Fair Set Selection: Meritocracy and Social Welfare*

Thomas Kleine Büning† Meirav Segal† Debabrota Basu‡ Christos Dimitrakakis†

25th February 2021

arXiv:2102.11932v1 [cs.AI] 23 Feb 2021

Abstract

In this paper, we formulate the problem of selecting a set of individuals from a candidate population as a utility maximisation problem. From the decision maker's perspective, it is equivalent to finding a selection policy that maximises expected utility. Our framework leads to the notion of *expected marginal contribution* (EMC) of an individual with respect to a selection policy as a measure of deviation from meritocracy. In order to solve the maximisation problem, we propose to use a policy gradient algorithm. For certain policy structures, the policy gradients are proportional to EMCs of individuals. Consequently, the policy gradient algorithm leads to a locally optimal solution that has zero EMC, and satisfies meritocracy. For uniform policies, EMC reduces to the Shapley value. EMC also generalises the fair selection properties of Shapley value for general selection policies. We experimentally analyse the effect of different policy structures in a simulated college admission setting and compare with ranking and greedy algorithms. Our results verify that separable linear policies achieve high utility while minimising EMCs. We also show that we can design utility functions that successfully promote notions of group fairness, such as diversity.

1 Introduction

Machine learning is now pervasively used to assist high-stake decision making, with significant consequences for individuals and organisations. We are especially interested in the case where the decision maker (DM) must select a *set* of individuals from a population so as to maximise expected utility. This case is especially interesting as maximising set utility can induce diversity without additional constraints for certain domains. For instance, according to Díaz-García et al. [2013], if we wish to maximise innovative ideas, we would prefer teams with gender diversity. We can thus assume that expected utility represents an appropriate notion of social welfare. In addition to maximising utility, the DM may also like their selection policy to be meritocratic. The first question that arises is how to define meritocracy for general utility functions. Here we suggest using a notion of *expected individual contribution* to social welfare to achieve meritocracy.

In particular, we consider the case where the DM maximises a utility function representing social welfare. This means that a policy maximising utility would be fair in terms of its aggregate benefit to society. However, this may conflict with notions of fairness relating to individuals or groups, and in particular egalitarianism and meritocracy.

Egalitarianism is a notion of fairness where individuals' chances of obtaining a reward are independent of personal characteristics. In the simplest case, we wish our decisions to be independent of *any* personal characteristics, which is achievable through a uniformly random lottery. More typically, we wish our decisions or the outcomes of our decisions to be independent of sensitive characteristics such as gender or ethnicity.

Meritocracy suggests that individuals with the highest "worth" obtain higher rewards. As long as there is an inherent, static measure of worth this is a well-defined notion. However, in many settings, including ours, worth is relative and dependent on circumstances. In particular, the relative contribution of an individual to utility depends on who else has been selected (and more generally the DM's policy) so it is a constantly shifting quantity.

*Submission to the archival-option track of FORC 2021.

†University of Oslo

‡Scool, Inria Lille-Nord Europe

Contribution. We define a notion of meritocracy for set selection problems and arbitrary utility functions relating to individual *expected marginal contributions* (EMC). This is analogous to marginal utilities for a basket of goods in economics, which must necessarily depend on both the *population* under consideration and the DM's *policy* for selection. This is in contrast to works on set selection based on ranking, which implicitly define meritocracy through a static worth for every individual. In particular, when the DM's policy is egalitarian, the EMC is identical to the Shapley value [Shapley, 1951]. A natural way to shift the policy towards meritocracy would then be to increase the probability of selecting individuals proportionally to their EMC. If we apply this algorithmic idea to any initial DM policy, we show that we arrive at a *policy gradient algorithm* for policies that are separably parameterised over the population. For these, maximising utility also achieves meritocracy, as the gradient is a linear transform of the EMC. For other policies, e.g. those relying on a threshold, this does not hold and meritocratic outcomes cannot be guaranteed when maximising utility. Then, we can measure the deviation from meritocracy in terms of the residual EMC. Experimentally, we analyse different policy structures in a college admission setting and illustrate the relation of meritocracy with respect to EMCs to policy structures and utility maximisation. We experimentally instantiate the policy gradient framework on linear and log-linear utility, and separable linear and threshold policies. We measure the expected utility and diversity of the solution to quantify the solution quality. We show that separable policies achieve expected utility, residual EMC, and diversity index that are close to optimal, whereas policies based on thresholding and ranking are significantly sub-optimal.

2 Related work

Ranking with pre-determined measures of "worth". Most work on fairness in set selection has focused on policies that rank individuals according to some pre-determined criterion [Kearns et al., 2017, Zehlike et al., 2017, Celis et al., 2017, Biega et al., 2018]. This approach satisfies meritocracy, since "better" individuals should be ranked higher, and higher ranked individuals are preferred over lower ranked ones. In particular, Kearns et al. [2017] consider a probabilistic ordering, generalising Dwork et al.'s notion of similar treatment to selection over multiple groups. More precisely, a person i in group j is preferred to a person i' in group j' only if their relative percentile ranking is higher. They extend this basic definition to different amounts of information available to the decision maker ranging from *ex ante* to *ex poste* fairness. In a similar vein, Singh and Joachims [2019] propose a fair ranking approach for Luce-Plackett ranking models. We instead approach this problem from a utility maximisation perspective, where meritocracy is defined as rewarding individuals according to their contribution to social welfare. This is in contrast to ranking methods such as the above, which implicitly assume a fundamental worth for individuals. In our setting, rather the contribution of each individual to the utility depends on who else is selected.

Several other papers assume the existence of latent individual worth with respect to which meritocracy can be measured. Kleinberg and Raghavan [2018] analyse a stylised parametric model of individual *potential* and Celis et al. [2020] consider interventions for ranking, where each individual has a latent utility they would generate if hired. Emelianov et al. [2020] also examine latent worth with variance depending on group (e.g. gender) membership.

Fair selection as utility maximisation. Our work is not the only one in the fairness literature seeking to select sets of individuals in a way that maximises utility. Recently, Kusner et al. [2018] considered the setting where the utility is linear and the performance of individuals in the selected set depends on who else is selected. They investigate how these interactions affect group fairness. While this is also possible in our setting if we use an appropriate model, we instead focus on the problem of defining meritocracy when there is no unique ranking of individuals. This is generally the case when the DM's utility is non-linear. While the problem of fair package assignment does consider non-linear utilities and targets efficiency and envy-freeness [Lahaie and Parkes, 2009], assignment problems do not readily transfer to the framework discussed in the paper.

Stoyanovich et al. [2018] consider diversity in terms of recommendations as a constraint in decision making. These methods rely on hard-coded thresholds on diversity, which are hardly available in reality and are too application-specific. Recently, Kilbertus et al. [2020] proves an impossibility result showing that decision making with *deterministic* thresholds (ref. Proposition 2) may not improve either fairness or utility but learning a *stochastic* fair policy can circumvent this. This result motivates us to consider stochastic policies that optimise a welfare-based utility.

Fair set selection. Fair set selection problems can also naturally be found in social choice. For example, in participatory budgeting, we aim to select a set of projects all of which have different costs, while respecting a general budget constraint [Aziz and Shah, 2020]. In elections or tournaments, we want to identify a set of winners [Moulin, 2016]. More specifically, in committee voting, we elect a committee of a given size, i.e. a subset of all candidates [Elkind et al., 2017, Lackner and Skowron, 2020]. All of these set selection problems are typically considering a set of candidates that have no features *per se*. It is assumed that there is a set of voters expressing preferences over the candidates. We can interpret those voters' preferences as features in analogy of our framework. Fair set selection is obtained through mechanisms that guarantee fairness axioms such as anonymity and justified representation. These axioms are often binary, i.e. are either satisfied or not. They also often deal with fairness with respect to the voters, while we concentrate on fairness towards the individual candidates.

