Policy Fairness in Sequential Allocations under Bias Dynamics

Meirav Segal [1] **Anne-Marie George** [1] **Christos Dimitrakakis** [1 2]

Abstract

This work considers a dynamic decision making framework for allocating opportunities over time to advantaged and disadvantaged individuals. Here, individuals in the disadvantaged group are assumed to experience a societal bias that limits their success probability. A policy of allocating opportunities stipulates thresholds on the success probability for the advantaged and disadvantaged group. We analyse the interplay between utility and a novel measure of fairness for different dynamics that dictate how the societal bias changes based on the current thresholds while the group sizes are fixed. Our theoretical analysis is supported by experimental results on synthetic data for the use case of college admissions.

1. Introduction

AI models for allocation problems have become prevalent in many applications, such as lending (Dastile et al., 2020) and policing (Lum & Isaac, 2016). These problems are characterized by a decision maker (DM) allocating limited resources among a population in order to maximize some objective. Recent efforts provide fair allocations by incorporating fairness constraints, guaranteeing a sufficient portion of resources to protected groups. While most of these endeavors focus on static settings (single allocation), we consider a sequence of allocations over time. This can introduce feedback effects, so that current decisions may change the future population distribution. In such cases, well-intended attempts to increase fairness might lead to negative long-term effects for the protected group (Liu et al., 2018). In this paper, we study how the underlying dynamics and our allocation policy affect the protected group in the short and long term.

[1]Department of Informatics, University of Oslo, Oslo, Norway [2]Computer Science Department, Université de Neuchâtel, Neuchâtel, Switzerland. Correspondence to: Meirav Segal <meiravs@ifi.uio.no>.

ICML 2022 Workshop on Responsible Decision Making in Dynamic Environments, Baltimore, Maryland, USA, 2022.

We focus on the example of college admissions. Higher education is a key element towards many career paths. As such, access to higher education is crucial for self fulfillment and financial security. Unfortunately, there are still sub-populations with reduced access to this opportunity due to societal biases: Discouraging environments, lack of role models and internalized stereotypes could lead to reduced chances of success through insufficient skill development or self-handicapping (Herrmann et al., 2016; Tyler et al., 2016). Hence, DMs may try to equalise group participation through affirmative action, such as setting a lower acceptance bar for disadvantaged groups. Such actions could potentially generate more role models and provide investment incentive, which might encourage further skill development. However, depending on the social dynamics, they might also have negative effects, as lowering the bar entails the admission of less qualified group members with reduced chances of graduation, leading to a lower success rate within this group and thereby increasing bias by reinforcing stereotypes.

We model this environment as a Markov Decision Process (MDP) and examine the purely utility maximising policy in terms of utility and fairness under different bias dynamics. To do so, we propose a measure for the fairness of a policy and analyse how affirmative action relates to this notion.

2. Related Work

Our setting is similar to that of Heidari and Kleinberg (2021), who formalized the process of intergenerational mobility between groups with different socioeconomic status. In their model, individuals may shift groups based on their performance in opportunities granted to them. For instance, individuals who are admitted to university and graduate successfully, have increased chances of earning higher salaries. In our setting, the groups are fixed, but our decisions can have an impact at the population level, by changing societal biases. Additionally, we consider two different dynamics, in order to see how sensitive our findings are to the policy and underlying model.

Others have explored the long-term effect of a policy on the fairness towards sub-populations. For example, a recent work has experimentally compared the long-term implications of different policies using simulations, with a different social welfare measure for each application (D'Amour et al.,

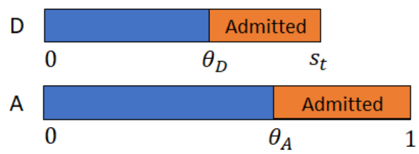


Figure 1. The success distribution and threshold effect on admittance for the disadvantaged (D) and advantaged (A) group.

2020). Moreover, the long-term influence of policies using affirmative action was examined (Mouzannar et al., 2019). Yet, other kinds of preferential treatments were not considered.

3. Use Case: College Admissions

We model a MDP in which the DM repeatedly allocates admission slots to a fixed proportion ($\alpha \in [0, 1]$) of the candidates. The DM's reward per round is the number of successful students and the DM's utility is the discounted sum of total rewards over time.

We consider a population partitioned into two disjoint groups: disadvantaged (D) and advantaged (A). We assume that these groups represent a constant affiliation (such as race) so there are no transitions between groups. More specifically, we assume that group fractions remain constant over time. The fraction of the population that is disadvantaged, ϕ, is hence a fixed parameter of the MDP. In addition, we assume that group affiliation has no influence over the innate ability of individuals, which is uniformly distributed for both groups ($a_i \sim U[0, 1]$ for individual $i \in A \cup D$).

