BIG DATA AND DIFFERENTIAL PRICING

February 2015

EXECUTIVE OFFICE OF THE PRESIDENT OF THE UNITED STATES
COUNCIL OF ECONOMIC ADVISERS

Contents

Executive Summary

Big data refers to the ability to gather large volumes of data, often from multiple sources, and with it produce new kinds of observations, measurements and predictions. Commercial applications of big data deserve ongoing scrutiny given the speed at which both the technology and business practices are evolving. One of the many questions raised by big data is whether companies will use the information they harvest to more effectively charge different prices to different customers, a practice that economists call price discrimination. Economics suggests that many forms of differential pricing, such as senior citizen discounts at the box office or tiered pricing for air travel, can be good for both businesses and consumers. However, the combination of differential pricing and big data raises concerns that some consumers can be made worse off, and have very little knowledge why. This report finds that many companies already use big data for targeted marketing, and some are experimenting with personalized pricing, though examples of personalized pricing remain fairly limited. While substantive concerns about differential pricing in the age of big data remain, many of them can be addressed by enforcing existing antidiscrimination, privacy, and consumer protection laws. In addition, providing consumers with increased transparency into how companies use and trade their data would promote more competition and better informed consumer choice.

Introduction

Understanding the customer is a core principle of good marketing, and over time companies have developed a wide variety of tools for doing so. These tools range from surveys to focus groups, rewards programs and the quarterly sales meeting. For many companies, big data and consumer analytics are an increasingly important part of this tool kit.

In the marketing context, big data refers to the ability to gather large volumes of data, often from multiple sources, and use it to produce new kinds of observations, measurements and predictions about individual customers. Much of what companies learn through big data is used to design products and services that deliver more value to the individual consumer. At the same time, if sellers can accurately predict what a customer is willing to pay, they may set prices so as to capture much of the value in a given transaction, especially when they face little competition.

This report considers the implications of big data and customer analytics for the American consumer, with a particular emphasis on how these tools might be used for differential pricing. The report strives to provide an economic perspective on these issues, focusing on the underlying technology, how it is used, the potential costs and benefits for both buyers and sellers, and the kinds of policies that can best promote efficiency, equity, and innovation in this space.

Big data clearly holds both promise and peril for the individual consumer. As the Executive Office of the President's 2014 report *Big Data: Seizing Opportunities, Preserving Values* recently observed, "It is one thing for big data to segment consumers for marketing purposes, thereby providing more tailored opportunities to purchase goods and services. It is another, arguably far more serious, matter if this information comes to figure in decisions about a consumer's eligibility for… employment, housing, health care, credit or education" (Podesta et al. 2014). This report explores these issues, beginning with an overview of the economics of differential pricing. It goes on to describe how sellers are using big data to create personalized marketing campaigns and pricing strategies, and how buyers are responding. The concluding section considers how big data and personalized pricing fit into our existing framework of antidiscrimination and consumer protection laws.

I. The Economics of Differential Pricing

Differential pricing, or what economists call "price discrimination," is the practice of charging customers different prices for the same product. While this sounds unfair, many forms of differential pricing generate few objections. For example, venues like movie theaters that charge a price of admission may offer discounts to particular groups, such as children, senior citizens, or members of the military. Business travelers often pay a higher price for the same plane ticket or hotel room if they purchase closer to the date of travel. And a variety of big-ticket items – products ranging from a new car to a university education – have a list price, but offer individualized discounts that vary from one customer to the next.

In each of these examples, the seller's goal is to raise prices for those willing to pay more, without losing another group of more price-sensitive customers. That is the general idea behind differential pricing – to set prices based on demand, or what customers are willing to pay, rather than costs.

Economics textbooks usually define three types of differential pricing. Personalized pricing, or first-degree price discrimination, occurs when a seller charges a different price to every buyer. Individually negotiated prices, such as those charged by a car dealer, are an example of personalized pricing. Quantity discounts, or second-degree price discrimination, occur when the per-unit price falls with the amount purchased, as with popcorn at the movie theater. Finally, third-degree price discrimination occurs when sellers charge different prices to different demographic groups, as with discounts for senior citizens.

Big data has lowered the costs of collecting customer-level information, making it easier for sellers to identify new customer segments and to target those populations with customized marketing and pricing plans. The increased availability of behavioral data has also encouraged a shift from third-degree price discrimination based on broad demographic categories towards personalized pricing. Nevertheless, differential pricing still presents several practical challenges. First, sellers must figure out what customers are willing to pay. This can be a complex problem, even for companies with lots of data and computing power. A second challenge is competition, which limits a company's ability to raise prices, even if it knows that one customer might be willing to pay more than another. Third, companies need to prevent resale by customers seeking to exploit price differences. And finally, if a company does succeed in charging personalized prices, it must be careful not to alienate customers who may view this pricing tactic as inherently unfair.

When differential pricing is possible, economic theory suggests that it can produce both costs and benefits. The main benefit is that when sellers have some market power, differential pricing allows them to expand the size of the market. For example, matinee prices encourage large families and music lovers on a tight budget to take in a show. If a theater were prohibited from using this type of differential pricing, it might decide to keep prices high and leave some seats empty. This would mean less profit for the theater and fewer people getting entertained.

Similarly, financial aid packages help universities bring in more tuition by charging the list price to those who can afford it, while educating more students who might be excluded if need or merit-based financial aid were prohibited. These forms of differential pricing typically generate few objections because they appeal to customers' sense of fairness – companies charge a bit more to the least price-sensitive customers, who can probably afford it, and a lower price to those who cannot.

