AEA Papers and Proceedings 2021, 111: 49–54
https://doi.org/10.1257/pandp.20211080

Measuring Racial Discrimination in Algorithms†

By David Arnold, Will Dobbie, and Peter Hull*

There is growing concern that the rise of algorithmic decision-making in many settings will lead to discrimination against legally protected groups (Angwin et al. 2016). This concern has fueled a rich theoretical literature in computer science, where algorithmic discrimination is formalized as the differential treatment of equally qualified individuals (Zafar et al. 2017, Berk et al. 2018). In the context of pretrial bail decisions, a risk assessment tool may be racially discriminatory if it recommends white defendants be released before trial at a higher rate than Black defendants with equal risk of pretrial misconduct.

Bringing the theory of algorithmic discrimination to data, however, is often hampered by a fundamental selection challenge. Data on an individual's qualification for treatment may only be available for individuals endogenously selected by an existing human or algorithmic decision-maker. In the pretrial setting, this "selective labels problem" arises because pretrial misconduct potential is only revealed among the defendants that a judge endogenously chooses to release before trial (Kleinberg et al. 2018, Lakkaraju et al. 2017). Such selection can both introduce bias in algorithmic predictions and complicate the measurement of algorithmic discrimination.

This paper shows how discrimination in algorithmic predictions can be measured in the context of pretrial bail decisions, extending methods we develop in Arnold, Dobbie, and Hull (2020) to measure discrimination in judicial decisions. We first show how the key selection problem is solved by estimating four race-specific parameters: the average pretrial misconduct potential in the population of white and Black defendants, along with the race-specific covariances of misconduct potential and algorithmic recommendations. In Arnold, Dobbie, and Hull (2020), we show how the average misconduct potential moments can be estimated and used to measure judge discrimination by extrapolating variation across quasi-randomly assigned bail judges. Here, we show how the race-specific covariances of misconduct potential and algorithmic recommendations can be similarly estimated and used to measure algorithmic discrimination. We illustrate our approach using data from the NYC pretrial system.

I. Empirical Framework

We consider a binary classification problem in which a population of individuals i is differentiated by their race $R_i \in \{w, b\}$ (here, either white or Black) and a latent variable $Y_i^* \in \{0,1\}$ that indicates their qualification for a binary treatment. In the pretrial context, $Y_i^* = 1$ indicates that defendant i would engage in pretrial misconduct (e.g., fail to appear in court or be rearrested for a new crime) if she were released before trial. We suppose an algorithm attempts to predict individual qualification from a vector of observables $\mathbf{X}_i$ and returns a binary treatment recommendation $T_i \in \{0,1\}$. In the pretrial context, the algorithm may use a wide range of defendant and case characteristics to predict pretrial misconduct potential and recommend pretrial release ($T_i = 1$) for defendants with low predicted risk.

Building on Arnold, Dobbie, and Hull (2020), we measure racial discrimination in the

*Arnold: University of California, San Diego (email: daarnold@ucsd.edu); Dobbie: Harvard Kennedy School and NBER (email: will_dobbie@hks.harvard.edu); Hull: University of Chicago and NBER (email: hull@uchicago.edu). Thanks go to our discussant, Avi Feller, and to Ashesh Rambachan and Jonathan Roth for insightful comments. Jeremy Albright and Alexia Olaizola provided excellent research assistance. The data we analyze are provided by the New York State Division of Criminal Justice Services (DCJS) and the Office of Court Administration (OCA). The opinions, findings, and conclusions expressed in this paper are those of the authors and not those of DCJS or OCA. Neither New York State, DCJS, or OCA assumes liability for its contents or use thereof.