The current state of the MDP is the current societal bias factor $s_t \in [0, 1]$. In our model, societal bias influences the skill development of the disadvantaged group members, leading to lower chances of success: The success probability for an advantaged group member is equal to their innate ability $p_A(i) = a_i$, while for a disadvantaged group member, the success probability is their ability multiplied by the current bias factor $p_D(j) = s_t a_j$. In effect, the current state is translated to the upper bound of the success distribution of the disadvantaged group, i.e., $p_D \sim U[0, s_t]$.

After observing the MDP state at time t, and the success probability distribution of candidates from either group, the DM selects a pair of admission thresholds $a_t = (\theta_D, \theta_A)$, one for each group. A candidate is admitted if their success probability is over the group's threshold. Having different thresholds allows for affirmative action policies, e.g., see Figure 1 where the threshold of the disadvantaged group θ_D is lower than that of the advantaged group θ_A.

The current action and state s_t determine the number of admitted and successful students, i.e., the reward $R(s_t)$.

This in turn determines the next MDP state through the transition dynamics. We assume that exogenous factors affect the bias state as well, at a constant level $\sigma \in [0, 1]$, and the effect of the DM's decisions accommodate the rest $(1 - \sigma)$. In this paper, we consider two types of dynamics:

1. Representation Dynamics: the future bias depends on the current fraction of selected students from the disadvantaged group compared to α. This corresponds to the notion of demographic parity: the action/decision should be independent of the group affiliation.

2. Relative Success Dynamics: the future bias depends on the current success probability of admitted students from the disadvantaged group, compared to that of the advantaged group. This corresponds to the notion of predictive parity (Chouldechova, 2017) or equality of opportunity(Hardt et al., 2016).

In each case, we use the corresponding bias factor as a measure of fairness. The transition dynamics are treated as fixed parameters of the MDP.

4. Utility Maximisation

We define the utility of the DM as a discounted sum of rewards achieved by acting according to a policy π from a given start state s_0:

$$U^{\pi}(s_0) = \sum_{t=0}^{T} \gamma^t R(s_t).$$

Where $\gamma \in [0, 1)$ is the discount factor, which determines the weight we give to future rewards. The closer γ is to 1, the more weight we place on future rewards. When the start state is unknown, we integrate over the possible start states through some starting distribution Q. In our case, we assume all states $s_t \in [\sigma, 1]$ are possible start states.

When the model is known, we can use approximate dynamic programming to obtain the utility maximising policy. Using discretization of the state and action spaces ($\#states = \#actions = 1000$), we apply policy iteration.

Using our model, we wish to observe the effect of different model parameters on the utility maximising policy, it's utility and fairness. Specifically, we would like to test our hypothesis, that for dynamic systems, there is not trade-off between fairness and utility, only a trade-off between short-term and long-term rewards. To this end, we first need to define a fairness measure for a policy.

5. Fairness as Consequentialism

In static settings, fairness is usually measured with respect to the action or decision. We can define the fairness of a state, which is closely tied to the action leading to that state.

To evaluate the fairness of a policy, we could consider the fairness of a state reached after a fixed time t_{max}. Yet, using this measure we cannot differentiate between policies along the entire time period. For example, we cannot differentiate two policies that reach an unbiased state after 20 steps, while one discriminated in every step but the last, and the other is discriminated in the first step but is unbiased afterwards.

Instead, we define the fairness of a policy as the weighted sum of state fairness scores, according to the expected path induced by the policy. Formally, let us assume we have a function $f : S \to \mathbb{R}$ such that for every state of the world, we get a real number indicating the fairness of that state. We define the fairness of a policy π according to weights w_t for the fairness scores of visited states s_t starting from start state s_0 as:

$$F^{\pi}(s_0) = \mathbb{E}_{\pi}[\sum_{t=0}^{T} w_t f(s_t)]. \tag{1}$$

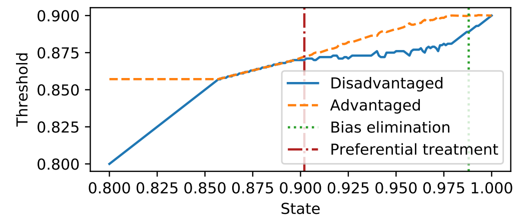
Note that, if a discount factor γ is used as a weight, the fairness of a policy $F^{\pi}(s_0) = \sum_{t=0}^{T} \gamma^t f(s_t)$ takes the same shape as the utility function where state fairness scores correspond to rewards.

