Harvard Journal of Law & Technology
Volume 22, Number 2 Spring 2009

DATA MINING AND ANTITRUST

*Douglas M. Kochelek**

TABLE OF CONTENTS

I. INTRODUCTION

In 2000, customers of Amazon.com discovered that the online retailer was varying the prices charged for DVDs depending on the identity of the purchaser.[1] Although Amazon discontinued what it described as a "price test"[2] after public outcry, Amazon's brief foray into first-degree price discrimination stands as a noteworthy example of the possibilities for price discrimination using aggregated data. In its price test, Amazon sought to use information it already had about its customers to predict higher prices that the customers would still be likely to pay.[3] Nearly a decade later, brokers of consumer information

* Harvard Law School, Candidate for J.D., 2009; B.S.E.E., Rice University, 2006.
1. *See* David Streitfield, *On the Web, Price Tags Blur*, WASH. POST, Sept. 27, 2000, at A1.
2. *Id.*
3. *See* Lior Jacob Strahilevitz, *Reputation Nation: Law in an Era of Ubiquitous Personal Information*, 102 NW. U. L. REV. 1667, 1732–33 (2008).

now sell terabytes of data for the purposes of market segmentation and other consumer analytics.[4]

Such aggregations of consumer data are ripe for applications of data mining technologies.[5] These technologies enable producers to recover more of the economic surplus created by a transaction with a particular consumer by facilitating the development of first-degree price discrimination schemes.[6] In contrast to a perfectly competitive market where producers capture only their marginal costs and consumers capture the entire economic surplus of a transaction, price discrimination allows producers to recover some or all of the economic surplus. Thus, effective first-degree price discrimination reduces the welfare of consumers compared to a competitive market.

In addition to effecting a redistribution of wealth, price discrimination incentivizes consumers to engage in aftermarket arbitrage. It also incentivizes producers to develop mechanisms to prevent such arbitrage and to invest in more effective price discrimination schemes. These changes in behavior waste resources that would otherwise be efficiently allocated in a competitive market. Furthermore, perfect price discrimination is impossible in real markets. Consequently, imperfect price discrimination imposes deadweight losses that would not occur in a competitive market.[7] The policy rationales advanced to justify antitrust doctrine recognize that each of these results is an economic loss. However, although the policies behind antitrust law tend to disfavor price discrimination, the doctrines do not typically proscribe such discriminatory conduct.

Part II of this Note examines the mechanics of current data mining technologies and distinguishes between uses that promote price discrimination and uses that serve other ends. Part III considers the economic effects of data mining technologies used to facilitate price discrimination. Part IV examines the policies and doctrines underlying the Robinson-Patman Act and the Sherman Act and argues that the policies that justify the Sherman Act are consistent with enforcement against data-mining-based price discrimination, although it is not available under present doctrine. Even if this conduct is not proscribed, the presence of data-mining-based price discrimination is indicative of the presence of other harms that are proscribed by the doctrine. Part V concludes that current antitrust policy and doctrine are mismatched and that, without legislative or judicial augmentation of the doctrine, data mining technology will likely pose a greater risk of future economic loss.

---

4. *See* IAN AYRES, SUPER CRUNCHERS 134–35 (2007); Acxiom: Data Products, http://www.acxiom.com/products_and_services/data_products/Pages/DataProducts.aspx (last visited May 15, 2009) (marketing databases of segmented consumer information).

5. *See* Streitfield, *supra* note 1.

6. *See* discussion *infra* Part III.

7. *See* discussion *infra* Part IV.B.1.a.

II. THE MECHANICS OF DATA MINING

Data mining refers to the process of extracting patterns from data.[8] For instance, a credit card issuer may mine transaction data to detect suspicious transactions to reduce credit card fraud;[9] astrophysicists may mine telescope data to select regions in space for more careful investigation;[10] or advertisers may mine consumer data to tailor advertising to particular demographics.[11] In each case, the knowledge sought (e.g., suspicious transactions) is obtained by examining prior data (e.g., past instances of fraud) to find relationships between data that is easy or inexpensive to observe (e.g., the location or dollar value of purchases) and the prior data. In the banking example, if purchases are made in a country that is distant from anywhere the credit card has been used in the past, the bank may flag the card as potentially stolen and may contact the owner. To determine how unusual the use should be before a call is made, the bank could compare a transaction with instances of fraud reported on other accounts. In this relatively simple way, a bank could make predictions about the legitimacy of future transactions on the basis of patterns of past credit fraud.[12]

Since large repositories of data rarely track only one variable like the location of a transaction, the designer of a real-world prediction system would want to use other similarly correlated variables available to him,[13] such as the amount of the transaction, the identity of the seller, and the date and time of the sale. To predict which transactions are most likely to be fraudulent, the bank would weigh and combine the variables in such a way as to minimize the error in its prediction of past fraudulent transactions.[14] The resulting mathematical model would permit the bank to make predictions about the fraudulent character of future transactions.

Although additional variables may be available for use in a prediction model, the inclusion of more variables in the model will not necessarily improve its accuracy. Additional variables that are redun-

---

8. *See, e.g.*, Michael J. Shaw et al., *Knowledge Management and Data Mining for Marketing*, 31 DECISION SUPPORT SYSTEMS 127, 128 (2001) ("Data mining is the process of searching and analyzing data in order to find implicit, but potentially useful, information.").

9. *See, e.g.*, American Express, Fraud Protection Center, http://www.americanexpress.com/cards/online_guarantee/ (last visited May 15, 2009).

10. *See* Steve Lohr, *Two New Ways to Explore the Virtual Universe, in Vivid 3-D*, N.Y. TIMES, May 13, 2008, at F3.

11. *See, e.g.*, Shaw et al., *supra* note 8, at 127–37.

12. *See* Philip K. Chan et al., *Distributed Data Mining in Credit Card Fraud Detection*, IEEE INTELLIGENT SYSTEMS, Nov./Dec. 1999, at 67, 67.