On the other hand, one cost of differential pricing is that it can produce incentives to inefficiently degrade product quality. Early railroad operators, for example, sought to increase profits by charging wealthy customers more for passage but did not have a good way of determining which customers were willing to pay more. As a response, some railroads chose to provide no roof on third-class carriages in order to increase the difference in quality (and price) between a first- and third-class ticket. The modern equivalent might be disabling features built into a car or a smartphone, or degrading the speed of an Internet connection, in order to create high and low-end versions of the same product (see box on versioning). In each case, sellers can profit at the expense of high-end customers by degrading the low-end product, even if it is costly to reduce overall quality.

WHAT IS VERSIONING?

Firms often produce multiple versions of a product to encourage consumers to *self-select* into groups that pay different prices, even when it would be more cost efficient to sell a single design. For example, at one time IBM sold two versions of its LaserPrinter Series E, where the only difference between them was a chip that made the low-priced version print more slowly (Deneckere and McAffee 1996). While intentionally disabling some features of a product to facilitate price-discrimination seems perverse, it can nevertheless increase welfare for both firms and consumers if it allows the seller to reach a larger number of customers who would not otherwise be served.

Versioning is especially prevalent with information goods such as books, films, or software because the costs of reproduction are typically small relative to the price. For these products, companies often release multiple versions over time (e.g. hardcover, paperback, and e-book) or add and remove features (e.g. bonus tracks or concurrent user limitations) as part of their product-line strategy.

It is difficult to predict how big data will influence the prevalence of versioning. If it becomes easier to predict individual customers' willingness to pay and charge different prices for an identical product, versioning may be replaced by personalized pricing. On the other hand, versioning has the benefit of reducing concerns about inequity that arise with personalized pricing, and big data may facilitate versioning strategies based on "mass customization," particularly for information goods that can be customized at relatively little incremental cost.

Another concern with differential pricing is that it transfers value from consumers to shareholders, which generally leads to an increase in inequality and can therefore be inefficient from a utilitarian standpoint. This is particularly true in settings where there is no competition, and few consumers would exit the market, even if a firm raised prices dramatically.

Ultimately, whether differential pricing helps or harms the average consumer depends on how and where it is used. In a competitive market with transparent pricing, the benefits are likely to outweigh the costs.[1] For example, while there is lots of differential pricing in airline ticket sales, the Internet has made it relatively easy for many travelers to compare prices and itineraries across airlines and to select the best deal for any given trip. Some studies even suggest that differential pricing can intensify competition relative to uniform pricing, by allowing high-margin sellers to compete more aggressively for price-sensitive customers who might otherwise buy from a lower-priced rival.[2]

Even in the absence of competition, consumers have several tools that can be used to undermine differential pricing. One of those tools is collective action, through negotiating group discounts. Another tool is arbitrage, or the ability to become a seller. For example, if a seller offers the same product in two cities (or through two distribution channels) at very different prices, an entrepreneurial buyer could purchase several copies in the low-price market and resell them at a profit in the high-price market, so long as the price difference exceeds the cost of transportation.

Ultimately, differential pricing seems most likely to be harmful when implemented through complex or opaque pricing schemes designed to screen out unsophisticated buyers. For example, companies may obfuscate by bundling a low product price with costly warranties or shipping fees, using “bait and switch” techniques to attract unwary customers with low advertised prices and then upselling them on different merchandise, or burying important details in the small print of complex contracts.[3] When these tactics work, the economic intuition that differential pricing allows firms to serve more price-sensitive customers at a lower price-point may even be overturned. If price-sensitive customers also tend to be less experienced, or less knowledgeable about potential pitfalls, they might more readily accept offers that appear fine on the surface but are actually full of hidden charges.

Finally, it is worth noting that the preceding discussion focused on “value-based” pricing, where prices reflect differences in consumers’ *willingness to pay* for a particular good or service. That

---

[1] In a survey of recent literature, Armstrong (2006) notes that competition generally allows consumers to better defend themselves from price-discriminating firms and keep prices low.

[2] See, for example, Corts (1998).

[3] "Economic Scene; The Usual Decorous Waltz between Prices and Sales Becomes a Lively Tango in the World of Online Sales." Hal Varian. *The New York Times*, December 19, 2002. Varian describes research by Ellison and Ellison (2009) showing that e-commerce sites advertise very low prices for products that, upon closer inspection, prove to be cheap knock-offs or missing key accessories. Houssain and Morgan (2006) find that buyers on e-Bay are attracted to lower product prices even when they are paired with much higher shipping fees, and may end up spending more in total when bidding on these products.

is what most economists mean by differential pricing. However, sellers may also charge prices that reflect differences in the *cost* of serving different groups of buyers. This type of "risk-based" pricing arises most commonly in insurance markets, where prices reflect the risk that an individual will experience the outcome covered by an insurance policy. Big data encourages risk-based pricing by enabling more fine-grained measurement of various risks, for example through tracking individual driving behaviors.[4]

Risk-based pricing can improve economic efficiency by discouraging risky behavior, such as when individuals with a history of traffic accidents are charged more for auto insurance. It can also make insurance more widely available by reducing adverse selection, which occurs when only high-risk individuals enroll at a given (uniform) price. At the same time, differential pricing in insurance markets can raise serious fairness concerns, particularly when major risk factors are outside an individual customer's control, with health insurance an obvious example.