† Go to https://doi.org/10.1257/pandp.20211080 to visit the article page for additional materials and author disclosure statement(s).

algorithmic recommendations by the implied treatment disparity among equally qualified white and Black individuals:

$$(1)\quad \Delta = E[E[T_i \mid R_i = w, Y_i^*] - E[T_i \mid R_i = b, Y_i^*]].$$

The inner difference in Δ compares the average recommendation T_i for white and Black individuals, holding fixed their qualification Y_i^*. The outer expectation averages this comparison over the marginal qualification distribution. We say that there is algorithmic discrimination against Black individuals when $\Delta > 0$, that there is algorithmic discrimination against white individuals when $\Delta < 0$, and that there is no white/Black algorithmic discrimination when $\Delta = 0$. In the pretrial context, $\Delta > 0$ means that the algorithm recommends that white defendants be released at a higher rate than Black defendants with equal misconduct potential, on average.

Our definition of algorithmic discrimination relates to the idea of "conditional procedure accuracy equality" or "equalized odds" in the computer science literature (Zafar et al. 2017, Berk et al. 2018). In the language of binary classification problems, this condition imposes the equality of true- and false-negative rates across race. Here, Δ is a weighted average of racial disparities in true-negative rates $\delta_r^T = \Pr(T_i = 1 \mid Y_i^* = 0, R_i = r)$ and false-negative rates $\delta_r^F = \Pr(T_i = 1 \mid Y_i^* = 1, R_i = r)$, where we interpret $Y_i^* = 1$ as an adverse state:

$$(2)\quad \Delta = (\delta_w^T - \delta_b^T)(1 - \bar{\mu}) + (\delta_w^F - \delta_b^F)\bar{\mu},$$

with weights given by the average qualification rate in the population, $\bar{\mu} = E[Y_i^*]$.[1]

This measure of algorithmic discrimination also aligns with the proposed definition of labor market discrimination in Aigner and Cain (1977), which compares the treatment of white and Black workers with the same objective level of productivity. We analogously compare the recommended release rates of white and Black defendants with the same objective potential for pretrial misconduct, Y_i^*. We show in Arnold, Dobbie, and Hull (2020) that measures like Δ capture a broad notion of discrimination arising from both accurate statistical discrimination and racially biased preferences or beliefs. We also show that $\Delta \neq 0$ can arise either because release decisions are directly based on race (i.e., R_i is included in the algorithmic input $\mathbf{X}_i$) or because release decisions are based on observable characteristics that are correlated with race (i.e., variables correlated with R_i are included in the algorithm's feature set $\mathbf{X}_i$).[2]

Estimating Δ is challenging when individual qualification Y_i^* is not directly observed. Often we observe a selected outcome $Y_i = D_i Y_i^*$, where $D_i \in \{0,1\}$ indicates the treatment decision of an existing human or algorithmic decision-maker. In the context of bail decisions, for example, pretrial misconduct potential Y_i^* is only observed among the defendants who are selected by a judge for release ($D_i = 1$). Individuals who are detained before trial ($D_i = 0$) cannot engage in pretrial misconduct, and so $Y_i = 0$. Such endogenous selection may both introduce bias in the algorithmic predictions and confound attempts to measure racial discrimination in the algorithmic recommendations.

Our approach to estimating Δ proceeds in two steps. We first show that the challenge of selectively observed qualification reduces to a challenge of estimating four moments that capture the average qualification rate for each race and how qualification covaries with the algorithmic recommendations within race. Specifically, the true- and false-negative rates that enter Δ can be written:

$$(3)\quad \delta_r^T = \frac{E[T_i(1 - Y_i^*) \mid R_i = r]}{E[(1 - Y_i^*) \mid R_i = r]} = \frac{E[T_i \mid R_i = r] - \rho_r}{1 - \mu_r}$$

and

$$(4)\quad \delta_r^F = \frac{E[T_i Y_i^* \mid R_i = r]}{E[Y_i^* \mid R_i = r]} = \frac{\rho_r}{\mu_r},$$

[1] Other notions of algorithmic fairness include the racial equality of only true-negative rates (Hardt, Price, and Srebro 2016) and the racial equality of both positive and negative predictive values (Zafar et al. 2017). We show in the online Appendix how our framework can be used to bring these alternative measures to data; see Kleinberg, Mullainathan, and Raghavan (2017) for a discussion of inherent trade-offs between them.

