# Equal Merit Does Not Imply Equality: Discrimination at Equilibrium in a Hiring Market with Symmetric Agents

**Serafina Kamp, Benjamin Fish**

University of Michigan Department of Computer Science Engineering
serafibk@umich.edu, benfish@umich.edu

**Abstract**

Machine learning has grown in popularity to help assign resources and make decisions about users, which can result in discrimination. This includes hiring markets, where employers have increasingly been interested in using automated tools to help hire candidates. In response, there has been significant effort to understand and mitigate the sources of discrimination in these tools. However, previous work has largely assumed that discrimination, in any area of ML, is the result of some initial unequal distribution of resources across groups: One group is on average less qualified, there is less training data for one group, or the classifier is less accurate on one group, etc. However, recent work have suggested that there are other sources of discrimination, such as relational inequality, that are notably non-distributional. First, we show consensus in strategy choice is a non-distributional source of inequality at equilibrium in games: We provide subgame perfect equilibria in a simple sequential model of a hiring market with Rubinstein-style bargaining between firms and candidates that exhibits asymmetric wages resulting from differences in agents' threat strategies during bargaining. Second, we give an initial analysis of how agents could learn such strategies via convergence of an online learning algorithm to asymmetric equilibria. Ultimately, this work motivates the further study of endogenous, possibly non-distributional, mechanisms of inequality in ML.

## 1 Introduction

Machine learning (ML) algorithms are marketed to make more efficient and data-driven decisions that improve human decision-making. These automated tools have become increasingly popular in recommendation systems, classification problems, and resource assignment amongst other areas (Berk et al. 2021; Calders and Verwer 2010; Kamiran and Calders 2012). With their growing popularity, it has become clear that these algorithms can engender discrimination against individuals or demographic groups (Angwin et al. 2016; Eubanks 2018; Miller 2015; O'Neil 2017).

In response, there is a growing research area on formalizing, measuring, and mitigating discrimination (Berk et al. 2021; Chouldechova and Roth 2018; Friedler et al. 2019; Mehrabi et al. 2021). Here, discrimination is typically taken to mean the *unequal distribution of resources among different groups* defined by sensitive attributes, such as race or gender (Arrow 2015; Barocas, Hardt, and Narayanan 2017; Hardt, Price, and Srebro 2016). Many fairness metrics have been developed across ML and economics to measure such resource differences. They cover group-level, individual-level, and causal measurements alongside notions of taste-based and statistical discrimination (Arrow 2015; Dwork et al. 2012; Hardt, Price, and Srebro 2016; Kamiran and Calders 2012; Kusner et al. 2017; Phelps 1972). Some examples of resources that could be distributed unequally include merit (i.e., relevant features or skills), data sets (i.e., distribution of training or ground truth labels, costs in acquiring data, or elasticity in demand for data), and opportunities (i.e., job offers, loan approvals, housing bids, or generally any ML prediction). In general, a fairness metric is satisfied if the corresponding resources it measures are approximately equal among groups or individuals of interest, a view of equality commonly referred to as distributive equality (Dworkin 2002). If there is no such inequality in resources – everyone to be classified has equal merit – then under this view we should expect no discrimination.

However, sources of discrimination in general, let alone in ML, need not be distributional (Abbasi et al. 2019; Birhane 2021; Elford 2017; Green 2022; Hoffmann 2019; Fish and Stark 2022; Kasirzadeh 2022; Schemmel 2012). Various other kinds of inequality, including relational (Anderson 1999), representational (Shelby et al. 2023; West, Whittaker, and Crawford 2019), status (Ridgeway 2014), and power (Kasy and Abebe 2021) have been proposed. Many of these are very difficult to measure, like relational equality, or still require existing distributional inequality, like when representational inequalities are caused by disparities in training data that an algorithm reproduces (e.g. searches for 'CEO' in image search engines (Kay, Matuszek, and Munson 2015; Lam et al. 2018)).

Given the recent work that recognizes the limitations of distributive equality (Birhane 2021; Fish and Stark 2022; Kasirzadeh 2022; Kasy and Abebe 2021), in this paper, we provide a non-distributional source of inequality that can be measured via a game-theoretic model: the strategies of the agents at equilibrium.

Throughout this paper, we use the setting of a hiring market where agents are a priori resource-symmetric: we assume no exogenous differences between agents in our market, including merit. Hiring markets are areas where ML is increasingly popular (Bogen and Rieke 2018; Raghavan et al. 2020), yet it has become clear that ML can perpetuate or engender discrimination here (Cavounidis and Lang 2015; Dastin 2018; Hu and Chen 2017). Further, in hiring markets, economic prospects are tied to both merit (i.e., having sufficient skills required to perform a job or being otherwise deemed by other agents to be deserving of a job) and how agents interact with others in the market. So, by assuming a priori resource-symmetric agents, we can isolate agent interactions via strategy choice as the only source of inequality at equilibrium.

We provide a model of a hiring market as a bargaining game to show the existence of asymmetric wages at equilibrium among agents with equal merit. Notably, by finding outcomes at equilibrium, the inequality will persist as the best-response for each agent in our market. Our bargaining game captures the wage negotiation process, an important process where agent interactions determine economic outcomes. We are interested in identifying a *source* of inequality at equilibrium in hiring markets, and wage negotiation must happen in any hiring context unless the firm and candidate are *a priori unequal* and the candidate is forced to accept any wage. We consider a two-sided market where firms are looking to fill a job position and candidates are applying for these positions. Once a firm and candidate are matched (i.e. the candidate applies for the firm's position), they participate in wage negotiation via a Rubinstein-style bargaining game (Osborne and Rubinstein 1990) over the surplus generated by the candidate's employment.

The asymmetric outcomes we find at equilibrium are driven by consensus among agents about their choice of strategy. Consensus over choice of strategy can be non-distributional when the consensus is not driven by exogenous disparities in merit or other such factors. As an example, consider social norms in negotiation: Women may follow social norms where they do not negotiate for higher salaries, as compared to men, which contributes to observed pay gaps between these genders (Ren, Xiu, and Hietapelto 2022). Here, the social norms may act as a non-distributional source of the observed discriminatory pay gaps and, further, adherence to discriminatory social norms may cause or exacerbate a gender social hierarchy (Ferrant, Pesando, and Nowacka 2014). Another example could be in a consensus among firms to use an algorithm for price setting (Weber 2023). If enough firms believe that the algorithm offers the true market value of their product, say, rental housing, then firms will all list the price given by the algorithm for their units and it will become the market value.

