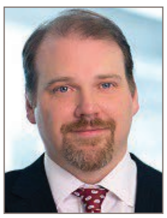
# Roundtable: Discussing the Big Picture on Big Data

***Editor's Note:*** *On November 2, 2018, five distinguished panelists sat down with Kevin Christensen and Kevin Yingling, Editors on* The Antitrust Source*, to discuss the competitive significance of data and pricing algorithms. The panelists provided their views on such topics as how to define the term "big data," why the topic of big data is receiving heightened attention in antitrust right now, how pricing algorithms trained on big data should be evaluated under the antitrust laws, whether data can form a barrier to entry, and how the collection of big data affects our thinking on the role of privacy in competition policy. The Roundtable was edited for publication.*


MODERATORS



**Kevin Christensen**
*Director and Founder, Christensen Economics LLC*

**Kevin Yingling**
*Senior Competition Counsel, Google*

PANELISTS

**James Cooper**
*Deputy Director for Economic Analysis, FTC Bureau of Consumer Protection*[1]

**Ai Deng**
*Lecturer at Johns Hopkins University and a Principal at Bates White Economic Consulting*

**Terrell McSweeny**
*Partner,Covington & Burling and former FTC Commissioner*

**Barry Nigro**
*Deputy Assistant Attorney General, Antitrust Division, U.S. Department of Justice*

**Darren Tucker**
*Darren Tucker, Co-head, Antitrust Group, Vinson & Elkins*


**KEVIN YINGLING:** Thank you for talking with us about big data issues. All of you have either written or spoken about big data, so we are sure to have a very insightful discussion during this roundtable, and we are very interested to hear your thoughts on these topics.

Let's start off with definitions so we have some clarity in what we're talking about. In some of the academic literature, you see big data described as having four or five Vs: volume, velocity, variety, value and, sometimes, veracity.

Is that a sensible way to talk about big data? How do you think about "big data" when using that term?

**JAMES COOPER:** Yes, I think that's right, the three to five Vs define it. Sometimes you also hear that a distinguishing feature of big data is that you're looking at nearly the entire population, rather than merely a sample of the population.

I think that aside from the quantity, one of the things that distinguishes, say, working with big data from a normal data set is that it's kind of rough. It's not well structured, and often, you're looking for associations between variables that emerge from the data, not necessarily for casual relationships.

[1] The views expressed here are solely those of Dr. Cooper and do not purport to represent those of the Federal Trade Commission or any individual commissioner.

When you're working with large data sets, it gives you the benefit of a lot of statistical power—the ability to detect small relationships that you can't with a smaller data set. But the problem is you also can find a lot of spurious correlations. You look at any large data set, and if you play with it long enough, just by random chance you're going to find associations.

Because of these issues, one must be able to apply some expertise when working with big data in order to derive the benefits. You have to know how to ask the right questions of the data, and interpret their answers.

*When you're working with large data sets, it gives you the benefit of a lot of statistical power—the ability to detect small relationships that you can't with a smaller data set. But the problem is you also can find a lot of spurious correlations.*

*—James Cooper*

**BARRY NIGRO:** One of my concerns with the term "big data" is that it's a label, and sometimes is used as a conclusion to indicate that what you're talking about is an entry barrier or something that facilitates monopolization and market power. I think when you label things and talk about them only in a general sense, there's a danger that you will misunderstand the competitive significance.

On the one hand, it's useful to put some parameters around what is meant by big data; on the other hand, I think it's always important to keep in mind specifically what you're talking about and the competitive significance of the asset on which you're focused. It's more productive to have a detailed discussion about the precise asset and its role in the competitive dynamics of the industry.

**TERRELL McSWEENY:** Big data is a term that gets thrown around to describe a lot of different features of data and the role it plays in the digital economy. In antitrust, we care a lot about specifics, and I think that's appropriate.

We should be very mindful of what the competitive significance of data is when examining it through an antitrust lens. Very often, when people talk about big data, they are generalizing about data and its role in the digital economy.

**AI DENG:** I do think the definition of big data can be very helpful as a guide. When we look at the competitive effects, one thing that's very important is to understand the role of big data in a particular industry. I often see people being unclear about whether big data is the output of an industry or a type of input that is used to improve their products and services. And I think that will be the first question that I would ask, if I am looking at competitive effects.

And if the big data turns out to be an industry output, it will be relatively straightforward to analyze. A lot of standard tools would apply. For example, we can ask how costly it is to procure, process, and host this data. But when the big data is used as an input, I think things are more subtle, and that is something we can discuss further today.

**DARREN TUCKER:** I think the four or five V definition is fine when you're making general comments about data or the evolution of technology. But when you're looking at the role of data in any particular industry or in a particular investigation or market study, it's very important to move beyond a generic definition of data and look at the nature of the data at issue and how it is actually used.

For example, is it used as an input or is it the final product? If it's used as an input, how important is it as an input into the final product? How widely available is the data to other potential entrants or companies in the industry? What's the value or difficulty in obtaining historical data? Historical data has a lot of value for certain industries. In other cases, historical data has very little, if any, value.

We should be careful to avoid some assumptions that are sometimes made in this area. For example, the fact that an incumbent may use a lot of data is not necessarily an indication as to the value of data to some other company or that data is a barrier to entry. Likewise, you often see an assumption that big data is used in a lot in digital online markets, but less so in traditional sectors. That's probably not a good assumption. For example, some kinds of online advertising use virtually no personal data, whereas some traditional industries like retailers or insurance rely very heavily on big data.

In short, it's important to move beyond assumptions about data and look carefully at its actual role and importance competitively for incumbents and for potential entrants.

*The current focus on big data and algorithm issues in the antitrust world is somewhat surprising to me because the antitrust agencies have a lot of experience in these areas, going back decades. These are not novel issues for the agencies.*

*—Darren Tucker*

**KEVIN YINGLING:** With the definitional work out of the way, let's turn to why there is so much interest in big data now. The FTC is having hearings at American University in November that are going to be covering the topic, and there have been a lot of published works and conferences addressing big data issues. Why is the topic getting so much attention right now?

**BARRY NIGRO:** One of the reasons it's getting a lot of attention is because people don't completely understand it. There's concern about how it's used and how it's acquired.

When I think about things that aren't well understood, I sometimes translate them into something with which I am familiar. If you think about the type of data that a platform may acquire about customer preferences, that may be analogous to—in the old days—company customer files that record preferences and customer requirements. It's hard for me to imagine, for antitrust reasons, the government or a court ordering that customer files be shared in order to give competitors equal access.

This goes back to the point I made earlier, when you're talking about data, the context is important; understanding where the data resides in the competitive ecosystem and how it's acquired and being used is critical. To some extent, there is a lack of familiarity and understanding of that now, which is why it's a topic of conversation.

