Long-term dynamics of data-driven targeted support for job seekers

Sebastian Scher, Know-Center Graz

June 21, 2021

1 Introduction

Research question: Assume we have a population of individuals, where the prospect on the labor market of each individual is controlled by the personal skill of that individual, described by a set of independent skill features. Additionally, each individual belongs to one of two groups, described by a protected attribute (e.g. gender). The average skill between the two groups is not the same, but still there is significant overlap between the groups. If an observer has knowledge about all skill features of an individual, they can with 100% accuracy compute the total skill of that individual, and knowledge about the protected attribute (i.e. which group the individual belongs to) does not yield any additional information with regard to the individual skill. We assume that the labor market (e.g. recruiters) do have access to all skill features, and thus make their decisions solely on the real personal skill of the individuals, as for them there is no additional information in the protected attribute.

We further assume that there is a public authority that helps individuals in improving their skill on the labor market, which we will call the Public Employment Service (PES). The improvement of the skill of the individuals by the PES is done with individualized services, based on the current skill on the labor market of that individual. Critically, however, this public authority does not have access to all skill features, but only to a subset. In addition, it has knowledge about the protected attribute. Since the total skill is not evenly distributed across the two groups, the knowledge of which of the protected groups the individuals belongs combined with historical data gives probabilistic information on the real skill: if the individual belongs to the group that has on average higher skill, the likelihood that this individual has high skill is larger than if it would belong the the other group, even if everything else remains the same. This additional information however has two potential problems: 1) it is only probabilistic and the resulting predictions are only accurate *on average*, and 2) it is based on a protected attribute - therefore, legal and/or ethical reasons might prohibit utilizing that information, as it results in different treatments based solely on the protected attribute.

The goal of this study is to investigate the long-term effects on the population if a public authority provides targeted help, based partly on protected attributes.

Does such targeted help reduce or strengthen existing group-inequalities? What is the long-term effect on general employment? What is the long-term effect on employment in each group? How does this compare to a system that is *not* using the protected attribute? Are there trade-offs between the - ethically problematic - inclusion of protected attributes in the targeting versus the global goal of high employment?

To address these questions, we use a combination of dynamical numerical modeling ans statistics/machine-learning. We use synthetic data and an individual skill model that is as simple as possible while being just sophisticated enough to reflect inequalities and the possibility of having either full of only part knowledge of the skill of an individual.

Additionally, we make - admittedly simplistic - assumptions on the labor market.

We develop and use two different overall models: the first model is a simple model that has a fixed population, whose increase of skill over time is influenced by the PES. The targeting of the PES is again influenced by the past mean skill of the two groups, and this past-skill is updated over time as the population changes. This model is generic and abstract, and could also be interpreted for example as a model for targeted education, and labor market dynamics are only included in a very abstract way. Still, this model is interesting, as it is easy to follow and allows to test our intuition, and already captures the basic dynamics over time.

The second model is more complex and considers a pool of job-seekers. This pool has an influx and an outflux, and considers not only the abstract skill of an individual, but additionally the time an individual needs to find a job. This is more related to actual job-markets.

A central aspect of this study is that we consider different scenarios of the PES and how it distributes its resources across individuals with different skills.

What this study is not: This study is not a real-world assessment of any real world public employment service or similar institutions.

- Things to add:
- efficiency argument, we assume there is only a certain amount of help or support available, and the PES wants to distribute in efficiently (whereas "efficient" needs to be defined), this is the goal of the EPS
- combination of statistics/ML and system dynamics

1.1 Literature overview

literature on AMS algorithm

literature on fairness in general
debiasing

[1] showed that simulation studies can be used to study fairness issues in changing systems, and that the results ("fair" or "not fair" with repsect to a certain definition of fairness) can differ from a static analysis. They analyzed simple settings of credit-scoring for loans, attention allocation for different sites, and college admissions.

References

[1] Alexander D'Amour, Hansa Srinivasan, James Atwood, Pallavi Baljekar, D. Sculley, and Yoni Halpern. Fairness is not static: Deeper understanding of long term fairness via simulation studies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, FAT* '20, pages 525–534, New York, NY, USA, January 2020. Association for Computing Machinery.

2 Methods

2.1 Personal skill model and data generation

The basis of this study is the following setting: we have a population of N individuals. Each individual has a personal skill s_{real} that is composed of two independent skill-features x_1 and x_2.

$$s_{real} = \frac{1}{2}\left(x_1 + x_2\right) \tag{1}$$

We call it "real" because we will differentiate it from observed and from predicted/assumed skill and from the skill that the labor market ("recruiters) assigns to the inviduals later on.

In addition, each individual has a binary protected attribute x_{pr} that can have values of 0 and 1. In reality this could for example be gender, but here it is used in an abstract way. Central here is that the definition if s_{real} does not contain x_{pr}.

