

Algorithmic Regulation and Consumer Welfare
Comments of
William Rinehart and Allison Edwards¹
FTC Docket Number: FTC-2018-0056

Executive Summary

As algorithms and artificial intelligence capabilities have advanced, there have been calls to regulate them. These calls take the form of various regulatory proposals, including creating a federal agency for algorithms, creating a safety board for algorithms, and forcing developers to make their algorithms transparent. As this comment highlights, however, crucial to discussions over algorithms and consumer welfare is the context in which they are implemented.

Algorithms Have to be Implemented

Model selection is a critical component of any algorithm, so it is no wonder that criticisms of risk assessment algorithms have focused on this aspect of the process. Error bars might reflect precision, but they tell us little about a model's applicability. More importantly, however, the implementation isn't frictionless. People have to use them to make decisions. Algorithms must be integrated within a set of processes that involve the messiness of human relations. Because of the variety of institutional settings, there is sure to be significant variability in how they come to be used. The impact of real decision-making processes isn't constrained only by the accuracy of the models, but also the purposes to which they are applied.

The implementation of pretrial risk assessment instruments highlights the potential variability when algorithms are deployed. These instruments can help guide judges when deciding whether a defendant be given bail and at what cost. The most popular of these instruments is known as the Public Safety Assessment, or simply the PSA, which was developed by the Laura and John Arnold Foundation and has been adopted in over 30 jurisdictions in the last five years.

The adoption of the PSA across regions helps to demonstrate just how disparate implementation can be. In New Jersey, the adoption of the PSA seems to have correlated with a dramatic decline in the pretrial detention rate.² In Lucas County, Ohio, the pretrial detention rate increased after the PSA was put into place. In Chicago, judges seem to be simply ignoring the PSA.³ Indeed, there appears to be little agreement on how well the PSA's high-risk classification corresponds to reality, as re-arrest can be as low as 10 percent or as high as 42 percent, depending on how the PSA is integrated in a region.⁴

In the most comprehensive study of its kind [comprehensive study of its kind](#), George Mason University law professor Megan Stevenson examined Kentucky after it implemented the PSA and

¹ William Rinehart is Director of Technology and Innovation Policy at the American Action Forum and Allison Edwards is a Policy Intern at the American Action Forum.

² Jon Schuppe, *Post Bail*, <https://www.nbcnews.com/specials/bail-reform>.

³ Frank Main, *Cook County judges not following bail recommendations: study*, <https://chicago.suntimes.com/news/cook-county-judges-not-following-bail-recommendations-study-find/>.

⁴ Sandra G. Mayson, *Dangerous Defendants*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2826600.

found significant changes in bail-setting practices, but only a small increase in pretrial release.⁵ Over time these changes eroded as judges returned to their previous habits. If this tendency to revert to the mean is widespread, then why even implement these pretrial risk instruments?

Although it was focused on pretrial risk assessments, Stevenson's call for a broader understanding of these tools applies to the entirety of algorithm research: Risk assessment in practice is different from risk assessment in the abstract, and its impacts depend on context and details of implementation. If indeed risk assessment is capable of producing large benefits, it will take research and experimentation to learn how to achieve them. Such a process would be evidence-based criminal justice at its best: not a flocking towards methods that bear the glossy veneer of science, but a careful and iterative evaluation of what works and what does not.

Algorithms are tools. While it is important to understand how well calibrated the tool is, researchers need to be focused on how that tool impacts real people working with and within institutions with embedded cultural and historic practices.

Tradeoffs in fairness determinations

Julia Angwin and her team at ProPublica sparked renewed interest in algorithmic decision-making when they dove deeper into a commonly used post-trial sentencing tool known as COMPAS.⁶ Instead of predicting behavior before a trial takes place, COMPAS purports to predict a defendant's risk of committing another crime in the sentencing phase after a defendant has been found guilty. As they discovered, the risk system was biased against African-American defendants, who were more likely to be incorrectly labeled as higher-risk than they actually were. At the same time, white defendants were labeled as lower-risk than was actually the case.

