

Implications of Big Data for Competition Policy and Practice

The Milton Handler Lecture
May 16, 2018

Hon. Katherine B. Forrest, U.S. District Court, Southern District of New York

I am honored to be here this evening and to have the opportunity to discuss evolving informational market forces that fundamentally alter the theory and practice of competition policy and law. There is nothing novel about restrictions on inter-firm exchange of competitively sensitive information. Horizontally competing firms cannot, for instance, share customer or pricing information with impunity. This has always been so – and has long informed anti-trust policy and practice. Many of you spend significant time advising clients as to what information may or may not be shared. My talk tonight concerns a changing informational landscape giving rise to competitive concerns.

It is now clear that big data has fundamentally altered the scale and velocity of information acquisition and the analytic capabilities to manipulate it. Firms with vastly expanded informational access have an enhanced ability to use such access to create competitive efficiencies as well as use it for more questionable competitive interactions.

The breadth of information now available adds a critical new dimension to our analysis of such interactions: the tried and true frameworks of price and output are inadequate to present a

comprehensive understanding of competitive conditions and/or firm conduct. The competitive implications of big data are real—and may go unrecognized by firms having particularly deep access. Firms may perceive their access and ability to manipulate vast data sets as simply a fortunate development—the march of technological progress. And, without malice aforethought, such a firm may engage in conduct that runs afoul of basic principles of antitrust law.

Setting policy or advising clients in this emerging competitive environment requires a multi-dimensional approach—one that takes into consideration that informational access both allows for firm-specific economic efficiencies, but may also lead to conduct that can have the effect of manipulating market dynamics.

This evening, I intend to begin a process of exploring a layered analytical approach that considers how this growth in informational access needs to figure into antitrust considerations. My basic premises are that a firm's data and its algorithmic ability to analyze such data—including that which it has, but equally importantly, that which it can get—are themselves products, multi-faceted commodities, that exist independently of the firm's more traditional products and services. The

data to which a firm has or to which it can gain access may or may not derive from the manufacturing of its special widget or creation of service offerings.

A second premise is that today most firms consider all manner of data harvested from their own conduct or acquired from a third-party as, so to speak, born in wedlock, and thus necessarily lawful and unproblematic. That is, firms expect that data they possess may be put to any use; and that includes data that they are able to publicly acquire from firms trading in big data sets pertinent to one or more industries or populations.

Data itself has, thus, quietly become a competitive force within all firms, and its usage is capable of causing anything from ripples to waves to tsunamis in market conditions. You might think of it this way: when analyzing a firm's products and its lines of business, it is time to consider not only what comes off the manufacturing line, but the information—the large data sets—that may be generated, available, or used anywhere in or by the firm. Today, the digital ability to capture and process such information may have little to do with product characteristics, a firm's unit sales, or customer lists. Instead, a

harvestable data set might include whether customers in particular zip codes are sophisticated or unsophisticated purchasers, whether they engaged in meaningful comparison shopping and if so, with which competitors; and whether those competitors offer substitutable widgets to the same customers in the same geographic area, with the same or similar terms and conditions. On the one hand, information can provide an extraordinary opportunity for efficient firm conduct; on the other, it carries known and unrecognized risks: it may be used to disrupt the well-functioning competitive process.

I want to now (1) define what I mean by big data and discuss specific areas of competitive concern, (2) offer a new definition of consumer welfare that takes informational issues into consideration, (3) suggest redefinitions of product markets and market power, and (4) describe potential impacts on competitive effects and exclusionary conduct.

A. Defining big data for purposes of competitive impact

Conceptually and practically, big data is the digital capture of vast quantities of information capable of algorithmic manipulation. Such information may include details about any firm that interacts with

buyers, sellers, or supply chain participants in a digitally networked manner; it similarly includes any and all information about the digitally-networked conduct of a single person, household, or population segment. It is not only current information, but also historical information converted or maintained in digital and algorithmically accessible format. It can include domestic or international data specific to an industry or generalized to the economy. Critically, algorithmic manipulation allows such data sets to be accessed and queried; the days of data “snapshots” or sampling are coming to an end.

Competitive utility of such data has unimaginable breadth. It can render competitive or anticompetitive conduct easier and harder to detect; and, as much of its value derives from processes occurring within servers and as the result of algorithmic manipulations, it is a machine and not human-driven role. Big meetings in smoke-filled conference rooms are unnecessary and even unhelpful. Informational exchange and usage is rendered incredibly quick, automatic, and nearly self-executing.

