

Is it Still Fair? A Comparative Evaluation of Fairness Algorithms through the Lens of Covariate Drift-Supplementary

Anonymous Author(s)*

CCS CONCEPTS

- Computing methodologies → Regularization; Multi-agent systems;
- Applied computing → Law; *Interactive learning environments*;
- Information systems → Clustering and classification.

KEYWORDS

Algorithmic Fairness, Covariate Drift, Fairness Algorithms, Robustness

ACM Reference Format:

Anonymous Author(s). 2024. Is it Still Fair? A Comparative Evaluation of Fairness Algorithms through the Lens of Covariate Drift-Supplementary. In *Proceedings of The 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'24)*. ACM, New York, NY, USA, 38 pages. <https://doi.org/XXXXXXXX.XXXXXXXX>

1 DCD VS. FAIRNESS

1.1 Adult Dataset

2 ROBUSTNESS OF FAIRNESS-AWARE-MODELS IN TERMS OF FAIRNESS AND PREDICTIVE PERFORMANCE

2.1 NWF Dataset

2.2 ITF Dataset

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SIGIR'24, July 14–18, 2024, Washington D.C., USA

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-XXXX-X/18/06
<https://doi.org/XXXXXXXX.XXXXXXXX>

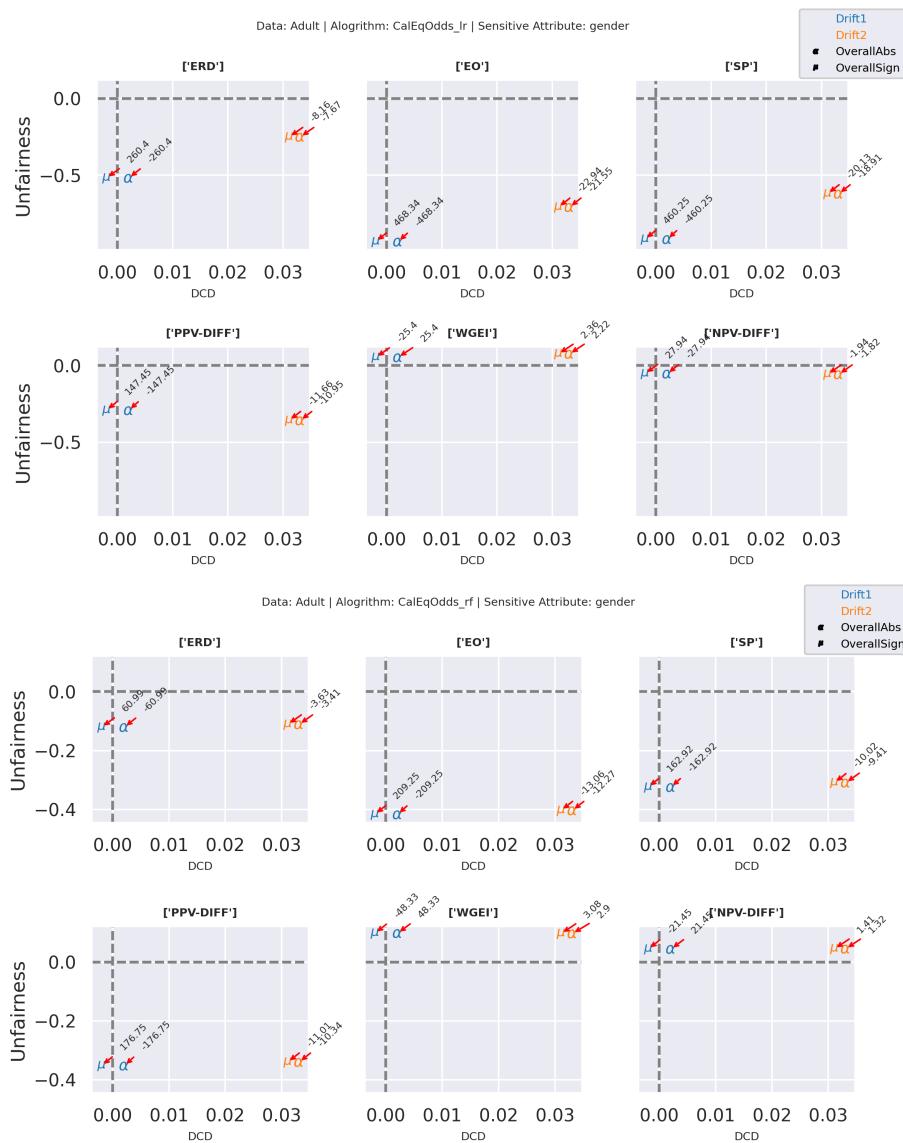
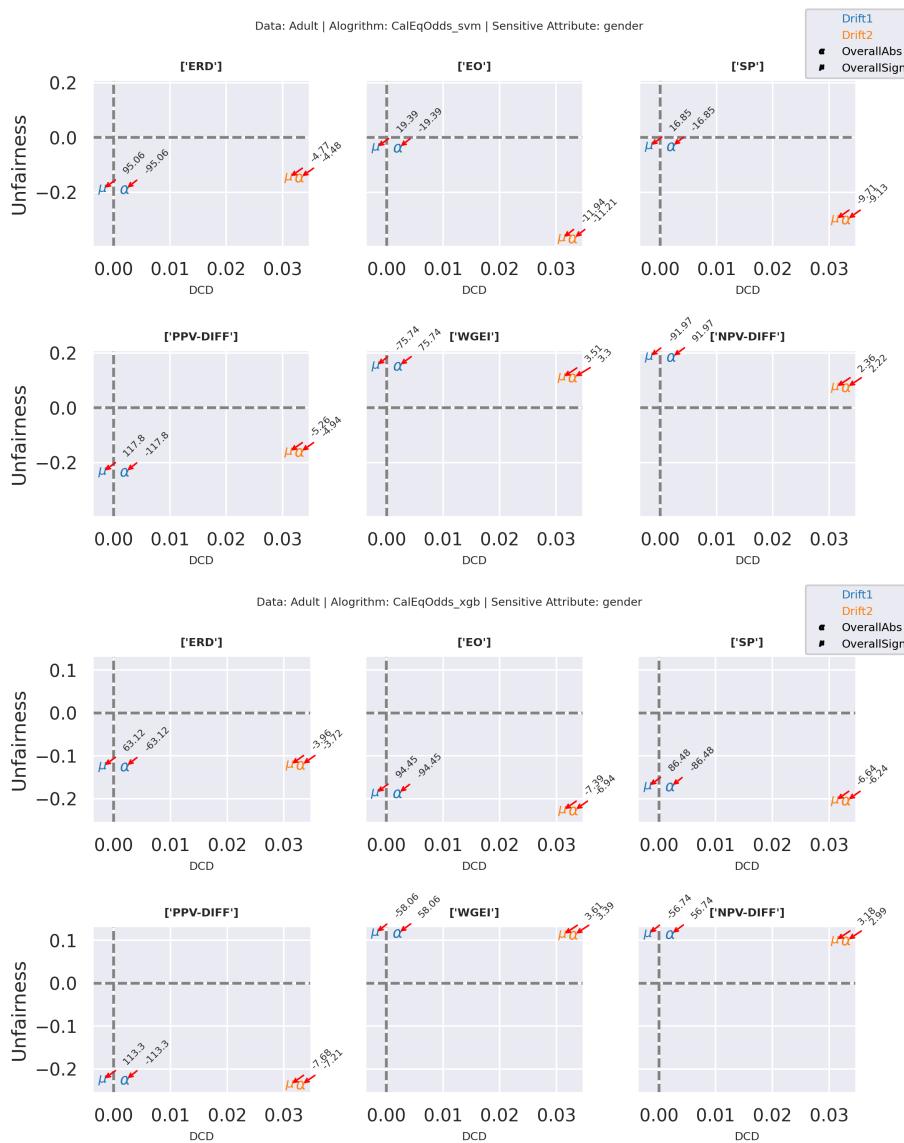


Figure 1: DCD vs unfairness for CalEqOdds for Adult dataset.

**Figure 2: DCD vs unfairness for CalEqOdds for Adult dataset.**

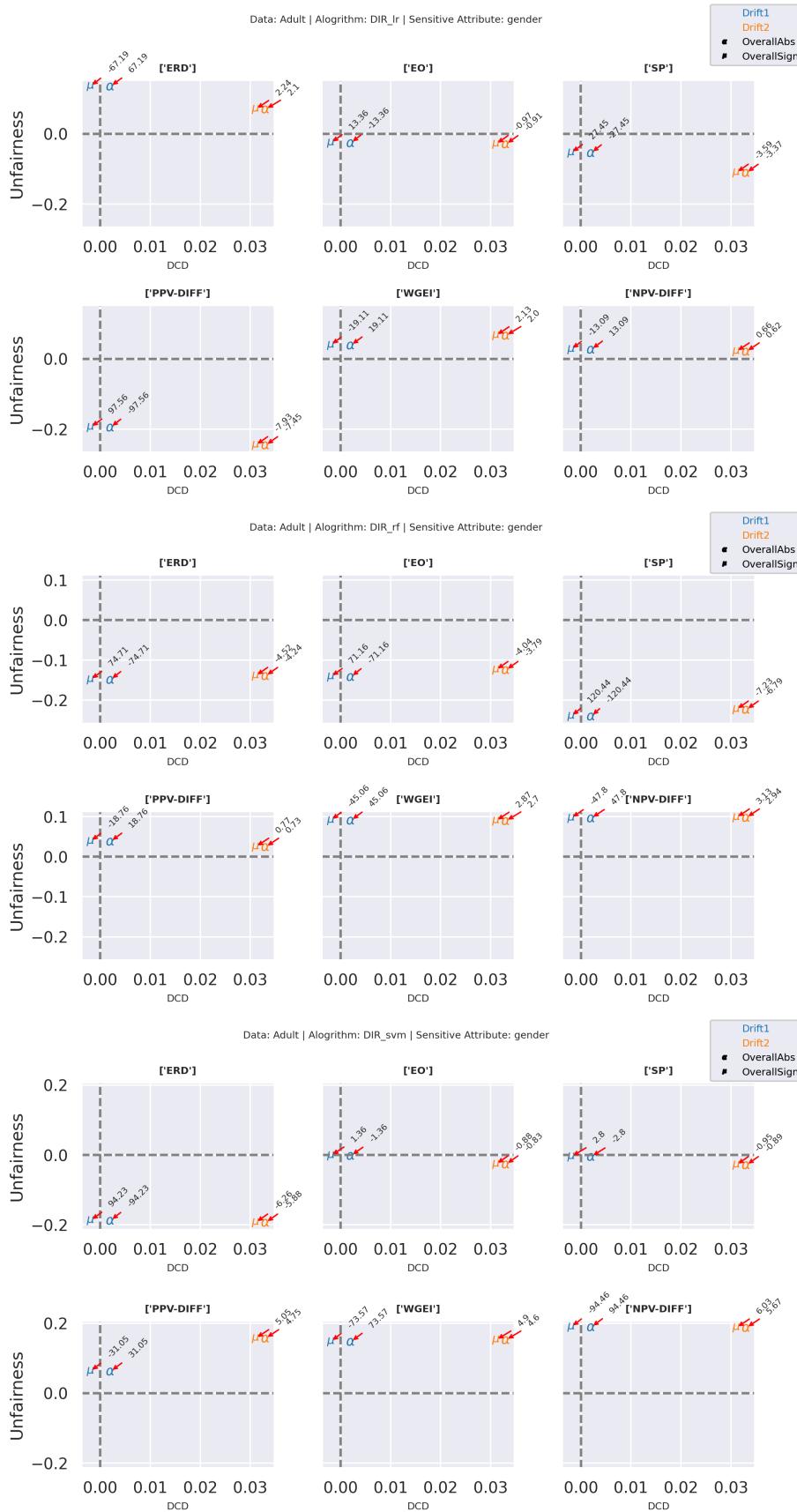


Figure 3: DCD vs unfairness for DIR for Adult dataset.

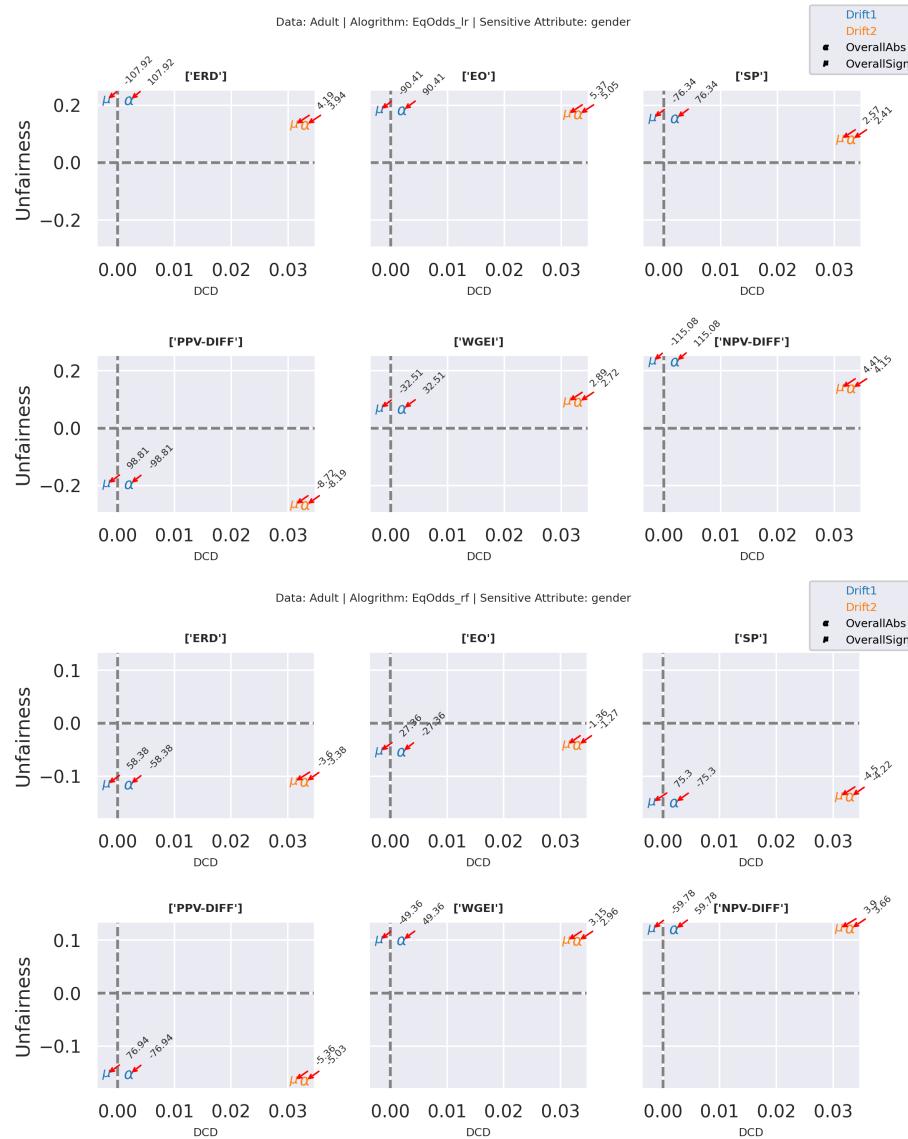


Figure 4: DCD vs unfairness for EqOdds for Adult dataset.