13. The designer would want to include both individually measured variables that correlate with the prediction and combinations of measurements — i.e., cross terms — that correlate with the prediction.

14. *See generally* CHRISTOPHER M. BISHOP, NEURAL NETWORKS FOR PATTERN RECOGNITION 1–32 (1995) (describing methods for error minimization in multiple variable statistics).

dant with those already included in the model[15] provide no greater predictive power. The extent to which one variable is redundant with another or a combination of others is measured by the correlation between them.[16] The greater the correlation between one variable and another variable or group of variables, the lesser the predictive power increased by introducing the additional variable into the model. Conversely, the less correlated two variables are, the greater the potential for the inclusion of the second variable to increase the predictive power of the model. In this way, the additional data must not be redundant with the data already included in the model. As a result, the accuracy of a predictive model for data mining depends on obtaining data that both correlates with the predicted variable and comes from non-redundant sources.

Separately from the inclusion of multiple uncorrelated variables, models can be significantly improved by developing them from larger data sets. As more samples (e.g., individual past instances of fraud) are included in a model, the accuracy of the model improves.[17] Weaker relationships between variables can be more readily identified from large data sets than from small ones. In this way, mining large sets of data has the potential to more accurately identify weaker relationships between variables and consequently provide more accurate predictions. Thus, transactions that take place in large volumes, such as online purchases of commodities, are better suited to data mining than low-volume transactions.

Because the practicability of data mining depends only on the existence of large volumes of diverse data that correlates with a feature sought to be predicted, its potential applications are immensely varied. However, since many applications clearly do not have economic — much less antitrust — ramifications, this Note will not develop their implementation any further. The particular application with which this Note is concerned is use of data mining techniques to determine a party's "pain point" in a transaction. The "pain point" is the most undesirable value at which a party will still engage in a transaction.[18] For instance, if a person were applying for a loan, the person's pain point

---

15. For instance, state would be redundant if zipcode was already included in the model because zipcodes do not overlap states. In practice, the distinction in this example may be overly simplistic because zipcodes are considerably more numerous than states, and relationships between large groups of zipcodes and predicted variables may go unnoticed unless those zipcodes are grouped into states.

16. *See generally* JOHN NETER, ET AL., APPLIED LINEAR STATISTICAL MODELS 285–95 (4th ed. 1996) (discussing the problem of multicollinearity and its effects on regression analysis).

17. *Cf.* 1 WILLIAM FELLER, AN INTRODUCTION TO PROBABILITY THEORY AND ITS APPLICATIONS 245 (3d ed. 1968) (applying the law of large numbers to explain that as sample sizes increase, the sample average approaches the population average); NETER ET AL., *supra* note 16, at 6–8 (noting that regression models presume the existence of a probability distribution of the response variable for each combination of predictor variables).

18. *See, e.g.*, AYRES, *supra* note 4, at 173.

would be the highest interest rate at which the person would take out the loan. For a lender, the borrower's pain point is the highest possible price that could be obtained in a particular transaction.

Accurate prediction of a party's pain point provides obvious negotiation advantages to the counterparty: the counterparty knows how far he can push the party without compromising the deal. As a result, parties have a strong incentive to develop techniques for predicting the pain points of their counterparties. Since a producer can use data mining to find correlations between consumers' pain points and characteristics of consumers that can be observed inexpensively,[19] the techniques have clear applications for pain point prediction. Producers have an incentive to aggregate and mine any and all data that may be sufficiently distinct to better predict consumers' pain points.

The inexpensively observable characteristics that are relevant to a particular pain point determination will depend on the specifics of the transaction. For instance, people living in Boston would probably be willing to pay a higher price for Red Sox paraphernalia than people living in New York, but both groups would probably exhibit similar demand for a gallon of milk. In addition to the correlation between geography and prices of certain goods, other less intuitive correlations may exist between inexpensively observable variables.[20] This is precisely the value of data mining: with large quantities of diverse data, mathematical algorithms can be applied to identify useful but unintuitive correlations between inexpensively observed characteristics that can be exploited to make significantly more accurate predictions.

It is worth noting that although typical data mining databases are large and span a diverse set of variables, such databases do not necessarily operate on personally identifiable information.[21] Although it is certainly possible for databases to contain such information,[22] or to contain so much information about a party to a transaction that the

---

19. Inexpensively observable variables (e.g., transaction location) should be distinguished from variables that are expensive to observe (e.g., whether a transaction is fraudulent) to emphasize that data mining is undertaken to make estimations about variables that would otherwise be prohibitively expensive or impossible to observe. If it were inexpensive to observe whether a transaction was fraudulent, then there would be no reason to engage in data mining for such a purpose.

20. Submissions in the competition for the Netflix Prize have identified meaningful relationships among ratings of movies that bear little ostensible relation. *See* Clive Thompson, *If You Liked This, You're Sure To Love That*, N.Y. TIMES, Nov. 23, 2008, § 6 (Magazine), at 74 (describing a meaningful relationship between ratings of *Joan of Arc*, *W.W.E.: SummerSlam 2004*, *It Had to Be You*, and *Bleak House*). For a description of the contest, see Netflix Prize, http://www.netflixprize.com (last visited May 15, 2009).

21. *See* Erika McCallister et al., *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)* 2-1 to -2 (Nat'l Inst. of Standards & Tech., Special Publication 800-122 (Draft), 2009), *available at* http://csrc.nist.gov/publications/drafts/800-122/Draft-SP800-122.pdf (defining personally identifiable information).

22. For example, a database of online orders including shipping addresses would contain personally identifiable information.

likelihood of two entities having the same characteristics is statistically insignificant,[23] any sort of privacy harm or risk existing because of the database is a separate harm from any economic harm effected by the use of the database for data mining purposes. The privacy harm comes from the possession or misuse of the personal information,[24] while any harm attributable to data mining would arise from the use of information, personally identifiable or not. Thus, while a database used for data mining may inflict privacy harms and data mining harms concurrently, the harms are distinct.[25]

III. THE ECONOMICS OF DATA MINING FOR PRICE DISCRIMINATION

When data mining is used to identify a consumer's pain point, it facilitates first-degree price discrimination.[26] Price discrimination, if perfect, is as efficient an allocation of resources as perfect competition. However, even in perfect first-degree price discrimination, producers and consumers incur secondary costs beyond those present in a competitive market.

