

Fairness under Competition

Why Competition Matters

In markets with multiple classifiers, ecosystem fairness is at least as important as individual fairness. We show that **individually fair classifiers** can produce **unfair ecosystems**.

We quantify losses via **correlations** and **overlap**. We also show that adjusting individual classifiers to be more fair can **backfire**.

Model Snapshot

Borrowers (x, a, y) , where x are observable traits, a is observable group membership, y is an unobservable outcome. Deserving borrowers: $y=1$. Utility $v(x, a, y, r)$ over r offers. • 0-1 utilities: borrowers only care about getting at least one offer. • 0-1-k utilities: borrowers get higher utility if more than one offer.

Lenders decide on loan offers using classifiers $c_\ell(x, a)$. **False-negative rates** β_ℓ is probability lender ℓ does not make offer to deserving borrowers.

Defining EOC

Equal Opportunity (EO): the group gap in false-negative rates is zero.

Level of Equal Opportunity under Competition (**EOC-level**) measures the group gap in $\Pr[at\ least\ one\ offer \mid Y=1]$, where level **zero** means ecosystem fairness. We also study **v-EOC-level**: group gap in expected utilities for deserving borrowers.

These metrics highlight **access to opportunity**, not just per-classifier parity.

Force 1: Correlation Imbalance

Fix two **EO** classifiers, where σ_i is standard deviation of classifier i and ρ_a is correlation between classifiers on group a . Then **EOC-level** is $\sigma_1\sigma_2|\rho_0 - \rho_1|$.

Worst case: $\min\{\beta_1, \beta_2\} - \max\{0, \beta_1 + \beta_2 - 1\}$.

When one group’s predictions are **highly correlated** and the other’s are not, errors compound unevenly, driving **systemic inequity**.

Competition between fair classifiers can lead to unfair outcomes.

Applying a fairness adjustment to individual classifiers can exacerbate unfairness.



RONEN GRADWOHL
EILAM SHAPIRA
MOSHE TENNENHOLTZ



Level of v-EOC

Under **0-1-k** utilities, v-EOC-level is $\sigma_1\sigma_2|(k - 2)(\rho_0 - \rho_1)|$. Worst case scales by $|k - 2|$. Interestingly, for $k < 2$ vs $k > 2$, the disadvantaged group can **flip**, motivating context-awareness.

More Competitors, More Risk

With n EO classifiers, worst-case EOC-level is $\min_i \beta_i - \max\{0, \sum_j \beta_j - 1\}$. If classifiers are symmetric and correlation is maximal in one group and minimal in the other, then EOC is $\beta - \beta^n$. Thus, **more competitors** can **worsen inequity**.

Force 2: Overlap Imbalance

When lenders serve different subsets of borrowers, overlap imbalance drives EOC-level. For uncorrelated classifiers, EOC-level depends on group a overlaps γ_a .

When $\beta_1 = \beta_2 = \beta$, EOC-level of two **EO** classifiers is $\beta(1 - \beta)|\gamma_0 - \gamma_1|$.

Adjustments Can Backfire

Post-processing to EO can **backfire**. By altering **correlations**, EOC-level can **increase**. We construct cases where EOC-level pre-adjustment is 0, yet after EO adjustment they become **positive**, despite improved **per-classifier fairness**.

Experiments show EOC-level **often** worsens after adjustment. Under several experiments, percent of instances EOC-level worsened was:

	300	1k	3k	10k	30k	100k
Exp. 1	[75.0, 82.2]	[68.0, 76.4]	[55.6, 64.2]	[49.4, 58.2]	[42.0, 50.6]	[26.2, 34.0]
Exp. 2	[75.6, 82.8]	[65.2, 73.8]	[51.8, 60.8]	[35.4, 44.2]	[25.8, 33.8]	[12.6, 19.0]
Exp. 3	[74.2, 81.2]	[63.4, 71.2]	[50.8, 59.6]	[35.6, 43.8]	[27.6, 36.2]	[14.2, 20.6]

When EOC-level worsened, it often worsened by a large factor:

	300	1k	3k	10k	30k	100k
Exp. 1	[10.8, 30.8]	[7.8, 16.3]	[7.6, 121.0]	[6.0, 19.9]	[7.3, 23.0]	[7.5, 31.0]
Exp. 2	[21.1, 111.8]	[8.3, 30.1]	[2.9, 9.4]	[1.8, 2.2]	[1.6, 1.8]	[1.3, 1.4]
Exp. 3	[14.3, 27.2]	[10.0, 74.2]	[8.2, 16.7]	[4.8, 32.4]	[2.5, 3.6]	[1.9, 4.3]

Extensions & Guidance

Extensions: • Many more experiments support findings. • Insights hold also under Equalized Odds and Demographic Parity.

Guidance: • Design **ecosystem-aware** audits. • Regulate **correlation/overlap**. • Optimize **joint** fairness.