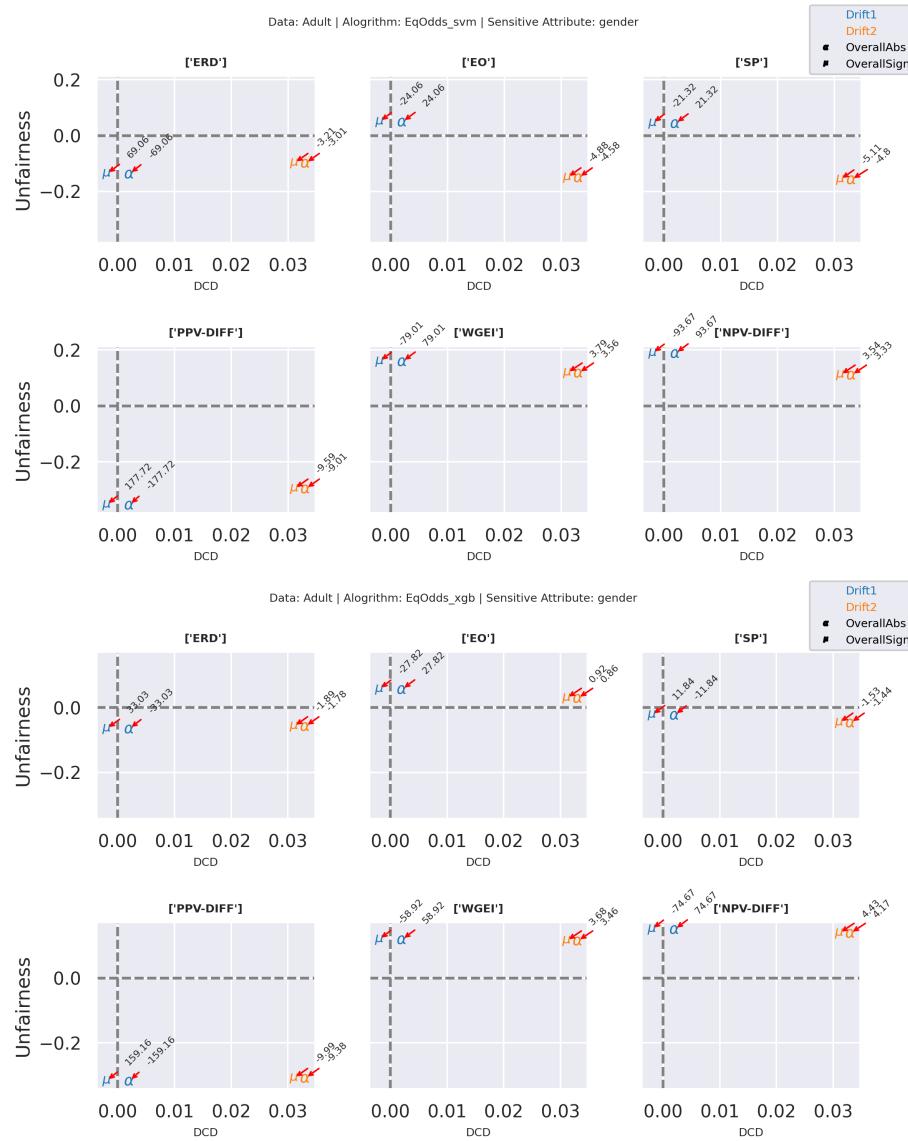


Figure 5: DCD vs unfairness for EqOdds for Adult dataset.

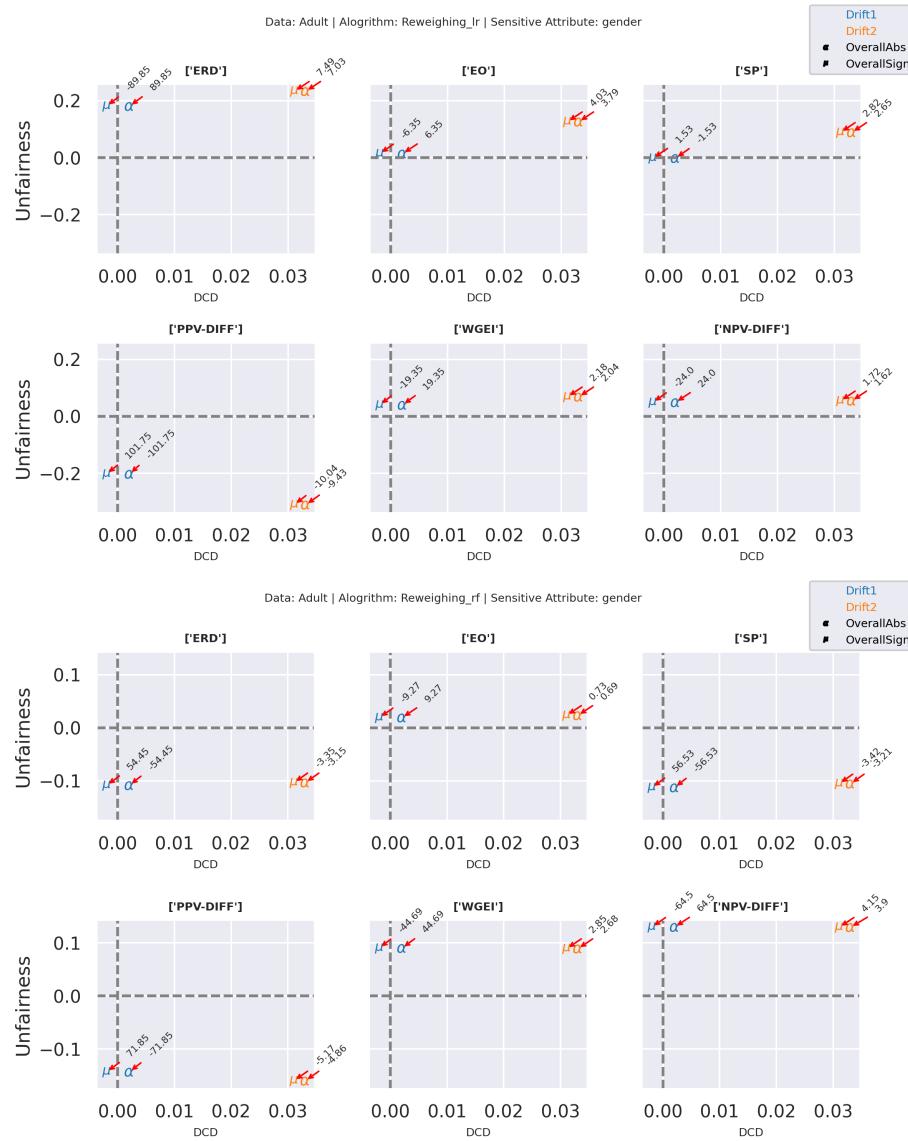


Figure 6: DCD vs unfairness for Reweighting for Adult dataset.

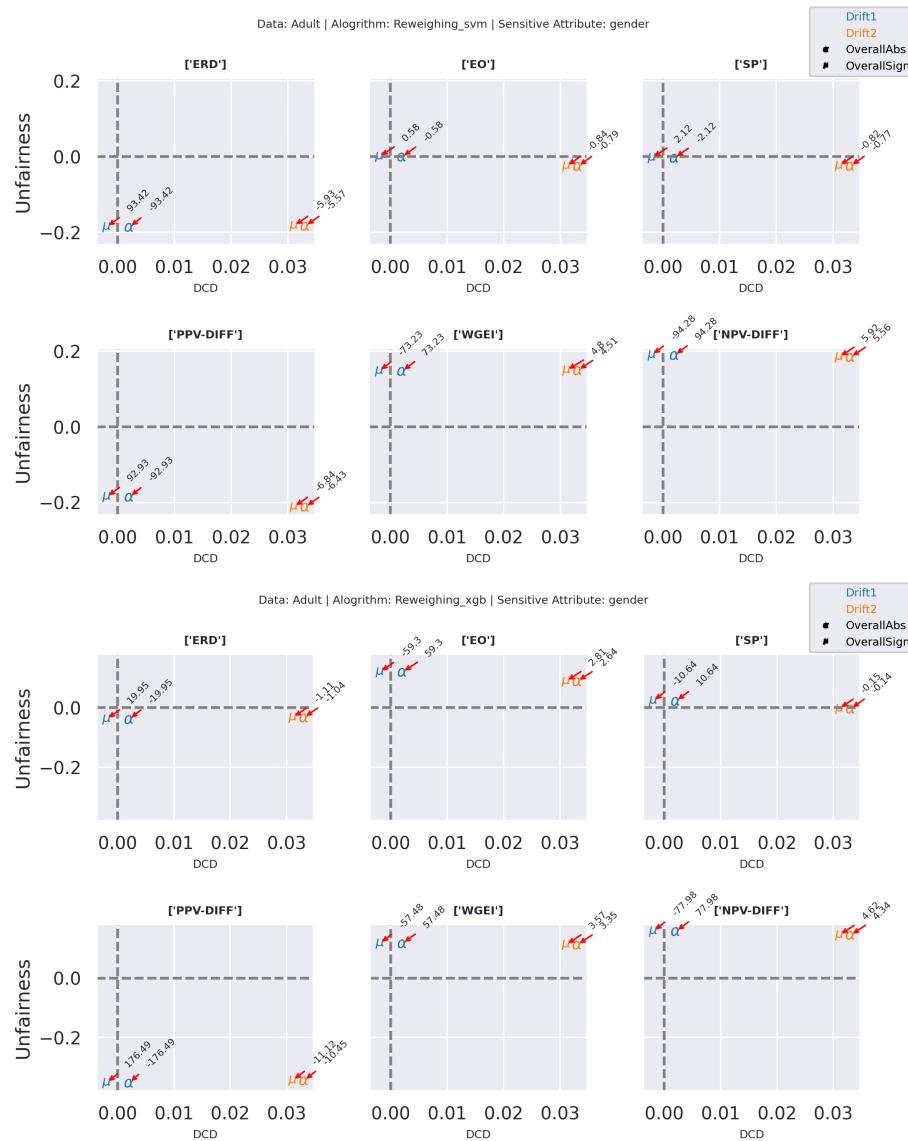


Figure 7: DCD vs unfairness for Reweighting for Adult dataset.

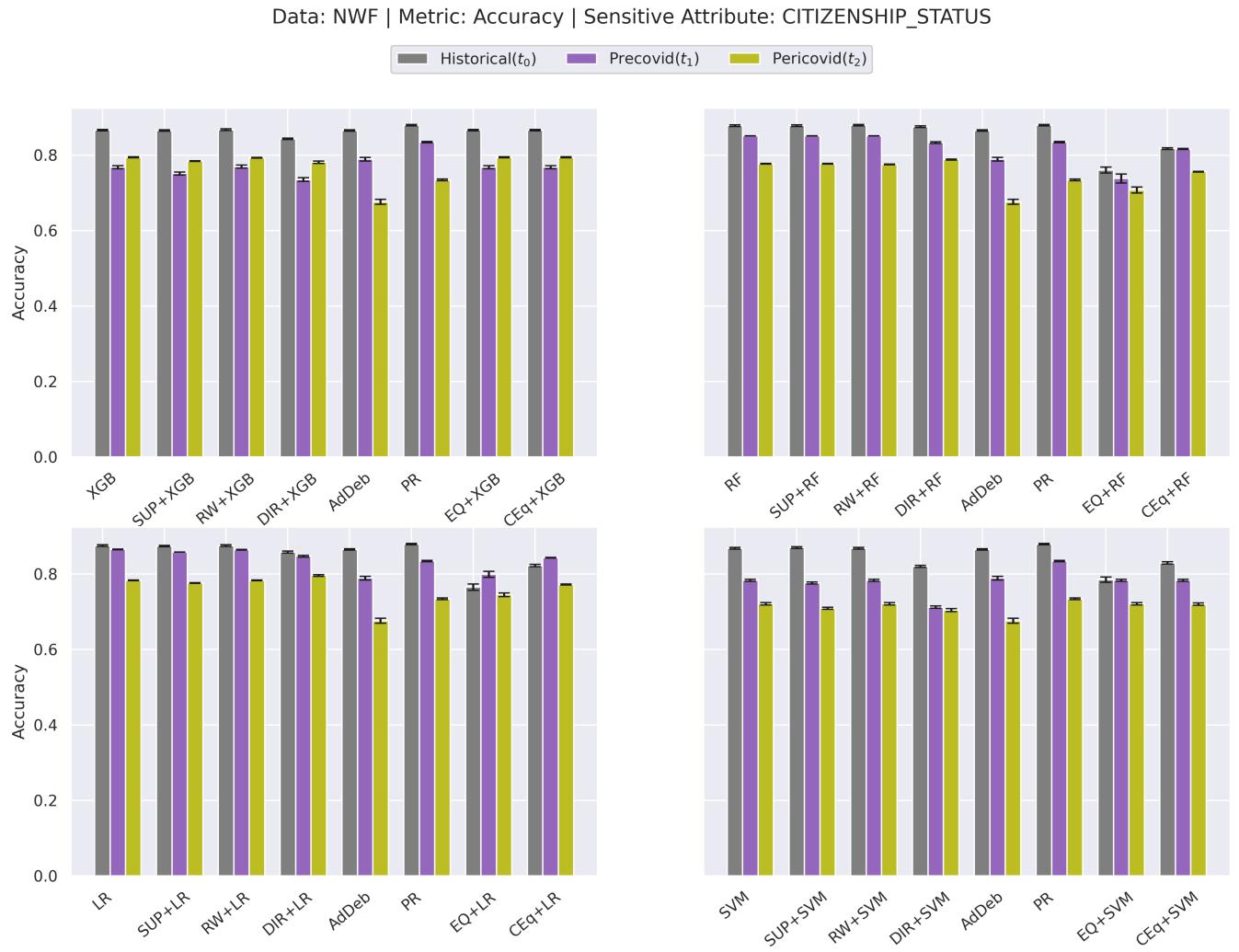


Figure 8: Accuracy for NWF dataset.