Most transactions in which parties engage have the potential to take place at a variety of prices.[27] There is typically a maximum price that a buyer is willing to pay and a minimum price that a seller is willing to accept. Obviously, either party would be willing to take a better price than that for which they would walk away from the deal. The difference between the two parties' limit prices constitutes the economic surplus of the transaction. This surplus is distributed between the two parties on the basis of the parties' agreed price for the transaction. Thus, the distribution of the surplus between the parties is a function of their negotiations in reaching their agreement.[28]

In analyzing the economic effects of data-mining-based price discrimination, it is significant that the prices consumers pay are largely beyond their control. Consumers are price-takers in most data mining

23. This possibility was realized in 2006 when AOL released a database of search queries for researchers from which users were able to be identified on the basis of their search data. *See, e.g.*, Michael Barbaro & Tom Zeller, Jr., *A Face Is Exposed for AOL Searcher No. 4417749*, N.Y. TIMES, Aug. 9, 2006, at A1.

24. *See generally* Will Thomas DeVries, *Protecting Privacy in the Digital Age*, 18 BERKELEY TECH. L.J. 283, 284–91 (2003) (discussing varying conceptions of privacy).

25. *See* FTC, Statement of Federal Trade Commission Concerning Google/DoubleClick FTC File No 071-0170 (Dec. 20, 2007), *available at* http://www.ftc.gov/os/caselist/0710170/071220statement.pdf (declining to prohibit Google's acquisition of DoubleClick, Inc. because of privacy concerns, but also concluding that the acquisition would not adversely affect non-price attributes of competition).

26. *See* William W. Fisher III, *When Should We Permit Differential Pricing of Information?*, 55 UCLA L. REV. 1, 4 (2007).

27. *See, e.g.*, ROBERT H. MNOOKIN ET AL., BEYOND WINNING 18–22 (2000) (discussing the zone of possible agreement that may exist between two parties to a negotiation).

28. *See, e.g.*, *id.*

transactions because their individual purchases comprise a small fraction of a producer's large volume of sales, and, as a result, they have little bargaining power.[29] If consumers were not price-takers, a producer's knowledge of a consumer's pain point alone would not permit first-degree price discrimination; some distribution of the surplus between producers and consumers could still take place through negotiation. In situations where consumers have bargaining power and are not price-takers, data mining would provide an informational advantage to the producer but would not necessarily result in a higher price.[30] However, because consumers are price-takers for many transactions where data mining is practicable — for instance, consumers do not expect to negotiate on the price of a book purchased from Amazon — the assumption that consumers are price-takers is generally valid for data-mining-influenced transactions.

In perfectly competitive markets,[31] consumers will pay the marginal cost of each producer's last unit of production.[32] This occurs because as producers compete to sell to consumers, they have an incentive to undercut their competitors' prices until they reach an equilibrium at the point just before they lose money on the sale, i.e., charge less than the cost of producing the good. Consumers value goods differently — some would be willing to pay more than the marginal cost of the good's production, but others would not. Accordingly, any consumer that would have paid a price higher than the marginal cost receives a benefit equal to the difference between the price he would have paid and the price he actually paid — the consumer surplus. Since at equilibrium the producers receive only their costs of production, there is no producer surplus. In this way, a perfectly competitive market allocates the entire surplus resulting from trade in a good to the consumers of the good.

If a producer secures a monopoly on production in a market, then he will be able to obtain a price higher than the cost of the good's production. The price that the producer chooses to charge will be that which maximizes the producer's profits: where marginal revenue equals marginal cost.[33] At such a price, fewer consumers — only those who value the good at, or more than, the higher monopoly price — will purchase the good. These consumers will receive a smaller consumer surplus than they would have recieved in a competi-

29. *See, e.g.*, ANDREU MAS-COLLEL ET AL., MICROECONOMIC THEORY 20–21 (1995).

30. *See* MNOOKIN ET AL., *supra* note 27, at 21–23 (explaining how asymmetric information can affect negotiations).

31. Perfectly competitive markets are typically characterized by a large number of producers who compete to sell the same good to a large number of consumers; there are no artificial barriers to producers' entering and leaving the market; firms behave so as to maximize their profits; and both producers and consumers have complete information. *See* PHILLIP AREEDA ET AL., ANTITRUST ANALYSIS 5 (6th ed. 2004).

32. *See id.* at 12–13.

33. *See id.* at 10–12.

tive market. The producer, on the other hand, will now receive a positive producer surplus because the monopoly price exceeds the cost of the good. In this way, a monopoly results in economic inefficiencies because it reduces the overall allocation of resources: some trades that would have occurred at lower prices do not occur. Furthermore, for those transactions that do occur in a monopoly, there is a transfer of benefit from consumers to producers relative to the competitive outcome.

A monopoly firm can increase both its profit and its output by offering its product at multiple prices. A firm that sells a good at two different prices — a higher price for consumers who will pay it and a lower price for all other consumers — will sell a greater quantity of the good and will receive a larger profit than it would with a single price. However, to maintain its discriminatory pricing, the firm must also invest in identifying workable price strata and segregating consumers into them. Additionally, the firm must establish mechanisms to prevent arbitrage between the low-paying consumers and the high-paying consumers. Without these mechanisms, consumers will expend resources in aftermarket arbitrage. Thus, a firm can increase its producer surplus through price discrimination, and this increase will reduce the deadweight loss of a single monopoly price,[34] but maintaining discriminatory measures requires other expenditures.

