Western

UNIVERSITY · CANADA

Fairness in Entity Matching and Blocking

Presenter: Mohammad Hossein Moslemi

Supervisor: Mostafa Milani

List of Contents

- **Context and Motivation**
- **Part 1: Blocking**
  - Quality of Blocking
  - Bias Measurement in Blocking
  - Experiments
- **Part 2: Matching**
  - Problem Definition
  - Calibration Algorithm
  - Experiments
- **Conclusion and Future Work**

Context and Motivation

Entity Matching

- Entity Matching (EM)
  - Identifies records referring to the same real-world entity across datasets.
- Step 1: Blocking
  - Groups similar records to eliminate unlikely pairs
  - Minimizes **computational overhead**
  - Often based on simple attributes (e.g., ZIP code, first letter of name)
- Step 2: Matching
  - Generates final matched pairs through ML or rule-based methods
- Why two steps?
  - Blocking reduces $O(n^2)$ comparisons
  - Matching ensures **precision** and **accuracy**

Fairness in Machine Learning

- ML models can amplify social biases present in data.
- **Sensitive attributes** (e.g., race, gender) should not unfairly influence outcomes.
- **Bias types**:
  - **Direct**: Sensitive attribute is explicitly used in prediction.
  - **Indirect**: A correlated feature causes unfair impact.
- **Real-world risk:** Unfair decisions in loans, healthcare, etc.
- **Goal:** Build **equitable** models with high **accuracy**.

Fairness in Entity Matching

- **Example of bias in EM**

Google's algorithm shows prestigious job ads to men, but not to women. Here's why that should worry you.

By Julia Carpenter

The Washington Post

Airline "no-fly" lists trample the rights of people of color. Seattle should not allow hotels to create "no stay" lists

Amy Roe, Former ACLU-WA Senior Writer

Published: Friday, July 19, 2019

ACLU Washington

Knowledge Gap

- **Blocking and Fairness**
  - Very limited research on fairness-aware blocking strategies.

- **Matching and Fairness**
  - Fewer studies compared to general ML fairness.
  - **Bias in similarity scores** is often overlooked.
  - Lacks methods to **reduce score-level bias** effectively.

Our Contribution

- **Part 1: Fairness in Blocking**
  - Defined a fairness metric specific to blocking.
  - Evaluated bias across multiple blocking methods.
  - Showed how blocking bias **propagates** to matching stage.

- **Part 2: Fairness in Matching**
  - Introduced a **score-based** fairness metric for matchers.
  - Developed a **post-processing calibration algorithm** to reduce bias.
  - Tailored solutions for different fairness definitions.

Part 1

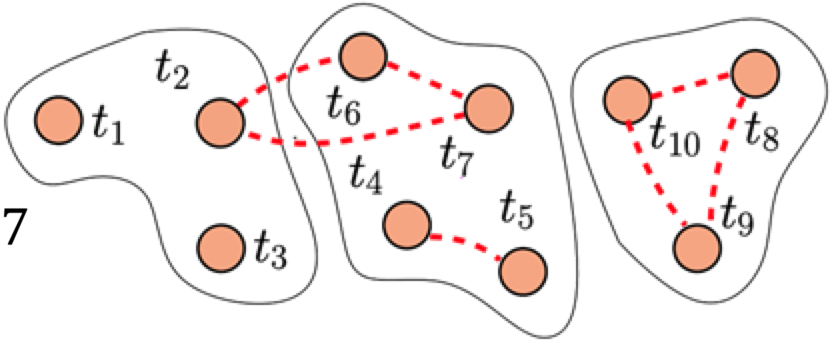
Blocking

Quality of Blocking

- **Blocking:** Groups similar records to avoid full pairwise comparisons.
- **Goal:** Maximize true matches, minimize unnecessary comparisons.

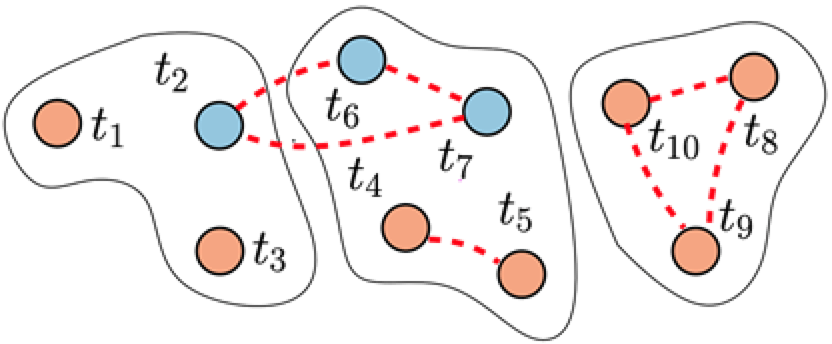
- **Metrics** (**P:** All pairs, **M:** True matches, **C:** Candidate set):
  - Reduction ratio (RR): $1 - \frac{|C|}{|P|}$
  - Pair completeness (PC): $\frac{|C \cap M|}{|M|}$