Data: NWF | Metric: Bal_Accuracy | Sensitive Attribute: CITIZENSHIP_STATUS

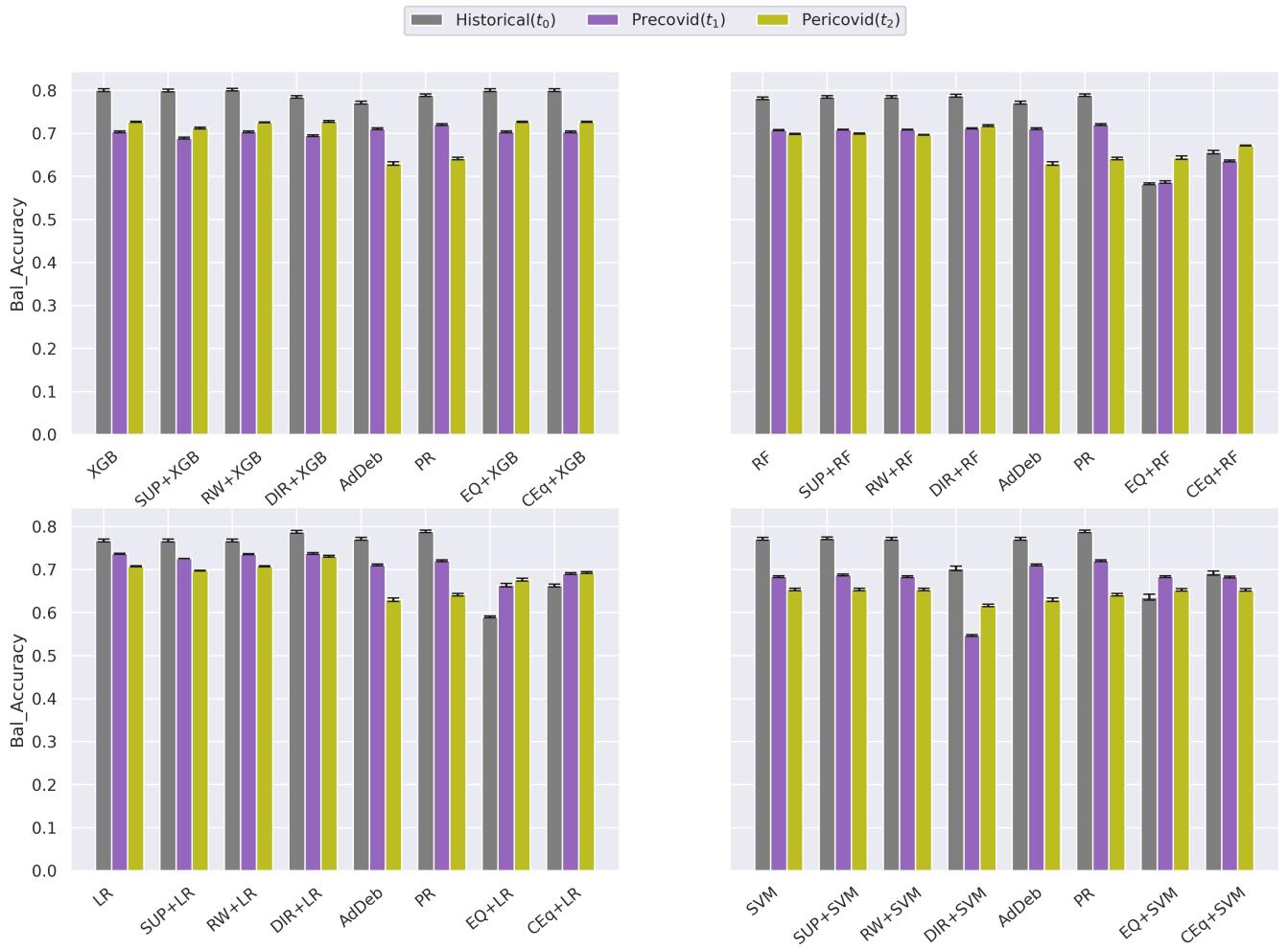


Figure 9: Balanced Accuracy for NWF dataset.

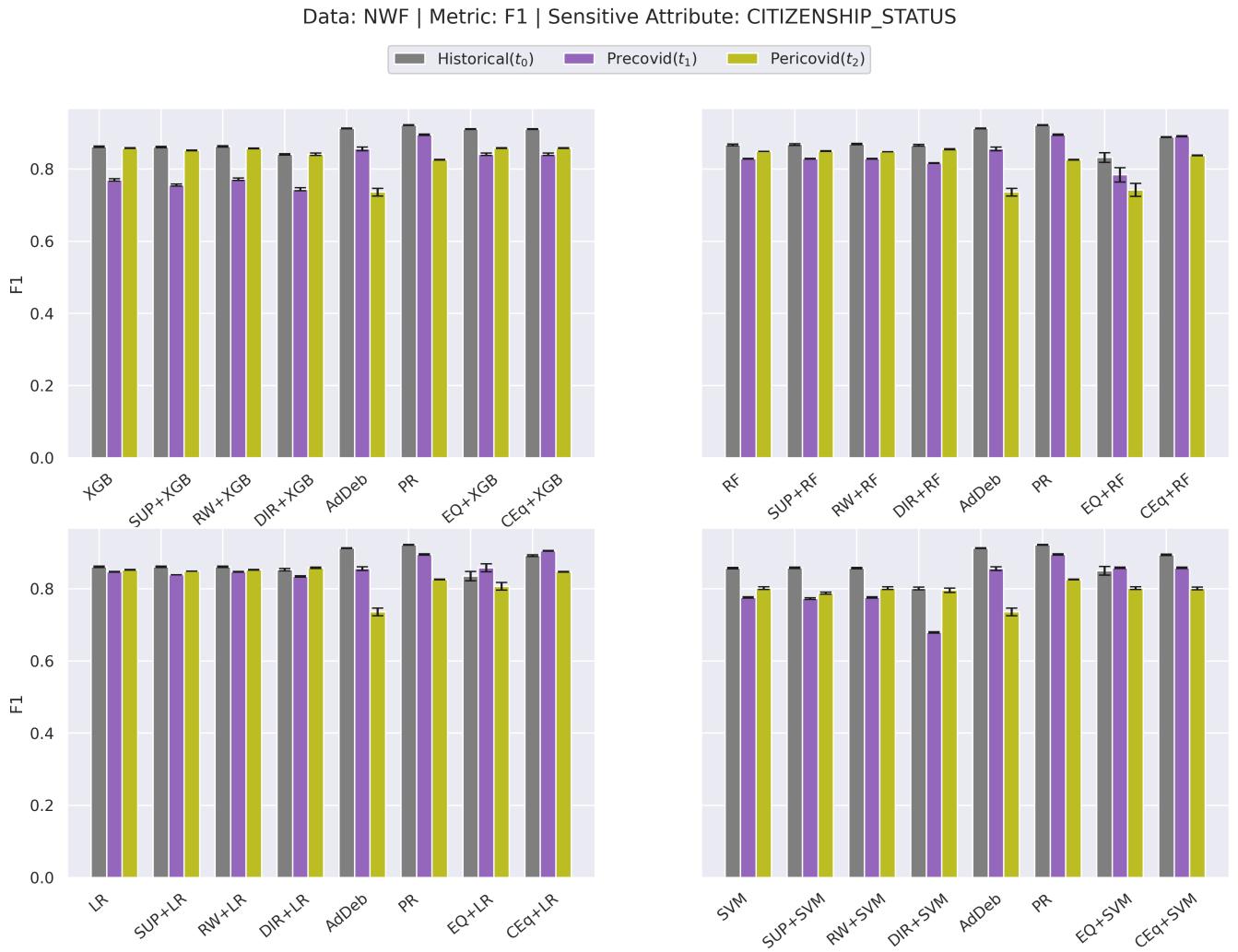


Figure 10: F1 for NWF dataset.

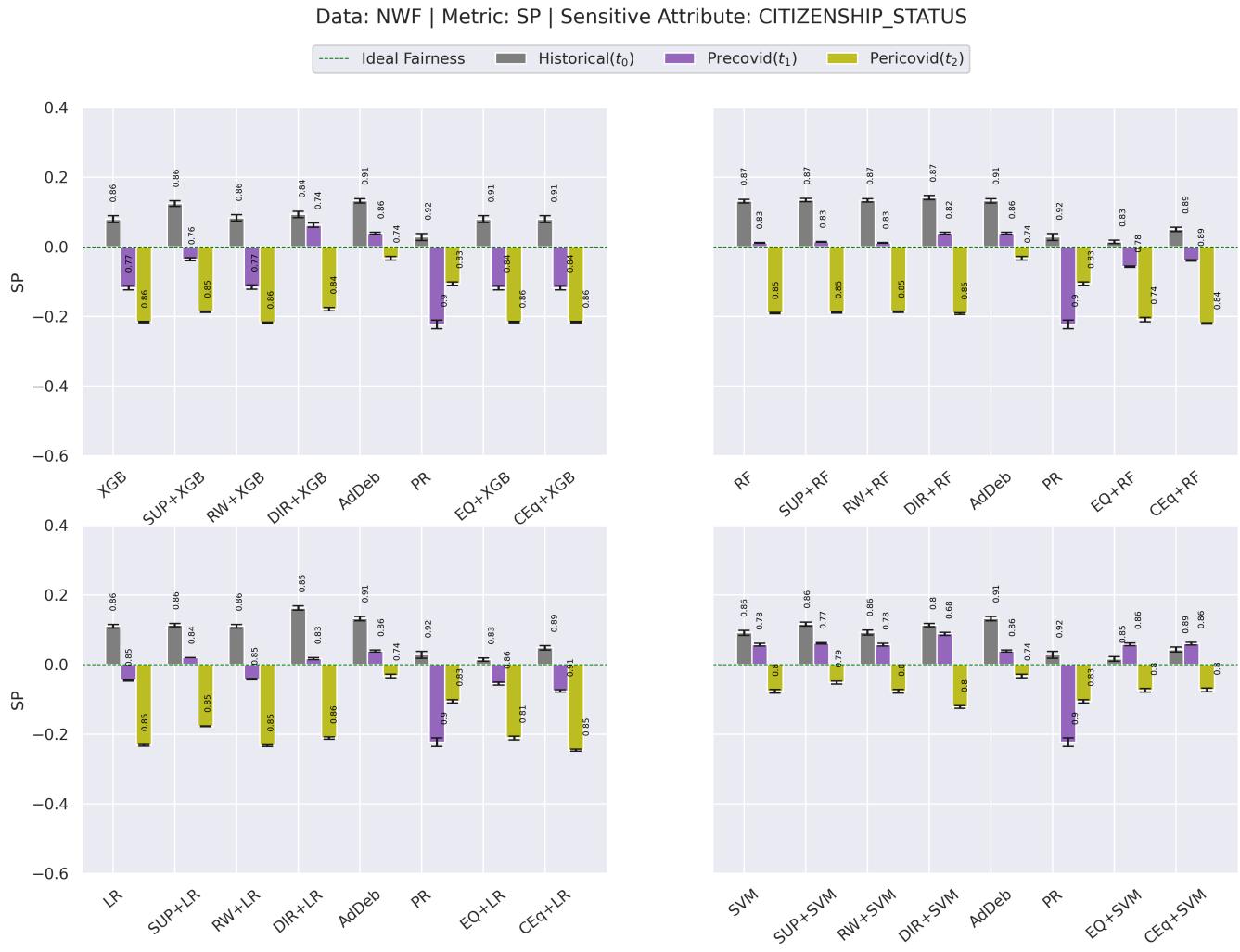


Figure 11: SP for NWF dataset.