Superficially, that seems like a simple problem to solve. Just add features to the algorithm that consider race and rerun the tool. If only the algorithm paid attention to this bias, the outcome could be corrected. Or so goes the thinking.

But let's take a step back and consider really what these tools represent. The task of the COMPAS tool is to estimate the degree to which people possess a likeliness for future risk. In this sense, the algorithm aims for calibration, one of at least three distinct ways we might understand fairness. Aiming for fairness through calibration means that people were correctly identified as having some probability of committing an act. Indeed, as subsequent research has found, the number of people who committed crimes were correctly distributed within each group. In other words, the algorithm did correctly identify a set of people as having a probability of committing a crime.

Angwin's criticism is of another kind, as Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan explain in "Inherent Trade-Offs in the Fair Determination of Risk Scores."⁷ The kind of fairness that Angwin is promoting might be understood as a balance for the positive class. To violate this kind of fairness notion, people would be later identified as being part of the class, yet they were predicted

⁵ Megan T. Stevenson, *Assessing Risk Assessment in Action*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3016088.

⁶ Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, *Machine Bias*, <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

⁷ Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan, *Inherent Trade-Offs in the Fair Determination of Risk Scores*, <https://arxiv.org/pdf/1609.05807.pdf>.

initially as having a lower probability by the algorithm. For example, as the ProPublica study found, white defendants that did commit crimes in the future were assigned lower risk scores. This would be a violation of balance for the positive class.

Similarly, balance for a negative class is the negative correlate. To violate this kind of fairness notion, people that would be later identified as not being part of the class would be predicted initially as having a higher probability of being part of it by the algorithm. Both of these conditions try to capture the idea that groups should have equal false negative and false positive rates.

After formalizing these three conditions for fairness, Kleinberg, Mullainathan, and Raghavan proved that it isn't possible to satisfy all constraints simultaneously except in highly limited special cases. These results hold regardless of how the risk assignment is computed, since "it is simply a fact about risk estimates when the base rates differ between two groups."

What this means is that some views of fairness might simply be incompatible with each other. Balancing for one kind of notion of fairness is likely to come at the expense of another. This tradeoff is really a subclass of a larger problem that is a central focus in data science, econometrics, and statistics. As Pedro Domingos noted,

You should be skeptical of claims that a particular technique "solves" the overfitting problem. It's easy to avoid overfitting (variance) by falling into the opposite error of underfitting (bias). Simultaneously avoiding both requires learning a perfect classifier, and short of knowing it in advance there is no single technique that will always do best (no free lunch).⁸

Internalizing these lessons about fairness requires a shift in framing. For those working in the AI field, actively deploying algorithms, and especially making policy, fairness mandates will likely create tradeoffs. If most algorithms cannot achieve multiple notions of fairness simultaneously, then every decision to balance for class attributes is likely to take away from efficiency elsewhere. This isn't to say that we shouldn't strive to optimize fairness. Rather, it is simply important to recognize that mandating of one type of fairness may necessarily come at the expense of a different type of fairness.

Algorithmic Pricing

Dynamic pricing, when sellers set prices using computer algorithms, is one area where the application of algorithms has become an important topic for competition authorities. To be clear, dynamic pricing is just another word for first-degree price differentiation, the act of charging each customer a different price.

Dynamic pricing strategies have been challenging to implement in traditional retail settings due to lack of data, such as competitors' prices, and physical constraints, such as manually relabeling prices on products. In contrast, e-commerce remains unconstrained by these limitations. Collecting real-time data on customers and competitors is straightforward.

In a highly cited study on this topic, Benjamin Schiller at Brandeis University found that,

⁸ Pedro Domingos, *A Few Useful Things to Know about Machine Learning*, <https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf>.

