
Neely Agin explains why the Supreme Court’s upcoming ruling on the fate of American Express’s anti-steering rules also creates a rare opportunity for the Court to ease the uncertainty over the application of the rule of reason.

Fuzzy Rules, Expressions Hair Design v. Schneiderman, and Ohio v. American Express

Updating his February 2017 article, Matthew Moloshok reviews the Supreme Court’s remand of Expressions Hair Design and its upcoming consideration of Ohio v. American Express and suggests the cases share a common flaw—that standards regulating “commercial speech” and “unreasonable restraints” may be too indefinite and inefficient to protect consumers against excessive credit card fees. Is it time to try direct and comprehensive regulation?

Raising the Bar: How Does China’s Tetra Pak Decision Measure Up to the ECJ’s Requirements for Loyalty Rebate Cases?

Fei Deng and Su Sun analyze China’s State Administration of Industry and Commerce’s Tetra Pak decision in view of the European Court of Justice’s Intel judgment. Comparing and contrasting these two rulings on loyalty rebates, the authors conclude that although they were generally consistent in principle, the ECJ’s judgment sets a higher standard than was contained in the SAIC’s decision in its reasoning for a finding of liability.

What Is So Significant About Statistical Significance?

Continuing our series explaining complex antitrust economics concepts to lawyers, Andrew Abere explicates for non-economists the concept of statistical significance of economic analysis, its limitations, and how it can be misinterpreted or misused. The article also highlights an ongoing debate on whether statistical significance is relied upon too heavily, or should be relied upon at all.

How to Economize Consumer Protection

Michael Baye and Joshua Wright advocate greater use of economic analysis in consumer protection enforcement, including both deception and unfairness cases. The authors discuss reasons for the relative absence of economic analysis in past cases, outline how economists can assess the extent to which a challenged business practice adversely affected consumers, and argue that economic analysis can assist government agencies in optimally allocating scarce resources to enforcement activities.

Book Review: Getting Economics Back in the Game

Gregory Leonard reviews a new book by Nobel Prize winner Jean Tirole—Economics for the Common Good. Among other things, Tirole proposes that economists should strive to make economic ideas comprehensible to a general audience and, in that way, regain the public’s respect for the profession and contribute to resolving the serious policy dilemmas facing the world.

Neely Agin

When advising clients on the potential antitrust risk of proposed agreements with competitors, customers, and suppliers and other business practices, antitrust attorneys often must evaluate whether the proposed arrangement is likely to be treated as per se illegal or evaluated under the more lenient rule of reason. If an arrangement is likely to be evaluated under the rule of reason, the attorney then must predict the likely outcome of such evaluation before advising her client on the level of antitrust risk. Although this task has always required extensive analysis, it has become increasingly theoretical, without clear guidance from the U.S. courts on how the rule of reason is applied in practice. This lack of guidance has resulted in business uncertainty and high legal costs.

The Supreme Court will have a rare opportunity to ease this uncertainty in the next few months when it decides the fate of American Express’s “anti-steering” rules. Following disagreement between the district court and Second Circuit regarding whether the plaintiffs met their burden of showing that the anti-steering rules are anticompetitive, the Supreme Court is being asked to articulate the proper application of the rule of reason standard. As a result, the Court may shed additional light on the practical application of the rule of reason test, and provide long-awaited guidance on which antitrust attorneys can rely when advising their clients, including the importance of market power in the analysis and the appropriate application of the burden-shifting framework.

The DOJ and States Challenge Credit Card Anti-Steering Rules Under the Sherman Act

The challenge to Amex’s anti-steering rules (also referred to as nondiscriminatory provisions or NDPS) began in 2010, when the Antitrust Division of the Department of Justice (DOJ) and 17 state attorneys general filed a lawsuit in the Eastern District of New York claiming the NDPs used by Amex, Visa, and MasterCard reduce competition. NDPs discourage merchants from “steering” customers to other credit card networks by prohibiting merchants from offering consumers discounts or other nonmonetary incentives to use credit cards less costly for merchants to accept, expressing preferences for any card, or disclosing to customers information about card costs to merchants. According to the complaint, enforcement of the NDPs in the defendants’ agreements with merchants violated Section 1 of the Sherman Act.

The DOJ and state plaintiffs alleged that the NDPs resulted in consumers paying more for their purchases because the NDPs prevent merchants from using discounts or incentives to encour-

age consumers to pay using a credit card that would cost the merchant less to accept. According to the DOJ, if merchants were able to use such incentives, card companies would have an incentive to reduce acceptance costs and merchants would, in turn, reduce the retail prices they charge to consumers.

On the same day that the DOJ filed its complaint, it announced that it had reached a proposed settlement with Visa and MasterCard. Under the settlement, MasterCard and Visa agreed to voluntarily rescind their NDPs and to refrain from blocking merchant steering. In particular, Visa and MasterCard agreed to allow merchants to offer consumers certain discounts and rebates for using particular card networks or cards and to communicate with consumers the cost incurred by the merchant when a consumer uses a particular card, network or other form of payment. Amex, however, chose not to settle and has continued to defend its NDPs all the way to the U.S. Supreme Court. Amex had a greater incentive to continue fighting the DOJ given Amex’s smaller market share and the fact that its business model relies heavily on revenue from merchant discount fees, which it uses to fund premium benefits for cardholders.

Applying the Rule of Reason, the District Court Found Amex’s NDPs Were Anticompetitive

Following a nearly two-month trial in 2015 with over 1,000 exhibits and more than thirty fact witnesses, the district court sided with the DOJ and state plaintiffs. The court found that the NDPs were vertical restraints and therefore should be analyzed under the rule of reason. Applying the rule of reason, the district court found that the NDPs constituted an unreasonable restraint of trade in violation of Section 1 of the Sherman Act. The court concluded that the NDPs created “an environment in which there is nothing to offset credit card networks’ incentives—including Amex’s incentive—to charge merchants inflated prices for their services. This, in turn, results in higher costs to all consumers who purchase goods and services from these merchants.”

In its analysis of the competitive effects of the NDPs, the district court defined the relevant market as general purpose credit and charge card network services provided to merchants. Relying on its interpretation of the Second Circuit’s opinion in United States v. Visa, the district court rejected Amex’s argument that the relevant market also should include the related market in which credit card issuers compete for cardholders. Instead, the district court found that although the credit card market is two-sided and that both sides are deeply interrelated, they are nevertheless separate and distinct markets. It then applied the rule of reason to analyze the legality of the NDPs solely in the context of the market for network services provided to merchants.

The district court then considered the alleged anticompetitive effects of the NDPs in the defined market. It noted that the plaintiffs could meet their initial burden to demonstrate anticompetitive effects in the relevant market either directly, by showing an actual adverse effect on competition in the relevant market caused by the NDPs, or indirectly, by establishing that market power plus other factors increased the likelihood of harm to competition (the so-called market power plus method).

---

5 Id. at 150.
6 Id. at 151 (citing United States v. Visa, 344 F.3d 229 (2d Cir. 2003)).
7 Id.
8 Id. at 168–69 (citing Tops Mkts., Inc. v. Quality Mkts., Inc., 142 F.3d 90, 96 (2d Cir. 1998)).
Evaluating whether the plaintiffs proved anticompetitive effects through the indirect method, the district court held that Amex had market power despite its market share of just 26.4 percent because there were only four key players in the market, barriers to entry were high, and “cardholder insistence” amplified Amex’s power. According to the court, Amex’s loyal cardholder base was critical to its finding of market power because their loyalty impeded merchants’ ability to resist anticompetitive behavior by Amex, including significant price increases. Because loyal Amex cardholders would choose to spend less or simply not shop at merchants that do not accept Amex, this cardholder insistence effectively precluded merchants from dropping Amex notwithstanding its relatively modest market share.  

The district court also considered the direct method of proof, and found that the NDPs harmed interbrand competition by preventing merchants from influencing the payment choices of their customers. It found that the NDPs prevented merchants from steering additional charge volume to low-cost networks, effectively short-circuiting the usual price-setting mechanism in the market for network services. As a result, there was no price competition among Amex and its rivals for network services because the NDPs removed the incentive of Amex or its competitors to compete against each other through lower prices. Because lowering prices would not increase transaction volume, a network would have no incentive to offer lower prices. Moreover, the NDPs also rendered it “nearly impossible for a firm to enter the relevant market by offering merchants a low-cost alternative to the existing networks.”  

To illustrate, the court pointed to Discover’s attempt in the 1990s to adopt a business model with lower fees and higher transaction volumes. The court noted that Discover’s attempt to change to a low-cost model failed because the NDPs prevented merchants from steering customers to Discover’s lower cost alternative.  

Finally, the district court determined that the NDPs had resulted in higher prices for merchants and consumers because they allow Amex and other credit card companies to charge higher prices without fear that those prices would be undercut by a competitor. Merchants then pass on these increased costs to consumers in the form of higher prices for the underlying products and services sold by the merchant. For Amex cardholders, these higher prices were at least partially offset by the premium rewards provided by Amex. However, higher retail prices affect all customers, including those who paid in cash or used a lower cost credit card, without any offsetting benefits. As a result, the court concluded that the NDPs are anticompetitive because they cause non-Amex cardholders effectively to subsidize the benefits received by Amex Cardholders.  

Having found that the government had met its initial burden of showing anticompetitive effects, the district court considered but ultimately rejected the pro-competitive justifications offered by Amex. Specifically, Amex argued that the NDPs were reasonably necessary to (1) preserve Amex’s differentiated business model and enable it to drive competition in network services, and (2) prevent free riding by merchants who, absent the NDPs, could associate themselves with Amex’s premium brand by representing that they accept Amex, but then steer customers to lower cost alternatives at the point of sale.  

The district court was unpersuaded, finding Amex’s argument that the NDPs were necessary to preserve its business model to constitute an admission that

9 Id. at 188–95.
10 Id. at 213.
11 Id. at 211–14.
12 Id. at 215–17.
13 Id. at 225.
the business model was fundamentally inconsistent with the goals of the Sherman Act because the model “could not survive if exposed to the full spectrum of interbrand competition.” The court also found that there was limited evidence in favor of Amex's free riding argument and that any procompetitive benefits from preventing free riding were not sufficient to offset the more pervasive harm to interbrand competition by the NDPs. After ruling in favor of the government, the court enjoined Amex from enforcing its anti-steering rules and required Amex to allow merchants to encourage customers to use other credit cards. Specifically, the order required Amex to permit merchants to offer discounts for the use of other cards, express a preference to customers for the use of other cards, provide customers with information about the costs merchants incur when a customer uses a particular card, and engage in conduct to encourage the use of other cards. The order also required Amex to take certain actions to inform merchants of their freedom to engage in steering activities and to adopt compliance measures to ensure that Amex employees understood that they could no longer block merchant steering.

Although the court of appeals accepted the district court’s factual findings, it determined the district court had erred in its (1) definition of the relevant market, (2) analysis of market power, and (3) application of the rule of reason test.

First, the Second Circuit found that the district court had erred in excluding the market for cardholders from its relevant market definition, thus considering only one side of a two-sided market. It found that by excluding cardholders, the district court improperly applied the hypothetical monopolist test, because the application of the test to a two-sided market must “consider the feedback effects inherent on the platform by accounting for the reduction in cardholders’ demand for cards (or card transactions) that would accompany any degree of merchant attrition” caused by an attempt by Amex to increase merchant fees.

Second, the Second Circuit held that the government had failed to prove that Amex had market power. Based on the district court’s findings of fact, the Second Circuit found that Amex had to compete fiercely for cardholder loyalty and that cardholder loyalty would “dissipate” if cardholders perceived that Amex offered lower value relative to its rivals. The Second Circuit reasoned that Amex’s need to compete on price through continual investment in cardholder benefits in order to maintain share showed that Amex lacked market power.

Finally, the Second Circuit found that the district court had erred in applying the rule of reason by failing to account for the procompetitive effects of the NDPs on the cardholder side of the market. This incorrect application of the rule of reason was, according to the Second Circuit, the direct result of the district court’s erroneous market definition. The Second Circuit reasoned that the district court’s analysis of anticompetitive effects should have considered the extent to which increased cardholder benefits would offset increased merchant fees. If, for example, increased

14 Id. at 227.
15 Id. at 235–38.
18 Id. at 203.
cardholder benefits helped Amex attract more affluent cardholders, then merchants might benefit from the increased fees because they would result in increased access to consumer spending. As a result, the Second Circuit held that the plaintiffs failed to meet their initial burden under the rule of reason because they had not demonstrated that the NDPs would have anticompetitive effects on the relevant market as a whole.¹⁹

**Several Plaintiff States Seek Certiorari**

On June 2, 2017, a subset of the plaintiff states filed a petition for a writ of certiorari asking the U.S. Supreme Court to review the Second Circuit’s decision in favor of Amex. The DOJ noticeably was absent from the petition despite leading the efforts against the NDPs for nearly seven years. Although the DOJ had participated in the plaintiff-appellee’s petition for a rehearing en banc by the Second Circuit, the DOJ dropped out of the case sometime between the Second Circuit’s denial of the petition for rehearing in January 2017 and the petition for the writ of certiorari five months later. Another surprising development occurred in August when, in a highly unusual move, the DOJ filed a brief with the Supreme Court opposing cert in its own case. Although the DOJ filed in opposition to cert, its brief was still highly critical of the Second Circuit’s opinion. The DOJ argued that although the Second Circuit was wrong on the merits, the states’ petition was not yet ripe for review. The DOJ explained that the novel market definition issues raised by the states’ petition should continue to percolate in the lower courts before being addressed by the Supreme Court. Notably, the DOJ’s decision not to join the plaintiff’s petition for cert and later to oppose the granting of cert came at a time when the DOJ was without Senate-confirmed leadership.

The parties differ in their interpretation of the Second Circuit’s description of the proper rule of reason analysis. The state petitioners, on the one hand, argue that the Second Circuit effectively upended the customary rule of reason analysis by placing the burden on the plaintiff to disprove any procompetitive benefits of the challenged restraint. They argue that the Second Circuit’s reasoning “effectively collapsed the staggered rule-of-reason framework into a single, all-encompassing step.”²⁰ Amex, on the other hand, argues that any supposed conflict with the traditional application of the rule of reason is illusory and the appearance of a conflict exists only as to the result of the district court’s erroneous market definition. According to Amex, the plaintiff always bears the initial burden of demonstrating competitive harm in the market as a whole. Because the district court improperly excluded the cardholder side of the market when it weighed the anticompetitive effects of the NDPs, it failed to properly apply the rule of reason test. Amex argues that the Second Circuit’s opinion does not collapse the traditional rule of reason analysis into a single step, but rather ensures that the first step of the analysis is properly applied to the entire relevant market instead of only a portion thereof.²¹ According to Amex, “[T]he interdependence of merchant and cardholder demand requires assigning burdens as the court of appeals did.”²² As the government’s expert testified: “[A]n assessment of market definition, market power and competitive effects should account for the two-sided nature of the market.”²³

¹⁹ Id. at 204–05.


²² Id. at 24.

²³ Id. (quoting expert testimony at Tr. 4018:13-19).
History of the Rule of Reason Analysis
The rule of reason is such a ubiquitous feature of antitrust law that nearly all practitioners could paraphrase it from memory. Simply stated, the rule of reason is a multi-step, burden-shifting framework to assess whether the anticompetitive effects of a restraint outweigh its procompetitive benefits. At the initial stage, the plaintiff bears the burden to demonstrate the anticompetitive effects of the restraint on the relevant market. If the plaintiff meets that burden, the burden shifts to the defendant to come forward with a procompetitive justification for the restraint. If the defendant offers legitimate procompetitive justifications, the plaintiff must then show that the restraint is not necessary to achieve the defendant’s procompetitive objectives or that the anticompetitive effects of the restraint outweigh its procompetitive benefits. The apparent simplicity of the analysis, however, belies its true complexity and the difficulties that courts have faced in applying it in practice. As a result, the disagreement over whether the Second Circuit’s decision represents a radical departure from a traditional rule of reason analysis or merely the application of that traditional analysis to a particular two-sided market should not be surprising.

The conceptual framework for the rule of reason analysis was first announced in *Standard Oil Co. v. United States*, when the U.S. Supreme Court held that the Sherman Act prohibited only “unreasonable” restraints of trade. The Court later articulated the basis of what is considered the modern version of the rule, stating that the analysis involves consideration of “the facts peculiar to the business to which the restraint is applied; its condition before and after the restraint was imposed; the nature of the restraint and its effect, actual or probable.” In the century since the rule of reason was first announced, the Court has offered significant guidance as to the types of restraints that should be analyzed under the rule of reason, those that should be subject to a “quick look,” and those that are per se illegal. The Court also has offered general principles that lower courts have applied in conducting a rule of reason analysis. Yet commentators generally agree that the Court has left “lower courts with no clear standards” in applying the rule of reason.

Lacking clear standards, the lower courts’ application of the rule often has been “unfocused” and so “wide-ranging” that it is “one of the most costly procedures in antitrust practice.” The specifics of how the test is applied vary from court to court and from circuit to circuit. For example, some courts speak of the need to “balance” the alleged pro and anticompetitive effects of a challenged restraint, despite the fact that the Supreme Court has “never actually conducted any balancing.” But as Professor Herbert Hovenkamp points out, while balancing “requires values that can be cardinaly measured and weighed against each other . . . . [T]he factors that are supposedly balanced in Sherman Act cases almost never fit this description.” Rather, district courts often consider a variety of subjective factors that arguably, but not unequivocally, have had a negative impact on competition, along with a defendant’s stated business reasons and other seemingly procompetitive effects of the business practice. Consideration of such subjective factors

---

24 221 U.S. 1, 58 (1911).
25 Chicago Bd. of Trade v. United States, 246 U.S. 231, 238 (1918).
28 Herbert Hovenkamp, *Antitrust Balancing*, 12 NYU J.L. & Bus. 369, 372 (2016) (“While the Supreme Court has both accepted and rejected defenses in rule of reason cases, it has only rarely described this analysis as ‘balancing,’ and it has never actually conducted any balancing.”) (footnote omitted).
29 Id. at 370.
cannot properly be viewed as balancing, as they cannot be compared using a standard metric. Instead, courts’ attempts to apply a balancing test often appear to be more of an ad hoc application of subjective judgment of the “appropriateness” of the business practice at issue and the nature and extent of any impact on the marketplace. Moreover, there are very few instances of courts actually engaging in balancing, likely because of the difficulty of measuring the opposing values at issue and the complete lack of guidance on the proper way to conduct balancing.