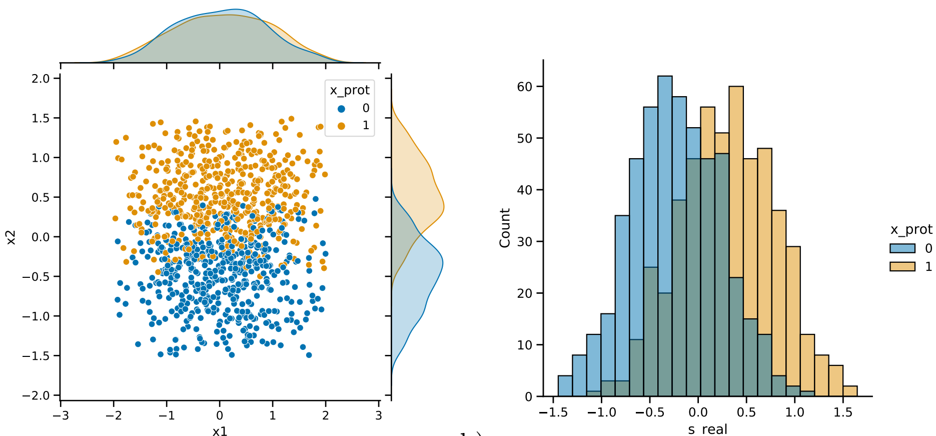
Further we assume that x_1 is completely independent of x_{pr}, but x_2 is correlated with x_{pr}.

$$x_{1=}\sigma_{trunc}\left(0,1\right) \tag{2}$$

$$x_{pr} = [0,1] \tag{3}$$

For generating our artificial population, we draw x_1 from a truncated normal distribution, and x_{pr} from the binary distribution $[0,1]$ with uniform probability.

Thus, the probability of an individual belonging to a particular group (with respect to the protected attribute) is 50% for both groups.



a) b)

Figure 1: Initial population. a) distribution of the two skill features, b) distribution of total skill, both split up according to the binary protected attribute

The second skill feature, x_2, is generated with

$$x_2 = \frac{1}{2}\left(\alpha_{pr} \cdot \left(x_{pr} - \frac{1}{2}\right) + \sigma_{trunc}\left(0,1\right)\right) \tag{4}$$

the factor $\frac{1}{2}$is subtracted from x_{pr} to ensure that it is centered around zero. When x_2 is generated this way, then individuals in group 0. have *on average* lower x_2, and therefore on average lower s_{real}. To reflect this, we will from now on call the group with individuals with $x_{pr} = 0$ the *underprivileged group*. Importantly, however, not all individuals in the underprivileged group have low x_2 and low s_{real}. There are individuals in the privileged group that have lower skill than some individuals in the underprivileged group, and there are individuals in the underprivileged group that have a skill that is above the population mean. The joint distribution of x_1 and x_2 and the distribution of s_{real} of the initial population is shown in fig. 1.

From the way x_1 x_2 are generated and the fact that, s_{real} per definition (e.q. (1)) can be completely inferred from x_1 and x_2 follow two central facts: given x_1 and x_2, there is no additional information contained in x_{pr} when one wants to infer s_{real}. If, however, one has 1) only access to x_1 and at the same time 2) information on the distribution of s_{real} over the two groups (e.g. the mean of s_{real} seperately for each group), and one wants to infer s_{real}, then including x_{pr} in addition to x_1 a statistical prediction system yields additional information, even though s_{real} is completely defined by x_1 and x_2. This will from the backbone of the whole study.

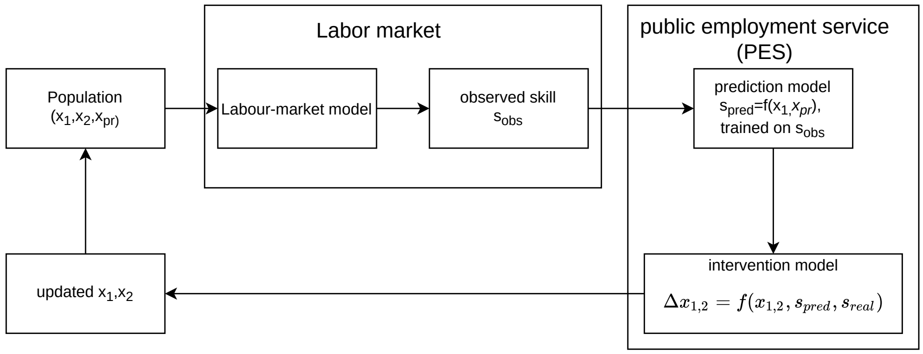


Figure 2: Sketch of the simple dynamical model

2.2 Simple dynamical model

The simple model assumes that the population is made up of a fixed pool of individuals, without any in- or outflux. The job-market and individual employment is modeled in very abstract and indirect way.