If a monopoly firm could determine and charge the value that each individual consumer placed on the good and could maintain the price discrimination scheme, then that firm could achieve perfect price discrimination.[35] In other words, the monopoly firm could charge each consumer the maximum value he would be willing to pay. Under such circumstances, each consumer who would have purchased the good at the competitive price (the marginal cost of the good) would still purchase the good. However, because the consumers would be paying the maximum amount they are willing to pay, the producer would receive the economic surplus of every transaction. In this way, in perfectly price discriminatory markets, resources are allocated as efficiently as in a competitive marketplace, but the benefit of the transaction goes to the monopoly producer rather than to the consumers.

A firm need not be a monopolist in order to effect discriminatory pricing. Firms acting in coordination, or acting in mutually recognized joint interest, may engage in discriminatory pricing. Such pricing will likely result in slightly smaller transfers and losses than monopolistic

---

34. *See* RICHARD A. POSNER, ANTITRUST LAW 80 fig.2 (2d ed. 2001) (illustrating the difference between a discriminating and non-discriminating monopoly and the effect on deadweight loss).

35. *See generally* HAL R. VARIAN, INTERMEDIATE MICROECONOMICS 445–47 (7th ed. 2006) (defining perfect price discrimination).

discriminatory pricing, but such oligopolistic behavior will still produce losses and costs in excess of those present in a competitive market.[36] In this way, oligopolies may also exploit data-mining-based price discrimination to enhance profits by capturing economic surplus from consumers while effecting economic losses similar to those caused by a monopolistic price discriminator.

Thus, a firm that wields monopoly or oligopoly power and charges consumers at their maximum willingness-to-pay can capture the surplus generated by the transactions. Even in real-world markets,[37] a price-discriminating producer can capture a significant share of the surplus that would have been captured by consumers in a competitive market.[38] Furthermore, because firms must invest in developing mechanisms of price discrimination and because consumers may expend resources in arbitrage, any discriminatory pricing scheme imposes secondary costs that a competitive market does not. Finally, because imperfect price discrimination is the actual result in real-world markets, deadweight losses from mis-priced and mis-allocated resources will occur that would not exist in competitive markets.[39] In this way, although price discrimination offers a more efficient allocation of goods than single-price monopolization, even perfect price discrimination results in an inferior outcome to that of a competitive market: producers rather than consumers receive the economic surplus, and secondary costs of maintenance, development, and arbitrage constitute waste.

IV. Antitrust Doctrine and Policy Pertaining to Price Discrimination

Two major statutes govern price discrimination in the antitrust realm: the Robinson-Patman Act's amendments to the Clayton Act[40] and the Sherman Act.[41] Although its provisions relating to price dis-

---

36. *Cf.* AREEDA ET AL., *supra* note 31, at 10–14 (explaining that the economic results in oligopoly markets fall between those of perfectly competitive markets and monopoly markets).

37. Real-world markets seldom satisfy all of the conditions for perfect competition. *See* AREEDA ET AL. *supra* note 31, at 8–9 (describing some limitations on perfect competition). Furthermore, price discrimination is never perfect. POSNER, *supra* note 34, at 80.

38. *See, e.g.*, W. KIP VISCUSI ET AL., ECONOMICS OF REGULATION AND ANTITRUST 344–49 (4th ed. 2005) (demonstrating how third-degree price discriminators can obtain higher profits even when total outputs decrease).

39. *See* HERBERT HOVENKAMP, FEDERAL ANTITRUST POLICY 576 (3d ed. 2005).

40. 15 U.S.C. § 13(a) (2006).

41. *Id.* §§ 1–2. Although the Clayton Act is also an applicable antitrust statute, this Note does not examine its applicability because the Clayton Act provides for relief for the same conduct and injuries reached by the other acts. *See* Clayton Act §§ 4, *id.* § 15 (providing injured persons with the power to sue for treble damages for injuries caused by violations of the antitrust laws); Clayton Act §§ 16, *id.* § 26 (providing injunctive relief for the same).

crimination appear more applicable, the Robinson-Patman Act neither provides for enforcement against data-mining-based price discrimination losses, nor do its policies suggest that it should. The policies underlying the Sherman Act, on the other hand, recognize the losses incurred in markets subject to data-mining-based price discrimination, but the doctrine does not currently permit enforcement. At the very least, data-mining-based price discrimination is evidence that recognized harms under the Sherman Act may be present.

A. The Robinson-Patman Act

Antitrust law generally views price discrimination through the lens of monopoly power and predatory pricing. Section 2 of the Clayton Act as amended by the Robinson-Patman Act of 1936 deals with price discrimination expressly:

> It shall be unlawful for any person engaged in commerce, in the course of such commerce, either directly or indirectly, to discriminate in price between different purchasers of commodities of like grade and quality . . . where the effect of such discrimination may be substantially to lessen competition or tend to create a monopoly in any line of commerce, or to injure, destroy, or prevent competition . . . .[42]

Because the Act applies only to transactions in "commodities of like grade and quality" which have the effect of substantially lessening competition, many transactions are beyond the reach of the Act.[43] Data mining, however, is most applicable to transactions in commodities;[44] the large data sets that are necessary for the algorithms are more readily obtainable for commodity transactions than they are for other transactions simply by virtue of the relative abundance of commodity transactions. In this way, the applicability of Robinson-Patman to data-mining-based price discrimination schemes depends on whether such schemes tend to create a monopoly or injure competition.

This Note limits its discussion to the conduct proscribed and harms avoided by the acts and does not address effects related to the procedural differences between the acts.

42. *Id.* § 13(a) (2006).

43. *See, e.g.*, ROGER D. BLAIR & DAVID L. KASERMAN, ANTITRUST ECONOMICS 296–97 (2d ed. 2009).

44. "Commodities" have been construed to encompass "goods, wares, merchandise, machinery and supplies." Columbia Broad. Sys., Inc. v. Amana Refrigeration, Inc., 295 F.2d 375, 378 (7th Cir. 1961) (concluding that a contractual right to sponsorship identification on broadcast television was not a commodity as contemplated by 15 U.S.C. § 13(a)).