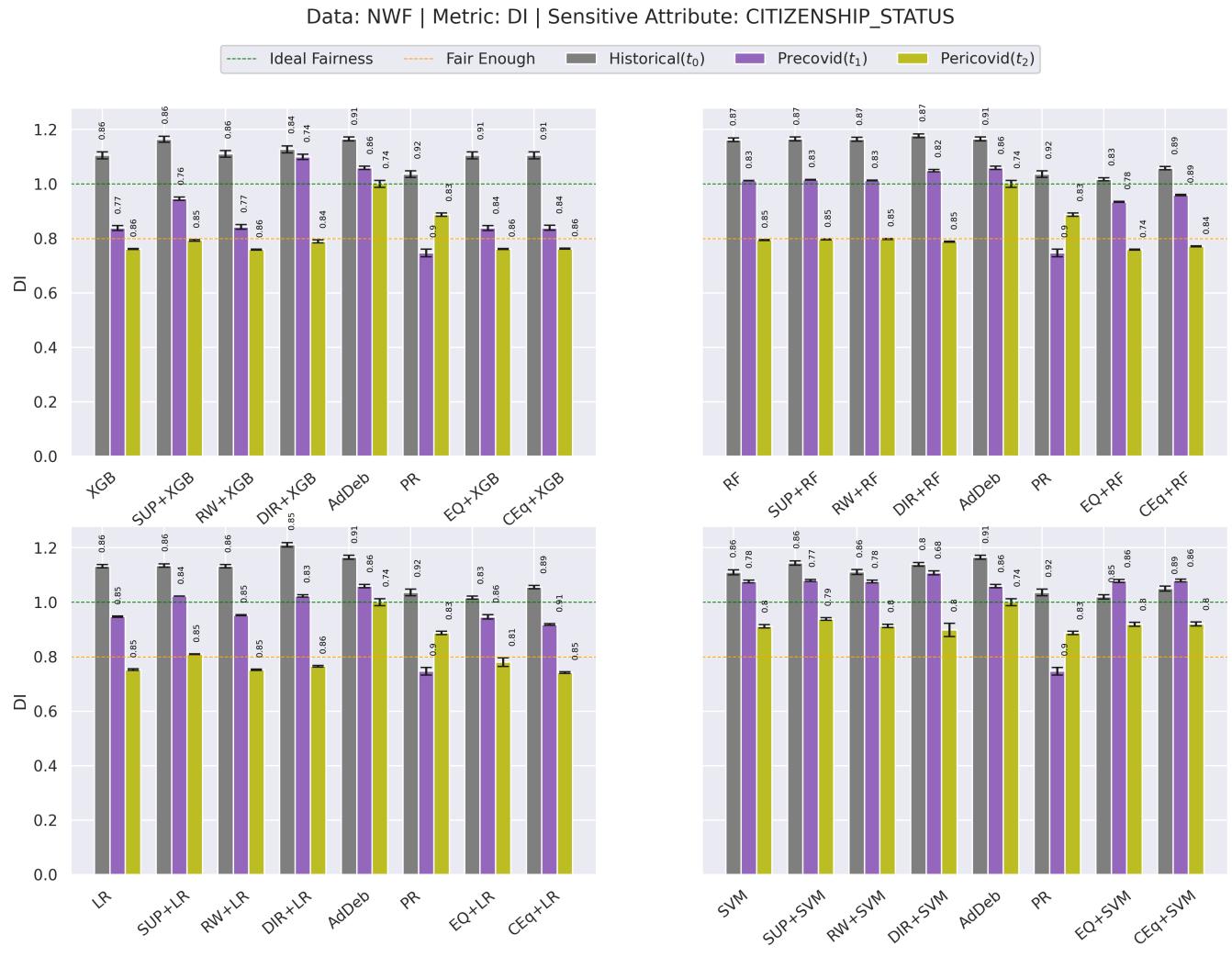


Figure 12: DI for NWF dataset.



Figure 13: EO for NWF dataset.

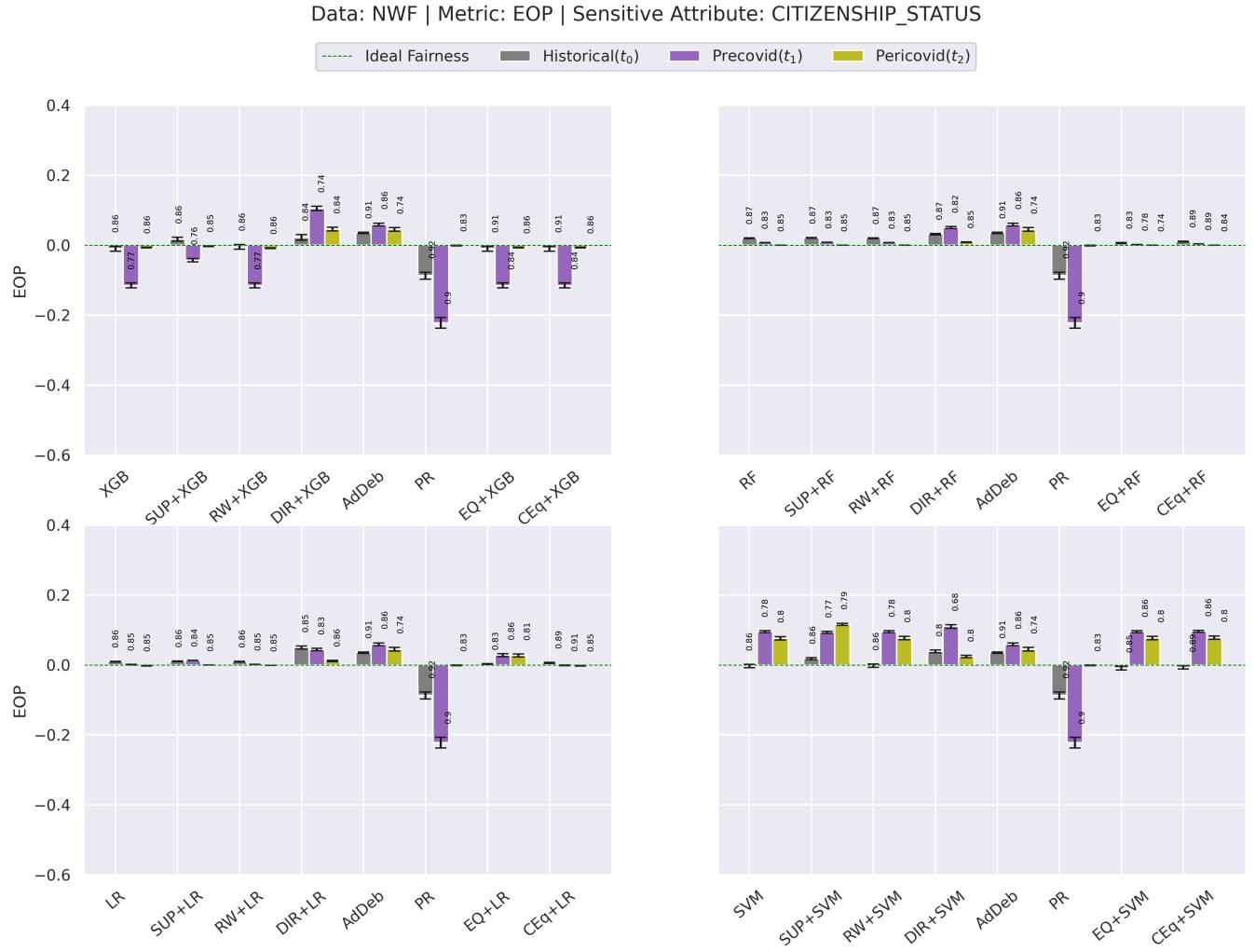


Figure 14: EOP for NWF dataset.

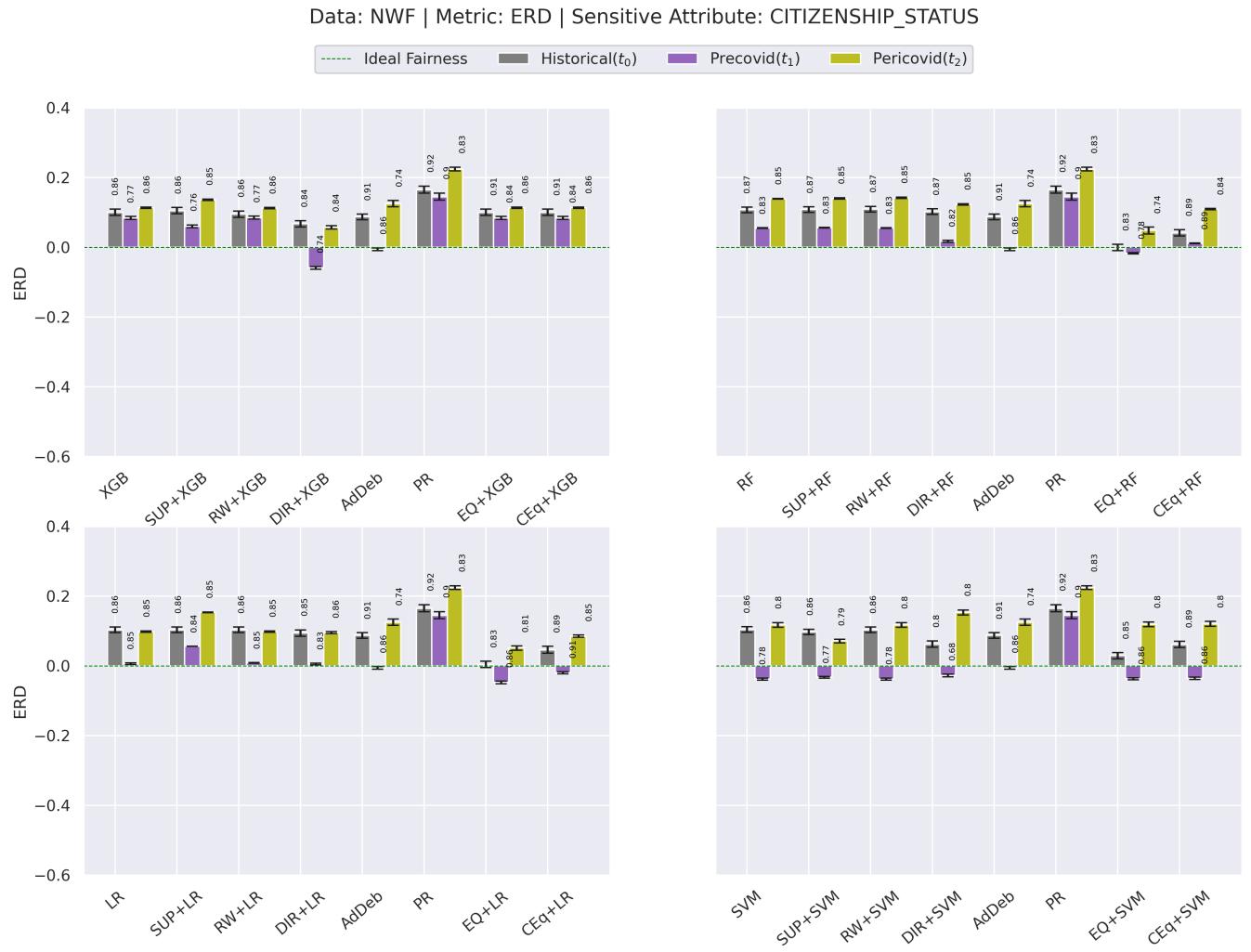


Figure 15: ERD for NWF dataset.

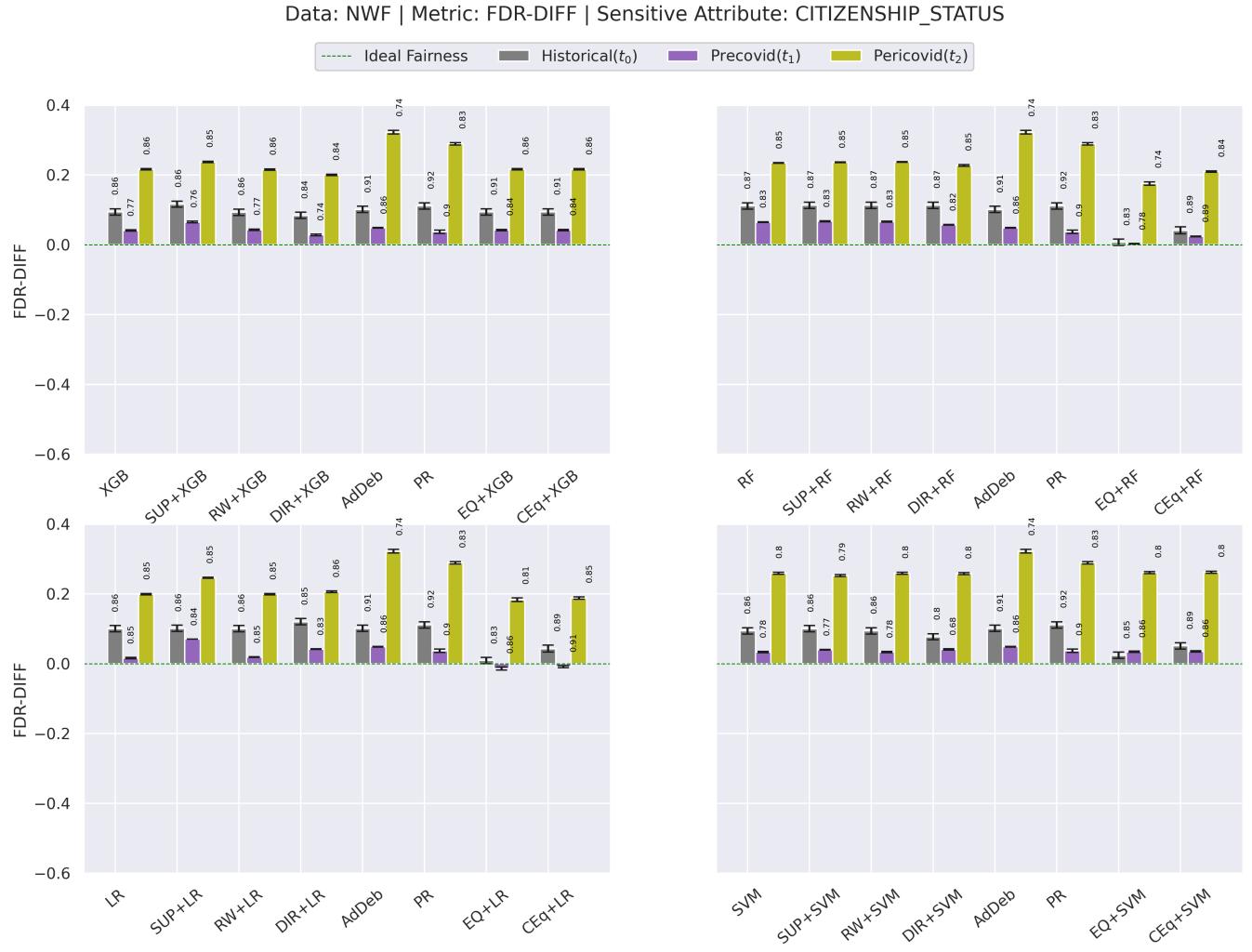


Figure 16: FDR-DIFF for NWF dataset.