[2] A finding of $\Delta \neq 0$ may indicate unlawful discrimination in many settings. For example, Title VII of the 1964 Civil Rights Acts prohibits employment decisions that have a disparate impact by race. In many other contexts, including bail decisions, the Equal Protection Clause of the Fourteenth Amendment prohibits the intentional unequal treatment of equally qualified white and Black individuals (Yang and Dobbie 2020).

where $\mu_r = E[Y_i^* | R_i = r]$ denotes the average qualification rate among race-r individuals and $\rho_r = E[T_i Y_i^* | R_i = r]$ denotes the race-specific second moment of algorithmic recommendations and individual qualification. The weights in Δ can further be written with $\bar{\mu} = \mu_w P_w + \mu_b P_b$, where $P_r = \Pr(R_i = r)$. Since these racial shares and the race-specific average recommendation $E[T_i | R_i = r]$ are directly estimable, these expressions show that the missing information in Δ is the four race-specific parameters in $\theta = \{\mu_w, \mu_b, \rho_w, \rho_b\}$. Algorithmic discrimination can thus be measured by estimating these four parameters, avoiding the need to measure and condition on each individual's qualification directly.

We next show how the four key parameters in θ (and thus algorithmic discrimination Δ) can be estimated by extrapolating reduced-form variation across as-good-as-randomly assigned decision-makers, such as bail judges in the pretrial setting. Under random assignment, each judge j makes treatment decisions D_{ij} among a comparable group of individuals i of each race. We can therefore estimate a series of judge-specific misconduct rates among the defendants of each race that a judge releases before trial, $\tilde{\mu}_{jR_i} \equiv E[Y_i | D_{ij} = 1, R_i] = E[Y_i^* | D_{ij} = 1, R_i]$, as well as the judges' race-specific release rates $\pi_{jR_i} \equiv \Pr(D_{ij} = 1 | R_i)$. In Arnold, Dobbie, and Hull (2020), we show how estimates of the race-specific average misconduct risk in these differentially selected samples can be extrapolated toward judges with high release rates to estimate the unselected average misconduct risk parameter $E[Y_i^* | R_i] = \mu_{R_i}$. Our insight here is that the same logic can be applied to estimate the unselected second moments, ρ_{R_i}, by estimating and extrapolating judge-specific released second moments, $\tilde{\rho}_{jR_i} \equiv E[T_i Y_i | D_{ij} = 1, R_i] = E[T_i Y_i^* | D_{ij} = 1, R_i]$, for each race.[3]

To build intuition for our estimation approach, consider a hypothetical "supremely lenient" bail judge j^* who releases nearly all defendants assigned to her of each race. This judge's race-specific release rates are close to one—that is, $\pi_{j^*R_i} \approx 1$—so by quasi-random assignment her race-specific released first and second moments are both close to the unselected moments: $\tilde{\mu}_{j^*R_i} \approx \mu_{R_i}$ and $\tilde{\rho}_{j^*R_i} \approx \rho_{R_i}$. The decisions of a supremely lenient and quasi-randomly assigned judge can therefore be used to estimate the parameters that enter our discrimination measure Δ. In the absence of a supremely lenient judge, these parameters can instead be extrapolated from the variation in $\tilde{\mu}_{jR_i}$ and $\tilde{\rho}_{jR_i}$ across quasi-randomly assigned judges j with high release rates. This approach is analogous to a standard regression discontinuity design, in which average potential outcomes are extrapolated to a treatment cutoff from nearby observations.[4] Here, selected moments are extrapolated from quasi-randomly assigned judges to the release rate cutoff of one to estimate unselected moments for each race. Estimates may, for example, come from the vertical intercept of linear, quadratic, or local linear regressions of the race-specific selected moment estimates on race-specific release rate estimates. Such extrapolations can be conducted flexibly without a model of judge decision-making.