The Robinson-Patman Act has been interpreted to regulate three classes of competitive injuries: primary-line injuries, secondary-line injuries, and indirect price discrimination injuries.[45] Primary-line injuries are those injuries suffered by the competitors of the discriminating firm.[46] The legal theory behind primary-line injuries is similar to that behind predatory pricing across geographic markets: a multimarket firm can reduce prices in one geographic market to eliminate competitors in that market while maintaining its revenues through continued profits in other markets.[47] Since the Supreme Court's decision in *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*,[48] the prerequisites to recovery on a claim of price discrimination under the Robinson-Patman Act have been essentially the same as for a claim of predatory pricing under section 2 of the Sherman Act.[49] Data-mining-based price discrimination harms consumers who face higher prices, rather than other producers. Furthermore, such discrimination is clearly not a form of geographic predatory pricing. Accordingly, the practice likely cannot be regulated under the primary-line theory of the Robinson-Patman Act.

Secondary-line injuries are suffered by purchasers of discriminatorily-priced goods who compete in a resale market.[50] In *FTC v. Morton Salt Co.*,[51] the Court held that volume discounts on table salt that were attained by only a few purchasers — large grocery chains — constituted illegal price discrimination because large buyers could secure a competitive advantage over smaller ones purely on the basis of their buying ability. More recently, in *Volvo Trucks North America, Inc. v. Reeder-Simco GMC, Inc.*,[52] the Court held that purchasers must be in actual competition for sales to the same customer in order for a competitive injury to exist.[53] If purchasers facing discriminatory prices must be in competition with one another for an injury under the Act to exist, then end-consumers of discriminatorily-priced commodity goods will not have suffered a competitive injury because they are not in competition for sales at all. For instance, typical purchasers from Amazon are not competing to resell books. In this way, secondary-line injuries likely cannot form a basis for regulation of data-mining-based price discrimination under the Robinson-Patman Act because the consumers are not competing in a resale market.

---

45. *See* BLAIR & KASERMAN, *supra* note 43, at 297.
46. *Id.*
47. *See id.*
48. 509 U.S. 209 (1993).
49. *See id.* at 221–24. "[P]rimary-line competitive injury under the Robinson-Patman Act is of the same general character as the injury inflicted by predatory pricing schemes actionable under § 2 of the Sherman Act." *Id.* at 221.
50. *See* BLAIR & KASERMAN, *supra* note 43, at 301.
51. 334 U.S. 37 (1948).
52. 546 U.S. 164 (2006).
53. *See id.* at 177.

Congress's attempts to regulate indirect price discrimination are codified in the remaining sections of 15 U.S.C. § 13. These provisions regulate payments of commissions to brokers and discriminatory allowances and promotional services to resellers.[54] None of these provisions can be plausibly applied to discriminatorily-priced sales to end-consumers. As such, no theory of application of the Robinson-Patman Act is likely to reach transactions between data mining price discriminators and their customers.

That no provision of the Robinson-Patman Act is directly applicable to price discrimination is consistent with the observations of commentators who have noted that the Act seems to discourage low prices in order to protect small resellers, rather than discriminatory prices.[55] In this way, neither the doctrine nor the underlying policy of the Robinson-Patman Act seems to favor its application to discriminatory pricing measures implemented through data mining by retailers of commodity goods.

B. The Sherman Act

Although the Robinson-Patman Act likely does not provide a means to regulate data-mining-based price discrimination, the Sherman Act's prohibitions against monopolization[56] and restraints of trade[57] provide a stronger basis for regulating data-mining-based price discrimination. Because similar data mining techniques among competitors can result in an oligopoly price, use of data mining techniques for price discrimination may violate the Sherman Act.

1. Policy

The policies upon which the Sherman Act rests generally oppose the waste and transfers that result from persistent price discrimination. Regardless of the particular theory of the purpose of antitrust regulation one accepts,[58] the economic effects of data-mining-based price discrimination suggest that such conduct ought to be proscribed by the Sherman Act.

---

54. *See* 15 U.S.C. § 13(c)–(e) (2006).
55. *See, e.g.*, HOVENKAMP, *supra* note 39, at 578–79.
56. 15 U.S.C. § 2.
57. *Id.* § 1.
58. *See* HOVENKAMP, *supra* note 39, at 48–49 (discussing the purposes of the Sherman Act). *See generally* THE POLITICAL ECONOMY OF THE SHERMAN ACT (E. Thomas Sullivan ed., 1991) (collecting articles proposing foundational policies for antitrust).

a. Allocative Efficiency

Many prominent antitrust commentators advocate that the purpose of the Sherman Act is the promotion of the efficient allocation of resources — i.e., allocative efficiency[59] — in the marketplace.[60] The proponents of an efficiency justification for antitrust law maintain that efficiency is the key social value that justifies the existence of antitrust policy.[61] Furthermore, because efficiency justifies antitrust policy, it also limits enforcement to those circumstances where efficiency is improved by eliminating or restricting anticompetitive behavior.[62] Under this rationale, the test for the propriety of antitrust regulation ought to be whether a particular enforcement mechanism improves allocative efficiency.

If a producer is able to engage in perfect price discrimination among consumers, then all consumers who would have been able to purchase in a competitive market will still be able to purchase goods from the producer, but at a price greater than or equal to that of the competitive market. Such a scheme has no deadweight loss. The only difference is that every purchaser pays the highest price he would be willing to pay rather than the producer's marginal cost of production. In this way, it appears that proponents of the allocative efficiency justification should be indifferent toward perfect price discrimination when compared to perfect competition since neither results in deadweight loss.

Even if the price discrimination scheme enacted by a data miner were perfect, the fact that consumers are receiving the same goods at different prices will induce consumers to engage in arbitrage. Aftermarket transactions between consumers will take place so long as the cost of the transaction is less than the difference in consumers' valuations of the good. Aftermarket arbitrage results in secondary transaction costs that are not present in a single-price — i.e., competitive — market because no consumer in the single-price market can obtain the commodity at a more favorable price than any other consumer. This arbitrage in discriminatorily-priced goods wastes resources that could be allocated otherwise in a competitive market.

