggCo and SprocketCo example, there is a thirty-four percent probability of observing at least one firm with a seventy percent or greater share. Hence, the \( p \)-value in this example is thirty four percent.

Ultimately, the hypothesis test boils down to using the \( p \)-value to decide which of the following two conclusions is most sensible:

(a) \textit{Not reject} \( H_0 \): Given the observed market shares and the number of contests, there is insufficient statistical evidence to reject the hypothesis that the two firms are, in fact, equally matched; or

or,

(b) \textit{Reject} \( H_0 \) \textit{in favor of} \( H_a \): Obtaining the observed market shares would be so improbable (unlikely) if the firms were truly equally matched, that it is more supportable through statistical evidence to conclude the firms are not equally matched.

The \( p \)-value is the likelihood of obtaining the observed share outcome if the null were true; therefore, the \( p \)-value serves as a natural way to choose between (a) and (b). A “large” \( p \)-value means that it is highly likely that, even if the firms are equally matched (that is, conclusion (a) above is true), one would observe the unequal pattern of wins that was actually observed.
Therefore, conclusion (a) is chosen as most plausible, and the null hypothesis of equally matched firms is not rejected. Alternatively, a “small” \( p \)-value means that if the firms are equally matched (that is, conclusion (b) is true), it would be very unlikely to observe the unequal pattern of wins that actually occurred; in this case, the analyst rejects \( H_0 \) in favor of \( H_a \), that is, in favor of conclusion (b).

The threshold question is then, what is the appropriate cutoff for deciding whether the \( p \)-value is “small” or “large”? Is the thirty-four percent (calculated) \( p \)-value in the CoggCo-SprocketCo example large enough to conclude that there is insufficient evidence of a dominant firm, or is it so small that we can reject the hypothesis that these are actually equally matched firms? The cutoff for deciding between the null (conclusion (a)) and alternative (conclusion (b)) hypotheses is called the significance level. If the calculated \( p \)-value (attained level of significance) is larger than the chosen threshold level of significance, then the null hypothesis is chosen. If the \( p \)-value is below the significance level, then \( H_0 \) is rejected in favor of \( H_a \). In effect, the significance level is the bound below which one concludes that the probability of obtaining the observed share is so implausibly small that there is statistical evidence to conclude that \( H_0 \) is not true.

There are three commonly accepted levels of significance: 1%, 5%, and 10%.\(^{12}\) A lower significance level implies more certainty that \( H_0 \) is not true before rejecting it in favor of \( H_a \). In the CoggCo and SprocketCo example, with a \( p \)-value of 0.34, there is insufficient statistical support to reject \( H_0 \) at any of these conventional levels. The observed shares, in and of

themselves, do not provide sufficient statistical evidence of weak dominance as defined above at any standard level of significance. The seventy percent market share, given the number of contests, is likely to be an artifact of a random outcome.

The general intuition of when one might reject $H_0$ is informative. For illustration, suppose once again that a seventy percent share for one firm is observed and the null hypothesis is that the firms were in fact equally matched. However, suppose that in this market, after observing a certain number of contest outcomes and performing the appropriate statistical test, the $p$-value were calculated to be 0.0000001. This means that, if the firms were actually equally matched (as under the null hypothesis), there would be only a 0.000001% chance of observing a firm with a seventy percent or higher share. With this calculated $p$-value, it is virtually impossible that, if the firms were truly equally matched, a seventy percent share would be observed. It is statistically supportable (and more reasonable) to conclude that at least one of the firms has an edge, and there is scientific support for rejecting $H_0$ in favor of $H_a$.

B. Errors in hypothesis testing

In general, statistical methods are applied when the truth is unknown because the researcher cannot know with certainty whether conclusions drawn from the hypothesis test “got it right.” The methods only provide statistical support for how probable it is that a test’s conclusion is erroneous. There are two types of errors that can arise in hypothesis testing: Type I and Type II errors:\footnote{See Wackerly, Mendenhal & Scheaffer, supra note 7, at 482–83.}
Type I: The null hypothesis is true but is incorrectly rejected (also called a false positive—the researcher has incorrectly rejected the null and accepted the alternative hypothesis).

Type II: The null hypothesis is false but is not rejected (also called a false negative—the researcher has incorrectly decided in favor of the null and rejected the alternative hypothesis).

The chosen significance level is the probability of a Type I error. In our example, it is the probability that one would conclude dominance when the firms are actually equally matched. For example, with a 5% significance level, there is a five percent chance of rejecting $H_0$ when it is actually true.

The probability of a Type II error is generally unknowable. Its calculation depends on the unknown true win probability. In application, it is most important to be mindful that there is an inherent tradeoff between Type I and II errors. Reducing the possibility of a Type I error (that is, choosing a low significance level) increases the probability of a Type II error (that is, a false negative—concluding $H_0$ when it is in fact false). Heuristically, choosing a significance level that seeks to mitigate a false finding of dominance makes it more difficult to reject $H_0$ and thus increases the probability of failing to conclude dominance when a firm in fact is dominant.

Although there is a tradeoff in Type I and Type II errors, one cannot in general choose a level of significance that determines the probability of Type II error; that is, one cannot, for example, choose a level of significance that allows only a five percent chance of erroneously concluding a dominant firm is nondominant. This is because the probability of Type II error depends on the unknown true win probability.
Observing more contests mitigates the chance for errors in hypothesis testing. The chance of a Type II error can be directly reduced by increasing the sample size. Strictly speaking, a Type I error is chosen by the experimenter and hence not directly reduced with a greater sample size. Nevertheless, more certainty is gained. The $p$-value can be interpreted as the probability of a Type I error anytime $H_0$ is rejected at the observed share level. (By definition, the $p$-value is the probability of the observed outcome if $H_0$ is actually true.) For example, in the ten-contest, seventy percent observed-share SprocketCo-CoggCo example above, an antitrust agency or other adjudicator that concluded dominance would be wrong 34% of the time under the initial assumption that the firms are equally matched. However, if the same seventy percent share is observed, but there were twenty-five contests, then the adjudicator that concluded dominance at that observed share would be wrong only 4.3% of the time.\textsuperscript{14}

C. Choosing the appropriate level of statistical significance

An essential choice when implementing the hypothesis test methods described above is choosing an appropriate level of statistical significance. Although this choice is somewhat of an art form, as noted above, there are three conventional levels of statistical significance generally accepted by statisticians and economists: 1%, 5% and 10%.\textsuperscript{15} Policymakers or other

\textsuperscript{14}With 25 contests, it is impossible for one firm to have exactly a seventy percent share (that is, 17.5 contest wins). If a firm wins 17 contests, its share is 68%; if it wins 18, the share is 72%. The probability of an observed share of 70% or more is the probability of observing one firm win 18 or more contests, which can be calculated with the binomial distribution.

\textsuperscript{15}Dominance or monopoly power is generally decided on a “more probable than not” standard, which is often interpreted to mean there is a greater than fifty percent probability that this
element of proof applies. See David Kaye, *Naked Statistical Evidence*, 89 Yale L.J. 601, 603 (1980). Practitioners may then question whether the conventional statistical levels mentioned above are relevant in this context. See, e.g., Michael D. Green, D. Michal Freedman & Leon Gordis, *Reference Guide on Epidemiology*, in *Reference Manual on Scientific Evidence*, supra note 10, at 358 n.67 (“A common error made by lawyers, judges and academics is to equate the [p-value] with the legal burden of proof.”). There is an extensive literature on the relationship between the scientific and legal standards of proof and their interplay; though a full summary is beyond the scope of this article, we highlight two broad issues.