- **Example:**
  - **P:** $\frac{10 \times 9}{2} = 45$, **C:** $3 + 6 + 3 = 12$, **M** = 7
  - **RR** ≈ 0.73, **PC** ≈ 0.71



Measuring Bias in Blocking

- **Problem**: Standard metrics (RR, PC) don't capture **blocking bias.**
- **Minority Pair:** A pair is minority if at least one record is minority.
- **Fairness Metrics:**
  - $\Delta\mathrm{RR} = \mathrm{RR_b} - \mathrm{RR_a}$
  - $\Delta\mathrm{PC} = \mathrm{PC_b} - \mathrm{PC_a}$



- **Example**:
  - Before: 21 Majority, 24 Minority pairs
  - After: 5 Majority, 7 Minority pairs
  - $\mathrm{RR_a} = 1 - \frac{7}{24} \approx 0.71,\ \mathrm{RR_b} \approx 0.76 \rightarrow \boldsymbol{\Delta\mathrm{RR} \approx 0.05}$
  - $\mathrm{PC_a} \approx 0.33,\ \mathrm{PC_b} = 1 \rightarrow \boldsymbol{\Delta\mathrm{PC} \approx 0.67}$

Experiments 1

Blocking

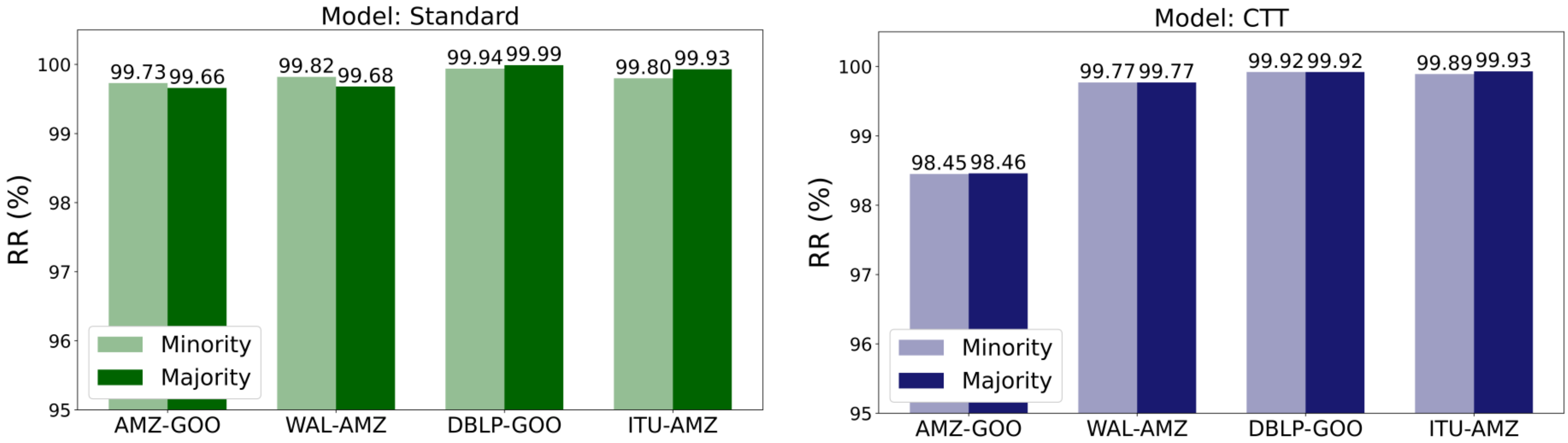
Experimental Setting

- **Datasets:**
  - 7 well established benchmarks in the literature
  - Each dataset includes a **sensitive attribute**.

- **Blocking methods:**
  - Evaluated 8 widely-used blocking techniques.

Bias Analysis Experiment

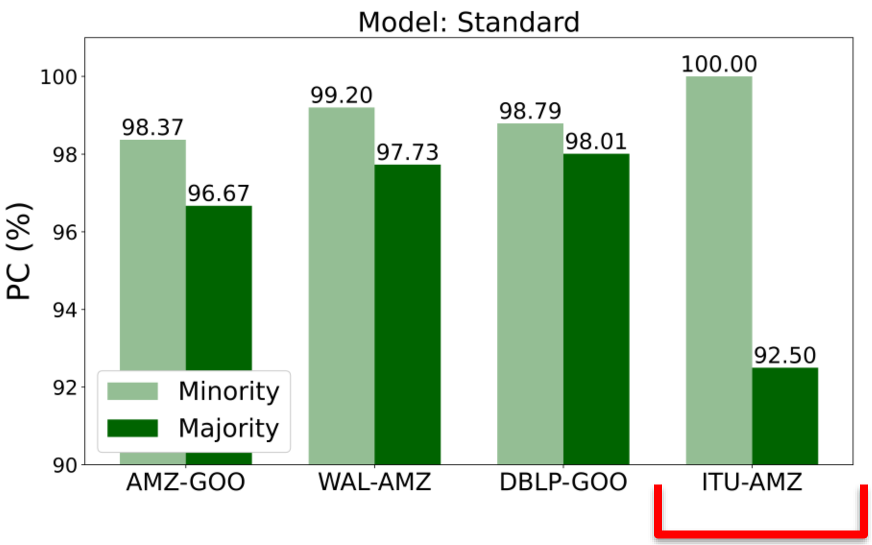
- Comparison of **reduction ratio** between Minority and Majority Groups across Models and Datasets.