3 Setting

We formulate the problem of selecting a set of individuals from a population from a general decision theoretic perspective, where the Decision Maker (DM) aims to select a subset of individuals maximising expected utility. Crucially, the utility function may be non-linear and depend on the set of selected individuals as a whole. Thus, the utility may not be decomposable into terms relating to each individual.

Specifically, we consider a population of N candidates $\mathcal{N}$. The DM observes the features of the population $\mathbf{x} \in \mathcal{X}$, where $\mathbf{x} \triangleq (\mathbf{x}_1, \ldots, \mathbf{x}_N)$, and then makes a decision $\boldsymbol{a} \in \mathcal{A}$ about the population using a (stochastic) policy $\pi(\boldsymbol{a} \mid \mathbf{x})$. We focus on the case where $\mathcal{A} = \{0,1\}^N$ and interpret a decision $\boldsymbol{a} = (a_1, \ldots, a_N)$ such that $a_i = 1$ stands for selecting the i-th individual and $a_i = 0$ stands for rejecting individual i. After the DM makes a choice, she observes an outcome $y \in \mathcal{Y}$. Typically, $\mathcal{Y}$ is a product space encoding outcomes for every individual in the population.

The utility of the DM is a function over subsets and outcomes $u : \mathcal{A} \times \mathcal{Y} \to \mathbb{R}$. The utility function $u(\boldsymbol{a}, \mathbf{y})$ represents social welfare for every possible combination of selected individuals and outcomes. Since we only have access to the observed features $\mathbf{x}$, the expected utility of an action $\boldsymbol{a}$ can be calculated by marginalising over outcomes $\mathbf{y}$:

$$U(\boldsymbol{a}, \mathbf{x}) \triangleq \mathbb{E}[u \mid \boldsymbol{a}, \mathbf{x}] = \sum_{\mathbf{y} \in \mathcal{Y}} \mathbb{P}(\mathbf{y} \mid \boldsymbol{a}, \mathbf{x})\, u(\boldsymbol{a}, \mathbf{y}). \tag{1}$$

Here, $\mathbb{P}(\mathbf{y} \mid \boldsymbol{a}, \mathbf{x})$ is a predictive model used by the DM to estimate the outcomes. For simplicity, in this paper we use a point estimate of the true conditional distribution. The expected utility of a policy π given population $\mathbf{x}$ is then

$$U(\pi, \mathbf{x}) \triangleq \mathbb{E}_\pi[u \mid \mathbf{x}] = \sum_{\boldsymbol{a} \in \mathcal{A}} \pi(\boldsymbol{a} \mid \mathbf{x})\, U(\boldsymbol{a}, \mathbf{x}). \tag{2}$$

While $U(\boldsymbol{a}, \mathbf{x})$ naturally induces a ranking over different subsets for a given population $\mathbf{x}$, it does not necessarily provide a ranking over individuals as each individual's utility may depend on the group selected alongside it.

The goal of the DM is to find a parameterised policy in the policy space $\mathit{\Pi} = \{\pi_\theta \mid \theta \in \Theta\}$ maximising expected utility. We distinguish two cases: in the first, the DM chooses a policy after observing the population, and in the second the policy is fixed before the population is seen.

Population-dependent policies. For a given population $\mathbf{x}$, the goal is to maximise the expected utility for a given population:

$$\boldsymbol{\theta}^*(\mathbf{x}) = \arg\max_{\boldsymbol{\theta} \in \Theta} U(\pi_{\boldsymbol{\theta}}, \mathbf{x}). \tag{3}$$

This is the typical setting where the selected set is chosen in a way so as to take into account all the information we have about the current population, and it will be the focus of this paper.

Population-independent policies. In some cases, e.g. when there have to be fixed rules for selecting individuals from a population, the DM may need to choose a policy before seeing a particular population. In this scenario, we can assume that the DM has access to some distribution β over possible populations, and the problem becomes maximising expected utility under this distribution:

$$\boldsymbol{\theta}^*(\beta) = \arg\max_{\boldsymbol{\theta}\in\Theta} \int_{\mathcal{X}} U(\pi_{\boldsymbol{\theta}}, \mathbf{x}) d\beta(\mathbf{x}). \tag{4}$$

Another advantage of such policies in terms of fairness is that, under some specific policy structures, each individual can be judged without taking into account who else is in the current population. Instead, individuals will be judged relative to their expected contribution over all possible populations.

4 Utility maximisation and marginal contributions

While maximising utility is appropriate from the point of view of social welfare, it is not necessarily meritocratic. In the following, we will link fairness to the expected marginal contribution of individuals. For certain policy structures, this is a linear transformation of the expected utility with respect to the policy parameters. In consequence, at any local maximum, the gradient and so the expected marginal contribution is zero. This can be interpreted as saying that social welfare would be reduced by changing the current allocation.

4.1 The expected marginal contribution

We are now going to formally introduce the expected marginal contribution of an individual under a given policy π. The marginal contribution of an individual j to a set $\boldsymbol{a}$ is given by $U(\boldsymbol{a}+j) - U(\boldsymbol{a})$, where we mean by $\boldsymbol{a}+j$ that individual j is being selected, i.e. $\boldsymbol{a}+j = (a_1, \ldots, 1, \ldots, a_N)$ for any $\boldsymbol{a} \in \mathcal{A}$. In our case, we do not have a fixed set $\boldsymbol{a}$ by use of which we could compare marginal contributions of two individuals. We therefore define the *expected marginal contribution* (EMC) of an individual j under a policy π as

$$\varDelta_j U(\pi, \mathbf{x}) = \sum_{\boldsymbol{a}\in\mathcal{A}} \pi(\boldsymbol{a} \mid \mathbf{x}) \big[U(\boldsymbol{a}+j, \mathbf{x}) - U(\boldsymbol{a}, \mathbf{x})\big]. \tag{5}$$

Informally, the EMC of individual j corresponds to the gain (or loss) in expected utility when a policy is modified so as to always pick individual j. The concept of marginal contributions has been studied in cooperative game theory, where the celebrated Shapley value [Shapley, 1951] constitutes a fair resource allocation based on marginal contributions. In particular, the Shapley value is characterised by four axioms of fair division, namely efficiency, symmetry, linearity, and the treatment of null players. As we show below, the EMC satisfies analogous axioms.

Properties. The EMC satisfies the following axioms of fair division, where we omit $\mathbf{x}$ for brevity.

1. Symmetry: If $U(\boldsymbol{a}+i) = U(\boldsymbol{a}+j)$ for all $\boldsymbol{a} \in \mathcal{A}$, then $\varDelta_i U(\pi) = \varDelta_j U(\pi)$ for any policy π.
2. Linearity: If U_1 and U_2 are two utility functions, then $\varDelta_j U(\alpha U_1 + \beta U_2)(\pi) = \alpha \varDelta_j U_1(\pi) + \beta \varDelta_j U_2(\pi)$ for any policy π.
3. Null player: If an individual does not contribute to any of the subsets, i.e. $U(\boldsymbol{a}+j) = U(\boldsymbol{a})$ for all $\boldsymbol{a} \in \mathcal{A}$, then $\varDelta_j U(\pi) = 0$ for any policy π.