In general, risk-based pricing favors less risky customers, whereas value-based pricing favors those who are more price-sensitive. Nondiscrimination policies are one way to promote fairness for high-risk buyers. However, those policies can re-introduce adverse selection problems unless they are accompanied by rules or subsidies that encourage low-risk buyers to remain in the market.[5]

The remainder of this report is primarily focused on value-based pricing, which has different economic motivations and implications from risk-based pricing. However, the impact of big data on risk-based pricing remains an important area for further discussion.

[4] See, for example, "So You're a Good Driver? Let's Go To the Monitor," Randall Stross, *New York Times*, November 24, 2012.

II. Big Data and Personalized Pricing

To understand how big data enables personalized pricing, it is useful to start with an overview of the technology. Computers have long been used to collect sales data, organize customer lists, and identify market segments. The features of big data that promise to make it more informative, however, are the increased scale of the underlying databases, the increasing variety of customer-level observations and measures, and the speed with which data are now harvested, traded, and deployed.

Two broad trends are driving the increased application of big data to marketing and consumer analytics. The first trend is the widespread adoption of new information technology platforms, of which the Internet and the smartphone are the most important. These platforms provide access to a wide variety of applications such as search engines, maps, blogs, and music or video streaming services. These new applications, in turn, create new ways for businesses to interact with consumers that produce new sources and types of data. For example, it is now possible to track:

- a user's location via mapping software
- their browser and search history
- whom and what they "like" on social networks like Facebook
- the songs and videos they have streamed
- their retail purchase history
- the contents of their online reviews and blog posts.

Sellers can also utilize these new types of information to make an educated guess about consumer characteristics. For example, some web sites use a computer's Internet Protocol (IP) address to guess the user's location.[6] Others might use the items in a virtual shopping basket to infer a buyer's gender, or the history of web sites that a user has browsed to guess at their income or health status.[7]

To some extent, this new ability to measure a consumer's digital footprint depends upon a company's relationship with that consumer. For example, when an Internet user visits a web site, the owner of the site may place a file called a "cookie" onto the user's computer, enabling the site to keep track of information about the user's interactions with the site. Over time, cookies can be used to build a long-term picture of an individual's Internet browsing history, and that information can be shared across sites (see box: What is a Cookie?). However, it is considerably easier to track customer behaviors on web sites or mobile applications that require users to create an account and log into that account with each use. In addition to simplifying online tracking, account holders typically provide personal information that a site can use to link them with other external information sources.

---

[6] IP is an acronym for Internet Protocol, and an IP address is the numeric identifier used to locate an individual computer on the Internet.
[7] "The Web's New Gold Mine: Your Secrets," Julia Angwin, *The Wall Street Journal*. July 30, 2010.

What is a Cookie?

A cookie is a small text file that a web site can place on a user's computer. Each time a user loads a particular web site, the cookie is sent to that site. This allows web sites to "remember" certain information, such as what pages a user has already visited, or whether they are currently logged in to the site. Internet browsers generally allow users to set various permissions that control whether cookies are allowed on their computer.

Cookies were created by programmers working for Netscape in 1994, and the name is a reference to "magic cookies" – a term used to describe a piece of data that a program receives and then retransmits unchanged. Today, cookies are used for a wide variety of purposes, most notably tracking customers across multiple sites in order to send them behaviorally targeted advertising.

Cookies are not regulated in the United States. However, in 2009 the European Union modified its e-Privacy directive to regulate cookies. In particular, the Directive told EU member states to pass laws requiring users to "opt in" or provide consent before placing a cookie on their computer.

The second trend driving the application of big data to marketing is the growth of the ad-supported business model, and the creation of a secondary market in consumer information. Companies like Google and Facebook, both of which earn much of their revenue by selling targeted marketing opportunities, demonstrate the commercial potential of ad-supported Internet platforms.[8] The ability to place ads that will be targeted to a specific audience based on their personal characteristics makes consumer information increasingly valuable to businesses. This has fostered a growing industry of data brokers and information intermediaries that buy and sell customer lists and other data used by marketers to assemble a digital profile of individual consumers.

Given sufficient data, sellers can try to predict how buyers will behave in response to different prices and pricing schemes. While randomized experiments are one approach, non-experimental tools and techniques for predicting consumer behavior are also evolving rapidly. Predictive modeling is not a simple problem. However, companies have large incentives to refine these tools, since even small improvements can have a large impact on profitability, particularly for companies with a large customer base. For example, a 2014 recent study by Benjamin Shiller estimates the increase in profits if Netflix were to use behavioral data for personalized pricing. He finds that differential pricing based on demographics (whereby Netflix would adjust prices based on a customer's race, age, income, geographic location, and family size) could increase profit by 0.8 percent, while using 5,000 web browsing variables (such as the amount of time a user typically spends online or whether she has recently visited Wikipedia or IMDB) could increase profits by as much as 12.2 percent.

[8] "Programmatic Bidding: Buy, Buy, Baby." *The Economist*. September 13, 2014.

The potentially large benefits of personalized pricing lead naturally to the question of whether many companies are actually engaging in the practice. The remainder of this section examines how both sellers and buyers are adapting to the rapid diffusion of big data in the context of personalized marketing.

What Sellers are Doing?

Although a few studies have tried to detect differential pricing online, current knowledge is mainly anecdotal. Companies naturally protect information about pricing strategies for competitive reasons, and perhaps also for fear of a customer backlash. Nevertheless, the anecdotes suggest that we have not yet entered an era of widespread personalized pricing. Rather, sellers are using online and offline pricing practices that fall into three broad categories: (1) exploring the demand curve, (2) steering and differential pricing based on demographics, and (3) behavioral targeting and personalized pricing.