First, statistical hypothesis testing does not directly calculate whether dominance is more probable than not. Statistical evidence is an input into that finding. Hypothesis testing helps us determine how reliable the observed share evidence is. Whether an observed share should be an input into a judge or jury’s finding of dominance is a function of the chosen levels of statistical significance.

Somewhat more technically, the preponderance standard and the p-value refer to two different conditional probabilities. Under this assumption, the preponderance of evidence is the probability that a firm is dominant, given the firm’s observed share:

\[ P(\text{Firm is Dominant}|\text{Observed share}) \].

The p-value considers a different conditional probability:

\[ P(\text{Observed share}|\text{Null Hypothesis}) \].

There is no easy relationship between the two. See, e.g., David Kaye, *Statistical Significance and the Burden of Persuasion*, 46 Law and Contemporary Problems 22 (1983). Stated another way, the p-value does not specify the likelihood that the firm is dominant. Rather, the p-value identifies the probability that an observed share above the dominance threshold arose
stakeholders can choose among these conventional levels of statistical significance or any other based on their own priors and tolerance regarding the importance of Type I or Type II errors. If policymakers are more concerned with Type I errors (false positives), then lower levels of statistical significance may be appropriate. If policymakers are more concerned with Type II because of random chance, assuming the firm’s true win probability is below the dominance threshold. If this probability is small (below the significance level), we conclude the hypothesis that its true win probability is below the dominance threshold is implausible; that is, the hypothesis that its true win probability is above the dominance threshold is more plausible.

In fact, classical statistics does not even allow one to define the probability that a firm is dominant. The reason is that dominance is defined as having a true win probability surpassing a specified threshold. Classical statistics treats this true win probability as a fixed but unknown value to be estimated, not as a random variable. Since the true win probability is not random, neither is the status (dominant versus nondominant) of the firm. Being nonrandom, there is no probability associated with the status of the firm. For further discussion of this issue, see, for example, Michael D. Green, *Science Is to Law as the Burden of Proof Is to Significance Testing*, 37 Jurimetrics 205, 222 (1995).

Second, it is important to note that if more than one element of evidence is subject to a hypothesis test, the overall significance level of whether the body of evidence is likely reliable is not the same as each of the individual significance levels. The overall level of significance on the overall body of evidence is lower than the individual significance levels (and its exact calculation can be quite complicated).

errors (false negatives), then an appropriate policy prescription appeals to higher levels of significance as the relevant thresholds.\textsuperscript{17}

To the extent there is litigation or some other adversarial situation involving dominance issues, it may be advisable to choose the significance level in such a way that lends the most credibility to a given party’s position. For example, consider a plaintiff in a litigation context that wishes to assert that a defendant is dominant. A plaintiff would want to reject the null hypothesis of no dominance. The most conservative position for this party would be to choose a lower level of significance. A lower level of significance makes it more likely the null hypothesis of no dominance would be accepted. Thus, a plaintiff could argue that it chose a standard unfavorable to a finding of dominance, but found dominance nonetheless. For defendants, choosing a higher level of significance may be appropriate for analogous strategic reasons.

\textit{D. Other thresholds of dominance}

The method discussed above can be used to test a range of alternative thresholds for dominance. We generically define dominance threshold as follows:

\begin{itemize}
\item The conventional wisdom is that European competition authorities tend to be more concerned with false negatives and U.S. authorities more concerned with false positives. For more on this issue and possible reasons why this may be the case, see Seth Sacher, The Past, Present and Future of Antitrust 11 GEO. J. INT’L AFFAIRS 115, 118–19 (2010). Below we note that our findings suggest this may not be a strictly correct characterization of European competition law, at least with respect to issues of dominance.
\end{itemize}
Definition 3—The dominance threshold is the cutoff-probability $\theta$ such that if a firm’s true probability of winning a given contract is anything greater than $\theta$, the firm is deemed dominant.\textsuperscript{18}

Suppose, for example, that dominance is a concern only if a firm has more than a sixty percent chance of winning a given contest. Then, the dominance threshold is $\theta = 0.6$, and the relevant null for the hypothesis test is as follows:

$$H_0: p_A = 0.6$$

The above $H_0$ necessarily implies that $p_B = 0.4$.\textsuperscript{19} The alternative hypothesis is that firm $A$ has greater than a sixty percent win probability and is therefore dominant:

\textsuperscript{18} The true probability of winning a given contest can also be called an “objective” probability or the true theoretical win probability in the following sense. It is a firm’s unknowable probability of winning a given contest in the absence of any randomness (in effect this is a latent variable). For example, suppose a firm has a true win probability of fifty percent. Then this would be the same as a fair coin, which has a true probability of being heads fifty percent of the time even though that outcome may not actually occur in a number of contests due to randomness. In the two-firm construct, testing for a true win probability other than fifty percent relaxes the assumptions of homogeneous, equally preferred products by firms with identical cost structures that were made in the CoggCo-SprocketCo example with a fifty percent win probability. Thus, considerations such as first-mover advantages, incumbencies, and cost differences are implicitly accounted for by such a change.

\textsuperscript{19} The null is written as an equality, but it implicitly incorporates the possibility that $p_A < 0.6$. The null is written as an equality because the relevant statistical methods only allow testing a null hypothesis that is an equality. The hypothesis test determines if there is statistical evidence
$H_a: p_A > 0.6$

Table 1: Two-Firm Test Critical Win Numbers
The minimum number of wins to conclude dominance

<table>
<thead>
<tr>
<th>Contests</th>
<th>Dominance Threshold</th>
<th>Level of Statistical Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10%</td>
</tr>
<tr>
<td>10</td>
<td>&gt; 50%</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>&gt; 55%</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>&gt; 70%</td>
<td>10</td>
</tr>
<tr>
<td>15</td>
<td>&gt; 50%</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>&gt; 55%</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>&gt; 70%</td>
<td>14</td>
</tr>
<tr>
<td>20</td>
<td>&gt; 50%</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>&gt; 55%</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>&gt; 70%</td>
<td>18</td>
</tr>
</tbody>
</table>

× Impossible to reject the null at this significance level.
Note: Critical wins determined by $p$-value calculated as probability that either of two firms achieved the observed share.

More generally, for a dominance threshold of $\theta$, the hypothesis test is

$H_0: p_A = \theta$

$H_a: p_A > \theta$

that the true win probability is larger than the largest nondominant win probability. Testing a smaller win probability would serve little benefit. For example, if the cutoff for dominance is a true sixty percent win probability or greater, and we test and reject (at a particular level of significance) $H_0: p_A = 0.5$, we would not have established dominance. This is because dominance is defined as a win probability of more than sixty percent, so to establish dominance we would still have to test the null of a larger win probability.
Assuming the null is that no firm exceeds the dominance threshold, table 1 displays the minimum number of observed wins necessary to conclude, at standard levels of significance, that a firm in a duopoly is dominant, for a given number of contests and a given dominance threshold.\textsuperscript{20} (Appendix A contains a more complete list of critical win probabilities; common software can also be used to implement the relevant test.\textsuperscript{21})

\textsuperscript{20} Inevitably, some firms that are dominant, as defined by having a true win probability surpassing the dominance threshold, will by random chance have an observed share below the threshold and appear nondominant. For example, assume the dominance threshold is seventy percent and a firm has a true win probability of seventy-one percent. It will sometimes be the case that such firms, due to random chance, will have an observed share below the seventy percent threshold.