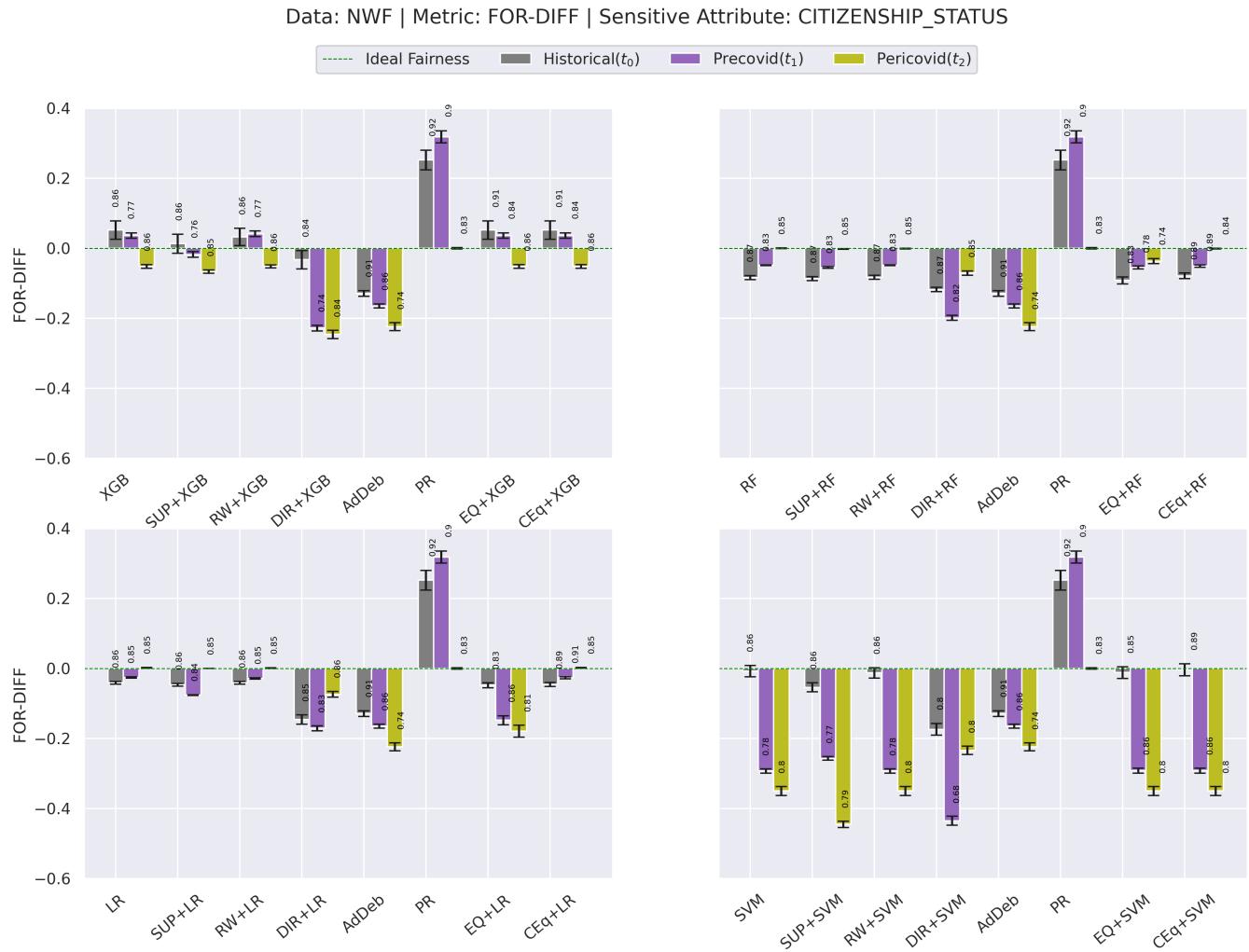


Figure 17: FOR-DIFF for NWF dataset.

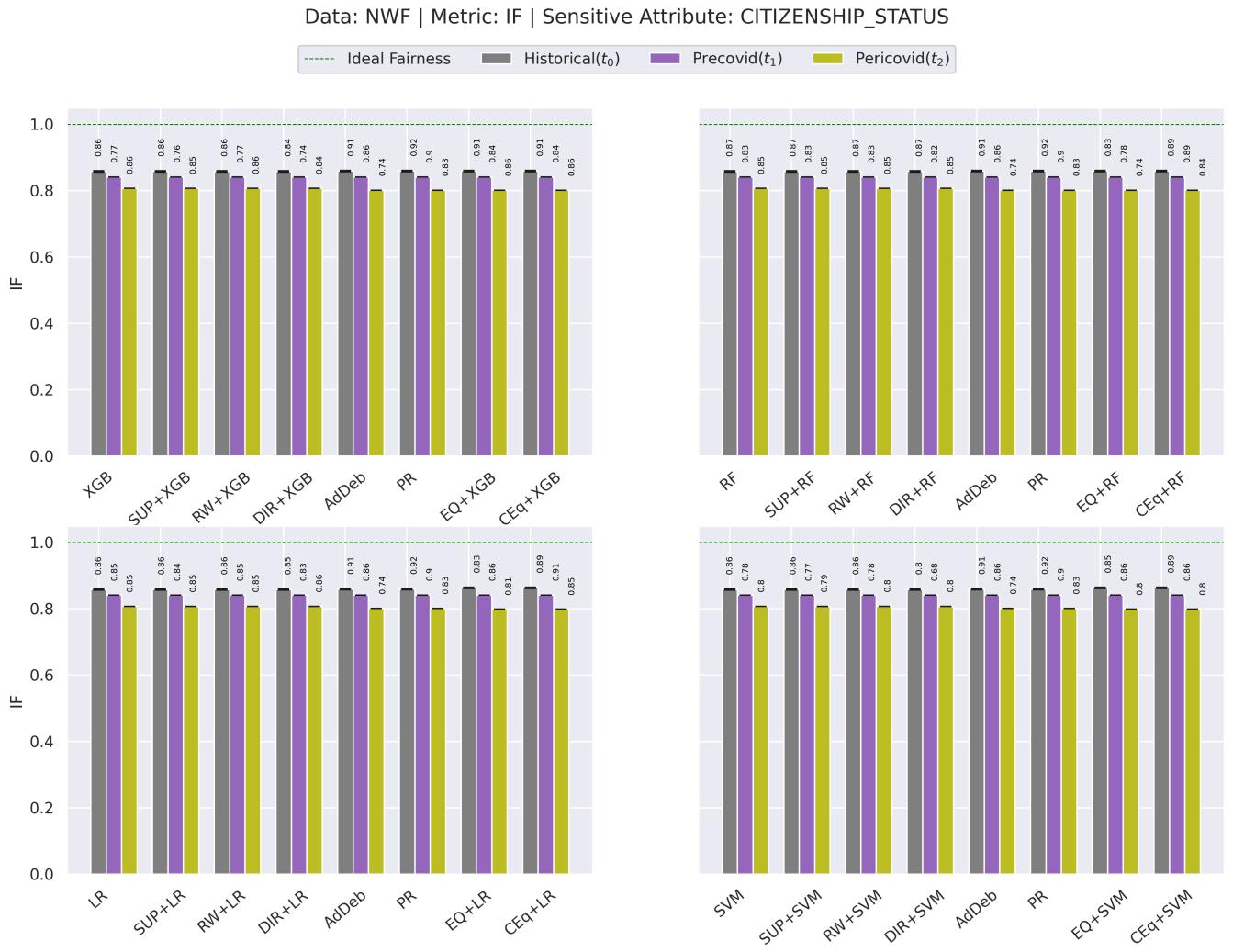


Figure 18: Consistency for NWF dataset.

Data: NWF | Metric: NPV-DIFF | Sensitive Attribute: CITIZENSHIP_STATUS

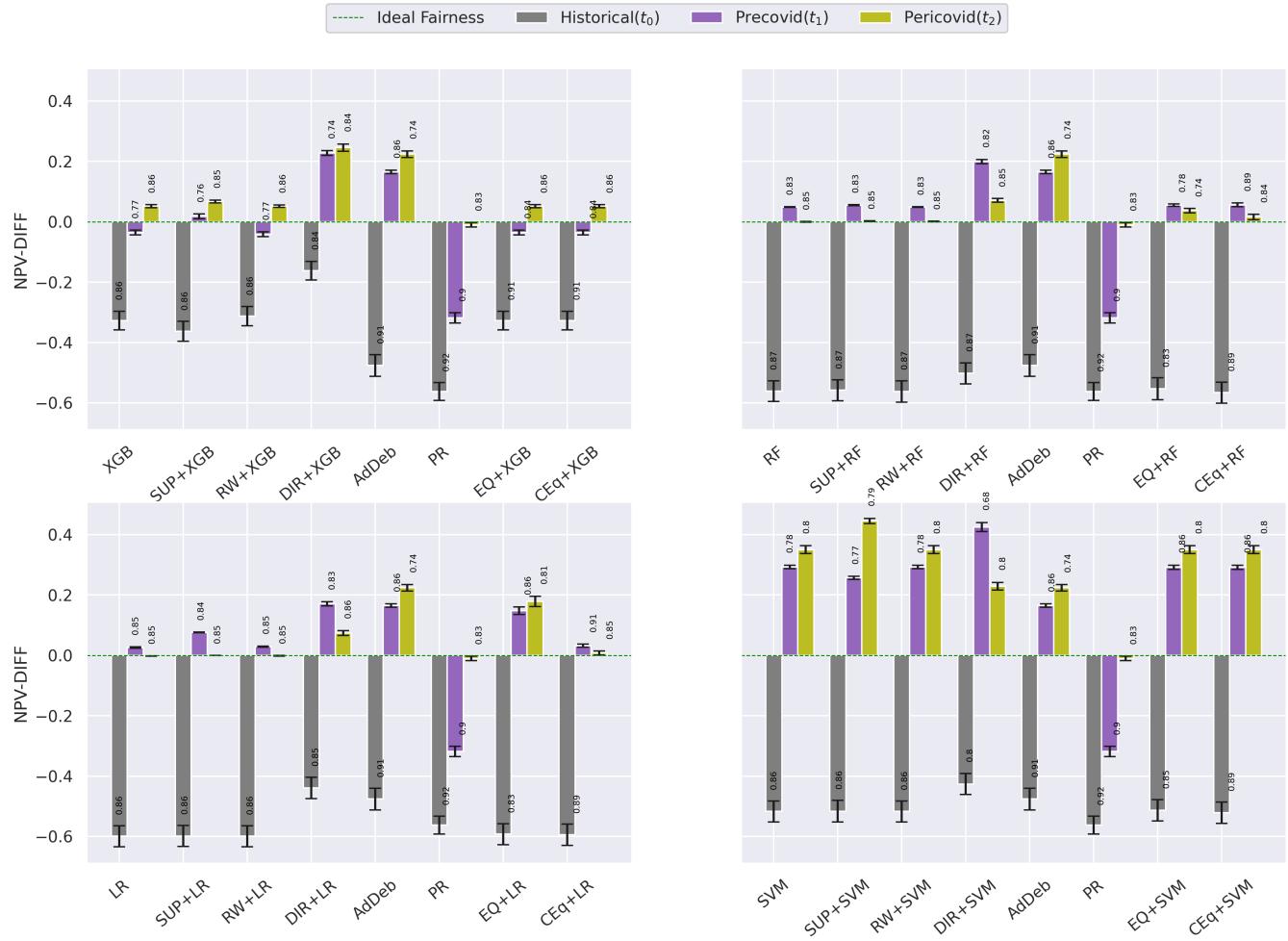


Figure 19: NPV-DIFF for NWF dataset.

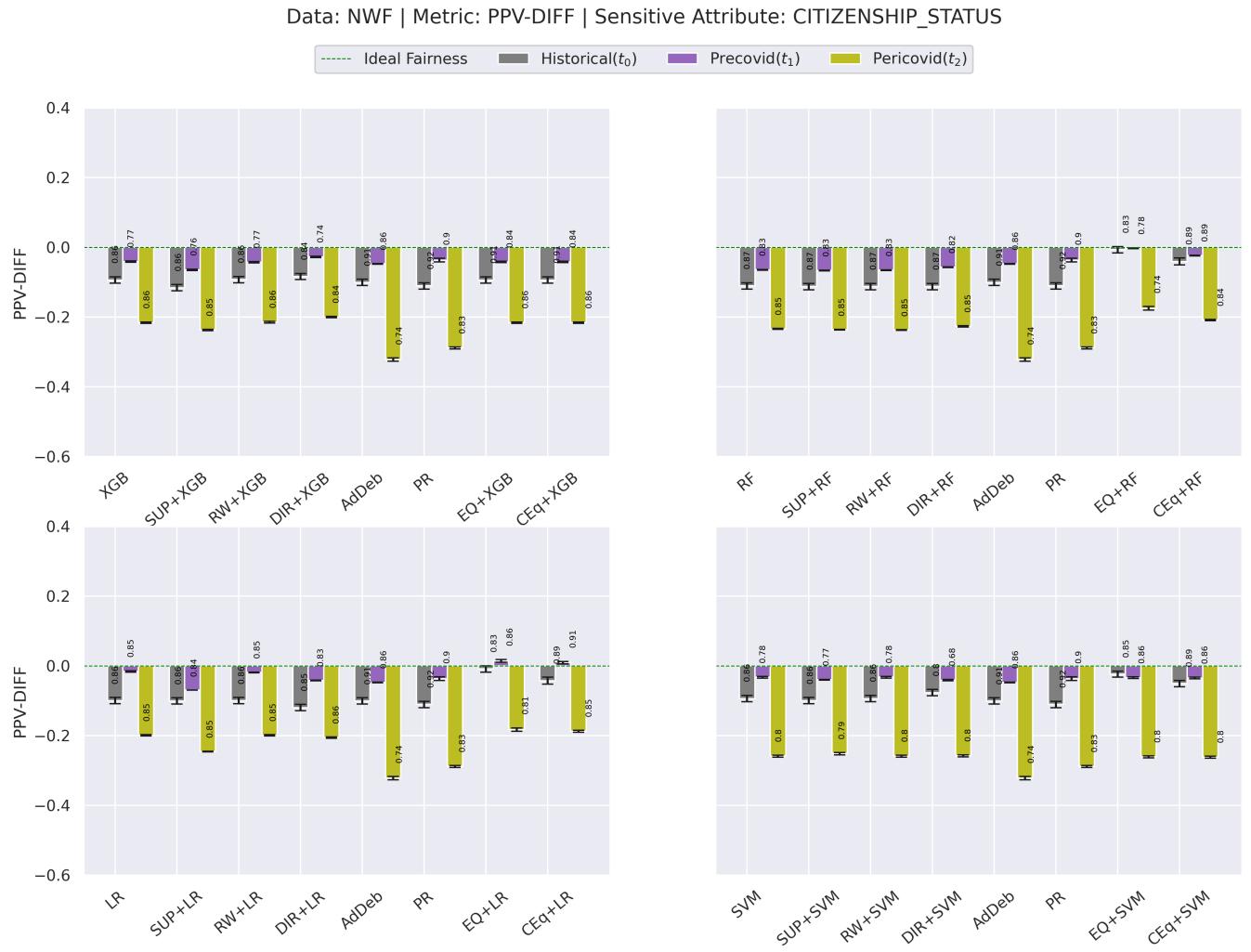


Figure 20: PPV-DIFF for NWF dataset.