Exploring the Demand Curve

Experiments are a powerful way of learning about demand and consumer behavior, even in the absence of big data. As a consequence, marketers often conduct "A/B Tests" that randomly assign customers to one of two possible price conditions. These experiments are technically a form of differential pricing, since they result in different prices for different customers, even if they are "nondiscriminatory" in the sense that all customers are equally likely to face the higher price.

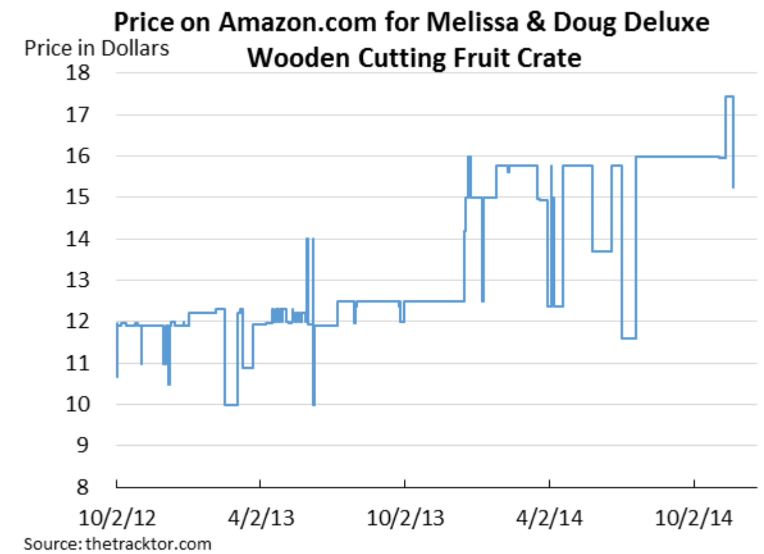
Offline businesses have long been able to explore the demand curve by testing prices in different stores or randomizing offers via direct mail. The Internet, however, provides a much better platform for running demand experiments quickly and effectively. For example, a recent study identified hundreds of thousands of "seller experiments" on eBay, where an identical item was listed multiple times by the same seller at different prices or with different auction parameters, presumably to learn how those variables influence demand for the underlying product (Einav et al 2011).

While such price experiments are common, they can still become controversial. For example, in 2000 users discovered that Amazon.com was conducting price tests and complained about paying different prices for the same DVD. Amazon's CEO Jeff Bezos apologized in a news release that same year indicating that the tests were random and promised that "We've never tested and we never will test prices based on customer demographics."[9]

Even if sellers do not wish to randomize prices across potential buyers at a point in time, it is often possible to collect similar data by raising and lowering prices for all customers over very brief time intervals. For example, if customers who arrive at a website at 10 am face a lower price than those who arrive at 10:15 am, and buy correspondingly more of a given product, the seller has discovered valuable information about the demand curve without technically offering different prices to its customers. To provide an example indicating that this phenomenon

[9] "Bezos Calls Amazon Experiment 'a Mistake.'" *Puget Sound Business Journal*. September 28, 2000.

happens, the following chart shows two years of prices for a particular children's toy on Amazon.com. While these frequent price changes are not necessarily experiments, they will certainly provide information about consumer demand for this item.



Steering

Steering is the practice of showing different products to customers in different demographic groups. In the online environment, steering occurs when a web site alters its search results based on information about a potential customer. Like third-degree price discrimination, steering uses information about potential buyers, but not at the individual level.

Both steering and differential pricing can happen with or without a customer's knowledge. As one example, Dell Computer's web site might offer the same laptop to different types of buyers (e.g. government, academic, individual, and business users) at different prices. This type of differential pricing is easily observed because buyers are asked to identify what group they belong to. On the other hand, customers may not be aware that some Internet retailers vary prices or steer customers to different products based on a computer's Internet address, which provides a proxy for their geographic location, or based on the type of device they are using.[10]

Computer scientists from Northeastern University recently found evidence of either steering or differential pricing at 4 out of 10 general merchandise web sites and 5 out of 5 travel web sites. Although this suggests widespread differential pricing, the same study found that these practices affected a very small number of products – less than 2 percent of the sampled products on each site in the study (Hannak et al 2014). Thus, the overall picture to emerge from

[10] For an example of differential pricing using IP addresses, see Valentino-Devries, Jennifer, Jeremy Singer-Vine, and Ashkan Soltani, "Websites Vary Prices, Deals Based on Users' Information." *The Wall Street Journal,* December 24, 2012. For an example of steering based on the type of operating system, see Martha C. White "Orbitz Shows Higher Prices to Mac Users." *Time,* June 26, 2012.

this study is that experimentation with steering and differential pricing is a common practice across web sites, but for a relatively limited set of products.

One reason why steering and third-degree discrimination are not more prevalent could be that it is difficult to infer much about Internet users without access to personal information. Variables such as the type of operating system, a user's IP address, or whether they are using a mobile device may contain some information about willingness to pay but are imperfect proxies at best. For this reason, much of the interest in big data and differential pricing has focused on data that captures individual users' browsing, search, purchase, or other behaviors.

Behavioral Targeting and Personalized Pricing

Behavioral targeting and personalized pricing use customer-specific information to target advertisements or tailor prices for a set of products. Historically, this kind of personalization required a human, such as a salesperson who could negotiate the price of each car or appliance. However, big data and electronic commerce have reduced the costs of targeting and first-degree price discrimination.