In addition, there's the obvious point that data has become more prominent in the economy with everything being digitized and the ability to store more and more information at a lower cost.

**JAMES COOPER:** I think that data is such a hot topic in consumer protection and antitrust because it's something that the large platforms have in common: they collect and use enormous amounts of consumer data. And their use of data has been front and center in some of the scrutiny they've come under recently on both the competition and consumer protection sides.

These discussions are one of the reasons the FTC is doing the series of hearings to look at the economy in the 21st century, the competition and consumer protection issues, and many of them revolve around data, not coincidentally because the largest companies in the economy tend to live on data.

**DARREN TUCKER:** The current focus on big data and algorithm issues in the antitrust world is somewhat surprising to me because the antitrust agencies have a lot of experience in these areas, going back decades. These are not novel issues for the agencies. The FTC and DOJ have been bringing what you might call big data cases at least since the '90s, if not earlier. And there's been quite a number of them. The same thing is true at the European Commission—the Google/DoubleClick case, for example.[2] The same is true in the algorithm space. There have been court cases

[2] Case M.4731, Google/DoubleClick (Mar. 11, 2008), http://ec.europa.eu/competition/mergers/cases/decisions/m4731_20080311_20682_en.pdf.

looking at potential Sherman Act Section 1 claims based on pricing algorithms in concentrated industries going back at least a decade.

Likewise, arguments that we should incorporate privacy or other non-competition considerations into antitrust are nothing new. The Supreme Court addressed this issue back in 1931 in *FTC v. Raladam Co.*[3]

So, it's a bit surprising to me that there's debate on these issues now. It would have made more sense to have this debate maybe 20 or 25 years ago when antitrust concerns about these issues would have been more significant. For example, big data analysis at that time was largely limited to governments and large businesses because the cost of collection, storage, and processing was dramatically higher. Same thing with algorithms. Smaller companies can offer much more sophisticated pricing today than they could 10 or 20 years ago, which can be disruptive to efforts of collusion.

**TERRELL McSWEENY:** I agree that big data is not actually a new thing in markets. Antitrust enforcers have confronted the competitive significance of data for a very long time. What is, arguably, different about data today is the Vs point, especially the volume and velocity of it.

Certainly, it's not by accident that people call the increasing digitization of previously analog businesses the fourth industrial revolution, because technology is changing many industries. Data is increasingly significant in a host of new ways in many areas.

So, in that sense, I think it's completely appropriate for competition enforcers around the world to engage in a dialogue about whether competition laws and frameworks are keeping pace with the technological change that is occurring in the marketplace. Here in the United States, in particular, that examination is coming at a time when there is also a broader debate over the role and purpose of antitrust law. I think the reason this conversation is happening among antitrust enforcers now is due to both of those things—the rapid evolution of technology and the largely political debate over the scope of antitrust law.

An important feature of our antitrust framework is that enforcers, civil society, and the bar continually engage in dialogue and examination of it. A strength of our framework is that it is adaptable and flexible, and ought to be able to evolve along with the dynamic marketplace.

**AI DENG:** As an economist and someone who deals with data every day, I always think about data, along with the ability to analyze it.

A lot of the interest is due to the supercomputing power that we have today, the revolution in machine learning, artificial intelligence. It may appear that all of a sudden with the help of big data, with the computational power, with the advances in the artificial intelligence and machine learning community, machines can do a lot more. And I don't think that's something we have seen in recent history. It is natural for people outside of the research circle to feel excited and also uncertain about what machines can do. And I think that uncertainty contributes to some of the interest as well.

**BARRY NIGRO:** I want to follow up on a point Darren was making, and that is that the availability of data and our ability to analyze it have grown. That raises a question whether the competitive significance of data has increased or decreased since the '90s. In other words, if you think about vol-

[3] 283 U.S. 643 (1931).

ume, one of the Vs, how many data points do you need to predict an outcome? At some point the marginal value of an additional data point declines rapidly; you don't need more than a handful of data points to predict with a fair degree of confidence a particular outcome. It's not clear to me, and this is a question of fact on which I don't have a view, whether having a large volume of data makes a material difference when it comes to a firm's ability to compete.

**AI DENG:** That's a good question. And it goes back to why I think the definition of big data can be really helpful. We can look at the four dimensions, and we would ask which Vs are the relevant features, because they're not always all relevant.

So, taking volume as an example, whether it is relevant is application- and industry-specific. For example, think about the standard machine learning tasks, say, image recognition, where you feed the machine a picture and then the machine will tell you what this picture is a cat or a dog. Typically, that kind of application would require a lot of data. Another example would be machine translation. You speak to the machine in one language and the machine would translate it into another.

If you look at other industries, let's say online advertising, my observation is that they are more interested in real-time data. So it's not necessarily the volume per se that matters. It's about how timely the data are.

I do think that you ought to look at it as a fact-based analysis. And there's not a single rule there.

**DARREN TUCKER:** For machine learning, labeled data is what's really valuable, not random data. Random data can still be useful, but labeled data is much, much more useful, and that typically involves human time to generate.

To borrow Ai's example, you could develop a machine learning program to identify cats and dogs in pictures. You will need to give it a series of pictures of different kinds of animals and objects to learn from. But first, some human has to go through and label which objects are the cats and dogs and which aren't so the program can learn from these examples.

There are a lot of ways to get labeled test data for machine learning applications. Sometimes there's publicly available data that companies can use, you can purchase it, or you can pay people to label data that you have.

As with other types of big data applications, machine learning often exhibits diminishing returns from the use of additional data. Also, the problems we are trying to solve using machine learning typically can also be addressed with techniques that do not involve machine learning. This reinforces the point that the need for data may vary considerably for different companies trying to solve the same problem.

There are also efforts to solve problems with machine learning using little or no data. That's called reinforcement learning. Machines essentially train themselves. A good example of that is Google's DeepMind team, which developed two different programs to play the Chinese game, Go. One of the programs learned from seeing how humans actually play. Another program learned just by playing itself over and over. And it turned out that the machine that was self-taught was the better player.

You're seeing competition agencies start to confront some of these issues. For example, in the Microsoft/LinkedIn decision,[4] the European Commission looked at what kind of data was needed

[4] Case M.8124—Microsoft/LinkedIn, Comm'n Decision (Dec. 6, 2016), http://ec.europa.eu/competition/mergers/cases/decisions/m8124_1349_5.pdf.

for CRM [customer-relationship management] machine learning. The Commission found that the data that LinkedIn had, while very valuable, was not essential and that other CRM suppliers could obtain relevant data from alternative sources.