II. Results

We apply our framework to measure algorithmic discrimination in the NYC pretrial system, one of the largest in the country. Our analysis is based on the universe of NYC arraignments made between November 1, 2008, and November 1, 2013. This sample consists of 595,186 cases assigned to one of 268 bail judges; we describe it in Arnold, Dobbie, and Hull (2020) and verify the key assumption of quasi-random assignment of judges to cases conditional on court and time effects. In this setting, an individual's qualification Y_i^* is her potential for pretrial misconduct (either a failure to appear in court or being arrested for a new crime), and D_i indicates whether or not individual i was released before trial. Observed pretrial misconduct is $Y_i = D_i Y_i^*$.

[3] This second set of extrapolations is not needed to estimate discrimination in a judge's own decisions, as in Arnold, Dobbie, and Hull (2020), since if $T_i = D_{ij}$, then $\rho_{R_i} = E[D_{ij} Y_i^* | R_i] = E[Y_i | D_{ij} = 1, R_i] \Pr(D_{ij} | R_i)$ is directly estimable for each judge j.

[4] Formally, this approach draws on recent advances in average treatment effect extrapolation with multiple discrete instruments (Brinch, Mogstad, and Wiswall 2017; Hull 2020) and a classic literature on identification "at infinity" in sample selection models (Heckman 1990, Andrews and Schafgans 1998).

Our baseline analysis estimates racial discrimination in algorithmic release recommendations that are based on machine learning predictions of pretrial misconduct potential. The predictions come from a gradient boosted decision tree estimated in the sample of released ($D_i = 1$) defendants, following Kleinberg et al. (2018). The features $\mathbf{X}_i$ include a number of characteristics of the current offense and prior criminal history but exclude demographic variables such as race, ethnicity, and gender. We form release recommendations of $T_i = \mathbf{1}\left[\hat{Y}_i < \tau\right]$ in the full sample of defendants, where $\hat{Y}_i$ is the algorithmic risk prediction of each defendant i and τ is a risk threshold. Our benchmark analysis sets τ to equalize the recommended average release rate $E\left[T_i\right]$ and the actual NYC release rate $E\left[D_i\right]$, though we explore a range of thresholds. The online Appendix gives further details on how we fit the algorithm.

Panel A of Figure 1 shows our extrapolation-based estimation of the key race-specific second moments ρ_w and ρ_b. We plot judge- and race-specific second moments of pretrial misconduct potential Y_i^* and algorithmic recommendations T_i among released white and Black defendants $\tilde{\rho}_{jR_i}$, along with race-specific local linear lines of best fit. We adjust these estimates by court and time fixed effects, which capture the level of quasi-random judge assignment. The vertical intercepts of the two lines in panel A of Figure 1, at one, are our estimates of the unconditional second moments. These estimates for white and Black defendants are similar, at $\rho_w = 0.226$ and $\rho_b = 0.213$. We estimate first moments of $\mu_w = 0.346$ and $\mu_b = 0.436$ by an analogous procedure in Arnold, Dobbie, and Hull (2020) with a corresponding visualization. As we show in the online Appendix, these four moments imply a stronger correlation between true misconduct potential and algorithmic release recommendations for Black defendants than for white defendants. The online Appendix also gives more detail on the parameter estimation procedure.

Panel B of Figure 1 uses the four extrapolated moment estimates to estimate algorithmic discrimination Δ. At the average release rate in New York City (73 percent), the algorithm yields a 7.9 percentage point disparity in the recommended release rates of white and Black defendants with the same potential for pretrial misconduct. The figure shows that this conditional disparity is a large share (76.0 percent) of the unadjusted release rate disparity in algorithmic recommendations (10.4 percentage points) and that algorithmic discrimination is found over a wide range of hypothetical release rates. We only fail to find a statistically significant level of algorithmic discrimination at very high thresholds, when the algorithm recommends releasing essentially all defendants.