The differing standards under the rule of reason that have resulted from the lack of guidance from the Supreme Court include even such basic questions as whether a defendant must possess market power for an alleged restraint to be condemned. Language from the Supreme Court’s decision in *FTC v. Indiana Federation of Dentists* that speaks of market power as a surrogate for anticompetitive effects has led some courts to hold that proof of market power is an alternative to proof of anticompetitive effects, while other courts have held that market power is a separate element that must be proven by the plaintiff. The Third Circuit has taken the approach that once a plaintiff demonstrates either market power or anticompetitive effects, the burden shifts to the defendant. Alternatively, the Seventh and D.C. Circuits require “substantial market power” as a threshold requirement for all rule of reason cases. The Second Circuit seems not to have decided on a consistent approach. In some instances, as the district court did here, Second Circuit courts treat market power as an alternative to meeting the plaintiff’s initial burden of showing actual anticompetitive effects. In other instances, the Second Circuit has held that market power must be demonstrated at the outset of the rule of reason analysis or the plaintiff’s case fails. This seemingly confused approach is one that the Second Circuit has at least at times acknowledged, stating that “[t]he precise role that market power plays in rule of reason analysis of horizontal combinations or conspiracies is a matter of some dispute.”

The *Amex* case also illustrates another key area of confusion for courts in applying the rule of reason—how to administer the test’s burden-shifting process. While the district court in *Amex* concluded that the plaintiffs had met their initial burden of proving anticompetitive effects and that the burden then shifted to *Amex* to prove any procompetitive justifications, the Second Circuit held that the plaintiffs failed to meet their initial burden of proving “adverse effect on competition as a whole in the relevant market.”

The states that sought certiorari argued that the Second Circuit’s decision incorrectly placed too high a burden on them by requiring the government to affirmatively disprove any procompetitive benefits from the NDPs. According to the state petitioners, this erroneously placed the initial burden of disproving procompetitive benefits on the plaintiffs, when, in fact, the burden should

---

30 See 476 U.S. 447, 460–61 (1986); see also Markham, Jr., supra note 26, at 645–48.
32 See *Hardy v. City Optical Inc.*, 39 F.3d 765, 767 (7th Cir. 1994) (“the rule in this Circuit is that a tying agreement is not actionable unless the defendant has substantial market power in the tying product”); *Rothery Storage & Van Co. v. Atlas Van Lines, Inc.* 792 F.2d 210, 221 (D.C. Cir. 1986) (holding that a defendant’s conduct cannot be found to violate the Sherman Act in a rule of reason analysis absent a finding of market power).
33 See *MacDermid Printing Sol’ns v. Cortron Corp.*, No. 15–589–cv (2d. Cir. Aug. 10, 2016); see also *Clorox Co. v. Sterling Winthrop, Inc.*, 117 F.3d 50, 56 (2d Cir. 1997).
34 See United States v. Visa USA, Inc. 344 F.3d 229, 238 (2d Cir. 2003) (“As an initial matter, the government must demonstrate that the defendant conspirators have ‘market power’ in a particular market for goods or services.”) (footnote omitted).
36 *American Express Co.*, 838 F.3d at 204–07.
shift to the defendant to prove procompetitive benefits once the plaintiff has proven anticompetitive effects. A group of 25 antitrust professors who filed an amicus curiae brief in support of the state petitioners argue that this essentially placed the burden on the plaintiffs to show a “net” harm to competition rather than just harm to competition.\(^{37}\) The professors state:

> [P]lacing the burden on plaintiffs to disprove offsetting procompetitive effects before they can make out even a \textit{prima facie} showing of an adverse effect on competition—as the Second Circuit did—fundamentally warps established Rule-of-Reason analysis. The plaintiff should not have to speculate as to what evidence of procompetitive effects a defendant might offer and then refute that hypothetical evidence to meet its initial burden.\(^{38}\)

Instead, the professors argue, proof of adverse effects on horizontal price competition must suffice to shift the burden to the defendant to prove any alleged procompetitive justifications.\(^{39}\) If the defendant meets its burden, then the burden should shift back to the plaintiff to demonstrate that anticompetitive effects were not outweighed by the procompetitive justifications.

In contrast, Amex argues that the Second Circuit correctly administered the burden-shifting test and that “a plaintiff always bears the burden to demonstrate competitive harm in the product market as a whole.”\(^{40}\) Amex also argues that under the Supreme Court’s precedent, the two-sided nature of the market necessitated the Second Circuit’s application of burden-shifting.\(^{41}\) In addition, Amex suggests that further guidance on how to apply the rule of reason analysis is not necessary. Amex argues that the Supreme Court has purposefully left open the particulars of how to apply the test in a specific case, leaving to the lower courts how to apply the rule in the “varying contexts presented to them.”\(^{42}\) Amex points to the decision in \textit{FTC v. Actavis, Inc.}, in which the Court stated: “We therefore leave to the lower courts the structuring of the present rule-of-reason antitrust litigation.”\(^{43}\) Amex’s view is that such analysis is fact-intensive and market-specific, and the Court therefore has chosen to leave the particulars of the analysis to the lower courts.

It seems likely that the Supreme Court will at least need to address in its decision whether its seeming lack of guidance on the rule of reason has been intentional, or whether it is the result of not yet having the right opportunity to decide how the test should be applied in practice. Perhaps the Court was purposeful in addressing, as a first step, when the \textit{per se} or rule of reason test should apply, and now will turn to the next step of clarifying how the rule of reason should be applied once it is determined that is the appropriate test. In any event, given the lack of consistency in its application by lower courts, many antitrust practitioners are hopeful that the Supreme Court will take the opportunity to “steer” courts in the right direction when reviewing Amex’s anti-steering rules.

---


\(^{38}\) \textit{Id.} at 14–15.

\(^{39}\) \textit{Id.} at 9.


\(^{41}\) \textit{Id.} at 24.

\(^{42}\) \textit{Id.} at 26.

\(^{43}\) \textit{Id.} at 26–27 (quoting \textit{FTC v. Actavis}, 133 S. Ct. 2223, 2238 (2013)).
Fuzzy Rules, Expressions Hair Design v. Schneiderman, and Ohio v. American Express

Matthew Moloshok

In my February 2017 article for The Antitrust Source¹ I addressed the Supreme Court’s anticipated ruling in Expressions Hair Design v. Schneiderman.² That case involves a New York statute that prohibits sellers from imposing a surcharge on customers who purchase goods or services with a credit card.³ Merchants claim the law infringes their First Amendment right to freedom of speech. The U.S. Court of Appeals for the Second Circuit⁴ upheld the statute, finding it prohibits conduct and says nothing about speech at all. That result seemed correct, particularly because treating the statute as a speech prohibition could threaten a wide swath of consumer protection regulations and even some longstanding antitrust rules.⁵

On May 29, 2017, the Supreme Court issued its ruling.⁶ It unanimously remanded the case for further proceedings to determine whether the no-surcharge law can survive “intermediate scrutiny” as a “commercial speech regulation” or “disclosure requirement.” By subjecting the statute to First Amendment analysis, the Court confirmed most of my fears that, over the long term, that approach may jeopardize many regulations that one would have thought policed conduct rather than speech.

More immediately, the decision to remand keeps merchants, card issuers, card networks, cardholders, other consumers, and governments in suspense. Perhaps they have gotten used to remaining in limbo. Litigation over various aspects of charges imposed by credit card payment systems, whether and how they may be passed on to consumers, and whether merchants can steer consumers to less-costly credit networks has been ongoing for 40 years.

The Court has accepted certiorari in yet another credit card case this Term, Ohio v. American Express Company.⁷ There, the United States and several states challenged anti-steering policies that the Amex network imposes on merchants as violating Section 1 of the Sherman Act. Amex charges relatively high fees and merchants want the ability to redirect consumers to cards that charge lower fees to the merchants; the anti-steering rules preclude that. The district court, apply-

³ N.Y. Gen. Bus. Law § 518 (“No seller in any sales transaction may impose a surcharge on a holder who elects to use a credit card in lieu of payment by cash, check, or similar means . . . .”).
⁴ Expressions Hair Design v. Schneiderman, 808 F.3d 118 (2d Cir. 2015), vacated, 137 S. Ct. 1144 (2017).
⁵ Moloshok, supra note 1, at 8–9.
⁶ Expressions Hair Design, 137 S. Ct. 1144.
looking the rule of reason, found that the Amex rules raised merchants’ costs, excluded rivals, harmed consumers, and so constituted an antitrust violation.8 The Second Circuit reversed, believing the lower court had ignored characteristics of Amex’s allegedly “two-sided platform” for merchants and cardholders. That is, because “cardholders benefit from holding a card only if that card is accepted by a wide range of merchants, and merchants benefit from accepting a card only if a sufficient number of cardholders use it,” the markets for cardholders and merchants are interdependent, and Amex has to compete on both sides.9 The circuit court ruled that to prove an antitrust violation the government had to show that harm across either market had not been outweighed by procompetitive features of Amex’s system in allowing it to address both.10 In theory, the Supreme Court will now resolve the issues or at least clarify the operation of the rule of reason for use on a remand.

Looking at the no-surcharge saga and the American Express saga together makes one wonder about the entire enterprise of challenging credit card practices through litigation. The fuzzy standards available through litigation—the “rule of reason” in antitrust and “intermediate scrutiny” in the First Amendment sphere—have not allowed courts to resolve the propriety of credit card fees, surcharging, and anti-steering practices in a coherent or efficient manner. The better approach may not be litigation at all, but direct and comprehensive regulation of these practices.

The Supreme Court’s Ruling in Expressions Hair Design and Its Ramifications

On March 29, 2017, the Supreme Court issued its ruling in Expressions Hair Design, unanimously (8-0) vacating the judgment below. A brief majority opinion directed that, on remand, the Second Circuit “analyze” the New York no-surcharge law “as a speech regulation”11 and determine whether the no-surcharge provision constitutes either “a valid commercial speech regulation” under Central Hudson Gas & Electric Corp. v. Public Service Commission of New York12 or “a valid disclosure requirement” under Zauderer v. Office of Disciplinary Counsel of Supreme Court of Ohio.13 The Court declined to answer these questions itself, as the case had been tried and had proceeded in the Second Circuit as a “facial” challenge to the statute but before the Supreme Court had been recast and limited to a challenge “as applied” to merchants who posted a single cash price and disclosed that they would charge an additional amount or percentage above the cash price for use of a credit card.

Justice Breyer concurred only in the judgment. Observing that all commerce requires speech, he suggested “it is often wiser not to try to distinguish between ‘speech’ and ‘conduct,’” but instead to focus on the communicative interest at stake, on a sliding scale of the speech’s impor-

---

9 American Express, 838 F.3d at 185–86.
10 Id. at 206–07.
12 The Supreme Court disposed of two related cases five days later. It granted certiorari, vacated, and remanded for reconsideration in light of Expressions Hair Design a case in which the Fifth Circuit had sustained Texas’s no-surcharge statute as a conduct regulation. Rowell v. Pettijohn, 816 F.3d 73 (5th Cir. 2016), vacated and remanded, 137 S. Ct. 1431 (2017). The Court denied certiorari, however, in a case where the Eleventh Circuit had invalidated the Florida no-surcharge statute. Dana’s R.R. Supply v. Attorney Gen., Florida, 807 F.3d 1235 (11th Cir. 2015), cert. denied, 137 S. Ct. 1452 (2017).
distance, with political speech receiving high protections and speech in “ordinary” business transactions very little. 14

Justice Sotomayor, joined by Justice Alito, concurred only in the remand. Justice Sotomayor expressed disappointment that, without a better understanding of the actual operation of the statute, the majority’s opinion “addresses only one part of one half of petitioners’ First Amendment challenge” amounting to “[a] quarter-loaf outcome . . . worse than none.” 15 She urged the Second Circuit to certify questions to the New York Court of Appeals to obtain a definitive ruling as to what price disclosures/price charging practices the statute prescribed or to give a narrowing interpretation. 16

The no-surcharge law now sits in limbo, and it is likely that lengthy litigation will continue to unfold. 17 On remand, the Second Circuit, following the path recommended by Justices Sotomayor and Alito, has certified a single question to the New York Court of Appeals (that state’s highest court): “Does a merchant comply with New York’s General Business Law § 518 so long as the merchant posts the total-dollars-and-cents price charged to credit card users?” 18 The New York Court of Appeals accepted the certified question on January 11, 2018, and directed briefing and argument. 19 The Second Circuit has not signaled how a “yes” or “no” response to the question will impact its deliberations. 20

The Supreme Court’s reasoning and principal holding in Expressions Hair Design confirmed my concern that many regulations one would have thought to regulate conduct—say, directing use of unit pricing, directing calculation of an “annual percentage rate,” or, in the antitrust realm, information exchanges among competitors and various vertical restraints—might instead be deemed to involve speech and thus become harder to sustain. 21 The New York statute described what looked like conduct—charging credit card users a price in excess of the posted price. But the Court reasoned that if the statute did not command the prices that merchants could charge, leaving price setting to the merchants, the state had only regulated how to describe the proposed transaction—essentially how the merchant communicated the difference between the posted

14 Expressions Hair Design, 137 S. Ct. at 1152 (Breyer, J., concurring in judgment).

15 Id. at 1153 (Sotomayor, J., concurring in judgment).

16 Federal law had, for a time, similarly banned surcharging credit card users but, unlike New York, Congress had spelled out what types of conduct amounted to impermissible surcharging: “[A] merchant could violate the [federal] surcharge ban only by posting a single price and charging credit card users more than that posted price.” Id. at 1147. Justice Sotomayor agreed this was one of three possible interpretations of the New York statute as well, the others being that a surcharge existed whenever the merchant charges customers different amounts based on method of payment, if the merchant fails to set out a price comparison in the form of a discount rather than in the form of a surcharge.


18 Expressions Hair Design, 877 F.3d 99, 100, 107 (2d Cir. 2017). That would conform to the former federal model. Expressions Hair Design, 137 S. Ct. at 1147. In certifying the question, the Second Circuit stated: “Not only do the First Amendment issues in this case not, to our mind, admit of easy answers, but, more importantly, these issues cannot be addressed without initially considering how [New York Gen. Bus. Law] § 518’s restrictions operate in practice.” Expressions Hair Design, 877 F.3d at 102.


21 Moloshok, supra note 1, at 4 & n.22, 8–9.
cash price and the credit price that would ultimately be charged. In my earlier piece I urged that would be a strained and surprising way to read the statute, yet that is where we find ourselves.

The majority’s opinion in Expressions Hair throws open the door to the argument that any consumer protection regulation short of government’s fully taking direct control of prices or of product or service quality does not regulate conduct but speech about the price or the quality. That covers pretty much every regulation anyone has ever encountered. Almost any regulation impacting the prices that may be charged or the quality of what may be sold could then be re-characterized as a speech restriction because those prices and product or service specifications can only be communicated through speech. If producers and merchants push their litigation along those lines, and if courts do not rein them in, then nearly every consumer protection regulation will have to face heightened scrutiny under Central Hudson or have to be justified as a disclosure mechanism under Zauderer, rather than under the “rational basis” test that is used for most economic regulation.22

The Uncertain Reach of Central Hudson and Zauderer

Subjecting consumer protection regulation to First Amendment standards of review will make it harder to regulate merchant conduct and create substantial burdens and confusion for courts due to the uncertain reach of those standards.