**TERRELL McSWEENY:** If I could just pick up on a point that Darren is making, an essential feature of data, even in very sophisticated use cases like for training machine learning systems and increasingly autonomous systems, is that it is often readily available.

Data can often be non-rivalrous, which means it can be used by multiple companies at the same time, it's publicly available, and relatively easy to gather. There are not necessarily a ton of barriers to the collection of the data that's required to train these systems.

That said, I think we are seeing antitrust enforcers flag the question of access as an area of inquiry. What I hope enforcers continue to strive for here is a very factual understanding of the competitive significance of data so that they can reach the right answers to those questions.

*Data can often be non-rivalrous, which means it can be used by multiple companies at the same time, it's publicly available, and relatively easy to gather.*

*—Terrell McSweeny*

**JAMES COOPER:** Jumping off what Barry and Ai said, I think it really is important to point out that it's not so much the size of the data set, it's what you do with it. Empirical work in economics has gotten much better over the past couple of decades, partly due to more data and more computing power, but also—and probably more importantly—because of clever research design that gets us closer to randomized control experiments. Having a good design is what allows one to draw causal inferences, which are so important in policy making.

To the point about big data and its competitive implications, you want to have access to good data, large amounts of data that are helpful in finding relationships and honing machine learning algorithms. But if you don't have the right team in place and are not asking the right questions, you're not going get useful answers.

For example, Hal Varian, Google's Chief Economist, wrote in *Big Data, New Tricks for Econometrics*, that he has found a random sample of 0.1 percent works for analysis of business data. Having a billion data points but not having the team or the ability to figure out how to do the right test to get at the right answer is pretty useless. The really clever team with the small amount of data is probably more likely to lead their firm in the right direction.

**KEVIN YINGLING:** Let's follow up on a topic that a number of you have already touched on, which is the issue of algorithmic collusion. Big data can be helpful in training these algorithms, and some commentators have suggested a concern about pricing algorithms and their ability to facilitate coordination or collusion.

Are those concerns legitimate? Even if there is no human coordination, do pricing algorithms using big data present a threat to competition?

**DARREN TUCKER:** The increasing use of pricing algorithms across industries and providers should be seen by antitrust authorities as a beneficial development, not as something that raises alarm bells. Pricing algorithms allow very rapid responses to competitive conditions and competitor actions. In other words, pricing algorithms facilitate efficiency in the marketplace, which we generally think of as procompetitive.

To turn to your question, a good starting point for assessing the potential for pricing algorithms to facilitate coordination is to look at what the empirical data show. Is there evidence that algorithms have actually facilitated collusion? Sophisticated pricing algorithms have been used for a few decades, at least by some firms. So, if we've had problems with algorithms facilitating either tacit or explicit collusion, we should have some evidence by this point. And we don't really have empirical evidence of those kinds of effects.

Now, there have been a handful of cases—for example, the DOJ and CMA [Competition and Markets Authority] brought the case against two poster sellers on Amazon Marketplace. But that was a fairly run-of-the-mill collusion case, where the parties got together as they normally do, agreeing not to compete on price, and the mechanism of collusion was the algorithm.

Likewise the Russian Federal Antimonopoly Service brought a case against the Russian subsidiary of a Korean manufacturer, involving vertical price fixing using an algorithm.[5] But again, the algorithm was just used to monitor retail pricing. Once the manufacturer saw the retailer deviating from its recommended price, the manufacturer just picked up the phone or sent an email saying, "Hey, raise your price."

So, we don't really see a lot of evidence of algorithms facilitating coordination in the way that some have hypothesized.

Another way to approach the question is to consider how the adoption of pricing algorithms may affect the potential for coordination, looking at the usual factors. Complex pricing algorithms tend to reduce market transparency, which will tend to reduce the ability to coordinate. The reason is that pricing algorithms allow much more dynamic, differentiated, rapid pricing responses, which rivals are not necessarily going to see. A sophisticated pricing algorithm might allow a company to change prices hourly and have personalized pricing and discounts, which will be very hard for rivals to monitor.

There was an article in the *Wall Street Journal* about how most retailers, airlines, hotels, and financial companies have all developed these things called CLVs, which are customer lifetime value scores.[6] These are proprietary scores each of these companies has developed using algorithms to measure the value of retaining a particular customer. Based on the score, we all get different prices, freebees, upgrades, and customer service—all of which is virtually impossible, if not impossible, for rivals to see.

In addition to decreased market transparency, the use of pricing algorithms may facilitate maverick behavior and reduce entry barriers due to the ability to price more accurately to particular customer segments. Pricing algorithms may also facilitate greater cost asymmetry and product differentiation, which makes coordination more difficult.

We are also seeing regulators using algorithms to identify problematic behavior. The FTC for many years has been monitoring wholesale and retail petroleum prices, using algorithms to identify unusual movement in the marketplace that might trigger closer attention. Similarly, the Brazilian and Korean competition agencies are monitoring specific industries for behavior that may look curious. I would not be surprised to see customers in some industries engaging in similar efforts to monitor their suppliers' pricing through algorithms.

**JAMES COOPER:** There are some interesting papers out there that posit various models or hypotheticals of how algorithmic pricing could lead to explicit or tacit collusion. But it's also important to remember that collusion at this scale is a really complex game-theoretic problem.

---

[5] FAS decision of 2 March 2018, No. 1-11-18/00-22-17: http://solutions.fas.gov.ru/ca/upravlenie-po-borbe-s-kartelyami/ats-14552-18 (in Russian).

[6] Khadeeja Safdar, *On Hold for 45 Minutes? It Might Be Your Secret Customer Score*, WALL ST. J., Nov. 1, 2018, https://www.wsj.com/articles/on-hold-for-45-minutes-it-might-be-your-secret-customer-score-1541084656? ("Your score can determine the prices you pay, the products and ads you see and the perks you receive.").

Computing power is certainty available, but I think the jury is still out concerning whether, as a practical matter, it would be feasible to set up algorithms that would be able to sort signals from noise and figure out what the other algorithms are saying and coordinate prices in a way that is harmful to consumers.

I would be careful moving down a road that would apply the antitrust laws to a situation where prices that are set by algorithms appear to be moving together absent some evidence of an agreement. In *Ethyl*[7] and *Boise Cascade*,[8] courts rejected the FTC's attempt to stretch Section 5 to cover the unilateral adoption of pricing practices, such as base point pricing or MFNs, under the theory that they facilitated tacit collusion in concentrated industries.

I would also echo some of the things that Darren talked about. Anything that reduces transaction costs is typically good for consumers. If you've got an algorithm that can make a pricing decision in a microsecond, which otherwise would take a marketing department a week or even a day to figure out, that's probably a good thing.

