ETH zürich

Fair Decisions Despite Imperfect Predictions

Niki Kilbertus Manuel Gomez-Rodriguez Isabel Valera
Bernhard Schölkopf Krikamol Muandet
(Max Planck Institute)

Guillaume Thiry - 21.04.2021

Agenda

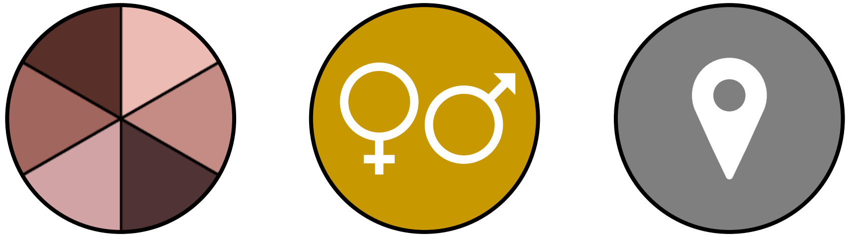
I. Context of the problem
II. Mathematical formalisation
III. Why deterministic policies don't work
IV. Stochastic policies and learning algorithm
V. Results
VI. Discussion on the paper

I. Context of the problem

Data-driven predictions today

Consequential decisions

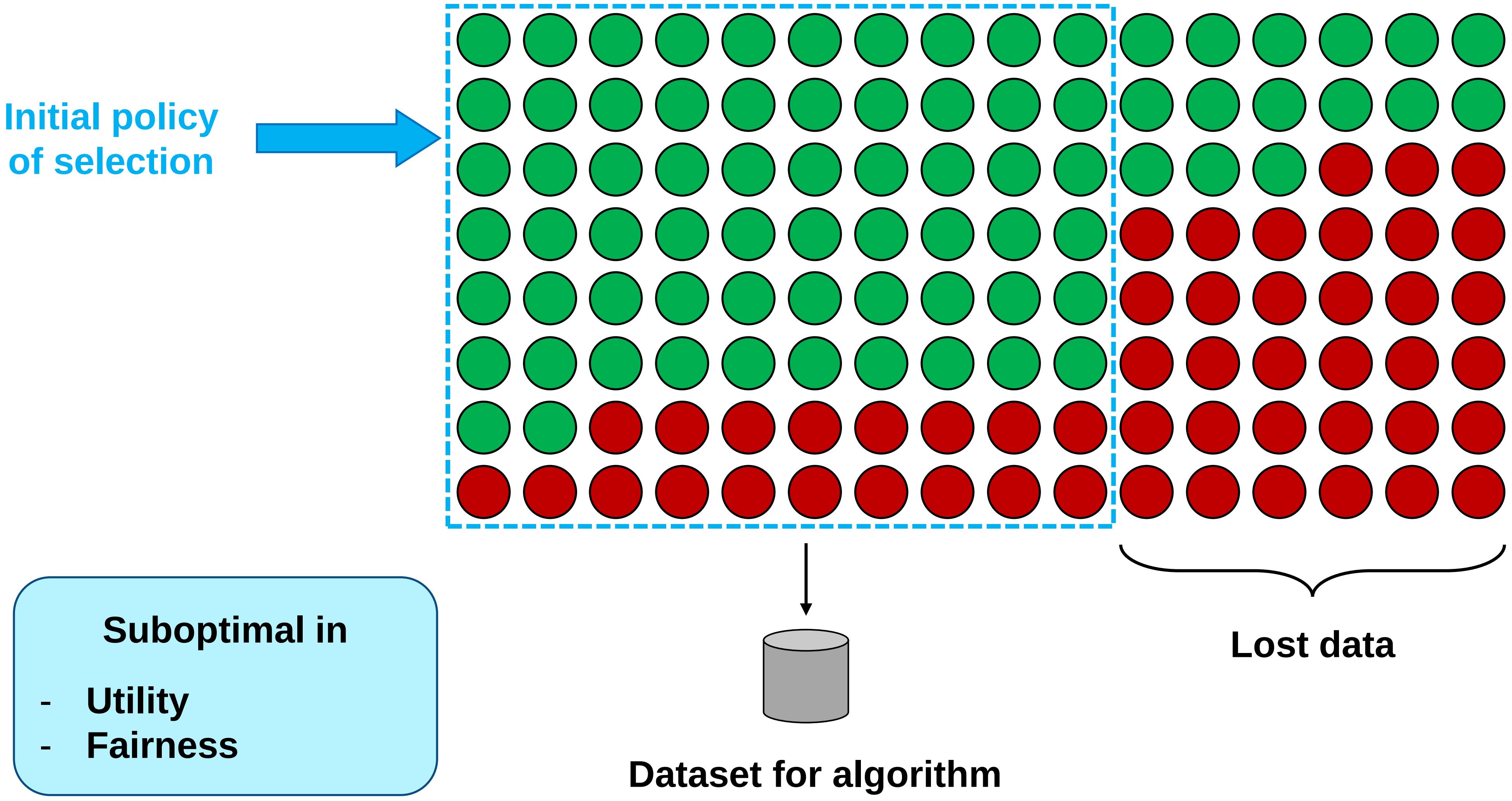
Discriminatory factors



→ **Utility**

→ **Fairness**

Imperfect dataset produced by selection



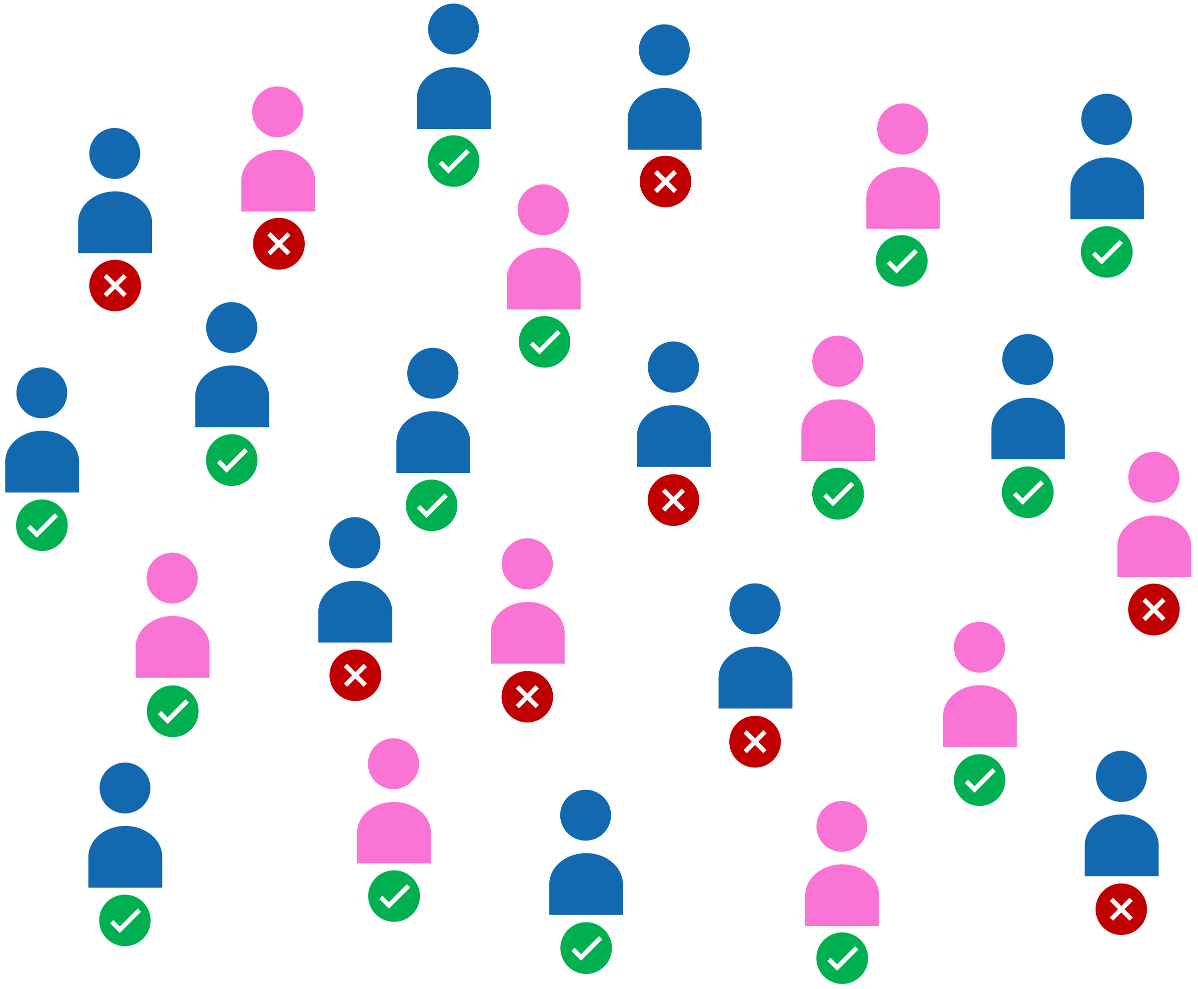
Suboptimal in

- **Utility**
- **Fairness**

What is fairness?