If we restrict ourselves to policies π with support equal to $\mathcal{A}$, then property 3 indeed becomes an equivalence. That is, if $\varDelta_j U(\pi) = 0$ for every policy π with support $\mathcal{A}$, then $U(\boldsymbol{a}+j) = U(\boldsymbol{a})$ for all $\boldsymbol{a} \in \mathcal{A}$.

These properties are - apart from *efficiency* which does not apply to the EMC - analogs to the axioms of fair division characterising the Shapley value. In fact, we obtain the Shapley value when assuming an egalitarian policy.

Remark 1. *When the selection policy π is egalitarian, i.e. $\pi(\boldsymbol{a}) = \frac{1}{(k+1)\binom{N}{k+1}}$ when $k = \sum_{i=1}^N a_i$, the EMC $(\varDelta_i U(\pi_{egal}))_{1\le i\le N}$ is equal to the Shapley value.*

4.2 Policy structures

We now demonstrate that the EMC defined in equation (5) is related to the policy gradient $\nabla_{\boldsymbol{\theta}} U(\pi_{\boldsymbol{\theta}})$ for specific types of policy parameterisations. Specifically, we consider policies that are separably parameterised over the population and show that the policy gradient is a linear transform of the EMC in specific cases. In the following, we will omit $\mathbf{x}$ and write $\pi(\boldsymbol{a})$ for $\pi(\boldsymbol{a} \mid \mathbf{x})$ for brevity.

4.2.1 Separable policies

We say that a parameterised policy is *separable over the population* $\mathcal{N}$ if $\boldsymbol{\theta} = (\theta_1, \dots, \theta_N)$ and there exists $Z(\boldsymbol{\theta})$ such that for every $i \in \{1, \dots, N\}$ the probability of selecting individual i takes form

$$\pi_{\boldsymbol{\theta}}(a_i) = \frac{g(a_i, \theta_i)}{Z(\boldsymbol{\theta})} \tag{6}$$

for some function g, and $\pi_{\boldsymbol{\theta}}(\boldsymbol{a}) = \prod_{i=1}^{N} \pi_{\boldsymbol{\theta}}(a_i)$. In slightly imprecise terms this means that the probability of selecting an individual i only depends on the i-th parameter θ_i apart from a common normalisation.

Softmax policies. A natural choice for policies that take form as in (6) are softmax policies of the following kind. For $\beta \geq 0$ and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_N)$, we define

$$\pi_{\boldsymbol{\theta}}(\boldsymbol{a}) = \frac{e^{\beta \boldsymbol{\theta}^T \boldsymbol{a}}}{\sum_{\boldsymbol{a}' \in \mathcal{A}} e^{\beta \boldsymbol{\theta}^T \boldsymbol{a}'}}.$$

Here, $\beta \geq 0$ is called the inverse temperature of the distribution. While such policies have the advantage that the probability of selecting a set is naturally constrained, they are clearly impractical for large scale experiments as calculating the denominator is computationally heavy.

However, we get a first glance at the close relationship between the policy gradient and the expected marginal contribution in form of the next lemma.

Lemma 1. *The gradient is a linear transform of the EMC. More precisely, for every $j \in \{1, \dots, N\}$:*

$$\nabla_{\theta_j} U(\pi_{\boldsymbol{\theta}}) = \beta \pi_{\boldsymbol{\theta}}(a_j = 1)\, \Delta_j U(\pi_{\boldsymbol{\theta}}),$$

where $\pi_{\boldsymbol{\theta}}(a_j = 1) \triangleq \sum_{\boldsymbol{a}: a_j = 1} \pi_{\boldsymbol{\theta}}(\boldsymbol{a})$ is the probability of selecting a set containing individual j under policy $\pi_{\boldsymbol{\theta}}$.

Note that the EMC constitutes the expected contribution of always adding individual j under the current policy, so if an individual is always added, then $\Delta_j U(\pi_{\boldsymbol{\theta}})$ is zero. More generally, the EMC of an individual j is decreasing as the probability of selecting individual j is increasing. Consequently, we can understand the factor $\pi_{\boldsymbol{\theta}}(a_j = 1)$ in Lemma 1 as a normalisation of the EMC opposing this effect. The relation of the EMC and the gradient is again illustrated at an example in Figure 1.

Linear policies. From a computational point of view, it is appealing to consider separable policies so that the probability of selecting individual i truly depends on the parameter θ_i only, i.e. $Z(\boldsymbol{\theta}) = \text{constant}$. Then, the probability of a decision about individual i is simply given by $\pi_{\boldsymbol{\theta}}(a_i) = \pi_{\theta_i}(a_i)$. Note that using such policies in Algorithm 1 does not render decisions about individual i independent from those about individual j, as the parameters $\boldsymbol{\theta}^*$ chosen by (3) depend on the whole population represented by the feature vector $\mathbf{x}$.

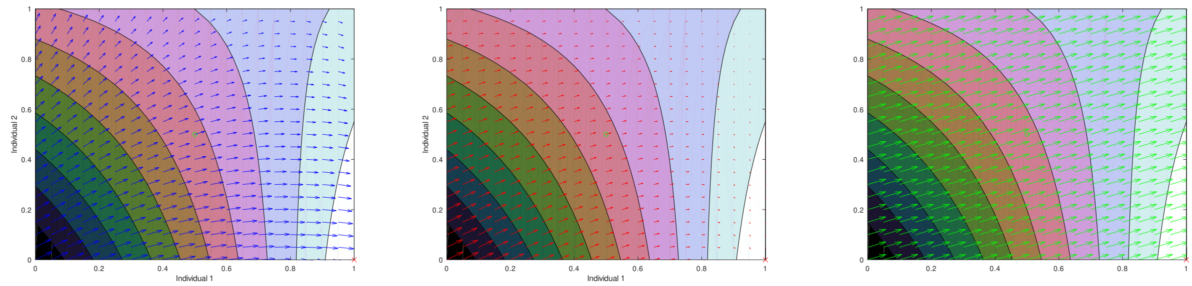
We consider separable linear policies as these have a particularly intuitive structure. These select individual i with probability θ_i, i.e.

$$\pi_{\theta_i}(a_i) = \theta_i \mathbb{I}\{a_i = 1\} + (1 - \theta_i) \mathbb{I}\{a_i = 0\},$$

where $\boldsymbol{\theta} = (\theta_1, \dots, \theta_N)$. Consequently, a separable linear policy assigns a decision $\boldsymbol{a} \in \mathcal{A}$ the probability

$$\pi_{\boldsymbol{\theta}}(\boldsymbol{a}) = \prod_{i=1}^{N} (\theta_i \mathbb{I}\{a_i = 1\} + (1 - \theta_i) \mathbb{I}\{a_i = 0\}).$$

We again observe that there is a natural link between the policy gradient and the EMC.



(a) Policy Gradient (b) EMC (c) Shapley value

Figure 1: Gradient, EMC, and Shapley value for two individuals, a log-linear utility, a linear separable policy, and a selection cost $c = 0.3$. Light colours denote higher utility, dark colours lower utility. The maximum is indicated by a red cross.

Lemma 2. *For separable linear policies, if* $\pi_{\theta_j}(a_j = 0) > 0$*, then* $\nabla_{\theta_j} U(\pi_{\boldsymbol{\theta}}) = \frac{\Delta_j U(\pi_{\boldsymbol{\theta}})}{\pi_{\theta_j}(a_j=0)}$.

We observe that the gradient takes a similar form as in Lemma 2, where in this case the normalising factor is the reciprocal of the probability of rejecting individual j under policy $\pi_{\boldsymbol{\theta}}$. In particular, note that under any uniform policy, the gradient is equal to the EMC after uniform scaling.