Data: NWF | Metric: WGEI | Sensitive Attribute: CITIZENSHIP_STATUS

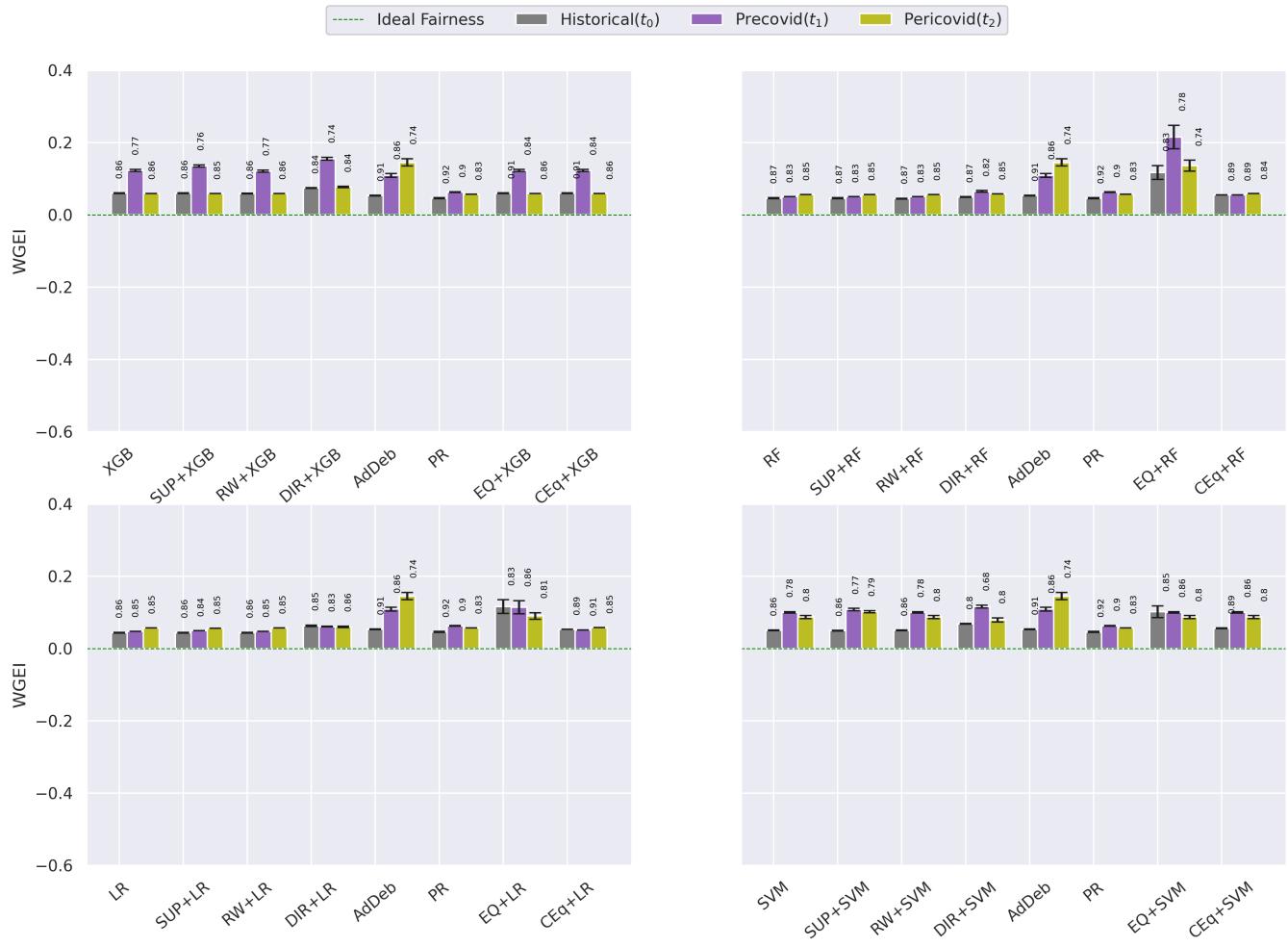
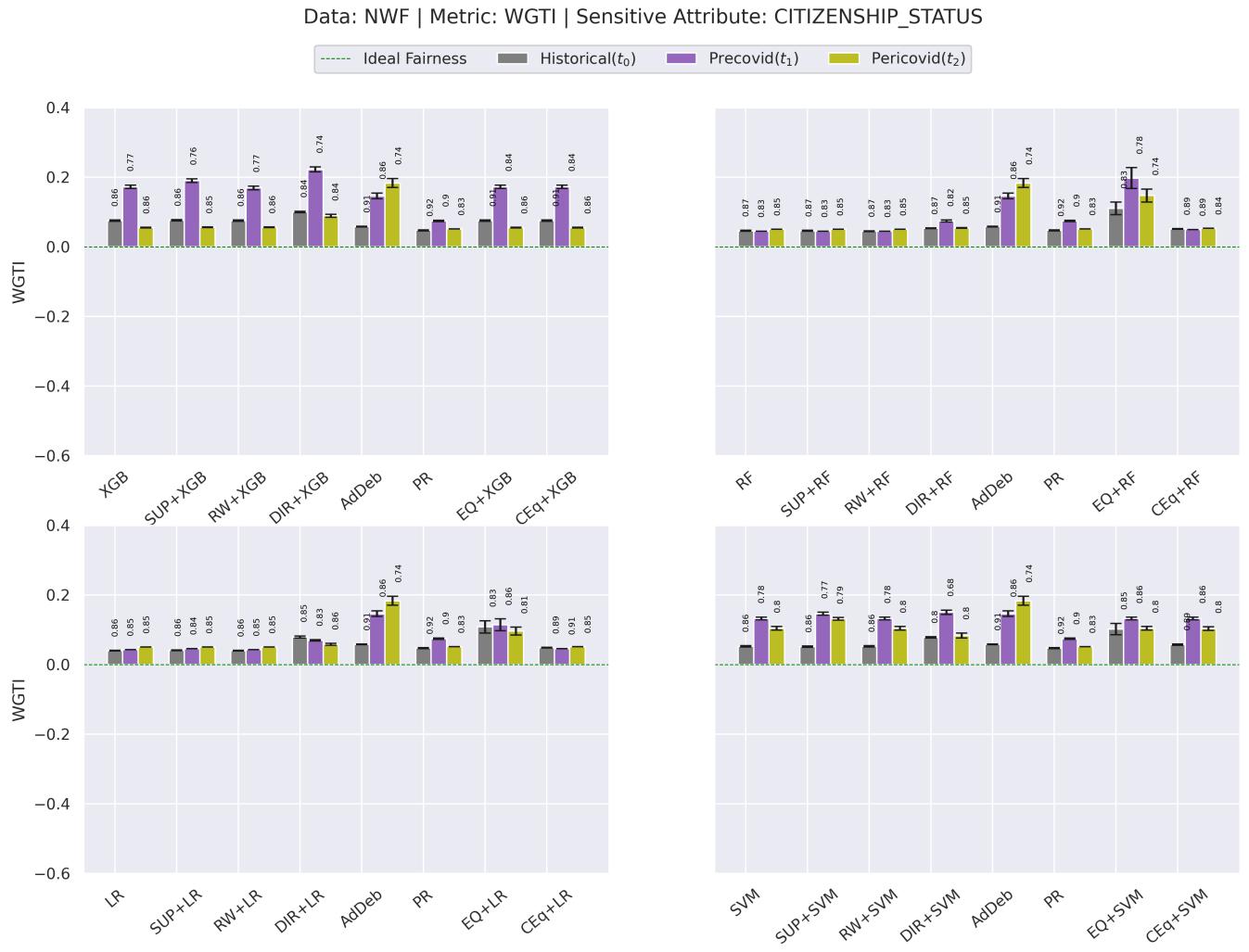


Figure 21: WGEI for NWF dataset.

**Figure 22: WGTI for NWF dataset.**

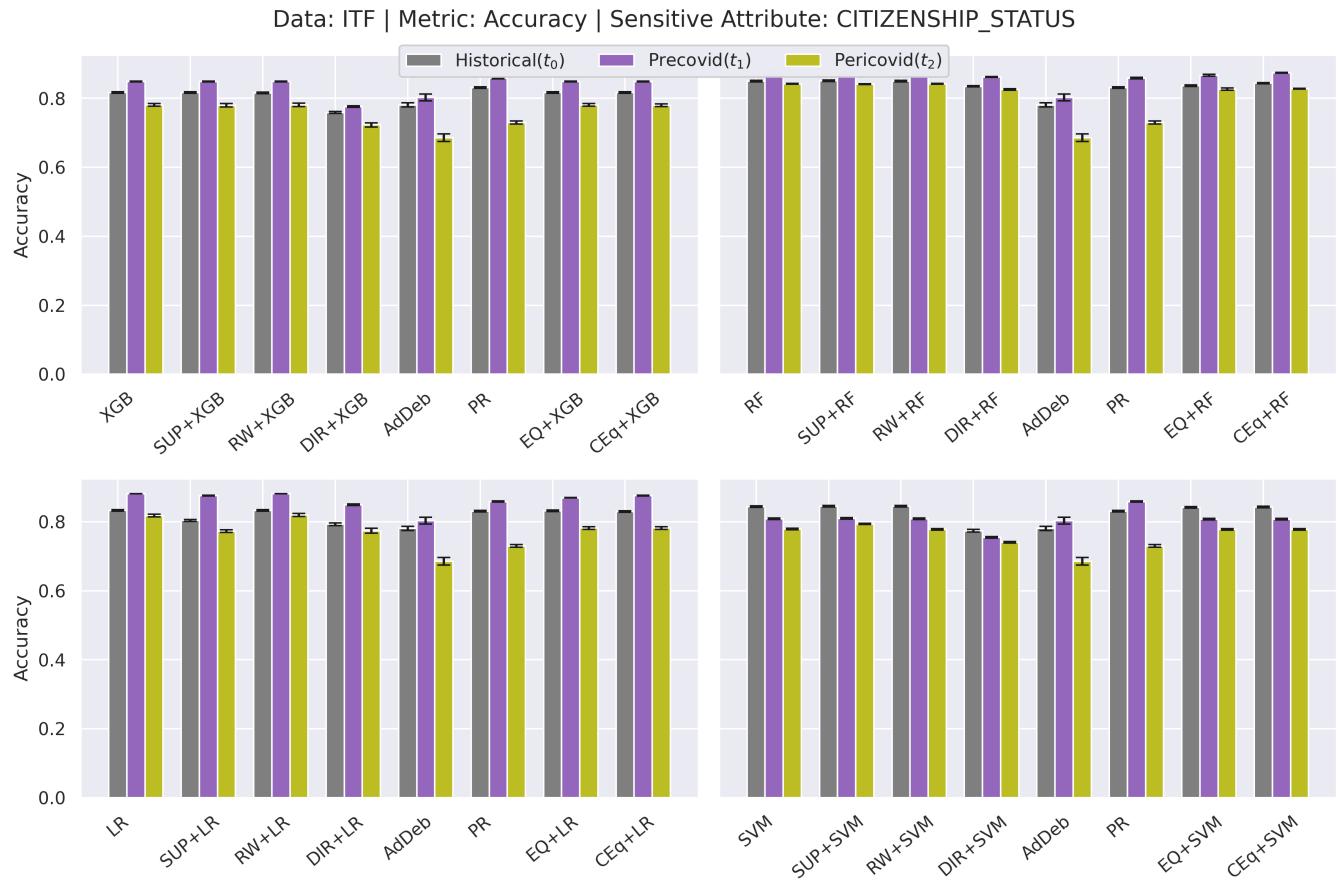


Figure 23: Accuracy for ITF dataset.

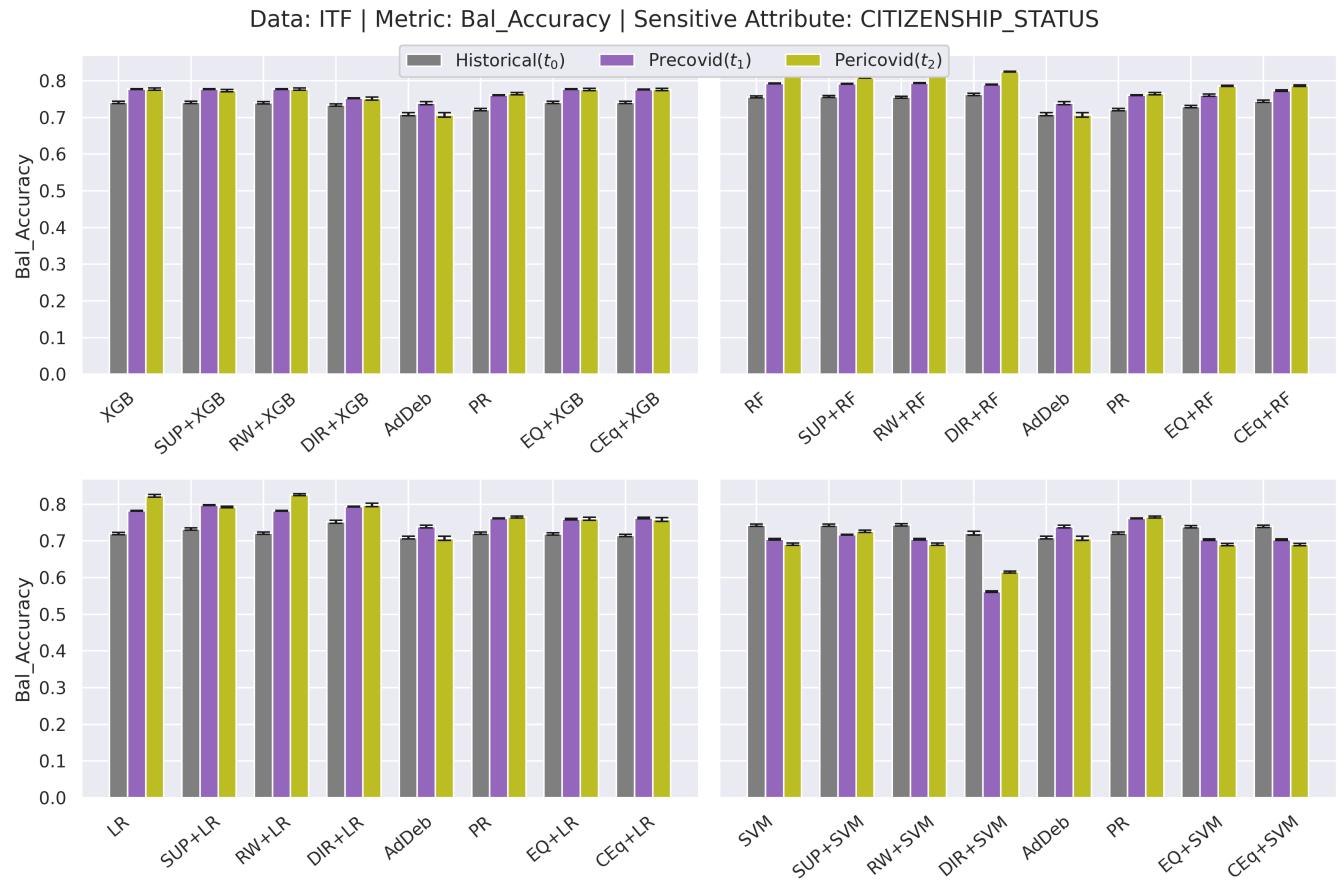


Figure 24: Balanced Accuracy for ITF dataset.