So I think we should have a presumption that algorithmic pricing is likely to reduce transaction costs and probably increase competition. Now, again, this is a very new area and one worthy of study. Indeed, this topic will be part of an upcoming FTC hearing at Howard University. But I think we shouldn't start off with the presumption that algorithmic pricing is likely to be harmful to competition based on conjectures. I'd rather start from the proposition that algorithm pricing is likely to be a good thing until evidence moves us off those priors.

*Based on this AI research, I would say there is already theoretical or experimental evidence that algorithmic tacit collusion could happen.*

*—Ai Deng*

**AI DENG:** Darren did a good job summarizing some of the issues related to algorithms. And it's true that we haven't seen a real case where you have autonomously colluding robots.

At the same time, I also think it's helpful—and this is something that I have been doing over the past two years—to look into the AI literature and figure out what the AI researchers have been doing in terms of developing algorithms that could potentially collude autonomously. What's interesting is that there is very active research going on in the machine learning and AI field, where algorithms that better coordinate and cooperate with opponents are being developed. In an article in the Fall issue of *Antitrust* magazine,[9] I summarize some of the papers, all of which came out in the past two years. Based on this AI research, I would say there is already theoretical or experimental evidence that algorithmic tacit collusion could happen. To give you one example, a group of AI researchers has recently developed an algorithm that is akin to a hub-and-spoke type algorithm. Their AI agents include a planning agent, so it's like the central hub. And this planning agent basically hands out punishments and rewards to make sure that the other agents learn to cooperate.

But there are several reasons I'm not overly concerned at this point. For one, there are many technical obstacles. I think James mentioned some—just the complexity of figuring out how to interact with others, how to elicit cooperation, presumably without explicit communication, without necessarily knowing other algorithm's intent and behavior, and all that. As a result, the existing experimental studies are still based on very simple frameworks, on a very limited set of possible strategies that agents can take. And I laid out many more reasons in my article.

**TERRELL McSWEENY:** If I could just follow up on that point. I agree that at this point, it's a highly the-

7 E.I. du Pont de Nemours & Co v. FTC (Ethyl), 729 F2d 128 (2d Cir. 1984).

8 Boise Cascade Corp. v. FTC, 637 F2d 573 (9th Cir. 1980).

9 Ai Deng, *What Do We Know About Algorithmic Tacit Collusion?*, ANTITRUST, Fall 2018, at 88.

oretical problem—a bit like sci-fi antitrust, if you will. But the technology is evolving very rapidly and, as we've been discussing, one of the transformative aspects is the velocity at which systems can change.

It's appropriate for competition enforcers to be on top of the literature in this area and for people who are in the antitrust bar to be thinking about it and to be helping clients navigate some of the potential challenges. For example, we know to be cautious about hub-and-spoke structures, and I think we understand much of the type of conduct that may be problematic. I think algorithms used for highly personalized pricing are probably unlikely to facilitate tacit coordination because of how challenging cheating might be in that environment, for example.

But in thinking about antitrust issues, it may be appropriate to consider what the long-term objectives are for a more autonomous system. If its objectives are very short term, then it will likely maximize price. But if it's very long term, and it's a very sophisticated system, then maybe the possibility is greater that it could edge towards coordination.

These are issues that competition lawyers should have on their radar. And, in addition to some of the research that Ai has just mentioned, some researchers have found it at least probable that algorithms in certain situations may end up engaging in coordinated behavior.

It's great to have a wealth of different points of view out there, and there's no agreement in the literature right now about it whatsoever, but I think it's an important area for everybody to have on their screen.

**BARRY NIGRO:** I think the facts and the context matter. Looking at algorithms generally, I don't see why they're inherently anticompetitive, especially to the extent they facilitate more efficient pricing. If the structure of the market is such that it's competitive, it's not obvious why an algorithm would cause competitors to engage in a practice of raising price rather than continuing to compete aggressively.

Now, it could be that the use of algorithms may or may not, depending on the circumstances, facilitate tacit coordination in a market where consolidation is increasing rather than decreasing. Imagine that a vendor has a pricing algorithm, and they promote it on the basis that it will increase revenue by 10 percent: if you use our pricing algorithm, and by the way, in your local market you have three competitors, each of which has purchased the software so you will benefit from having it as well, and all of you will be able to raise revenue and be able to raise profits as a result of using the same pricing software. That raises some interesting questions as to whether it is a good thing, a bad thing, legal, or illegal. As when we were talking about big data, it's hard to generalize; you need to think about the context in which the question is arising in order to properly understand and evaluate it.

**DARREN TUCKER:** Barry's interesting hypothetical resembles a hub-and-spoke type of conspiracy theory. We have an existing framework to evaluate those types of claims, as we do for other potential algorithm-based concerns.

Beyond the merger context, I don't think the Sherman Act can or should reach tacit collusion, whether it's facilitated through algorithms or otherwise. As a general matter, parallel adoption of pricing algorithms is going to be outside the reach of the antitrust laws, even if that leads to interdependent pricing.

Even in markets with interdependent pricing, use of algorithms should, in my view, defeat an inference of conspiracy under *Matsushita*,[10] even in the presence of plus factors. If your pricing

[10] Matsushita Elec. Indus. Co. v. Zenith Radio Corp., 475 U.S. 574 (1986).

is based on an algorithm, that shows pretty definitively that you are in fact, pricing independently, even if you wind up with pricing that looks very similar to your competitors. A case on point is the LTL shipping services antitrust litigation from about a decade ago,[11] which involved use of identical software by two companies in the industry to determine fuel price surcharges, a highly concentrated industry, and parallel behavior. The court dismissed the case because there was no evidence of conspiracy; having two separate algorithms that reached the same outcome was insufficient.

**JAMES COOPER:** Yes, just reacting a little bit to what Darren and Barry said, I certainly would agree that antitrust should stay away from just conscious parallelism without an agreement. But going back to Barry's hypothetical, I think this is why antitrust is really up for the job of dealing with algorithmic pricing.

You could imagine a situation where there is a software producer who says: "Hey, I've got this algorithm, and I've fixed the problem of coordination. Just plug it in and you will be colluding with your competitors. Your profits will go up!" But this scenario isn't really that much different than say DOJ's *Apple*[12] case or the *Toys "R" Us*[13] case the FTC brought under a hub-and-spoke theory. You would probably develop the same sort of evidence where you have the software provider giving assurances that everyone is on board.

And I think that the important thing here again is that you can find an agreement under those sets of facts. So, to reduce the risk of false positives, you want to be careful about going down a road where now that we have algorithms, we don't need to worry so much about showing agreement.