Therefore, it is crucial to study the conditions under which strategy consensus drives inequality at equilibrium, including via the choice to use algorithmic decisions. This paper addresses this by (1) demonstrating endogenous consensus about outside options among equal merit agents which creates wage inequality at equilibrium in a bargaining game and (2) showing initial results for how an online learning algorithm would converge to such strategies.

Section 3 gives our first result. We create a resource-symmetric market with the following assumptions :1) every pair of agents split equal surplus they are equally entitled to, ensuring equal merit; 2) every agent has an endogenous outside option so that every agent has an a priori chance to have bargaining power; 3) the game itself is symmetric so that no agent has a priori more bargaining power than any other, e.g. by getting to propose an offer first; 4) all agents are identical up to type: The only difference between agents are their strategies. We show that differences in credible threats can be a mechanism for non-distributional discrimination by showing that there exist multiple strategy profiles in subgame perfect equilibrium (SPE) in the bargaining game with asymmetric payoffs. A set of strategies is in SPE when there are no beneficial deviations from the strategies on or off the equilibrium path.[1] As such, all threats used by agents must be credible, i.e., it must be rational for them to take any action specified in their strategy. We highlight two kinds of asymmetric outcomes at equilibrium: One where a group of candidates gets more of the surplus than another group of candidates and another where the firms get more of the surplus than the candidates.

Section 4 gives our second result. Our interest here is in understanding when it is possible for online learning algorithms to converge to equilibrium strategies in a bargaining game. In this paper, we initiate a study of how learning algorithms can learn the kinds of strategies that constitute non-distributional inequality. As is typical in learning in games, we model strategy learning in bargaining games as an online optimization problem. We use a simplified version of the above bargaining game to one that has only two players and finitely many rounds. This approach does not require access to specific wage setting algorithms used in practice, which are often proprietary and may be subject to frequent changes as the technology evolves. Unfortunately, these bargaining games are not convex and learning is hard in general for non-convex games, but we are still able to show that Follow the Regularized Leader (FTRL) is no regret for our bargaining game. [2] This motivates why agents might use FTRL to learn strategies. We then show the existence of parameter settings of $\ell_1$-regularized FTRL where the algorithm converges to Nash equilibria (NE) with unequal outcomes. Here, the initialization of the algorithm and the choice of how ties are broken in the optimization problem dictate when agents converge to asymmetric NE. Crucially, our agents still have equal merit after learning how to bargain, yet we show it is possible for agents to receive significantly different wages.

In Section 2 we review related work. Sections 3 and 4 give our main results. This paper establishes the importance of understanding non-distributional sources of inequality, particularly from consensus in strategy choice in games. We highlight additional opportunities for future work throughout. All proofs for this paper can be found in the appendix of the full version (Kamp and Fish 2024).

[1] See for example Tadelis (2013) for an introduction to the game theoretic concepts used in this work.

[2] Strategies learned from a no regret algorithm can be considered the "ground truth" in online optimization.

# 2 Related Work

## 2.1 Discrimination Against Agents of Equal Merit

In economics, statistical discrimination demonstrates how initially equal levels of merit *become unequal at equilibrium* because one group of workers finds they are less likely to be hired by a firm, so they end up investing less in relevant skill sets (Phelps 1972). Further, taste-based discrimination allows for accurate but discriminatory outcomes among groups of workers with equal merit due to exogenous differences in a firm's utility function for hiring each group due to direct animus (Arrow 2015). Instead, we show that discrimination can still arise among agents of equal skills and symmetric utility functions.

Demographic parity is a fairness metric that may find an accurate algorithm to be discriminatory (Kamiran and Calders 2012). However, this metric randomly assigns resources (i.e., positive predictions) so that they are equal among desired groups and *the mechanism that created the inequality is ignored*. Other related notions of bias in ML are *historical bias* (Hellström, Dignum, and Bensch 2020; Mehrabi et al. 2021; Rajkomar et al. 2018; Roselli, Matthews, and Talagala 2019), *feedback loops* (Adam et al. 2020; Lum and Isaac 2016; Malik 2020), *label bias* (Dai and Brown 2020; Jiang and Nachum 2020), and *representational harms* (Abbasi et al. 2019; Buddemeyer, Walker, and Alikhani 2021; Cheng et al. 2023; Curry, Robertson, and Rieser 2020) . Each of these notions is either describing a distributional source or describes a mechanism other than strategy choice to create asymmetric outcomes among agents of equal merit.

## 2.2 Discrimination Analysis in Markets

We build off the work of Fish and Stark (2022) on non-distributional sources of discrimination in a bargaining game, but we provide fully symmetric agents, a stronger notion of equilibrium, and a different non-distributional mechanism: Rather than beliefs about outside options alone, we also look at an agent's ability to make threats during bargaining.

There are several works which highlight undesirable outcomes of bargaining models, including disparities in wages or employment rates (Cavounidis and Lang 2015; Fang and Moro 2011; Fernandez and Glazer 1989; Hu and Chen 2017; Lax 1989; Ulph and Ulph 1998). However, these works assume exogenous distributional inequality in the form of disparity in skills, costs across workers, or bargaining power while we show that discrimination arises *even when we control for exogenous differences*.

There are several works that consider bargaining games with multiple equilibria (Dwork et al. 2024; Hyde 1997; Shaked et al. 1994), as we do here. Multiple equilibria are useful for demonstrating the existence of discrimination, because it enables the possibility of both an unfair outcome and a better alternative that's still incentive-compatible. However, these works all have some asymmetry between agents including exogenous costs, who gets to propose first, and outside option values (Dwork et al. 2024; Hyde 1997; Shaked et al. 1994). Our work directly extends the scenario modeled by Ponsati and Sákovics (1998) by endogenously modeling an agent's outside option, allowing for more than two agents in our market, and most importantly, imposing symmetry between the bargaining agents. Similarly to us, Agranov, Cotton, and Tergiman (2018) demonstrate the existence of asymmetric outcomes at SPE with initially symmetric agents. However, their model does not include initially equal bargaining power, including outside options.