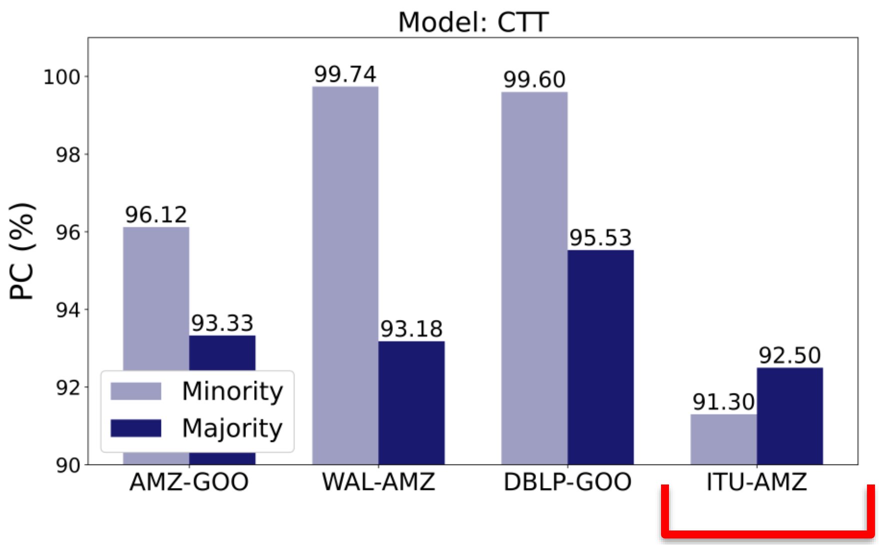


Western

Bias Analysis Experiment

- Comparison of **pair completeness** between Minority and Majority Groups across Models and Datasets.





Bias Propagation Experiment

Model	AMZ-GOO
StdBlck	1.70 (98.37, 96.67)
QGram	-1.01 (95.66, 96.67)
XQGram	6.16 (94.49, 88.33)
Suffix	16.01 (89.34, 73.33)
XSuffix	$\boxed{18.15}$ (84.82, 66.67)
AUTO	8.98 (88.98, 80.00)
CTT	2.79 (96.12, 93.33)
GRAPH	**0.62** (93.95 , 93.33)

PC Bias on AMZ-GOO Dataset

	Positive Rate Bias
QGram	4.42×10^{-3}
XSuffix	8.11×10^{-3}

Propagated bias on a perfect matcher

Takeaways

- Blocking reduces complexity, but can **introduce bias**
- Biases in blocking can **propagate** to downstream matchers
- **Blocking methods** vary in bias even on the **same dataset**
- A single method shows varying bias **across datasets**

→ Choose blocking methods based on both **quality** and **fairness.**

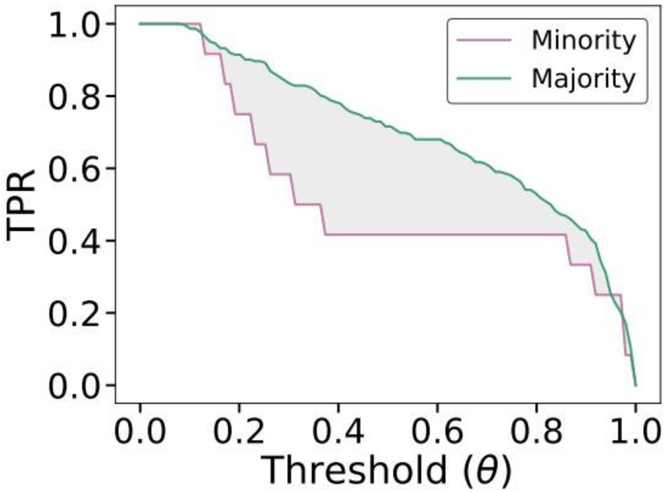
- Moslemi, Mohmmad Hossein, Harini Balamurugan, and Mostafa Milani. "Evaluating Blocking Biases in Entity Matching." 2024 IEEE Big Data, 2024.