**AI DENG:** I made the same point in the forthcoming *Antitrust* magazine article I mentioned earlier. In terms of what antitrust agencies and even private parties can do, there is a lot we can learn from the current literature. Keep in mind that, as I just said, you do see evidence in the technical literature that algorithms are being developed to tacitly cooperate. But at the same time, developing such algorithms turns out to be very challenging, for many reasons.

One lesson we can draw from that is the following. Suppose you have some rogue developers who are trying to develop an algorithm to elicit tacit coordination. The chances are that there is a paper trail, so that even without technical expertise, it may still be possible to for us to uncover a lot of the evidence. It is common and in fact critical in any research to document research design and findings. And such documents from the algorithm developers, I think, are going to be extremely helpful in discovery or an investigation. You could potentially subpoena such documents and see exactly what they were trying to achieve when they were designing the algorithms.

Another type of document I also find relevant is the one that others just mentioned, the marketing materials. If you just developed a collusive algorithm and you market it to your customers by saying, "Look, if you buy this, you're going to raise your profits because my algorithm learns to cooperate with your competitors"—again, you don't need technical expertise to uncover and interpret such documents.

It's also worth emphasizing that cartels come in different shapes and forms. In a lot of the industries, some collusive agreements, whether they are among humans or machines, are going to

---

[11] *In re* LTL Shipping Servs. Antitrust Litig., No. 1:08-MD-01895-WSD, 2009 WL 323219 (N.D. Ga. Jan. 28, 2009).

[12] United States v. Apple, Inc., 791 F.3d 290 (2d Cir. 2015).

[13] Toys "R" Us, Inc. v. FTC, 221 F.3d 928 (7th Cir. 2000).

manifest themselves through what we call plus factors and super-plus factors. For example, let's say we have an industry that relies heavily on sales representatives. A cartel trying to either fix prices or allocate markets may need to instruct the sales reps to maintain pricing and stop going after others' business. This change is something one could observe ex post. And it does not matter whether the instruction comes from a machine or a human. I argued in the article that firms may have a hard time placing the blame on a machine in that case.

**KEVIN CHRISTENSEN:** We came close to an answer earlier but let me pose the question more directly: Can big data be a barrier to entry?

**TERRELL McSWEENY:** I think that we discussed, earlier in the conversation, how important facts are to that assessment. When it comes to antitrust law, facts and specificity matter. And the fact is that antitrust enforcers have been able to assess data as a barrier to entry when it is in fact a barrier to entry. I'm thinking here of cases like Nielsen-Arbitron[14] and Reed Elsevier-ChoicePoint.[15] We've seen both DOJ and the FTC accurately assess the competitive significance of data and make the right call in several cases in the last decade. We've also seen them accurately assess—as in the Microsoft/LinkedIn transaction—when data, even very large amounts of it, is not a barrier to entry.

**DARREN TUCKER:** I would agree with Terrell's remarks. I do think that a lot of the debate in this area is on a somewhat different topic though, which is the claim that big data leads to a feedback loop. The argument goes that the collection of data will allow incumbents to improve their products, which attracts more people, particularly in markets with network effects, which allows them to collect even more data and so on and so on. The end result is that entry becomes extraordinarily difficult, if not impossible, because of the incumbent's permanent data advantage. I think there are a lot of problems with that theory, and I don't think that the agencies have bought into this theory. But you do see a vigorous debate in the literature on this point.

There are a number of reasons to be skeptical of this feedback loop theory. For one thing, data, by itself, doesn't guarantee a good quality product, as we discussed before. There are lots of other inputs into a good product. For an online service, for example, there is engineering talent, there's capital, there's responsiveness, there's an attractive user interface, and probably most important of all, just having a good idea.

Another reason to be skeptical of the feedback loop theory is that little, if any, data may be needed to enter and gain scale. Even when data is necessary, you can typically purchase it. There are hundreds of data brokers selling a wide range of personal and other data.

End users multihome. People typically use multiple search engines, multiple social networks, and multiple online retailers. Producers multihome as well. For example, a website may use a dozen or more different advertising and analytic services observing a user's browsing session on the website. Both user and producer multihoming facilitates getting data to potential entrants or smaller competitors.

Platforms are typically differentiated. We typically don't see like-for-like competition in the digital space, which means that new entrants can rely on differentiated data, to the extent that there's even a need for data.

---

14 Statement of the Fed. Trade Comm'n, In the Matter of Nielsen Holdings N.V. and Arbitron Inc., FTC File No. 131-0058 (Sept. 20, 2013), https://www.ftc.gov/system/files/documents/cases/140228nielsenholdingstatement.pdf.

15 Analysis of Agreement Containing Consent Order to Aid Public Comment, In the Matter of Reed Elsevier and ChoicePoint, FTC File No. 081-0133 (Sept. 16, 2008), https://www.ftc.gov/sites/default/files/documents/cases/2008/09/080916reedelseviercpanal.pdf.

As Barry and Ai mentioned earlier, there are diminishing returns to data, so even if an incumbent has a lot of data, that doesn't necessarily mean that it has much of an advantage over a new entrant.

We typically see congestion in two-sided markets where a lot of the concerns about big data arise. So, where you have platforms with lots of buyers and sellers for example, some sellers will want to go to a new platform, just so they can get noticed more easily, rather than fight for attention in a very crowded platform.

As the Supreme Court pointed out in its recent *Amex* decision,[16] some markets with network effects don't behave all that differently from regular markets. They may, for example, involve one-way indirect network effects, in which case, there's not a strong feedback loop.

And finally and perhaps most importantly, the empirical evidence doesn't support the feedback loop theory. Instead, we keep seeing, over and over, the big online incumbents get overthrown by some clever new company. We've seen that with social networks, search engines, dating sites, online retailers, transportation services, and so forth.

**JAMES COOPER:** I would just echo much of what everyone else has said. These questions are very fact specific. Data are non-rivalrous. But we can solve the public good problem if we make it exclusive or proprietary.

I would want to caution against punishing success. Network effects exist because you've got a successful product, so you draw lots of people to your platform and the more people, the more data, the better your platform becomes. But once you start saying, "You've become so successful and your data are so good that we are going to force you to share it," then we have to worry about dynamic incentives. So when we think of data as a barrier to entry we have to be careful to balance static and dynamic competition concerns.

The FTC will be exploring some of these interesting and important questions at American University as part of its Competition and Consumer Protection hearings.

**BARRY NIGRO:** I agree with James that we don't want to punish success. More specifically, where a firm develops data that's valuable and gives it a competitive advantage, requiring that firm to then have to share the data because it's helped the firm succeed would undermine the incentive to invest in new assets to compete for the market. In the short run, it would potentially create more competition within the market, but the dynamic competition that James referenced would be muted or potentially lost if investors taking on risk knew that once they succeeded, they would have to turn around and share the origin of that success with firms that didn't take on that risk.

