FTC Hearing: Big Data, Competition & Privacy
Panel: How companies use data
6 November 2018
Remarks by Marianela Lopez-Galdos
Computer & Communications Industry Association

Introduction

Good afternoon. My name is Marianela Lopez-Galdos, and I am the director of competition & regulatory affairs at the Computer and Communications Industry Association. Before I start with my remarks, I would like to thank the FTC for giving me the opportunity to speak here today about big data, competition & privacy, and more broadly, would like to congratulate the FTC for putting together these hearings.

It is my understanding that the question we are trying to respond in this particular panel is that of how companies put data to use. So, I thought, it would be interesting to discuss the role data has for data-driven businesses and use some examples to this end. With this in mind I would like to highlight 3 main ideas, and expect to have the opportunity to get into more details during our discussions with the rest of panelists.

1.   Data is not essential, ideas are

The first idea that I bring to this discussion is how, for companies to flourish and compete, data was not, and still isn't essential. What I mean by this is that many of the successful companies that use data do not exist because of the data they accumulate. Data successful companies, just as in the old brick and mortar world, are born out of an idea. Without an idea that brings a new or better product or service to the market, data-driven companies would not exist. So ideas come first.

As simple as obvious as this might sound, this fact is often forgotten when discussing the role that data plays for companies. It is undeniable that companies that invest in R&D, including data analytics, can innovate and evolve by creating value. However, it is important to differentiate this idea from the notion that digital companies exist because of their access to data.

In fact, if we look back to recent history, we may find recent examples of how innovative winners without data can gain success.   Companies like Snapchat and Slack gained hundreds of millions of users despite starting with no data. A more recent example is Handshake - a LinkedIn competitor that focused on college students and recent grads - that has used their new idea to convince 14 million users to join their platform.

Two leading economists, Lambrecht and Tucker, that have been participating in these hearings put it very clearly:

> The history of the digital economy offers many examples, like Airbnb, Uber and Tinder, where a simple insight into customer needs allowed entry into markets where incumbents already had access to big data.

Therefore, when discussing how companies use data, it is important to bear in mind that companies exist and succeed prior to those companies accessing data. And only as these companies succeed, data allows them to improve that original product and/or service that they brought to the markets, or even bring new products and services if they invest in R&D and innovation.

Furthermore, it is not necessarily true that the more volume of data that companies can process the better results they might get. In fact, recent studies show how access to data has limited diminishing returns to companies.

Stanford University conducted a study to analyze whether increased accumulation of data improves the outcomes of the analysis performed on such data. The Stanford Dogs Dataset contains images of 120 breeds of dogs from around the world.[1] Researchers used this dataset for classifying breeds of dogs in images, and calculated the mean accuracy for identification as the number of images in the dataset increased. The results showed that additional access to data provided diminishing returns to the accuracy of classification results.

It shouldn't be surprising that at some point data doesn't yield better results -- think about your own data. It can become stale quickly: for example, data about your interest in purchasing a car is not useful after you purchase the car.

As such, proposals to classify data as an essential input are unfounded and rest on a misunderstanding of the concept of data that is, among other things, non-rivalrous. The clarification of the non-essential nature of data takes me to the second idea.

2.   Innovation vs. market positioning drives the digital economy

The second thought that I would like to share with you is the role that data analytics-based innovation plays in allowing companies to discover and develop opportunities for value creation.

If we accept that access to data has diminishing returns to companies, then it becomes easier to understand that the mere accumulation of data will not help a company drive competitors out of the market. So how do companies use data to compete?

---

[1] Stanford Dogs Dataset, *available at* http://vision.stanford.edu/aditya86/ImageNetDogs/.

Companies invest in what we could call "data-driven R&D" to look for new opportunities to build their companies. In the value-creating process, data analytics and the capacity to process data becomes fundamental for companies to innovate.

In the digital economy, innovation, rather than market positioning, is more relevant in the determination of winners and losers. This is how we have evolved from IBM´s linear computers to new quantum computing, or Machine Learning. If companies want to remain competitive, they have to innovate.

These examples are also helpful to illustrate how in the data driven economy, because innovation plays such a crucial role, new ideas together with the deployment of data intensive technologies such as machine learning, will revolutionize markets as we know them.

3. Data needs will diminish over time

This leads me to the third and final idea which is that as technologies advance, the need for big data will diminish. We are already observing this trend in the real world with the development of new technologies.

A good example of these new technologies is how engineers are trying to apply "zero shot learning" to improve visual recognition. "Zero shot learning" refers to the process by which a machine learns how to recognize objects in an image without any labeled training data to help in the classification. In other words, "Zero shot learning" helps machines categorize objects that they have never seen before. By applying this technology a machine may recognize objects, e.g. distinguish a zebra from a horse, without the need to employ so much data. So as these experiments progress, less data will be needed.

Similarly, with the development of synthetic data that holds no personal information and cannot be traced back to any individual, confidentiality and privacy are being protected as technology advances.

4.   Conclusion

Let me conclude by noting the need to preserve the incentives for companies to experiment with data, especially in light of the progress that data intensive technologies such as machine learning have experienced in the last years. Despite this progress, we are only at the beginning, at the nascent moment of the lifecycle of these new technologies. Therefore, authorities' intervention concerning data should be limited only to those instances where there is a risk of consumer harm.

The incentives for companies to continue to advance in the field of AI and machine learning have become, one of the most important features that will determine our future. This is why intervening data-driven markets without evidence of harm to consumers or competition where effects of such harm is evaluated would be detrimental to innovation.