Algorithmic Harm in Consumer Markets[*]

Oren Bar-Gill[†]   Cass R. Sunstein[†]   Inbal Talgam-Cohen[‡]

Abstract

Machine learning algorithms are increasingly able to predict what goods and services particular people will buy, and at what price. It is possible to imagine a situation in which relatively uniform, or coarsely set, prices and product characteristics are replaced by far more in the way of individualization. Companies might, for example, offer people shirts and shoes that are particularly suited to their situations, that fit with their particular tastes, and that have prices that fit their personal valuations. In many cases, the use of algorithms promises to increase efficiency and to promote social welfare; it might also promote fair distribution. But when consumers suffer from an absence of information or from behavioral biases, algorithms can cause serious harm. Companies might, for example, exploit such biases in order to lead people to purchase products that have little or no value for them or to pay too much for products that do have value for them. Algorithmic harm, understood as the exploitation of an absence of information or of behavioral biases, can disproportionately affect members of identifiable groups, including women and people of color. Since algorithms exacerbate the harm caused to imperfectly informed and imperfectly rational consumers, their increasing use provides fresh support for existing efforts to reduce information and rationality deficits, especially through optimally designed disclosure mandates. In addition, there is a more particular need for algorithm-centered policy responses. Specifically, algorithmic transparency—transparency about the nature, uses, and consequences of algorithms—is both crucial and challenging; novel methods designed to open the algorithmic "black box" and "interpret" the algorithm's decision-making process should play a key role. In appropriate cases, regulators should also police the design and implementation of algorithms, with a particular emphasis on exploitation of an absence of information or of behavioral biases.

[†]Harvard Law School, Cambridge, MA, USA.
[‡]Technion–Israel Institute of Technology, The Henry and Marilyn Taub Faculty of Computer Science, Haifa, Israel.