The Central Hudson test “is significantly stricter than the rational basis test . . . .”23 It allows the regulation of truthful speech about lawful commercial activity24 only if the government’s interest is “substantial,” its regulation “directly” advances the government’s interest, and is “narrowly drawn” so as to be “not more extensive than is necessary to serve that interest.”25 Similarly, Zauderer allows states to mandate disclosures as part of commercial speech only to impart “purely factual and uncontroversial information about” the goods or services “reasonably related to the State’s interest in preventing deception of consumers.”26

While it seems easy to state the Central Hudson and Zauderer tests, lower courts confess that they find the tests hard to understand and administer.27 They cannot agree if they are separate tests at all: although Zauderer seems to provide a separate, more regulation-friendly standard, and the Supreme Court referred to the tests separately in Expressions Hair Design, many courts and judges have believed that the Central Hudson test must be satisfied before the government can mandate disclosures.28 Courts also cannot agree on the standard to use to determine if

24 If the commercial speech concerns unlawful activity or is fraudulent it may be proscribed entirely. In re R.M.J., 455 U.S. 191, 203 (1982).
26 Zauderer, 471 U.S. at 651.
27 See, e.g., Nat’l Ass’n of Mfrs. v. SEC, 800 F.3d 518, 524 (D.C. Cir. 2015) (noting “the flux and uncertainty of the First Amendment doctrine of commercial speech, and the conflict in the circuits regarding the reach of Zauderer . . . .”).
28 Compare Dana’s Railroad Supply, 807 F.3d at 1250 (treating Zauderer disclosures as a less restrictive alternative to banning speech on the subject, as a means to satisfy Central Hudson), and Am. Meat Inst. v. U.S. Dep’t of Agric., 760 F.3d 18, 33 (D.C. Cir. 2014) (Kavanaugh, J., concurring) (“Zauderer is best read simply as an application of Central Hudson, not a different test altogether.”), with American Meat Institute, 760 F.3d at 29 (Rogers, J., concurring) (“Zauderer does not reformulate the Central Hudson standard but rather establishes a different standard based on the ‘material differences between disclosure requirements and outright prohibitions on speech,’” (quoting Zauderer)); Expressions Hair Design, 877 F.3d at 103 (but acknowledging: “[T]he Supreme Court has never clearly specified a governing framework that determines when Zauderer’s less-exacting standard should apply instead of Central Hudson’s intermediate scrutiny.”).
compelled disclosures involve “purely factual and uncontroversial” information.29

Meanwhile, despite the Supreme Court’s repeated caution that courts must find regulations narrowly tailored so long as they promote the government’s interest in a way that “would be achieved less effectively absent the regulation,” and need not use the least restrictive means,30 lower courts often invalidate statutes they believe broader than necessary—at least that has been true of the decisions invalidating no-surcharge laws as speech restrictions under the Central Hudson and Zauderer standards.31

The prohibition against surcharging credit card users seems to advance directly at least two governmental interests: (1) a conduct interest in promoting credit card use to reduce use of cash (which, among other things, helps assure better sales tax collections, improves productivity through faster check-outs, facilitates returns of merchandise, and broadens access to online marketplaces), and (2) a disclosure interest in assuring that consumers can compare prices apple-to-apple, knowing that in all instances they will be presented with the highest price and then evaluate their possible savings, rather than being tempted by a lower price only to have to figure out the added cost of paying with a credit card later or be held up at the register.32

The risk of deception that consumers face absent a uniform way that allows them to know the highest price they will be asked to pay seems “obvious.”33 Adding a credit card surcharge at check-out is not akin to, say, having sales tax calculated and added at check-out. Sales tax is prescribed by state law, both as to amount and as to the requirement to pay it. In contrast, not every

---

29 Compare Nationwide Biweekly Admin., Inc. v. Owen, 873 F.3d 716, 732 (9th Cir. 2017) (“uncontroversial . . . refers to the factual accuracy of the compelled disclosure, not to its [disturbing] subjective impact . . .”) (citations and quotations omitted), with National Association of Manufacturers, 800 F.3d at 530 (unduly controversial to require disclosures that require “an issuer to confess blood on its hands . . .”) (citations and quotations omitted). See also American Meat Institute, 760 F.3d at 34 (Kavanaugh, J. concurring) (“[i]t is unclear how we should assess and what we should examine to determine whether a mandatory disclosure is controversial.”).


31 Expressions Hair Design v. Schneiderman, 975 F. Supp. 2d 430, 447 (S.D.N.Y. 2013) (Rakoff, J.) (invalid because “far broader than necessary”), vacated, 808 F.3d 118, vacated, 137 S. Ct. 1144; Dana’s Railroad Supply, 807 F.3d at 1250 (invalid because “available less restrictive alternatives are legion”). None of these decisions expressly considered whether the legislatures could have deemed the suggested alternatives “less effective” for its purposes. The Second Circuit is free to apply its own analysis to the issues because the Supreme Court did not comment on, much less approve, how those decisions applied the Central Hudson and Zauderer tests.

In January of this year, the Ninth Circuit weighed in on the issue. See Italian Colors Rest. v. Becerra, 878 F.3d 1165 (9th Cir. 2018), affirming, “as applied” to a merchant posting a single price and adding a surcharge, the lower court’s 2015 grant of summary judgment invalidating the California no-surcharge statute. Finding that the Supreme Court’s ruling in Expressions Hair Design requires the California statute to be deemed to prohibit speech not conduct, the Ninth Circuit held the statute fails the Central Hudson test because the state could not show the statute directly advanced its interests, where it had not presented evidence “that surcharges posed economic dangers that were in fact real before [enactment] or that [the statute] actually alleviates these harms to a material degree” and was not “narrowly tailored,” where the state could have banned deceptive surcharging, required affirmative disclosures pre-sale, or simply used ordinary rules prohibiting unfair business practices or misleading advertising. Id. at 1177–78. Remarkably, the Ninth Circuit does not mention Zauderer, even though the Supreme Court’s ruling in Expressions Hair Design had raised the question directly.

32 Cf. Her Majesty’s Treasury, Explanatory Memorandum to the Payment Services Regulation 2017, 2017 No. 752 § 7.16 (explaining that a new United Kingdom ban on retailers’ surcharging “is intended to . . . create a clearer picture for consumers in which they know the full price of the product/service they are purchasing upfront and are confident that there will be no additional charges when they come to pay using a particular payment instrument.”), http://www.legislation.gov.uk/uksi/2017/752/pdfs/uksiem_20170752_en.pdf.

33 “When the possibility of deception is as self-evident as it is in this case, we need not require the State to ‘conduct a survey of the . . . public before it [may] determine that the [advertisement] had a tendency to mislead.’” Zauderer, 471 U.S. at 652–53 (citation omitted); Milavetz, Gallop & Milavetz, P.A. v. United States, 559 U.S. 229, 251 (2010) (Zauderer forecloses argument that the state cannot compel disclosures before having proof of prior actual deception). The standard under Central Hudson seems harder to satisfy. See Edenfield v. Fane, 507 U.S. 761, 770–71 (1993) (overturning ban on personal solicitation by accountants where state had not pointed to studies or anecdotal evidence in support of its law).
merchant will impose a surcharge, and those who do will vary in their policies as to how the upcharge is disclosed, the amounts they will add, or the methods they will use to calculate a surcharge. The suggestion that ordinary unfair business practice and misleading advertising laws might adequately serve to address the legislature’s goals ignores the interest in promoting credit card usage entirely, and in all events returns to my original concern: importing First Amendment analysis significantly impedes consumer protection because enforcement under those enactments would equally have to satisfy the Central Hudson/Zauderer standards, whatever they may be.

Ohio v. American Express and the Rule of Reason
First Amendment litigation brings ill-fitting standards to fundamentally economic transactions. Antitrust litigation, however, has shown itself a precise means of rooting out anticompetitive practices in the credit card industry, at least where it has confronted practices that result from horizontal agreements among credit card issuers. But antitrust litigation has been less successful in addressing credit card practices that arise from vertical arrangements, such as the anti-steering rules at issue in Ohio v. American Express.

The predominant antitrust thinking since the 1970s views vertical restraints as generally pro-competitive—an approach that makes it hard, costly, and complicated to challenge them under the rule of reason. Congress banned contracts, combinations, and conspiracies in restraint of trade, but the judiciary long ago limited that to “unreasonable” restraints.” The saga of Ohio v. American Express indicates this may be too fuzzy a standard to yield reliable results when applied to particular vertical practices. Indeed, as Allan L. Shampine recently noted, because the rule of reason is (and I would suggest can only be) laid out in “broad strokes” there is “a great deal of uncertainty as to the specifics of its application” when determining what counts as an unreasonable restraint.

These problems are on full display in Ohio v. American Express. Absent Congressional definition, courts typically look to economic learning to decide what is unreasonable. But there appears to be no consensus view among economists of how to evaluate the economic effects, costs, and benefits of credit card payment systems as currently established. Nor does there appear to be a consensus view among economists as to how to understand and evaluate “platforms” or two-

---

34 See Italian Colors, 878 F.3d at 1178.
35 See Adam M. Samaha & Roy Germano, Are Commercial Speech Cases Ideological? An Empirical Inquiry, 25 WM. & MARY BILL RTS. J. 827, 831 (2017) (“[I]t has never been clear that constitutional judicial review in this field has been worth the price in litigation costs and policy flexibility, or whether judges are sufficiently able to address emerging challenges such as business or regulator leveraging of big data analytics and cognitive nudges.”).
36 Bruce D. Sokler, Robert G. Kidwell & Farrah Short, What Have Merchants Gained from Payment Card Antitrust Litigation? ANTITRUST ADVISORY (Aug. 3, 2016) (summarizing some accomplishments and setbacks), https://www.mintz.com/legal-insights/alerts/articletype/articleview/articleid/3548/what-have-merchants-gained-from-payment-card-antitrust-litigation. Among other things, through those actions, MasterCard and Visa have entered consent orders that remove their prohibitions against surcharging and steering. But these accomplishments do not cap their fees, and so consumers still bear these costs when passed on through surcharges.
sided markets, or whether credit card networks even properly should be considered as two-sided markets. So even if the courts had already developed clear rules on applying the rule of reason (and they have not to date), they would have difficulty reaching reliable results in this context. That the district court and Court of Appeals would disagree on these issues, in light of these uncertainties, should surprise no one. Whether the Supreme Court can actually make better sense of it is perhaps wishful thinking if the economic learning is in turmoil, and the rule of reason itself provides insufficient guideposts to select appropriate standards.

The Regulatory Experiment

Society, not just the merchants and card networks, has a great interest in getting the issues of merchant surcharges and anti-steering rules resolved, and in the manner that best protects consumers. Credit cards are a prominent feature of the economy and commercial activity. The antitrust and First Amendment litigation logjams do not appear to be resolving the issues, and certainly are not directly addressing the ultimate costs borne by consumers. There is a third approach that deserves consideration—direct regulation, only on a more comprehensive basis than previously applied. This seems preferable to leaving courts struggling to apply the fuzzy rules available to them under the First Amendment and the rule of reason.

Expressions Hair Design confirms, once again, that government can set prices, and this may be a highly appropriate circumstance in which to do so. Congress has, if only intermittently and inconsistently, entered the fray before to ban surcharging temporarily, to allow its ban on surcharging to expire, and to require that credit card issuers allow merchants to offer discounts for use of cash. The federal government has also limited the interchange (“swipe”) fees charged in debit card transactions (but not in credit card transactions) conducted through the four-party payment systems, such as Visa and MasterCard (but not the fees charged by three-party payment systems, such as Amex and the Diner’s Club division of Discover).


41 “If a restraint is reasonably necessary to achieve procompetitive effects, those effects must be balanced against the restraint’s anticompetitive effects,” yet “[t]he Supreme Court has provided little guidance about how this balancing process should be conducted . . . .” 1 ABA SECTION OF ANTITRUST LAW, ANTITRUST LAW DEVELOPMENTS 79–80 (8th ed. 2017). Indeed, the United States, while asserting the Second Circuit’s decision in Ohio v. American Express was in error, nonetheless had originally opposed the petition for certiorari due to the dearth of precedent and the impact the lack of precedent might have on the proofs presented in the trial court. Brief for the United States in Opposition, No. 16-1454, at 30–36 (Aug. 21, 2017), http://www.scotusblog.com/wp-content/uploads/2017/08/16-1454-BIO.pdf.


43 Cash Discount Act § 201, 95 Stat. at 144.


But these efforts have been a patchwork, and regulating only one portion of the interchange fee/surcharge issue leads to scrambled means of evasion or an imposition on consumers, as experience in Australia and the European Union has demonstrated. Both Australia and the European Union have therefore decided to take much more comprehensive steps through regulation, although they have elected to take markedly divergent approaches.46

As part of a 2003 reform of its credit card sector, Australia essentially capped credit card interchange fees and simultaneously lifted all bans against merchants surcharging for credit card use.47 While it expected that merchants would thereby be able to recoup the fees, studies by the regulators showed that merchants often charged far more than their actual costs.48 Effective January 1, 2013, Australia therefore decided to limit merchants to a “reasonable” mark-up but declined to be more specific, in view of the differing characteristics of various merchants and credit card networks.49

Concluding that merchants abused the wiggle room this allowed, Australia adopted a new Competition and Consumer Amendment (Payment Surcharges) Act 2016 under which, if a merchant chooses to impose a surcharge, it cannot charge the customer any “more than what it costs the business to process the payment,” and that rule applies to all merchants, “regardless of their size.”50 The new Australian surcharge limit applies to all point of sale electronic fund transfer systems (debit and prepaid), MasterCard (credit, debit, and prepaid), Visa (credit, debit, and prepaid), and “American Express cards issued by Australian banks,”51 i.e., cards issued through a four-party payment system. While that might seem to leave Amex’s traditional three-party payment system cards free of regulation, it appears that Amex removed anti-steering provisions from its Australian contracts more than a decade ago,52 and has given an undertaking that its proprietary Amex cards will comply with the key provisions of the regulation.53 Thus, as matters now stand, Australian merchants can steer to lower-charging cards and also surcharge, but only by the amount they have actually been charged by the credit card company. Europe, which also opted for comprehensive regulation, has gone in a different direction—while it has prohibited anti-steering rules, it has largely banned merchants surcharging their customers.

46 For a more comprehensive summary than space here allows, see Ann Wardrop, Dealing with Excess: Regulatory Perspectives on Surcharging for Payment, 36 Univ. of Queensland L.J. 99 (2017).
49 Id. at 9 (difficulties in defining reasonable surcharges), 17–19 (advocating against prescribing “reasonable” surcharges), 21–25 (adopting Standard No. 2).
Its Second Payment Services Directive (PSD2),\(^{54}\) effective January 13, 2018, requires Member States to prohibit merchant surcharging in transactions where the interchange fee has been capped by regulations.\(^{55}\) The United Kingdom has adopted regulations to implement this.\(^{56}\)

As might be expected, without the ability to surcharge, merchants seek other ways to recoup the costs (or more) from their customers, such as across-the-board “service” fees, raising shelf prices, and requiring use of other payment media (cash, debit cards).\(^{57}\) One might despair then of pleasing all constituents, or protecting consumers or merchants perfectly, even through regulation.

But direct legislation and regulation appear to offer several advantages as compared with litigation—expertise (banking and consumer regulators having a better understanding of the issues than lay judges); speed (even when rulemaking is challenged it can often be concluded in the span of a few years, rather than decades\(^{58}\)); and clarity (the rule can clearly define what conduct is and is not permitted, as well as explain what choices have been made in a manner understandable to all constituents, thereby ending economists and lawyers talking past each other)—and thereafter enable (relatively) simple administration as compared with ongoing court proceedings. This seems a significant improvement as compared with the fuzzy rules of the First Amendment and antitrust’s rule of reason.

Will a legislative/regulatory fix favor credit card networks, issuers, and banks—the usual regulatory capture? No doubt credit card issuers and banks are heavy hitters in the worlds of lobbying and administrative rule-making, but so are merchant and consumer groups, and it is not as if both sides were not heavy hitters in the courts as well.

Will a legislative/regulatory fix ossify the card payment industry as compared with a more case-specific, granular approach allowed by (say) the rule of reason? It is a risk, but repeal, clarification, and amendment are all available, with more flexibility than trying to obtain reconsideration of a judicial ruling, relief from a consent order, or the years and years of discovery, expert testimony, and appeals needed to address an alleged rule of reason violation, by which time technology may have moved on anyway and disruptive new payment systems will have entered the market. After 40 years of slogging through litigation and uncertainty, it now may be well worth the full bore regulatory experiment.●

---


55 Id., preamble (paragraph 66) and art. 62(4). In paragraph 28 of the preamble to Regulation (EU) 2015/751, http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32015R0751&from=EN, the EU recognized that Member States could exempt three-party payment systems (like Amex) from fee caps during a “transitional period.” Theoretically, this could allow merchants handling Amex cards to try to surcharge. As to circumstances in which Amex should be deemed to be operating a “four party” system, where fees are capped, see Case 304/16, American Express Co. v. The Lord Commissioners of Her Majesty’s Treasury, ECLI:EU:C:2017:524 (Opinion of the Advocate General, July 6, 2017), followed by Cases C-304/16 and C-643/16, The Queen on application of American Express Company v. The Lord Commissioners of Her Majesty’s Treasury, ECLI:EU:C:2018:66 and ECLI:EU:C:2018:67 (First Chamber Feb. 7, 2018).


58 Such was the speed with which the litigation concerning the Durbin Amendment, capping certain debit card interchange fees, was resolved. See supra note 45.
Raising the Bar: How Does China’s Tetra Pak Decision Measure Up to the ECJ’s Requirements for Loyalty Rebate Cases?

Fei Deng and Su Sun

In November 2016, about ten months prior to the European Court of Justice’s (ECJ) Intel judgment, China’s State Administration of Industry and Commerce (SAIC) fined the Swedish firm Tetra Pak $97 million, after almost five years of investigation, for allegedly abusing its dominant position in the aseptic carton packaging equipment, materials, and service markets in China. The SAIC published an unprecedentedly lengthy 47-page decision, in which it condemned three types of business practices allegedly adopted by Tetra Pak: tie-in sales, exclusive dealing, and loyalty rebates. Exclusive dealing and tying are explicitly specified in Articles 17(4) and (5) of China’s Antimonopoly Law (AML), respectively, as prohibited behaviors of a dominant firm that constitute abuse. Thus, the SAIC had an easier job establishing these offenses after showing that Tetra Pak had dominance in the relevant markets, relative to finding illegal conduct in loyalty rebates, which are not specified in the AML.

Loyalty rebates are not explicitly addressed in the AML. The most applicable clause might be Article 17(2), which states that a dominant firm is prohibited from “selling commodities at a price below cost without proper reason” as well as the catch-all Article 17(7), identifying “other behaviors that are considered by the Antimonopoly Enforcement Agency as abuse of the dominant position in the market.” With the uncertainty in the law, how the SAIC analyzed loyalty rebates has drawn significant attention from antitrust practitioners.