Loyalty programs provided some of the first applications of big data to personalized pricing. When a buyer joins a loyalty program, they typically provide some personal information and consent to a seller tracking their purchases. In return, the buyer typically receive some type of benefit such as seat upgrades or free flights from an airline frequent flier program, or price discounts on specific items from a grocery store. Sellers use loyalty programs to customize their marketing. Some retailers also partner with companies that aggregate data from loyalty programs and use it to create customized coupons, which are printed on the back of receipts that a customer receives at the cash register or point of sale. Firms that specialize in this type of personalization claim that data-driven analysis can increase the redemption rates on such coupons from 1 percent for non-personalized coupon offers to as much as 25 percent for highly-targeted coupons.[11]

Widespread adoption of the Internet and smartphones is producing many new types of behavioral data that can be used for targeted marketing. For example, online advertisers use browsing data collected from ad networks and third party cookies to send users targeted ads. Thus, someone who spends lots of time at a travel web site reading about Hawaii might begin to see advertisements for flights to Honolulu or hotels in Waikiki when they are later reading the news on another site in the same network. Many web services also use behavioral data to make customized recommendations. For example, Amazon recommends books, Netflix recommends movies, and Apple and Google recommend music based on customers' purchase and consumption patterns.

While behavioral targeting has been widely reported, there is relatively little evidence of personalized pricing on the Internet. For example, the Northeastern study described above

---

[11] "Catalina Marketing Aims for the Cutting Edge of 'Big Data.'" Doug Henschen, *InformationWeek,* September 6, 2011.

found no evidence of personalization based on user-specific information at any of the 10 general merchandise web sites in the study and could only discern behavioral discrimination at 1 out of 5 travel sites. Likewise, recent reports on the data broker industry produced by the Federal Trade Commission in 2014 and the Government Accountability Office in 2013 contained no specific references to examples of personalized pricing.

The relative scarcity of personalized pricing examples suggests that companies are moving slowly or remaining quiet, perhaps due to fears that consumers will respond negatively, but also because the methods are still being developed. For example, a new story about using big data to identify and target customers who are pregnant produced a number of headlines containing terms like "creepy."[12] At the same time, the story led practicing statisticians to point out that marketers are typically very cautious with this type of prediction, because when big data suggests that a customer is 75 percent pregnant, the 25 percent risk of marketing to a "false positive" customer is unacceptably high.[13]

The fact that companies have not yet embraced personalized pricing does not mean that sellers share the same general concerns as their customers. Consumer groups typically object to the fact that many buyers are only vaguely aware of what a website might know about them or how that information could be used. They also argue that in settings where the cost of a mistake is high, individuals should be able to ensure the accuracy of their data. Sellers appear more concerned with the reliability of statistical models and the potential reputational impact of their pricing practices. Nevertheless, companies' reputational concerns may reflect a belief that consumers will view personalized pricing as unfair if the data are not reliable, if the buyer did not opt-in to the data collection process (as with loyalty programs), or if price differences are not framed as discounts (as with coupons).

What Buyers are Doing?

While much of the interest in big data and differential pricing has focused on sellers, buyer behaviors are also revealing. In particular, three broad trends suggest that concerns about big data and personalized pricing are not stifling consumer activity on the Internet. Those trends are: (1) the rapid growth of electronic commerce, (2) the proliferation of consumer-empowering technologies, and (3) the slow uptake of privacy tools.

E-Commerce

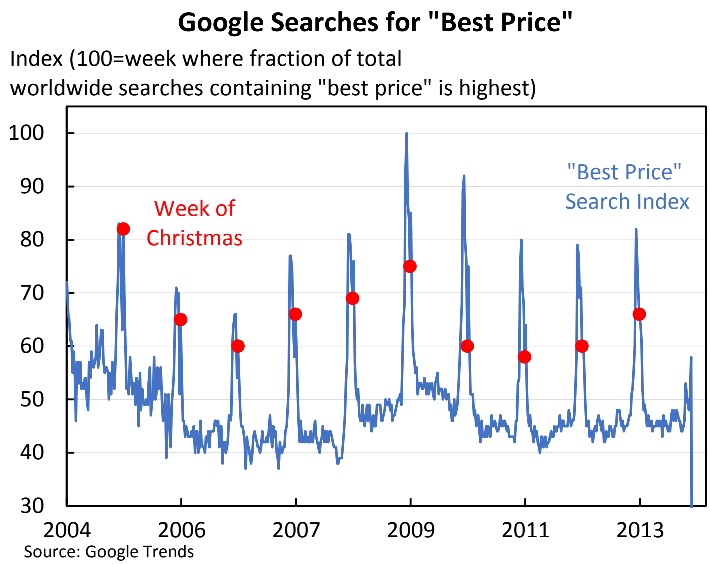
Americans are using the Internet to shop in rapidly growing numbers, suggesting that consumers believe they are getting a good deal on the Internet, regardless of any differences in the pricing practices of online and offline retailers. The U.S. Census estimates that e-commerce has increased from 2 percent of total U.S. retail sales in 2004 to 6 percent in 2014. Moreover, electronic commerce revenue is currently growing at a rate of 16 percent per year in the United States, more than three times the 5 percent growth rate in overall retail sales (Borgie 2014).

---

[12] "How Companies Learn Your Secrets." Charles Duhigg, *The New York Times,* February 16, 2012. "What's Even Creepier Than Target Guessing That You're Pregnant?" Jordan Ellenberg, *Slate,* June 9, 2014.
[13] "Big Data: Are we making a big mistake?" Tim Harford, *Financial Times*, March 28, 2014.