Figure A1 in the online Appendix shows that this finding of algorithmic discrimination is not driven by the specific machine learning algorithm we use to predict pretrial misconduct risk. We obtain similar estimates of the second moments ρ_w and ρ_b, and correspondingly similar estimates of algorithmic discrimination Δ, using simpler regression-based predictions of pretrial misconduct risk inspired by a widely used pretrial risk assessment tool. At the baseline release rate of 73 percent, we find a 6.7 percentage point disparity in the recommended release rates of white and Black defendants with the same potential for pretrial misconduct, which is again a large share (73.6 percent) of the unadjusted release rate disparity in algorithmic recommendations.

The online Appendix further shows how our estimates of the parameters in θ can be used to compute alternative measures of algorithmic discrimination and decompose our Δ estimates into racial disparities in the algorithm's true- and false-negative rates δ_r^T and δ_r^F. While both disparities are positive at the baseline release rate, only the disparity in false-negative rates (i.e., the release rate disparity between white and Black defendants with misconduct potential) is statistically distinguishable from zero at conventional levels. This suggests that an alternative measure of "inequality of opportunity" (Hardt, Price, and Srebro 2016), based only on the disparity in true-negative misconduct rates, may fail to detect a broader notion of discrimination.

Finally, in the online Appendix, we compare our selection-corrected measure of algorithmic discrimination with a more naïve estimate computed on the selected sample of released defendants. This comparison reveals the extent of confounding by selective labels. At the baseline NYC release rate, the selected estimate is lower by 1.2 percentage points, a difference just at the margin of conventional statistical significance levels. Thus, while in theory the selective labels problem can induce bias in observable measures of algorithmic discrimination, we find

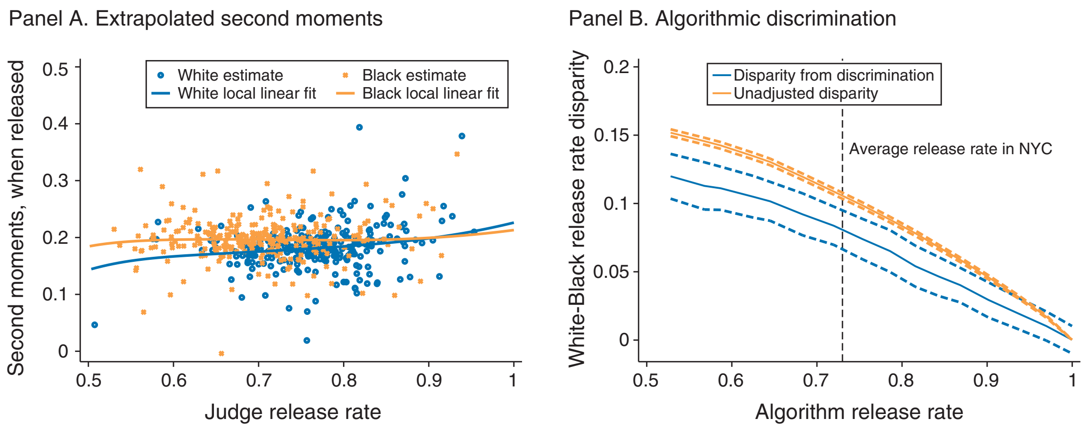

Figure 1. Estimating Algorithmic Discrimination

Notes: Panel A plots race-specific release rates for the 268 judges in our sample against race-specific second moments of pretrial misconduct and algorithmic recommendations among released defendants. All estimates adjust for court-by-time fixed effects (the level at which judges are as-good-as-randomly assigned), and recommendations are calibrated to the average release rate in New York City. The two curves plot the fitted values of race-specific local linear regressions that inverse-weight by the variance of the estimated released second moments and use a Gaussian kernel with race-specific rule-of-thumb bandwidths. Panel B plots estimates of algorithmic discrimination and unconditional racial disparities in algorithmic recommendations for different average release rates. The discrimination estimates use the extrapolated second-moment estimates from panel A and corresponding estimates of the first moments from Arnold, Dobbie, and Hull (2020). Dashed lines indicate pointwise 95 percent confidence intervals obtained from a bootstrapping procedure. See the online Appendix for details.

by computing Δ in this setting that the scope for such bias is small.