Loyalty rebates have been in the spotlight around the world in recent years. A number of U.S. cases have drawn much attention, and there have been extensive discussions and intense debates in this area. The ECJ’s recent Intel judgment resulted in additional heated discussion on this subject. While the Intel judgment tried to clarify a number of legal issues, arguably some ambiguities remain. However, the ECJ’s Intel judgment raised the bar by making it more difficult

---

1 SAIC Competition Enforcement Public Notice No. 10 (2016) [hereinafter Tetra Pak Decision], http://www.saic.gov.cn/fldyfbzdj/jzzfgg/201703/t20170309_232289.html. The SAIC is a ministry-level administrative agency that has a broad range of mandates in regulating the marketplace in China, including investigating non-merger and non-price related monopolization conduct. Such antitrust investigations are carried out by its Anti-monopoly and Anti-unfair Competition Enforcement Bureau. As is commonly done in antitrust writings relating to China, we use SAIC as the acronym for its antitrust enforcement bureau.

2 Examples of cases that reached the circuit courts include: Eisai Inc. v. Sanofi-Aventis U.S., LLC, 821 F.3d 394 (3d Cir. 2016); ZF Meritor, LLC v. Eaton Corp., 696 F.3d 254 (3d Cir. 2012); and Allied Orthopedic Appliances v. Tyco Health Care Group LP, 592 F.3d 991 (9th Cir. 2009). For discussions of these and other loyalty rebates cases in the United States, see, e.g., Jonathan M. Jacobson & Daniel P. Weick, Countering Exclusion: The Complainant’s Obligation, 81 ANTITRUST L.J. 423 (2017); Derek W. Moore & Joshua D. Wright, Conditional Discounts and the Law of Exclusive Dealing, 22 GEO. MASON L. REV. 1205 (2015); and Richard M. Steuer, Musthavedness, 81 ANTITRUST L.J. 447 (2017).


to condemn loyalty rebates provided by a dominant firm in Europe, in essence requiring a rule of reason analysis of the rebate programs to evaluate their capacity to foreclose rivals and potential offsetting efficiency gains, as opposed to finding a per se violation of the EU competition law. In particular, the ECJ laid out the economic analyses that are required of the European Commission (EC) to prove liability:

[T]he Commission is not only required to analyse, first, the extent of the undertaking’s dominant position on the relevant market and, secondly, the share of the market covered by the challenged practice, as well as the conditions and arrangements for granting the rebates in question, their duration and their amount; it is also required to assess the possible existence of a strategy aiming to exclude competitors that are at least as efficient as the dominant undertaking from the market. . . . The analysis of the capacity to foreclose is also relevant in assessing whether a system of rebates which, in principle, falls within the scope of the prohibition laid down in Article 102 TFEU, may be objectively justified. It has to be determined whether the exclusionary effect arising from such a system, which is disadvantageous for competition, may be counterbalanced, or outweighed, by advantages in terms of efficiency which also benefit the consumer . . . .

In this article, we look back at the SAIC’s Tetra Pak decision in light of the ECJ’s Intel judgment. We start with a comparison between the two cases, the SAIC’s Tetra Pak investigation and the EC’s Intel investigation, the latter of which resulted in a decision issued by the EC in 2009 and the ECJ judgment in 2017. We compare how dominance was shown and how loyalty rebates were approached in the two decisions. We then explore how the SAIC’s Tetra Pak decision measures up to the ECJ’s requirements for a finding of liability in loyalty rebates cases. In particular, we focus on a test that was at the heart of the Intel case and implied in the Tetra Pak case, as well as the analysis of the exclusionary effects in the two cases. By comparing these two decisions, we do not mean that the ECJ’s decision should govern where the SAIC, an enforcement agency in a different jurisdiction, will be headed in the future, but rather, in the context of antitrust globalization, we explore what one would expect to see if a similar standard were followed in adjudicating such cases in China.

**Dominance**

In its Tetra Pak decision, the SAIC first defined three relevant product markets based on demand and supply substitution analyses in accordance with China’s Guidelines on the Delineation of Relevant Markets: aseptic carton packaging equipment, technology services of aseptic carton packaging equipment, and aseptic carton packaging materials.

For each of these three relevant product markets, the SAIC determined that Tetra Pak was dominant based on the examination of four factors, as required by Article 18 of China’s AML: (1) Tetra Pak’s market share and the level of competition it faced in the relevant market; (2) Tetra Pak’s ability...
ity to control the relevant market, including an examination of the contractual conditions Tetra Pak imposed on buyers; (3) buyers’ reliance on Tetra Pak for the relevant products; and (4) barriers to entry.

Unlike Intel’s position in the relevant market in EU’s Intel case, where AMD was the only practical competitor, Tetra Pak faced competition from a number of (much) smaller rivals in each relevant market. Despite finding Tetra Pak’s dominance in the three relevant markets, the SAIC acknowledged that Tetra Pak’s shares in these markets, though above 50 percent, were decreasing in recent years. This latter observation on Tetra Pak’s diminishing market shares does cast some doubt on the competitive impact of its alleged anticompetitive conduct, which we will discuss below.

**Nature of Conduct: “Fidelity-Inducing” Rebates vs. Exclusivity Rebates**

While the EC’s allegations against Intel are based on exclusivity rebates, i.e., rebates conditional on a customer purchasing all or almost all of its volume from Intel, the SAIC’s allegations against Tetra Pak were based on “fidelity-inducing” rebates, i.e., loyalty rebates, which are less restrictive as they generally lack the explicit condition that a customer has to purchase all or almost all of its volume from the dominant firm.

Specifically, there are two types of “fidelity-inducing” rebates that the SAIC found problematic—“retroactive cumulative rebates” and “individualized targeted rebates.” The former are described as setting cumulative rebates (i.e., rebates applied not only to the additional volume currently being purchased, but to the total volume of all past purchases within a certain period) of increasing size as the threshold increases. The SAIC alleged that Tetra Pak offered this type of rebate not only on individual product lines but also on a combination of various product lines of packaging materials. The latter type of rebate is described as a rebate conditional on a customer purchasing a certain amount or a percentage of volume from Tetra Pak within a certain time period, with the targeted volume (specified in amount or percentage) set specifically for that individual customer.

Fidelity-inducing rebates are typically less restrictive than exclusivity rebates. However, if the dominant firm has good knowledge about its customers’ demands, which is often the case, it could set up fidelity-inducing rebates with carefully constructed targets and incentives, so that they become in essence exclusivity rebates. Nonetheless, there are circumstances where the dominant seller may not want to achieve total exclusivity. From an economist’s perspective, it makes sense to analyze fidelity-inducing rebates and exclusive rebates in the same manner.

Even for exclusivity rebates, the ECJ’s Intel judgment seems to endorse a rule of reason approach if the alleged dominant firm submits evidence “that its conduct was not capable of restricting competition and, in particular, of producing the alleged foreclosure effects.” In comparison, in the SAIC’s Tetra Pak decision, the SAIC took the rule of reason approach analyzing the facts and effects of the exclusivity agreement Tetra Pak signed with a key raw packaging materi-
als provider that prohibited it from supplying such materials to third parties. Though this exclusivity agreement was with an upstream supplier instead of a downstream customer, it does show SAIC's willingness to conduct a rule of reason analysis in examining exclusivity agreements, which is to be commended in our opinion given the potential economic efficiencies inherent in such exclusivity provisions that should be taken into account.

**As Efficient Competitor (AEC) Test**

It appears that the AEC test plays a key role in both the EC's *Intel* decision and the SAIC's *Tetra Pak* decision. The purpose of the test is to see whether the dominant firm’s specific conduct, for example, the use of loyalty rebates, is capable of excluding a competitor that is as efficient as the dominant firm. Although the EC did not consider an AEC test necessary, its discussion of the test in its *Intel* decision was extensive, accounting for almost a third of the total length of the decision.

In the SAIC’s *Tetra Pak* decision, the SAIC concluded that Tetra Pak’s loyalty discount programs “caused foreclosure to its competitors in the short run, and resulted in their inability to compete with Tetra Pak at the same or similar costs.” While the basic rationale and formula applied by both agencies seem largely consistent, the EC’s *Intel* decision provides more details on the inputs to the calculations, whereas the SAIC’s *Tetra Pak* decision does not. However, as explained in ECJ’s *Intel* judgment, how exactly the AEC test should be carried out can be a subject of intense dispute between the private party and the agency, and may lead to different conclusions. In addition, whether or not the AEC test is an appropriate test to evaluate loyalty rebates in the first place is still a question under debate.

Although not explicitly labeled as an AEC test, the SAIC’s *Tetra Pak* decision explains the basis for its finding of the anticompetitive effects of Tetra Pak’s loyalty rebates as follows: First, it lays out the basic rationale and formula for the evaluation. Assuming that a customer’s demand for packaging materials is Q, it could get a rebate of d (%) if it purchased all its volume from Tetra Pak, but could only get a discount of d – H9004(d %) if it purchased only part of its demand from Tetra Pak, specified as Q – Q2. By choosing not to purchase all of its volume from Tetra Pak, the customer receives a smaller rebate on the portion of volume that stays with Tetra Pak, and the total amount of the loss equals (Q – Q2)*H9004d. In order for the customer to be willing to do so, this amount has to be “made up” on the volume that the customer purchases from the competitor, Q2. In other words, in order for a competitor to obtain part of this customer’s volume, Q2, it is not enough for this competitor to offer a discount of d, but in addition it has to compensate the customer for the

---

16 *Intel Commission Decision*, supra note 7, ¶ 1155.

17 Id. ¶¶ 1002–1576.

18 *Tetra Pak Decision*, supra note 1, at 46.

19 *Intel ECJ Judgment*, supra note 4, ¶ 132.


21 *Tetra Pak Decision*, supra note 1, at 37–42.

22 Although the illustration provided by the SAIC is based on exclusivity rebates, similar arithmetic is applicable to “fidelity-inducing” rebates.
additional loss in discount that the customer incurs by switching, which equals \((Q - Q_2)^*d/Q_2\). Therefore, the total discount offered by the competitor would be \(d + (Q - Q_2)^*d/Q_2\).

Second, the SAIC decision explains that the SAIC conducted statistical analyses for the top 30 customers of Tetra Pak and found that a competitor would have to offer as high as twice the discount offered by Tetra Pak in order to obtain a “small volume of contestable demand.”23 The decision further states that “the contestable demand is limited due to the specific market conditions in this case, thus requiring the competitor’s matching price to be very low.”24 However, the decision does not provide any information on how low the competitor’s matching price has to be, and whether the level is below some cost measure. Judging from the decision, one may infer that the SAIC’s application of the AEC test may not be as complete as that of the EC in the 2009 Intel decision as explained below. The SAIC’s Tetra Pak decision concludes that the rebates “made it difficult for a competitor to participate in the competition, even possibly driving the competitor out of the competition, therefore inducing the customer to choose Tetra Pak, foreclosing competitors, and eliminating and restricting market competition.”25

In contrast, in the EC’s 2009 Intel decision, in order to determine whether Intel’s exclusivity rebates were capable of causing anticompetitive foreclosure in the relevant market, the EC conducted an AEC test, positing a hypothetical firm that wants to enter the market and is at least as efficient as Intel. It tested whether there was a price this firm could charge to the Original Equipment Manufacturers (OEMs) that would allow it to stay in business and to compensate its OEM customers for losing rebates that they would have gotten from Intel.26 Specifically, the cost benchmark against which the EC compares this firm’s price, to judge whether this firm would be able to stay in business, is average avoidable cost, for which the EC uses Intel’s Cost of Goods Sold as a proxy.27 Intel contended that the EC wrongly carried out the AEC tests, for example, with respect to cost measures,28 forgone rebates,29 and contestable shares.30 Intel put forth its own AEC tests.31 The ECJ ordered the General Court to retry the case, taking into consideration Intel’s line of argument regarding EC’s errors in regards to the AEC test.32

Comparing the AEC tests implied in the SAIC’s Tetra Pak decision and explicitly conducted in the EC’s Intel decision, the EC decision provides more details, including discussions of the cost measures that are not present in the SAIC decision. Perhaps more importantly, the EC decision describes the counter arguments raised by the alleged offender and provides detailed responses to such arguments, which offers a balanced and transparent disposition.

23 Tetra Pak Decision, supra note 1, at 42.
24 Id. at 45.
25 Id.
26 Intel Comm’n Decision, supra note 7, ¶ 1003.
27 Id. ¶¶ 1037, 1043.
28 Id. ¶¶ 1045–1050.
29 Id. ¶¶ 1160–1164, 1195–1197.
31 Id. ¶¶ 1038, 1040.
32 Intel ECJ Judgment, supra note 4, ¶ 138–147.
Exclusionary Effect of Loyalty Rebates

The ECJ considers the analysis of the capability of restricting competition to be relevant in assessing whether loyalty rebates should be prohibited under the EU law.\(^{33}\) One possible interpretation of this judgment is that “capability” of restricting competition, as opposed to actual foreclosure effect, arguably sets a low threshold that would only require a certain level of plausibility that the conduct could restrict competition.\(^{34}\) Based on this interpretation of the low threshold, the SAIC’s *Tetra Pak* decision seems to have provided ample evidence to support such capacity to foreclose. However, focusing too much on theoretical capability runs the risk of conflicting with empirical observations that may indicate a lack of actual foreclosure effects.

As explained above, it appears from the SAIC’s *Tetra Pak* decision that the SAIC presumed a finding of foreclosure effects, after concluding that the rebates offered by Tetra Pak required the competitors to offer a lower, and sometimes much lower, price. The conclusion seems inconsistent with other facts that suggest a lack of actual foreclosure effects. For example, in the earlier part of its decision, the SAIC listed ten other competitors in China’s aseptic carton packaging materials market, and stated that Tetra Pak’s share had been decreasing over the period of 2009–2013, albeit it was above 60 percent in all years.\(^{35}\) There is also no mention of a single instance of a competitor’s exit from the marketplace due to the loyalty rebates, or tying and exclusive dealing practices alleged in the SAIC decision. Indeed, the market facts seem to suggest that Tetra Pak’s rebates could have been a procompetitive response to increasing competition.

It is possible that the foreclosure effects work through raising rivals’ costs by limiting rivals’ scale of operation, but not by eliminating the rivals from the market. In the *McWane* case, the Federal Trade Commission issued a decision finding that McWane unlawfully monopolized the domestic fittings market through its Full Support Program, which foreclosed potential entrants from accessing distributors.\(^{36}\) The FTC decision was upheld by the Eleventh Circuit.\(^{37}\) In that case, McWane’s rival and a potential entrant, Star Pipe Products, actually entered the market. But McWane’s exclusive dealing policy was considered to have raised Star Pipe Products’ distribution costs and prevented it from achieving the minimum efficient scale.\(^{38}\) Though the SAIC’s *Tetra Pak* decision mentions the loyalty discounts’ effects on competitors’ capacity utilization and the ability to expand,\(^{39}\) there is certainly room for better development of this argument, perhaps with quantitative evidence and analysis. After all, procompetitive price responses by a dominant firm will also have the effect of limiting competitors’ capacity utilization.

Additionally, in contrast to the EC’s 2009 Intel decision, the *Tetra Pak* decision does not offer detailed customer-specific analysis of the dominant firm’s attempt to delay or stop the customer from switching to a competitor using its loyalty discounts, which would have been more consis-
tent with a rule of reason approach supported by the ECJ for proving that the conduct is “capable of restricting competition.” 40

Conclusion

Loyalty rebates have been in the antitrust headlines worldwide in recent years, and such cases continue to emerge. 41 Both antitrust enforcement agencies and the courts in different jurisdictions need to tackle these difficult issues. With no clear consensus on the assessment criterion in sight, it is desirable for the agencies and courts to take a prudent approach.

The ECJ’s Intel decision provides some clarification and support for a rule of reason approach that requires clear and detailed showing of factual evidence and economic analysis. Against this backdrop and in retrospect, the SAIC’s Tetra Pak decision was generally consistent with the principles advocated in the ECJ’s Intel decision, but on the other hand there is clearly room for improvement in providing more details of its analysis and in showing actual exclusionary effects.

40 Intel ECJ Judgment, supra note 4, ¶ 149.

What Is So Significant About Statistical Significance?

Andrew E. Abere

Expert economic testimony has become increasingly important in antitrust cases. For good reason, many of the most crucial issues in antitrust cases hinge on empirical economic evidence, such as levels and trends in prices, market shares, sales, margins, profitability, and shipment patterns. Further, more sophisticated analyses, such as econometric and statistical studies, now play a larger and more central role in antitrust litigation. It is common in such cases for experts to report and be questioned about whether the results of such studies are “statistically significant.” For example, if an economist presented an estimate of overcharges that was not statistically significant, a question would almost certainly arise as to whether the overcharge estimate was reliable enough to form a basis for an award of damages. This article explains, from a layperson’s perspective, what exactly statistical significance is (and is not), and explores the growing debate over its relevance in scientific research and legal cases.¹

Surprisingly, the concept of statistical significance dates back only about 100 years. It was developed in the (dismal) science of economics, but ignored at first. During the course of the 20th century, however, it began to gain greater acceptance not only in economics, but also in other social sciences as well as in applied sciences. The concept also began to become an important issue in the presentation of legal evidence. For example, it was relied upon by courts in employment discrimination cases in the 1970s. Fast-forwarding to today, the concept is widely used not only in antitrust and consumer protection cases, but also in many other areas of the law. It is also, however, widely misinterpreted and misused.

What Is Statistical Significance?

Consider the example of a clinical drug trial. Suppose a pharmaceutical company develops a new anti-obesity drug. Two of its researchers test the effectiveness of the drug (“efficacy”) by randomly assigning a group of obese individuals into two study groups. The first is called the “treatment group,” which means that its members are given the drug. The second is the “control group,” whose members are given a placebo. The researchers calculate and compare the differences in weight changes at the beginning and end of the trial between the two groups. (This is sometimes referred to as the “difference in differences” approach.) They find the members of the treatment group lost an average of 30 pounds while the members of the control group lost an average of 10 pounds.