Questions may be asked of any data set – and those questions may be the simple ones we imagine: does this industry or set of firms, this

population of consumers, this household, this person, acquire, want to acquire, or can he or she be made to acquire, X product? Will that firm or person pay a particular price based on prior purchasing patterns? Do those patterns rely on the customer base of a firm, or the demographic characteristics of consumers? The questions may also be more complex: is this population likely to want, need, or be receptive to a product or service no one has ever heard of? Is there any “must have” data for firms seeking to achieve a significant position in this area? Is exclusionary conduct relating to data access lawful?

Let me provide a few examples of the utility of acquired or harvested and analytically manipulated data sets. I’ll start with a few that are more obvious and then move to several of a newer vintage.

First, a firm’s unparalleled access to data enables targeted advertising at a level that allows a firm to know more about us than many of the people closest to us; this advertising has the potential to create demand through manipulated, perceived need.

In addition, a firm’s unparalleled access to data can further enable it to obtain an unerring first mover advantage in new product areas based on predictive modeling.

It allows a firm to achieve or maintain a competitive advantage, even dominance, by tying other firms or consumers to “sticky” experiences (think Facebook, Linked In, Amazon, Apple, Uber, but also other platforms).

It allows a firm to engage in price discrimination at a sophisticated level based on known and predictive buying patterns. Big data allows consumer-directed price discrimination differentiating between demographics as well as those exhibiting certain buying patterns, or between those with varying digital presences. In short, different prices can be presented simultaneously to different market participants.

It allows a firm to gain unparalleled insight into purchasing patterns and prices a buyer has paid or is likely to consider paying.

It allows a firm to create a market opportunity or enter a market early and dominate in areas in which that firm did not traditionally compete (thus rendering firms with data access into all manner of nascent competitors).

It allows a firm to observe, analyze, and act on incremental pricing behavior, in a way that is similar to that of flash securities traders.

It allows a firm too provide data on an exclusive basis to a firm that may acquire it simply to shelve it.

It allows a firm to engage in price fixing through “understandings” gleaned from digital knowledge heretofore unattainable. For instance, a firm may be technically able to observe numerous diverse price elements through machine-related processes and determine or predict prices charged by other market participants through algorithmic processing; conforming pricing under these circumstances could well be problematic.

B. Redefining consumer welfare

The potential economic utility of big data—both good and bad—pushes us to redefine consumer welfare. For tonight’s purposes, let me describe consumer welfare as a consumer’s ability to acquire goods or services based on market forces free from unreasonable external manipulation. Such an ability provides consumers with fair access to

available goods and services that are necessary to and/or enhance their quality of life.

Competition policy and law is premised on a view that maintaining robust and competitive conditions between firms, in which exclusionary behavior is discouraged and even penalized, and in which dominance is controlled or prevented when possible and carefully monitored when naturally occurring or nascent, enhances consumer welfare.

In an era in which data can heavily influence or even control firm conduct or consumer behavior, observed price and output effects are no longer the only factors by which to analyze competition and consumer welfare. Observed price alone may obscure an array of behind-the-scenes conduct that is increasingly sophisticated, enabling discrimination at granular levels, and that can be exchanged in non-traditional ways. Thus, using observed price and output effects to define the parameters of conduct that is welfare-enhancing or harmful no longer has the same utility it once did. An additional driver, how firms can and do acquire, collect, use, and trade in vast quantities of

information, must also be recognized as directly impacting on consumer welfare.

Tonight I offer a revised definition of consumer welfare that takes into account issues relating to the competition implications of big data. What I suggest is that we must understand consumer welfare and the health of the competitive process, as including freedom from unreasonable manipulation of captive data sets.

Algorithmic data manipulation that reduces free market decision making reduces overall welfare. To be clear, access to data is often welfare enhancing—we depend on it for well-functioning markets. However, since big data has potential manipulative effects, one must consider whether particular uses impact consumer welfare.

Let me give two additional examples of big data's potential impact on consumer welfare. The first is that use of data by a firm that has deep access, inherent network effects, and dominant market position (think Amazon, Facebook, or Linked In) may inhibit new entry. Consumer welfare may be harmed when data is harvested, analyzed, and used in a manner intended to or having the effect of precluding competitive choice or creating such networked attachment that

consumers no longer have a real ability to engage in freely-chosen purchasing behavior.

Another example of welfare-reducing conduct is, as I previewed above, granular and highly sophisticated price discrimination based on a known demographic or a willingness to share data in exchange for price effects. In this regard, consumers with a robust digital presence may have different pricing options than those without—leading to data-driven consumer “leave behind” or “jump ahead”.