4.2.2 Threshold policies

A different class of policies emerges when instead of separably parameterising over the population, we parameterise over the feature space $\mathcal{X}$. One natural example is a policy of logistic type when $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{|\mathcal{X}|})$ and the probability of selecting individual i is

$$\pi_{\boldsymbol{\theta}}(a_i = 1 \mid \mathbf{x}_i) = \frac{e^{\boldsymbol{\theta}^T \mathbf{x}_i}}{1 + e^{\boldsymbol{\theta}^T \mathbf{x}_i}}.$$

These policies can be viewed as threshold policies since an individual i with feature vector $\mathbf{x}_i$ such that $\boldsymbol{\theta}^T \mathbf{x}_i > 0$ is being selected at least 50% of the time. In practice, these policies can be projected to the next closest vertex in the simplex so as to attain a deterministic threshold, namely, individual i is being selected only if $\boldsymbol{\theta}^T \mathbf{x}_i > 0$.

4.3 A policy gradient algorithm

In order to find the optimal policy for a fixed population, i.e. to solve (3), we propose to use a policy gradient algorithm. The general pseudo code is given in Algorithm 1.

Algorithm 1 Policy gradient algorithm

Input: A population $\mathcal{N}$ with features $\mathbf{x}$ and a utility function u.
Initialise $\boldsymbol{\theta}_0$, threshold $\delta > 0$, learning rate $\eta > 0$
while $\| \boldsymbol{\theta}_{i+1} - \boldsymbol{\theta}_i \| > \delta$ **do**
    Evaluate $\nabla U(\pi_{\boldsymbol{\theta}}, \mathbf{x})$ using data $\mathbf{x}$
    $\boldsymbol{\theta}_{i+1} \leftarrow \boldsymbol{\theta}_i + \eta [\nabla U(\pi_{\boldsymbol{\theta}}, \mathbf{x})]_{\boldsymbol{\theta} = \boldsymbol{\theta}_i}$
    i ++
end while
return $\pi_{\boldsymbol{\theta}_{i+1}}$

From Figure 1, we observe that the policy gradients and the EMCs have almost similar direction except for the edges where the probability of selecting an individual is almost 1. This illustrates the effect of the 'normalising' factors emerging in Lemmas 1 and 2. We also observe that the Shapley value and the gradients are similar near the point $(0.5, 0.5)$, i.e. when the policy is almost egalitarian. This validates our arguments of EMC reducing to the Shapley value under egalitarian policy.

4.4 Deviation from meritocracy as the residual EMCs

We do not state meritocracy in our setting as a optimisation constraint, but we are interested in the natural link between utility maximisation and meritocratic fair decisions. For this reason, we wish to quantify the *deviation from meritocracy* of policies in terms of the EMC. We suggest to use the cumulative positive EMC under a policy π:

$$\mathrm{Res}(\pi) = \sum_{i=1}^{N} \max(0, \mathit{\Delta}_i U(\pi)).$$

Here, large values for $\mathrm{Res}(\pi)$ indicate strong deviation from fully meritocratic decisions. Note that we decide to only account for positive EMCs as a negative EMC indicate negative potential contributions of individuals. From an individual's point of view, the EMC of an individual under a given policy represents her potential that remains unrecognized by the DM.

From Figure 1b, we observe that as the policy reaches the optimal solution, i.e. the probability of selecting the individual with higher contribution (individual 1) becomes higher, the residual EMCs are getting close to zero. This is demonstrated by the arrow pointing downwards in the bottom right corner, showing that individual 2 has negative expected contribution under a policy selecting individual 1 with high probability. We also observe that residual EMCs are always close to zero (decreasing arrow size) when both individuals are being selected with probability close to 1.

5 A Case Study: College Admissions

In this section, we perform an empirical case-study on fair set selection of individuals in a college admission system to test different algorithms with regards to achieved utility and our notion of meritocracy. The goal of the experiments is primarily to analyse the effect of different policy structures and secondarily to demonstrate that utility maximisation is sufficient for promoting group fairness notions (e.g. diversity or statistical parity).

5.1 Data and Experimental Setup

Data generation. Data is generated through simulation. From this, we first generate records for $4,000$ students, which represent the admitted students to the faculty from previous years. Each student is represented by a feature vector including a high school GPA for three specific fields (e.g. humanities or science), gender and high school. We assume that students come from 10 different schools and that the school to some extent depends on the individual's ethnicity (being majority, large minority, or small minority). In addition, each student is provided with a graduation result (numerical between 0 and 1), which we use to train a prediction model. The graduation grade is based on two latent variables: 1) talent: an innate quality, normally distributed across the population, and 2) skill: based on talent, but can be developed or hindered. Skill development is affected by school quality and societal gender biases. For evaluation, we generate another 500 students from the simulator from who we wish to select a subset. A detailed description of the simulator is given in the Appendix A.

Modelling graduation using regression. As we wish to estimate the success for each student, we train a linear regression model on the data D from the $4,000$ students to predict the graduation results of the students observed in the decision phase. We refer to these predictions as *predicted outcomes*. The trained predictor $\mathbf{f}$ takes the data $\mathbf{x}$ of students from this year as input and predicts the outcome $y_i \in [0, 1]$ of the i-th individual: $\mathbf{f}(i) \triangleq \mathbb{P}(y_i \mid \mathbf{x}, D)$.

Algorithms for set selection. We examine the behaviour of separable linear and threshold policies. For threshold policies, we use the high school GPA as a numerical three dimensional feature and add the gender and school as one-hot encoded categorical features. For both policy structures, we allow the policy gradient algorithms 250 updates to converge. In addition, we consider uniform set selection as a trivial lower baseline and for a robust benchmark, implement stochastic greedy set selection (see e.g.

Qian et al. [2018]) which essentially performs dynamic ranking of individuals according to marginal contributions.

Even though we believe that ranking is not particularly suitable for this problem, we can use *predicted outcomes* of an individual to rank them. We implement the cross-population noisy-rank algorithm of Kearns et al. [2017], which produces a probabilistic ordering over the population based on individuals' relative percentiles. In its original form, the noisy-rank algorithm requires us to only select an individual i if all higher ranked individuals are being selected as well. To combine this with a utility function, we select individuals in the order induced by the ranking until an individual has negative contribution. We will refer to this in our experiments as *noisy-rank*.

5.2 Designing utility functions

From a utilitarian social-welfare point of view, the DM is interested in the *outcomes* of her decisions. This requires having a utility function that depends on the actual outcomes - in the university scenario, for instance the graduation result encoded in some numerical value. In addition, the DM might have a cost associated with the selection of a student, so that it might only be favourable for the DM to select a student, if the gain in utility from adding the student to the set exceeds the cost c of selecting her. In case of *linear utility*, the DM's utility function $u : \mathcal{A} \times \mathcal{Y} \to \mathbb{R}$ then takes form

$$u(\boldsymbol{a}, \mathbf{y}) = \sum_{i \in \boldsymbol{a}} y_i - c \cdot |\boldsymbol{a}|,$$

where we slightly abuse notion when writing $i \in \boldsymbol{a}$ to mean that i is being selected under decision $\boldsymbol{a}$, i.e. $i \in \{k : a_k = 1\}$.