There are two ways to deal with this possibility. First, one can set up the null hypothesis to assume the firm has a dominant share and then test whether the null can be rejected in favor of nondominance. In this case, a firm would be dominant if its true win probability $p$ is equal to (or larger than) the threshold $\theta_d$. The null hypothesis would be $H_0: p = \theta_d$, and the firm would be assumed dominant under the null. The alternative would be $H_a: p < \theta_d$. A conclusion of nondominance would be reached only with sufficient statistical evidence of a true win probability below the dominance threshold. Unfortunately, this methodology is akin to a guilty-until-proven-innocent standard, because the firm is presumed under the null to be dominant unless the statistical evidence is sufficiently strong to conclude nondominance.

The other approach is to use the observed share to formulate a confidence interval (CI) for the true win probability. The downside is that a CI cannot establish that a firm with an observed share below the threshold is dominant; the utility of this approach would be in gauging
In the spirit of the term critical value used in hypothesis testing, the minimum number of wins to conclude dominance is called the critical win number. Note that as the number of contests increases, the percentage of wins needed for concluding dominance decreases for a given dominance threshold.

Definition 4—The critical win number is the minimum number of wins needed to conclude dominance for a given number of contests.

how much information the observed share conveys. A wide CI implies little can be said about the true win probability, whereas a narrow CI allows a degree of confidence that the firm is indeed not dominant. See infra note 34 for more on confidence intervals.

21 In general, the following Excel command can be used to calculate the p-value, which can then be compared to the chosen level of significance:

\[ p = 2 \times \text{BINOM.DIST}(W - 1, n, \theta, 1) - \text{BINOM.DIST}(W - 1, n, 1 - \theta, 1) \]

where

\[ W = \text{the observed number of wins of the firm observed to have a majority share}; \]
\[ n = \text{the number of contests}; \]
\[ \theta = \text{the dominance threshold}. \]

22 Hypothesis tests use a critical value, which is a value on a distribution with a probability equal to the significance level (typically located in the tail of the distribution). If the observed outcome (usually standardized to have the same mean and variance as the relevant distribution) is greater than the critical value, the null hypothesis being tested is rejected. Similarly, any observed win outcome larger than the “critical win number” rejects \( H_0 \).
III. ONE-FIRM OR TWO-FIRM TEST? THE IMPORTANCE OF EXOGENOUS INFORMATION

In testing for dominance, one needs to draw upon exogenous information in order to draw policy-relevant conclusions. In this section, we discuss this consideration. Simply put, a firm that has a minority win probability can, by sheer luck, win a majority of contracts.\textsuperscript{23} Therefore, if the null hypothesis that no firm is dominant is rejected, one needs to draw upon exogenous information before concluding that the majority-observed firm is in fact dominant. Exogenous information could include one firm having a superior cost structure or a first-mover advantage. However, such evidence is generally not obvious. For example, cost data is notoriously difficult to deal with (as evidenced by predatory pricing cases) and evidence regarding a first mover advantage may largely be subjective and very difficult to prove.\textsuperscript{24}

\textsuperscript{23} As an intuitive example, suppose forty names are placed in a hat: nineteen males and twenty-one females. Five names are selected at random. There is a nontrivial chance that the sample drawn will contain a majority of males even though males have a minority chance of being selected.

\textsuperscript{24} We would also note that in a dominance or monopolization case, the practices at issue are often the claimed cause of dominance, such as exclusive dealing, loyalty discounts, certain types of tying arrangements, or conditional rebates. One might want to consider these practices as possible reasons for an a priori expectation that majority-observed firm could never have a minority win probability. However, we would caution against this to the extent that firms that are not being accused of dominance engage in such practices as well. For example, the Federal Trade Commission investigated allegations that Google’s practice of “universal search” was an abuse of a dominant position, but other, smaller search engines implemented the same
To illustrate the implications of exogenous information, assume two-firms. Further, assume for simplicity that the dominance threshold is an unrealistic fifty-one percent. With this threshold, the hypothesis test for dominance is:

\[ H_0: p_A = 0.51 \]
\[ H_a: p_A > 0.51 \]

The null implicitly implies \( p_B = 0.49 \) and the alternative hypothesis implies \( p_B < 0.49 \).

Define the firm observed to have the larger share as the majority-observed firm. Suppose the majority-observed firm realizes eight wins out of ten contests. Under \( H_0 \), there is a 6.2% chance of observing firm \( A \) with an eighty percent or greater share. However, under \( H_0 \) it is almost as likely to observe firm \( B \) with an eighty percent share, a probability of 4.8%. In other words, based solely on the observed data, there is little statistical evidence that the majority-observed firm is the one with the higher true win probability\(^{25} \); the majority-observed firm may very well be the firm with the lower true win probability.

---

\(^{25}\) See supra note 18 for a discussion of how we define true or objective win probability.
This consideration is important to calculating the $p$-value and hence to conclusions derived from the hypothesis test. The $p$-value in this case (that is, the attained level of significance) is the probability of observing an eighty percent or greater share if the null hypothesis is actually true. This probability can be calculated according to either of two approaches:

- **One-firm test:** Assume it is impossible that firm $B$ won a majority of contests. Then, what is the probability that firm $A$ gets an eighty percent or larger share if the null is true? or

- **Two-firm test:** What is the probability that either firm $A$ or firm $B$ gets an eighty percent share or larger if the null is true?

The two-firm test is the method applied in the previous section; its $p$-value is the probability that either firm—$A$ or $B$—wins eighty percent or more of the contests if $H_0$ is actually true. In other words, the two-firm test acknowledges the possibility that firm $B$, with less than a fifty percent chance of winning a given contest, “got lucky” and won the majority of contests. This $p$-value is the sum of the two probabilities calculated above, which is 11%. Under this approach to calculating the $p$-value, at a 10% significance level, one would *not* reject $H_0$ in favor of a dominance finding. However, the one-firm test explicitly rules out the possibility that firm $B$ could have won a majority of the contests. Therefore, the $p$-value under the one-firm test is the probability that firm $A$ won eighty percent or more of the contests, which is 6.2%. With this $p$-value, $H_0$ is rejected in favor of the dominance finding at a 10% significance level.

The one-firm test dismisses entirely (in fact does not even consider) the possibility that firm $B$ could have obtained the observed majority share, hence the categorization as a “one-firm” test: Only firm $A$ could be the majority-observed firm. The two-firm test acknowledges the
possibility that either firm A or firm B (with just shy of a fifty percent chance of winning a given contest) might have won the majority of contests (that is, a two-firm test because there are two firms that are potentially the majority-observed firm).

As demonstrated in the example above, calculating the \( p \)-value with the one-firm test results in lower \( p \)-values and hence makes the observer more likely to reject \( H_0 \). Table 2 illustrates this point. The table presents the critical win numbers for various dominance thresholds, numbers of contests, and significance levels, using the one-firm test. (See appendix A for a list of critical win numbers, through thirty contests, for fifty, sixty, and seventy percent dominance thresholds. Common software can also be used to implement the hypothesis test.\(^26\))

\(^26\) In general, the following Excel command can be used to calculate the one-firm test’s \( p \)-value and then compare it to the desired significance level:

\[
=1-\text{BINOM.DIST}(W - 1 , n , \theta , 1)
\]

with

\( W = \) the observed number of wins of the majority-observed firm;

\( n = \) the number of contests;

\( \theta = \) the dominance threshold.
**Table 2: One-Firm Test Critical Win Numbers**
The minimum number of wins to conclude dominance

<table>
<thead>
<tr>
<th>Contests</th>
<th>Dominance Threshold</th>
<th>Level of Statistical Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10%</td>
</tr>
<tr>
<td>10</td>
<td>&gt; 50%</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>&gt; 55%</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>&gt; 70%</td>
<td>10</td>
</tr>
<tr>
<td>15</td>
<td>&gt; 50%</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>&gt; 55%</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>&gt; 70%</td>
<td>14</td>
</tr>
<tr>
<td>20</td>
<td>&gt; 50%</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>&gt; 55%</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>&gt; 70%</td>
<td>18</td>
</tr>
</tbody>
</table>

*Impossible to reject the null at this significance level. Note: Critical wins determined by p-value calculated as probability assuming it is impossible that Firm B could have won majority of contests.*

Table 2 can be compared to table 1, in which the critical win numbers for the same thresholds, numbers of contests, and significance levels were calculated using p-values under the two-firm test.