Part 2:

Matching

Binary vs. Score-Based Matching

- Prior work treated EM as a binary task:
  - Reducing bias at a fixed threshold
- Fair at one threshold, highly biased at another



- Threshold Adjusting is crucial:
  - **No-fly lists:** Lower threshold → More detection, more false positives
  - **Finance:** Higher threshold → Avoids wrongful merges, protects privacy & security

Traditional Fairness Measures

- There are many different fairness measures. Three major ones:

- **Demographic Parity (DP)**
  - Independence of prediction from groups
  - $\hat{Y} \perp\!\!\!\perp A$

- **Equal Opportunity (EO)**
  - Independence of prediction from groups in true matches
  - $\hat{Y} \perp\!\!\!\perp A \mid Y = 1$

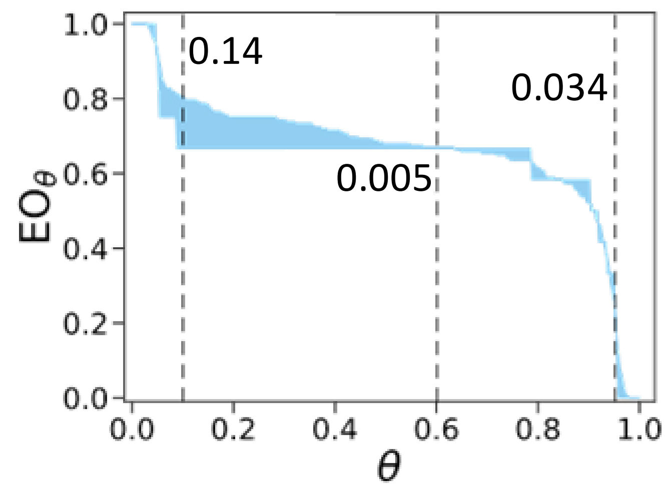
- **Equalized Odds (EOD)**
  - Independence of prediction from groups in true matches non-matches
  - $(\hat{Y} \perp\!\!\!\perp A \mid Y = 1)$ and $(\hat{Y} \perp\!\!\!\perp A \mid Y = 0)$

Score Function Fairness Measures

- Traditional fairness measures are threshold-dependent and can be misleading.

- Score bias for Φ (PR, TPR, …)
  - Averaging bias of Φ across all thresholds.

$$bias(s, \Phi) = \int_0^1 |\Phi_b(s, \theta) - \Phi_a(s, \theta)| \, d\theta$$



Problem of Fair Entity Matching

- Consider score function s and performance metric Φ

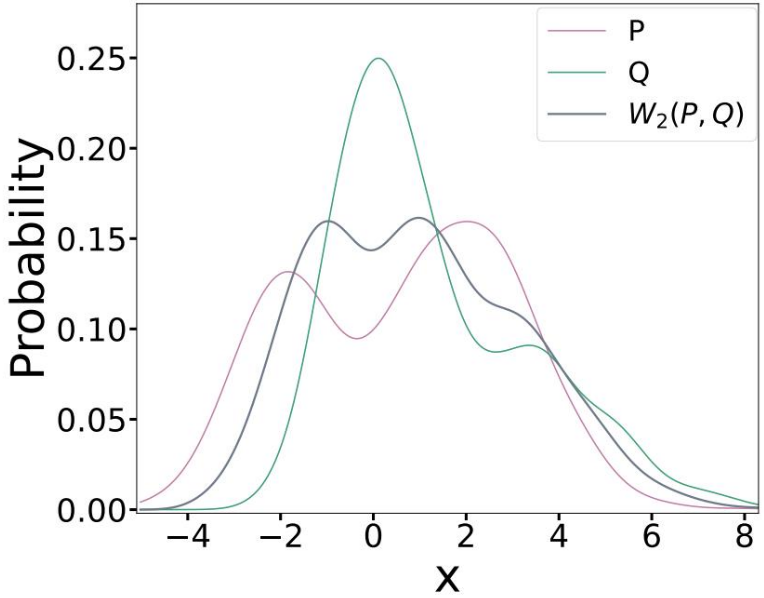
- **FairScore:** Find new score function s* with bias(s*, Φ) < δ and minimal deviation from s.

$$s^* = \arg\min_{s' \in \mathcal{S}_{fair}} risk(s', s)$$

$$risk(s', s) = \mathbb{E}\Big[|s'(X) - s(X)|\Big], \qquad \mathcal{S}_{fair} = \Big\{s \mid bias(s, \Phi) \leq \delta\Big\},$$

Solution: Score Calibration

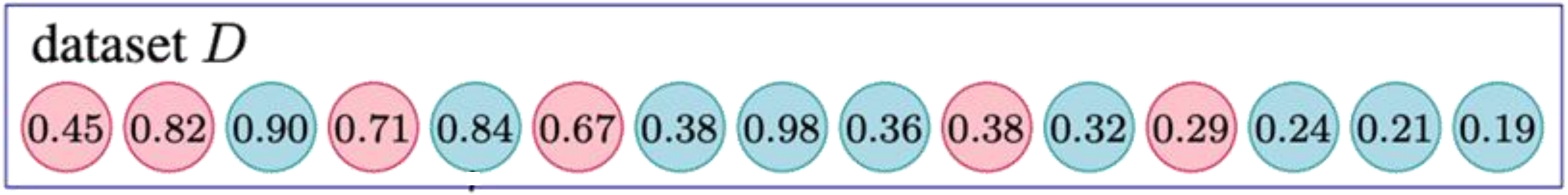
- Minority and majority scores have different distributions.
- We aim to align them with minimal change to scores:
  - **Wasserstein barycenter:** A Central probability distribution.