People applying for a loan:

- More men than women
- But women tend to repay their loan more often



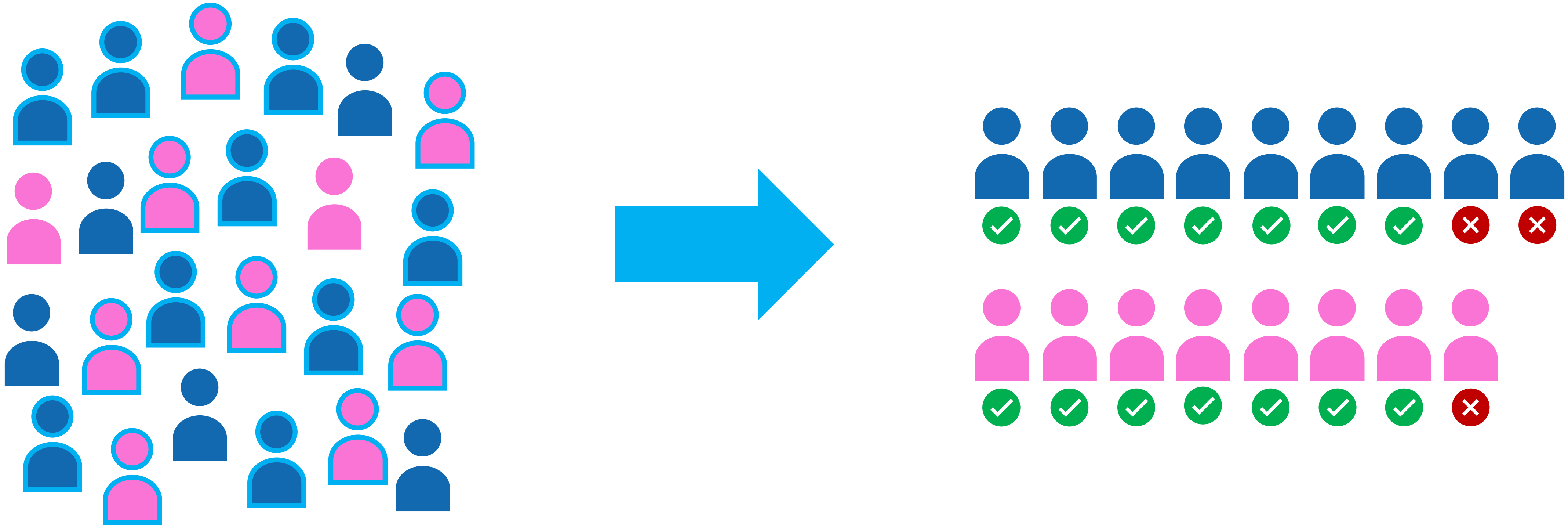
What is fairness?

Maximising utility without fairness

Only the utility matters, differences in selectivity among the groups are not relevant

What is fairness?

Maximising utility without fairness



→ A woman has more chance of being accepted than a man
→ But utility is maximal

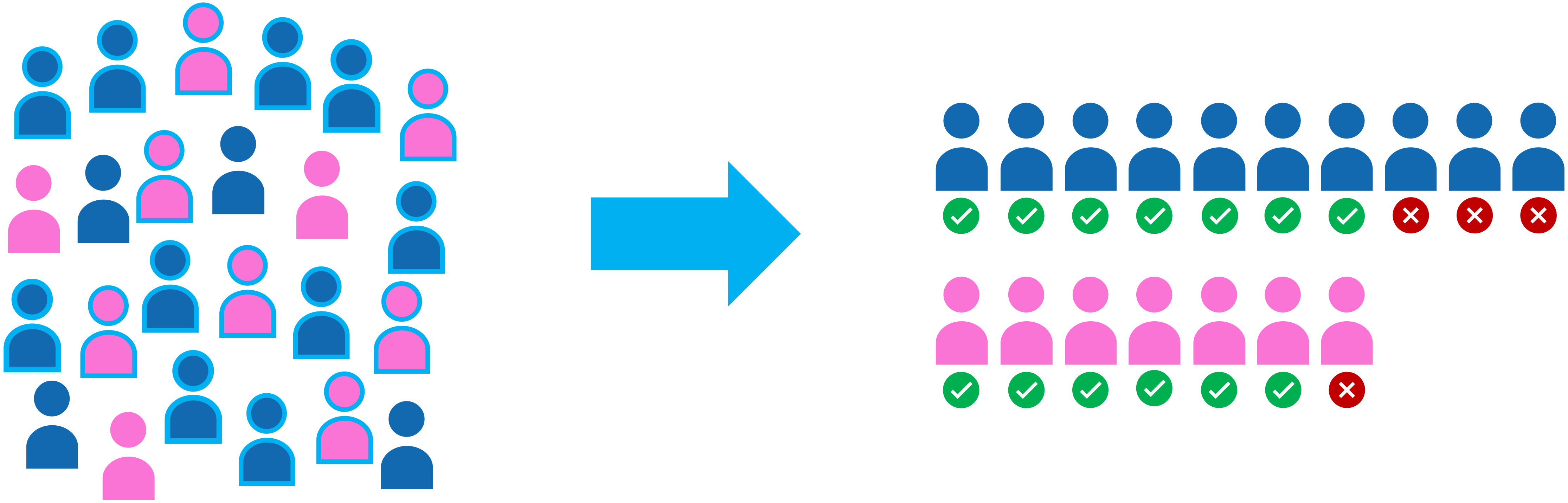
What is fairness?

Fairness by Demographic Parity

The probability of being accepted
has to be the same in both groups

What is fairness?

Fairness by Demographic Parity



→ A woman has as many chance of being accepted as a man
→ But the utility is suboptimal

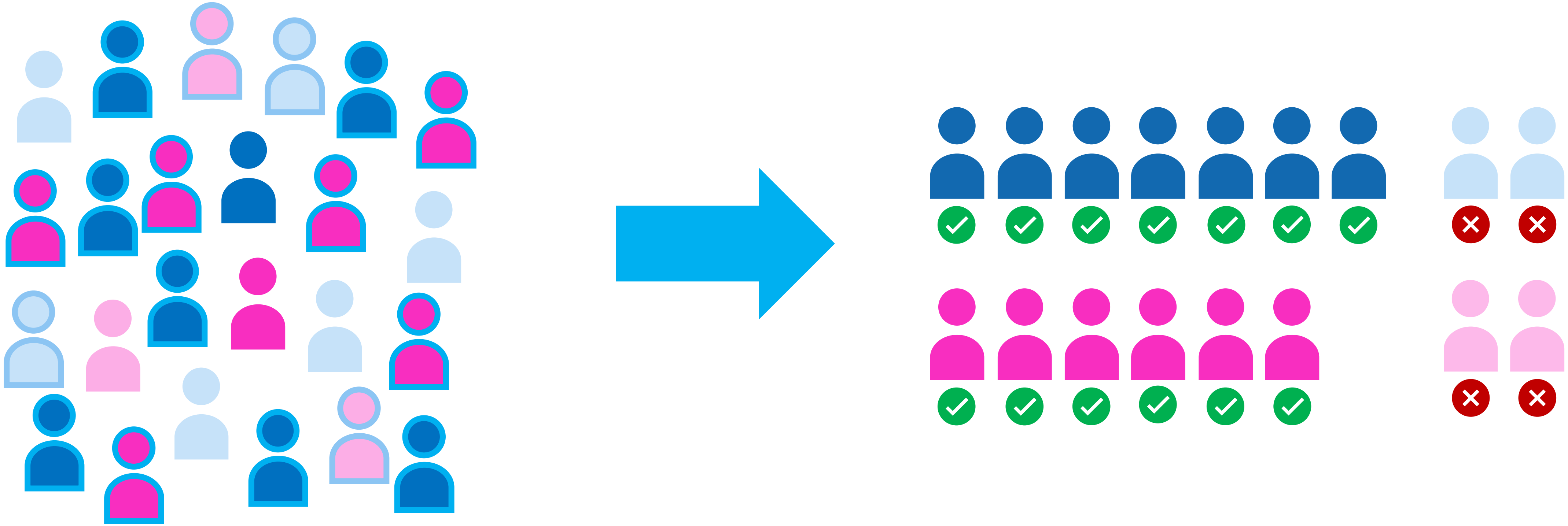
What is fairness?

