Scalable Approaches to Transparency and Accountability in
Decisionmaking Algorithms

Remarks at the NYU Conference on Algorithms and Accountability

Commissioner Julie Brill
February 28, 2015

Thank you, Kathy, for the introduction, and to NYU and the Information Law Institute for hosting this fascinating conference. Kathy, you asked us to address the rise of algorithmic decision-making in the context of consumer-level commerce. I'd like to give an overview of the issues from my perspective as a Commissioner of the FTC – that is, as a government official charged with protection consumers -- and then Oren and Natasha can go into some depth with a couple of very interesting examples.

My main message is that the data that feeds these algorithms and the outputs from the algorithms can be deeply significant for consumers, more so than the algorithms themselves, and all of us – technologists, academics, advocates, businesses, and policymakers – have roles to play in giving consumers more meaningful control over both ends of the algorithmic equation.

The first thing I want to do is to bring more clarity to what we mean by "decisions." There are decisions, and then there are *decisions*. Understanding the range of decisions is important, because different types of decisions both raise different consumer protection challenges and fall under different laws and regulations.

At one end of the spectrum are marketing decisions. The advertising industry has developed a huge machinery to collect data about consumers' activities, put the data into profiles, and segment consumers according to observed and inferred characteristics. All of this data goes into automated decisions about which ads to show to consumers. Those decisions can, in turn, have a significant influence on the choices that consumers make. That's the whole point of marketing.

Yet important questions arise even in the marketing realm. If consumers see ads for high interest payday loans but not entry-level, lower-cost offerings from traditional banks, these consumers might understandably be enticed to click through the ads for payday loans, and potentially obtain loans that are more expensive than they might obtain based on a full and fair evaluation of their financial status. The FTC Act prohibits unfair or deceptive acts or practices in marketing, but generally speaking, the laws in the U.S. generally don't reach the types of data that companies use in marketing decisions, or the results that pop out of the algorithms that determine which ads to show to consumers.

At the other end of the spectrum are decisions about *eligibility* and *costs* of credit, insurance, employment, housing, and other services and opportunities of basic importance to most people's everyday lives. The history of the credit reporting industry holds some important insights into the benefits and harms of today's algorithmic decisions about consumers.

Credit reports are the grease that keeps the consumer economic wheel turning. Prior to the advent of credit reports, consumers obtained loans if they knew their local banker, or had a social reputation that preceded them into his office. Not every town had a George Bailey!

But in the 1950s and 1960s, as credit bureaus increased the amount of personal information held in their growing data bases, unease about the amount of information that credit bureaus held – as well as its accuracy and its use – also increased. Congress recognized that the data collected by third parties and the use of this data in critical decisions in consumers' lives may lead to efficiencies, but this newly developing system could also create serious risks for consumers if data is too old, incomplete, or simply wrong. As a result, in 1970, Congress enacted a robust law, the Fair Credit Reporting Act (FCRA), that gives consumers important rights to access, dispute, and correct information in their credit reports. The FCRA also restricts the purposes for which credit bureaus can sell credit reports, and requires users of credit reports to provide certain notifications when a report is the basis of an adverse decision, among other things.

Over time, industry developed ways to boil information in credit reports down to a single number – a credit score – intended to serve as a heuristic device that places the consumer on a scale of "riskiness". And companies began to expand the use of credit scores beyond the credit context; for instance, insurers started to use credit scores to set auto insurance premiums. Questions of fairness about the use of credit scores in these non-credit contexts began to emerge again. For instance, people started to ask what data is used to calculate the scores? Do scores have disparate impacts on certain groups of consumers? These concerns led Congress in 2003 to order the Federal Trade Commission to study credit score-based auto insurance premiums.

The Commission spent four years thoroughly studying this issue. Its report found that credit scores did not serve as a proxy for characteristics such as race or sex in this context. The report provided much needed transparency, particularly about fairness of the use of credit scores for auto insurance, which was important information for all stakeholders – policymakers, consumers, and companies – to know.