Consumer Technology

In the online environment, consumers have a variety of tools that may help them find a better price if they are concerned about steering or differential pricing. The simplest of these tools is the search engine. For example, the following chart shows the weekly frequency of worldwide Google searches that include the phrase "best price." The strong seasonality suggests that consumers are using the Internet for price comparisons during the run-up to holiday season, and perhaps more so during the recession.



Price comparison and price-tracking sites are another type of buyer-empowering technology. Price comparison sites, such as Travelocity, Expedia, or PriceGrabber, show the current price of similar items at many different online sellers at a point in time, making it easier for buyers to see the full range of prices. Price tracking sites provide the price that a single online seller offers for one item over time and can issue alerts to a potential buyer if the price falls below some threshold.[14]

The Internet has also strengthened the ability of arbitrageurs to undermine differential pricing by making it easier for buyers to become sellers. In particular, if a seller charges widely divergent prices to customers in different markets, it creates an opportunity for someone to buy at the low price and undercut the seller in the high-priced market. For example, a student named Supap Kirtsaeng created a substantial business out of importing foreign edition textbooks from Thailand and re-selling them on eBay. Websites like Amazon and Craigslist create similar opportunities. In Kirtsaeng's case, textbook sellers challenged his legal right to import the copyrighted works, thereby highlighting a more general point – our intellectual

[14] Search engines, price comparison, and price-tracking sites all require some effort on the part of the user (even with price alert services) and therefore work somewhat like coupons, by rewarding those who are most motivated to seek a bargain. Of course, users might reasonably wonder whether a given tool is truly offering the best possible unbiased comparisons, or if the site itself might be engaging in a certain amount of steering or targeted marketing.

property and contract laws, trade rules and safety regulations can all influence consumer bargaining power by placing limitations on re-sale.[15]

Privacy Tools

Finally, it is worth noting the relatively low adoption rates for some widely available consumer privacy tools. For example, most browsers allow users to set policies for accepting different types of cookies used to track their web surfing behavior. While there does not appear to be any systematic study of this behavior, a test by one Internet Service Provider (ISP) found that 96 to 97 percent of its users allow some cookies and 85-90 percent allow third-party cookies.[16] Similarly, users can opt out of targeted advertising on many web sites by clicking on a small triangular icon but often decline to do so.[17]

Privacy advocates suggest that users are unaware of these tools and unaware that advertisers are gathering and aggregating data about their Internet use. On the other hand, this could be "rational ignorance" on consumers' part, reflecting a view that the cost of engaging with details of privacy settings outweighs the benefits gained. Evidence on this topic is mixed, perhaps reflecting the rapidly evolving technical and commercial landscape. Some older surveys support the claim that many Internet users don't understand how their data are being gathered and used (Turrow, Feldman and Meltzer 2005). However, a 2014 survey by the Pew Research Center finds that 55 percent of Americans are "willing to share some information about [themselves] with companies in order to use online services for free" at the same time as 64 percent believe that government could do more to regulate what advertisers do with their personal information (Pew Research Center 2014).

---

[15] Kirtsaeng ultimately prevailed in the Supreme Court. See *Kirtsaeng v. John Wiley and Sons, Inc.*, in the Supreme Court of the United States, 11-697.

[16] The ISP study is described at http://webmasters.stackexchange.com/questions/58210/do-many-users-turn-off-cookies. A report issued in 2000 by Pew Research Institute suggests that 10 percent of users disabled cookies, and a press release by an internet marketing firm just one year later put the figure at 0.68 percent (see http://www.prnewswire.com/news-releases/cookie-rejection-less-than-1-percent-on-the-web-according-to-websidestory-82321002.html)

[17] According to Chris Babel of TRUSTe, only 0.00015% of the users who see the "Ad Choices" icon take advantage of the option.

III. Toward a Policy Framework

Economists typically see value-based pricing as a tool for expanding the size of the market by charging more to those willing to pay and less to those who are not. While this practice may seem unfair to those who pay a higher price, buyers often have tools for fighting back, especially in a competitive market. Moreover, anecdotal evidence suggests that while many sellers use the Internet to explore the demand curve, few make widespread use of personalized pricing, at least for the moment.

Nevertheless, the combination of big data and differential pricing does raise serious concerns. For example, big data may facilitate discrimination against protected groups, and when prices are not transparent, differential pricing could be conducive to fraud or scams that take advantage of unwary consumers. This final section of the report considers how big data and differential pricing relate to existing antidiscrimination and consumer protection laws that might address these issues.

Antidiscrimination

Big data naturally raises concerns among groups that have historically been victims of discrimination. Given hundreds of variables to choose from, it is easy to imagine that statistical models could be used to hide more explicit forms of discrimination by generating customer segments that are closely correlated with race, gender, ethnicity, or religion. Moreover, the term "price discrimination" may lead to concerns about economic injustice, even if the profit motive is different from, and in many cases fundamentally inconsistent with, the sort of prejudice that our antidiscrimination laws seek to prohibit.

To understand how big data is likely to affect historically disadvantaged groups, it helps to distinguish between disparate treatment and disparate impact. In the marketing context, disparate treatment occurs when a seller uses race, religion, or some other consumer characteristic as a proxy for demand. This practice can be efficient, producing gains for both seller and buyer, even if it is annoying (or worse) for members of a targeted group. The premise of big data is that when marketers have a wide variety of behavioral data to choose from, they will find imperfect proxies such as race or religion to be less useful. In other words, big data aims to reduce the rate of "false positive" cases that potentially make disparate treatment a problem.