Wasserstein barycenter of P and Q

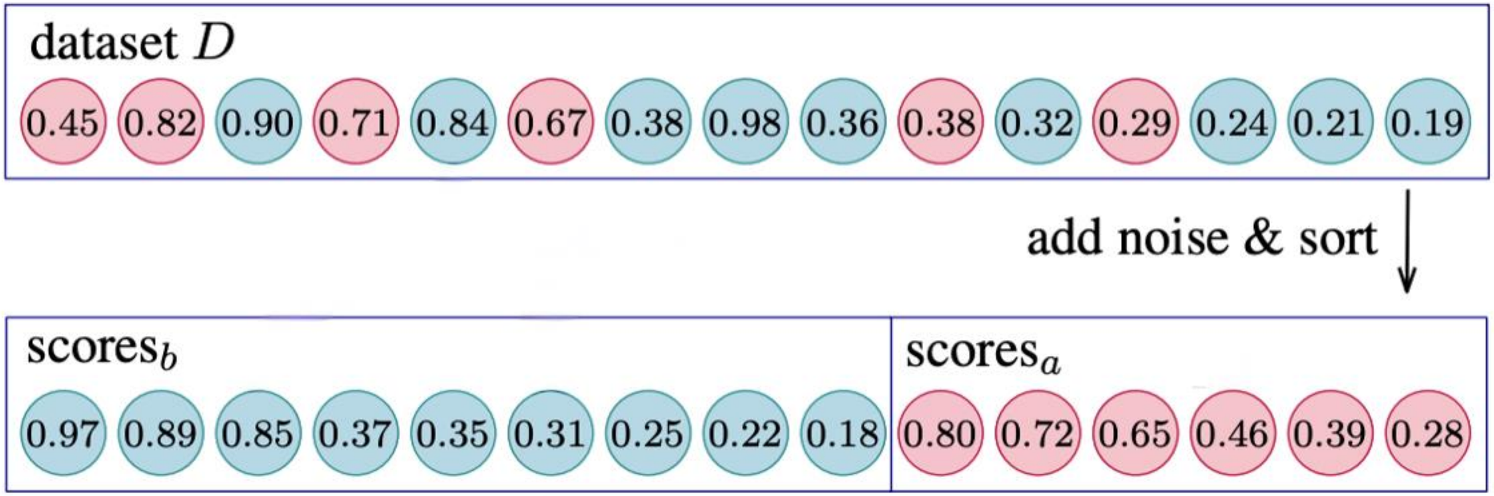
Score Calibration

- What is query point and Dataset D
- Example
  - 6 Minority in Red
  - 9 Majority in Blue
  - Query point score is 0.34 and majority



Score Calibration

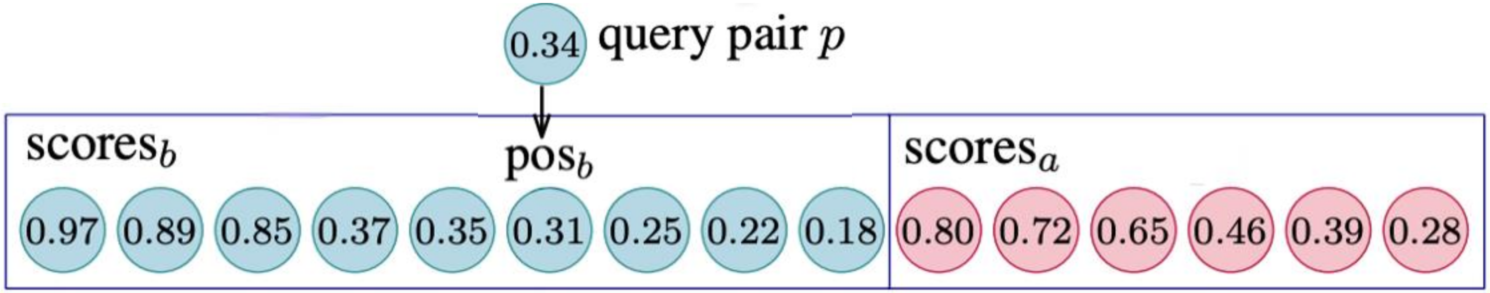
1. Sort minority and majority scores in dataset D. Add noise for continuity. Majority size is n_b, Minority size is n_a.