¹ There is much on the topic that is beyond the scope of this article. For readers looking to deepen their knowledge of the subject, see David H. Kaye & David A. Freeman, Reference Guide on Statistics, in Reference Manual on Scientific Evidence 3d ed. (2011). See also Phillip Johnson, Edward Leamer & Jeffrey Leitzinger, Statistical Significance and Statistical Error in Antitrust Analysis, 81 ANTITRUST L.J. 641 (2017) (offering a more technical discussion of the issue and suggestions for how more individualized guidelines for the use of statistical significance might be created and applied).
One researcher (the “believer”) says the drug works because participants who were given the drug lost, on average, 20 more pounds than those who were not given the drug. The other (the “skeptic”) says the drug does not work, because the difference in weight loss could be the result of influences other than the drug. What might those influences or sources of error be? To answer that question it is first useful to understand why researchers study samples.

Researchers often want to estimate some numerical fact, known as a “parameter,” about a class of individuals, known as the “population.” For example, what is the average personal income in the United States? Studying the whole population is usually impractical, so researchers instead estimate average personal income using a “statistic,” which is computed from a “sample,” or subset of the population. Here, the two drug researchers know the estimated effect of the drug from a sample of obese individuals. What they want to know, however, is the reliability of this statistic to predict the weight change for the entire population. The statistic is an estimate of what weight loss would be across the entire population (the parameter), but may not be equal to the parameter for a number of reasons.

Factors that Can Impact Reliability

First, there may be selection bias. For example, the participants who agree to be part of the trial may “self-select” and be individuals for whom weight loss drugs have been more effective than for the population as a whole. In this case, the observed difference may be due to differences in the participants rather than the drug. Second, there may be measurement error. For example, the scale (or scales) used to measure the participants’ weights may measure weight with some error. In this case, the observed difference may be due to differences in the readings of the scale (or scales) rather than the drug. Third, there may be chance sampling error (or in less technical language, the “luck of the draw”). I focus on this error in the following discussion.

Why does sampling error exist? It is often the case that there is variation among the individuals in the population, such that the characteristics of the sample are likely not to be the same as the characteristics of the entire relevant population. For example, since personal income varies across individuals, it is unlikely that the average personal income of a sample of people in the United States will equal the average personal income of all of the people in the United States. If a researcher surveyed everyone in his neighborhood, he would know with certainty the average personal income of people in the neighborhood (if that neighborhood were the population of interest), but using that neighborhood’s average to estimate the average personal income of the entire country would likely be pretty inaccurate. This is one reason why designers of surveys worry about getting responses from people from many different walks of life and parts of the country.

How would sampling error be at work in the clinical drug trial example? Suppose the drug is ineffective, such that it is no better than the placebo at helping obese individuals lose weight. Suppose also that despite the participants being assigned to the treatment and the control groups through a random process, the people in the treatment group are more active (and burn more calories as a result) than the people in the control group. In this case, the observed difference in weight loss may be due to the higher level of activity of the participants in the treatment group rather than the drug. One can easily imagine that other factors for which the researchers did not control, such as diet, gender, age, and genetic makeup, might also have played a role in weight loss.

---

2 Measurement error can also be a problem even when researchers study the entire relevant population rather than just a sample drawn from the population.
The reader might ask, “Why not just control for these other factors and eliminate any ‘luck of the draw?’” In practice, researchers do control for these other factors. In the drug trial example, the researchers could collect information on observable factors such as diet, activity level, gender, age, and genetic makeup and control for those in an attempt to eliminate the effect of their roles in weight loss.³

There may remain, however, unobservable factors for which researchers cannot control. For example, suppose the drug is ineffective, such that it is no better than the placebo at helping obese individuals lose weight. Suppose also, though, that the placebo effect may vary across individuals and that it is not possible for the researchers to observe the variation in the placebo effect in the participants. As a result, they cannot control for variation in the placebo effect as they might do for diet or activity level (which can be observed). If the researchers happen by chance to get a lot of individuals in the treatment group who respond very strongly to the placebo effect, then they may overestimate the effectiveness of the drug across the population as a whole.

Testing the Hypothesis that the Result Was Not Due to Chance

How can the believer persuade the skeptic the observed result is not due to the “luck of the draw?” One way is through hypothesis testing, which is a procedure for testing the reliability of an estimated value for the studied parameter (e.g., weight loss for obese individuals). There are several steps in the test. The first step is for the researchers to state two hypotheses—the “null hypothesis,” which states that the observed difference (20 pounds) reflects chance variation, and the “alternative hypothesis,” which states that the observed difference (20 pounds) is real. The believer would like to persuade the skeptic to abandon, or “reject,” the null hypothesis.⁴ The skeptic should reject the null hypothesis only when evidence convinces the skeptic it is false.

But how convincing must the evidence be before the researchers conclude there really is an effect? This leads to the next step in the test, which is for the researchers to determine when it is reasonable to reject the null hypothesis that the observed 20 pound difference reflects chance variation. There are two important points about this decision. The first is that this decision is arbitrary—there is no scientific standard to determine this criterion. The second important point about the decision is that whatever criterion is used, it should be set prior to conducting the study.

Researchers have developed and employed various criteria for rejecting the null hypothesis. For example, the researchers should likely reject the null hypothesis (i.e., accept that the 20 pound difference is “real”) when the difference observed in the sample would be very unlikely (say, with a probability of 5 percent) to be the result of chance variation. The logic of setting a fairly high threshold for proof is that there are then only two possible explanations for the observed data:

1. A rare event occurred.
2. The null hypothesis is false.

Statistical decisions are based on the premise that “rare events never happen to me,” but certainly they do sometimes, and it is important to be aware of this possibility. Regardless of the chosen threshold, there is always some possibility that the 20 pound difference is not the result of the drug but is due to chance variation. In this case, the null hypothesis will be rejected even though it should not. This risk is called the “Type I Error Rate.” Type I errors are also known as “false pos-

³ Of course, it is important that researchers correctly control for these other factors. Failure to do so (such as “misspecification” in a regression model) may result in estimates that are unreliable, even when researchers study the entire relevant population.

⁴ The names are often somewhat confusing, since the “alternative” hypothesis may be the main one, that is, the one the researcher wants to prove. The “null” hypothesis is what most people would consider an alternative explanation.
itives." Assuming a 95 percent threshold for rejecting the null hypothesis, this error rate will be 5 percent, or 1 out of 20.

Once researchers have chosen the decision criterion for rejecting chance (such as 5 percent) they next obtain the data from the sample (or samples). The researchers then use these data to compute the "test statistic," which allows the data from the sample (or samples) to be compared with what would be expected under the null hypothesis. One common test statistic used in antitrust cases is the "t-statistic," but there are others. I will spare the reader the technical details, but in short, the t-statistic is computed in part by comparing the difference between the observed result (a difference of 20 pounds between the two samples) and what is expected under the null hypothesis (a difference of zero pounds between the two samples).

The researchers then compute the "p-value," or the observed significance level, which is the probability of finding the observed result under the null hypothesis (zero difference) and then they compare this p-value to the chosen threshold significance level for rejection set in the second step of the test. The final step in the test process is for the researchers to draw a conclusion.

I return to the clinical drug trial example to show how this works in practice. The null hypothesis is that average weight loss from the drug is zero and that any difference between the control and treatment groups is solely due to chance. The alternative hypothesis is that the average weight loss from the drug is not zero. 5 Suppose the researchers choose the threshold for rejection, or significance level, to be 5 percent (that is, they are willing to accept a false positive risk of 5 percent). They obtain the data from the clinical trial and compute the "p-value," which is the probability of obtaining the observed result if the null hypothesis (zero difference) were true. They reject the null hypothesis if the p-value is less than the 5 percent significance level, and conclude the observed result is "statistically significant." For example, if the probability of obtaining the observed result (the difference of 20 pounds) is 2 percent when the null hypothesis is true, then the researchers "reject the null" and say the observed difference is "statistically significant" at the 5 percent level. On the other hand, if the probability of obtaining the observed result is 20 percent when the null hypothesis is true, then they "fail to reject the null" and say the observed difference is "not statistically significant" at the 5 percent level.

So, to be clear, statistical significance refers to determining whether the observed result is due to chance (the "luck of the draw") or other uncontrolled influences.

Setting the Level for Statistical Significance
As noted above, the significance levels are to be chosen in advance. 6 The choice of level, however, is arbitrary. Generally accepted levels are 1%, 5%, and 10%. Many factors affect the choice of significance level. One common factor that affects the choice is the size of the sample. Some researchers advocate using a more stringent threshold (say 1% instead of 5%) when the sample size is large, but again, the choice is arbitrary. 7

---

5 This is an example of a "two-sided" test, which researchers use when the estimate may vary from the comparison value (here, zero) in two directions (greater than or lesser than). Researchers use a "one-sided" test when the estimate may vary in one direction (greater than or lesser than, but not both).

6 This is to prevent researchers from setting the significance level to fit the results.

7 There is no such thing as a "statistically significant sample size." Significance applies to hypotheses about statistics computed from the sample, but not to the sample itself, or the size of the sample. Tests of statistical significance about a single parameter can be conducted on a sample of two or more observations. There is not, however, an easy answer to the question of what should be the sample size.
The reader might ask, “Why not set the Type I error rate as small as possible, say, at very close to 0 percent?” One pragmatic answer is that it may be impractical to obtain data precise enough to obtain statistically significant results given such a strict threshold. That is, an imperfect answer may be better than no answer at all. More importantly, though, there are two types of errors, Type I and Type II, and there are trade-offs between them. As noted above, a Type I error rejects the null hypothesis when it is true (false positive). Using the analogy of a criminal case, a Type I error occurs when an innocent person is convicted.

The other type of error, Type II, fails to reject the null hypothesis when the alternative hypothesis is true, also known as a “false negative.” Using the analogy of a criminal case, a Type II error occurs when a guilty person is acquitted. What is the trade-off between these two? If the significance level (the probability of a Type I error) is set at 0 percent, then the probability of a Type II error is 100 percent. Again, using the analogy of a criminal trial, if we do not want to have any chance of convicting an innocent person whatsoever, then we have to acquit everyone, which means all guilty people will be acquitted as well. The practical strategy is to balance the risks of the two types of the errors, taking into account the costs of each.

A hypothesis testing procedure identical to that of the clinical drug trial example is often employed in antitrust cases. Consider the familiar example of an alleged cartel. A typical question in such a case is whether customers were injured as a result of the alleged cartel. In such cases the researcher, usually an economist, wants to estimate a parameter, such as whether the average price paid by the population of purchasers increased as a result of the alleged cartel’s attempt to increase prices.

Often the economist will compare a treatment group (say, customers who purchased during the period the alleged cartel was in operation) with a control group (say, customers who purchased before or after the period the alleged cartel was in operation) in terms of outcomes (say, prices paid). Unlike the researchers for the pharmaceutical company, however, economists involved in antitrust cases do not have the luxury of running controlled experiments. Rather, their data typically come from the real world. They try, however, to make these data behave as if they come from a controlled experiment using techniques such as econometric methods to control for other observable factors that may affect outcomes (say, supply and demand) in the way that randomization or other methods do in the clinical drug trial for observable factors (say, for diet, activity level, age, gender, and genetic makeup). In both cases, though, the data are often from samples drawn from a population. In the clinical drug trial example, the treatment and control groups were comprised of samples of obese individuals in the population. This is often the case in alleged cartel cases, and other antitrust cases as well. Suppose a product was first sold in 1990, and a cartel of sellers was alleged to have been in operation for ten years from 2007 through 2016, after which the product was no longer sold. The customers who purchased the product at any time from 1990 through 2016 comprise the population. The treatment group is comprised of customers who purchased during the alleged cartel period. Suppose data are available for customers who purchased before the alleged cartel period only for the ten years from 1997 through 2006. If the customers who purchased during this ear-

---

8 In some cases, the treatment group will contain some or all of the customers in the control group (for example, if a customer purchased during the alleged cartel period as well as before or after that period).

9 Again, of course, it is important that economists correctly control for these other factors. Failure to do so (such as “misspecification” in a regression model) may result in estimates that are unreliable, even when researchers study the entire relevant population.
lier period are the control group, then it is clear their purchases are only a sample of all purchases from the entire period before the alleged cartel was in operation because the data exclude purchases in the population between 1990 and 1996.10

In the same way the researchers for the pharmaceutical company calculated and compared the differences in weight changes between the two groups, the economist calculates and compares the differences in prices paid between the two groups. The null hypothesis is that the difference between the average price paid during the alleged cartel period and outside the cartel period is zero. The alternative hypothesis in the case of an alleged cartel of sellers is that the average price is greater during the alleged cartel period than outside the alleged cartel period.

Suppose an economist finds that the difference between the average price paid by customers in the control group and the average price paid by customers in the treatment group was $20 per unit (after accounting for other supply and demand factors thought to affect price). For purposes of the example, I will call the $20 per unit difference in average price the “estimated overcharge.”

Suppose the economist chooses (in advance) the threshold for rejection, or significance level, at 5 percent (that is, the economist is willing to accept a false positive risk of 5 percent). The economist computes the p-value, and rejects the null hypothesis if the p-value is less than the significance level, concluding that the observed result is “statistically significant,” that is, unlikely due to chance.

For example, if the probability of obtaining the observed result (the estimated overcharge of $20 per unit) is 2 percent when the null hypothesis is true, then the economist “rejects the null” and says the observed difference is “statistically significant” at the 5 percent level. On the other hand, if the probability of obtaining the observed result is 20 percent when the null hypothesis is true, then he “fails to reject the null” and says the observed difference is “not statistically significant” at the 5 percent level.11

What Statistical Significance Is Not

As noted above, statistical significance refers to the likelihood of whether the observed result is due to chance (the “luck of the draw”). Unfortunately, the concept of statistical significance is widely misinterpreted and misused. While there are many ways in which these misinterpretations and misuses occur, I focus here on two.

The first is that some researchers erroneously conclude statistical significance indicates whether the null hypothesis is true. For example, some consider the p-value to be the probability of the null hypothesis being false (or “wrong” or “incorrect”). To illustrate, consider the clinical drug trial example. The believer says the drug works because participants who were given the drug lost, on average, 20 more pounds than those who were not given the drug. The skeptic says the drug does not work, because the observed difference is just due to chance. Suppose the skeptic seeks to persuade the believer by conducting a hypothesis test. Suppose the researchers choose the threshold for rejection, or significance level, to be 5 percent (that is, they are willing

10 Similarly, suppose instead the alleged cartel was in operation from 1990 through 2006, but was inactive from 2007 through 2016. The control group would be comprised of customers who purchased during the period from 2007 through 2016. Suppose again data are available only for customers who purchased before the alleged cartel period for the ten years from 1997 through 2006. The treatment group would be comprised of customers who purchased during this period. Again, it is clear their purchases are only a sample of all purchases because the data exclude purchases in the population between 1990 and 1996.

11 This is an example of a “one-sided test.” The economist is generally only interested in whether the average price paid by customers in the treatment group was greater, not less, than the average price paid by customers in the control group.
to accept a false positive risk of 5 percent). They obtain the data from the clinical trial and compute the “p-value,” which is the probability of obtaining the observed result if the null hypothesis (zero difference) were true. They reject the null hypothesis if the p-value is less than the 5 percent significance level, and conclude the observed result is “statistically significant.” Suppose the p-value is 2 percent. The researchers conclude the difference in the average weight loss between the two groups (20 pounds) is statistically significant at the 5 percent level, and, as a result, the believer claims victory.

That the result is statistically significant at the 5 percent level, however, does not mean the probability of the null hypothesis (zero difference) being false is 5 percent or less (or, equivalently, the probability of the null hypothesis being true is 95 percent or greater). The sample result is an attempt to derive a best estimate of what the result would be in the more general population, and the p-value simply works to reassure one that this result was not a false positive. However, it cannot tell one how good one’s “estimate” is or tell one with certainty whether the null hypothesis (zero difference) is true or false. When researchers conduct a hypothesis test they assume the null hypothesis is true, and then either reject it or fail to reject it.12 There is no “probability the null hypothesis is true” or “probability the null hypothesis is false.”

How then, the reader might ask, can we know if the null hypothesis (zero difference) is true or not? In theory, researchers would study the entire population.13 As noted above, however, studying the whole population is usually impractical. So, an unpleasant reality is that we may never know the truth.14

The second way in which statistical significance is misinterpreted and misused is that some researchers erroneously conclude statistical significance provides information about the practical importance of the observed effect. Consider the example above. Is the difference in the average weight loss between the two groups (20 pounds) of practical significance (importance)? The p-value is of no use in answering this question. The researchers would need to consider other criteria to answer this question. For example, the weight loss might be considered of practical significance if it is greater than some threshold level, such as 5 percent of average body weight measured at the beginning of the trial.

In economics, economists refer to this practical significance as economic significance. Here too, the p-value does not provide any evidence about the economic significance of an observed effect. Further, there is not even a formal definition of economic significance. For example, the reader may be familiar with the concept of a “small but significant and non-transitory increase in price,” or “SSNIP,” used in the hypothetical monopolist test set forth in the Horizontal Merger Guidelines published by the Department of Justice and the Federal Trade Commission. The agencies make it clear that, while they often use a SSNIP of 5 percent, what constitutes a “small but significant” increase in price depends upon a number of factors, and the agencies may use a price increase that is smaller or larger than 5 percent.15

Practical significance (including economic significance) and statistical significance can each exist without the other. Effects that are large in importance or magnitude may often not be statis-

---

12 More formally, the p-value is a “conditional probability,” that is, the probability of obtaining the observed result if the null hypothesis (zero difference) were true.

13 Again, measurement error can also be a problem even when researchers study the entire relevant population rather just a sample drawn from the population.