Disparate impact occurs when some practice has an adverse impact on a protected group, even if the practice was not intended to be discriminatory. For high-stakes markets such as credit and employment, laws like the Civil Rights Act of 1964 and the Fair Credit Reporting Act (FCRA) are used to prevent disparate outcomes.[18] In principle, big data could lead to disparate impacts by providing sellers with more variables to choose from, some of which will be correlated with

[18] Many provisions of these laws have been extended to prohibit discrimination on the basis of age, disabilities, or genetic information.

membership in a protected class. However, big data also provides new tools for detecting problems, both before and perhaps after a discriminatory algorithm is used on real consumers. For example, it is often straightforward to conduct statistical tests for disparate impact by asking whether the prices generated by a particular algorithm are correlated with variables such as race, gender or ethnicity. Put differently, in markets where it is important to prevent disparate impact, big data can be used to enforce existing antidiscrimination laws more effectively, thereby obviating the need for broader restrictions on its use.

Finally, it is important to keep in mind that if historically disadvantaged groups are more price-sensitive than the average consumer, profit-maximizing differential pricing should work to their benefit. This argument does, however, come with two caveats. First, it assumes competition is held constant. Minority groups may pay higher prices because they are underserved, so that sellers who do business with them face less competition and can more easily raise prices. But in that case, policy should focus on encouraging competition rather than limiting differential pricing. Second, as described above, risk-based pricing generally favors less risky customers, as opposed to those who are more price sensitive. This suggests that policies to prevent inequitable application of big data should focus on risk-based pricing in high-stakes markets such as employment, insurance, or credit provision. While the antidiscrimination provisions of existing laws such as the FCRA and Civil Rights Act should apply to these settings, ongoing scrutiny is warranted given the rapid changes in both technology and business practices.

Consumer Protection

In a competitive market with transparent pricing, value-based pricing is unlikely to harm the average consumer, who can easily compare offers and switch sellers. However, when sellers obfuscate by bundling a low product price with costly warranties or shipping fees, use "bait and switch" tactics to attract customers with false promises, or bury important details in the small print of complex contracts, differential pricing can cross the line into fraudulent behavior. In such cases, Section 5 of the Federal Trade Commission Act generally provides the FTC with sufficient authority to prohibit "deceptive acts or practices."

Some consumer advocates suggest that we should go further and limit the use of personalized pricing to offline settings or require its disclosure to buyers.[19] Economic reasoning suggests that differential pricing, whether online or offline, can benefit both buyers and sellers, as described above. Thus, we should be cautious about proposals to regulate online pricing – particularly if we believe that online markets are particularly competitive. Proposals to require seller disclosure are less worrisome. In general, rules requiring disclosure that prices or search results may vary across users should be based on a comparison of the compliance costs and the expected benefits from increased transparency. However, in assessing the benefits, we should recognize that buyers have strong incentives to seek out a good price, and that they may come to expect personalized pricing in specific online settings, just as they do offline.

---

[19] "Websites that Charge Different Customers Different Prices: Is their 'price customization' illegal? Should it be?" Anita Ramasastry, *FindLaw,* June 20, 2005.

A final issue that is often conflated with differential pricing, but represents a valid concern in its own right, is privacy. While many questions about privacy rights fall outside the scope of this report, it is important to note that big data involves the aggregation, sale, and use of large amounts of personal information, often in ways that individual consumers know very little about.[20] Much of this activity facilitates personalized tracking and targeting, which create value by helping sellers better identify and serve buyers' needs. However, many users may not want some kinds of information, such as visits to health or finance-related web sites, to be tracked without their explicit consent. In some cases, individuals may also have a greater interest in the accuracy of their information than the companies handling it.

Concerns over privacy, data quality, and fairness are especially salient when considering risk-based pricing. In particular, big data may cause risk-based pricing strategies to target consumers based on factors outside their own control. The challenge for policy in this area will be to promote the application of big data where it can discourage excessive risk-taking and help solve adverse selection problems, while preventing unfair discrimination against consumers who have little control over newly-measureable risk factors.

One way to limit unfair or inaccurate applications of big data might be to give consumers greater control over their information. Data brokers claim that strong property rights over personal information could produce large transaction costs that would undermine valuable applications of big data. Economic theories also suggest that such property rights would not fully resolve the privacy problem. For example, adverse selection issues could re-emerge through voluntary disclosure of information.[21] Nevertheless, a property rights approach to privacy seems particularly appealing where big data leads to concerns about fairness in the application of risk-based pricing strategies, and information intermediaries may have insufficient incentives to ensure the accuracy of personal information.

In practice, various self-regulatory efforts, such as the Individual Reference Services Group, which lasted from 1997 to 2001, have failed to produce consistent rules or widely adopted norms regarding consumer access to or control over the information used in personalized marketing. Thus, in their recent reports on the activities of data brokers and information resellers, both the Federal Trade Commission and the Government Accountability Office have suggested a need to rethink existing frameworks for regulating consumer privacy and the acquisition and use of big data in the marketing context.

---

[20] "Getting to know you. Everything people do online is avidly followed by advertisers and third-party trackers." *The Economist*, September 13, 2014.

[21] Katz and Hermalin (2006) discuss other cases where property rights in personal data need not lead to efficient outcomes and provide a more general model of privacy as secrecy.