---

59. The literature distinguishes allocative efficiency from distributive efficiency. *See, e.g.*, HOVENKAMP, *supra* note 39, at 49–50. The former refers to the allocation of resources so that utility is optimized across society, while the latter refers to normative concerns relating to the distribution of wealth within society. *See id.* ("A policy is purposefully distributive only if it is adopted instead of a policy believed to be more efficient, because the adopted policy distributes wealth in a way that the policy maker finds more appealing.").

60. *See, e.g.*, ROBERT H. BORK, THE ANTITRUST PARADOX 107–15 (1978) (emphasizing the importance of offsetting increases in deadweight loss with increases in efficiency elsewhere); POSNER, *supra* note 34, at 2 (characterizing the efficiency justification as the "consensus view").

61. *See* POSNER, *supra* note 34, at 2.

62. *See id.*

Where there are opportunities for arbitrage, a price-discriminating firm may attempt to impose controls to limit this behavior. Although such measures are more easily imposed in markets for non-commodity goods,[63] firms may also be able to impose them in markets for commodity goods. For example, digital rights management technology for electronically distributed music could be used to prevent arbitrage.[64] These measures are developed and deployed only for the purpose of maintaining discriminatory prices. As with the transfer costs associated with aftermarket arbitrage, producers' expenditures to maintain price discrimination create waste that is not present in a competitive market.

Firms that price discriminate must also expend resources to develop and deploy means to do so.[65] In the context of data-mining-based price discrimination, these costs include the acquisition of the requisite data sets, the development and application of algorithms to make pricing predictions, and the implementation of infrastructure to actually apply the discriminatory pricing to the firm's customers. Since firms would not make these additional investments in a competitive market,[66] these costs constitute additional wasted resources of price discrimination.

Moreover, perfect price discrimination is not actually possible in a real marketplace,[67] and implementations of price discrimination may lead to deadweight loss. Although imperfectly competitive markets may exhibit deadweight losses, the losses in these markets result from prices that are elevated above the marginal cost. That is, only the consumers who valued the good the least are affected. In imperfectly price-discriminatory markets, deadweight losses may occur anytime a firm's profit-maximizing algorithm overestimates a consumer's valuation of a good. Thus, instead of confining deadweight losses to the consumers who value the good the least, discriminatorily-priced markets may fail to satisfy demand of high-valuing consumers. These unsatisfied consumers constitute the deadweight loss associated with imperfect price discrimination. In this way, real-world price discrimination can potentially produce more deadweight loss than real-world competition, but the comparison between particular instances of real-

63. *See, e.g.*, United States v. United Shoe Mach. Corp., 110 F. Supp. 295, 298–99 (D. Mass. 1953), *aff'd per curiam*, 347 U.S. 521 (1954) (leasing discriminatorily-priced machinery to prevent resale).

64. *See, e.g.*, Timothy K. Armstrong, *Digital Rights Management and the Process of Fair Use*, 20 HARV. J.L. & TECH. 49, 50–51 (2006). Digital rights management technology generally restricts purchasers of digital media, such as music, from selling those electronic copies to others. *See id.* Unlike rights-managed music, a compact disc can be sold to another consumer.

65. *See* POSNER, *supra* note 34, at 85–86.

66. *See id.*

67. *Id.* at 80.

world competition and real-world price discrimination depends on the circumstances.

In sum, although perfect price discrimination and perfect competition result in the same number of goods consumed, price discrimination leads to wasted resources on arbitrage transfers, arbitrage prevention, and discriminatory price determination and implementation. Furthermore, because perfect price discrimination exists only theoretically, imperfect price discrimination creates deadweight losses that may exceed deadweight losses in imperfectly competitive markets. In this way, the allocative efficiency justification for the Sherman Act suggests that the Act should proscribe data-mining-based price discrimination because the practice creates waste and losses that are only possible because of the monopoly or oligopoly circumstances in a market.

b. Consumer Welfare Maximization

Although some commentators describe "consumer welfare maximization" as synonymous with allocative efficiency,[68] others have used the term to refer to a distributive goal behind antitrust law. Professor Robert Lande has argued that antitrust laws were passed to prevent transfers of wealth away from consumers by firms exerting market power.[69] A monopoly results in the transfer of wealth from consumers to producers because a portion of the surplus of the transaction, which a competitive market would allocate to consumers, accrues instead to the monopolist.[70] The same transfer occurs to a lesser extent in oligopoly.[71] These redistributive effects typically exceed the deadweight loss of monopoly, which is caused by a decreased equilibrium quantity compared to a competitive market, by a factor between two and forty.[72] The merit of such a transfer depends entirely on the welfare context in which one evaluates the transaction between the producer and the consumer. Lande argues that Congress's enactment of the antitrust statutes constituted an entitlement of consumers to a competitive economic system and the allocation of surpluses to consumers implicit therein.[73]

If one accepts such a distributive policy justification for antitrust law, then the analysis of data-mining-based price discrimination is relatively straightforward. Commentators like Lande disfavor the transfer of wealth from consumers to producers inherent in monopo-

68. *See* BORK, *supra* note 60, at 51.

69. *See* Robert H. Lande, *Wealth Transfers as the Original and Primary Concern of Antitrust: The Efficiency Interpretation Challenged*, 34 HASTINGS L.J. 65, 69–70 (1982).

70. *See id.* at 74–75.

71. *See supra* text accompanying note 36.

72. *See* Lande, *supra* note 69, at 75.