14 Another unpleasant reality is that we may never know whether we made an error.

tically significant, especially when sample sizes are small. Effects that are small in importance or magnitude may often be statistically significant, especially when sample sizes are large. Advances in technology that have allowed for the capture, storage, and analysis of more data have led to larger sample sizes in both scientific research and in legal cases. As a result, estimated effects, however small, are often statistically significant when the number of observations in a sample is in the thousands or millions.

Is Statistical Significance Still Significant?

For decades some researchers, including some economists, have questioned whether statistical significance is relied upon too heavily, or whether it should be relied upon at all. In recent years, their concerns have received more attention. One often expressed concern is that statistical significance tells one only about the evidence and the experiment and nothing about the real world. It is the magnitude of the effect, or the practical significance, they contend, that tells one something about the real world. In other words, they argue that statistical significance is nice, but not necessary.

To illustrate, consider the clinical drug trial example. The believer says the drug works because participants who were given the drug lost, on average, 20 more pounds than those who were not given the drug. The skeptic says the drug does not work, because the observed difference is just due to chance. Suppose the skeptic seeks to persuade the believer by conducting a hypothesis test. Suppose further the skeptic finds the difference in the average weight loss between the two groups is not statistically significant at the chosen level and claims victory.

The believer says no, it is the size of the effect that matters. The believer says it is clear the drug works because not only did the group that took the drug lose, on average, three times as much weight as the group that did not (30 pounds versus 10 pounds), but a weight loss of 30 pounds is meaningful for an obese individual. For example, if all the participants in the treatment group weighed 300 pounds at the start of the trial, then they lost an average of 10 percent of their body weight by the end of the trial. The believer says statistical significance would be nice, but it is not necessary to conclude that the drug works.

Is statistical significance necessary? The answer to this question may depend on the answer to another question: “What are the costs of being wrong (false positives)?” For example, what are the costs of giving obese individuals an ineffective weight loss drug that was (seemingly) effective in the trial? The reader might imagine the costs may well be significant, not only in terms of the money cost of the drug but the time lost not pursuing another course of action that would be effective for weight loss. For example, suppose there is an older drug proven to be effective for weight loss, but not as (seemingly) effective as the new drug was in the trial. Suppose the older drug cannot be taken in combination with the new drug, so that obese individuals would need to choose between taking one drug or the other. Those that choose the new drug will have lost the opportunity to take the older proven drug. They may learn later that the new drug was ineffective and then switch to the older drug, but in the meantime will have incurred the costs of carrying extra weight.

---

16 See, for example, the discussion in Deirdre N. McCloskey & Stephen T. Ziliak, The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives (2008).

17 Other costs may include any negative side effects of the new drug.
Consider the example of the alleged cartel. Suppose an economist finds that the difference between the average price paid by customers in the control group and the average price paid by customers in the treatment group was $20 per unit (after accounting for other supply and demand factors thought to affect price). Again, the $20 per unit difference in average price can be referred to as the “estimated overcharge.” The purchasers say they were injured because the alleged cartel was effective in increasing prices. The sellers say the purchasers were not injured because the estimated overcharge of $20 per unit is just due to chance. Suppose the sellers seek to persuade the purchasers by conducting a hypothesis test. Suppose further the sellers find the estimated overcharge is not statistically significant at the chosen level and claim victory. The purchasers say no, all that matters is that the estimated overcharge is positive (that is, greater than zero). They say statistical significance would be nice, but it is not necessary to conclude that the alleged cartel was effective in increasing prices.

Why might the estimated overcharge of $20 per unit not be statistically significant if the alleged cartel was effective in increasing prices? One reason may be the sample sizes are too small to provide statistically significant results. This might be the case, for example, when purchasers buy the good or service in question infrequently, and where the length of the sample period for the control or treatment group of purchasers is not long enough to provide a sufficiently large sample. A recent exchange between Laila Haider and Muneeza Alam on one side, and Kenneth Flamm and Michael Naaman on the other, over the issue of testing using “sub-regressions” at the class certification stage provides a glimpse of how the debate over statistical significance is being played out in actual antitrust cases.¹⁸ For the most part, however, the discussion has been over how much weight to give a finding (or lack thereof) of statistical significance rather than whether it is admissible at all.

In the antitrust context, what are the costs of say, erroneously awarding damages (in the sense of a Type I error) to purchasers in the form of estimated overcharges in an alleged cartel case when in fact they have suffered no injury (that is, a false positive)?¹⁹ Again, the reader might imagine the costs of such an error may well be significant. For example, the risk of an erroneous award may deter firms from undertaking procompetitive forms of cooperation that will increase total economic welfare. The cost to society will be the difference in welfare between the situation with the cooperation and the situation without the cooperation. Depending on the next best alternatives available to the firms, the cost could be significant.

Recall, though, there is also the Type II error (that is, a false negative). In this case, this would be to erroneously fail to award damages to purchasers in the form of estimated overcharges in an alleged cartel case when in fact they have suffered injury. Again, the reader might imagine the costs of such an error may well be significant. For example, the risk of an erroneous failure to award damages may motivate firms to undertake anticompetitive forms of cooperation that will decrease total economic welfare. The cost to society will be the difference in welfare between the situation without the cooperation and the situation with the cooperation. Depending on the next best alternatives available to the firms, the cost could be significant. Again, the practical strategy is to balance the risks of the two types of errors, taking into account the costs of each.


¹⁹ The absence of injury may not necessarily be due to the absence of a violation. For example, the alleged cartel members may have violated the law by agreeing to fix prices, but the agreement may not have resulted in any injury to purchasers in the form of overcharges due to “cheating” by one or more of the members, or competition from non-cartel members.
Conclusion

The concept of statistical significance is now accepted in a wide variety of legal cases, in particular, in antitrust cases in which econometric and statistical studies have increasingly played a larger and more central role. There are, however, growing concerns that statistical significance has been relied upon too heavily in scientific research.

Consider the statement issued by the American Statistical Association (ASA) in 2016 in response to these concerns about reliance on statistical significance. The ASA articulated six principles:

1. P-values can indicate how incompatible the data are with a specified statistical model.
2. P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
3. Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.
4. Proper inference requires full reporting and transparency.
5. A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.
6. By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.

The ASA concluded:

Good statistical practice, as an essential component of good scientific practice, emphasizes principles of good study design and conduct, a variety of numerical and graphical summaries of data, understanding of the phenomenon under study, interpretation of results in context, complete reporting and proper logical and quantitative understanding of what data summaries mean. No single index should substitute for scientific reasoning.

It will be interesting to see whether these concerns about statistical significance will continue to find their way into antitrust (and other) cases, if, and how, economists, lawyers, judges, and juries address them, and whether statistical significance remains significant.

---


21 Id.

22 Id.
How to Economize Consumer Protection

Michael R. Baye and Joshua D. Wright

Courts, private parties, and federal agencies extensively rely upon economists to help evaluate the merits of, and potential harm from, alleged violations of the Clayton Act, the Sherman Act and, more generally, business practices that are alleged to run afoul of antitrust law. At their core, the U.S. antitrust laws are about protecting consumers from harm stemming from abuses of competition and the competitive process. For instance, at the Federal Trade Commission—which has dual missions of consumer protection and competition—the deployment of staff economists and the use of economic analysis is standard practice on the competition side of the house. Bureau of Competition staff lawyers and the economists in the Bureau of Economics commonly work together to use economic analysis to identify, analyze, and prosecute cases. These analyses often involve statistical and econometric techniques that facilitate decision-making based on scientific evidence of likely harm to competition or consumers.

In contrast, the typical approach to a consumer protection matter relies upon a combination of surveys and subjective opinions to establish the facts relevant to a consumer protection dispute—for example, the existence and magnitude of consumer harm arising from an alleged inadequate disclosure. This is true for both deception and unfairness cases. The FTC’s deception cases address instances where a material statement, practice, or omission may mislead reasonable consumers. Unfairness cases, meanwhile, target other instances where a defendant’s behavior may cause substantial consumer harm that was not reasonably avoidable by consumers and not outweighed by attendant benefits. Both sets of cases lend themselves to valuable input from economists—unfairness cases, in fact, require a cost-benefit analysis.

Even in deception or unfairness cases where the FTC ties payments for consumer restitution or disgorgement to ill-gotten gains—i.e., to actual consumer harm\(^1\)—evidence is often limited to subjective survey evidence rather than objective economic evidence that the practice harmed consumers in the actual marketplace.\(^2\) While plaintiffs, defendants, and parties under investigation utilize economists with increasing frequency, it remains relatively uncommon to use econom-

---

\(^1\) See, e.g., FTC v. Gem Merch. Corp., 87 F.3d 466, 470 (11th Cir. 1996) (“We conclude that section 13(b) permits a district court to order a defendant to disgorge illegally obtained funds.”); FTC v. BurnLounge, Inc., 584 F. App’x 315 (9th Cir. 2014) (affirming in part and remanding in part the disgorgement award for clarification of the calculation of consumer harm attributable to certain packages and fees); FTC v. Wells, 385 F. App’x 712, 713 (9th Cir. 2010) (“The FTC may seek full restitution from any individual who caused consumer harm through unfair practices.”) (citing FTC v. Pantron I Corp., 33 F.3d 1088, 1101–04 (9th Cir.1994)).

\(^2\) We note that, in practice, the FTC’s monetary remedies are limited to the disgorgement of “ill-gotten gains” or the amount of restitution required to restore consumers to the welfare they would have enjoyed but-for the deceptive or fraudulent act. See supra note 1. When a product is purely a sham (e.g., snake oil), it hardly takes an economist to determine that every penny the seller collected was “ill-gotten” and that full restitution requires returning all of monies received back to consumers. As discussed below, many consumer protection cases are more complex and require more sophisticated economic analysis to determine (a) how many consumers were deceived by a particular business practice, and (b) how much of the money collected by the seller must be returned to consumers to make them whole.
In our experience, many seemingly benign consumer protection settlements induce asymmetries in the marketplace that put the settling firm at a competitive disadvantage. Regardless of whether one's goal is to protect consumers or defend one's client, doing so requires accounting for these effects. More broadly, economics provides tools which, when properly utilized, can help improve the allocation of scarce resources—at agencies and beyond—to better serve and protect consumers and competition.

In this article, we identify some reasons why economics has often been an afterthought in consumer protection matters, why this is changing, and some of the adverse effects that arise from the gap between the use of economists in antitrust and consumer protection. We then provide a handful of examples illustrating how economic analysis can be used to inform decisions and help ensure that well-intentioned efforts by plaintiffs to protect consumers do precisely that.

---

3 LabMD, Inc. v. FTC, 678 F. App’x 816, 820–21 (11th Cir. 2016) (“The FTC’s ruling did not point to any tangible harm to any consumer, because there is no evidence that any consumer suffered a harm such as identify theft or physical harm. Instead, the FTC found actual harm here due to the sole fact of the 1718 file’s [which contained 1,718 pages of sensitive personal information for roughly 9,300 patients] unauthorized disclosure. The FTC also found that consumers suffered a ‘privacy harm’ that may have affected their reputations or emotions . . . . Alternatively, the FTC found that the unauthorized exposure of the 1718 file was likely to cause substantial injury.”).

4 FTC v. D-Link Sys., Inc., No. 3:17-cv-00039-JD, 2017 WL 4150873, at *5 (N.D. Cal. Sept. 19, 2017) (“The FTC does not identify a single incident where a consumer’s financial, medical or other sensitive personal information has been accessed, exposed or misused in any way, or whose IP camera has been compromised by unauthorized parties, or who has suffered any harm or even simple annoyance and inconvenience from the alleged security flaws[,]”.

---
The Underutilization of Economics in Consumer Protection Cases: Causes and Effects

The genesis of economics’ relative absence in consumer protection enforcement is understandable. In earlier years, consumer protection enforcement focused heavily upon clear fraud cases. Understanding that fraudulent conduct (involving a sham product, say) offers zero benefit to consumers hardly requires an advanced economic degree. For example, prohibiting a seller from making money purely as the result of obviously false claims about its brand of “snake oil” will reduce the number of unwanted and uneconomic purchases, while simultaneously allowing producers of legitimate products to make credible claims and succeed. In these clear-cut cases, one hardly requires sophisticated economic analysis to determine “ill-gotten gains” or the amount of restitution to make consumers whole—the amount of ill-gotten gains from sales of an obviously sham product (i.e., one that provides no value to the consumer) is the entirety of the amount collected by the seller.5

In traditional unfairness cases, the FTC initially challenged only behavior that was clearly harmful to consumers and lacked any justifiable benefits. As with fraud, the challenged conduct in such cases was more often than not clearly harmful, such as documented fraudulent charges.6 Rigorous cost-benefit analysis was generally not necessary to conclude the conduct at issue made consumers worse off.7

The FTC’s unfairness enforcement, however, is no longer confined to such clear-cut cases. Rather, enforcement based upon unfairness authority by the FTC—and others, like state attorneys general—has expanded to include cases where an assessment of tradeoffs is critical.8 For example, enforcement challenges to disclosure and product design decisions that offer substantial benefits are increasingly common. Approaching these and similar business practices as a traditional fraud case poses a serious risk of creating unsound policy and reducing consumer welfare.9

---

5 It is important here to note the difference between a finding by the FTC of deception and its ability to obtain monetary redress through disgorgement or restitution. The FTC can issue a cease and desist order to prevent firms from lying even if there is no real harm—and this may be a good thing—but obtaining disgorgement requires more. Namely, disgorgement requires some demonstration of ill-gotten gains (i.e., consumer harm) arising from the defendant’s conduct, and “full restitution” requires the FTC collect no more than this amount. See supra note 1; Dissenting Statement of Commissioner Joshua D. Wright, Cardinal Health, Inc., FTC File No. 101-0006 (Apr. 17, 2015), https://www.ftc.gov/public-statements/2015/04/dissenting-statement-commissioner-joshua-d-wright-cardinal-health-inc.

6 See, e.g., Complaint for Permanent Injunction & Other Equitable Relief at 6, FTC v. Jesta Digital, LLC, No. 1:13-cv-01272 (D.D.C. Aug. 20, 2013) (alleging that “Jesta charged consumers who did not click on the subscribe button and charged consumers for products they did not order.”).

7 In addition, there was traditionally some amount of internal hostility toward integrating economists into consumer protection cases more closely. See Paul A. Paulten, A History of the FTC’s Bureau of Economics 109 n.393 (AAI Working Paper No. 1503, 2015), https://www.antitrustinstitute.org/sites/default/files/FTC%20Bureau%20of%20Economics%20History_0.pdf (“The ‘economist as subordinate’ view was explicit and evident in the latter 1990s on the consumer protection side of the agency, when certain attorney managers would not allow economist participation in meetings with the respondents or with outside third-parties.”).

8 See, e.g., Complaint, Apple, Inc., FTC No. 112-3108 (Jan. 15, 2014), https://www.ftc.gov/sites/default/files/documents/cases/140115Apple cmpt.pdf. As developed below, ignoring the tradeoffs in cases like Apple iTunes can lead to FTC actions that condemn conduct whose harms are greatly outweighed by their benefits. See infra Section II.B; Dissenting Statement of Commissioner Joshua D. Wright, Apple Inc., FTC File No. 112-3108 (Jan. 15, 2014), https://www.ftc.gov/sites/default/files/documents/cases/140115Appletestamentwright_0.pdf [hereinafter Wright Apple Dissent].

9 See Elise M. Nelson & Joshua D. Wright, Judicial Cost-Benefit Analysis Meets Economics: Evidence from State Unfair and Deceptive Trade Practices Laws, 81 ANTITRUST L.J. 997 (2017) (measuring how economists versus courts analyzed unfairness claims under federal and state law and explaining the “primary reason to require cost-benefit analysis as a prerequisite to finding business conduct unlawful is to ensure consumer protection enforcement does not chill conduct that makes consumers better off even if there is some small harm experienced by consumers”).
As challenges to product design decisions and disclosure practices of digital platforms become increasingly common, the onus should shift to the FTC to properly assess the effects of modern business conduct. ¹⁰

In many cases involving the disgorgement of “ill-gotten gains” resulting from inadequate disclosures or other deceptive practices, plaintiffs completely sidestep the issue of whether any real consumers’ decisions were actually impacted by the alleged conduct, and do not rigorously link the alleged wrongdoing to consumer harm, restitution, or disgorgement. The FTC’s own guidance does not do much to resolve important questions like (1) what constitutes a “significant minority” of consumers sufficient to support a deception claim, (2) how the agency will analyze the materiality of allegedly deceptive statements, (3) when that materiality will be presumed,¹¹ and (4) what constitutes a “reasonable estimate” of consumer harm when the FTC seeks disgorgement of ill-gotten gains.¹² Thus, the FTC perspective has often been that the focus should not be upon providing economic evidence proving that an advertiser made a deceptive claim that actually harmed consumers, but rather upon the defendant’s attempt to offer proof that consumers were not impacted by the claims.¹³ To make a comparison to antitrust law, deception is often analyzed in a manner similar to per se illegal naked restraints of trade, whereas the unfairness standard is more analogous to rule of reason treatment.¹⁴ It is not surprising, especially with respect to the application of its deception authority, that the FTC has not developed an analytical approach to proving the consumer harm arising from deception directly.¹⁵ It has not often needed to do so in order to prevail in matters involving, for instance, sham products.


¹³ See, e.g., Kraft Inc. v. FTC, 970 F.2d 311, 319, 324 (7th Cir. 1992) (noting that the FTC deemed that “Kraft’s consumer surveys were insufficient to rebut [the FTC’s] inference and in particular criticize[d] Kraft’s survey methodology because it offered limited response options” to respondents); see also James C. Cooper & Joshua D. Wright, The Missing Role of Economics in FTC Privacy Policy, in CAMBRIDGE HANDBOOK OF CONSUMER PRIVACY (Jules Polonetsky, Evan Selinger & Omer Tene eds.) (forthcoming 2018).