Conclusions

Economists have studied differential pricing for years, and while big data seems poised to revolutionize pricing in practice, it has not altered the underlying principles. In particular, value-based pricing generally benefits sellers who earn more profit and buyers who would otherwise be priced out of the market, at the expense of less price-sensitive customers who end up paying a higher price. Risk-based pricing favors the least costly customers, as opposed to the most price-sensitive but can produce similar market-expanding benefits in the presence of adverse selection and may also discourage excessive risk-taking in some settings, although it can raise serious fairness concerns in others. Competition, collective purchasing, and arbitrage can all limit sellers' ability to implement differential pricing.

How might big data change differential pricing in practice? In many ways, it remains too early to tell. Broadly speaking, big data seems likely to produce a shift from third-degree price discrimination based on broad demographic categories towards personalized pricing and individually targeted marketing campaigns. A review of the current practices suggests that sellers are now using big data and digital technology to explore consumer demand, to steer consumers towards particular products, to create targeted advertising and marketing offers, and in a more limited and experimental fashion, to set personalized prices. At the same time, buyers are making use of the Internet and the variety of choices and tools it provides to ensure that they get a good deal.

Concerns with differential pricing are often entangled with related concerns about competition or consumer privacy. In many cases, policy should be able to address these issues individually. In particular, where differential pricing harms a protected group or crosses the line into fraudulent behavior, our existing antidiscrimination, competition, and consumer protection laws provide a variety of tools to correct the problem. However, given the speed at which both the technology and business practices are evolving, commercial applications of big data deserve ongoing scrutiny, particularly where companies may be using sensitive information in ways that are not transparent to users and fall outside the boundaries of existing regulatory frameworks.

References

Armstrong, Mark. 2006. "Recent Developments in the Economics of Price Discrimination." *Advances in Economics and Econometrics: Theory and Applications: Ninth World Congress*.

Borgie, Lana. 2014. "Trends in Producer Prices between E-commerce and Brick-and-mortar Retail Trade Establishments." *Bureau of Labor Statistics* 3, no. 18. http://www.bls.gov/opub/btn/volume-3/trends-in-producer-prices-between-e-commerce-and-brick-and-mortar-retail-trade-establishments.htm.

Cackley, Alicia Puente. 2013. "Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace." September. *United States Government Accountability Office.*

Corts, Kenneth. 1998. "Third-Degree Pricing Discrimination in Oligopoly: All-Out Competition and Strategic Commitment" *The RAND Journal of Economics* 29, no. 2.

"Do Many Users Turn off Cookies?" Webmasters Stack Exchange. http://webmasters.stackexchange.com/questions/58210/do-many-users-turn-off-cookies.

Deneckere, Raymond J., and R. Preston McAfee. 1996. "Damaged Goods." *Journal of Economics & Management Strategy* 5, no. 2. http://www2.econ.iastate.edu/faculty/langinier/teaching2005/615-fall2005/paper-hom2/damagegoods.pdf.

Einav, Liran, Theresa Kuchler, Jonathan D. Levin, and Neel Sundaresan. 2011. "Learning from Seller Experiments in Online Markets." National Bureau of Economic Research Working Paper No. 17385. http://www.nber.org/papers/w17385.

Ellison, Glen and Sarah Fischer Ellison. 2009. "Search, Obfuscation and Price Elasticities on the Internet." *Econometrica* 77, no. 2.

Hannak, Aniko, Gary Soeller, David Lazer, Alan Mislove, and Christo Wilson. 2014. "Measuring Price Discrimination and Steering on E-commerce Web Sites." http://www.ccs.neu.edu/home/cbw/pdf/imc151-hannak.pdf.

Hermalin, Bejamin and Michael L. Katz. 2006. "Privacy, Property Rights and Efficiency: The Economics of Privacy as Secrecy." *Quantitative Marketing and Economics* 4, no. 3.

Hossain, Tanjim, and John Morgan. 2006. "...Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on EBay." *Advances in Economic Analysis & Policy* 6, no. 2. http://faculty.haas.berkeley.edu/rjmorgan/ebay.pdf.

Pew Research Center. 2014. "Public Perceptions of Privacy and Security in the Post-Snowden Era." November. http://www.pewinternet.org/2014/11/12/public-privacy-perceptions/

Podesta, John, Penny Pritzker, Ernest J. Moniz, John Holdren, and Jeffrey Zients. 2014. "Big Data: Seizing Opportunities, Preserving Values." *Executive Office of the President*. http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf.

Varian, Hal. 1997. "Versioning Information Goods." *Digital Information and Intellectual Property*. http://people.ischool.berkeley.edu/~hal/Papers/version.pdf.

Varian, Hal. 1985. "Price Discrimination and Social Welfare." *The American Economic Review* 75, no. 4. http://www.jstor.org/stable/1821366.

Ramirez, Edith, Julie Brill, Maureen K. Ohlhausen, Joshua D. Wright, and Terrell McSweeny. 2014. "Data Brokers: A Call for Transparency and Accountability." *Federal Trade Commission*. http://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

Shiller, Benjamin Reed. 2014. "First-Degree Price Discrimination Using Big Data." http://benjaminshiller.com/images/First_Degree_PD_Using_Big_Data_Apr_8,_2014.pdf.

Turow, Joseph, Lauren Feldman, and Kimberly Meltzer. 2005. "Open to Exploitation: America's Shoppers Online and Offline." *A Report from the Annenberg Public Policy Center of the University of Pennsylvania*. http://repository.upenn.edu/cgi/viewcontent.cgi?article=1035&context=asc_papers.

Government Accountability Office. 2013. “Information Resellers: Consumer Privacy Needs to Reflect Changes in Technology and the Marketplace.” http://www.gao.gov/assets/660/658151.pdf.