When the start state is unknown, we integrate over the possible start states to measure the fairness (and utility) of a policy. That is, we compute

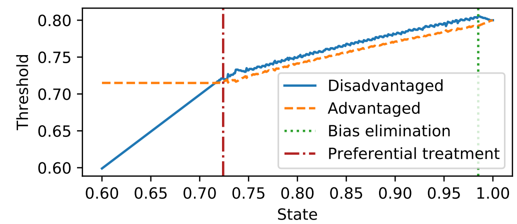
$$F^{\pi} = \int_S \mathbb{E}_{\pi}[\sum_{t=0}^{T} w_t f(s_t)]dQ(s) \tag{2}$$

This fairness measure allows us to identify the fairest and most unfair policy and set the utility maximising policy of the DM in relation to these.

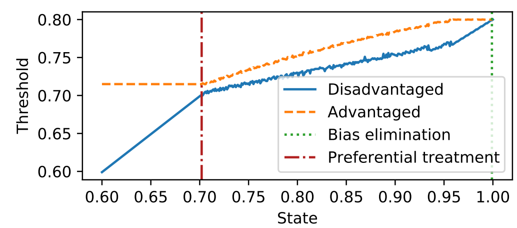
In our model, the fairness of a state corresponds to the value of the bias factor. When it equals 1, the success probability of the disadvantaged group members is not impaired, and we see this state as completely unbiased. The lower the bias factor, the more unfair the state. We choose the same discount factor γ as used in the utility function as weights for the fairness measure. Furthermore, the considered dynamics are deterministic and any policy considered is deterministic, such that we can omit the expectation in the fairness measure in Equation (2). Since we discretise the state space, we only need to compute a sum instead of the integral.



(a)



(b)



(c)

Figure 2. Utility maximising policy for $\phi = 0.3$: for each state we plot the action $a = (\theta_A, \theta_D)$ according to the policy. (a) representation dynamics ($\sigma = 0.8, \alpha = 0.1, \gamma = 0.1$), (b) relative success dynamics ($\sigma = 0.6, \alpha = 0.2, \gamma = 0.2$) and (c) combined dynamics ($\sigma = 0.6, \alpha = 0.2, \gamma = 0.2$).

6. Experimental Results

First, we present the utility-maximising policies for three transition functions - representation dynamics, relative success dynamics and a combined dynamics (average of the two). Then we evaluate the utility-fairness trade-off for these policies.

6.1. Utility Maximising Policy

As we can see in Figure 2, for all three types of dynamics there is an interval of high-bias (i.e., low state values) in which no one from the disadvantaged group is admitted (the acceptance threshold for the disadvantaged group appears as equal to the upper bound of the success distribution, but it could also be considered as equal to the acceptance threshold for the advantaged group, as both lead to the same outcome). Interestingly, only for the representation dynamics, this is followed by an interval of higher state values in which thresholds for advantaged and disadvantaged group are equal.

All dynamics display a clear tipping point (lower bound) on state values from which on a preferential treatment is implemented, i.e., the action sets different thresholds for the two groups, which increase the future bias factor. For representation dynamics, this preferential treatment is in the known form of affirmative action, i.e., setting a lower acceptance threshold for the disadvantaged group. The same preferential treatment is demonstrated by the combined dynamics. Yet, for the relative success dynamics, preferential treatment actually means increasing the threshold for the disadvantaged group. While this action would appear as discriminatory in a static setting, it is in fact an action with positive long-term effects for the disadvantaged group.

For both, the representation and the relative success dynamics, there exists a small interval of high state values for which an unbiased state can be reached within only one time step (see green line in in Figure 2). It seems not to be possible to eliminate the bias within one step for the utility maximising policy if the bias dynamic is a combination of representation and relative success.

6.2. Utility-Fairness Trade-off

In this experiment, we analyse policies which maximise utility over an infinite horizon, for different discount factors γ and dynamics. The same γ is used for measuring both utility and fairness. In figure 3 we can see the utility and fairness (both representation and relative success) in a number of different cases. In figure 3a these measures are computed for horizon of only 1 step for the representation dynamics. Meaning, the utility is the immediate reward and the policy fairness is simply the state fairness after one action. For different discount factors, it seems that we can tune the utility-fairness trade-off: for lower values we get higher rewards with lower fairness, while for larger discount factors we get lower rewards with higher fairness. Yet, when we observe these measures for a longer horizon (utility and fairness measured for 150 steps, normalised by $1-\gamma$) in figure 3b, we can see that the utility and representation fairness are aligned and increase with the discount factor. Similar results are achieved for relative success dynamics (figure 3c) where and the combined dynamics (figure 3d), where utility is aligned with both fairness measures. Thus, we can conclude that the trade-off in this dynamic system is not between fairness and utility, but between short-term and long-term rewards.