¹⁴ Cooper & Wright, supra note 13; J. Howard Beales, III, Director, Bureau of Consumer Protection, Fed. Trade Comm’n, The FTC’s Use of Unfairness Authority: Its Rise, Fall, and Resurrection, at 9 (May 30, 2003), http://www.ftc.gov/public-statements/2003/05/ftcs-use-unfairness-authority-its-rise-fall-and-resurrection (“[T]he primary difference between full-blown unfairness analysis and deception analysis is that deception does not ask about offsetting benefits. Instead, it presumes that false or misleading statements either have no benefits, or that the injury they cause consumers can be avoided by the company at very low cost. In other words, deception analysis essentially creates a shortcut, assuming that, when a material falsehood exists, the practice would not pass the full benefit/cost analysis of unfairness, because there are rarely, if ever, countervailing benefits to deception.”).

¹⁵ See J. Howard Beales, III & Timothy J. Muris, FTC Consumer Protection at 100: 1970s Redux or Protecting Markets to Protect Consumers?, 83 Geo. Wash. L. Rev. 2157, 2204 (2015) (discussing the FTC’s failure to follow the standard used in its Policy Statement in deception cases); Cooper & Wright, supra note 13.
Underutilization of economic analysis can also impact consent agreements and their negotiation. Untethering consumer protection from economic analysis increases the likelihood that case decisions—whether to abandon, settle, or litigate—are based in part on speculative or subjective opinions where scientific or economic evidence could result in a more informed view and provide unique insights compared to other forms of fact-finding. Providing rigorous economic evidence also increases the likelihood that both the plaintiff and the defendant confront the relevant issues. Engagement with economic evidence increases the likelihood that consent agreements are closely tied to the underlying issues and hence proportionate to the actual effects on consumers. It also allows the Commission to focus on obtaining remedies that are most likely to benefit consumers and least likely to have unintended, adverse consequences on competition.

**Expanding Use of Economic Analysis in Consumer Protection Cases**

In this section, we provide a detailed explanation of how economic tools and insights can help the FTC to identify and prosecute deception, or for defendants to successfully rebut such claims. It is important to note that, while it may be possible to establish that a practice is misleading or deceptive without economic analysis, a reasonable estimate of consumer harm is required in order to extract payments by the perpetrator (e.g., restitution or disgorgement of ill-gotten gains). We highlight several other ways in which economics can help the Commission and those affected by its decisions to make better-informed decisions, including by identifying and analyzing unfairness, ensuring that actions are in the public interest, quantifying the “risk” or “likelihood” of consumer harm, conducting relevant but-for analysis, and enhancing agency resource allocation.

**Identifying Deception.** According to the FTC’s Policy Statement on Deception, “Certain elements undergird all deception cases. . . . First, there must be a representation, omission or practice that is likely to mislead the consumer. . . . Second, [FTC Staff] examine the practice from the perspective of a consumer acting reasonably in the circumstances. . . . Third, the representation, omission, or practice must be a ‘material’ one.” The typical starting point for assessing consumer protection claims—particularly for allegations of deception or inadequate disclosures—is to use survey techniques in an attempt to determine whether a “significant minority” of “reasonable consumers” might fail to understand material terms, or otherwise that the “net impression” of the ad is to deceive consumers.

While survey methods (if properly conducted) may provide context for the first element of deception cases, they can be severely limited in their application to the second and third elements. Surveys, by their very nature, are unable to measure the actual circumstances of consumers making choices in the face of deceptive practices. Survey results can only imperfectly replicate the true economic incentives and informational context of the consumers whose choices are at issue. The behavior of survey respondents may significantly differ from the choices made in the actual marketplace by “reasonable consumers.”

Furthermore, these same limitations of surveys may hamper the ability of a researcher to infer whether an act or practice is “material” in the context that actual consumers make decisions—that is, whether it “is likely to affect the consumer’s conduct or decision with regard to a product or service.” The mere fact that information was inadequately disclosed, for instance, does not imply

---


17 Id. Note that consumers do not make purchase decisions in a “vacuum,” and survey data may fail to capture information that the typical consumer obtains prior to purchasing a product or service.
that all or even most consumers were harmed; indeed, they might have made the same choices or decisions in a “but-for world” with better information.\^18 Likewise, although one may attempt to use a survey to infer the actions of consumers (or their understanding of a disclosure), the survey setting is a very poor substitute for analysis of data on the actual choices that consumers made when acting upon the information they received from an alleged deceptive act or practice in their regular daily lives. Because surveys measure choices (or the understanding) of survey respondents rather than actual consumers (who, unlike survey respondents, condition purchase decisions on the actual cost they would have to pay for the good or service and ancillary information), they may not accurately reflect behavior in actual markets.\^19

In contrast to the survey methods typically applied to investigations of potentially deceptive practices, economics offers useful insights into the question of whether the observed choices of the actual consumers who purchased the product associated with false or misleading information are likely to have been significantly impacted by deception. Additionally, economics permits one to accurately estimate the amount of restitution required to fully compensate consumers from harm stemming from the harmful act. These same economic tools (rather than surveys) are typically used to examine the merits and calculate damages in antitrust cases.

Where data on the choices of consumers are available, econometric methods may offer insight into all three elements of the FTC Staff’s investigation of deceptive practices. As will be discussed in more detail below, economists have developed numerous, rigorously studied empirical methods to understand the impact of changes to consumers’ information sets (e.g., comparing an actual world with allegedly inadequate disclosures and a “but-for world” with more detailed disclosures) on their decisions. Here we highlight three broad categories of analyses which, in our experience are often helpful in establishing (or rebutting) consumer harm in, for example, consumer protection matters involving allegations of deception, inadequate disclosures, and/or unsupported performance claims. We stress that we adopt a simple pedagogical approach to illustrate concepts; in practice, a variety of econometric and statistical tools are required in order to conduct reliable statistical tests for the concepts presented below.

As a general matter, there are many alternative approaches, and the specific approach used will depend on the industry, the nature of the alleged deceptive practice, and the quality of data available. Likewise, a variety of different types of data may be used to formally test whether alleged behavior impacted consumer decisions or resulted in harm or injury to consumers.

\^18 As one example, consider that a promotion may inadequately disclose the “full price,” but the “full price” might be so attractive that consumers would have made the exact same decision in a “but-for world” with adequate disclosure. Suppose a special promotion at a new Mercedes dealership clearly discloses that its first—and only its first—customer will be offered a Mercedes at a special price of $1,000, but fails to disclose a $500 processing fee. Because the $500 processing fee is unlikely to impact buyers’ decisions, the omission of the $500 fee is not economically material (e.g., the dealer does not enjoy any ill-gotten gains, and the consumer would have made exactly the same decision in a “but-for” world with adequate disclosure). In this example, none of the monies collected by the perpetrator were ill-gotten; the same amount would have been collected in a but-for world without deception, so consumers need not receive any restitution payments to be made “whole.” Note that, as a general matter, it is not efficient to use scarce enforcement resources to prevent deceptive practices that do not impact consumer behavior or adversely impact consumer welfare. Unlike fraudulent practices, where the nature of the deception may be criminal, the FTC’s enforcement of Section 5 violations extends to business practices that are not unambiguously deceptive. Although halting unambiguously deceptive practices (e.g., selling “snake oil”) may be desirable, disgorgement of monies for an alleged deceptive practice in which no one has been harmed is not likely to be in the public interest, given scarce agency resources. Indeed, doing so may have a chilling effect on legitimate business practices regarding the dissemination of valuable information.

The Before-After Approach. First, suppose that a firm with an established product sold through its website initiated a new pricing strategy on January 1, 2016. This new pricing strategy did not change the actual price of the product, but reduced the price displayed on its landing page and added a separate “handling fee” in small print that is alleged to have been inadequately disclosed. The allegations appear to be consistent with survey results showing that 20 percent of respondents failed to notice the handling fee. It is possible that the survey is flawed owing to the absence of relevant filter questions, the use of leading questions, or the absence of a control group. In any event, survey results alone are insufficient to directly address the issue of consumer harm in the actual market; economic analysis provides a more direct method of examining the issues.

Figure 1

Economic theory indicates that, if consumers failed to notice the handling fee, they would have erroneously believed the price was lower than it actually was. By the law of demand, this implies that, if the allegations are correct, the firm would have enjoyed an increase in sales, other things equal. Even if a survey appears to indicate that 20 percent of respondents indicated that they did not notice the handling fee, actual data on sales is more reliable for examining the potential impact of the firm’s alleged failure to disclose on the behavior of actual consumers. Indeed, given historical data on the firm’s sales, one may conduct tests to examine the allegations using data on actual consumer choices in the marketplace. For example, suppose the data patterns for the company’s sales are similar to those in Figure 1. In this case, there does not appear to be any material difference in consumer behavior before and after the alleged inadequate disclosure (e.g., before and after January 1, 2016). Expressed differently, these data are inconsistent with the hypothesis that the alleged inadequate disclosure was economically material to the decisions of actual consumers.

See id.

For exposition, this example assumes that there have not been any other changes to the market, apart from the disclosure, that could affect consumer behavior. As discussed below, it may be important to account for other exogenous factors when examining the impact of the change in disclosure on consumer behavior.
One can formally test this hypothesis using regression techniques to control for other factors that might influence consumer decisions, but for purposes of this article, Figure 1 illustrates how actual data on consumer behavior may be used, in principle, to refute or reject an allegation that 20 percent of consumers suffered economic harm because they failed to notice the handling fee. Unlike a survey, the results in Figure 1 are based on actual consumer behavior and not subject to concerns about leading questions, whether one is testing “consumer memory,” etc.

In contrast to Figure 1, it is possible that actual data might display a pattern like that in Figure 2. Here, notice the company's units sold increasing following January 1, 2016, and thus the economic data appear consistent with the survey results. In other words, the jump in units sold immediately following the date at which the company started using the disclosure at issue is consistent with the hypothesis that some consumers did not observe the fine print regarding the handling fees and purchased the product that they would not otherwise have purchased. However, also notice that, in the scenario in Figure 2, the size of the jump is on the order of 1.5 percent (units sold jump from approximately 10,000 to 10,150, a 1.5 percent increase). Depending upon the nature of the product, this might indicate that far less than 20 percent of customers were actually harmed by the inadequate disclosure.

Figure 2

We note that analyses based on data such as that in Figure 1 and Figure 2 hold other things constant—things like changes in consumer incomes, changes in the prices of substitutes, and other factors (such as the amount of TV advertising the firm conducts to induce consumers to visit its website in the first place). It turns out that this is not a real limitation, however, because econometric techniques (including regression analysis) may be used to control for other factors that might have influenced the company’s sales.

To illustrate, suppose the company's underlying online sales data is as displayed in Figure 2, but that the company also invested more in TV advertising, that this change (which is not alleged to be deceptive) induced more consumers to visit the company's website, and that the allegedly deceptive disclosure only appeared on the website. The associated increase in visitors to the firm's website is displayed in Figure 3.

In comparing Figure 2 and Figure 3, there are obviously confounding effects raising the possibility that the increase in online sales observed in Figure 2 did not stem from inadequate dis-
closure, but from an increase in web traffic (as a result of the benign increase in TV advertising).

One simple method of distinguishing between these two possibilities is to examine data on the “conversion rate”—that is, the ratio of the company’s sales to website visits. If the allegations are correct, the inadequate disclosure would have resulted in an increase in the conversion rate. That is, for any given number of visitors, the firm would sell more units because consumers would be unaware of a portion of the price. In contrast, if the conversion rate did not change following January 1, the data are inconsistent with the allegation that consumers were harmed as a result of inadequate disclosure.

Figure 4 shows that the conversion rate, which is the ratio of the numbers in Figures 2 and 3, was unaffected by the change in disclosure. The conversion rate in Figure 4 is around 50 percent, indicating that about half of those consumers visiting the company’s website end up purchasing its product. Thus, using the “before-after approach” based on the data in Figure 2 and Figure 3
leads to a conclusion that the allegedly deceptive practice did not significantly impact the purchase decisions of actual consumers.

In concluding our discussion of the before-after approach, we note that, if the ultimate pattern of data (once one uses regression analysis to control for other potential factors) appears as in Figure 2 rather than as in Figure 4, there are two potential reasons for the observed patterns. One possibility is that a small fraction (e.g., 1.5 percent) of consumers did not read the fine print, but the remaining fraction (e.g., 98.5 percent) did. Another possibility is that, even though 20 percent of consumers might not have understood the hidden fee, only 1.5 percent of consumers would have made a different decision in a “but-for world” in which they did understand the fee (e.g., the fee was adequately disclosed).

Thus, while economics cannot shed light on the normative question of what constitutes a “significant minority” of consumers, it can help quantify the number or percentage of consumers who were potentially harmed by the alleged act. Even in instances where an advertisement is deemed to have deceived a significant minority of consumers, one may use consumer data and econometric analysis to determine how many consumers were impacted by the deceptive act—in this case, 1.5 percent—and reliably determine the extent of consumer injury and/or restitution.

The Treatment-Control Approach. In consumer protection matters where historical data on consumer purchases are not available, or where the practice at issue has been ongoing and therefore there is no “before” period, the treatment-control approach may be used to examine the impact of an alleged act on consumers. To illustrate, suppose that a firm markets a brand of wine in City A using an “organic” label. It sells the exact same product at the exact same price in City B; however, as a result of differences in local regulations, the product does not carry the organic label in City B.

Suppose there is an allegation that the firm was unable to substantiate its organic label and, as a consequence, consumers in City A have been harmed. In this example, it is as if wine in City A has been “treated” with an organic label, and wine in City B (the “control” wine) has not. For purposes of this example, one might initially assume that consumers in these two cities are identical in terms of their overall consumption of alcohol, tastes for wine versus beer or spirits, and so on. In practice, this assumption is not necessary because regression analysis allows one to control for these sorts of heterogeneities.

Taking the non-substantiation claims at face value, one may use available data on sales in City A and City B to examine potential consumer harm. As an initial matter, one may observe that average sales of the wine at different stores in City A tend to exceed sales at stores in City B. In this case, the patterns of consumption would be consistent with the allegations; in City A (where the “organic label” treatment was applied), sales are higher because consumers in City A were duped into believing there was more value to the product than the company can actually substantiate.

While this assumes the only difference between City A and City B are the labels, one could control for other factors (including population size and demographic factors such as age or gender differences) that might lead to differences in wine consumption in City A and City B for purely benign reasons. Regression analysis permits one to control for these factors in much the same way that one can control for age and other characteristics in a study that examines the impact of smoking on mortality. Unlike experimental studies, the “control group” here is based on an observational study.22

22 Thus, when we refer to a control group, we refer to an “observational study” in which one controls for “confounding variables by statistical methods.” See David H. Kaye & David H. Freedman, Reference Guide on Statistics, in Reference Manual on Scientific Evidence 221 (3d ed. 2011).
In the case at hand, suppose that one conducts such an analysis to obtain estimates of wine consumption in City A and City B that hold all characteristics constant except for the labeling. This permits one to statistically test whether consumers in City A purchased more wine than they would have purchased in a but-for world in which the company did not utilize the organic label. For example, one might find that sales in City A are about 1 percent higher than in City B after one controls for other factors that influence consumption, but the difference is not statistically significant. In this case, the data would be inconsistent with the allegation that consumers were harmed as a result of the deceptive organic label.

The Difference-in-Difference Approach. The difference-in-difference approach combines the intuition underlying the before-after approach and the treatment-control approach. For example, suppose that a car dealer (Dealer A) changed the manner in which it disclosed its warranty on cars on January 1, 2016, and a nearby dealer of similar cars (Dealer B) did not modify its disclosure. It is alleged that Dealer A’s new warranty is deceptive because reasonable consumers would infer that it covers all maintenance, even though fine print indicates it does not. In effect, the allegation is that the warranty did not actually change on January 1, 2016, but the manner in which it was disclosed to consumers changed, deceptively. In this instance, we have both before (pre-January 1) and after (post-January 1) data on sales by the treatment company (Dealer A) and a control company (Dealer B).

Hypothetical data are depicted in Figure 5. Notice that both dealer’s sales are trending upward before January 1, and there is a difference in sales levels. After January 1, both dealer’s sales increase suddenly, but Dealer A’s increase appears larger than Dealer B’s.

A difference-in-difference approach allows one to disentangle the effect of a specific treatment on a group from other confounding effects occurring at the same time. For example, a new popular model of car may be launched on January 1, 2016. In this case, economic theory predicts Dealer A’s sales would have increased regardless of the allegedly deceptive change in warranty disclosure. A properly applied difference-in-difference approach may be able to identify the increase in sales specifically attributable to the warranty disclosure. In this stylized example, one
may observe that the difference in sales between Dealer A and Dealer B, shown in Figure 6, shrinks from 20 cars to 10 cars after January 1 once the analysis accounts for factors like new model introductions on January 1st. Thus, this evidence is consistent with Dealer A’s change in disclosure leading to 10 more car sales per month. In reality, a regression model of a difference-in-difference approach may apply additional controls to account for other potential explanations of the different pattern of sales between the dealers.

Figure 6

Although purchase data were central to all three forms of analysis discussed above, data and information on other factors that might impact consumer behavior may easily be accounted for using standard economic and statistical techniques. Additionally, other data—such as lost sales, returns or cancellations—may be relevant for assessing the quality of information available to consumers before purchase. Likewise, greater detail (more granular levels of observation) tends to provide for additional forms of variation that an econometrician can exploit to identify other benign factors that impact consumer decisions.