73. *See id.* at 76.

lies; price discrimination amplifies this transfer, making it increasingly disfavored. Thus, practices that effect discriminatory pricing, be they by a monopoly or an oligopoly, ought be proscribed by the Sherman Act.

c. Business Fairness

Beyond the efficiency and distributive rationales, other commentators have proposed that the antitrust laws were enacted to promote more abstract values of business fairness. Professor Louis Schwartz notes that, in addition to protecting overall competition, the Sherman Act also prohibits conspiracies among competitors that may overwhelm a smaller competitor.[74] According to Schwartz, both these protections and the Robinson-Patman Act's concern for price discrimination evince a congressional concern for the non-economic goal of "justice" in the sense of "fair and equal treatment of persons in like situations."[75] Schwartz advocates that these non-economic factors should be considered as part of the antitrust analysis and should shift the burden of persuasion that the behaviors are justified to the suspect firm.[76]

If antitrust policy is based on congressional intent to effect "justice" in the form of like treatment for similarly situated firms, then antitrust policy almost certainly disfavors data-mining-based price discrimination. Pricing designed to exploit purchasers' individual pain points does not comport with a notion of justice that requires like treatment for those similarly situated. If, alternatively, one were to permit "justice" to be done within data-mined classes of customers because such consumers were "similarly situated," then one could also permit conspiracies against competitively insignificant firms because they are not similarly situated to the larger firms. This result is inconsistent with Schwartz's observation. Thus, if antitrust law disfavors inequitable treatment among similarly situated consumers, it ought to proscribe data-mining-based price discrimination.

d. Small Business Favoritism

Professors George Stigler and Herbert Hovenkamp, among others, have suggested that the antitrust statutes were passed as protectionist measures for small business interests against larger, more

74. *See* Louis B. Schwartz, *"Justice" and Other Non-Economic Goals of Antitrust*, 127 U. PA. L. REV. 1076, 1078 (1979).
75. *Id.*
76. *Id.* at 1080.

efficient firms.[77] If the Sherman Act exists to protect small businesses from larger, more efficient firms, then its applicability to discriminatory pricing schemes depends largely on which firms have the access and the capability to exploit the technology. Since some amount of market power is required to effect discriminatory pricing, the presence of data-mining-based price discrimination likely would be evidence of market power under such a theory of antitrust law but would not constitute a separate harm. If, however, data-mining-based price discrimination were available to all market participants, then an antitrust policy favoring small business interests likely would be neutral toward the practice because of its equal availability. Of course, if the benefit of data-mining-based price discrimination increased with the size of the firm, then such a policy might discourage data-mining-based price discrimination depending on the size of the firm and the effect of the practice on the firm's competitors. In this way, under a small business protection rationale for antitrust law, the appropriateness of proscription should depend on the particular circumstances: a harm to small businesses is not always present. However, when such harms are present, the rationale likely favors proscription of the practice.

2. Monopoly Doctrine

Section 2 of the Sherman Act prohibits unlawful monopolization, attempts to monopolize, and conspiracies to monopolize trade.[78] Although data mining technologies can facilitate price discrimination within a market that is already monopolized, the technologies cannot themselves effect a monopoly. In a competitive market, if one producer were to raise prices discriminatorily, other competitors would undercut him or new entrants would enter the market. However, data-mining-based price discrimination could be used by a monopolist in one market to extract increased profits to fund predatory pricing in other markets in which he participates. Under such circumstances, the law would proscribe the conduct effecting the predatory harm in the competitive market rather than the discriminatory conduct that created the economic loss.

For the reasons already described, data mining technologies are able to facilitate price discrimination within a monopoly market.[79] If,

---

77. *See* HOVENKAMP, *supra* note 39, at 48–52 (comparing rationales behind the Sherman Act and concluding that promotion of small business interests is the most plausible); George J. Stigler, *The Origin of the Sherman Act*, 14 J. LEGAL STUD. 1, 7 (1985).

78. 15 U.S.C. § 2 (2006). The Act has not been interpreted to prohibit monopolies that are the most efficient configuration of a particular market. *See* HOVENKAMP, *supra* note 39, at 275–76. The Act prohibits activities that monopolize, but not the attainment of monopoly by means that are not exclusionary. *See* AREEDA ET AL., *supra* note 31, at 368.

79. *See* discussion *supra* Part III.

however, the discriminating firm is not already a monopolist or does not already possess market power, the introduction of discriminatory pricing behavior will not benefit the firm. So long as competitors are present in the market with available capacity, any attempt by a firm to price discriminate will only cause consumers to purchase from the firm's competitors instead. It is only if the competing firms also raise their prices to discriminatory levels that customers will actually pay the discriminatory prices.[80] Thus, a firm that is not already a monopoly will not be able to monopolize a market by instituting discriminatory pricing. As such, section 2 of the Sherman Act does not prevent the economic harms caused by data-mining-based price discrimination because price discrimination either arises in a market that is already monopolized or is more likely to invite competition than to reduce it.

Although discriminatory pricing cannot establish a monopoly within a competitive market, discriminatory pricing can extract greater monopoly profits from an already-monopolized market. Those profits can be used to subsidize predatory pricing behavior in a separate competitive market. Such predatory pricing behavior can be used to establish a monopoly by eliminating competition. Under section 2, however, the relevant harm is the eventually elevated prices in the predatory market,[81] not the economic harm in the monopolized market nor the primary-line discrimination, which falls under the Robinson-Patman Act.[82] In this way, although the Sherman Act proscribes some predatory pricing conduct, these proscriptions do not address the losses imposed on already-monopolized markets.

Despite the lack of a Sherman Act proscription against price discrimination because of harms inflicted on discriminated markets, the presence of price discrimination within a market is itself evidence that the discriminating firm has some amount of market power. Because such pricing schemes can only be maintained by firms with monopoly power — otherwise competitors would undercut the firm or new firms would enter the market — persistent data-mining-based price discrimination probably supports an inference that the discriminating firm wields some amount of market power.

3. Oligopoly Doctrine

Section 2 of the Sherman Act prohibits monopolies and attempts to monopolize a market; by its terms, it applies only to single-firm

80. *See infra* Part IV.B.3 (discussing oligopoly behavior).
81. *See* HOVENKAMP, *supra* note 39, at 340.
82. *See* Brooke Group Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209, 222–24 (1993).

domination of a market and does not reach oligopolies.[83] Oligopolies can, however, be regulated under section 1 of the Sherman Act if they rise to the level of a "conspiracy" to restrain trade.[84] Since data mining technologies facilitate discriminatory pricing based on inexpensively observable consumer characteristics, the same technologies can be employed by multiple firms to establish parallel discriminatory pricing schemes.[85] Yet, because courts require plaintiffs in section 1 cases to present evidence that "tends to exclude the possibility" of independent action,[86] it is unlikely that a successful case could be marshaled against price-discriminating data miners.