I don't think requiring firms to share their data in that context is a costless exercise. In contrast, if you're looking at a merger, where a firm is acquiring data that is a critical input to competitors, then we have our usual tools, which we have used for many years, to evaluate that acquisition and assess whether a remedy is required. In that case, it is just another input. I don't think there's anything different about it from other types of critical inputs. The antitrust tools that we have are well suited to undertaking that analysis.

**AI DENG:** We talked about the definition of big data and whether big data is a barrier to entry. Sometimes I feel we also should be clear about what an entry barrier really is. If you look at the economic literature, there are actually a dozen different definitions of what an entry barrier is. It's

---

16 Ohio v. Am. Express Co., 138 S. Ct. 2274 (2018).

my observation that, in practice, when most of us talk about entry barriers, we have in mind some types of costs, that is, how costly it is to enter or how much time it takes to enter. The two are very often highly correlated. Very often it's both time-consuming and capital-intensive to enter. But whatever definition we're using, I think we've got to be clear about it. Once we are clear on that, we can then turn to the question of whether big data is a barrier to entry. In my mind, there's no "yes" or "no" answer to that question.

And in some cases, I agree with Barry and others that standard tools would apply. A leading example in my mind would be, say, in an industry where big data is the product or the output itself. Take the Thomson/Reuters[17] case—that is one where both companies are data vendors. They collect, process, and host economic data and then sell the data to the customers. In such a case, I think, to assess whether the data presents an entry barrier, you can ask standard questions, such as how costly it is to accumulate those data, how much does the human capital cost, how costly are the hardware and software, how much time is that going to take. At least conceptually, I don't see that we need to take a significantly different approach.

Now, when the big data is used as inputs to improve products and services, the situation tends to be more interesting and can be more subtle. I have a decision tree in my mind when I think about approaching this situation. The first layer of the decision tree would be to look at the four features of big data and ask whether any of those are relevant to the particular industry we're studying. Let's say two out of the four are the most relevant features; then we'd ask the standard questions again about those features. For example, if the volume is relevant, then how costly is it to obtain the volume? If velocity is the key feature of that industry, then how costly is it to have the ability to capture real time data?

If the conclusion is that, yes, those features could serve as potential barriers to entry, then we move down to the next layer of the decision tree and ask whether those features are necessary for entry. The analogy I have in mind is to think of big data as ingredients available to a chef. Many of us have watched cooking shows where you are given some random ingredients and your job is to cook a great dish. And different chefs will have different ways to approach a dish. You don't need to necessarily use exactly the same ingredients to make a great one.

An entrant may enter into the market with a very good idea that could remove or circumvent some of those Vs relevant to the incumbent. For example, they may not need to use as large a data set as incumbents may have to use. There is an active research area in machine learning whose aim is to find ways for machines to learn without using as much data as is typically needed. One example of such an approach in the machine learning field is called transfer learning. People are interested in the idea because they wonder, what if we don't have this much data? How far can I go with limited data? This is still an active research area in machine learning and AI.

Of course, we should recognize that predicting whether someone could replace the big data with a big idea is typically not easy. And we need to be aware of that limitation.

In any event, I think having a framework such as the one I described can be very helpful when analyzing this question.

**DARREN TUCKER:** I agree that we need to look at this issue on a market-by-market, case-by-case basis, but we have had a number of cases involving whether data is a barrier to entry. So we can look to for patterns in terms of what the agencies are finding.

---

[17] Complaint, United States v. The Thomson Corp., No. 08-CV-00262 (D.D.C. Feb. 19, 2008).

Cases where the agencies have found data to be an entry barrier typically have involved commercial data that was proprietary and difficult to source, and customers often required historical data and a reputation for quality. On the other hand, the agencies have consistently found that personal data, or user data, collected over the internet or through other means do not constitute a barrier to entry. EU and U.S. decisions on point there include Google/DoubleClick and Facebook/WhatsApp.[18]

Another pattern we can see at this point is that where the agencies have required relief in big data cases is where data is the product, as opposed to an input. Ai made a similar point referencing the Thomson/Reuters case. Other examples include CoreLogic/DataQuick,[19] CCC/Mitchell,[20] and Dun & Bradstreet/QED.[21] In contrast, the U.S. agencies and the European Commission have not required divestitures where data was used as an input.

Even where there have been concerns around data being a particularly important input, the agencies have not required a standalone divestiture of data. A good example is the Nielsen/Arbitron case,[22] where the FTC found audience data to be quite important to the end product, yet the divestiture involved the overall measurement platform.

**KEVIN CHRISTENSEN:** There have been lots of conversations and research discussing whether or not antitrust is the right venue to address privacy concerns. So, I pose that question to you: is antitrust the right venue to address privacy concerns?

**JAMES COOPER:** I'll give you an unequivocal "no." We often talk at the FTC about how the Bureau of Competition and the Bureau of Consumer Protection can work together. But I think this is one place where privacy is probably best left in the domain of Consumer Protection.

I think there are two core problems with moving privacy into the domain of antitrust. And before I continue, let me say that I think the law is clear that courts would not consider privacy directly as a value on a par with consumer welfare in an antitrust analysis. What I'm talking about is smuggling privacy into antitrust as a metric of competition, akin to quality.

So first, antitrust tackles non-price competition, but I think the notion that a reduction in privacy can be analogized to a reduction in quality—or a concomitant increase in price—breaks down under close scrutiny. Suppose an automaker charges the same price for a car but leaves the air conditioner out. It immediately increases its profit from a reduction in quality because its costs are lower. But if a company takes more data from you, they aren't immediately more profitable. In fact, they've invested in the collection and perhaps analysis. The only way to make collecting more data a winning strategy is to monetize it. They do this by using it themselves to offer more tailored products, driving more people to their platform so they can increase ad revenues, or by selling it to others so they can advertise more effectively.

---

18 Case COMP/M.7217—Facebook/WhatsApp, Comm'n Decision (Oct. 3, 2014), http://ec.europa.eu/competition/mergers/cases/decisions/m7217_20141003_20310_3962132_EN.pdf; Case M.4731, Google/DoubleClick (Mar. 11, 2008), http://ec.europa.eu/competition/mergers/cases/decisions/m4731_20080311_20682_en.pdf; Statement of Fed. Trade Comm'n, Google/DoubleClick, FTC File No. 071-0170, https://www.ftc.gov/system/files/documents/public_statements/418081/071220googledc-commstmt.pdf.