Finally, Kuksov (2024) also considers the question of how an asymmetric equilibrium arises in a symmetric market based on search costs and uncertainty in payoffs that may be extremely high, but we show that inequality still arises when there is a fixed surplus that agents bargain over.

## 2.3 Critiques of Distributional Equality

There are a few works that consider alternative sources and presentations of discrimination beyond distributive equality. Several works in political philosophy discuss the shortcomings of a purely distributional view of equality and identify relational equality as a way to address the gaps (Anderson 1999; Elford 2017; Schemmel 2012). In the literature on machine learning, Birhane (2021) draws attention to the complexities of fairness questions and discusses how relational ethics might be a useful framework. Kasy and Abebe (2021) similarly recognize the failures of current fairness metrics to capture certain forms of discrimination and they model the ways in which power affects the outcome of an algorithm. Neuhäuser et al. (2023) investigate the effects of homophily preferences during the growth of networks on the visibility of minority populations. However, these works are not focused on modeling particular sources of non-distributional inequality, particularly in markets.

## 2.4 Piecewise Linear Optimization

Our learning setting requires online optimization over piecewise linear utility functions. While there has been lots of work on online optimization for convex functions [3], there has been much less work in online piecewise linear function optimization (Balcan, Dick, and Vitercik 2018; Cohen-Addad and Kanade 2017). This work makes some restrictive assumptions about the adversarial setting. In our work, we focus on bargaining games rather than arbitrary piecewise linear utility functions, but assume an adversary with different assumptions. We show that a simple discretization of our action space leads to a no regret algorithm with respect to the optimal continuous strategy. Kroer and Sandholm (2015) study optimal discretizations of continuous action spaces in extensive form games, but their work does not consider strategies learned under a no-external-regret framework as we do.

# 3 Discriminatory Equilibria via Threats

In this section, we describe a bargaining game where agents are able to use credible threats to set their strategies in a way that produces discriminatory outcomes at equilibrium. This

[3]See Hazan et al. (2016) for an introduction to online convex optimization.

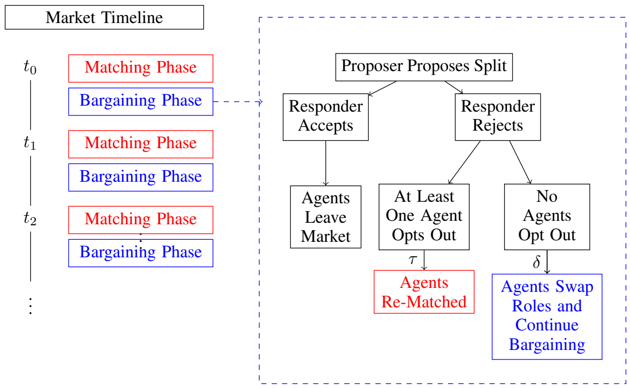


Figure 1: The progression of the market over time is shown on the right with one round of the bargaining phase expanded on the left.

section first sets up our symmetric market, and then characterizes equilibrium outcomes in our first main theorem.

## 3.1 The Market Model

Our market is an extensive-form game where candidates and firms are matched and bargain over the split of the surplus generated as a result of the candidate's employment. To control for exogenous differences between agents, we make the normative assumption that all firms and candidates are equally skilled and entitled to the surplus. The market timeline and the bargaining game can be visualized in Figure 1 and we go through the details below.

In this market, there are discrete time steps where there is a matching and a bargaining phase. During the matching phase, unmatched firms and candidates are matched with some probability which represents a candidate applying for a job with the firm. We say two agents are the same type if they play the same strategy and assume there is a constant probability of matching with a strategy of each type.

Once two agents are matched, they enter the bargaining phase. Here, firms and candidates participate in an extensive form Rubinstein-style bargaining game (Osborne and Rubinstein 1990) to determine the split of the surplus, which we normalize to 1. To ensure agents have a priori equal bargaining power, we insist a newly matched agent gets to propose a split of the surplus first with probability $\frac{1}{2}$ (or the other agent proposes first).

Given agent $i$ is the proposer and agent $j$ the responder, agent $i$ will propose a split of the surplus $(1-y, y)$ for some $y \in [0,1]$. As the responder, agent $j$ can either reject or accept the proposal. If they accept, they both leave the market. If they reject, Both agents have a chance to opt out of the match if the responder rejects the offer (which without loss of generality can happen simultaneously, as in Ponsati and Sákovics (1998)) and otherwise in the next round the proposer becomes the responder and vice versa and a discount factor $\delta$, $0 < \delta < 1$, is incurred. So, if agent $i$'s offer of $y$ is accepted by agent $j$ at round $k \geq 1$, then, agent $i$ gets utility $\delta^k(1-y)$ and agent $j$ gets $\delta^k y$. When either agent chooses to opt out, both agents must pay a cost of $0 \leq \tau \leq 1$ and are sent back into the market pool to be re-matched in some future time. Here, $\tau$ captures the waiting and matching cost incurred when an agent decides to re-enter the market.

The agents' strategies are given as automata with a base state and a threat state. Each state consists of the offer an agent would make in each round as the proposer, the amount they would accept in each round as the responder, and whether they opt out if a proposal is rejected. The agents start using the base state and move to the threat state if their bargaining opponent deviates from the equilibrium path. The strategy table of each agent can be found in the appendix of the full version of this paper.

For convenience, we make a large market assumption: Our market remains a constant size with the same distribution of firm's and candidate's strategies at all time steps. For example, whenever two agents leave the market, we can assume two agents with identical strategies enter in the next time step.

## 3.2 Disparity at Equilibrium

In this section, we will focus on the case of one type of firm and two types of candidates which we will call $c_1$ and $c_2$ candidates. Since we have only two kinds of bargaining matches, let $p$ be the probability of a firm matching with a $c_1$ candidate and otherwise they match with a $c_2$ candidate.

We now state our first main theorem which gives the range of payoffs each type of agent can receive at SPE. Notably, there are SPE where payoffs are asymmetric between firms and candidates and between candidate types.