The DM might also be interested in promoting *diversity* in the study body, e.g. by admitting candidates from different backgrounds or demographics. For instance, the DM may wish to admit approximately the same number of female and male students. In our case, we can define group types, $\mathcal{T} = \{\text{gender}, \text{schools}\}$, and groups of such type, e.g. $\mathcal{G}_{\text{gender}} = \{\text{male}, \text{female}\}$. This allows us to define utilities accounting for the demographic background of applicants. A natural choice for a diversity promoting utility is the *log-linear utility over groups*:

$$u(\boldsymbol{a}, \mathbf{y}) = \sum_{T \in \mathcal{T}} \sum_{G \in \mathcal{G}_T} \log \Big(\sum_{i \in \boldsymbol{a} \cap G} y_i \Big) - c \cdot |\boldsymbol{a}|$$

as it naturally promotes egalitarian selection among groups.

In practice, we do not have access to the actual outcomes $\mathbf{y}$ of individuals and therefore use the trained predictor $\mathbf{f}$ to estimate the utility of a set by $U(\boldsymbol{a}, \mathbf{x}) = \mathbb{E}_{\mathbf{y} \sim \mathbf{f}}[u(\boldsymbol{a}, \mathbf{y})]$ (cf. Section 3).

5.3 Measuring diversity

In ecology, economics, and political science, quantifying diversity in form an index is studied historically [Simpson, 1949, Hill, 1973, Laakso and Taagepera, 1979, Jost, 2006, Chao et al., 2016]. In general, if individuals are chosen from C communities (or groups) according to some demographic feature, the *expected diversity* of order $l \in \mathbb{R}^+$ is defined as

$$\mathrm{Div}_l \triangleq \left(\sum_{i=1}^{C} p_i^l \right)^{1/(1-l)},$$

where p_i is the probability of selecting an individual from the i-th community. These are called Hill numbers in ecology [Hill, 1973]. Here, we use diversity of order $l = 2$, which is also called the inverse Simpson's index [Simpson, 1949], where higher values of Div_l represent higher diversity. This specific quantification of diversity is used to represent the effective number of parties in an election process [Laakso and Taagepera, 1979] or to represent the effective number of different species in an ecosystem [Chao et al., 2016]. We adopt this well studied diversity measure to quantify the expected number of groups having representatives in the selected set of individuals. As we have different types of groups in our situation, we take the sum of Div_2 over all group types (i.e. gender, school).

5.4 Performance Analysis

We report the results of the experiments for linear and log-linear utility in Tables 1 and 2 ,respectively. The experiments are conducted with a population of 500 students, and a selection cost per individual of 0.5 and 0.2 for linear and log-linear utility, respectively. We evaluate the algorithms according to the *true* outcomes of individuals generated by the simulator which are unavailable to the algorithms.

Table 1: Expected utility and residual EMCs w.r.t. actual outcomes for *linear utility* ($c = 0.5$), as well as inverse Simpson's index.

Algorithm	$U(\pi)$	$\text{Res}(\pi)$	Div_2
Linear	24.53	2.00	10.01
Threshold	13.03	8.83	10.00
Greedy	25.07	1.62	10.14
Noisy-rank	1.25	26.80	6.47
Uniform	-4.10	14.10	10.65

Table 2: Expected utility and residual EMCs w.r.t. actual outcomes for *log-linear utility* ($c = 0.2$), as well as inverse Simpson's index.

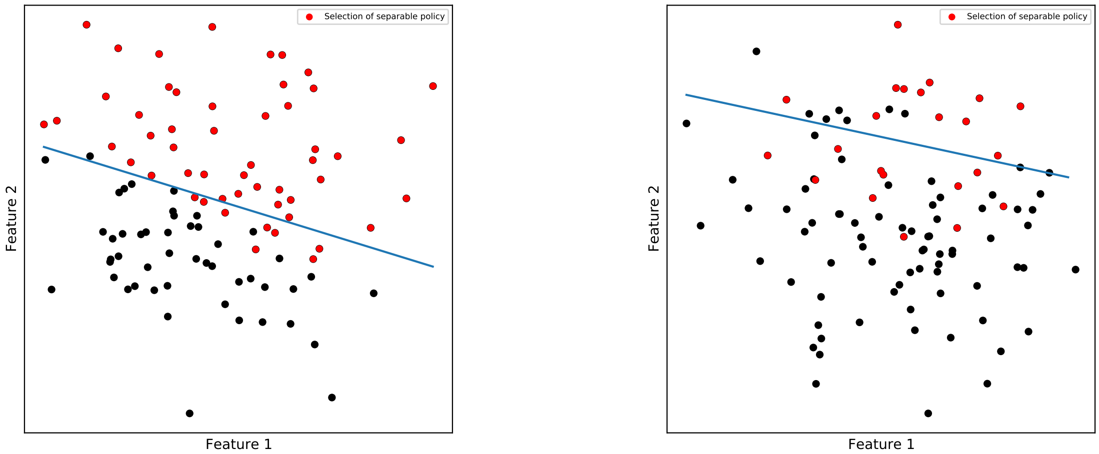
Algorithm	$U(\pi)$	$\text{Res}(\pi)$	Div_2
Linear	16.64	0.59	11.42
Threshold	12.34	1.69	10.24
Greedy	17.03	0.06	11.73
Noisy-rank	11.83	37.04	9.75
Uniform	7.63	8.91	10.71

From Tables 1 and 2, we observe that closeness to optimal expected utility corresponds to low residual EMCs. The residuals for the separable linear policy and the stochastic greedy algorithm are non-zero due to prediction errors of the trained model. With respect to the predicted outcomes, both algorithms achieve almost zero residuals for both utility types (cf. further results in Tables 4 and 5 in Appendix D). Thus, the residuals of the separable linear policy and the greedy algorithm can be seen as a deviation from meritocracy caused by the DM's imperfect predictions.

The threshold policy remains highly unstable as we do not project the policy to the next closest vertex in the simplex, which results in relatively low utility and high residuals. The aim of a threshold policy is to provide simple criteria for admission as these are attractive from the point of view of transparency and explainability. In order to guarantee such clarity and explainability, we want our policy to make decisions about individuals in the original feature space without mapping to higher dimensions. As a result, the threshold policy struggles whenever individuals are in the optimal set (i.e. have positive contribution), but are surrounded by sub-optimal individuals with similar features. This is visualised in Figure 2, where we conducted the experiments with a simplified simulator having only two grade features (apart from gender and school) for illustrative purposes.

Tables 1 and 2 show that meritocratic fairness notion of Kearns et al. [2017] is too restrictive in our setting as selection is stopped too early. Consequently, the residual EMCs are even higher than for the uniform policy as the majority of the population could have positively contributed but were not considered. The comparison between noisy-rank and the uniform selection also yields the observation that while close to optimal utility generally implies low residual EMCs, low utility does not necessarily imply high residuals. For instance, the policy that is selecting the whole population always achieves zero residuals while achieving low utility in general.

When comparing the inverse Simpson indices Div_2 for both utilities, we see that the log-linear utility successfully increases the diversity in the selected set compared to the linear utility. Here, the stochastic greedy algorithm and the separable linear policy achieve almost maximal diversity of 12, i.e. when every group has the same probability of selection. This is representative of the possibilities when using utility maximisation as a framework for group fair set selection as we could clearly design utilities promoting other group fairness notions such as statistical parity. In particular, when applying the log-linear utility and choosing students according to their marginal contribution we encode their demographic background implicitly into their "merit". This approach of utility design to promote affirmative actions is may be more appealing to the DM than designing different admission criteria for different groups.



(a) Linear utility

(b) Log-linear utility

Figure 2: The selection of the separable linear policy (red circles) and the soft decision boundary $\boldsymbol{\theta}^T \mathbf{x}$ determined by the threshold policy (blue line) illustrated in the feature space of 100 students generated with two numerical grade features (x-axis and y-axis) as well as gender and school. Each circle represents a student by means of her feature vector. Every student above the line has > 0.5 probability of selection under the threshold policy.