19 CoreLogic, Inc., FTC Docket No. C-4458 (2014).

20 FTC v. CCC Holdings, Inc., 605 F. Supp. 2d 26 (D.D.C. 2009).

21 Dun & Bradstreet Corp., FTC Docket No. 9342 (2010).

22 Nielsen Holdings N.V. & Arbitron, Inc., FTC File No. 131-0058 (2013).

My point is, at the end of the day, collecting data is about selling something to a consumer. So there's always a benefit tied into the data collection. Now, because tastes and preferences are heterogeneous, some may be privacy sensitive and find increased use of data harmful on net but others may find the bargain—more data for more content and more relevant ads—a net positive. So, unlike quality, which is a vertical dimension, a decision by a firm to take more data is more like positioning oneself in a horizontal space.

As a hypothetical, suppose two firms merge and one has really stringent privacy protections and the other less so. Further suppose that they decide that the new entity will adopt the less privacy-protecting model—the ad-supported application—rather than the paid application.

I would argue that this scenario isn't like a reduction in quality in the same way as if two automakers merged and then decided to exercise market power by maintaining price but reducing the quality of their cars. In the privacy context, because privacy preferences vary, some consumers may prefer that new firm, some may not. Again, it's more like repositioning in a horizontal dimension than in a vertical dimension.

I think the other really big problem of introducing privacy as a dimension of competition is that it introduces a whole range of subjectivity into the analysis. While this will help make sure economists and lawyers continue to make a good living, it will inject more uncertainty into the regulatory system, which is always bad for business. Antitrust analysis can be quite complex, but nonetheless, you have a metric. You have price and output. As you move away from those objectively verifiable metrics, analysis starts to become more subjective and less predictable. Privacy is in the eye of the beholder and so could you imagine that the subjectivity that would go into the Commission, the DOJ, or a court deciding whether a merger that eliminates a privacy-protective firm is harmful to consumers.

Additionally, when you introduce uncertainty and subjectivity, you introduce a lot of dissipative expenditures on rent seeking. As you expand the dimensions over which courts and agency officials can make antitrust enforcement decisions, you'll necessarily draw resources into convincing courts and agency officials to act, or not act, over these dimensions.

**BARRY NIGRO:** I can imagine circumstances in which privacy is a basis on which firms are competing. In that circumstance, it could be relevant to some extent to the analysis of competitive effects, if privacy was something that consumers valued and it differentiated competition among firms in a significant way. Just like any other point of differentiation, traditional antitrust analysis could take that into account, along with other relevant factors, in evaluating whether the firms are close competitors or not.

I don't think there's anything unique about accounting for privacy in that context. I agree with James that privacy for the sake of privacy is better managed using other tools, to the extent that it is a value society wants to promote.

**TERRELL McSWEENY:** I agree with what Barry's saying. I think the appropriate threshold question for enforcers is whether privacy is a dimension of competition in the markets that are before them. And if it is, then are there quality or innovation effects from a loss of that competition? Are they offset?

It just can't be assumed that competition on privacy is actually occurring. There must be some evidence of it. The challenge is that we quickly get into a relatively perilous zone for competition enforcers because it's hard to engage in competition analysis based on what you think consumers may want. You can't create competition and privacy features and services where none exist, even if you think it would be good to have it—that very quickly leads into the broader privacy policy debate. The debate over whether or not there is an appropriate level of privacy regulation is happening in the U.S. and other parts of the world. We've seen implementation of the GDPR in Europe, which is a very strong privacy regulatory framework, and is having a lot of different effects. The complexities within those policy debates really underscores how multifaceted a problem privacy actually is, which is why competition enforcers should be careful about straying too far into the privacy zone.

It's also important for competition enforcers to be mindful of the tension that may start to exist between privacy regulatory frameworks and competition frameworks. Competition is important to innovation. Innovation and competition both require, to a certain extent, the flow of data in marketplaces. IP privacy regulation results in locking down data, which could result in new barriers of entry forming around data, and the competitive significance of data could change, depending on the policies that are adopted. That might chill innovation.

Some of the most harmful policies, in my view, are data localization policies, which further cabin data in certain countries. And in that space, I think competition enforcers have an advocacy role to play to help a policymakers get the balance right between privacy, data protection, and innovation.

**DARREN TUCKER:** We've touched on two discrete issues relating to big data and privacy.

James was principally addressing whether traditional privacy or consumer protection concerns should be incorporated into antitrust analysis. So, for example, if a merger resulted in an accumulation of data, would that by itself be sufficient grounds to block the transaction, even if there was no competitive effect from the data accumulation? I fully agree with James that there are good reasons not to incorporate these privacy considerations into competition analysis. There are different views of what privacy is, and I have no idea how you would balance competition concerns such as price and output with privacy concerns.

If you did incorporate loss of privacy as a potential concern in antitrust enforcement, you'd presumably have to also consider a gain of privacy as a potential efficiency. So if a transaction was otherwise anticompetitive and the merging companies committed to provide greater privacy protections, does that mean that the deal then goes through, even though it's anticompetitive under a traditional competition viewpoint?

But, thankfully, we don't have to get to those difficult questions because this is one of those rare areas where the law is crystal clear. It is black letter law that traditional consumer protection concerns, such as privacy, cannot be incorporated into competition analysis. Over 80 years ago, the FTC found that the use of deceptive advertising by the seller of a supposed obesity cure was an unfair method of competition.[23] And the Supreme Court reversed, finding that consumer protection concerns were beyond the FTC's unfair methods of competition jurisdiction. That decision led the FTC to get its consumer protection authority.

We've seen the Supreme Court repeat this point time and time again. In *Philadelphia National Bank*,[24] *Indiana Federation of Dentists*,[25] and *Professional Engineers*,[26] the Supreme Court has

[23] FTC v. Raladam Co., 283 U.S. 643 (1931).

[24] United States v. Philadelphia Nat'l Bank, 374 U.S. 321 (1963).

[25] FTC v. Indiana Fed'n of Dentists, 476 U.S. 447 (1986).

[26] Nat'l Soc'y of Prof'l Eng'rs v. United States, 435 U.S. 679 (1978).

said that the only thing you're supposed to look at in an antitrust case is the effect on competition. You're not allowed to look at other economic or social values, as legitimate as they may be in other contexts. You see the same thing in the U.S. Merger Guidelines.

This is not a U.S.-specific view. In the Facebook/WhatsApp case, the European Commission said it didn't have the authority to consider privacy considerations in the course of a merger review. The European Court of Justice said the same thing in the *Asnef-Equifax* case.[27] So, there's a general consensus on this point.

Then there's the second issue that Barry and Terrell addressed, which is whether privacy is an element of competition that antitrust agencies should consider? That's a distinct question. And I agree that they should.