**Theorem 1.** *If $\tau \leq \frac{\delta^2}{1+\delta}$, then for any $p \in [0,1]$ and any $w_1, w_2 \in [0,1]$ that satisfy*

$$w_k \leq \frac{1}{2}\left(\frac{1+\delta-2\tau}{1-\tau}\right) \text{ for } k \in \{1,2\}, \quad (1)$$

$$w_1 \geq \frac{1}{2}\left(\frac{1-\delta+2\tau(1-p)w_2}{1-\tau p}\right), \quad (2)$$

$$\text{and } w_2 \geq \frac{1}{2}\left(\frac{1-\delta+2\tau p w_1}{1-\tau(1-p)}\right), \quad (3)$$

*there exists an SPE where the firms obtain an expected payoff of $pw_1 + (1-p)w_2$, the $c_1$ candidates get an expected payoff of $1-w_1$ and the $c_2$ candidates get an expected payoff of $1-w_2$ at equilibrium.*

These results hold because there is a range of first round offer values that are immediately accepted as a result of "threats" made in the second round where each first round offer in that range corresponds to an SPE, so long as the threats use *the actual outside option of each agent*. So, all that is required is that agents in the market have a consensus in their strategies about each others' outside options to establish the range of strategy profiles in SPE. Such a consensus drives agent acceptance of a discriminatory outcome at equilibrium despite their equal merit.

Though $c_1$ and $c_2$ candidates need not explicitly agree on each other's outside options through the choice of a strategy, there is a dependence between their outside options because the outside option of the firm relative to one group depends on what the other group is willing to accept from or offer to the firm. Hence, Theorem 1 sets up the bounds of the payoffs to each group of candidates (related to their outside options via $\tau$) in terms of the payoff to the other. Therefore, asymmetric outcomes exist when there are gaps between payoffs that satisfy assumptions (1) - (3) in the theorem statement.

### 3.3 Discussion

To find gaps in payoffs, first suppose $w_2 > w_1$ such that $c_1$ candidates are getting a higher payoff than $c_2$ candidates at equilibrium. Then, we can fix $w_2$ at its upper bound given by assumption (1) and find the difference with the lower bound of $w_1$ given by assumption (2). The largest such gap is given by $\frac{\delta-\tau}{1-\tau p}$. When $\tau = \frac{\delta^2}{1+\delta}$, the gap size grows with $p$ given a fixed $\delta$ and with $\delta$ given a fixed $p$. Notably, a payoff gap persists between $c_1$ and $c_2$ candidates even as $p$ approaches 0. Additionally, even when candidate groups are equally sized (i.e., $p = 0.5$), a larger $\delta$ can cause the gap to be larger than $\frac{1}{2}$ such that candidate groups are getting significantly different splits of the surplus. [4]

In this market, no agent has any advantage over the others in terms of information, power, or merit in the bargaining game at the start and yet, as we have shown, it is possible for agents to choose strategies that are in SPE where one type of candidate receives a greater split of the surplus than the other. As such, this market is susceptible to discrimination via credible threats without any initial asymmetric advantage among any of the agents. Although fairness metrics based on distributive equality could detect this kind of discrimination, they would aim to correct the discrimination by equalizing the split of the surplus. However, our model shows that equal entitlement to the surplus did not prevent the discriminatory outcome such that an intervention to equalize resources may not last.

Now consider another case, where $w_1 = w_2 = \frac{1}{2}(\frac{1+\delta-2\tau}{1-\tau})$ and this creates a payoff gap of $\frac{\delta-\tau}{1-\tau}$ between the firm and both kinds of candidates. Note that this gap grows with $\delta$ and approaches 1 for a fixed $\tau < 0.5$. Here, the firms' strategy does not depend on the candidate type and the candidates receive the same split of the surplus at the end. However, the firms then get more of the surplus than the candidates even though we assumed that the firms and candidates are equally entitled to the surplus, indicating that the firms were able to acquire more bargaining power than the candidates at equilibrium.

At this point, we have described how our non-distributional source of credible threats in a strategy can produce a number of discriminatory outcomes and power advantages at equilibrium, but we do not say how agents might *learn* to play a strategy in one of these equilibria. So, we turn to learning in bargaining games in Section 4.

## 4 Discriminatory Equilibria via Learning

In this section, we initiate a study of online learning in bargaining games. We provide conditions where an online learning algorithm converges to an unequal NE to demonstrate how learning algorithms can produce discriminatory outcomes by learning from the outcome of a repeated game, rather than from distributional inequalities in training data. Demonstrating this does not require the full model introduced in Section 3, so for the sake of simplicity, we use a two-player bargaining game with finitely many rounds of bargaining. We also motivate the use of this algorithm by showing conditions under which the algorithm is no regret. This implies agents may want to employ the algorithm regardless of which algorithm the other agent uses.

### 4.1 Simplified Bargaining Game

In this section, we consider a bargaining game between two agents: $P$ is the agent who proposes first and $R$ is the agent who responds first. Let $\mathcal{G}^{(n)}$ be a bargaining game with $n < \infty$ rounds of bargaining. In each round, a proposer makes an offer in $[0, 1]$ to the responder and the responder simultaneously specifies an acceptance threshold in $[0, 1]$ indicating the minimum offer they would accept. If the agents don't agree (i.e., the offer is less than the acceptance threshold), the proposer becomes the responder and vice versa in the next round. Agents no longer have an outside option, so if no deal is made within $n$ rounds, both agents get 0.

We represent the strategy of each agent $i \in \{P, R\}$ as an $n$-dimensional vector $s_i \in [0, 1]^n$, representing what they would do at each round of bargaining. Let $s_{i,k}$ be the strategy of agent $i$ at round $k$ of bargaining. Since two agents alternate being proposer and responder, if $i$ is the proposer at round $k$, $s_{i,k} \in [0, 1]$ is an offer to agent $j$ (their opponent) and otherwise $s_{i,k} \in [0, 1]$ is an acceptance threshold. Given a strategy profile $(s_P, s_R)$, the utility payoff function to each agent in game $\mathcal{G}^{(n)}$ matches the utility described in Section 3, though agents get 0 utility if no offer is accepted at any round. With this setup, note that $\mathcal{G}^{(1)}$ is the well-studied ultimatum game. [5]

The strategy profiles that are in pure NE are those where $s_{P,k} = s_{R,k}$ for the round $k$ where each agent is getting maximum possible utility, given their opponent's strategy. We will say the *value* of the equilibrium is $s_{P,k}$, i.e., the surplus split that the responder in round $k$ accepts.