The number of wins needed to conclude dominance for a given threshold-contest-significance combination is often lower and never higher in table 2 than in table 1. This reflects the smaller p-value obtained under the one-firm test. Nevertheless, the issue tends only to matter for dominance thresholds near fifty percent. As seen in tables 1 and 2, for a dominance threshold probability of sixty percent or greater, there is almost no discernible difference between the two
approaches in terms of the critical win numbers. This reflects the fact that under higher dominance thresholds, $p_A$ is larger and $p_B$ is smaller under the null, which implies that it becomes less likely that firm B will achieve a majority-observed share.

The fact that the two-firm test requires no assumption regarding whether the majority-observed firm is the firm with the higher true win probability or the firm with the lower true win probability comes at a cost. The logic in the two-firm test implies that, even when rejecting $H_0$ in favor of $H_a$, there remains some possibility that the majority-observed firm is in fact the minority true win probability firm B. In other words, rejecting $H_0$ does not necessarily lead to the conclusion that the majority-observed firm is the one with the higher true win probability, so much as it suggests that there is some firm that exceeds the dominance threshold. Therefore, to draw policy conclusions under the two-firm test, as with the one-firm version, an exogenous reason must be used to support the belief that the majority-observed firm is dominant. If the null hypothesis is rejected under the two-firm test, the analyst must rely on alternative information regarding which firm is dominant in order to conclude dominance is even possible. With the one-firm test, the assumption is embedded in the hypothesis test; whereas with the two-firm test conclusions about which firm is dominant are made after the null hypothesis is rejected.

---

27 The differences in the $p$-values calculated under the logic in the one- and two-firm tests, and hence the difference in the critical win numbers, are smaller for higher threshold definitions of dominance. For example, with ten contests and an observed share of seventy percent, with dominance defined according to $H_a: p_A > 0.5$, the $p$-values are 0.17 and 0.34, using the logic in the one-firm and two-firm tests, respectively. With $H_a: p_A > 0.6$, the analogous $p$-values are 0.382 and 0.437, respectively.
While both approaches require exogenous information in order to draw conclusions about which firm is dominant, there are reasons to prefer the two-firm test. First, the one-firm test, by requiring an assumption about which firm is dominant before implementing the hypothesis test appears to be at odds with the logic behind the null hypothesis, that is, that bids are won randomly with a specified win probability.

Second, it would appear the two approaches differ in situations where the difference matters most. When the threshold is low, the threat of a false positive is greatest; it is here that the two-firm test is less likely to result in a false positive because it has the relatively lower \( p \)-value. At higher thresholds, the difference tends to disappear. When the dominance threshold is high, there is less likelihood of a false positive and the tests result in similar \( p \)-values. We would also note that the proposed analysis can function as a “screen” in small sample situations. The two-firm test can be implemented before resources are invested in determining the importance of such exogenous considerations. It is only after the hypothesis of no dominance is rejected that an investment in resources to determine whether the observed majority firm is in fact dominant need be made.

Although from a practical standpoint it may not always matter much whether the one-firm or two-firm test is chosen, in general the two-firm version may be more appropriate since it accounts for the higher risk of false positives when the threshold is low and considers the possibility the majority-observed firm could be either the firm with the higher true win probability or the firm with the lower true win probability. This belief should not discount reasons for choosing the one-firm test. If there are concerns with false negatives there might be legitimate objections to the use of the two-firm test. However, there are alternative and arguably more appropriate ways to adjust the analysis to account for concerns regarding false positives.
and false negatives. Nevertheless, as noted above, there may be strong exogenous reasons to
assert that the majority-observed firm could never be the firm with the lower true win
probability.

B. Implicit trade-off in Type I and II Error from the one-firm versus two-firm test

Decisions regarding the format of a hypothesis test create implicit choices about errors in
hypothesis testing. This tradeoff occurs when choosing between the one-firm and two-firm tests,
as well. Because, for the same observed shares, the two-firm test yields larger $p$-values than the
one-firm test, the former implicitly increases the chance of a Type II-like\textsuperscript{28} error: failing to
conclude dominance when in fact a firm is dominant. On the other hand, the one-firm test, with
its smaller $p$-value, is more likely to have a Type I-like error: concluding dominance when in
fact none exists.

IV. THREE FIRMS

A. The one-firm test applied to three or more firms: Strong priors on which firm achieved a
majority-observed share

\textsuperscript{28} Strictly speaking, Type I and Type II errors refer to errors in a specific null hypothesis.

Because the one-firm and two-firm tests have different null hypotheses, the tradeoff between the
two tests cannot be described strictly as one between Type I and Type II error. The concepts are
similar, though, in that heuristically the hypotheses of interest are the null of “nondominance”
versus the alternative of “dominance.” With this intuitive treatment of the hypotheses, Type I
and Type II error have the same intuitive, heuristic meaning as with formally specified null and
alternative hypotheses.
In section III we discussed the need for exogenous information regarding which firm could have won a majority of contests in order to draw policy-relevant conclusions in a market with two firms (that is, whether it is the firm with the higher true win probability or the firm with the lower true win probability). It was stated that if there is sufficient information to assert it is highly improbable for a firm with a minority true win probability to win a majority of contests, the analysis should implement the one-firm test. The same logic applies to a market construct with three or more firms: If the analyst dismisses all possibility that a firm with a minority chance of winning could have gotten lucky and won a majority of contests, she should apply the one-firm test described in section II. The analysis is considerably more complicated when there is not such strong evidence. The remainder of this section discusses testing for dominance without such strong priors.

B. Bounding p-values with three firms

With three firms, there are infinite null and alternative hypotheses. In this subsection, we describe how choosing among them can have a fundamental impact on scientifically supportable outcomes.

Suppose there are three firms: A, B, and C. The complication is, given a dominance threshold $\theta$, it is ambiguous how to assign true win probabilities to firms B and C under the null hypothesis. For example, suppose the dominance threshold is fifty percent; then, under $H_0$ firm A is assumed to have a fifty percent win probability. Given this, the win probabilities for firms B and C might be:

$$p_B = 0.25; p_C = 0.25$$

or
\[ p_B = 0.49; p_C = 0.01 \]

or any other combination of \( p_B \) and \( p_C \) such that \( p_B + p_C = 0.5 \). The choice of \( p_B \) and \( p_C \) in \( H_0 \) can matter to the \( p \)-value and hence to conclusions drawn from the hypothesis test. That is, for the same observed outcome, a different allocation of \( p_B \) and \( p_C \) in the null yields a different \( p \)-value. Each different allocation in the null is, in fact, a different test.