2.2.1 labor market model

We use two different definitions of the labor market: a biased, and and biased one. Each ones assigns an effective skill . In the unbiased model labor-market prospect is simply the skill s_{real}:

$$x_{eff} = x_{real} \tag{5}$$

, thus the labor market has oracle-access to the real skill and does not alter it in any way. Note that this makes the - quite unrealistic - assumption that the labor-market 1) has perfect information on the skill of individuals and 2) is discrimination free in terms of individual fairness (individuals are not treated differently based on their protected attribute). Important to point out that this labor market is *not* discrimination free under definitions of fairness based on group-fairness: as the labor-market does not even have access to the protected attribute, it can by definition not ensure group-fairness constraints, such as requiring that from each group the same fraction of people is hired, which in this abstract model would mean that the computation of s_{real} would need to be modified to ensur that each group has the same mean skill.

For the biased labor market model, the effective skill is partly dependent on x_{pr}:

$$x_{eff} = \begin{cases} x_{real} + \alpha_{lb} & x_{pr} = 1 \\ x_{real} - \alpha_{lb} & x_{pr} = 0 \end{cases} \tag{6}$$

For the unprivileged group the skill is made smaller due to the bias, and for the privileged group larger, by the same biase-parameter α_{lb}

2.2.2 Prediction model

The prediction model is the statistical model used by the hypothetical PES. It is used to predict the labor-market prospects of an individual. In this abstract setting the labor-market prospect is, as mentioned above, s_{real}. The prediction model classifies individuals on having below or above average s_{real}. In a real setting this would be estimated from historical data (potentially updated from time time, or at each timestep). Here it is estimated from the current distribution of s_{real} over the two population groups.

As mentioned before, the basis of this study is that the PES has access to an incomplete set of skill-features only, namely solely to x_1, and additionally access to x_{pr}. To estimate (predict) the prospect class (above or below average s_{real}) from this, logistic regression is used to create the main (full) prediction model:

$$P\left(s_{real} > \gamma | x_1\right) = \frac{1}{1 + e^{-(\alpha_1 x_1 + \alpha_2 x_{pr} + \beta)}} \tag{7}$$

whereas the parameters α_1, α_2 and β are estimated from the current population, and γ is the threshold set for dividing the low and the high prospect class.

Additionally, we use a second prediction model that does not use x_{pr}

$$P\left(s_{real} > \gamma | x_1\right) = \frac{1}{1 + e^{-(\alpha_1 x_1 + \beta)}} \tag{8}$$

Per definition of the data generation process, this base model has lower accuracy than the full prediction model.

In the initial timestep, γ is set to the mean of s_{real}, which equals to zero by definition of how the data is generated. For the following timesteps, we use two different strategies in the experiments: 1) *constant* γ, thus the threshold between the low and the high prospect class does not change, and 2) *adaptive*, where γ is set to the population-mean of s_{real}.

2.2.3 Intervention Model

The intervention model describes how the effect that the PES has on the individuals. For the simple model, we define the support that the PES provides in a singe timestep as a change in the individual skill features x_1 and x_2 of an individual. We make the change dependent on the current value, with decreasing increments as the skill feature grows, approaching the limits set by the constants x_1^{max} and x_2^{max}:

$$x_1^{t+1} = max\left(x_1^t + k_1\left(x_1^{max} - x_1^t\right), x_1^t\right) \tag{9}$$

$$x_2^{t+1} = max\left(x_2^t + k_2\left(x_2^{max} - x_2^t\right), x_2^t\right) \tag{10}$$

The model parameters k_1 and k_2 define how fast x_1 and x_2 grow. For simplification, we set $k = k_1 = k_2$,and thus both skill features have the same growth rate. The value of k is central to our study, as it defines how the intervention models affects different people. To this end, we make k dependent both on the real prospect class C_r, and prospect class C_{pr} predicted by the prediction model. The fact that the growth rate is made depended on the predicted class reflects the idea of targeted help for different prospect classes, and that a prediction model is used for this. The real world idea is not only that different prospect classes receive a different *quantity* of help, but also a different *quality* that is better suited for that prospect group. Therefore, we also make k dependent on the real prospect group of each individual, as arguably if a certain type of help is better suited for the low than for the low prospect group, than this will have the adverse effect that for a incorrectly classified person it will be especially unsuited.

Since both C_r and C_{pr} are binary, this leads to a $2x2$ matrix k_{ij}:

	predicted low	predicted high
real low	k_{11}	k_{12}
real high	k_{21}	k_{22}

With different values for k_{ij} we can now define different scenarios. The difference between $k_{11}$and k_{22}defines how different the effect is of the intervention model is for the two different prospect classes, as intended by the PES. The difference between k_{11} and k_{21} and between k_{12} and k_{22} defines how individuals are adversely affected if the are incorrectly classified by the prediction algorithm, and receive the type of help that is in fact intended for the other group.