Fairness by Equality of Opportunity

Individuals able to repay need to have equal probability of acceptance in both groups

What is fairness?

Fairness by Equality of Opportunity



→ A woman able to pay has as many chance of being accepted as a man able to pay
→ Utility is still suboptimal

II. Mathematical formalization

Mathematical formalisation

$X \subseteq \mathbb{R}^d$	Dataset features
$S = \{0,1\}$	Sensitive feature (race, gender…)
$Y = \{0,1\}$	Ground truth label
$\pi : X \times S \longrightarrow \mathcal{P}(\{0,1\})$	Policy (or decision rule)
$P(\boldsymbol{x}, s, y) = P(y\|\boldsymbol{x}, s)P(\boldsymbol{x}, s)$	Ground truth distribution

$$\mathrm{d} \sim \pi(d|\boldsymbol{x}, s)$$

decision made by the policy

$$\begin{cases} d = 0 & \text{loan refused} \\ d = 1 & \text{loan granted} \end{cases}$$

$$y \sim P(y|\boldsymbol{x}, s)$$

final outcome

$$\begin{cases} y = 0 & \text{default} \\ y = 1 & \text{loan repaid} \end{cases}$$

Utility

Utility under P:

$$u_P(\pi) \coloneqq \mathbb{E}_{x,s,y\sim P,d\sim\pi(x,s)}[yd - cd]$$

where $c \in [0,1]$ reflects economic considerations:

$$\textbf{Utility gain} \quad \begin{cases} 1\text{-}c & \text{if loan repaid} \\ \text{-}c & \text{if default} \end{cases}$$

Fairness

f-benefit for group s:
(s = 0 or 1)

$$b^{s}_{P(\pi)} \coloneqq \mathbb{E}_{x,y\sim P(x,y|s),d\sim\pi(x,s)}[f(d,y)]$$

$$f : \{0,1\} \times \{0,1\} \longrightarrow \mathbb{R} \qquad \begin{cases} f(d,y) = d & \text{demographic parity} \\ f(d,y) = d \cdot y & \text{equality of opportunity} \end{cases}$$

Fairness criteria: $b^{0}_{P(\pi)} = b^{1}_{P(\pi)}$

Some observations

Under perfect knowledge of $P(y|x,s)$, the policy maximizing the utility under fairness criteria is a deterministic threshold rule:

$$\pi^*(d=1|x,s) = 1[P(y=1|x,s) \geq c_s]$$

with cost factors c_0, c_1 being chosen to respect the fairness criteria

→ If no fairness criteria, just take $c_0 = c_1 = c$

→ π (probability) is either 0 or 1 so the decision is deterministic

→ **We can only have a finite approximation of the ground truth !**

Some observations

But it is even worse in reality, as an initial selection policy $\boldsymbol{\pi_0}$ will select a weighted distribution:

$$P_{\pi_0}(\boldsymbol{x}, s, y) \propto P(y|\boldsymbol{x}, s)\, \pi_0(d = 1|\boldsymbol{x}, s) P(\boldsymbol{x}, s)$$

not constant for all $(\boldsymbol{x}, s)$

→ This creates samples from the ground truth that are not i.i.d

III. Why deterministic policies don’t work

Mathematical setup

We can reformulate the problem with fairness constraints as an unconstrained problem with an additional penalty term:

$$v_P(\pi) \coloneqq u_P(\pi) - \frac{\lambda}{2}\left(b_P^0(\pi) - b_P^1(\pi)\right)^2$$

- Can we directly maximize $v_P(\pi)$ under the induced distribution P_{π_0}?

Propositions about deterministic policies

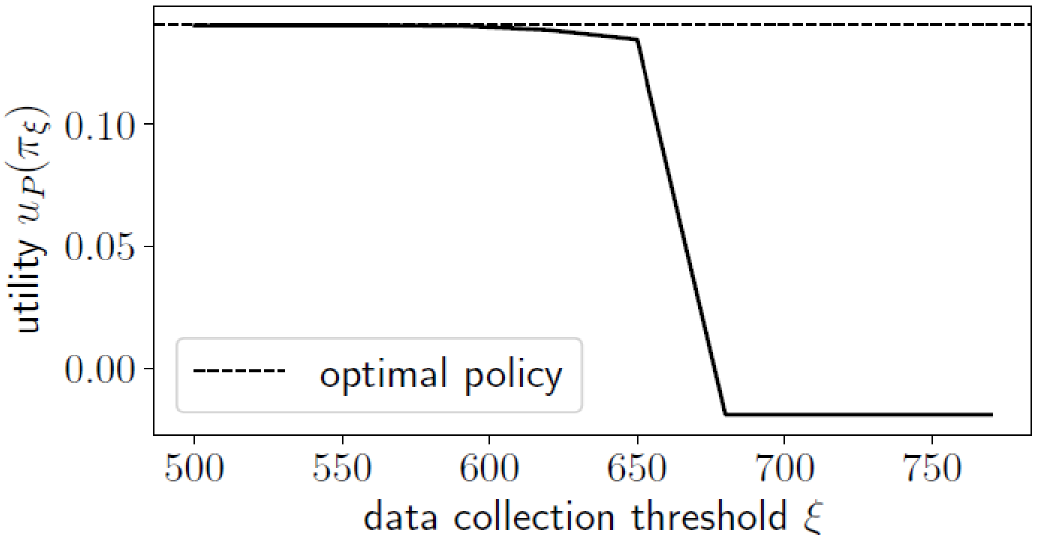
Example:

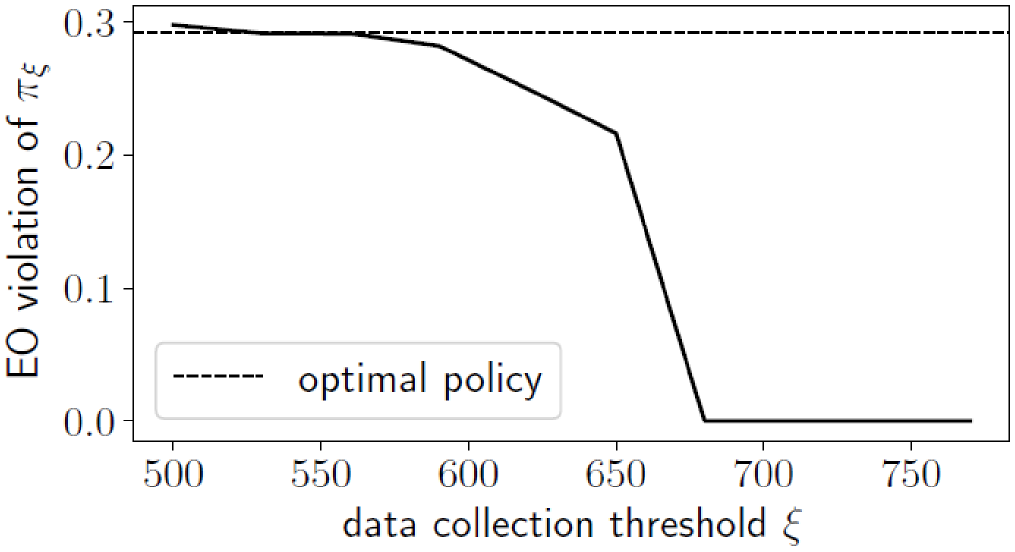
FICO scores $x \in X \coloneqq \{300, \dots, 820\}$