III. Conclusion

Algorithmic discrimination is an increasingly widespread concern in many settings, but its measurement is often hampered by a fundamental selection challenge. We show that this challenge can be overcome by estimating four race-specific parameters involving algorithmic recommendations and an individual's selectively observed qualification. We further show that these parameters can be estimated by extrapolating variation across quasi-randomly assigned decision-makers. We illustrate our approach in the NYC pretrial setting, where we find significant discrimination in algorithmic release recommendations that do not directly use information on defendant race. This discrimination persists across a wide range of recommendations and in both a sophisticated machine learning algorithm and simpler regression-based predictions. Comparing our discrimination estimates to more naive measures, we find minimal scope for selection bias in this setting.

We conclude by noting that the methods we develop to study racial discrimination in algorithmic bail decisions may prove useful for measuring unfairness in several other high-stakes settings, both within and outside of the criminal justice system. One key requirement is the quasi-random assignment of decision-makers, such as judges, police officers, employers, government benefits examiners, or medical providers. A second requirement is that an individual's qualification for treatment is measurable among a subset of individuals that the decision-maker endogenously selects. By mapping these settings to the quasi-experimental approach in this paper, researchers can overcome the fundamental selection challenge in bringing a large theoretical literature on algorithmic fairness to data.

REFERENCES

Aigner, Dennis, and Glen G. Cain. 1977. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review* 30 (2): 175–87.

Andrews, Donald W.K., and Marcia M.A. Schafgans. 1998. "Semiparametric Estimation of the Intercept of a Sample Selection Model." *Review of Economic Studies* 65 (3): 497–517.

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. "Machine Bias." *ProPublica*, May 23. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Arnold, David, Will Dobbie, and Peter Hull. 2020. "Measuring Racial Discrimination in Bail Decisions." NBER Working Paper 26999.

Berk, Richard, Hoda Heidari, Shahin Jabbari, Michael Kearns, and Aaron Roth. 2018. "Fairness in Criminal Justice Risk Assessments: The State of the Art." *Sociological Methods & Research* 50 (1): 3–44.

Brinch, Christian N., Magne Mogstad, and Matthew Wiswall. 2017. "Beyond LATE with a Discrete Instrument." *Journal of Political Economy* 125 (4): 985–1039.

Hardt, Moritz, Eric Price, and Nathan Srebro. 2016. "Equality of Opportunity in Supervised Learning." In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, edited by Daniel D. Lee, Ulrike von Luxburg, Roman Garnett, Masashi Sugiyama, and Isabelle Guyon, 3323–31. Red Hook, NY: Curran Associates, Inc.

Heckman, James J. 1990. "Varieties of Selection Bias." *American Economic Review* 80 (2): 313–18.

Hull, Peter. 2020. "Estimating Hospital Quality with Quasi-Experimental Data." Unpublished.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2018. "Human Decisions and Machine Predictions." *Quarterly Journal of Economics* 133 (1): 237–93.

Kleinberg, Jon, Sendhil Mullainathan, and Manish Raghavan. 2017. "Inherent Trade-Offs in Algorithmic Fairness." *Proceedings of Innovations in Theoretical Computer Science*.

Lakkaraju, Himabindu, Jon Kleinberg, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2017. "The Selective Labels Problem: Evaluating Algorithmic Predictions in the Presence of Unobservables." In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 275–84. New York: Association for Computing Machinery.

Yang, Crystal, and Will Dobbie. 2020. "Equal Protection under Algorithms: A New Statistical and Legal Framework." *Michigan Law Review* 119 (1): 291–396.