### 4.2 Learning to Bargain via FTRL

Learning bargaining strategies in our market is an *online optimization* problem where each agent commits to a strategy at each time step and then learns the utility of their strategy. The standard goal for learning in such settings is to suffer no external regret in the limit, where external regret is the difference in utility between the chosen strategies at each time step and the single best strategy in hindsight given an adversary choosing utility feedback at each time step. An online learning algorithm over a game is *no regret* if its external regret is sublinear in $T$, the number of time steps it runs, for any sequence of adversarial strategies chosen by an opponent (Hazan et al. 2016). This is a very common notion of

[4] Qualitatively similar results hold for markets with $m$ kinds of firms and $n$ kinds of candidates for any finite $m$ and $n$, and these results can be found in the appendix of the full version of this paper.

[5] Note that for $\mathcal{G}^{(1)}$ any strategy profile in NE is also in SPE. See (Debove, Baumard, and André 2016; Tadelis 2013) for studies of the Ultimatum game.

accuracy in online optimization because in general it's not possible to achieve any stronger notion.

Follow the Regularized Leader (FTRL) is a simple algorithm used in many online *convex* optimization problems that achieves no regret in games when, for example, the utility functions are convex and the strategy space is discrete (Shalev-Shwartz et al. 2012). We start by motivating the use of FTRL by showing when it is no regret here as well. However, the utility functions for $\mathcal{G}^{(n)}$ are notably non-convex (or concave), but rather piece-wise linear with jump discontinuities. Moreover, the strategy space is continuous rather than discrete. So we use Algorithm 1, which discretizes the strategy space $[0,1]^n$ before running FTRL. This ensure that the *expected* utility function is convex with respect to mixed strategies over the discretized space. Then, we can use the previous guarantees of FTRL in online convex optimization to give conditions under which Algorithm 1 is a no regret learner, even with respect to the optimal strategy in the continuous strategy space.

Algorithm 1 is parameterized by the number of time steps $T$, a positive integer $D > 1$ for discretizing the strategy space, a learning rate $M$, and $p \geq 1$ (since we use the $\ell_p$ norm as the regularizer). At each time step, agent $i$ receives *full feedback* $u_i^{(t)}$, i.e. a vector indexed by $\mathcal{S}^n$ that gives the payoff had agent $i$ played $s \in \mathcal{S}^n$ given their opponent's actual strategy at time $t$. Then, the FTRL update step is used to update agent $i$'s *mixed* strategy based on the expected utility $i$ gets from their strategy at time $t$.

Even though this algorithm always returns a mixed strategy over a finite space with precision $1/D$, we'd like to show that this Algorithm 1 is a no regret learner with respect to the optimal strategy over the *continuous* space $[0,1]^n$. However, if the adversary is allowed to specify strategies with arbitrary precision while the learner isn't, the adversary can take advantage of the additional precision to ensure the learning agent gets linear regret. So, we assume the adversary also must discretize $[0,1]$ into bins of size at least $\frac{1}{D}$ before choosing their strategy, though they are allowed to choose a different discretization for each of the $n$ rounds of bargaining. Then, the following proposition establishes sufficient conditions for no regret:

**Proposition 3.** *In game $\mathcal{G}^{(n)}$, for any agent $i \in \{P, R\}$, when $D = T$, $M = O(1/\sqrt{T})$, and $p = 2$, Algorithm 1 is no regret with respect to the optimal continuous strategy in hindsight for agent $i$.*

For our NE convergence results, we use $p = 1$ for ease of analysis, and we leave for future work extending our convergence results to other choices of regularizer. The use of the $\ell_1$ regularizer implies that each $w_i^{(t)}$ is a pure strategy, which is not the case for other regularizers. It may also be the case that multiple pure strategies have the same objective value, so in the following analysis, we break ties by choosing the largest tied acceptance threshold for the responder and the largest tied offer for the proposer in each round.

In the following sections we show (1) convergence to NE occurs when both agents learn to bargain via Algorithm 1 in $\mathcal{G}^{(1)}$ and at least in some cases in $\mathcal{G}^{(2)}$ and (2) there exist initial conditions such that agents converge to a NE with *asymmetric payoffs*. Since we have assumed that agents are equally entitled to the surplus here, these results show the importance of understanding how initial conditions and algorithm design choices, like a tie-breaking mechanism, contribute to possibly discriminatory outcomes at NE.

Algorithm 1: Discretized FTRL with $\ell_p$ regularizer

**Input:** $T, M, D, p, i \in \{P, R\}, \mathcal{G}^{(n)}$
$\mathcal{S} \leftarrow \{0, \frac{1}{D}, \frac{2}{D}, \ldots, \frac{D-1}{D}, 1\}$
$w_i^{(1)} \in \Delta(\mathcal{S}^n)$ {*pure* initial strategy}
$\alpha_i \in \Delta(\mathcal{S}^n)$ {*pure* strategy reference point}
**for** $t = 1$ to $T$ **do**
  Play $w_i^{(t)}$ and observe the feedback $u_i^{(t)}$
  Update $w_i$:

$$w_i^{(t+1)} = \underset{w \in \Delta(\mathcal{S}^n)}{\arg\max} \langle w, \sum_{\tau=1}^{t} u_i^{(\tau)} \rangle - \frac{\frac{1}{p}\|w - \alpha_i\|_p^p}{M}$$

**end for**

## 4.3 Convergence to Nash Equilibrium

In this subsection, we show a range of conditions under which Algorithm 1 converges to NE for $n = 1$ and $n = 2$. We use a strong notion of convergence where there exists a time step $t'$ such that the strategy profiles each agent plays are at NE and that for all $t \geq t'$, they remain at that NE.

Theorem 4 first characterizes convergence in $\mathcal{G}^{(1)}$ which reflects the intuition that a responder should accept any nonzero offer, so the convergence value is the lowest offer the responder sees in the first few time steps.