Initial policy π_0 based on a threshold $\xi \in X$

Collected dataset $\mathcal{D}^{(\xi)}$ to learn a model $Q_\xi(y = 1|x)$

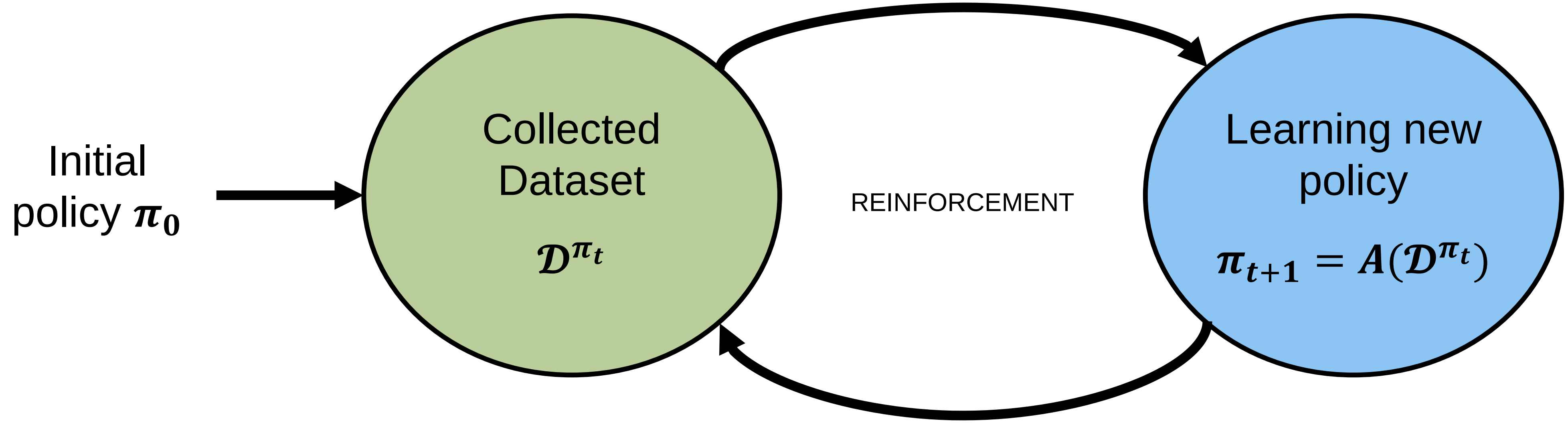
Resulting policy $\pi_\xi(d = 1|x) = \mathbf{1}[Q_\xi(y = 1|x) > c]$





Propositions about deterministic policies

- Can we perform a sequential policy learning task that will converge to the optimal solution?



→ Problems if the update rule A is non-exploring (error based)

IV. Stochastic policies and learning algorithm

Why stochastic policies?

A fully randomized initial policy would explore all the ground truth:

$$\pi_0(d = 1|\boldsymbol{x}, s) = \frac{1}{2} \quad \forall\, \boldsymbol{x}, s \qquad\qquad P_{\pi_0} = P$$

Then, if our class of policies is rich enough, we will be able to converge to π^*

→ **But unethical and unefficient initial policy**

A policy π is called an **exploring policy** if $\pi(d = 1|\boldsymbol{x}, s) > 0$ for any measurable subset $X \times S$ of with positive probability under P.

Learning a stochastic policy

logistic policy

$$\pi_\theta(d = 1|x, s) = \sigma(\Phi(x, s)^T \theta)$$

$\Phi : \mathbb{R}^d \times \{0,1\} \longrightarrow \mathbb{R}^m$ is a fixed feature map

$\sigma(z) = \frac{1}{1+e^{-z}}$ is the sigmoid function

$\theta \in \Theta \subset \mathbb{R}^m$ is the parameter we want to learn

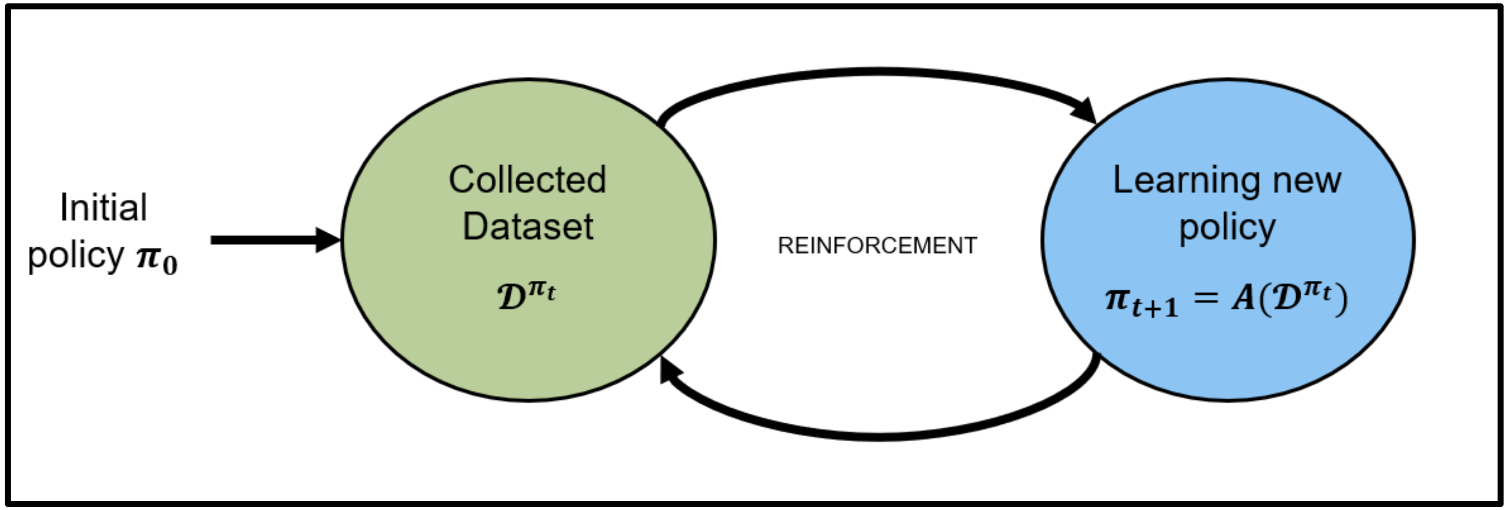
semi-logistic policy

$$\tilde{\pi}_\theta(d = 1|x, s) = 1[\Phi(x, s)^T \theta \geq 0] + 1[\Phi(x, s)^T \theta < 0]\, \sigma(\Phi(x, s)^T \theta)$$

Learning a stochastic policy

Learning method: Stochastic Gradient Ascent (SGA)

$$\theta_{i+1} = \theta_i + \alpha_i \cdot \nabla_\theta v_P\,(\pi_\theta)\Big|_{\theta=\theta_i}$$



Learning a stochastic policy

Algorithm 1 CONSEQUENTIALLEARNING: train a sequence of policies $\pi_{\boldsymbol{\theta}_t}$ of increasing $v_P(\pi_{\boldsymbol{\theta}_t})$.

Require: Cost c, time steps T, decisions N, iterations M, minibatch size B, penalty λ, learning rate α.

```
θ_0 ← INITIALIZEPOLICY()
for t = 0, ..., T − 1 do                                  ▷ time steps
    D^t ← COLLECTDATA(θ_t, N)
    θ_{t+1} ← UPDATEPOLICY(θ_t, D^t, M, B, α)
return {π_{θ_t}}_{t=0}^T

function COLLECTDATA(θ, N)
    D ← ∅
    for i = 1, ..., N do                                  ▷ N decisions
        (x_i, s_i) ~ P(x, s) and d_i ~ π_θ(x_i, s_i)
        if d_i = 1 then                                   ▷ positive decision
            D ← D ∪ {(x_i, s_i, y_i)} with y_i ~ P(y | x_i, s_i)
    return D                                              ▷ data observed under π_θ

function UPDATEPOLICY(θ', D, M, B, α)
    θ^(0) ← θ'
    for j = 1, ..., M do                                  ▷ iterations
        D^(j) ← MINIBATCH(D, B)                           ▷ sample minibatch
        ∇ ← 0, n_j ← 0
        for (x, s, y) ∈ D^(j) do                          ▷ accumulate gradients
            d ~ π_{θ^(j)}(x, s)
            if d = 1 then
                n_j ← n_j + 1
                ∇ ← ∇ + ∇_θ v(π_θ, π_{θ'})|_{θ=θ^(j)}
        θ^(j+1) ← θ^(j) + α ∇/n_j
    return θ^M
```