Most generically, the null hypothesis can be written as follows:

\[ H_0: p_A = 0.5; p_B = \varepsilon; p_C = 0.5 - \varepsilon \]

for any \( 0 \leq \varepsilon \leq 0.25 \). Table 3 illustrates how \( \varepsilon \) matters to the \( p \)-value. In all cases, \( p_A = 0.5 \), the number of contests is assumed to be ten, and the observed share is assumed to be eighty percent. As can be seen, the \( p \)-value is largest when \( p_B = 0.5, p_C = 0 \); in other words, the two-firm test applied to three firms maximizes the \( p \)-value. The \( p \)-value decreases as the win probabilities attributable to firms \( B \) and \( C \) converge and become equal. A simple proof shows that \( p_B = p_C = 0.25 \) minimizes the \( p \)-value and \( p_B = .5, p_C = 0 \) maximizes it (see appendix B).

Intuitively, as firms \( B \) and \( C \) get further from a fifty percent chance of winning a given contest, the less likely it is that either one has a chance of “lucky streak” resulting in an observed majority share.

29 This ambiguity does not arise with two firms because the win probability of firm \( A \) determined the win probability of firm \( B \), that is, \( p_B = 1 - p_A \).

30 Technically, \( \varepsilon \) can be as large as 0.5, but, for any \( \varepsilon > 0.25 \), the null is for practical purposes the same, with \( p_B \) and \( p_C \) reversed.

31 The table was calculated with the binomial distribution, using the Excel command presented \textit{infra} note 37.
The implication for a policy prescription is that the choice of the allocation of \( p_B \) and \( p_C \) in \( H_0 \) matters to the results of the test. For example, with

\[
H_0: p_A = 0.5; \ p_B = 0.49; \ p_C = 0.01
\]

the null is not rejected at the 10% significance level (the \( p \)-value equals 0.103). However, under the following construct:

\[
H_0: p_A = 0.5; \ p_B = 0.4; \ p_C = 0.1
\]

the null is rejected at the 10% significance level (\( p \)-value equals 0.067).

It can be demonstrated with numerical analysis (see appendix B) that this result is true for other null hypotheses, observed shares, and numbers of contests. That is, for any given dominance threshold \( \theta \), the \( p \)-value is bounded by the following two formulations of the null hypothesis:

- The \( p \)-value is maximized with the \textit{two-firm test}:

  \[
  H_0: p_A = \theta; \ p_B = 1 - \theta; \ p_C = 0
  \]

- The \( p \)-value is minimized with the \textit{equalized three-firm test}:
\[ H_0: p_A = \theta; \quad p_B = \frac{1 - \theta}{2}; \quad p_C = \frac{1 - \theta}{2} \]

The version of the null that minimizes the \( p \)-value is defined as the “equalized three-firm test.”

Below, basic guidance is provided regarding choice of the \( p \)-value approach for various stakeholders in an antitrust context. First, however, the next subsection discusses the alternative hypothesis in tests with three firms.

**C. The alternative hypothesis with three firms**

A well articulated hypothesis test requires a null and alternative hypothesis. If a particular null hypothesis is rejected (at a specified level of significance), it must be rejected for an alternative hypothesis. Defining an appropriate alternative hypothesis beyond the one-firm construct\(^{32}\) is nontrivial.

First, consider the simple two-firm test applied to three firms. As noted above, this test results in an upper bound on the \( p \)-value. To set ideas, assume the dominance threshold is a true win probability of fifty percent (the logic for any other threshold is the same) and that, in accordance with the two-firm test, firm C’s assumed win probability under the null is zero. Fully specified:

\[ H_0: p_A = 0.5; \quad p_B = 0.5; \quad p_C = 0 \]

\(^{32}\) As discussed in section IV.A, in the presence of strong priors that only one particular firm (called firm A above) could have achieved a majority-observed share, the issues raised in this section are nonapplicable. With such priors, the analysis implicates the one-firm test: With a dominance threshold of \( \theta \), the null hypothesis is \( H_0: p_A = \theta \) and the alternative is \( H_a: p_A > \theta \).
Three equal signs in the null indicates a *compound* test, that is, more than one equality. Under this construct, the null is untrue if any of the equal signs do not hold. As such, there are multiple possible alternative hypotheses, including the following:

\[ H_a: p_A = 0.5; p_B < 0.5; p_C > 0 \]

\[ H_a: p_A > 0.5; p_B < 0.5; 0 \leq p_C < 0.5 \]

For illustration, assume ten contests and an observed eighty percent share (the same stylized fact pattern assumed in table 3.) The appropriate economic question is whether it is sensible to reject \( H_0 \) in favor of \( H_{a^*} \) or \( H_a \). This question is fundamental to correctly characterizing the testable hypothesis: clearly, \( H_{a^*} \) does not provide statistical evidence of dominance—since no firm would have more than a fifty percent share. On the other hand, \( H_a \) does provide such evidence.

Economic logic rules out \( H_{a^*} \) as an alternative. Rejecting \( H_0 \) in favor of \( H_{a^*} \) implies a belief that the true win probabilities described by \( H_{a^*} \) are more plausible than those described by \( H_0 \). Recall that the null hypothesis is rejected whenever the \( p \)-value is less than the significance level. If \( H_0 \) is rejected, its \( p \)-value must be less than the significance level. However, as demonstrated in table 3, all allocations of true win probabilities defined by \( H_{a^*} \) have smaller \( p \)-values than \( H_0 \) and would also be rejected at the same significance level as \( H_0 \). It makes no sense to reject \( H_0 \) in favor of an alternative that would also be rejected at the same significance level. As such no allocation of win probabilities as described by \( H_{a^*} \) can be a viable alternative to \( H_0 \).

Alternatively, \( H_a \) is an economically meaningful alternative to \( H_0 \). The \( p \)-values for the true win probabilities defined by \( H_a \) are all larger than \( H_0 \)’s. All true win probabilities defined by \( H_a \) are more likely to produce the observed outcome than \( H_0 \).

Next, consider the lower-bound construction of the null, the equalized three-firm test:
This null is also compound. As with the upper-bound construct, the only economically meaningful alternatives to this null are sets of win probabilities that are more likely to generate the observed shares than $H_0$ (that is, that produce a larger $p$-value). Under this standard, there are multiple plausible alternatives, including the following:

\[ H_a: p_A > 0.5; \text{and that the win probability of A or B or both is less than 0.25} \]

or

\[ H_a: p_A = 0.5 \text{ with either } (p_B > 0.25 \text{ and } p_C < 0.25) \text{ or } (p_B < 0.25 \text{ and } p_C > 0.25) \]

The first alternative provides evidence of a dominant firm, whereas the second does not. The clear implication is that there are plausible alternatives to $H_0$ that do not provide evidence of dominance. As such, the null hypothesis formulated above yields no unique alternative hypothesis with which to conclude dominance. That is, rejection of this formulation of the null does not provide evidence of dominance. Nevertheless, a hypothesis test based on this construction of the null has its value: nonrejection of $H_0$ suggests a lack of evidence for dominance. We discuss this point further below.

**D. Guidance for practitioners: which null to assume?**

In the relevant settings in which claims about dominance are made, two categories of agents can generally be described: (1) parties presenting before adjudicators (plaintiffs and defendants); and (2) the adjudicators themselves (agencies, judges, tribunals, mediators, and arbitrators). The lines can sometimes be blurred. For example, a competition agency can play the role of a nonadjudicator in an adversarial setting in which it asserts that a particular player or
group of players is dominant. Plaintiffs and defendants can be parties to a formal litigation. However, these roles can also be ascribed to parties appearing before an agency. Thus, a complaining firm or other market participant can be viewed as a plaintiff, while the target firm or firms can be viewed as defendants.