6 Discussion and Future Work

We have provided a first look into how one can define meritocracy in a general set selection scenario, where individuals contribute different amounts depending on the DM's policy. We have arrived at a natural definition relying on driving the expected marginal contribution of individuals to zero. Any policy satisfying this property cannot be improved locally by adding or removing individuals from the set. For separable policy structures, such policies can be obtained easily through greedy maximisation or gradient ascent.

However, in some cases we are interested in policies which make decisions only based upon individual characteristics, and without making an explicit comparison between individuals. It is possible to use policy gradient methods to find policies that maximise utility, but we have shown experimentally that these (at least for the case of a linear parametrisation) lead to inferior outcomes both in terms of meritocracy and utility.

For completeness, we have also compared our methods experimentally with an algorithm from the fair ranking literature, *noisy-rank*, in a college admission simulation. This method seeks not to admit "worse" individuals before "better" ones, and we have adapted it to our setting by admitting everyone in the order ranked by the algorithm until the DM's utility could not be further improved. Perhaps unsurprisingly, this method performed poorly, even though it used the same predictive model as the other approaches.

In future work, it will be interesting to examine policies that select fixed admission rules before actually seeing. In some cases, e.g. when there have to be fixed rules for selecting individuals from a population at the beginning of the process, the DM may need to choose a policy before seeing a particular population. Then we can assume that the DM has access to some distribution β over possible populations, and the problem becomes maximising expected utility under this distribution.

Another advantage of such policies in terms of fairness is that, under some specific policy structures, each individual can be judged without taking into account who else is in the current population. Instead, individuals will be judged relative to their expected contribution over all possible populations, and this will provide a plausible method for extracting an individual worth, which could imaginably be useful in some scenarios.

In many real-world applications, however, including the university admission scenario, the problem is more complicated, as there are multiple universities and students. The setting then becomes a matching problem, for which other concepts of fairness, such envy-free allocation may well be applicable.

We leave this question for future work.

Acknowledgements

Many thank you to Rachel Cao, Yang Liu, David Parkes and Goran Radanovic for long discussions about fairness in the set selection setting that partly served to inspire this line of work. We would also like to thank Anne-Marie George for discussions on the relation of this work to social choice and envy free allocation more generally, as well as Mauricio Byrd for contributing to early validation experiments.

References

Haris Aziz and Nisarg Shah. Participatory budgeting: Models and approaches, 2020.

Asia J Biega, Krishna P Gummadi, and Gerhard Weikum. Equity of attention: Amortizing individual fairness in rankings. In *The 41st international acm sigir conference on research & development in information retrieval*, pages 405–414, 2018.

L Elisa Celis, Damian Straszak, and Nisheeth K Vishnoi. Ranking with fairness constraints. *arXiv preprint arXiv:1704.06840*, 2017.

L. Elisa Celis, Anay Mehrotra, and Nisheeth K. Vishnoi. Interventions for ranking in the presence of implicit bias, 2020.

Anne Chao, Chun-Huo Chiu, and Lou Jost. Phylogenetic diversity measures and their decomposition: a framework based on hill numbers. *Biodiversity Conservation and Phylogenetic Systematics*, page 141, 2016.

Cristina Díaz-García, Angela González-Moreno, and Francisco Jose Saez-Martinez. Gender diversity within r&d teams: Its impact on radicalness of innovation. *Innovation*, 15(2):149–160, 2013.

Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*, pages 214–226, 2012.

Edith Elkind, Piotr Faliszewski, Piotr Skowron, and Arkadii Slinko. Properties of multiwinner voting rules. *Social Choice and Welfare*, 48(3):599–632, 2017.

Vitalii Emelianov, Nicolas Gast, Krishna P Gummadi, and Patrick Loiseau. On fair selection in the presence of implicit variance. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 649–675, 2020.

Mark O Hill. Diversity and evenness: a unifying notation and its consequences. *Ecology*, 54(2):427–432, 1973.

Lou Jost. Entropy and diversity. *Oikos*, 113(2):363–375, 2006.

Michael Kearns, Aaron Roth, and Zhiwei Steven Wu. Meritocratic fairness for cross-population selection. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1828–1836. JMLR. org, 2017.

Niki Kilbertus, Manuel Gomez Rodriguez, Bernhard Schölkopf, Krikamol Muandet, and Isabel Valera. Fair decisions despite imperfect predictions. In *International Conference on Artificial Intelligence and Statistics*, pages 277–287, 2020.

Jon M. Kleinberg and Manish Raghavan. Selection problems in the presence of implicit bias. *CoRR*, abs/1801.03533, 2018. URL http://arxiv.org/abs/1801.03533.

Matt J. Kusner, Chris Russell, Joshua R. Loftus, and Ricardo Silva. Causal interventions for fairness, 2018.

Markku Laakso and Rein Taagepera. “effective” number of parties: a measure with application to west europe. *Comparative political studies*, 12(1):3–27, 1979.

Martin Lackner and Piotr Skowron. Approval-based committee voting: Axioms, algorithms, and applications, 2020.

Sébastien Lahaie and David C Parkes. Fair package assignment. *Auctions, Market Mechanisms and Their Applications*, 2009.

Hervé Moulin. *Handbook of Computational Social Choice*. Cambridge University Press, 2016. doi: 10.1017/CBO9781107446984.

Chao Qian, Yang Yu, and Ke Tang. Approximation guarantees of stochastic greedy algorithms for subset selection. In *IJCAI*, pages 1478–1484, 2018.

Lloyd S Shapley. Notes on the n-person game—ii: The value of an n-person game. 1951.

Edward H Simpson. Measurement of diversity. *nature*, 163(4148):688–688, 1949.

Ashudeep Singh and Thorsten Joachims. Policy learning for fairness in ranking. In *Advances in Neural Information Processing Systems*, pages 5427–5437, 2019.

Julia Stoyanovich, Ke Yang, and HV Jagadish. Online set selection with fairness and diversity constraints. In *Proceedings of the EDBT Conference*, 2018.

Meike Zehlike, Francesco Bonchi, Carlos Castillo, Sara Hajian, Mohamed Megahed, and Ricardo Baeza-Yates. Fa* ir: A fair top-k ranking algorithm. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 1569–1578, 2017.

A Description of the Simulator

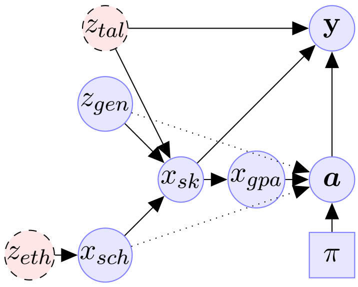


Figure 3: A simulator for the college admission and graduation with different schools, skills, gender and a DM policy. In dashed lines are relations that are only included in the prediction model and are not part of the historical decision process.

Students are admitted depending on the faculty's admission policy, which can use their grade information as well as which school they came from if necessary. We simulate a faculty with a specific admission profile. They prefer candidates whose grades match their profile, in practice, with a higher inner product between grade and faculty profile. This score is added with uniform noise to account for other admission factors such as a personal statement. The default admission process is that the top k matches are selected. Their graduation grade $\in [0, 1]$ depends on the faculty admission profile, their skills at the time of application and their innate talent in each topic[1]. A parameter γ determines the ratio between the talent and skill contributions to the graduation grade.

The *ethnicity* affects which high school a student will attend, with different high schools having different qualities. For that reason, high schools develop different skills to different amounts according to their profile.