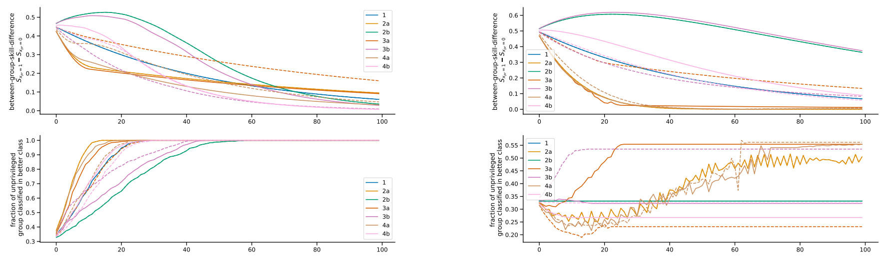
The values for the different entries of k_{ij} define how - in the abstract setting of our model - “attention” or “resources” are distributed across the different groups. In order to make the different scenarios better comparable, we constrain the absolute values in such a way that the geometric mean of all 4 entries must be the same in all scenarios. The gemoetric mean is the usually used mean functinon for variables related to growth-rates and similar. We set it to 1/50. This number was chosen more or less arbitrarily so that a typical simulation with constant decision function reaches its equilibrium on the order of 100 timesteps. Changing the scale would only change how much happens in a single timestep of the simulation, and not the overall behavior (with higher k, we will get the same result earlier, with smaller k, one needs to run the model for more timesteps). Since also the unit of a timestep is abstract and cannot be translated to a real world time, this does not change the results or their interpretation.

For better readability, the k-values presented in the text and the plots are all scaled with a factor of 50, so for the k-matrix for scenario 1 all entries will be displayed as 1.

We will use the following scenarios:

1. “Agnostic”: $k_{11} = k_{12} = k_{21} = k_{22}$. This is the base scenario, where all classes receive both the same quantity and quality of help, and this also has the same effect, independent of the actual labor-market prospects.

2. no targeting, but class-dependent effect: $k_{11} = k_{12}$, $k_{21} = k_{22}$, two subscenarios: a) $k_{11} > k_{22}$, b) $k_{11} < k_{22}$. In these scenarios, the PES privides



a) b)

Figure 3: Upper panels: time evolution of the between-group skill difference (mean skill of privileged group - mean skill of underprivileged group), lower panels: time evolution of fraction of individuals in from the underprivileged group that are classified as belonging to the high-prospect group. a) with constant decision function, b) with adaptive decision function.All for unbiased labor market.

the same help to everyone, but this helps works better (a) or worse (b) for individuals with (real) low prospects.

3. targeting, but no class-dependent effect. In this scenario k is only dependent on C_{pr}, not on C_r, $k_{11} = k_{21}$, $k_{12} = k_{22}$
4. targeting with class dependent effect.a)

The scenarios are summarized in table 1.

2.3 Complex Model

not yet implemented, not yet decided whether we actually need it.

2.4 Metrics

3 Results

4 Discussion and conclusion

What if only "decision support", and the decision by the model is only input to a human? It is known that humans tend to rely on such systems (https://journals.sagepub.com/doi/10.1518/001872097778543886)

Coda and Data Availability

The software for this study was written in Python and is published with this paper. It allows full reproduction of the results of this study as well as further

Table 1: overview of scenarios, defined by the k-matrix of the intervention model

1) no targeting, no class-dependent effect

	predicted high	predicted low
real high	1	1
real low	1	1

2a)no targeting, class-dependent effect (more on lowprospect group)

	predicted high	predicted low
real high	2	2
real low	1/2	1/2

2b)no targeting, class-dependent effect (more on highprospect group)

	predicted high	predicted low
real high	1/2	1/2
real low	2	2

3a) targeting (more on lowprospect group), no class-dependent effect

	predicted high	predicted low
real high	2	1/2
real low	2	1/2

3b) targeting (more on highprospect group), no class-dependent effect

	predicted high	predicted low
real high	1/2	2
real low	1/2	2

4a)

	predicted high	predicted low
real high	$4 \cdot 8^{-\frac{1}{4}}$	$8^{-\frac{1}{4}}$
real low	$8^{-\frac{1}{4}}$	$2 \cdot 8^{-\frac{1}{4}}$

4b)

	predicted high	predicted low
real high	$2 \cdot 8^{-\frac{1}{4}}$	$8^{-\frac{1}{4}}$
real low	$8^{-\frac{1}{4}}$	$4 \cdot 8^{-\frac{1}{4}}$

experimentation with the model parameters.