As noted in the two-firm discussion, to the extent claims about dominance take place in an adversarial context, one criterion for choosing an appropriate test is whether it lends credibility to a given party’s position. For plaintiffs or others that wish to assert dominance, this implies using a test that maximizes the $p$-value. As discussed above, this means the plaintiff should choose the two-firm test, which assigns under the null a zero win probability to either firm $B$ or $C$. This is the most “conservative” approach for a party attempting to prove dominance. The two-firm test yields the highest possible $p$-value and therefore it is the least likely to support a claim of dominance. If this test succeeds in rejecting $H_0$ (at a particular level of statistical significance), there is arguably strong statistical evidence to support a dominance claim.

Conversely, for defendants (or others wishing to assert there is not dominance), the test that provides the most credible support for a claim of no dominance would appear to be one that minimizes the $p$-value. This is the equalized three-firm test, which has equal shares distributed across firms $B$ and $C$ as described above. This test produces a test statistic that is most likely to reject the null hypothesis of equal competitors. If such a test is unable to reject the null hypothesis, this would appear to be the most credible stance for defendants (that is, defendants employed a test most likely to reject a hypothesis of no dominance, but nonetheless could not reject that hypothesis).
However, as raised in the previous section, if the null hypothesis is rejected (based on a particular level of statistical significance), what is an appropriate alternative hypothesis? As noted above, the only credible alternative hypotheses are ones with a higher \( p \)-value than the null itself. For the plaintiff, since it is recommend that she choose the null with the highest possible \( p \)-value, the only plausible alternatives are ones that evidence dominance.

Alternatively, for the defendant, it is recommended that she choose the null with the lowest possible \( p \)-value, which is the equalized three-firm test. With this test there are alternatives with a higher \( p \)-value that do not imply dominance. However, in this context this should not be troubling. It is the goal of the defendant to provide evidence calling a finding of dominance into question. If she has chosen a null most adverse to her position, it is of little import that there are alternatives that still support her position.

Now consider the choices confronting adjudicators. Some adjudicators may be most concerned with false positives (that is, rejecting a finding of equal competitors when the rivals are actually equal competitors). As with plaintiffs, such stakeholders may wish to choose the test that maximizes the \( p \)-value, the two-firm test, which assigns a zero win probability to firm \( B \) or \( C \). This makes it more likely to accept the null and, therefore, lowers the likelihood of false positives.

Other adjudicators are more concerned with false negatives (not finding dominance when rivals are actually not equal competitors). It may seem like such stakeholders should implement the test that produces the lowest \( p \)-value, such as the equalized three-firm test described for the defendants above. However, this is incorrect. If the null is rejected, since the adjudicator is going to use this as evidence of dominance, the fact that there can be alternative hypotheses with higher \( p \)-values that still show no dominance is much more troubling than in the defendant
case. In this context, it is recommended that the two-firm test, which produces the highest \( p \)-value, be used and that these adjudicators address their concerns through adjustments to the significance level.

It was noted in footnote 17 that, in part, the European Commission also considers the possibility of asymmetric shares between the firm of concern and its competitors (that is, the larger the gap in shares between the firm of concern and its competitors, the more likely it is that that firm is considered dominant). The pattern regarding the shares attributed to firms \( B \) and \( C \) and the \( p \)-value detailed above provides a possible explanation for this concern.

Policy prescriptions regarding the gap between firm \( A \) and its competitors may actually reflect a concern with false positives. If there is not a large gap between the firm of concern and at least one of its competitors, then the more likely it is that at least one firm has a threshold exceeding the dominance threshold purely by chance (that is, a higher \( p \)-value). On the other hand, the greater this gap, the lower the \( p \)-value and the less likely it is that one firm exceeds the threshold purely by chance. Focusing on the latter situation decreases the likelihood of a false positive (incorrectly condemning a firm as dominant).

While the conventional wisdom holds that European competition authorities are more focused on avoiding false negatives than false positives, the above result suggests that part of the European Commission’s case “screening” methods may point in the opposite direction. Further, since the case screening reduces the chances of false positives, this may support making choices later in the analysis, such as the choice of significance levels and dominance thresholds, that lower the chances of false negatives.

\textit{E. Caution in interpreting nonrejection of }H_0
The null hypothesis assumes no firm is dominant. In that sense, if the null hypothesis were “accepted,” (that is, not rejected at a particular level of significance) this would be evidence of a nondominant position. However, that logic is not precisely correct. Statisticians never accept the null hypothesis. Rather, they say that a hypothesis test “fails to reject $H_0$” at a particular level of significance. In other words, hypothesis testing works in a framework akin to an “innocent until proven guilty” standard. Such a standard can never statistically demonstrate positive evidence that a firm is nondominant. This point raises two important questions: (1) what can be learned if the null hypothesis is not rejected; and (2) is there a statistical method that can provide positive evidence that a firm is not dominant? For the second point, narrow confidence intervals might provide statistical evidence of nondominance, but they are beyond the scope of this article.

With respect to question (1), even though the hypothesis test cannot demonstrate that a firm is not dominant, a nonrejection of $H_0$ is nevertheless useful in itself; a nonrejection reveals how to interpret the data. A firm that appears to have a strong dominant position is statistically indistinguishable from a nondominant firm. That is, not rejecting $H_0$ tells us when there is no scientific justification to conclude dominance from the observed outcomes, however asymmetric they may appear. In effect, in the framework established above, not rejecting may simply mean

---

33 Analogously, in a court case, the choice of verdicts is between guilty and not guilty, not between guilty and innocent.

34 See, for example, WACKERLY, MENDENHAL & SCHEAFFER, supra note 7, at 380–95 and 480–82, for a discussion of the relationship between hypothesis testing procedures and confidence intervals.
the sample size is too small to generate statistically supportable conclusions, which is precisely
the concern being addressed.

There is an important caveat in interpreting a non-rejection of $H_0$ as positive evidence in
favor of $H_0$. Consider a regression measuring a post-merger price increase. An insignificant
result may merely reflect a small sample issue or noisy data; had the analyst obtained a larger
sample the result might not have been insignificant. In a sense, an insignificant result can be the
consequence of an analyst not trying very hard to find good data. Stated bluntly, an
insignificant result can be the outcome of an intentionally selected sample, poor quality analysis,
a misspecified model, a nonincrease in price, or any number of other explanations.

In the context of our analysis, an insignificant result is more informative because, by
construct, all available data are being used; that is, all available win outcomes (the population of
observable outcomes) are included in the analysis. In fact, there is no possible way for an
analyst to obtain more or better data, because no more data pertaining to the matter at issue
exists. As such, an insignificant result (at the designated level of statistical precision) in the
context developed herein positively implies a lack of evidence of dominance. Indeed, while
adjudicating a matter, it is important to be vigilant that when parties assert small sample issues as
a defense regarding dominance or market power issues that they have truly captured all
attainable observations.

\textit{F. r-firm test}

The three-firm case can be extended more generally as an equalized r-firm test. Suppose
there are $r$ firms. The test that results in maximization of the $p$-value is the two-firm test (applied
to $r$ firms):
Using numerical methods similar to those in appendix B, it can be shown that this formulation of the null maximizes the $p$-value for every observed share value larger than fifty percent for up to 100 observed contests, for up to five firms, over a wide swath of dominance thresholds $\theta$.\textsuperscript{35}

The test that results in minimization of the $p$-value is:

$$H_0: p_A = \theta; p_B = 1 - \theta$$

for $i \in \{B, C, D, \ldots\}$ where the set $\{B, C, D \ldots\}$ includes all firms except $A$. Using the same numerical methods in the previous paragraph, it can be shown that this test minimizes the $p$-value.\textsuperscript{36} Common software can be used to implement the hypothesis test.\textsuperscript{37}

\textsuperscript{35} The code is available from the authors upon request.