But this was a special case. The uses of credit scores that the FTC examined fell within the existing regulatory framework of the FCRA. An increasing range of algorithmic scores and decisions fall outside of any such framework. The FTC identified a few of them in our May 2014 report on data brokers. We highlighted so-called "risk mitigation" services as sources of potentially significant decisions about consumers that are not subject to the specific protections of the FCRA. What do these services do? They answer questions like "Is this consumer who she claims to be?" and "Is the purchase that this consumer is attempting to make likely to be fraudulent?" While some uses of these "risk mitigation" scores may fall under the FCRA, an important set of them does not.

Another important development is coming about through the Internet of Things. There are currently around 25 billion connected devices on the planet, and in five years it is estimated that there will be 50 billion. Much has been said, including by me, about the potential societal benefits, as well as the data security and privacy risks that will come with the burgeoning world of connected devices and sensors. But today I would like to talk about how these relatively

cheap, connected devices and their role in measuring what we do, will have consequences that go beyond innovation, efficiency and risks to privacy and security. Many decision-makers have traditionally relied on third parties for data about individual consumers to make decisions. Now, with more sensors in more places, some decision-makers could collect and analyze the data from these sensors to make their own decisions. For example, an auto insurer that once relied on credit scores to help set premiums could simply ask consumers to install a sensor in their vehicles to help the insurer monitor the consumer for risky driving behavior. Consumers that fail to accept the sensor could face higher premiums. The crucial twist from my law enforcement perspective is that the FCRA – with its access, accuracy, and other requirements – doesn't apply to the information that the insurance company collects on its own behalf.

As the Internet of Things makes it easier to take a "DIY" approach to traditional credit reporting, the limitations of the FCRA will become more pronounced.

Transparency and accountability are important across the full range of decisions that I've illustrated. But I don't think it's realistic to rely on the approach that the FTC took to understand credit-based auto insurance pricing. As I said, it took the FTC nearly four years to study the ramifications of using credit scores for auto insurance underwriting. The FTC – and all other federal agencies for that matter – simply do not have the capacity to study every score out there. To borrow a phrase from the technology community, this approach "doesn't scale."

What is a more scalable approach? If we're going to bring greater transparency and accountability to the use of algorithms across the very broad range of settings that I have discussed, we need to think in terms of roles for everyone – companies, technologists, consumers, and policymakers.

Let me begin with consumers. I have long been a big proponent of giving consumers more control over their personal information. Consumers should be able to exercise appropriate control over information that goes into the pipelines that feed the algorithms that end up having an effect on their lives, particularly those pipelines that are not now consumer facing, as Solon pointed out. For example, consumers ought to be able to tell data brokers that they do not want to have their information used for marketing purposes. Consumers should also have the ability to correct information that is used for risk mitigation and other comparably substantive decisions. And industry should provide consumers with usable tools – like centralized portals – to allow them to exercise this control. Both the White House and the full FTC included my recommendations in recent reports on big data and data brokers.

Some in industry are taking steps to provide greater transparency and control to consumers, but we have a very long way to go here. I believe we need legislation to address these issues, but industry can and should do more right now. Companies can provide consumers with creative UIs to provide more meaningful, usable access to their data.

But transparency is not the whole answer; consumers cannot do this themselves. Companies should do more to determine whether their own data analytics result in unfair, unethical, or discriminatory effects on consumers. For example, what if a company analyzing its own data, in an effort to identify "good" versus "troublesome" customers, ends up tracking

individuals along racial or ethnic lines. A *Harvard Business Review* article argues that this kind of result isn't just possible, but inevitable. All companies should think carefully about where "value-added personalization and segmentation" ends, and harmful discrimination begins.

As many have mentioned today, technologists have a key role to play. They have the skills to make data access tools that are easy for consumers to use, and they have the technical insights that are necessary to determine whether specific analytics practices pose risks of excluding, or otherwise placing at a disadvantage, groups defined according to sensitive traits. As others have said, this isn't a simple task, but it is critically important that technologists focus on it now.

Let me conclude with a quick word about the FTC's role. We will, of course, bring enforcement actions in appropriate cases. But I also believe the agency should play a constructive role by encouraging best practices that reflect input from a broad range of stakeholders. The FTC has lots of relevant expertise that could help lead the way to practical solutions right now.

Thank you.