Achieving Socio-Economic Parity through the Lens of EU AI Act
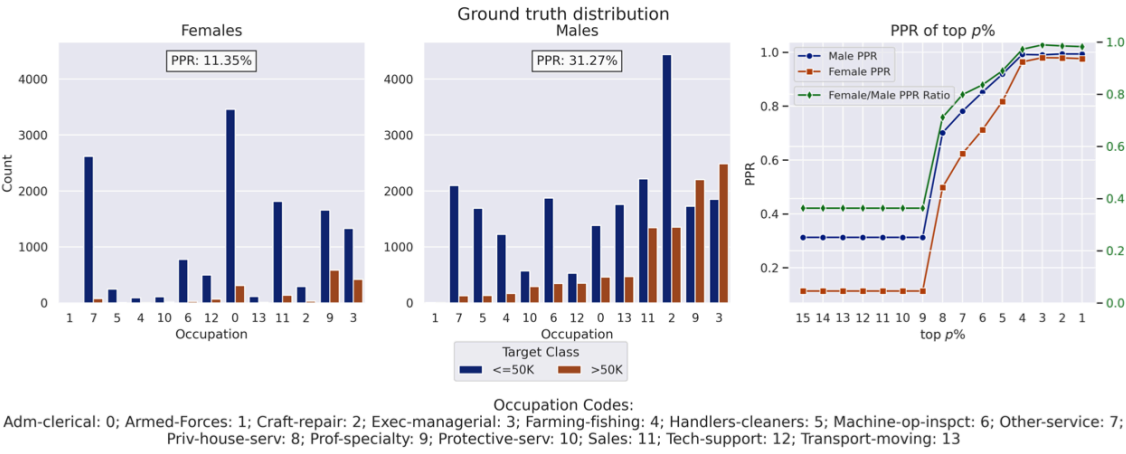
Arjun Roy[1], Stavroula Rizou[2], Symeon Papadopoulos[2], Eirini Ntoutsi[1]

[1] RI CODE, Universität der Bundeswehr München, Germany
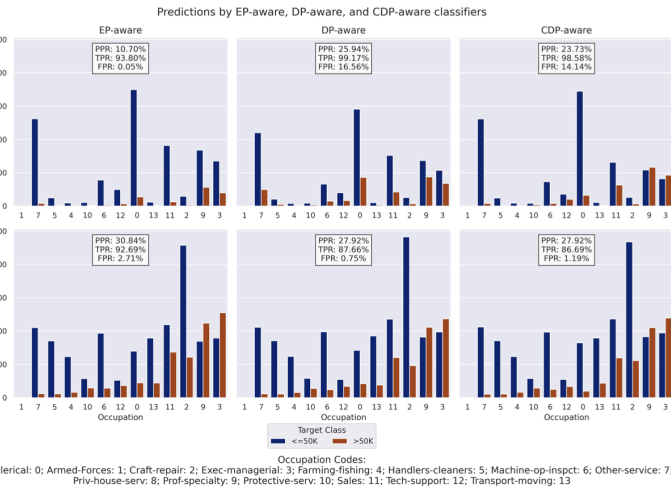[2] Information Technologies Institute, CERTH, Greece

Main Paper

Disparities are Deeply Rooted in Socio- Economic Status (SES)

PPR scores refer to the % of positive labels in the ground truth.
PPR of top p% represents the population with top p% capital gain

Occupation Codes:
Adm-clerical: 0; Armed-Forces: 1; Craft-repair: 2; Exec-managerial: 3; Farming-fishing: 4; Handlers-cleaners: 5; Machine-op-inspct: 6; Other-service: 7; Priv-house-serv: 8; Prof-specialty: 9; Protective-serv: 10; Sales: 11; Tech-support: 12; Transport-moving: 13

- Despite advancements in AI fairness, most **fail to capture the compounded effects of socio-economic privilege**.

- **Gap** in positive label (high income) distribution between "Males" and "Females" **narrows as Socio-economic-status (SES) privileges increases**.

- In Adult data, **Positive class ratio for females surpasses 80% in the top 5% of capital gain** (a wealth proxy), but drops below 40% for the bottom 92%.

- Intersectional vulnerabilities amplify systemic inequities, conflicting with equalization principles that aim to **account for disadvantages beyond an individual's control**.

- While positive action (e.g., distribution of false positives i.e., FPR) helps close demographic gaps, additional **preference for underprivileged groups** is essential for equity.

- Furthermore, **accounting for individual-level factors** (e.g. working hours) in the distribution of positive actions is needed to enhance fairness with regard to equalization principles.