Identifying and Analyzing Unfairness. Under the Commission’s Policy Statement on Unfairness, an unfairness case requires that consumer injury be “substantial,” not “trivial or merely speculative,” and that it “not be outweighed by any offsetting benefits.” In other words, the Policy Statement itself requires a cost-benefit analysis of the allegedly unlawful conduct to ascertain whether the harms it causes are greater than the benefits it offers. This should ingrain an economic analysis within unfairness cases, and it provides a clear path through which economists can meaningfully contribute to case development. It is, in fact, difficult to overstate the value of economic analysis in hard cases like unfairness allegations involving privacy and data security. Both privacy and data security cases can involve economic tradeoffs more subtle than the typical deception case involving fraud or other conduct without redeeming economic virtue. Economists can be helpful in thinking through several important issues that arise in such unfairness cases.


24 See Cooper & Wright, supra note 13; Wright, supra note 10; Nelson & Wright, supra note 9.
Each of the techniques described above with respect to deception could also be deployed to enhance Commission analysis of unfairness cases. They could help more precisely to identify the costs and benefits at issue, thereby allowing the FTC more accurately to conduct the required cost-benefit analysis. For instance, consider that Application Platform A is being investigated under an unfairness theory for failing expressly to inform consumers that a new payment option it was introducing, which would charge a consumer only once a month (rather than at the time of each purchase), would cost $0.05 per month. Assume, further, this information was available in billing policy statements and in an updates section, but not explicitly called to the consumer’s attention. This option might benefit some consumers, by essentially extending them a temporary credit (particularly for those using debit cards to pay), but might harm others who valued this option at less than $0.05 per month. Application Platform B instituted a similar policy—but also expressly disclosed it—six months before the Commission began its investigation. Depending upon the data available, the FTC could use before-after, treatment-control, or difference-in-difference approaches to evaluate how many consumers opted to use this payment option when expressly informed of the cost versus how many opted to use it when the information was available only in the billing and updates pages, and also how purchasing trends may or may not have altered across platforms, to better understand the implications of the changes and the relative effects across each platform.25

Consider further the impact of privacy settlements on incentives to make disclosures in privacy policies. In Nomi, for example, the FTC relied upon a statement in Nomi’s privacy policy pledging to “[a]lways allow consumers to opt out of Nomi’s service on its website as well as at any retailer using Nomi’s technology.”26 Based upon this statement, the Commission concluded Nomi engaged in unlawful conduct because it failed to always provide an in-store opt-out option or to inform consumers that a specific retailer would be using the service—despite the fact that Nomi always offered an easy opt-out option on its webpage. The Commission and Nomi entered into a 20-year consent agreement to resolve the matter, which requires Nomi to “not misrepresent in any manner, expressly or by implication” the options by which consumers can control the collection, use, disclosure, or sharing of their information or regarding the notice Nomi would provide consumers regarding how it would use that data. It further subjects Nomi to various requirements regarding maintaining materials, making these materials available to the Commission, and other reporting requirements for several years.

Here, economic analysis could help in a variety of ways. One important question is whether consumers were actually harmed and, if so, to what extent. Economists could assess whether consumers’ behavior would have been changed but-for the allegedly unlawful conduct—that is, were they actually relying upon Nomi’s statement that they could opt out in stores or were they sufficiently able to opt out online.27

Importantly, economists can also play a valuable role in providing input on potential consent agreement provisions. For example, one concern with the consent agreement in Nomi was that, because the FTC imposed no ex ante legal obligation on the firm to adopt a privacy policy (let

25 See generally Wright Apple Dissent, supra note 8 (explaining why the FTC’s decision in the Apple iTunes matter failed to reflect a rigorous assessment of consumer harm in the unfairness setting).


alone to offer a simple, always-available online opt-out option), the decision would chill similarly situated firms from adopting one in the first instance—a result that would be likely to make consumers worse off. Such agreements might particularly discourage start-ups like Nomi from adopting any (and especially comprehensive) privacy policies for fear of incurring costly FTC investigations and burdensome, years long requirements for which they simply do not have resources.

Economists can also identify and analyze potentially countervailing benefits in unfairness cases and their magnitude relative to potential harms. This is a role economists traditionally play in competition cases and are well suited for. In the absence of thorough analysis of potential benefits of allegedly unfair conduct, the agency is susceptible to the error of allocating agency resources to prohibit conduct that is of net benefit to consumers.

Consider the FTC’s recent settlements in the various “in-app purchase” cases. The FTC alleged it to be unfair that one of Apple’s product features opened a 15-minute window after entry of a password during which users could make additional purchases without being prompted to re-enter the password. The FTC challenged Apple’s business practice as unfair because it charged the consumer without obtaining express informed consent during the window, equating the conduct to a failure to disclose.28 While the disclosure level certainly may have resulted in some harm to consumers, there is undoubtedly also consumer benefit that arises from the customer experience without additional disclosures and with the 15-minute window as designed. A careful cost-benefit analysis would be required not only to satisfy the Commission’s own unfairness standard, but also to ensure that the FTC’s action resulted in a net benefit to consumers. While a formal discussion of the economic and statistical tools available to conduct such cost-benefit analyses is beyond the scope of this article, we note that these techniques are used in “rule of reason” antitrust cases; indeed, economists routinely quantify the procompetitive benefits of potentially harmful practices, such as vertical restraints or information sharing agreements.

Ensuring that Actions Are in the Public Interest. Enforcement actions are not costless. Economic analysis can help target Commission enforcement to its best uses. The Commission has limited resources to devote across investigations, enforcement actions, and actual litigations. Economics can help the agency to enhance its enforcement efforts by identifying those areas where the Commission’s scarce resources can best protect consumers. For instance, economic analysis can help illuminate which matters under investigation involve the most consumer harm and for which of those matters agency enforcement would offer the most benefits.

Of course, comprehensive economic analysis will not be warranted in every case. But closely incorporating economic analysis into consumer protection enforcement more generally would elucidate when in-depth economic investigations are and are not useful. Economic analysis has helped us to learn where best to target resources on the antitrust side—for instance, merger retrospectives help us to better understand and predict outcomes and to identify where intervention has helped or hurt, leading the Commission to focus heavily upon, among other things, hospital mergers. Just as economic analysis has helped the Commission target its antitrust resources, it can assist in putting consumer protection resources to their best uses.

Conclusion

Consumer protection enforcement efforts are fundamentally important to a well-functioning economy, but can also have adverse, unintended consequences on competition that ultimately harm

---

the very people they are designed to protect. Deeper integration of economic thinking and economic analysis into consumer protection matters offers value to all stakeholders—agencies, private plaintiffs, state attorneys general, and defendants—from case development to litigation and settlement.
Book Review
Getting Economics Back in the Game

Jean Tirole
Economics for the Common Good
Princeton University Press 2017

Reviewed by Gregory K. Leonard

Jean Tirole, chairman of the Toulouse School of Economics and winner of the 2014 Nobel Prize, is one of the world’s foremost economists. In an age when most economists specialize narrowly, Tirole has, over the course of his career, made significant contributions to an impressively broad number of areas within economics.1 Now, in a new book entitled Economics for the Common Good, Tirole has taken on a formidable task: convincing non-economists that economics can—and should—play an important role in developing policies and institutions that further the “common good.” Antitrust practitioners familiar with Tirole’s work in industrial organization will not be surprised to find that competition policy and related subjects such as innovation and regulation are central to Tirole’s vision and are given extensive treatment in the book.

Tirole is troubled by widespread rejection of economic thinking, particularly among members of the populist movements that have arisen around the globe: “Nowadays, people with expert knowledge are often dismissed. Populist politicians and media in particular show complete contempt for elementary economic mechanisms.”2 He recognizes that economists are partly to blame for this unhappy state of affairs. Some economists sequester themselves in their ivory towers, refusing to participate in policy debates where their input could be valuable; other economists are all too willing to provide superficial soundbites to the media that are counterproductive in their simplicity and sometimes not even within their area of economic expertise.

Tirole fears that, unchecked, the move away from economic thinking will result in policies that are disastrous for the common good. Too often, discussion of a policy is based on superficialities rather than a complete assessment of the policy’s actual benefits and costs, including a comparison to potential alternative policies. He points out that indignation often enters into policy discussions, but is a poor guide: ignoring that there are tradeoffs because making such tradeoffs seems “immoral” (e.g., consider a potential market for organ donation) is just sticking one’s head in the sand.

So, how can economics contribute to sound policy-making? Tirole identifies a number of ways. Economists are good at identifying adverse unintended consequences that a well-intended policy may have. He gives an example of how the French government prohibited the opening of large

---


2 JEAN TIROLE, ECONOMICS FOR THE COMMON GOOD 481 (2017).
supermarkets, a policy intended to restrain the market power of supermarket chains. But the unintended consequence was that it restrained competition by preventing entry (both store expansion by existing chains and new chains), which in turn resulted in higher prices for consumers—the exact opposite of what was intended. Economists are trained to consider all of the effects of a policy, including effects that are small for each individual in a given group but significant in the aggregate. The voice of such groups is often lost in the political process in comparison to what may be a small number of individuals for whom the effects are large. Economists “insist[] on reality rather than fairytale,” which counters wishful thinking and other cognitive biases that would otherwise favor poor economic policies. Finally, economists do not let emotion get in the way of reasoned thinking about tradeoffs. Tirole asks, “Isn’t it absurd to spend a large sum on saving one life when, for the same amount of money, dozens of others could be saved? Yet, the financial nature and apparent cynicism of such considerations shock people, who refuse to engage with them.”

Tirole’s proposed approach for regaining the public’s respect for the profession is for economists to strive to “mak[e] economic ideas comprehensible to a general audience,” thereby providing the public with a basis to understand which policies would serve the common good. This will make it harder for the false promises of economically unsound policies to gain political support and also make it easier for governments to implement sound policies, the benefits of which are hard for the economically naïve to perceive. This book is a useful step in that direction.

An early section of Economics for the Common Good is devoted to gently convincing readers that much of what they think they know about economics is wrong. Uniquely for a general interest book on economics, it tries to build a connection with the reader by explaining in detail the process by which economic research is conducted, from theoretical model-building to empirical testing to peer review. To dispel the caricature of economics as a static field clinging to abstract mathematical models detached from reality, Tirole explains how economics has, in fact, substantially expanded the range of human preferences and motivations that it encompasses. For example, moving beyond the basic assumption of the individual as homo economicus, economists have studied how preferences for things like altruism can be incorporated into economic models, which in turn provides a deeper understanding of many real-world phenomena. For example, paying a fee to blood donors could decrease, rather than increase, donations. Why? Because when donors are paid, donation loses its value as a signal of virtue. Given that many people have preferences for appearing virtuous to others (or themselves), they could be less inclined to donate blood after its virtue-signaling value has been diminished by the introduction of payments for donations.

Having established that economics has something useful to contribute, Tirole sets forth a set of principles that should guide economists’ participation in public policy debates. While this section may seem to be aimed at his fellow economists, it also helps the general public evaluate which economists they should trust: an economist worth listening to is one who has expertise in the area under discussion, makes economic judgments, not political pronouncements, and exhibits humility in acknowledging limitations on the state of economic knowledge and uncertainties in the analysis.

3 Id. at 19.
4 Id. at 40 (footnote omitted). Economists generally are in favor of understanding the tradeoffs involved when making a decision. Which route is chosen ultimately may be a political decision that incorporates factors such as notions of morality, but this decision will be better informed if the tradeoffs are known.
5 Id. at 481.
exhibits humility in acknowledging limitations on the state of economic knowledge and uncertainties in the analysis. Some of these principles are also applicable to economists who serve as expert witnesses in antitrust proceedings. For example, Tirole suggests that economists should never make a statement in a public policy setting that they would not also make in a seminar in front of other economists. This echoes the Supreme Court’s statement in Kumho Tire that an expert should “employ[] in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.”

Tirole views the current institutions for policy-making in Europe and the United States as problematic, with politicians driven to avoid grappling with hard solutions by their incentives to be re-elected. Instead, they adopt feel-good, but ineffective (and sometimes even counterproductive) policies. He favors making elected officials responsible only for setting high-level principles, with relatively independent administrative agencies being responsible for implementing detailed policies in line with those high-level principles. Such agencies can be staffed with experts capable of performing the serious analysis and thinking that is necessary to produce sound policy. Being independent, they would not be subject to the same political pressures as elected officials. Tirole points to competition authorities as exemplars of the independent expert agency. However, it is unlikely this approach would win over the populists whom Tirole wishes to persuade, given their suspicions concerning “experts” and fears that independent agencies take on a life of their own (e.g., “the deep state”).

In trying to further the common good, should we count on the market or the state? Tirole answers forcefully that both are necessary and that indeed, they are complementary in the sense that each serves as a remedy for the other’s limitations. With their decentralized nature, markets excel at producing outcomes that reflect the information held by market participants. States that have attempted to take over decision-making for firms or industries generally have failed spectacularly—the former Soviet Union, China under Mao, Cuba, and, most recently, Venezuela are powerful examples. On the other hand, markets are subject to certain failures, such as externalities (e.g., pollution) and internalities (e.g., cognitive biases of consumers). Tirole argues that it is the role of the state to correct such failures, for example, by using tradeable emissions permits in the case of pollution externalities. He emphasizes that state interventions in the market should recognize the state’s informational limitations as compared to market participants. Thus, tradeable emissions permits are preferred to state-imposed industry-by-industry caps on emissions because, with the former, the market sorts out which firms have the highest marginal value of being able to pollute, something the government would be unlikely to do well given its limited information. Tirole notes that studies have shown that the performance of industry-by-industry emissions caps has been vastly inferior to that of tradeable permits on a cost per unit of emission reduction basis.

A difficult question is defining what we mean by the “common good.” Tirole notes that it is ultimately a value judgment, rather than one based on economics, and a judgment on which people may differ. He suggests that reaching agreement on a definition may be easier if our contemplation takes place behind a Rawlsian “veil of ignorance,” i.e., in a hypothetical state that exists

---


8 James Madison’s argument for a republican form of government, rather than direct democracy, runs along similar lines. The Federalist No. 10 (Madison).
before we know how abilities, endowments, etc. will be allocated across people. For example, behind the veil, prior to knowing either how much pollution we each will create or how much we each will suffer at the expense of others’ pollution, it may be easier to agree on the idea that the common good should require a polluter to pay for the externality it creates. Similarly, the concept of social insurance may be easier to agree to behind the veil, when each person may yet be the recipient of bad luck (e.g., poor health), than ex post when allocations are known. Where economists come in, Tirole notes, is helping to understand what is economically feasible, given people’s incentives, preferences, and the information available to them and to the state, and how to develop policies that will further the common good, once defined.

In the second half of the book, Tirole turns to important policy issues of the day. He does not always offer concrete answers, but explains how economics can help us to understand the problems better and begin thinking about solutions that will actually be effective. He starts with chapters on the “macroeconomic challenges” presented by, respectively, climate change, the labor market, the state of the European Union, and the financial crisis of 2008. He then turns to the “industrial challenge,” where he addresses competition policy and industrial policy. Speaking to the populists, Tirole notes that fostering competition is important, not only for prices, efficiency, and innovation, but also to avoid corruption and “crony capitalism.”

Tirole devotes two chapters to the digital economy. He offers an excellent explanation of how digitization has been revolutionary, in particular by creating an environment that has supported development of two-sided platforms. Platforms are distinguished from traditional firms by their matching of individual buyers and sellers. On the one hand, the information available through digitization makes such matching possible; on the other hand, the vast amounts of information about potential matches would overwhelm platform users without the platform’s assistance in finding a good match. A nugget the book provides that may be surprising to many non-economists is that digitization and platforms actually can lead to higher prices for some products. An example is used books. Prior to the existence of platforms to identify matches, sellers were unlikely to encounter buyers with large valuations, leading them to sell at low prices to the low valuation buyers they encountered. The platform puts them in touch with buyers who have high valuations. Of course, prior to the platforms, high valuation buyers also had trouble finding sellers with the books they wanted. So, having the platform is better not only for sellers, but also for buyers—despite the higher prices.

Tirole explains why competition policy for platforms can be tricky—for example, low prices on one side of the market may be a poor indicator of predatory pricing by the platform because the low prices may be designed to build users on one side of the market, which will ultimately result in more users on the other side of the market and therefore the ability to charge high prices on the other side. He also notes competition issues that arise in the platform context, such as an incumbent platform tying a complementary product to the platform to deter entry of a supplier that might start with the complementary product that interacts with the incumbent’s platform and later use this product as a springboard to introduce a competing platform.

Tirole also offers chapters on intellectual property, explaining, among other things, how to design patent pools to solve the royalty stacking problem while avoiding anticompetitive coordinated interaction, and regulation. Changes in the way natural monopolies are regulated around the world has been a significant success story for economics, and one in which Tirole was cen-

---

trally involved. For example, Tirole (with frequent collaborator Jean-Jacques Laffont) showed that an effective way to regulate a firm with multiple products is to impose a cap on the average price of the firm’s products, but then let the firm, which has better information than the regulator, choose a structure of prices subject to the cap. The firm will have the incentive to choose prices for the products that satisfy Ramsey pricing, i.e., the price of each product will be proportional to the inverse of its demand elasticity, which is efficient. Price caps in general also provide strong incentives for the firm to reduce its costs.

Economics for the Common Good achieves Tirole’s goals of showing non-economists what economics can do and pointing ways forward on the serious policy dilemmas facing the world. Perhaps equally important is that a certain decency and sincerity pervade the book. One is left with the thought that if policies were designed and implemented by an army of Tiroles, we would see dramatic improvements in economic outcomes and the common good. Unfortunately, it seems that decency and sincerity are on the wane. Will Economics for the Common Good fall on deaf ears? Or will we see a movement back towards decency: keeping an open mind, assuming others have honorable intentions until proven otherwise, and engaging in honest debate? If the latter, Economics for the Common Good provides a roadmap for how, together, we can move the world to a better place for people of all political stripes.


Laffont & Tirole, Competition in Telecommunications, supra note 10, at 66–73.