**Theorem 4.** *For any setting of the initial conditions $w_r^{(1)}, w_p^{(1)}, \alpha_r, \alpha_p \in \mathcal{S} \setminus \{0, 1\}$, Algorithm 1, parameterized by $M > 2D$ and $p = 1$, converges to a NE in $\mathcal{G}^{(1)}$ whose value is $\min\{w_{r,1}^{(1)}, w_{p,1}^{(1)}, \alpha_r\}$.*

However, in games with $n \geq 2$, both agents have a chance to propose and agents can make strategic *threats* about how they will act in the second round of bargaining, which the other agent will observe in this full feedback model. Below we show convergence occurs for at least some initial conditions in the game $\mathcal{G}^{(2)}$.

**Theorem 5.** *Suppose the following initial conditions hold:*

$$1 - w_{r,1}^{(1)} \geq \delta w_{r,2}^{(1)},$$
$$w_{p,1}^{(1)} > \delta(1 - w_{p,2}^{(1)}), \text{ and}$$
$$\alpha_{i,1} > w_{j,1}^{(1)}.$$

*Then Algorithm 1, parameterized by $D > \frac{1}{1-\delta}$, learning rate $M > 2D$, and $p = 1$, converges to a NE in $\mathcal{G}^{(2)}$.*

The fact that we can characterize convergence in these cases means learning in games $\mathcal{G}^{(n)}$ for $n \geq 2$ may be of independent interest and we leave full analysis for future work.

### 4.4 Disparity at Equilibrium

We now highlight a range of initial conditions where agents end up with asymmetric payoffs at equilibrium after learning strategies for $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ via Algorithm 1 with $p = 1$. Recall our assumption that all agents are equally entitled to the surplus, so cases with asymmetric payoffs are potentially discriminatory.

First, for $\mathcal{G}^{(1)}$, Theorem 4 characterizes many possible NE convergence values which depend on the initial conditions. Specifically, any case where $\min\{w_{r,1}^{(1)}, w_{p,1}^{(1)}, \alpha_r\} \neq 0.5$ will converge to a NE with asymmetric payoffs. Next, for $\mathcal{G}^{(2)}$, we start by using Theorem 5 to illustrate how agents can converge to almost maximally disparate outcomes by only changing the initial conditions:

**Example 6.** *Suppose both a proposer $f$ and responder $c$ use Algorithm 1, parameterized by $D > 10, M > 2D, p = 1, \delta = 0.9$ to learn in $\mathcal{G}^{(2)}$. If $w_f^{(1)} = (0.5, w_{f,2}^{(1)} > 1 - \frac{0.5}{0.9}), w_c^{(1)} = (\frac{1}{D}, 1)$, and $\alpha_{i,1} > w_{j,1}^{(1)}$ for $i \in \{f, c\}$, then the algorithm converges to a NE whose value is $\frac{1}{D}$. If, instead, $c$ uses the initial strategy $w_{c'}^{(1)} = (\frac{D-1}{D}, \frac{1}{D \cdot 0.9})$, then the algorithm converges to a NE whose value is $\frac{D-1}{D}$.*

Suppose there exists one firm $f$ and two candidates, $c_1$ and $c_2$, and each candidate is independently learning to bargain with the firm. Then, Example 6 shows that even if $f$ starts as the proposing agent and offers $0.5$ to both candidates in the first round, then, $c_1$ and $c_2$ can end up in different and asymmetric equilibria with $f$. If $c_1$ uses the strategy $w_c^{(1)}$ they get a payoff of $\frac{1}{D}$ and if $c_2$ uses the strategy $w_{c'}^{(1)}$ they get a payoff of $\frac{D-1}{D}$ for any $D > 10$.

Empirically, we find that $Algorithm$ 1 converges to NE even outside the conditions of Theorem 5, and are typically asymmetric equilibria. We implemented Algorithm 1 using CVXPY (Diamond and Boyd 2016) to give empirical convergence results for a range of initial conditions. The parameters we use are $T = 300, D = 16, M = 40, p = 1$. Since we use $p = 1$ for the regularizer, we ensure only pure strategies are chosen at each time step by rounding to 5 decimal places and breaking ties as described above.

We first chose a subset of strategy values from $\{0, \frac{1}{D}, \ldots, 1\}$ and fix each reference point. Then, we ran the algorithm for each possible pairing of initial strategies of the proposer and the responder from this subset. Empirically, we observe that the algorithm converges to NE within $T = 300$ time steps for all of these initial conditions, and the color of each cell in Figure 2 corresponds to the average payoff to the proposer using a fixed strategy over all possible responder strategies at the final converged NE point. Figure 2 shows that high first round initial offers result in a fairly stable payoff for the proposer, low initial first round offers have greater variability. In particular, the proposer gets less than half of the surplus on average if they also have a low initial second round acceptance threshold, but the proposer gets close to or more than half of the surplus if they have a high initial second round acceptance threshold. Additional simulation results are in the appendix of the full version of this paper.

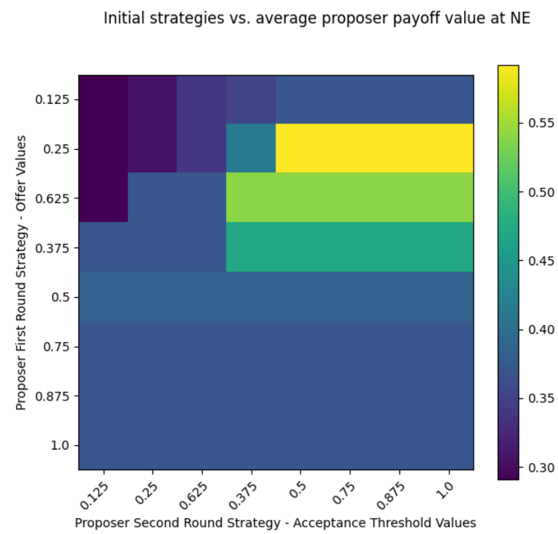


Figure 2: Simulation results for NE outcomes of agents learning strategies in $\mathcal{G}^{(2)}$. The setting is $T = 300, M = 40, D = 16, p = 1, \alpha_P = (0.125, 0.375), \alpha_R = (0.375, 0.875)$. The initial strategy of the proposer varies from $\{\frac{1}{D}, \frac{3}{D} \ldots, \frac{D-1}{D}\}$ in both strategy dimensions. The color of each cell represents the average payoff to the proposer playing that initial strategy over the initial strategies the responder plays from the same set.