Data: ITF | Metric: F1 | Sensitive Attribute: CITIZENSHIP_STATUS

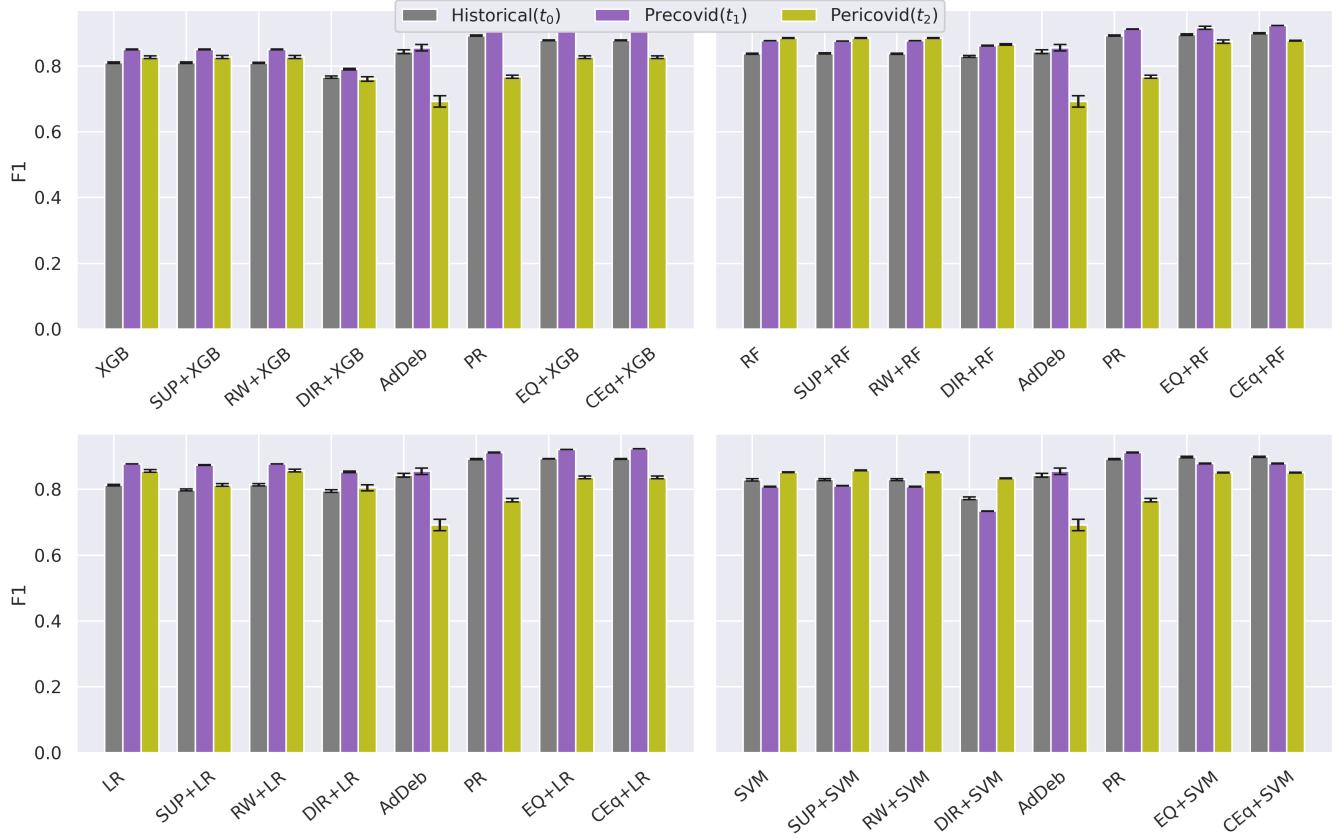


Figure 25: F1 for ITF dataset.

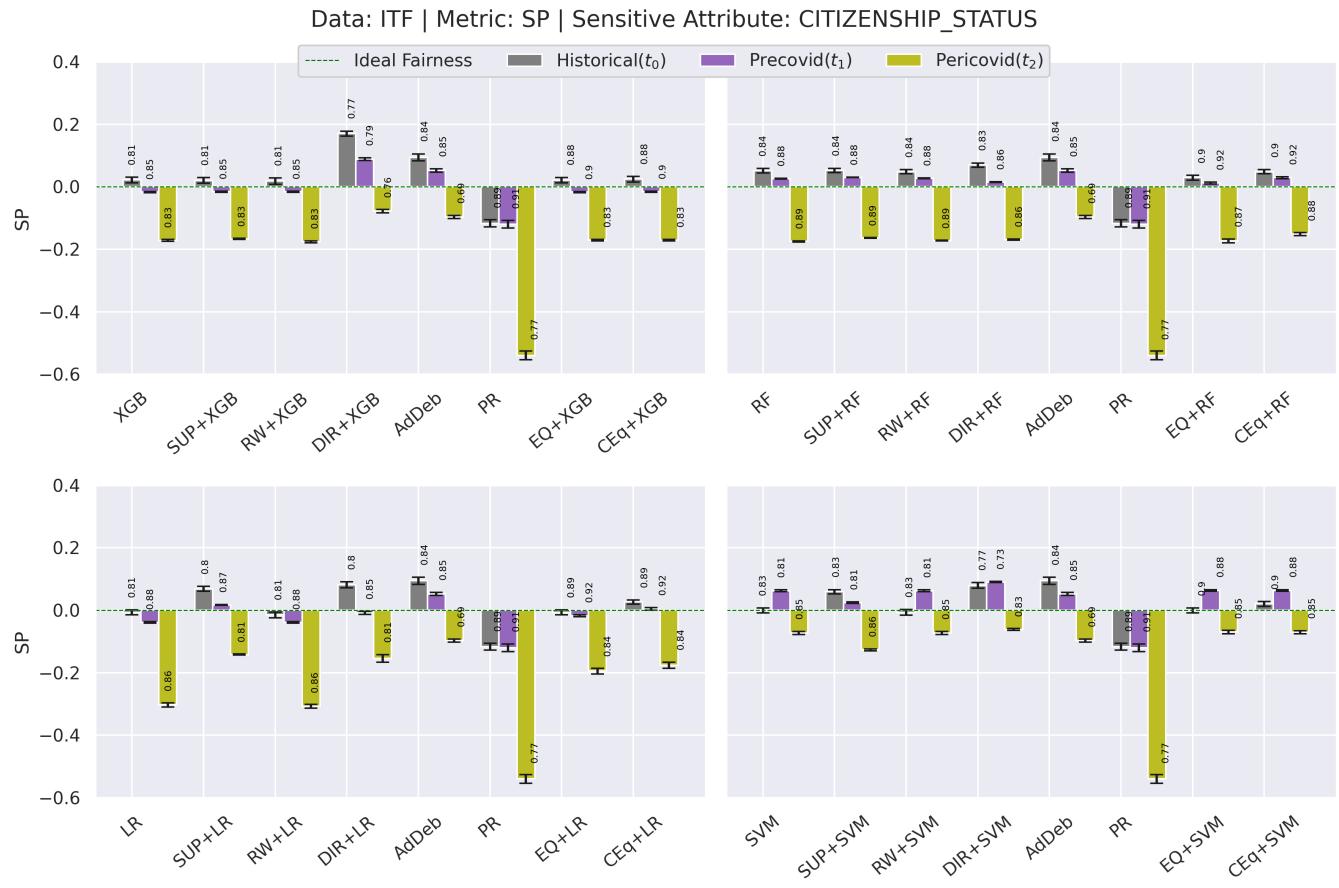


Figure 26: SP for ITF dataset.

Data: ITF | Metric: DI | Sensitive Attribute: CITIZENSHIP_STATUS

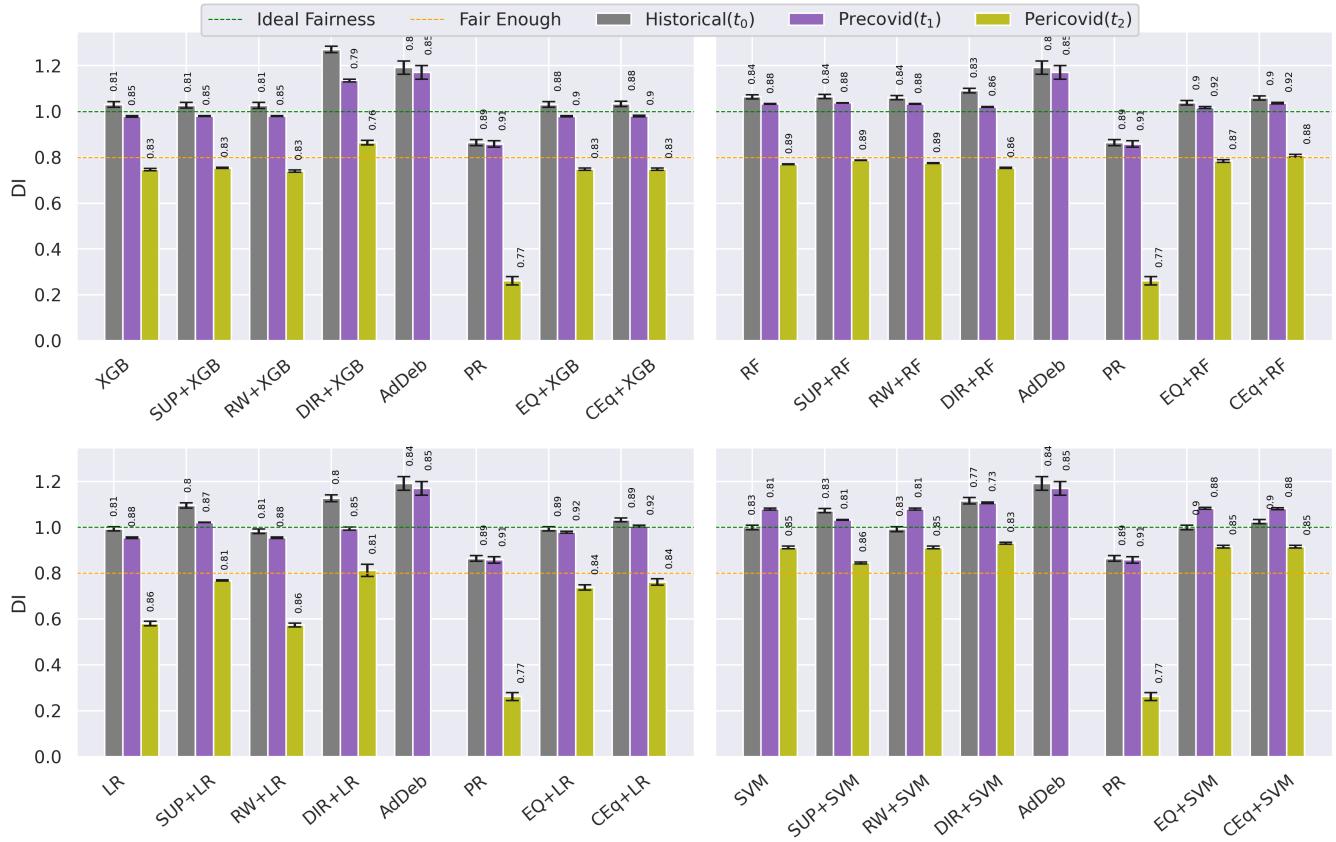


Figure 27: DI for ITF dataset.