Zafar, Muhammad Bilal, Isabel Valera, Manuel Gomez Rodriguez, and Krishna P. Gummadi. 2017. "Fairness beyond Disparate Treatment & Disparate Impact: Learning Classification without Disparate Mistreatment." In *Proceedings of the 26th International Conference on World Wide Web*, 1171–80. Geneva: International World Wide Web Conferences Steering Committee.

This article has been cited by:

1. E Jason Baron, Joseph J Doyle, Natalia Emanuel, Peter Hull, Joseph Ryan. 2024. Discrimination in Multiphase Systems: Evidence from Child Protection. *The Quarterly Journal of Economics* **139**:3, 1611-1664. [Crossref]
2. Joshua Grossman, Julian Nyarko, Sharad Goel. 2024. Reconciling Legal and Empirical Conceptions of Disparate Impact: An Analysis of Police Stops Across California. *Journal of Law and Empirical Analysis* **1**:1. . [Crossref]
3. William M. Casey, Jennifer E. Copp, Stephen Demuth. 2024. Disparities in the Pretrial Process: Race, Ethnicity, and Citizenship. *Justice Quarterly* **41**:2, 268-290. [Crossref]
4. Louis Lippens. 2024. Computer says 'no': Exploring systemic bias in ChatGPT using an audit approach. *Computers in Human Behavior: Artificial Humans* **2**:1, 100054. [Crossref]
5. Adrien Raizonville, Xavier Lambin. 2024. From black box to glass box: algorithmic explainability as a strategic decision. *SSRN Electronic Journal* **11**. . [Crossref]
6. Ozge Demirci. 2024. Can Gender-Blind Algorithmic Pricing Eliminate the Gender Gap?. *SSRN Electronic Journal* **133**. . [Crossref]
7. E. Jason Baron, Richard Lombardo, Joseph Ryan, Jeongsoo Suh, Quitze Valenzuela-Stookey. 2024. Mechanism Reform: An Application to Child Welfare. *SSRN Electronic Journal* **93**. . [Crossref]
8. David Arnold, Will Dobbie, Peter Hull. 2024. Building Non-Discriminatory Algorithms in Selected Data. *SSRN Electronic Journal* **133**. . [Crossref]
9. Vanessa G. Perry, Kirsten Martin, Ann Schnare. 2023. Algorithms for All: Can AI in the Mortgage Market Expand Access to Homeownership?. *AI* **4**:4, 888-903. [Crossref]
10. Anirban Basu. 2023. Use of race in clinical algorithms. *Science Advances* **9**:21. . [Crossref]
11. E. Jason Baron, Joseph John Doyle, Natalia Emanuel, Peter Hull, Joseph Ryan. 2023. Racial Discrimination in Child Protection. *SSRN Electronic Journal* **30981**. . [Crossref]
12. Fabrizio Gilardi. Digital Technology, Politics, and Policy-Making **7**, . [Crossref]
13. J. Aislinn Bohren, Peter Hull, Alex Imas. 2022. Systemic Discrimination: Theory and Measurement. *SSRN Electronic Journal* **13**. . [Crossref]
14. Vanessa Gail Perry, Kirsten Martin. 2022. Algorithms for All: Has Digitalization in the Mortgage Market Expanded Access to Homeownership?. *SSRN Electronic Journal* **62**. . [Crossref]
15. Brendan O'Flaherty, Rajiv Sethi, Morgan Williams Jr. 2022. Profiling in Criminal Justice. *SSRN Electronic Journal* **111**. . [Crossref]
16. Faizaan Teizoon Kisat. 2021. Loan Officers, Algorithms, & Credit Outcomes: Experimental Evidence from Pakistan. *SSRN Electronic Journal* **30**. . [Crossref]
17. Annie Liang, Jay Lu, Xiaosheng Mu. 2021. Algorithmic Design: Fairness Versus Accuracy. *SSRN Electronic Journal* **133**. . [Crossref]
18. Megan Stevenson, Jennifer L. Doleac. 2019. Algorithmic Risk Assessment in the Hands of Humans. *SSRN Electronic Journal* **133**. . [Crossref]