Existing Gold standard fairness notion fails to capture SES-privileges

Occupation Codes:
Adm-clerical: 0; Armed-Forces: 1; Craft-repair: 2; Exec-managerial: 3; Farming-fishing: 4; Handlers-cleaners: 5; Machine-op-inspct: 6; Other-service: 7; Priv-house-serv: 8; Prof-specialty: 9; Protective-serv: 10; Sales: 11; Tech-support: 12; Transport-moving: 13

Comparing the outcomes on **Adult dataset** given by classifiers constrained by popular fairness notions, we observe:
- **Equal Opportunity** (EP-aware) classifier **struggles with class imbalances** by emulating ground truth, widening male-female disparities

- **Demographic Parity** (DP-aware) classifier narrows the disparities.
  - PPR Female ≈ PPR Male
  - blind distribution of positive decisions raises concerns

- **CDP-aware** classifiers (**Demographic Parity conditioned on non-protected attribute**) are often seen as the baseline for equitable allocation across subdivisions (e.g., occupation groups). They:
  - Promote a **wider and fairer distribution of positive actions**
 - Are widely **adopted for policy-sensitive applications**

- CDP fails the equalization principle: **Privileged groups benefits**
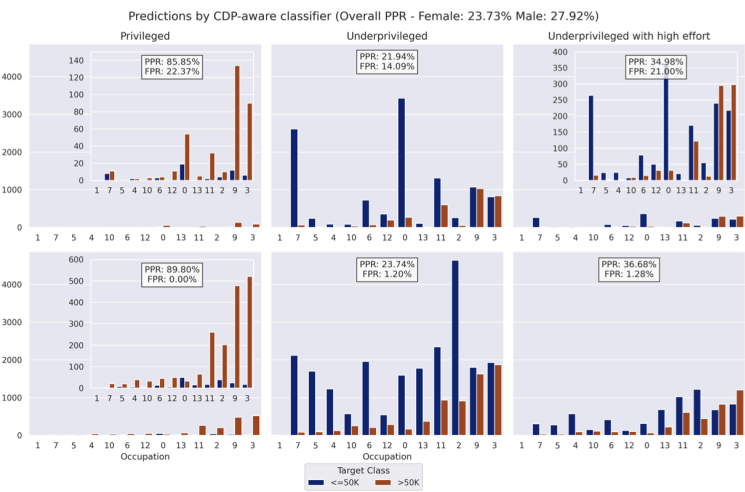
- FPR for females privileged **>** under-privileged
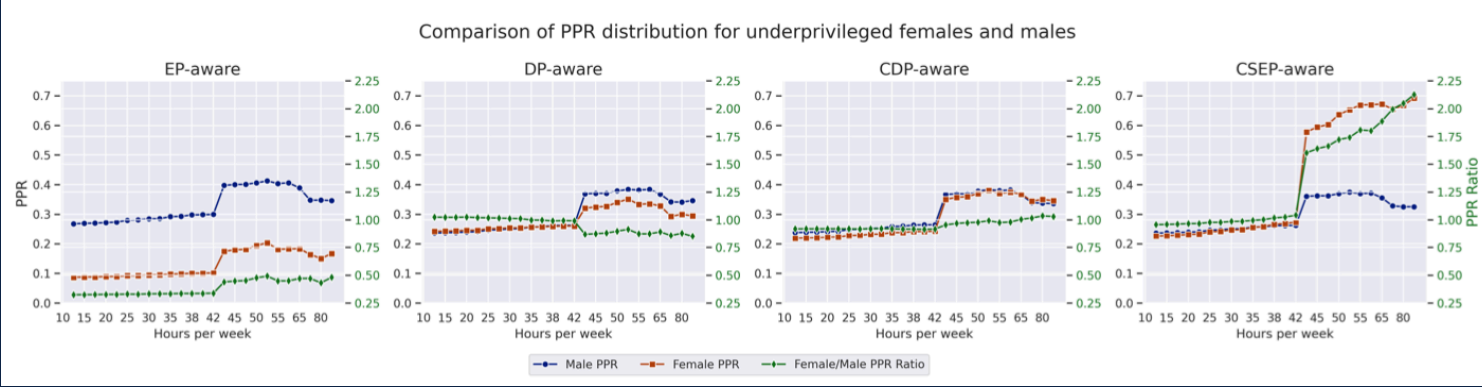- PPR for females privileged **>>** under-privileged

Socio-Economic Parity (SEP) and Conditional Socio-Economic Parity (CSEP)