7. Discussion

This analysis indicates that the effect of affirmative action on long-term fairness strongly depends on the dynamics. For representation dynamics, the use of affirmative action would reduce societal bias, while for relative-success dynamics, it would be sub-optimal in terms of both utility and policy fairness. Note that consequential fairness may not be the only fairness notion that should be considered, and that our model does not take into account any possible societal outcomes of normalizing preferential treatment. Moreover, we only consider here the effect of one kind of intervention (setting different admission thresholds), but based on the analysis of the dynamics, the DM could apply other measures, e.g., increasing success of admitted students by providing private tutors.

In this model, as in many real-world cases, fairness and utility are aligned in the long-term, because the current state of the world diverged from meritocracy due to historical biases. This analysis sheds light on the true trade-off for decision makers in such cases, which is between short-term and long-term rewards.

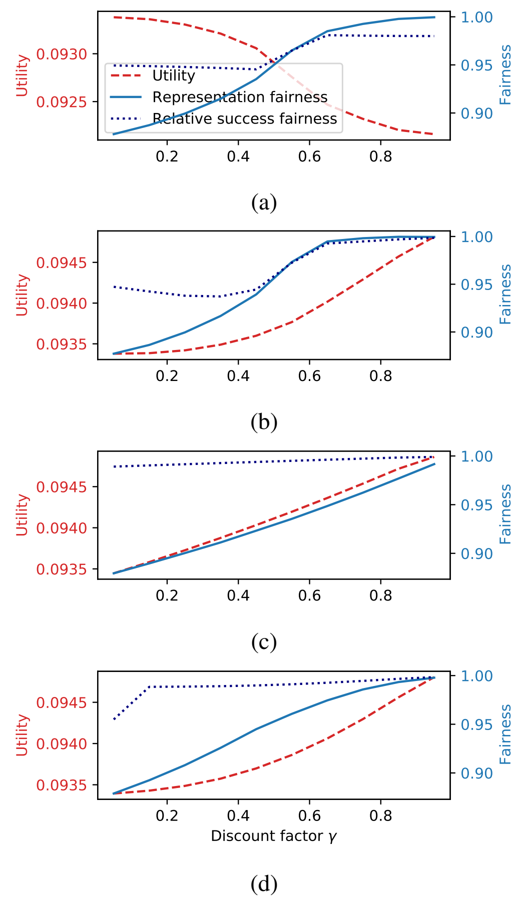


Figure 3. Utility, representation fairness and relative success fairness for utility-maximising policies under different dynamics trained for 10 different discount factors ($\phi = 0.3, \sigma = 0.8, \alpha = 0.1$). (a) Representation dynamics for horizon of 1 (immediate reward and state fairness) (b) Representation dynamics for horizon of 150. (c) relative success dynamics for horizon of 150 steps. (d) combined dynamics for horizon of 150.

Acknowledgements

This work was supported by the Research Council of Norway under project number 302203.

References

Chouldechova, A. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data*, 5(2):153–163, 2017.

D'Amour, A., Srinivasan, H., Atwood, J., Baljekar, P., Sculley, D., and Halpern, Y. Fairness is not static: deeper understanding of long term fairness via simulation studies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 525–534, 2020.

Dastile, X., Celik, T., and Potsane, M. Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, 91:106263, 2020.

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

Heidari, H. and Kleinberg, J. Allocating opportunities in a dynamic model of intergenerational mobility. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 15–25, 2021.

Herrmann, S. D., Adelman, R. M., Bodford, J. E., Graudejus, O., Okun, M. A., and Kwan, V. S. The effects of a female role model on academic performance and persistence of women in stem courses. *Basic and Applied Social Psychology*, 38(5):258–268, 2016.

Liu, L. T., Dean, S., Rolf, E., Simchowitz, M., and Hardt, M. Delayed impact of fair machine learning. In *International Conference on Machine Learning*, pp. 3150–3158. PMLR, 2018.

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

Mouzannar, H., Ohannessian, M. I., and Srebro, N. From fair decision making to social equality. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pp. 359–368, 2019.

Tyler, K. M., Thompson, F. A., Gay, D. E., Burris, J., Lloyd, H., and Fisher, S. Internalized stereotypes and academic self-handicapping among black american male high school students. *Negro Educational Review*, 67(1-4): 5, 2016.