By its terms, section 1 of the Sherman Act forbids contracts, combinations, or conspiracies in restraint of trade.[87] Some form of agreement is necessary in order for a violation to occur.[88] The Court has made it clear that conscious parallelism between firms is not in itself illegal when firms effectively share monopoly power and set prices by recognizing shared economic interests.[89] In this way, firms must do more than simply recognize and rationally respond to shared economic interests in order to establish a section 1 violation.

Although conscious parallelism is beyond the reach of section 1, the Court has recognized that there is often no direct evidence of an agreement among competitors to fix prices.[90] As such, courts have recognized the need for so-called "plus factors" that can supplement evidence of parallel behavior to support an inference of a conspiracy.[91] These plus factors include evidence of conspiratorial motivation, acts against self-interest, poor economic performance, traditional evidence of a conspiracy, or other evidence of an agreement.[92]

---

83. *See, e.g.*, Consol. Terminal Sys. v. ITT World Commc'ns, 535 F. Supp. 225, 228–29 (S.D.N.Y. 1982); Am. Tel. & Tel. Co. v. Delta Commc'ns Corp., 408 F. Supp. 1075, 1106 (S.D. Miss. 1976).

84. *See* 15 U.S.C. § 1 (2006); *cf.* Theatre Enters., Inc. v. Paramount Film Distrib. Corp., 346 U.S. 537, 541 (1954) (indicating that "conscious parallelism" is insufficient for establishing conspiracy under the Sherman Act).

85. *See* discussion *supra* Part III.

86. Matsushita Elec. Indus. Co. v. Zenith Radio Corp., 475 U.S. 574, 588 (1986) (quoting Monsanto Co. v. Spray-Rite Serv. Corp., 465 U.S. 752, 764 (1984)).

87. 15 U.S.C. § 1.

88. BLAIR & KASERMAN, *supra* note 43, at 235.

89. *See* Brooke Group Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209, 227 (1993).

90. *See* Local Union No. 189, Amalgamated Meat Cutters v. Jewel Tea Co., 381 U.S. 676, 720 (1965).

91. *See, e.g.*, Petruzzi's IGA Supermarkets, Inc. v. Darling-Del. Co., 998 F.2d 1224, 1232–33 (3d Cir. 1993) (requiring a plaintiff to demonstrate plus factors to exclude the possibility of independent action by the defendant); *see also Matsushita*, 475 U.S. at 588 (requiring that a plaintiff show that the inference of a conspiracy is reasonable in light of the competing inference of independent action).

92. *See* 6 PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW § 1434 (2d ed. 2003).

Assuming that two firms choose to implement data-mining-based price discrimination schemes independently, none of the traditional evidence of a conspiracy or other evidence of an agreement will be available to support a claim under section 1. Although firms may not use identical algorithms, their input data and output prices likely will be similar. Data mining services for particular purposes are widely available, and data brokers exist that sell enormous databases precisely for the purpose of data mining.[93] As such, two firms data mining independently likely will have comparable predictions of a particular consumer's pain point. Such a priori pricing is similar to the pre-announcements of pricing changes found in some tacit agreements.[94] In this way, although the firms are not actively sharing information, the firms' use of similar means of prediction and the availability of the same source data for purchase suggests that an elevated oligopoly price may persist.

An argument may be posed that a single firm has no incentive to institute discriminatory pricing because in the presence of competitors with excess capacity or in a market with low barriers to entry, the price discriminator's customers would simply buy elsewhere. As the argument goes, a firm would not engage in price discrimination in a competitive market unless its major competitors were also going to do so. If, however, the capacity of the other suppliers to a market is not sufficient to satisfy all the demand at the competitive price, and if there are non-negligible barriers to entry, then one firm's introduction of data-mining-based price discrimination could be reasonable as an independent action because the firm would be able to raise prices to some extent.[95] Under such circumstances, it is foreseeable that a firm would institute data-mining-based price discrimination independently of its competitors.

As such, deployment of data-mining-based price discrimination is consistent both with independent action and with agreement. Where it appears that firms are tacitly colluding, a court may view marketwide data-mining-based price discrimination in an oligopoly to be circumstantial evidence of an agreement. However, because there is no inherent characteristic of data mining that makes its adoption by one or several firms in a market an agreement under section 1, the Sherman Act likely does not reach the conduct even though the conduct may impose oligopoly losses on the market.

---

93. *See, e.g.*, Acxiom: Data Products, *supra* note 4.

94. *See* BLAIR & KASERMAN, *supra* note 43, at 238.

95. *See* 1 ABA SECTION OF ANTITRUST LAW, ANTITRUST LAW DEVELOPMENTS 238–41 (5th ed. 2002) (explaining the relationship between market power and barriers to entry, including capacity limitations).

V. CONCLUSION

Data-mining-based price discrimination schemes fall into a gap between antitrust doctrine and the policies underlying the doctrine. As the costs of implementing such schemes decrease and data sources become more robust and diverse, market participants will be more likely to employ such techniques to extract additional profits. As a result, if the doctrine is not augmented — judicially or legislatively — then it is likely that producers of commodity goods will become able to capture larger shares of the economic surplus of transactions. Under any policy rationale for antitrust, transfers of wealth away from consumers, resources wasted on development and preservation of discriminatory pricing schemes, and the deadweight loss associated with imperfections suggest that antitrust law should discourage such practices. Unless antitrust doctrine adapts to the economic losses potentially imposed by data-mining-based price discrimination, increased deployment of the technology may reduce consumer welfare, waste resources, and reduce allocative efficiency in exchange for increased producer profits that are insufficient to justify their cost.