main loop

draw a data point from the distribution and take a decision

add the point to the collection set

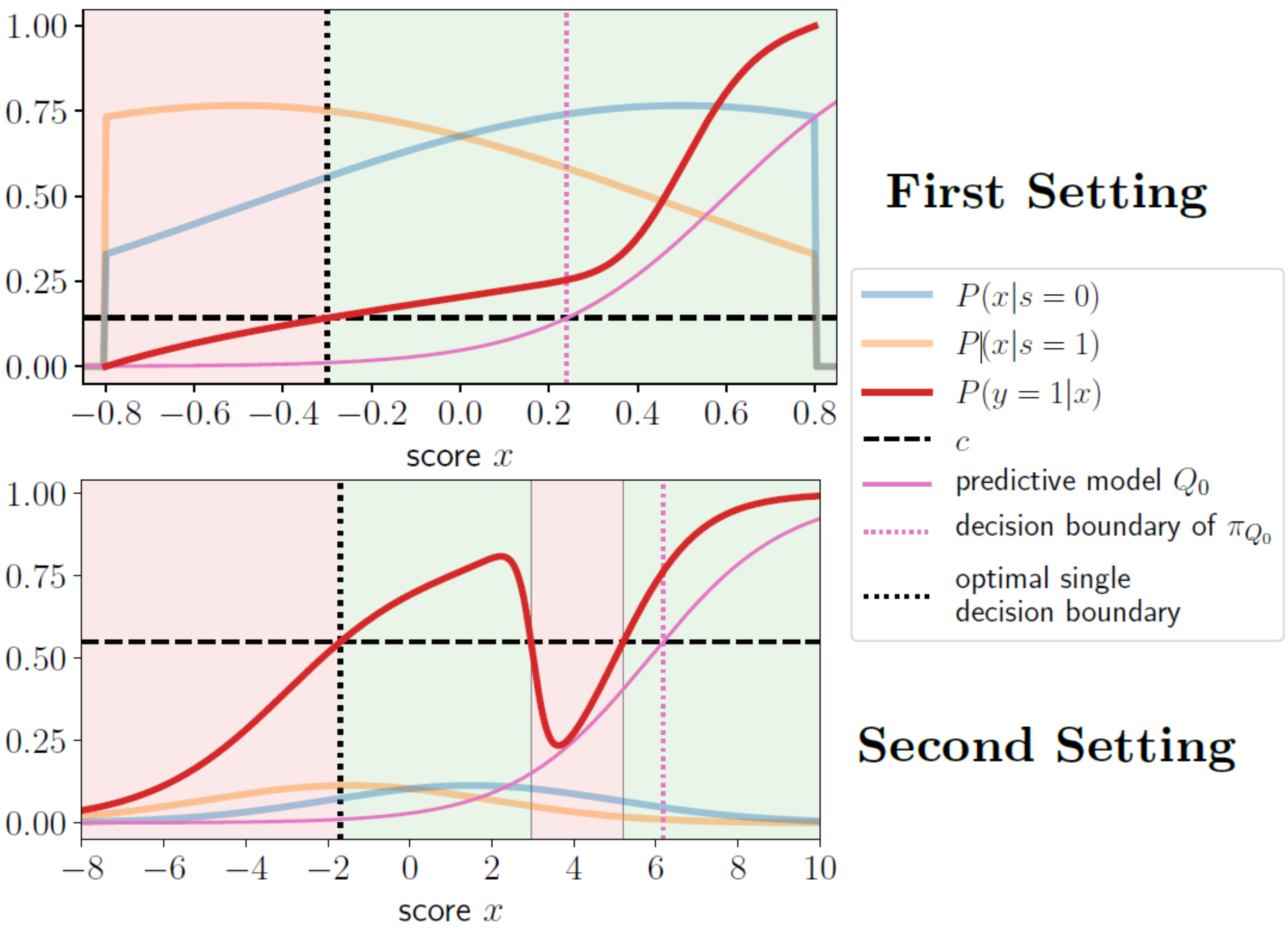
use of minibatches

computation of the gradient for the batch

update of the parameter by SGA

V. Results

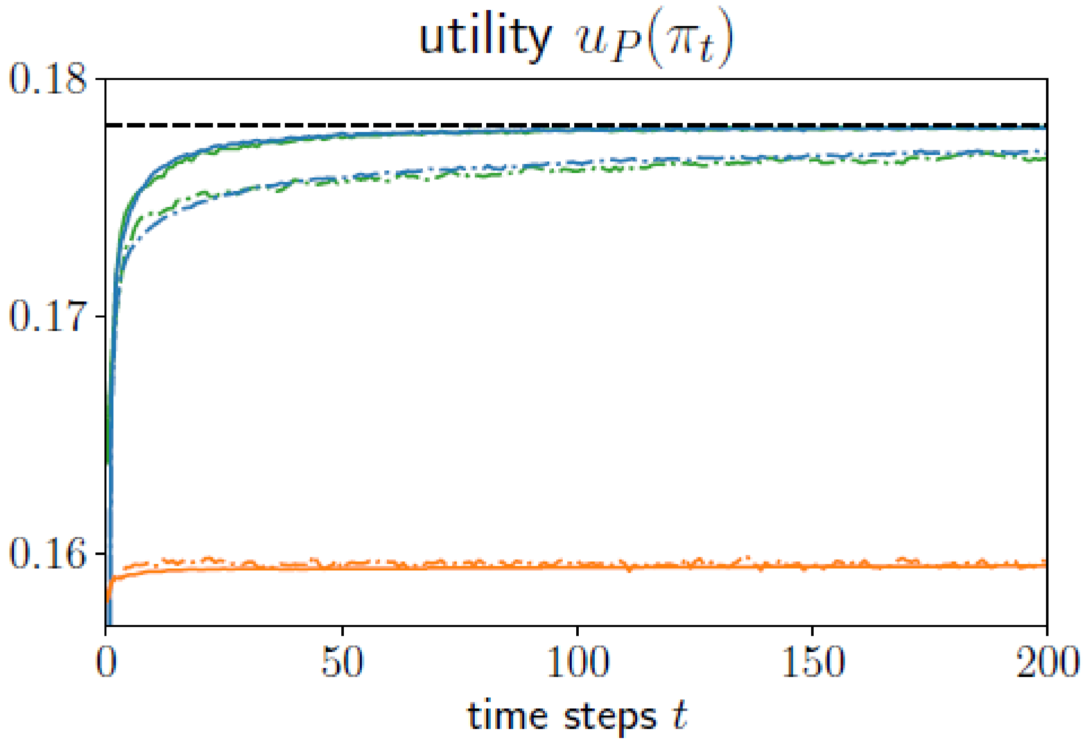
Synthetic data

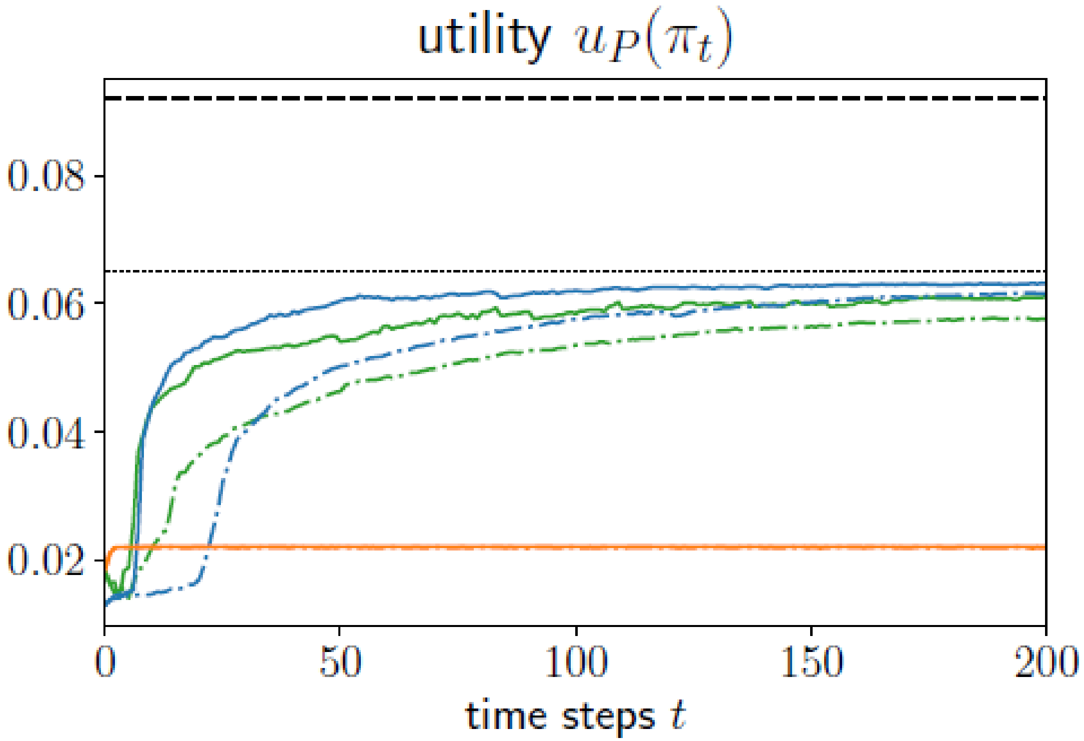


Synthetic data