Again, we have some Supreme Court guidance. In *LinkLine*,[28] the Supreme Court said there's no distinction, as far as antitrust is concerned, between price and non-price effects. Consistent with that, we've seen the FTC, DOJ, and EC bring cases on the basis of a reduction in quality, such as *Promedica*,[29] *H&R Block*,[30] and the FTC's Intel[31] and Google[32] investigations.

In theory, as Barry said, you could have concerns about a merger between firms that compete closely on the basis of privacy. But, in practice, this kind of case is going to be extraordinarily rare. There haven't been any cases along these lines to date despite interest by some regulators in bringing such a case, and even cases alleging quality diminution are relatively rare compared to more traditional effects, such as price and output.

There are a lot of practical difficulties here. For one thing, we don't have any economic or empirical evidence that increased concentration will lead to a diminution of privacy. More broadly, there's no correlation between concentration and quality.

It's interesting that many of the cases where there have been calls to block a merger on the basis of a loss of privacy competition are often the cases that are the least likely to have privacy effects. The classic example is Facebook/WhatsApp. There were consumer and privacy groups urging enforcement under Section 7 on the view that Facebook was likely to reduce WhatApp's much more robust privacy protections. But in fact, a standard unilateral effects analysis tells us a transaction like that is very unlikely to lead to a reduction in privacy competition because they're not close competitors on the basis of privacy competition. They're distant competitors.

A more plausible example of a reduction in privacy competition would be a merger between firms that provide very strong privacy protections. Even in that scenario, antitrust concerns would only arise if the merging parties competed in the same relevant market, repositioning on the basis of privacy was unlikely, and consumers viewed privacy as an important element of competition.

And even for the very rare acquisition where we have legitimate concerns about the loss of competition on the basis of privacy, the anticompetitive effect is probably not going to be less privacy, it's going to be higher prices or reduced quality. That's because it's hard to reduce privacy. There are regulations in the U.S. and now particularly in Europe that make it very difficult for companies to reduce their privacy protections. By contrast, it's relatively easy to raise prices or spend less money innovating.

---

27 Case C-238/05 Asnef-Equifax, ECLI:EU:C:2006:734 (Eur. Ct. Justice).

28 Pacific Bell Tel. Co. v. linkLine Commc'ns, Inc., 555 U.S. 438 (2009).

29 ProMedica Health Sys., Inc. v. FTC, 749 F.3d 559 (6th Cir. 2014).

30 United States v. H&R Block, Inc., 833 F. Supp. 2d 36 (D.D.C. 2011).

31 Intel Corp., FTC Docket No. 9341 (2010).

32 Google Inc., FTC File No. 111-0163 (2013).

**JAMES COOPER:** Yes, I completely agree that privacy could be a dimension of competition, and in theory could be challenged by the agencies. I agree with that, but I still think looking at privacy as a dimension of competition raises a serious subjectivity problem. In the case of two firms that are competing closely on privacy, how would you determine that they're competing closely on privacy? And let's say that they come together and now we just have one firm that's really privacy-sensitive as opposed to two, what does that really mean for consumers? What if the combined firm stops being as privacy-protective? How do you measure market share? How do you measure the extent to which privacy in a marketplace has been diminished?

So, again, in the realm of theory, you can certainly imagine a scenario where we have evidence that two firms were competing on privacy, and that it was important to consumer decision-making. But in practice, I think you will end up with a great deal of subjectivity, which to me is really a first order concern.

**BARRY NIGRO:** Both the FTC and the Department of Justice have to go to court and explain to a judge why a transaction will reduce competition. If there isn't relatively strong evidence in the documents and the testimony that it's an issue, you're likely not going to see a case.

The question is whether there's evidence that it's a meaningful dimension of competition between the firms, whether, as a result, they're closer competitors than the other firms in the market, and the competitive significance of it. I don't think there are likely to be many cases like that, but who knows? It's like any other qualitative factor that affects whether firms are close competitors.

**AI DENG:** It is often raised that one challenge in thinking about privacy as a type of quality is that there's a lot of heterogeneity, i.e., different people value their privacy differently, and it's really hard to put a number on it, unlike a price.

For example, I use two social media platforms, and one has a very stringent privacy restriction. I don't see any posts or replies by anybody who I'm not connected with. I also use Facebook, where I can see posts by people I don't know. Do I value the privacy offered by the other platform? On the margin, maybe I do. But by how much? I have not figured it out. I think that speaks to this type of heterogeneity and challenge. I know some had proposed the idea that we might just need to better educate consumers, give them more information about what privacy means, what the companies are using the information for, just so consumers know the value of the information they're sharing. But some have also argued that this is not going to be that useful.

I am sure there will be a lot more discussions and research to come. But when we think about whether privacy is a competition issue, I believe it is relevant to understand how consumers value it.

**DARREN TUCKER:** I agree that there is a wide range of views of privacy among consumers. And some consumers view privacy, at least in certain contexts, as a negative. There are over 10 million people that have a public Facebook account, for example, and there are lots of other social networking services that are designed to broadcast personal information as widely as possible.

In the context of a merger review or other antitrust proceeding looking at the importance of privacy competition to consumers, it may be challenging for authorities to accurately gauge consumers' privacy preferences. There is often a wide differential between what consumers will say in a survey, for example, about their privacy preferences versus their actual behavior. So, the same consumer who will say that he feels strongly about protecting his privacy and doesn't want firms to collect his personal information, will then turn around and fill out a card to enter a contest, providing lots of personal information for a prize effectively worth a few cents.

And so it's important to make sure you have good data in this context, and looking at revealed preferences is generally going to be superior to survey-based approaches.

**TERRELL McSWEENY:** We're not going to solve the privacy policy debate in this discussion. You're making a relevant point, but I think the privacy policy debate is far, far broader and more multifaceted, which is why the competition tool is not necessarily the right tool to address it. As we discussed, enforcers quickly get into that gray area that we've been navigating, and it's confusing and hard to figure out how to achieve the right balance.

That said, I think there is also a case to be made that there's a bit of a market failure when it comes to privacy data policy because of the well-documented information asymmetries that exist in the marketplace. Most people don't have a great sense of what the deal is when they're exchanging their data. They're getting a better sense, and we're seeing it result in people starting to have a different relationship with technology.

I think we all ought to be very concerned about an erosion in trust in technology because it could affect adoption. For innovation to flourish and for competition to flourish, consumers need to have confidence in technology. So these are important areas for discussion, debate, and engagement.

**KEVIN YINGLING:** That's a good place to end. Thank you all for spending your morning with us discussing big data. It's been a great conversation and we really appreciate it. ●