Current literature touches on some of these issues – framing them differently in important ways—as incursions on consumer privacy. No doubt privacy issues carry impacts on consumers – many negative. But in terms of competition law, the question for tonight is how big data may manipulate firm and consumer interactions with the marketplace for goods and services.

Now, lest my remarks convey a solely negative picture, let me be clear that big data has welfare enhancing aspects. For instance, a positive effect could be identifying product characteristics consumers find most useful, or anticipating and avoiding supply chain bottlenecks that allow smooth maintenance of output levels and price equilibrium.

Of course there are also enormous health impacts that big data can have: identifying epidemiological issues, responses and the like.

Pharma companies surely find these of great interest.

C. Redefining product markets and market power

Once we redefine consumer welfare we must carry the implications of that definition through to conceptualizing what constitutes a competitive effect.

Traditionally, the existence or impact of competitive effects is often correlated to a firm's position in the market: does the firm have sufficient market power to create or enhance a welfare-reducing market condition?

In the world we now inhabit, when data sets allow for unseen but real manipulative impact on competition and consumer welfare, the concept of market power also needs redefinition. Is it still appropriate to define market power or dominance in relation to interchangeable products, when a firm's real competitive impact is based instead on how it manipulates data? Firms with the capacity to harvest and utilize big data now have two products—their nominal one (for example, the widgets they manufacture), and their data set and related processes for

algorithmic analysis. A competitor group therefore encompasses not only those engaged in making or providing substitutable goods or services, but those who have or possess similar data sets or analytic capacities. In this regard, the potential commoditization of a data set works as a proxy to define a competitive universe.

This definition means that what used to be a single-product firm now has another product line: its data and analytic capabilities. Thus, firms that might not traditionally be considered competitors may find themselves in competition with each other.

What does this mean? It means that defining markets based on manufactured products may be insufficient, and therefore that measuring competitive effects based on price and output may similarly be insufficient. An anticompetitive effect thus includes data manipulation that has a defined ability to unreasonably impact market behavior. My remarks are not suggesting abandonment of price or output analyses to measure competitive effects. But, in this informationally-driven world, price and output analyses are no longer sufficient measures to define market impact.

In theory as well as practical application, what constitutes economically efficient behavior—what is procompetitive—must also evolve. A firm’s decision to follow a path to making more and better widgets, expanding market reach, or improving quality, is only a piece of what will improve that firm’s overall competitive prospects and what will create market efficiencies. Modeling firm conduct that is categorized as economically rational or irrational should now include analyses of how a firm’s data has been effectively harvested and used.

D. Redefining potential impacts on competitive effects and exclusionary conduct

In our redesigned framework, individual or collective firm conduct that seeks to utilize data to reduce independent decision making, create unparalleled dominance, or preclude timely and effective entry, would be subject to scrutiny and potential redress. Exclusionary conduct that prevents wide access to certain data sets may, however, not be the real problem; and therefore more access may not provide the real solution. I do not view data itself as an essential facility; on the other hand, the algorithm through which it is run, and the mining techniques applied to it, may be.

With all of this said, non-exclusive access to algorithms is not necessarily the best solution. In this context, non-exclusivity may act to proliferate rather than reduce market manipulation.

E. Practical Issues

Let me turn now to certain practical questions: what, if any, implications does all of this have for policy and practitioners?

Critically, when regulators, policy makers, and advisors are analyzing firm behavior, the concept of what may harm or help ensure robust competition and best serve consumers, needs to expand. For example, we must ask:

First, in what ways can firms cause anticompetitive price effects through data manipulation?

Second, in what ways is price discrimination an acceptable or unacceptable market outcome of data manipulation?

Third, what are the competitive implications for firms in one line of business to share non-price, and non-output-related data sets?

Fourth, how do we include an analysis of informational access and manipulation in merger analysis? For instance, what role does it play in the merger of firms that previously would have been considered non-

horizontal competitors but that we now recognize as having complementary data sets and strong analytical capabilities?

Let me end with a few final practical points. Firms increasingly and appropriately recognize data as a crown jewel. Elevating data to such status informs strategic decision making in terms of acquisition, harvesting, and use. Regulators need to understand and think through the implications of this. The speed at which all of this is moving challenges legislative and regulatory processes. Separately, firms may want to consider self-policing in the form of codes of conduct. Such codes could, for instance, recognize, that data manipulation can negatively impact free market competition.

Conclusion

In conclusion, we are at an inflection point that requires fundamental alterations in legal theory. I have offered one way of considering the issues arising from unparalleled informational access and usage, but of course there are many ways to think about such issues. More important than the absolute correctness of any emerging views is that we begin the dialogue that recognizes the important changes are occurring around us.