The *gender* has two effects. First, different schools develop skills to a different extent depending on gender. Second, a general gender bias from society causes a positive, negative or neutral effect on skill development for different topics. This effect is the same for all schools. At the end of high-school, students obtain a grade based on their skill.

The notations in Figure 3 are as follows:

- z_{eth}—Ethnicity, z_{gen}—Gender, z_{tal}—Talent: independent variables
- x_{sch}—School: determined by Ethnicity (to simulate schools in geographical areas with different ethnicity distribution)
- x_{sk}—Skill: determined by Talent, School and Gender
- x_{gpa}—Grade: determined by School, Skill
- $\boldsymbol{a}$—Admission: determined by Grade, School, and policy
- $\mathbf{y}$—Graduation: determined by Skill and Talent (for admitted students)
- π—Admissions Policy: chosen by the DM

The utility of the DM $u(\boldsymbol{a}, \mathbf{y})$ should be related to the action $\boldsymbol{a}$ and the outcome $\mathbf{y}$.

B Modelling of the simulation

Though we generate the data using the simulator of Figure 3, the policies do not have access to the simulator. Instead, they can use a model built on data from the simulator that has been collected using a simple threshold policy as described in Section A.

[1]Note that we assume there are no biases affecting the performance of individuals once they are admitted.

Even though the graduation of individuals depends only on talent and skill, we do not have access to these variables. Thus, we employ a regression model using school grades, the schools and the gender of individuals. Linear regression was selected as the prediction model based on the validation scores obtained for several regression models, as shown in table 3.

Table 3: Mean squared error on validation set for different regression models. All models are trained with the default sklearn hyperparameters. Training set included 10000 samples and validation set of 500 samples, averaged over 5 repeated experiments.

Model	Mean Squared Error
Linear Regression	**0.0050686**
Bayesian Regression	0.0050690
Ridge Regression	0.0050737
Multi-layer Perceptron Regressor	0.0050967
Random Forest Regressor	0.0053579
Lasso Regression	0.0304602

Moreover, we only have labels (graduation grades) for individuals who have been admitted.

As shown in Fig. 4, the prediction error of the model increases with γ (talent skill ratio). This is caused due to the indirect relation between the latent variable talent and the observed features (school grades).

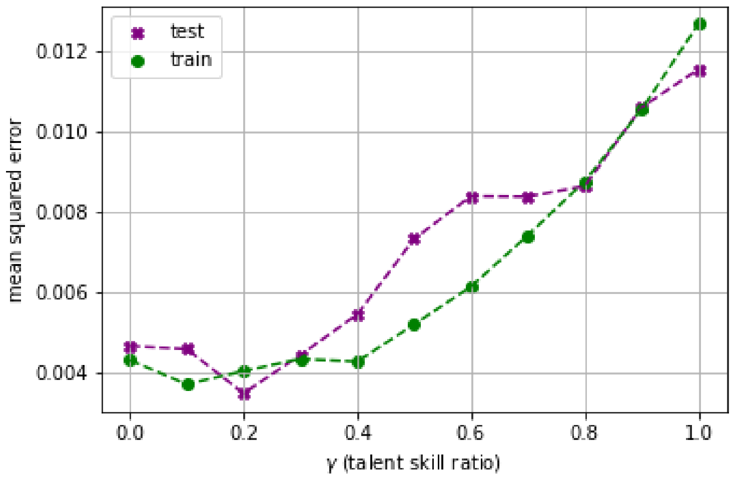


Figure 4: Mean squared error of the linear regression model, both on the train set and the test set, according to varying γ (talent skill ration). For $\gamma = 0$, the final graduation grade is based on skill alone, while for $\gamma = 1$ it is entirely based on the talent. These scores are averaged over 5 repeated experiments with training set of 10000 samples and test set of 500 samples.

B.1 Details of Dataset Generation

For each student the school comes from a discrete distribution returning one of ten schools when sampled. There is a different school distribution depending on the student's ethnicity. Ethnicity is directly sampled from a discrete distribution with probabilities $[0.7, 0.25, 0.05]$. In this case we use a total of three ethnicity, to simulate a majority group, large minority and small minority. The gender is sampled from a Bernoulli distribution with $p = 0.5$. Talent has three dimensions (for three topics), each one is sampled from a normal distribution with predefined mean and standard deviation. The school and gender define biases for each topic. These biases are added to the talent which serve together as the mean of a normal distribution from which the skill is sampled. The school grades are determined by the skill level combined with uniform noise in a specific range.

C Proofs

C.1 Proof of Lemma 2

Denote the set of decisions rejecting individual j by $\mathcal{A}_{-j} \triangleq \{\boldsymbol{a} : a_j = 0, a_i \in \{0,1\} \text{ for } i \neq j\}$. We omit writing $\mathbf{x}$ here and calculate the gradient to be

$$\begin{aligned}
\frac{\partial}{\partial \theta_j} U(\pi_{\boldsymbol{\theta}}) &= \frac{\partial}{\partial \theta_j} \sum_{\boldsymbol{a} \in \mathcal{A}} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) U(\boldsymbol{a}) \\
&= \sum_{\boldsymbol{a} \in \mathcal{A}} \frac{\partial}{\partial \theta_j} \pi_{\theta_j}(a_j) \prod_{i \neq j} \pi_{\theta_i}(a_i) U(\boldsymbol{a}) \\
&= \sum_{\boldsymbol{a} \in \mathcal{A}} (\mathbb{I}\{a_j = 1\} - \mathbb{I}\{a_j = 0\}) \prod_{i \neq j} \pi_{\theta_i}(a_i) U(\boldsymbol{a}) \\
&= \sum_{\boldsymbol{a} \in \mathcal{A}_{-j}} \prod_{i \neq j} \pi_{\theta_i}(a_i) U(\boldsymbol{a} + j) - \sum_{\boldsymbol{a} \in \mathcal{A}_{-j}} \prod_{i \neq j} \pi_{\theta_i}(a_i) U(\boldsymbol{a}) \\
&= \sum_{\boldsymbol{a} \in \mathcal{A}_{-j}} \prod_{i \neq j} \pi_{\theta_i}(a_i) [U(\boldsymbol{a} + j) - U(\boldsymbol{a})] \\
&= \frac{1}{\pi_{\theta_j}(a_j = 0)} \sum_{\boldsymbol{a} \in \mathcal{A}_{-j}} \pi(\boldsymbol{a}) [U(\boldsymbol{a} + j) - U(\boldsymbol{a})] \\
&= \frac{1}{\pi_{\theta_j}(a_j = 0)} \sum_{\boldsymbol{a} \in \mathcal{A}} \pi(\boldsymbol{a}) [U(\boldsymbol{a} + j) - U(\boldsymbol{a})] \quad \text{(for } \boldsymbol{a} \text{ with } a_j = 1 \text{ the difference is zero)} \\
&= \frac{\Delta_j U(\pi_{\boldsymbol{\theta}})}{\pi_{\theta_j}(a_j = 0)}.
\end{aligned}$$

We used the equality

$$\sum_{\boldsymbol{a} \in \mathcal{A}} \mathbb{I}\{a_j = 1\} \prod_{i \neq j} \pi_{\theta_i}(a_i) U(\boldsymbol{a}) = \sum_{\boldsymbol{a} \in \mathcal{A}_{-j}} \prod_{i \neq j} \pi_{\theta_i}(a_i) U(\boldsymbol{a} + j)$$

in the 4th equality of the proof and for completeness, rigorously prove it here. The LHS is equal to

$$\sum_{\boldsymbol{a} \in \mathcal{A} : a_j = 1} \prod_{i \neq j} \pi_{\theta_i}(a_i) U(\boldsymbol{a}) = \sum_{\boldsymbol{a} \in \mathcal{A} : a_j = 1} \prod_{i \neq j} \pi_{\theta_i}(a_i) U(\boldsymbol{a} + j).$$