- To address SES-driven disparities while accounting for individual-level factors we introduce the notions **Socio-Economic Parity (SEP)** and the stricter **Conditional SEP.** The basis on which the notions are established are:
  - Chances of favourable outcomes should not be hindered for socio-economically underprivileged individuals
  - Individuals in underprivileged subgroups who put higher effort beyond a threshold should be given higher chance of positive rewards proportional to their effort
  - Positive action for privileged subgroups should be minimized in order to reduce the influence of privilege
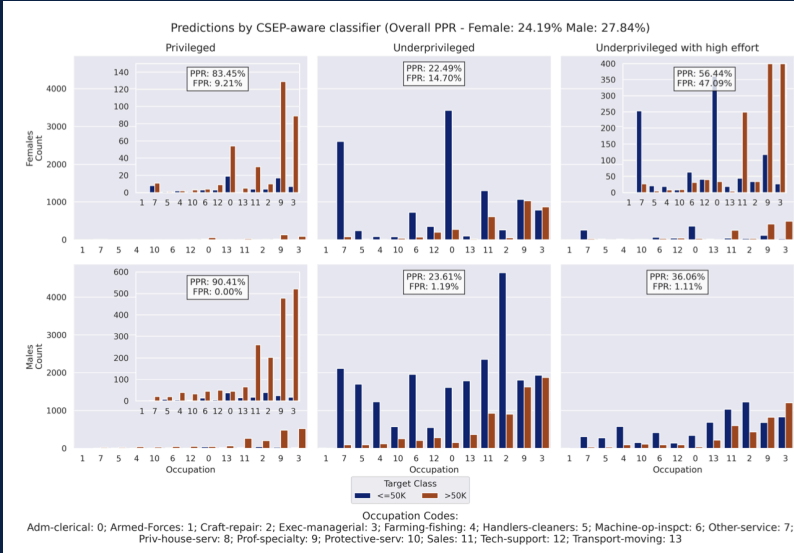
Empirical Study shows:
- **CSEP reallocates positive outcomes for underprivileged females** based on their effort, without inflating error rates for other groups
  - **PPR for underprivileged females increases** from <0.3 to >0.6 with greater effort; for those with the highest effort, PPR approaches 0.7

  - **Female-to-male PPR ratio nearly doubles** (from 0.98 to 2.1) for underprivileged **high-effort females**, mitigating both overall demographic disparities and those within vulnerable subgroups

Framework for fairness-aware notions: addressing SES bias in light of the AI Act

The proposed approach supports key AI Act provisions by:
- **Conformity assessment** (Art. 43): Identification and mitigation of unfair discrimination (Art. 10(2)(f),(g)), guiding AI system design to prevent or mitigate risks (Art. 9(5)), and enhancing accountability.

- **Fundamental Rights Impact Assessment** (FRIA, Art. 27): Enabling context-based evaluation and measurement of discrimination risks affecting protected groups (Art. 27(1)(d)), supporting risk mitigation, and informing comprehensive management responses.

Detection of a high-risk AI system
Identification of risk level of the AI system (article 10 par. 2 (f), (g))

1. Identification and mitigation of unfair discrimination   (article 10 par. 2 (f), (g))
2. Representative and ad hoc datasets (article 10 par. 3 and 4)
3. Technical documentation maintenance (article 11)
4. Record-keeping of logs (article 12)
5. Human oversight (article 14 par. 4 (b))
6. Mitigation measures against (re)production of bias (article 15 par. 4)
7. Fundamental Rights Impact Assessment & Conformity Assessment (article 27, article 43)

Conclusions and Future Directions

- Our framework provides a strong foundation for addressing financial exclusion by focusing on (un)privileged groups.

- The critical balance achieved by SEP is supported by measurable evidence and legal explanations for all decisions, ensuring fairness and transparency.

- The concept of indirect discrimination helps identify and address inequalities between privileged and underprivileged groups within SES.

- Our empirical results show that SEP rewards high-effort underprivileged subgroups while retaining comparable error rates across the population — demonstrating its potential to narrow socioeconomic gaps without sacrificing performance.

- Enhanced SES protection under the EU AI Act is needed, reflecting its growing role in algorithmic fairness and policy compliance.

- Leveraging statistical data from sociological and economic studies can further guide SES-aware decisions in AI, strengthening fairness and equity in practice.

This research work was funded by the European Union under the Horizon Europe MAMMOth project, Grant Agreement ID: 101070285. UK participant in Horizon Europe Project MAMMOth is supported by UKRI grant number 10041914 (Trilateral Research LTD). The work of Arjun Roy is also supported by the EU Horizon Europe project STELAR, Grant Agreement ID: 101070122.