\textsuperscript{36} Due to computational constraints, the minimization-maximization for markets with more than five firms was not implemented. However, given the proven results for five or fewer firms, the results likely also hold for more than five firms.

\textsuperscript{37} For the two-firm-test, which maximizes the $p$-value, the following Excel command calculates the $p$-value:

$$=2\text{-BINOM.DIST}(W - 1,n,\theta,1) - \text{BINOM.DIST}(W - 1,n,1-\theta,1)$$

with

$W$ = the observed number of wins of the majority-observed firm;

$n$ = the number of contests;

$\theta$ = the dominance threshold.

For the equalized three-firm test, which minimizes the $p$-value, the following Excel command calculates the $p$-value:
V. CONCLUSION

Beyond the well-known issues related to market definition, entry, and efficiencies, small sample issues are an additional reason why caution may be warranted before drawing conclusions about dominance in thinly traded markets. More specifically, in the presence of small sample issues, high shares can lead to statistically unsupportable conclusions about market dynamics and firm dominance.

First, whether small sample issues can confound findings of dominance in a duopoly market structure was considered. The appropriate tests vary depending on the practitioners’ exogenous evidence regarding the firm observed to have a majority share. In most circumstances, however, we suggest the practitioner use the two-firm test, which explicitly allows for the firm with a minority win probability to, by random chance, win a majority of contests. We also noted that the relative weight placed by the analyst on false positives and false negatives can affect the choice of test.

Next, a multifirm market structure was considered. If, based on exogenous evidence, it is deemed highly improbable for a firm with a minority win probability to have won a majority of contests, we suggest using the one-firm test. Conversely, in the more likely scenario where the

\[
\text{BINOM.DIST}(W - 1, n, \theta, 1) - (r - 1) \times \text{BINOM.DIST}(W - 1, n, (1 - \theta)/(r - 1), 1)
\]

with

\[
\begin{align*}
    r &= \text{the total number of firms}; \\
    W &= \text{the observed number of wins of the majority-observed firm}; \\
    n &= \text{the number of contests}; \\
    \theta &= \text{the dominance threshold}.
\end{align*}
\]
analyst is not able to rule out this possibility, we propose a multifirm test. It was found that the
$p$-value varies in a systematic way with the testable hypothesis considered. Specifically,
assumptions regarding the null hypotheses for the firms with minority win probabilities have a
nontrivial impact on the analysis. We bound the $p$-values and propose one test for the maximum
$p$-value and one for the minimum; we discuss when either is appropriate, noting that the
implications may depend on relative concern with false positives and negatives or a party’s
position in an adversarial context. This variance may provide insight into why European
competition authorities tend to put emphasis on asymmetries in shares in addition to a dominant
market position. It was also emphasized that parties can assert small sample issues as a defense
only when they have truly compiled all attainable observations.

Dominance is not the only area in which small sample concerns and techniques have
relevance in industrial organization antitrust applications. For example, it is not uncommon to
evaluate market evidence where potentially competing firms exhibit little geographic overlap
suggesting that these firms are more complementary than competitive in nature. Such an
argument can be tenuous if there are limited “contests” in each of the observed geographic areas.
Further, conclusions regarding high concentration and high diversion ratios can also be
confounded by small sample issues.
Appendix A: Tables of Critical Win Numbers

The code that generated the tables in this appendix is available from the authors upon request.

Table A1: Critical Win Probabilities for One-Firm Test:

*Strong exogenous reason to assert the majority-observed firm could never be the firm with the lower true win probability*

| Dominance threshold: | >50% | | >60% | | >70% |
|----------------------|------|---|------|---|------|---|
| Sig. level:          | 10%  | 5% | 1%   | 10% | 5% | 1% |
| # of contests        |      |    |      |    |    |    |
| 5                    | 5    | 5  | 5    |    |    |    |
| 6                    | 6    | 6  | 6    |    |    |    |
| 7                    | 7    | 7  | 7    | 7  | 7  | 7  |
| 8                    | 8    | 8  | 8    | 8  | 8  | 8  |
| 9                    | 9    | 9  | 9    | 9  | 9  | 9  |
| 10                   | 10   | 10 | 10   | 10 | 10 | 10 |
| 11                   | 10   | 10 | 11   | 11 | 11 | 11 |
| 12                   | 12   | 12 | 12   | 12 | 12 | 12 |
| 13                   | 13   | 13 | 13   | 13 | 13 | 13 |
| 14                   | 14   | 14 | 14   | 14 | 14 | 14 |
| 15                   | 15   | 15 | 15   | 15 | 15 | 15 |
| 16                   | 16   | 16 | 16   | 16 | 16 | 16 |
| 17                   | 17   | 17 | 17   | 17 | 17 | 17 |
| 18                   | 18   | 18 | 18   | 18 | 18 | 18 |
| 19                   | 19   | 19 | 19   | 19 | 19 | 19 |
| 20                   | 20   | 20 | 20   | 20 | 20 | 20 |
| 21                   | 21   | 21 | 21   | 21 | 21 | 21 |
| 22                   | 22   | 22 | 22   | 22 | 22 | 22 |
| 23                   | 23   | 23 | 23   | 23 | 23 | 23 |
| 24                   | 24   | 24 | 24   | 24 | 24 | 24 |
| 26                   | 26   | 26 | 26   | 26 | 26 | 26 |
| 27                   | 27   | 27 | 27   | 27 | 27 | 27 |
| 28                   | 28   | 28 | 28   | 28 | 28 | 28 |
| 29                   | 29   | 29 | 29   | 29 | 29 | 29 |

Note: Blanks cells are where it is impossible to reject null at stated significance level.
<table>
<thead>
<tr>
<th># of contests</th>
<th>&lt;50% 10% 5% 1%</th>
<th>&gt;60% 10% 5% 1%</th>
<th>&gt;70% 10% 5% 1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6 6</td>
<td>6 6</td>
<td>6 6</td>
</tr>
<tr>
<td>7</td>
<td>7 7 7</td>
<td>7 7 7</td>
<td>7 7 7</td>
</tr>
<tr>
<td>8</td>
<td>7 8 8 8 8</td>
<td>7 7 7 7 7</td>
<td>7 7 7 7 7</td>
</tr>
<tr>
<td>9</td>
<td>8 8 8 8 9</td>
<td>8 9 9 9 9</td>
<td>8 9 9 9 9</td>
</tr>
<tr>
<td>10</td>
<td>9 9 9 10 9</td>
<td>9 9 10 10 10</td>
<td>9 9 10 10 10</td>
</tr>
<tr>
<td>11</td>
<td>10 10 10 10 10</td>
<td>10 10 10 10 10</td>
<td>10 10 10 10 10</td>
</tr>
<tr>
<td>12</td>
<td>10 10 10 10 10</td>
<td>10 10 10 10 10</td>
<td>10 10 10 10 10</td>
</tr>
<tr>
<td>14</td>
<td>11 12 12 12 12</td>
<td>11 12 12 12 12</td>
<td>11 12 12 12 12</td>
</tr>
<tr>
<td>15</td>
<td>12 12 12 12 12</td>
<td>12 12 12 12 12</td>
<td>12 12 12 12 12</td>
</tr>
<tr>
<td>16</td>
<td>12 13 13 13 13</td>
<td>13 13 13 13 13</td>
<td>13 13 13 13 13</td>
</tr>
<tr>
<td>17</td>
<td>13 13 13 13 13</td>
<td>14 14 14 14 14</td>
<td>14 14 14 14 14</td>
</tr>
<tr>
<td>18</td>
<td>13 14 14 14 14</td>
<td>14 14 14 14 14</td>
<td>14 14 14 14 14</td>
</tr>
<tr>
<td>19</td>
<td>14 15 15 15 15</td>
<td>15 15 15 15 15</td>
<td>15 15 15 15 15</td>
</tr>
<tr>
<td>20</td>
<td>15 15 15 15 15</td>
<td>16 16 16 16 16</td>
<td>16 16 16 16 16</td>
</tr>
<tr>
<td>21</td>
<td>15 16 16 16 16</td>
<td>16 16 16 16 16</td>
<td>16 16 16 16 16</td>
</tr>
<tr>
<td>22</td>
<td>16 17 17 17 17</td>
<td>17 17 17 17 17</td>
<td>17 17 17 17 17</td>
</tr>
<tr>
<td>23</td>
<td>16 17 17 17 17</td>
<td>18 18 18 18 18</td>
<td>18 18 18 18 18</td>
</tr>
<tr>
<td>24</td>
<td>17 18 18 18 18</td>
<td>18 18 18 18 18</td>
<td>18 18 18 18 18</td>
</tr>
<tr>
<td>25</td>
<td>18 18 18 18 18</td>
<td>19 19 19 19 19</td>
<td>19 19 19 19 19</td>
</tr>
<tr>
<td>26</td>
<td>18 19 19 19 19</td>
<td>20 20 20 20 20</td>
<td>20 20 20 20 20</td>
</tr>
<tr>
<td>27</td>
<td>19 20 20 20 20</td>
<td>20 20 20 20 20</td>
<td>20 20 20 20 20</td>
</tr>
<tr>
<td>28</td>
<td>19 20 20 20 20</td>
<td>21 21 21 21 21</td>
<td>21 21 21 21 21</td>
</tr>
<tr>
<td>29</td>
<td>20 21 21 21 21</td>
<td>22 22 22 22 22</td>
<td>22 22 22 22 22</td>
</tr>
<tr>
<td>30</td>
<td>20 21 21 21 21</td>
<td>22 22 22 22 22</td>
<td>22 22 22 22 22</td>
</tr>
</tbody>
</table>