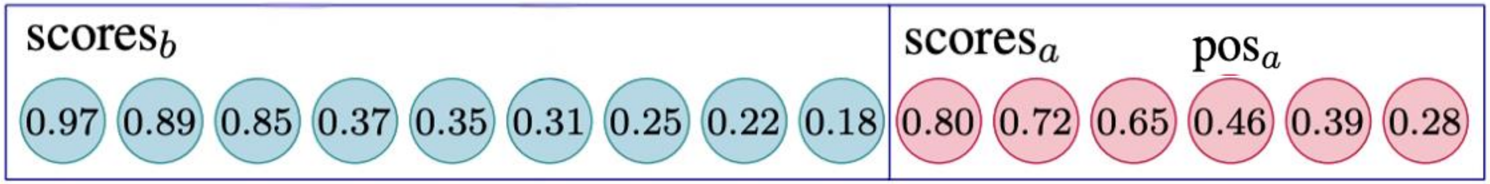
Western

Score Calibration

2. Locate the query's rank in its group (6th from top out of 9)

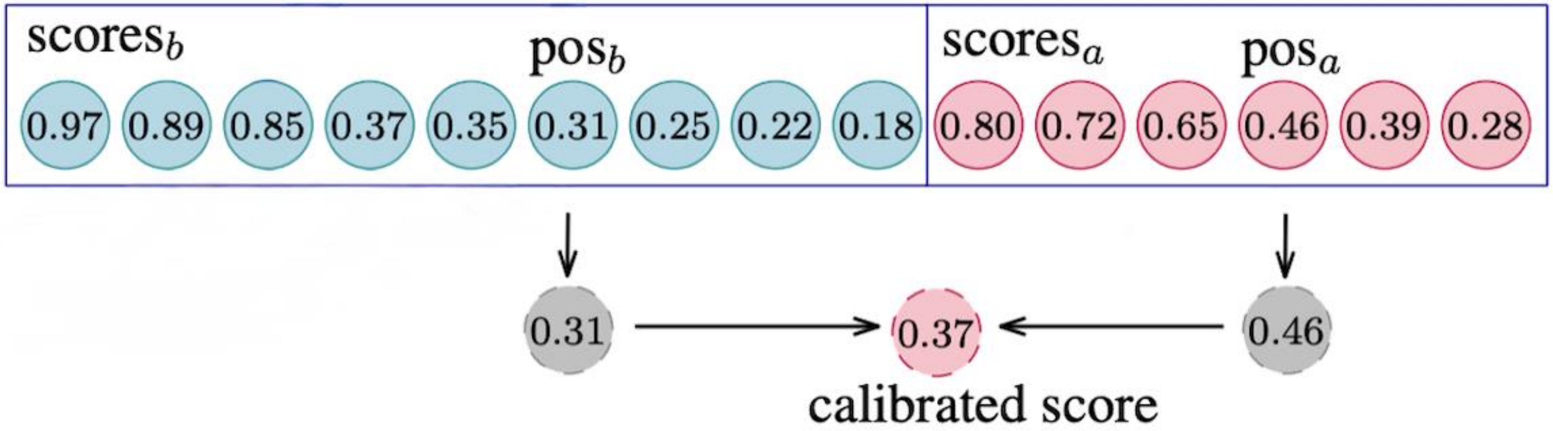


3. Transfer that rank to the other group
   – 6th out of 9 → 4th out of 6



Score Calibration

4. Take the values at the matched positions in both groups
5. Calibrated score is a **weighted average** based on **group sizes**.



Theoretical Insights

- It was an approximation not an exact computation
- **Given** initial score s, optimal score s*, calibrated score ŝ, and dataset size n, the bounds are:

$$bias(\hat{s}, PR) = O(n^{-1}),$$
$$risk(s^*, \hat{s}) = O(\log(n)^{-1/2}).$$

- **Intuition:** As the size of D increases:
  - Bias reduces at a rate of $\frac{1}{n}$
  - Calibrated score gets closer to s* at $\frac{1}{\sqrt{\log(n)}}$

- Detailed proof in the thesis

Conditional Score Calibration

- Aligning score distributions removes DP bias:
  - Same positive rate at all thresholds

- Doesn’t remove EO or EOD, as they rely on true labels.
  - **Solution:** Calibrate using pairs with the same label as the query point.

- Labels for query point or dataset D may be unknown.
  - **Solution:** Estimate labels using a threshold that best splits scores in D.