[W]eb browsing behavior substantially raises the amount by which person-specific pricing increases variable profits relative to second-degree PD - 2.14% if using all data to tailor prices, but only 0.14% using demographics alone. Expressed in total profits, this difference appears more striking - 12.2% vs. 0.8%. Web browsing data hence make first-degree PD more appealing to firms and likely to be implemented, thus impacting consumers.⁹

But static pricing has its problems as well. Retailers charge uniform prices although there is variance in consumer income by location, and so, uniform prices effectively are a substantial increase on the prices paid by poorer households. In total, uniform pricing costs firms an average of 7 percent of profit.¹⁰ Wealthier consumers, who would be willing to sustain higher prices, pay around 9 percent lower prices and lowest income group bear prices that are 0.7 percent higher than they would with flexible prices.

While many worry about ever-inflating prices online, the reality on the ground is far more complex. Early studies of online price dispersion, much as theory had suggested, found that the range of prices for a good is far narrower with more sellers.¹¹ More recent work conflicts this original study, and instead suggests that industry prices have become more narrowly confined for all industries, not just those with more sellers.¹² Combined, the widespread use of dynamic pricing seems to have led to an understatement of inflation by 1.3 percent.¹³

Competition authorities have been active in this space for some time. The 1994 case involving Airline Tariff Publishing Company (ATPCO) serves as an example. ATPCO collects and distributes the vast bulk of airline pricing information, and thus provided a space where the airlines could discuss in private price-fixing efforts. These private discussions facilitated pervasive coordination of the airlines to reach overt price-fixing agreements and was resolved when all of the parties entered into a consent decree in 1994.¹⁴

More recently, the DoJ charged a company, Trod Ltd., for using a pricing algorithm to fix the prices of certain posters sold in Amazon Marketplace.¹⁵ One conspirator programmed its algorithm to find the lowest-price offered by a non-conspiring competitor for a poster and set its poster price just below the non-conspiring competitor's price. The second conspirator then sent its algorithm to match the first conspirator's price, and therefore, the two conspirators eliminated competition among themselves. In the end, those involved plead guilty of violating Section 1 of the Sherman Act and agreed to pay a \$20,000 criminal fine.

⁹ Benjamin Reed Shiller, *First-Degree Price Discrimination Using Big Data*, http://www.brandeis.edu/economics/RePEc/brd/doc/Brandeis_WP58R2.pdf.

¹⁰ Thomas Franck, *Retailers are charging the same prices across US, boosting income inequality, new research shows*, <https://www.cnbc.com/2017/11/09/retailers-are-charging-the-same-prices-across-us-boosting-income-inequality-new-research-shows.html>.

¹¹ Michael R. Baye, John Morgan, and Patrick Scholten, *Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site*, <http://faculty.haas.berkeley.edu/rjmorgan/small&large.pdf>.

¹² Alberto Cavallo, *Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers*, <https://as.vanderbilt.edu/econ/seminars-and-research/seminarpapers/cavallo.pdf>.

¹³ Austan Goolsbee and Peter Klenow, *Internet Rising, Prices Falling: Measuring Inflation in a World of E-Commerce*, https://bfi.uchicago.edu/sites/default/files/research/WP_2018-35.pdf.

¹⁴ Severin Borenstein, *Rapid Price Communication and Coordination: The Airline Tariff Publishing Case (1994)*, http://global.oup.com/us/companion.websites/fdscontent/uscompanion/us/pdf/kwoka/9780195322972_09.pdf.

¹⁵ Department of Justice, *Online Retailer Pleads Guilty for Fixing Prices of Wall Posters*, <https://www.justice.gov/opa/pr/online-retailer-pleads-guilty-fixing-prices-wall-posters>.

Conclusion

Advocates who push for regulation of all algorithms need to be cautious. Understanding the internal logic of risk assessment tools is not the end of the conversation. Without data of how they are used, it could be that these algorithms entrench bias, uproot it, or have ambiguous effects. To have an honest conversation, we need to understand how they nudge decisions in the real world. Most algorithms pose little or no risks to safety or consumer welfare and should thus be exempted from regulations. Light touch regulations are the best way to ensure future innovation of algorithms and not to disrupt consumer welfare.