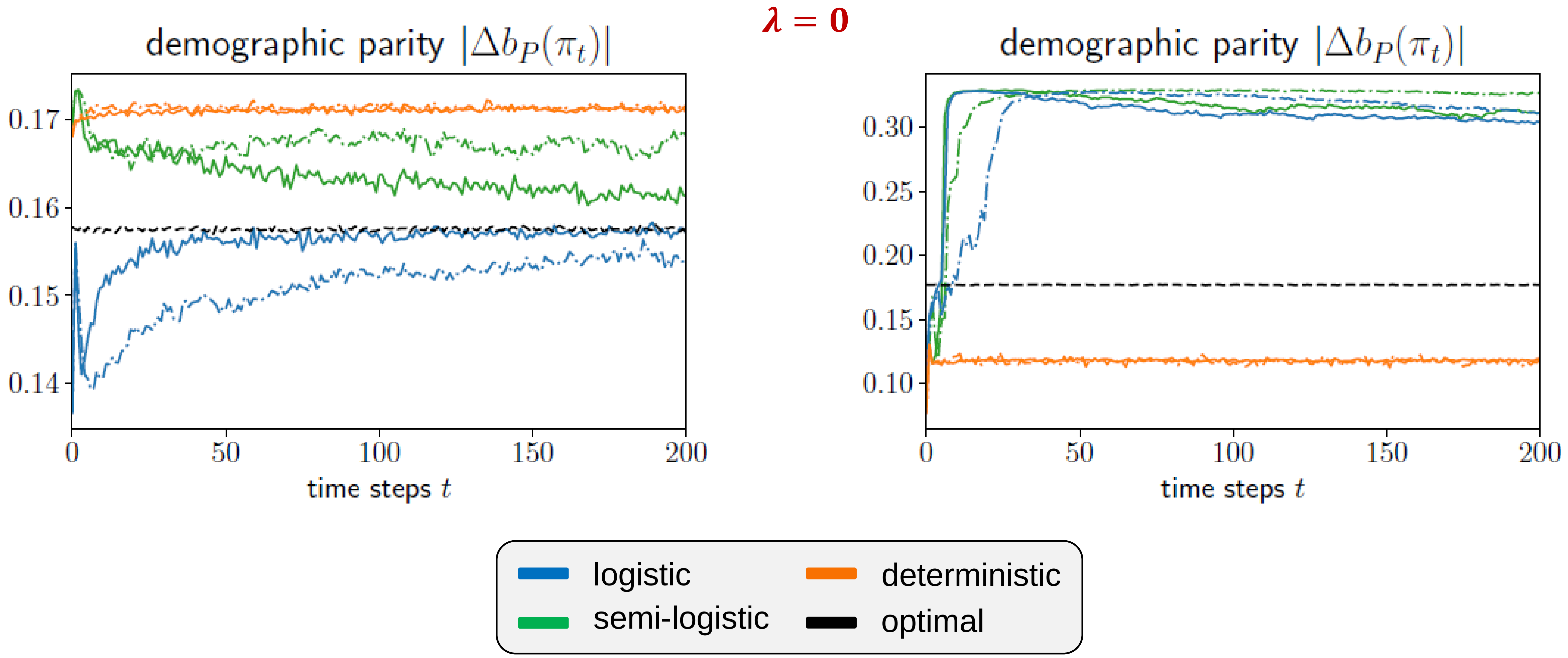
$$v_P(\pi) := u_P(\pi) - \cancel{\frac{\lambda}{2}\left(b_P^0(\pi) - b_P^1(\pi)\right)^2}$$

$\boldsymbol{\lambda = 0}$



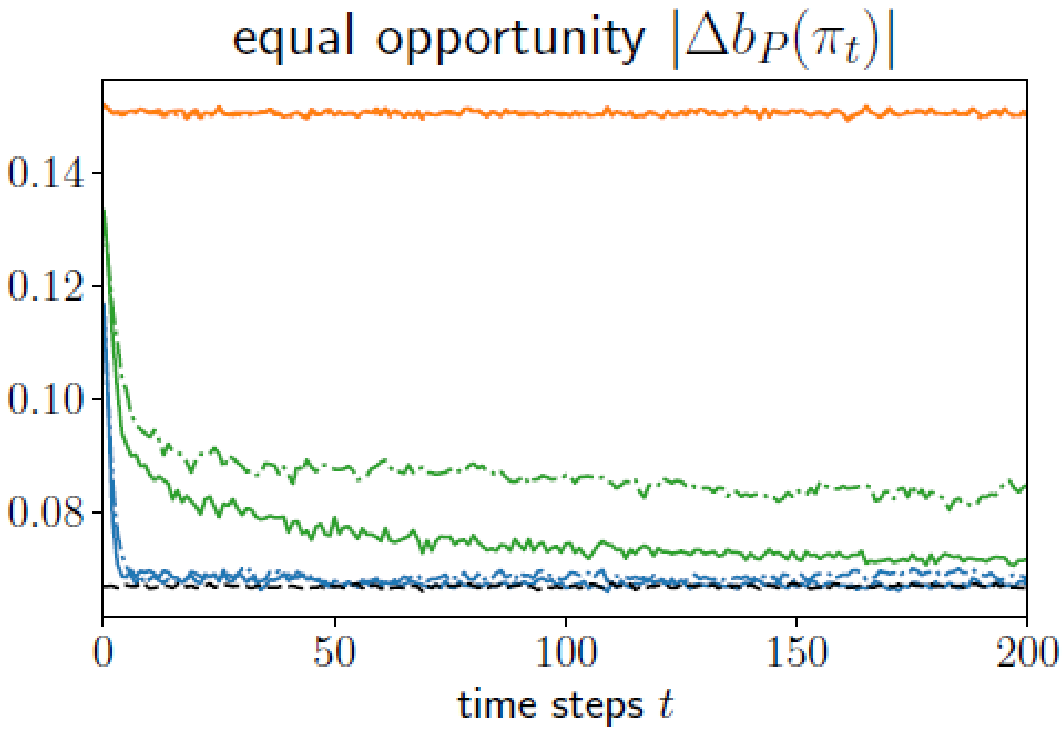


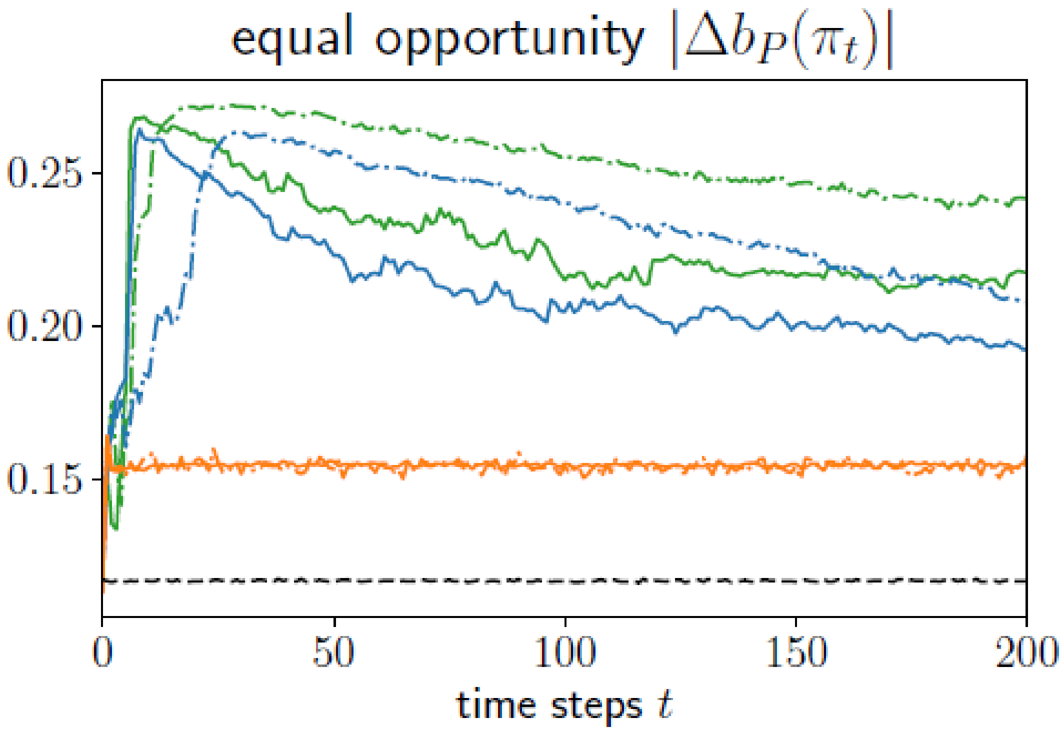
Synthetic data



Synthetic data

$\lambda = 0$





logistic
semi-logistic
deterministic
optimal

Real data

COMPAS dataset (criminality data)