Note: Blanks cells are where it is impossible to reject null at stated significance level.
Appendix B

Analytical Proof for the Win Probabilities that Minimize and Maximize the $p$-value

Assume

(i) three firms, $A$, $B$, and $C$;

(ii) the probabilities that firms $A$, $B$, and $C$ win a given contest are $p_A$, $p_B$, and $p_C$, respectively;

(iii) ten competitions;

(iv) an eighty percent observed share;

(v) a dominance threshold of fifty percent.

The minimization-maximization problem is choosing the allocation of the win probabilities $p_B$ and $p_C$ in $H_0$. (In $H_0$, the [assumed] true win probability for $p_A$ is fixed at 0.5, and firm $A$ is irrelevant to the minimization/maximization problem.) Letting $B$ and $C$ denote the obtained shares of firms $B$ and $C$, respectively, and letting $N_B$ and $N_C$ denote the number of wins of firms $B$ and $C$, respectively:

$$P(B \geq 0.8) + P(C \geq 0.8) = P(N_B \geq 8) + P(N_C \geq 8)$$


Using the binomial distribution and the fact that $p_C = 0.5 - p_B$:

(A1)
\[ P(B > 0.7) + P(C > 0.7) \]

\[
= \left\{ \frac{10!}{8! \cdot 2!} p_B^8 (1-p_B)^2 + \frac{10!}{9! \cdot 1!} p_B^9 (1-p_B) + p_B^{10} \right\} \\
+ \left\{ \frac{10!}{8! \cdot 2!} (0.5-p_B)^8 (0.5+p_B)^2 + \frac{10!}{9! \cdot 1!} (0.5-p_B)^9 (0.5+p_B) + (0.5-p_B)^{10} \right\}
\]

Differentiating the probability in (A1) yields

(A2) \[ 360 (1-p)^2 p^7 - 360 (0.5-p)^7 (0.5+p)^2 \]

The expression (A2) has one real solution, at \( p = 0.25 \). The second derivative of (A2) is

\[
7.03125 - 45 p_B^3 + 630 p_B^2 - 1575 p_B + 7560 p_B^5 - 12960 p_B^7 + 6480 p_B^9
\]

which is positive at \( p = 0.25 \), thereby establishing \( p = 0.25 \) as a minimum. Moreover, we can establish a maximum at \( p_B = 0.5 \): It is simple arithmetic to verify that (A2) evaluated at \( p = 0.26 \) is positive. Because (A1) is a polynomial, it is continuous, and therefore, (A1) cannot be negative anywhere where \( p > 0.25 \). (In order for the [A2] to be negative at some point where \( p > 0.25 \), the expression would have to be equal to zero at some point \( p > 0.25 \). But we established that the only root to the [A2] is at \( p = 0.25 \), so the expression cannot equal zero at any point \( p > 0.25 \) and hence cannot be negative at any point \( p > 0.25 \).) Because (A2) is the derivative of the probability in (A1) and is always increasing for \( p > 0.25 \), the probability in (A1) is maximized at the corner where \( p = 0.5 \).

**Numerical Minimization-Maximization for Other Observed Shares, Numbers of Contests, and Dominance Thresholds**

Generically, the null hypothesis can be written

\[ H_0: p_A = 0.5; p_B = \epsilon; p_C = 0.5 - \epsilon \]
for some $0 \leq \epsilon \leq 0.5$. Suppose there are $n$ contests. Let $B$, $C$, $N_B$, $N_C$ be as defined in the previous footnote; and let $A$ and $N_A$ be the analogous definition for firm $A$. The probability of observing a firm with a share of $s_0$ or more is, if $s_0 > 0.5$,

$$P(A \geq s_0) + P(B \geq s_0) + P(C \geq s_0)$$

Let $k_0$ be the observed number of wins (i.e., so $s_0 = k_0/n$). Then, the $p$-value is

(A3) $P(A \geq s_0) + P(B \geq s_0) + P(C \geq s_0) = P(N_A \geq k_0) + P(N_B \geq k_0) + P(N_B \geq k_0)$

or the probability of observing a firm win $k_0$ or more contests out of $n$. Given that $P_A = 0.5, p_B = \epsilon, P_C = 0.5 - \epsilon$, (A3) can be found with the binomial distribution:

(A4) $\sum_{k=k_0}^{n} \frac{n!}{k!(n-k)!} \{0.5^k 0.5^{n-k} + (0.5 - \epsilon)^k (0.5 + \epsilon)^{n-k} + \epsilon^k (1 - \epsilon)^{n-k}\}$

The expression (A4) is difficult to minimize (with respect to $\epsilon$) with calculus, so we use numerical methods (that is, a grid search). The results from the grid search are that, for every value of $n$ from 10 to 100, and for every $k$ (such that $\frac{k}{n} > 0.5$),

(A5) $\frac{n!}{k!(n-k)!} \{0.5^k (0.5)^{n-k} + (0.5 - \epsilon)^k (0.5 + \epsilon)^{n-k} + \epsilon^k (1 - \epsilon)^{n-k}\}$

is minimized when $\epsilon = 0.25$. Therefore, (for every value of $n$ from 10 to 100, and for every $k$ [such that $\frac{k}{n} > 0.5$]) every term in the sum (A4) is minimized when $\epsilon = 0.25$, and hence the sum (A4) itself is minimized when $\epsilon = 0.25$. Similarly, the terms in (A5) are always maximized when $\epsilon = 0.5$ and hence the sum in (A4) is also maximized when $\epsilon = 0.5$. The Stata code for the numerical (grid-search method) minimization is available electronically from authors upon request.