Experiments 2

Matching

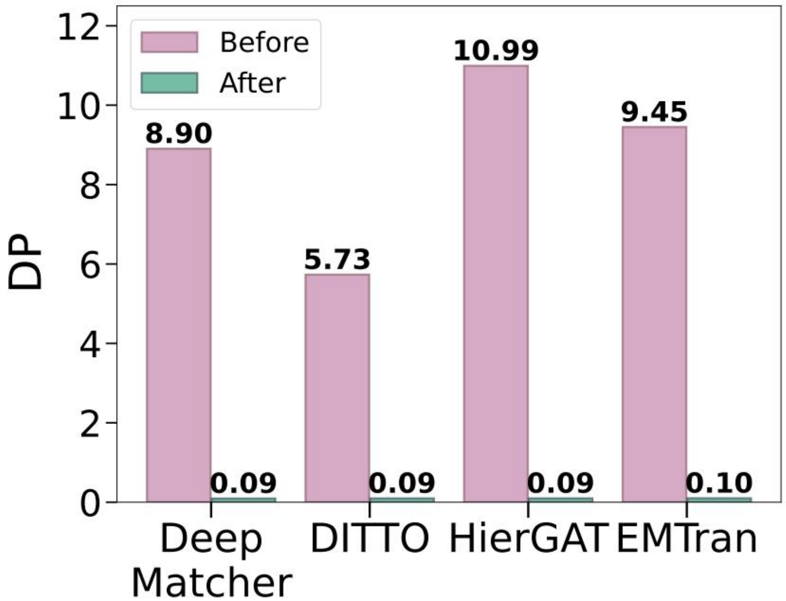
Western

Experimental Setting

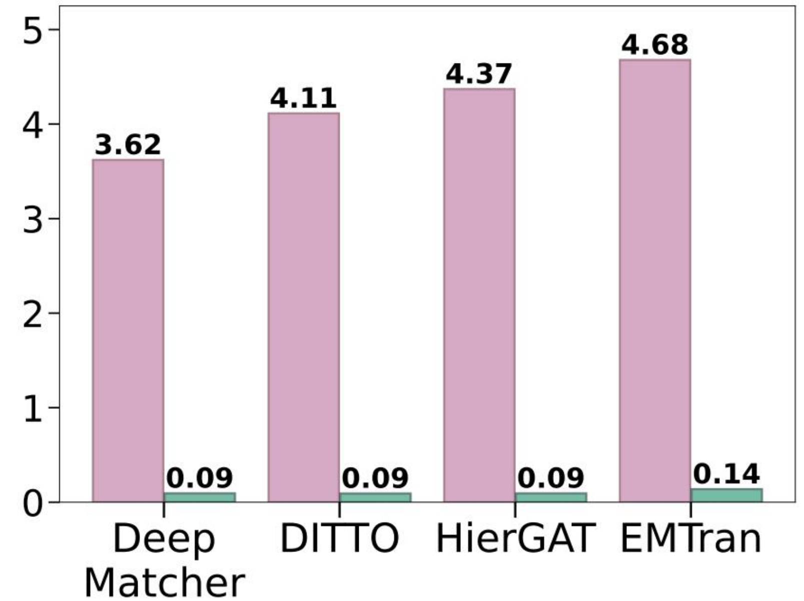
- **Datasets:**
  - Same as the blocking part