### 4.5 Discussion

The purpose of this section is to show that online learning algorithms could converge to asymmetric equilibria in bargaining games. First, Proposition 3 demonstrates that no regret learning is possible for our game which justifies that learning algorithms in this and similar settings are likely to be successful.

In Theorems 4 and 5, we were able to show conditions where Algorithm 1 converges to a NE and we show *how* the initial conditions and the tie-breaking mechanism determine which equilibrium an algorithm converges to. Then, Example 6 and Figure 2 illustrate examples of asymmetric outcomes. It may seem trivial to avoid these cases by setting the relevant parameters carefully. However, our simplified algorithm where we have control over the entire design is not so realistic. Notably, there are possibly analogies to the real-world for why different agents may use different parameters in practice. We can think of $w_f^{(1)}, w_c^{(1)}$ and $\alpha_f, \alpha_c$ as the "first offers" and "private valuations"of a firm and candidate, respectively. These values possibly depend on things like information about outside options of the bargaining game, aggression of the negotiator, and past bargaining experiences.

As such, characterizing the convergence of Algorithm 1 for arbitrary parameter initializations is a start to understanding how to model the complex, possibly non-distributional, considerations of decision makers in bargaining games so that we may understand how to mitigate discriminatory effects.

## References


Abbasi, M.; Friedler, S. A.; Scheidegger, C.; and Venkatasubramanian, S. 2019. Fairness in representation: quantifying stereotyping as a representational harm. In *Proceedings of the 2019 SIAM International Conference on Data Mining*, 801–809. SIAM.

Adam, G. A.; Chang, C.-H. K.; Haibe-Kains, B.; and Goldenberg, A. 2020. Hidden risks of machine learning applied to healthcare: unintended feedback loops between models and future data causing model degradation. In *Machine Learning for Healthcare Conference*, 710–731. PMLR.

Agranov, M.; Cotton, C.; and Tergiman, C. 2018. The Use of History Dependence in Repeated Bargaining.

Anderson, E. S. 1999. What is the Point of Equality? *Ethics*, 109(2): 287–337.

Angwin, J.; Larson, J.; Mattu, S.; and Kirchner, L. 2016. Machine bias. In *Ethics of data and analytics*, 254–264. Auerbach Publications.

Arrow, K. J. 2015. The theory of discrimination. In *Discrimination in labor markets*, 1–33. Princeton University Press.

Balcan, M.-F.; Dick, T.; and Vitercik, E. 2018. Dispersion for data-driven algorithm design, online learning, and private optimization. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, 603–614. IEEE.

Barocas, S.; Hardt, M.; and Narayanan, A. 2017. Fairness in machine learning. *Nips tutorial*, 1: 2017.

Berk, R.; Heidari, H.; Jabbari, S.; Kearns, M.; and Roth, A. 2021. Fairness in criminal justice risk assessments: The state of the art. *Sociological Methods & Research*, 50(1): 3–44.

Birhane, A. 2021. Algorithmic injustice: a relational ethics approach. *Patterns*, 2(2): 100205.

Bogen, M.; and Rieke, A. 2018. Help wanted: An examination of hiring algorithms, equity, and bias.

Buddemeyer, A.; Walker, E.; and Alikhani, M. 2021. Words of wisdom: Representational harms in learning from AI communication. *arXiv preprint arXiv:2111.08581*.

Calders, T.; and Verwer, S. 2010. Three naive bayes approaches for discrimination-free classification. *Data mining and knowledge discovery*, 21: 277–292.

Cavounidis, C.; and Lang, K. 2015. Discrimination and worker evaluation. Technical report, National Bureau of Economic Research.

Cheng, M.; De-Arteaga, M.; Mackey, L.; and Kalai, A. T. 2023. Social norm bias: residual harms of fairness-aware algorithms. *Data Mining and Knowledge Discovery*, 37(5): 1858–1884.

Chouldechova, A.; and Roth, A. 2018. The frontiers of fairness in machine learning. *arXiv preprint arXiv:1810.08810*.

Cohen-Addad, V.; and Kanade, V. 2017. Online optimization of smoothed piecewise constant functions. In *Artificial Intelligence and Statistics*, 412–420. PMLR.

Curry, A. C.; Robertson, J.; and Rieser, V. 2020. Conversational assistants and gender stereotypes: Public perceptions and desiderata for voice personas. In *Proceedings of the second workshop on gender bias in natural language processing*, 72–78.

Dai, J.; and Brown, S. M. 2020. Label bias, label shift: Fair machine learning with unreliable labels. In *NeurIPS 2020 Workshop on Consequential Decision Making in Dynamic Environments*, volume 12.

Dastin, J. 2018. Amazon scraps secret AI recruiting tool that showed bias against women. In *Ethics of data and analytics*, 296–299. Auerbach Publications.

Debove, S.; Baumard, N.; and André, J.-B. 2016. Models of the evolution of fairness in the ultimatum game: a review and classification. *Evolution and Human Behavior*, 37(3): 245–254.

Diamond, S.; and Boyd, S. 2016. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83): 1–5.

Dwork, C.; Hardt, M.; Pitassi, T.; Reingold, O.; and Zemel, R. 2012. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*, 214–226.

Dwork, C.; Hays, C.; Kleinberg, J.; and Raghavan, M. 2024. Equilibria, Efficiency, and Inequality in Network Formation for Hiring and Opportunity. *arXiv preprint arXiv:2402.13841*.

Dworkin, R. 2002. *Sovereign virtue: The theory and practice of equality*. Harvard university press.

Elford, G. 2017. Survey article: Relational equality and distribution. *Journal of Political Philosophy*, 25(4).

Eubanks, V. 2018. *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Fang, H.; and Moro, A. 2011. Theories of statistical discrimination and affirmative action: A survey. *Handbook of social economics*, 1: 133–200.

Fernandez, R.; and Glazer, J. 1989. Striking for a bargain between two completely informed agents.

Ferrant, G.; Pesando, L. M.; and Nowacka, K. 2014. Unpaid Care Work: The missing link in the analysis of gender gaps in labour outcomes. *Boulogne Billancourt: OECD Development Center*, 20.

Fish, B.; and Stark, L. 2022. It's Not Fairness, and It's Not Fair: The Failure of Distributional Equality and the Promise of Relational Equality in Complete-Information Hiring Games. In *Equity and Access in Algorithms, Mechanisms, and Optimization*, EAAMO '22. New York, NY, USA: Association for Computing Machinery. ISBN 9781450394772.

