Massachusetts Institute of Technology
Sloan School of Management
Catherine Tucker
Sloan Distinguished Professor of Management
Professor of Marketing
Chair MIT Sloan PhD Program

Management Science Area
Phone: (617) 252-1499
E-mail: cetucker@mit.edu
cetucker.scripts.mit.edu/

TO: Marsha Blackburn, Chairman of the Subcommittee on Communications and Technology, Robert E. Latta, Chairman of the Subcommittee on Digital Commerce and Consumer Protection
FROM: Catherine Tucker
DATE: January 2018

Responses to additional questions arising from Testimony on `Algorithms: How Companies' Decisions about Data and Content Impact Consumers'

Dear Congresswoman Blackburn and Congressman Latta,

In response to your letter dated December 15, 2017, I would like to offer additional answers to the supplementary questions submitted by members of Congress.

 The Honorable Robert E. Latta

1.   *Your research has demonstrated that economic factors can provide for different outcomes than perhaps even the creators of the relevant algorithms might have intended. Assumptions might be built into the algorithm about what does or does not count, but the output might not be reliable or intended due to unforeseen factors not built in the model. Can a regulatory technique be fashioned to solve these types of problems, or is it another way of saying that highly complex systems at this juncture in time will tend to produce unexpected results?*

I think you are exactly right to hone in on the question of what regulatory technique may be most appropriate at this time of uncertainty and transition.

My research so far has been mainly focused on the question of whether "algorithmic transparency" is sufficient or necessary as a regulatory regime for algorithms. My research suggests it is neither sufficient nor particularly helpful as a policy emphasis. The reason algorithmic transparency is not sufficient is that there are many cases where the data that the algorithm feeds on, not the algorithm itself, is what causes bias (or at least the appearance of bias). Just looking at the algorithmic code (supposing that was even viable or possible) would not allow regulators to identify instances of bias in such cases. The reason it is not helpful is that "hard coding" of bias or discrimination into a code is rare, from my experience in talking to many technology companies. It is simply not the case that programmers add lines to their code where they instruct the algorithm to treat

groups differently on the basis of race or gender. Instead, a more appropriate area of concern is a focus on the complex interactions of algorithms with data and human behavior, with a particular focus on studying outcomes relative to their non-algorithmic counterfactual.

2.   *What impact have prescriptive regulations, such as those promulgated under the Children's Online Privacy Protection Act, had on investment and creation of content for children?*

We have a new paper on the question of children's privacy in mobile applications. It has not been released yet, but I hope to have a copy up online by February and will email your staff to ensure the committee has access. This paper focuses on mobile applications on smartphones that are targeted at the under-5s. We find two main things:

> 1) A surprising number of applications targeted at toddlers or preschoolers collect highly personalized data (including precise location data).
> 2) Many of these applications are developed by developers outside the US, and in particular the most intrusive applications often originate from developers based in Asia and countries like Ukraine.

One interpretation of these results is that potentially developers within the US are not developing apps for children because of concerns over legal compliance, and that absence has attracted foreign developers who are not constrained by any concerns for children's privacy. Consequently, we may inadvertently have a situation where our children's privacy laws may have led to worse privacy practices in the apps on the market.

The Honorable Gregg Harper

1. Can you tell us some of the considerations consumers make when deciding to exchange private information for services, and the degree to which existing disclosure rules factor into those decisions?

This has been a great deal of research into this question over the last few years, which I probably cannot do justice to except for saying that this seems to be very context dependent. My own research highlights that more sophisticated or technologically-savvy consumers often are unwilling to share personal data, unless it is framed in terms of an economic exchange whereby there is some gain (even if it is negligible) in doing so. My research also suggests that the effect of disclosure rules depends on the extent to which they are accompanied by a parallel sense of control for the consumer. Simply receiving information about the potential risks of data disclosure can be off-putting to consumers, unless they are offered (even a slight) sense of control at the same time.

The Honorable Michael C. Burgess

1.   In your testimony, you state that "algorithms may appear biased." In your research, how does algorithmic bias manifest itself - how do you measure issues of bias or fairness?

My research has been focused on algorithms where the outcome may appear biased, but the bias reflects market outcomes, rather than human bias as such. The way we measured this apparent bias was by seeing whether or not women or men were more likely to see ads for jobs in Science, Technology, Engineering and Math. We found that women were less likely to see such ads, not because of any direct bias or because the algorithm predicted that women were less likely to respond to the ad. Instead, it was because women are such a desirable demographic that they cost more to advertise to, and the algorithm, in its attempt to save the advertiser money, showed fewer ads to expensive female eyeballs.

This particular example, I think, is useful in illustrating how hard it is to say whether an algorithm is "biased" or "fair" or "discriminatory." Instead, we have a well-meaning algorithm trying to be cost-effective, which inadvertently leads to an outcome where women see fewer job ads in a way we may find as a society undesirable. I would hesitate to call this 'bias,' but instead think of it as an example of the occasional inadvertent consequences of well-intentioned algorithms leading to outcomes that are less than desirable.

2.   If an algorithm tends to produce results that were not intended by its creator, what is the likelihood that fact will be discovered and corrected?

I think this will be context dependent. If I were to speculate, my guess is that firms who are developing specialized algorithms for obviously sensitive areas (such as predictive policing, predictive sentencing, enhancing hiring decisions) will be more likely to conduct audits and ensure that their results are not inadvertently distorted.

My concerns would instead focus on firms that are developing algorithms whose client base is broad enough that they may not be aware that there will be particular cases or situations where algorithmic bias may matter. One example of this is the advertising industry. On the whole, we don't really care as a society who sees a particular shoe ad. On the other hand, there are isolated cases where we do care who sees advertising - for example, we might care if discriminated-against racial groups were more likely to see ads from predatory lenders, and we might care if women are less likely to see ads for high-paying jobs than men.

The Honorable Adam Kinzinger

1. Given that companies tend to have extensive and rather transparent privacy policies, does more disclosure tend to make consumers more reluctant to use a particular service or site?

My research suggests that in general more disclosure can have a chilling effect. This is partially because consumers can find it off-putting, but also because complying with disclosure requirements can impose costs on firms, meaning they may not offer that particular service. The exception to this is when disclosure is accompanied by a parallel sense of control for consumers. In such instances, consumers are encouraged to use a technology. So for example, if a company has a set of disclosures but also communicates to a consumer that they retain control or ownership of their own data, that can increase the chance of a customer using a service or website.

I hope these responses are helpful. Please let me know if you have any concerns or if there is anything I can clarify.

Yours sincerely,

Catherine Tucker