- **Matching methods**
  - 5 state-of-the-art methods

Calibration Performance on DP



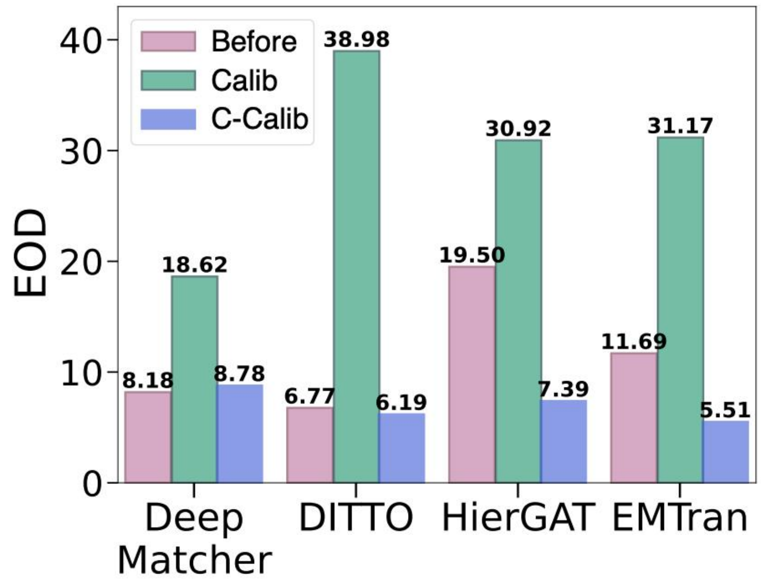
AMZ-GOO



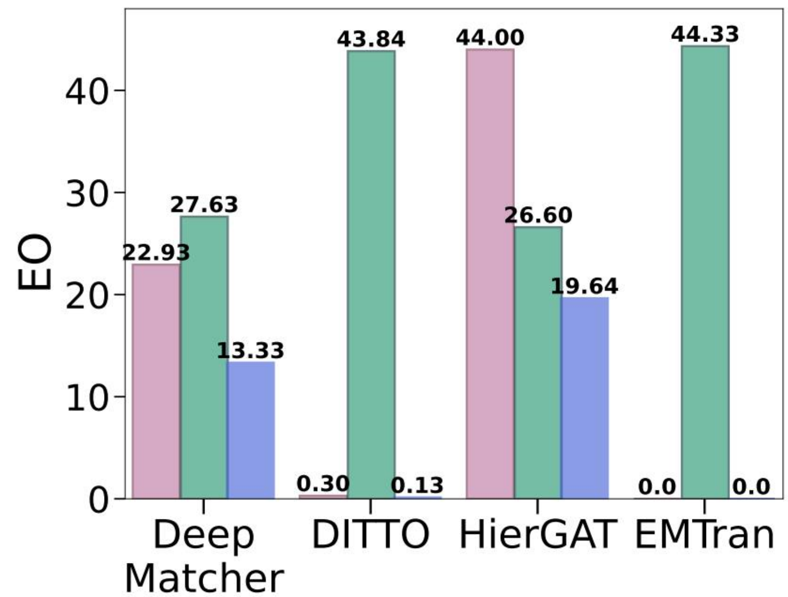
DBLP-ACM

- Risk change is minimal; details come after conditional calibration bias results.

Conditional Calibration Performance

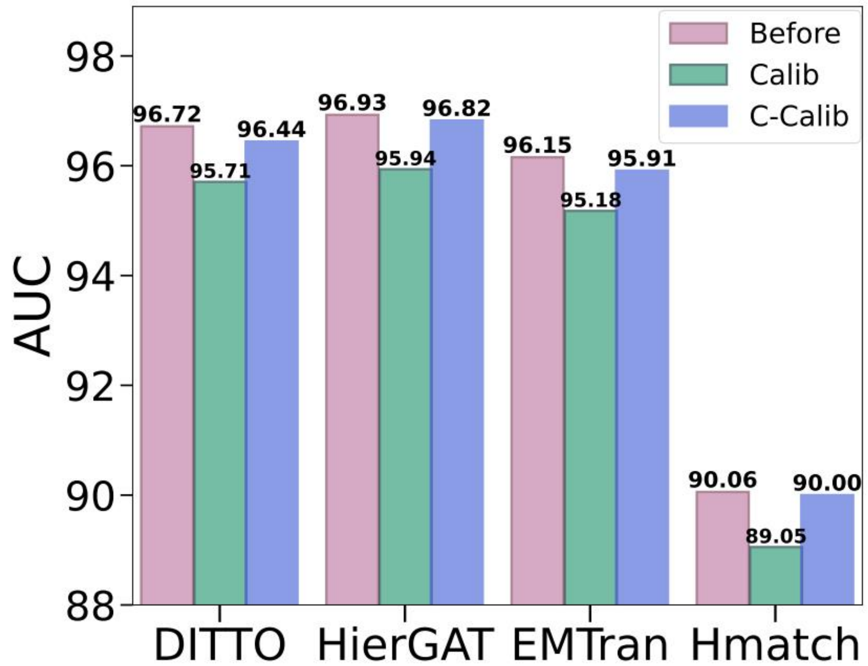


AMZ-GOO



BEER

Effect on Risk



AMZ-GOO

Takeaways

- Calibration works well for DP, but not for label-based metrics.

- We propose **Conditional Calibration**, which handles this much better.

- Risk (AUC) impact is minimal, especially with conditional calibration, as it accounts for labels, causing fewer changes.

- Moslemi, Mohammad Hossein, and Mostafa Milani. "Threshold-independent fair matching through score calibration." GUIDE-AI at SIGMOD. 2024.

Conclusion and Future Work

Western

Conclusion

- Studied fairness in both **blocking** and **matching** steps of EM
- Designed new **fairness metrics** for blocking and analyzed bias in blocking methods
- Proposed post-processing methods to fix **score bias**
- Improved fairness with little impact on accuracy

Future Work

- **Fairness in Blocking:** Design methods to reduce bias in blocking
- **Beyond Post-Processing:** Try pre- and in-processing bias reduction
- **Theory:** Build stronger theoretical foundations for conditional calibration

References

- S. Nima, et al. "Through the Fairness Lens: Experimental Analysis and Evaluation of Entity Matching. In *VLDB* 2023
- E. Chzhen, et al. Fair regression with wasserstein barycenters. In NeurIPS2020
- C. Dwork, et al. Fairness through awareness. Innovations in theoretical CS conf., 2012.
- M. Ebraheem, et al. Distributed representations of tuples for ER. VLDB, 2018.
- M. Cuturi and et al. Fast computation of wasserstein barycenters. In ICML, PMLR, 2014.
- S. Nilforoushan, et al. EM with auc-based fairness. In IEEE Big Data, 2022.

Thank you!

Western

UNIVERSITY · CANADA