Now, let $\boldsymbol{a}$ and $\boldsymbol{a}'$ only differ in the j-th element, namely $a_j = 1$ and $a'_j = 0$. Clearly,

$$\prod_{i \neq j} \pi_{\theta_i}(a_i) = \prod_{i \neq j} \pi_{\theta_i}(a'_i) \quad \text{and} \quad U(\boldsymbol{a} + j) = U(\boldsymbol{a}' + j),$$

which yields

$$\sum_{\boldsymbol{a} \in \mathcal{A} : a_j = 1} \prod_{i \neq j} \pi_{\theta_i}(a_i) U(\boldsymbol{a} + j) = \sum_{\boldsymbol{a} \in \mathcal{A}_{-j}} \prod_{i \neq j} \pi_{\theta_i}(a_i) U(\boldsymbol{a} + j).$$

C.2 Proof of Lemma 1

Define $Z(\boldsymbol{\theta}) = \sum_{\boldsymbol{a} \in \mathcal{A}} e^{\beta \boldsymbol{\theta}^T \boldsymbol{a}}$. Note that

$$\frac{\partial}{\partial \theta_j} \log Z(\boldsymbol{\theta}) = \frac{1}{Z(\boldsymbol{\theta})} \sum_{\boldsymbol{a} : a_j = 1} e^{\beta \boldsymbol{\theta}^T \boldsymbol{a}} = \beta \sum_{\boldsymbol{a} : a_j = 1} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}),$$

and set $\pi_{\boldsymbol{\theta}}(\boldsymbol{a} \text{ with } a_j = i) \triangleq \sum_{\boldsymbol{a}:a_j=i} \pi_{\boldsymbol{\theta}}(\boldsymbol{a})$ for $i \in \{0,1\}$. In the following, we omit the factor β as it will be nothing but a constant factor in the calculations. Straight forward calculation now yields

$$
\begin{aligned}
\frac{\partial}{\partial \theta_j} U(\pi_{\boldsymbol{\theta}}) &= \sum_{\boldsymbol{a}\in\mathcal{A}} \frac{\partial}{\partial \theta_j} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) U(\boldsymbol{a}) \\
&= \sum_{\boldsymbol{a}\in\mathcal{A}} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) \frac{\partial}{\partial \theta_j} \log(\pi_{\boldsymbol{\theta}}(\boldsymbol{a})) U(\boldsymbol{a}) \\
&= \sum_{\boldsymbol{a}\in\mathcal{A}} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) \frac{\partial}{\partial \theta_j} (\boldsymbol{\theta}^T \boldsymbol{a} - \log Z(\boldsymbol{\theta})) U(\boldsymbol{a}) \\
&= \sum_{\boldsymbol{a}\in\mathcal{A}} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) \big(a_j - \sum_{\boldsymbol{a}':a'_j=1} \pi(\boldsymbol{a}')\big) U(\boldsymbol{a}) \\
&= \sum_{\boldsymbol{a}:a_j=1} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) U(\boldsymbol{a}) - \pi_{\boldsymbol{\theta}}(\boldsymbol{a} \text{ with } a_j = 1) U(\pi_{\boldsymbol{\theta}}) \\
&= (1 - \pi_{\boldsymbol{\theta}}(\boldsymbol{a} \text{ with } a_j = 1)) \sum_{\boldsymbol{a}:a_j=1} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) U(\boldsymbol{a} + j) - \pi_{\boldsymbol{\theta}}(\boldsymbol{a} \text{ with } a_j = 1) \sum_{\boldsymbol{a}:a_j=0} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) U(\boldsymbol{a}) \\
&= \pi_{\boldsymbol{\theta}}(\boldsymbol{a} \text{ with } a_j = 0)\, e^{\theta_j} \sum_{\boldsymbol{a}:a_j=0} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) U(\boldsymbol{a} + j) - \pi_{\boldsymbol{\theta}}(\boldsymbol{a} \text{ with } a_j = 1) \sum_{\boldsymbol{a}:a_j=0} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) U(\boldsymbol{a}) \\
&= \pi_{\boldsymbol{\theta}}(\boldsymbol{a} \text{ with } a_j = 1) \sum_{\boldsymbol{a}:a_j=0} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) U(\boldsymbol{a} + j) - \pi_{\boldsymbol{\theta}}(\boldsymbol{a} \text{ with } a_j = 1) \sum_{\boldsymbol{a}:a_j=0} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) U(\boldsymbol{a}) \\
&= \pi_{\boldsymbol{\theta}}(\boldsymbol{a} \text{ with } a_j = 1) \sum_{\boldsymbol{a}:a_j=0} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) [U(\boldsymbol{a} + j) - U(\boldsymbol{a})] \\
&= \pi_{\boldsymbol{\theta}}(\boldsymbol{a} \text{ with } a_j = 1) \sum_{\boldsymbol{a}\in\mathcal{A}} \pi_{\boldsymbol{\theta}}(\boldsymbol{a}) [U(\boldsymbol{a} + j) - U(\boldsymbol{a})] = \pi_{\boldsymbol{\theta}}(\boldsymbol{a} \text{ with } a_j = 1)\, \mathit{\Delta}_j U(\pi_{\boldsymbol{\theta}}).
\end{aligned}
$$

D Additional results from experiments

In addition to Tables 1 and 2, we include comprehensive tables from our experiments comprising the performance of the algorithms with respect to actual outcomes as well as predicted outcomes:

Table 4: Expected utility $U(\pi)$ and residual EMCs $\mathrm{Res}(\pi)$ w.r.t. *actual* outcomes, expected utility $U_{\mathbf{f}}(\pi)$ and residual EMCs $\mathrm{Res}_{\mathbf{f}}(\pi)$ w.r.t. *predicted* outcomes for *linear utility*, as well as inverse Simpson's index.

Algorithm	$U(\pi)$	$\mathrm{Res}(\pi)$	$U_{\mathbf{f}}(\pi)$	$\mathrm{Res}_{\mathbf{f}}(\pi)$	Div_2
Linear	24.53	2.00	23.70	0.31	10.01
Threshold	13.03	8.83	10.41	7.21	10.00
Greedy	25.07	1.62	23.98	0	10.14
Noisy-rank	1.25	26.80	1.01	23.00	6.47
Uniform	-4.10	14.10	-11.47	12.00	10.65

Table 5: Expected utility $U(\pi)$ and residual EMCs $\mathrm{Res}(\pi)$ w.r.t. *actual outcomes*, expected utility $U_{\mathbf{f}}(\pi)$ and residual EMCs $\mathrm{Res}_{\mathbf{f}}(\pi)$ w.r.t. *predicted outcomes* for *log-linear utility over groups*, as well as inverse Simpson's index.

Algorithm	$U(\pi)$	$\mathrm{Res}(\pi)$	$U_{\mathbf{f}}(\pi)$	$\mathrm{Res}_{\mathbf{f}}(\pi)$	Div_2
Linear	16.64	0.59	16.30	0.38	11.42
Threshold	12.34	1.69	11.65	1.00	10.24
Greedy	17.03	0.06	16.69	0	11.73
Noisy-rank	11.83	37.04	11.97	33.00	9.75
Uniform	7.63	8.91	6.67	0.30	10.71