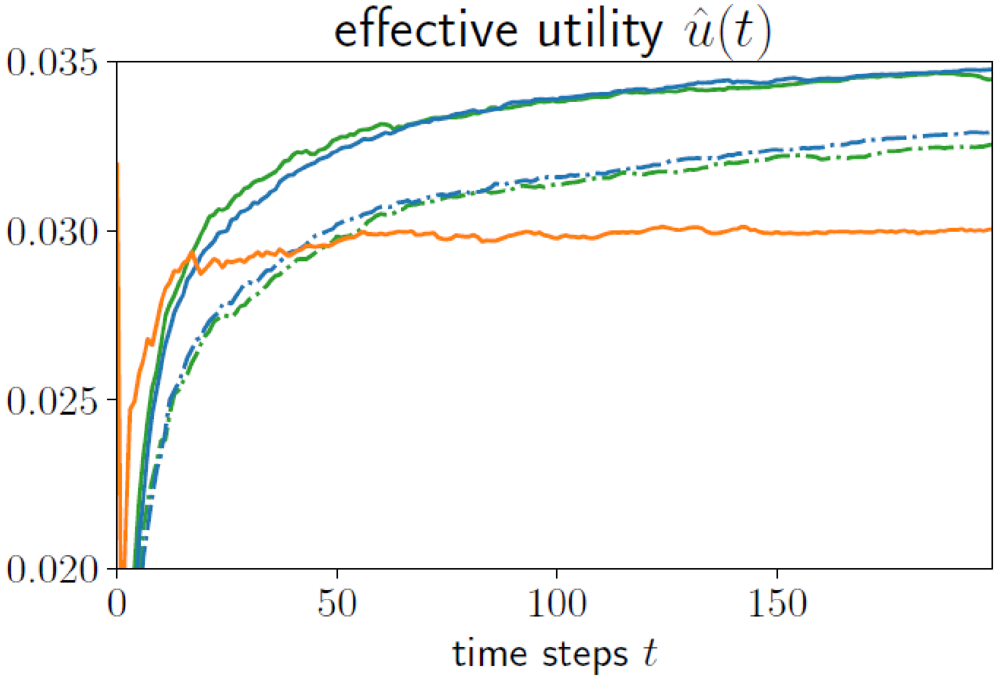
- Features: demographic, criminal history
- Sensitive feature: “white” or not

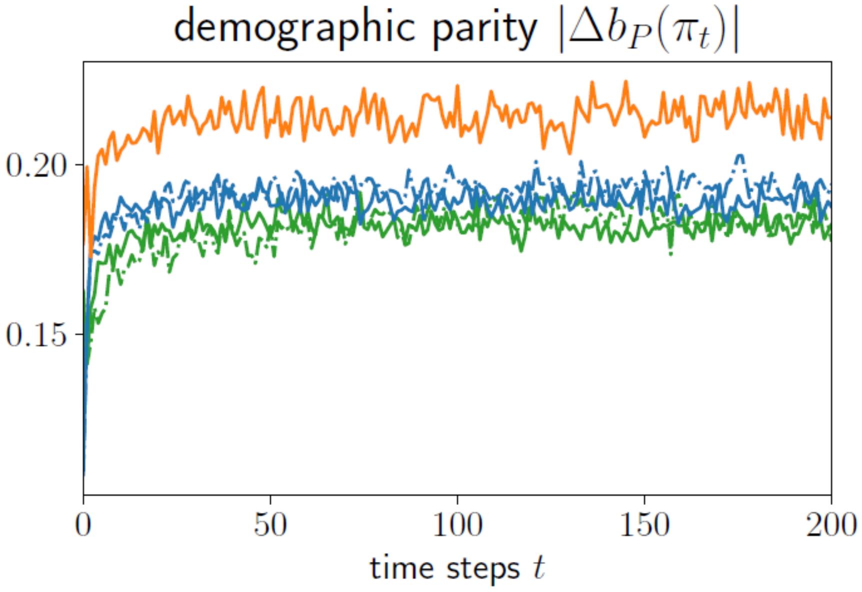
- 80% data for the training (used over many time steps)
- 20% remaining for the testing