Friedler, S. A.; Scheidegger, C.; Venkatasubramanian, S.; Choudhary, S.; Hamilton, E. P.; and Roth, D. 2019. A comparative study of fairness-enhancing interventions in machine learning. In *Proceedings of the conference on fairness, accountability, and transparency*, 329–338.

Green, B. 2022. Escaping the impossibility of fairness: From formal to substantive algorithmic fairness. *Philosophy & Technology*, 35(4): 90.

Hardt, M.; Price, E.; and Srebro, N. 2016. Equality of opportunity in supervised learning. *Advances in neural information processing systems*, 29.

Hazan, E.; et al. 2016. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4): 157–325.

Hellström, T.; Dignum, V.; and Bensch, S. 2020. Bias in Machine Learning–What is it Good for? *arXiv preprint arXiv:2004.00686*.

Hoffmann, A. L. 2019. Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society*, 22(7): 900–915.

Hu, L.; and Chen, Y. 2017. Fairness at equilibrium in the labor market. *arXiv preprint arXiv:1707.01590*.

Hyde, C. E. 1997. Multiple equilibria in bargaining models of decentralized trade. *Economic Theory*, 283–307.

Jiang, H.; and Nachum, O. 2020. Identifying and correcting label bias in machine learning. In *International conference on artificial intelligence and statistics*, 702–712. PMLR.

Kamiran, F.; and Calders, T. 2012. Data preprocessing techniques for classification without discrimination. *Knowledge and information systems*, 33(1): 1–33.

Kamp, S.; and Fish, B. 2024. Equal Merit Does Not Imply Equality: Discrimination at Equilibrium in a Hiring Market with Symmetric Agents. arXiv:2412.15162.

Kasirzadeh, A. 2022. Algorithmic fairness and structural injustice: Insights from feminist political philosophy. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, 349–356.

Kasy, M.; and Abebe, R. 2021. Fairness, equality, and power in algorithmic decision-mhickmaning. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 576–586.

Kay, M.; Matuszek, C.; and Munson, S. A. 2015. Unequal representation and gender stereotypes in image search results for occupations. In *Proceedings of the 33rd annual acm conference on human factors in computing systems*, 3819–3828.

Kroer, C.; and Sandholm, T. 2015. Discretization of continuous action spaces in extensive-form games. In *Proceedings of the 2015 international conference on autonomous agents and multiagent systems*, 47–56.

Kuksov, D. 2024. Endogenous Inequality in Decentralized Two-Sided Markets. *Marketing Science*.

Kusner, M. J.; Loftus, J.; Russell, C.; and Silva, R. 2017. Counterfactual fairness. *Advances in neural information processing systems*, 30.

Lam, O.; Broderick, B.; Wojcik, S.; and Hughes, A. 2018. Gender and jobs in online image searches.

Lax, D. A. 1989. Commentary on Understanding Pennzoil v. Texaco: Market Expectations of Bargaining Inefficiency and Potential Roles for External Parties in Disputes between Publicly Traded Companies. *Va. L. Rev.*, 75: 367.

Lum, K.; and Isaac, W. 2016. To predict and serve? *Significance*, 13(5): 14–19.

Malik, N. 2020. Does machine learning amplify pricing errors in housing market?: Economics of ml feedback loops. *Economics of ML Feedback Loops (September 18, 2020)*.

Mehrabi, N.; Morstatter, F.; Saxena, N.; Lerman, K.; and Galstyan, A. 2021. A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54(6): 1–35.

Miller, C. C. 2015. When algorithms discriminate. *The New York Times*, 9(7): 1.

Neuhäuser, L.; Karimi, F.; Bachmann, J.; Strohmaier, M.; and Schaub, M. T. 2023. Improving the visibility of minorities through network growth interventions. *Communications Physics*, 6(1): 108.

O'Neil, C. 2017. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Osborne, M. J.; and Rubinstein, A. 1990. *Bargaining and markets*. Academic press.

Phelps, E. S. 1972. The statistical theory of racism and sexism. *The american economic review*, 62(4): 659–661.

Ponsati, C.; and Sákovics, J. 1998. Rubinstein bargaining with two-sided outside options. *Economic Theory*, 11: 667–672.

Raghavan, M.; Barocas, S.; Kleinberg, J.; and Levy, K. 2020. Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, 469–481.

Rajkomar, A.; Hardt, M.; Howell, M. D.; Corrado, G.; and Chin, M. H. 2018. Ensuring fairness in machine learning to advance health equity. *Annals of internal medicine*, 169(12): 866–872.

Ren, Y.; Xiu, L.; and Hietapelto, A. B. 2022. Gender Differences in Asking for Pay Raises: The Role of Explicit Rules.

Ridgeway, C. L. 2014. Why status matters for inequality. *American sociological review*, 79(1): 1–16.

Roselli, D.; Matthews, J.; and Talagala, N. 2019. Managing bias in AI. In *Companion Proceedings of The 2019 World Wide Web Conference*, 539–544.

Schemmel, C. 2012. Distributive and relational equality. *Politics, philosophy & economics*, 11(2): 123–148.

Shaked, A.; et al. 1994. Opting out: bazaars versus 'Hi Tech'markets. *Investigaciones Economicas*, 18(3): 421–432.

Shalev-Shwartz, S.; et al. 2012. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2): 107–194.

Shelby, R.; Rismani, S.; Henne, K.; Moon, A.; Rostamzadeh, N.; Nicholas, P.; Yilla-Akbari, N.; Gallegos, J.; Smart, A.; Garcia, E.; et al. 2023. Sociotechnical harms of algorithmic systems: Scoping a taxonomy for harm reduction. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 723–741.

Tadelis, S. 2013. Game Theory; an Introduction.

Ulph, A.; and Ulph, D. 1998. Labour markets, bargaining and innovation. *European Economic Review*, 42(3-5): 931–939.

Weber, B. C. 2023. Hub-and-Spoke Conspiracies: Can Big Data and Pricing Algorithms Form the Rim? *SMU Sci. & Tech. L. Rev.*, 26: 25.

West, S. M.; Whittaker, M.; and Crawford, K. 2019. Discriminating systems. *AI Now*, 1–33.