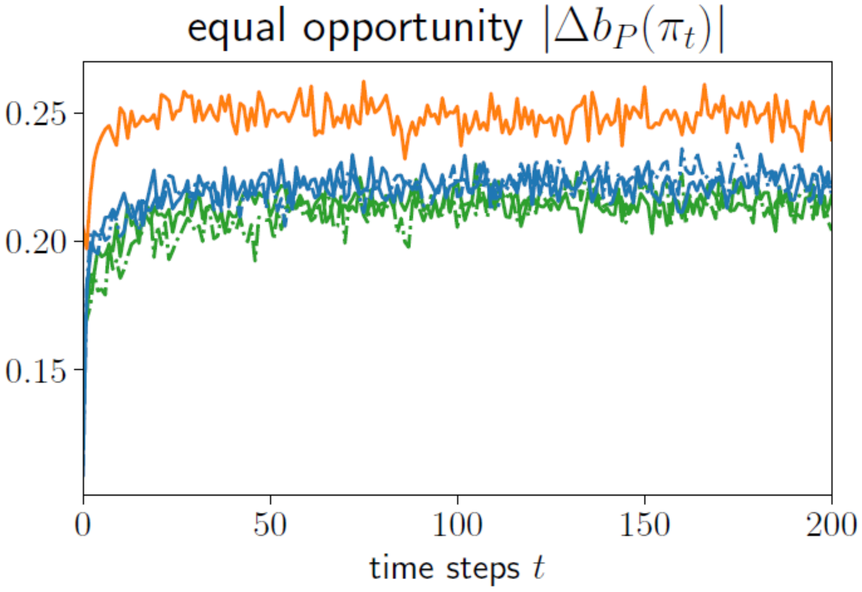
Real data

- logistic
- semi-logistic
- deterministic

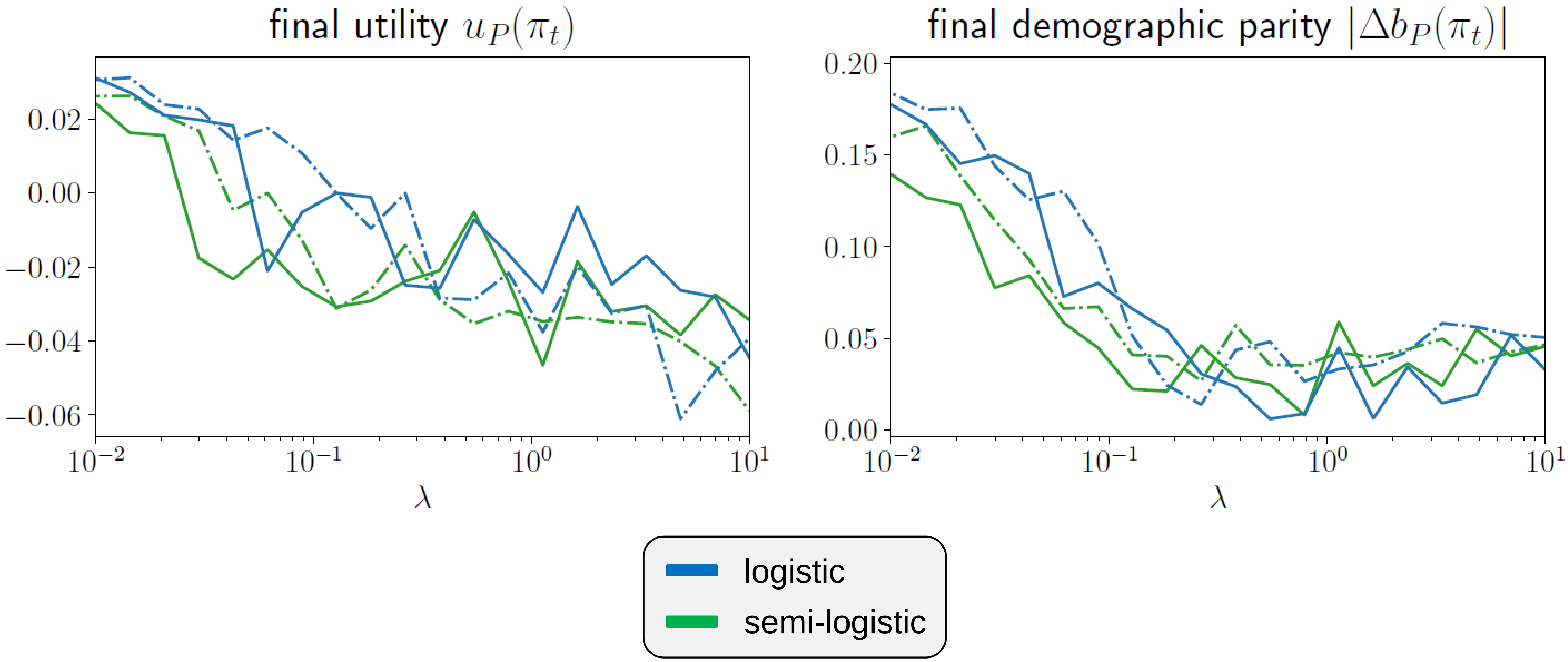
$\lambda = 0$





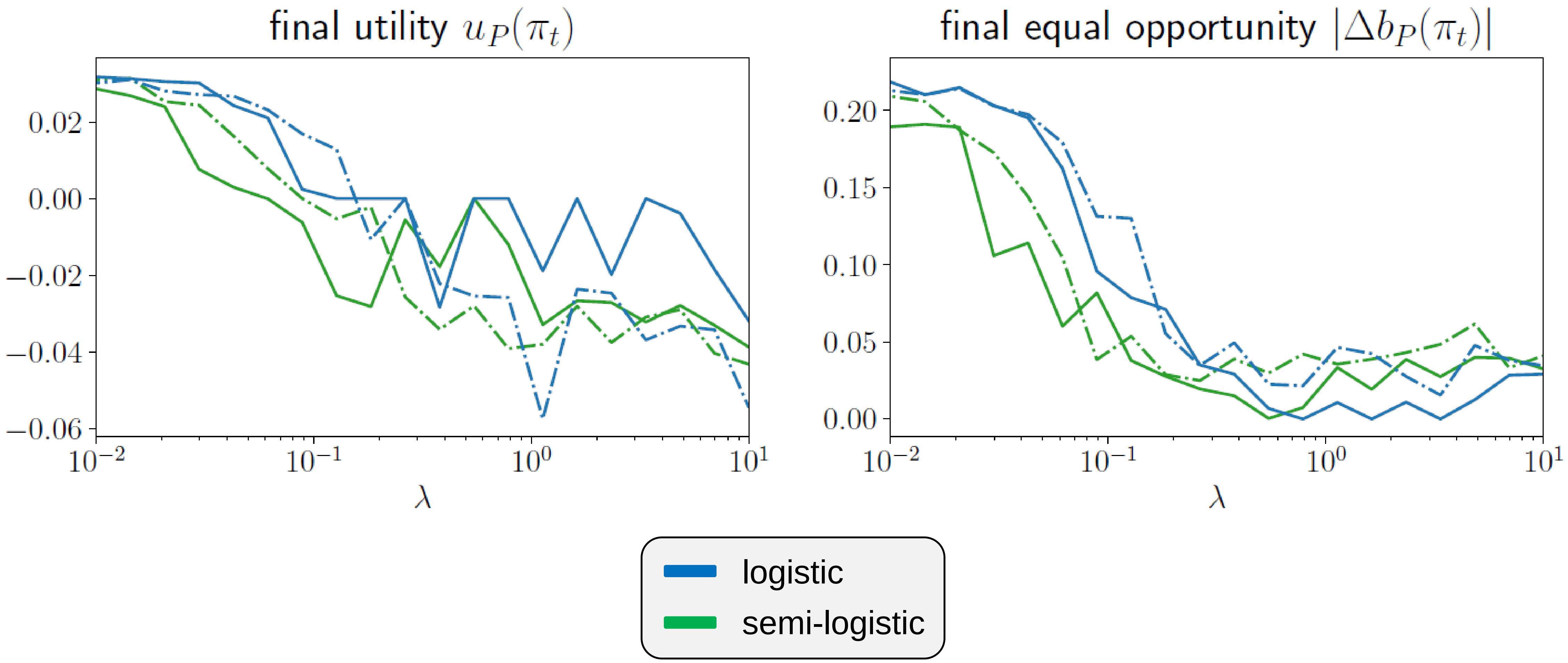


Real data



final utility $u_P(\pi_t)$
0.02
0.00
−0.02
−0.04
−0.06
10^{-2}
10^{-1}
10^{0}
10^{1}
λ
final demographic parity $|\Delta b_P(\pi_t)|$
0.20
0.15
0.10
0.05
0.00
10^{-2}
10^{-1}
10^{0}
10^{1}
λ
logistic
semi-logistic

Real data



final utility $u_P(\pi_t)$
0.02
0.00
−0.02
−0.04
−0.06
10^{-2}
10^{-1}
10^{0}
10^{1}
λ
final equal opportunity $|\Delta b_P(\pi_t)|$
0.20
0.15
0.10
0.05
0.00
10^{-2}
10^{-1}
10^{0}
10^{1}
λ
logistic
semi-logistic

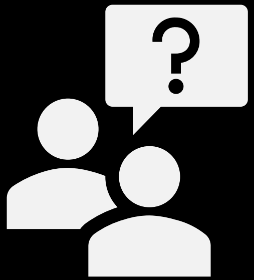
VI. Discussion

My take on the paper?

- All the proofs and more results available in Appendixes
- Code available
- Great idea to treat the problem of fairness taking into account the problem of imperfect dataset
- Promising practical results

- Main story sometimes hard to follow
- The results are mainly theoretical
- Needs to be tested on more datasets
- The solution is not viable in practice

Thank you for your attention !



Do you have
any question?

Annexes

Some observations

Under finite approximation $Q(y|x,s)$ of the ground truth (finite dataset), we can take:

$$\pi_Q(d=1|x,s) = 1[Q(y=1|x,s) \geq c]$$

with $Q(y=1|x,s) \approx P(y=1|x,s) - \delta_s$ and $\delta_s = c_s - c$

→ Incorporate the fairness criteria in the distribution

→ **Unfortunately this is suboptimal both in utility and fairness**

Propositions about deterministic policies

Maximize $v_{P_{\pi_0}}(\pi_Q)$ under the induced distribution P_{π_0}?

Proposition 1:

If there exists a subset $\mathcal{V} \in X \times S$ of positive measure under P such that $P(y = 1|\mathcal{V}) \geq c$ and $P_{\pi_0}(y = 1|\mathcal{V}) < c$, then there exists a maximum $Q_0^* \in \mathcal{Q}$ of $v_{P_{\pi_0}}$ such that $v_P\left(\pi_{Q_0^*}\right) < v_P\left(\pi_{Q^*}\right)$.

Propositions about deterministic policies

Non-exploring update rule:

No individual who has received a negative decision under the old policy would receive a positive decision under the new policy.

→ Error based learning algorithms are non-exploring

Proposition 2:

A deterministic threshold policy $\pi \neq \pi^*$ with $\Pr[\pi(x, s) \neq y] = 0$

will never converge to π^* under error based learning algorithm.

Formulas for an exploring policy:

$$u_{P_{\pi_0}}(\pi, \pi_0) \;\coloneqq\; \mathbb{E}_{x,s,y\sim P_{\pi_0},d\sim\pi(x,s)}\left[\frac{d(y-c)}{\pi_0(d=1|x,s)}\right]$$

$$b^{s}_{P_{\pi_0}}(\pi, \pi_0) \;\coloneqq\; \mathbb{E}_{x,s,y\sim P_{\pi_0},d\sim\pi(x,s)}\left[\frac{f(d,y)}{\pi_0(d=1|x,s)}\right]$$

$$v(\pi^*) = \sup_{\pi\in\Pi\setminus\{\pi*\}}\left\{u_{P_{\pi_0}}(\pi,\pi_0) - \frac{\lambda}{2}\left(b^{0}_{P_{\pi_0}}(\pi,\pi_0) - b^{1}_{P_{\pi_0}}(\pi,\pi_0)\right)^2\right\}$$

$$\nabla_\theta u_P(\pi_\theta) = \mathbb{E}_{x,s,y\sim P_{\pi_0},d\sim\pi_\theta(x,s)}\left[\frac{d(y-c)\nabla_\theta\, log\,\pi_\theta}{\pi_0(d=1|x,s)}\right]$$

$$\nabla_\theta b^{s}_{P}(\pi_\theta) = \mathbb{E}_{x,s,y\sim P_{\pi_0},d\sim\pi_\theta(x,s)}\left[\frac{\mathrm{f(d,y)}\nabla_\theta\, log\,\pi_\theta}{\pi_0(d=1|x,s)}\right]$$