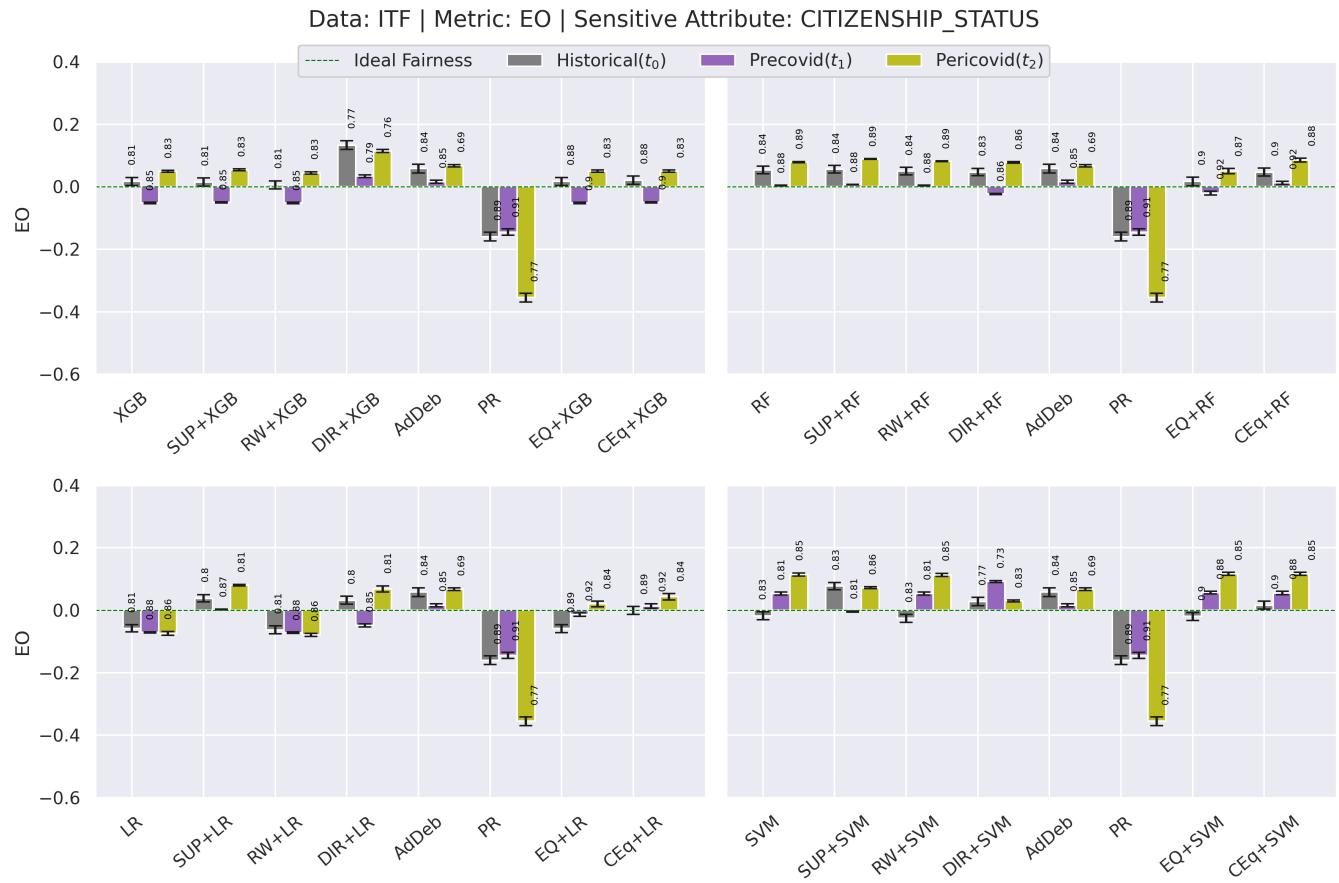


Figure 28: EO for ITF dataset.

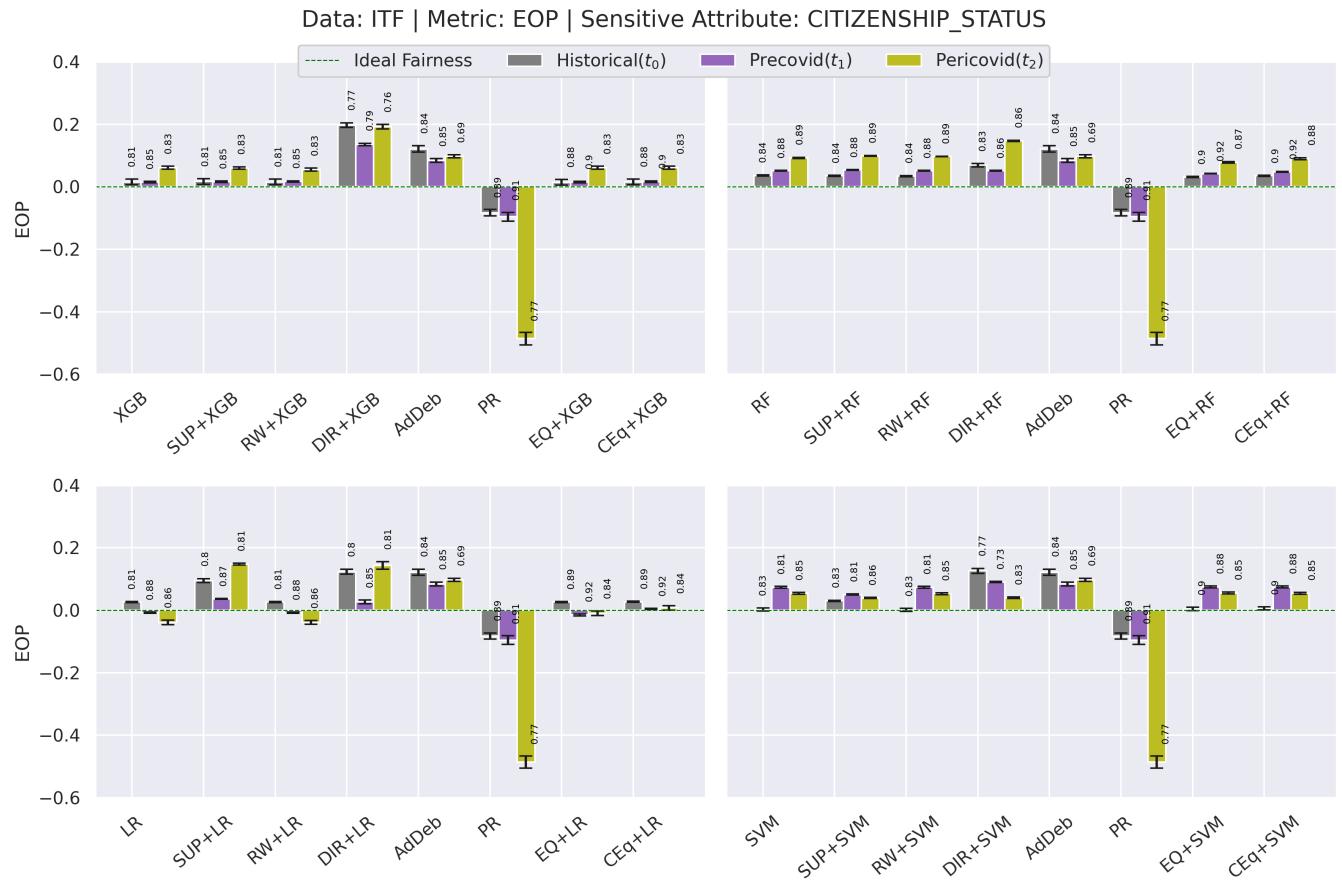


Figure 29: EOP for ITF dataset.

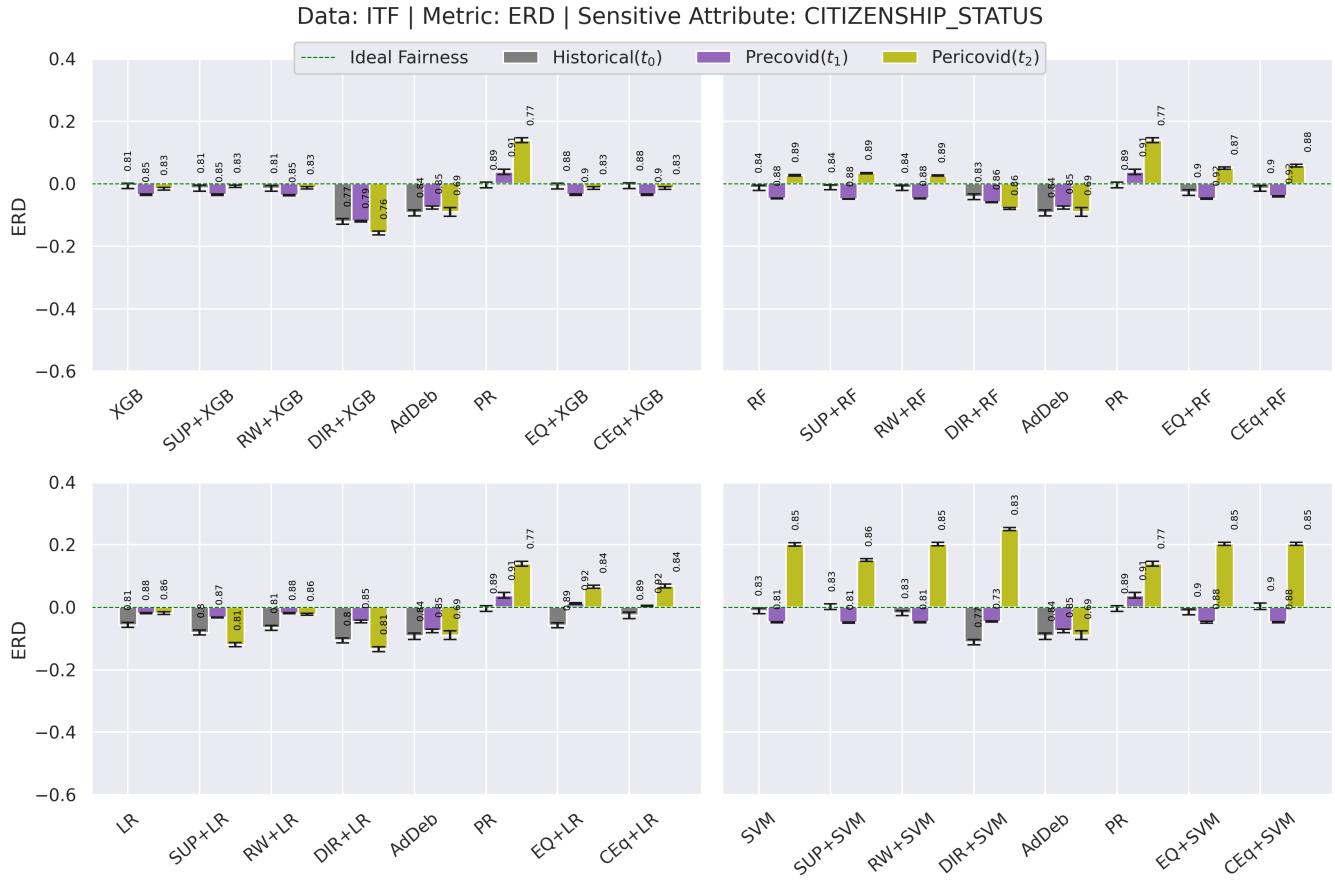


Figure 30: ERD for ITF dataset.

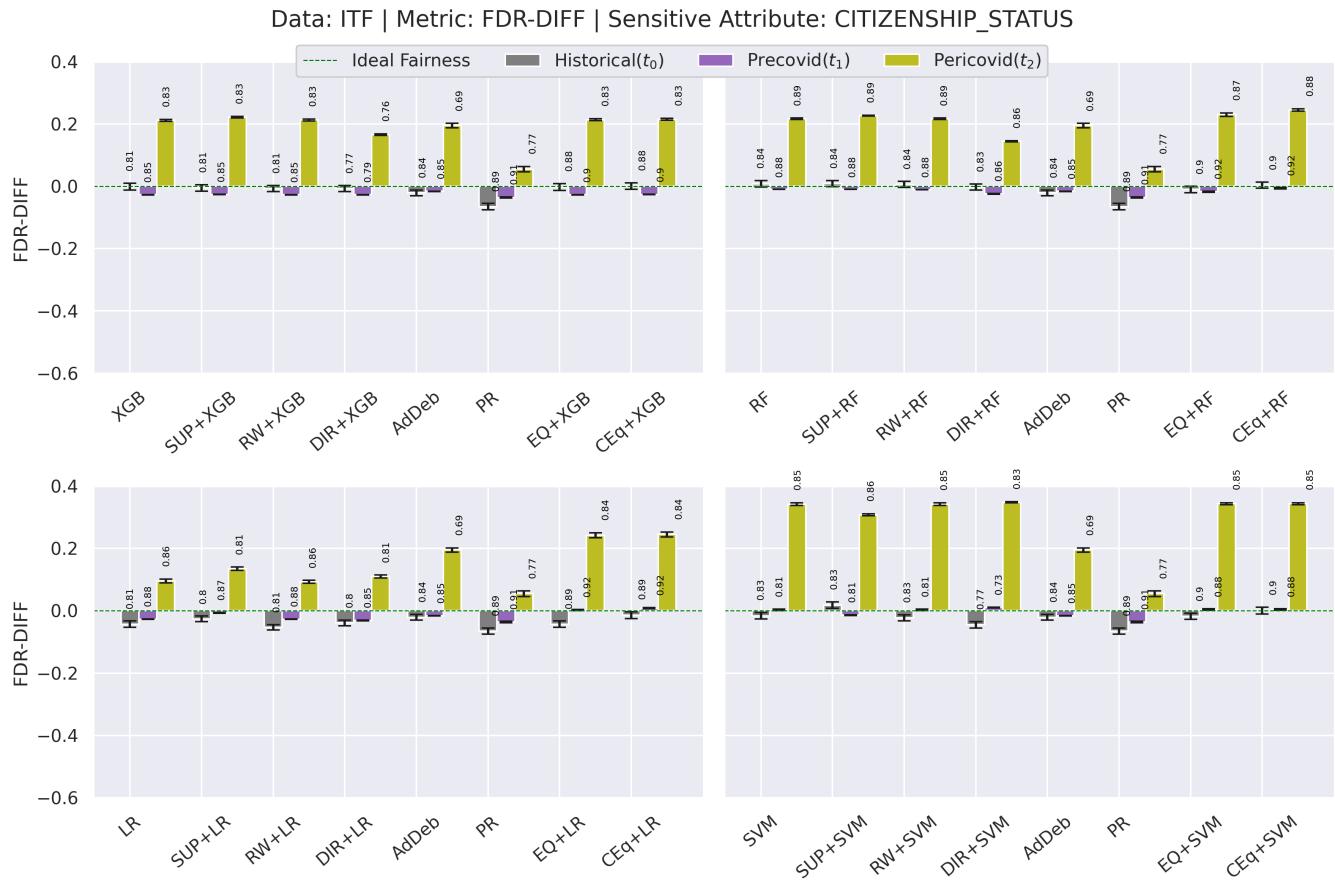


Figure 31: FDR-DIFF for ITF dataset.

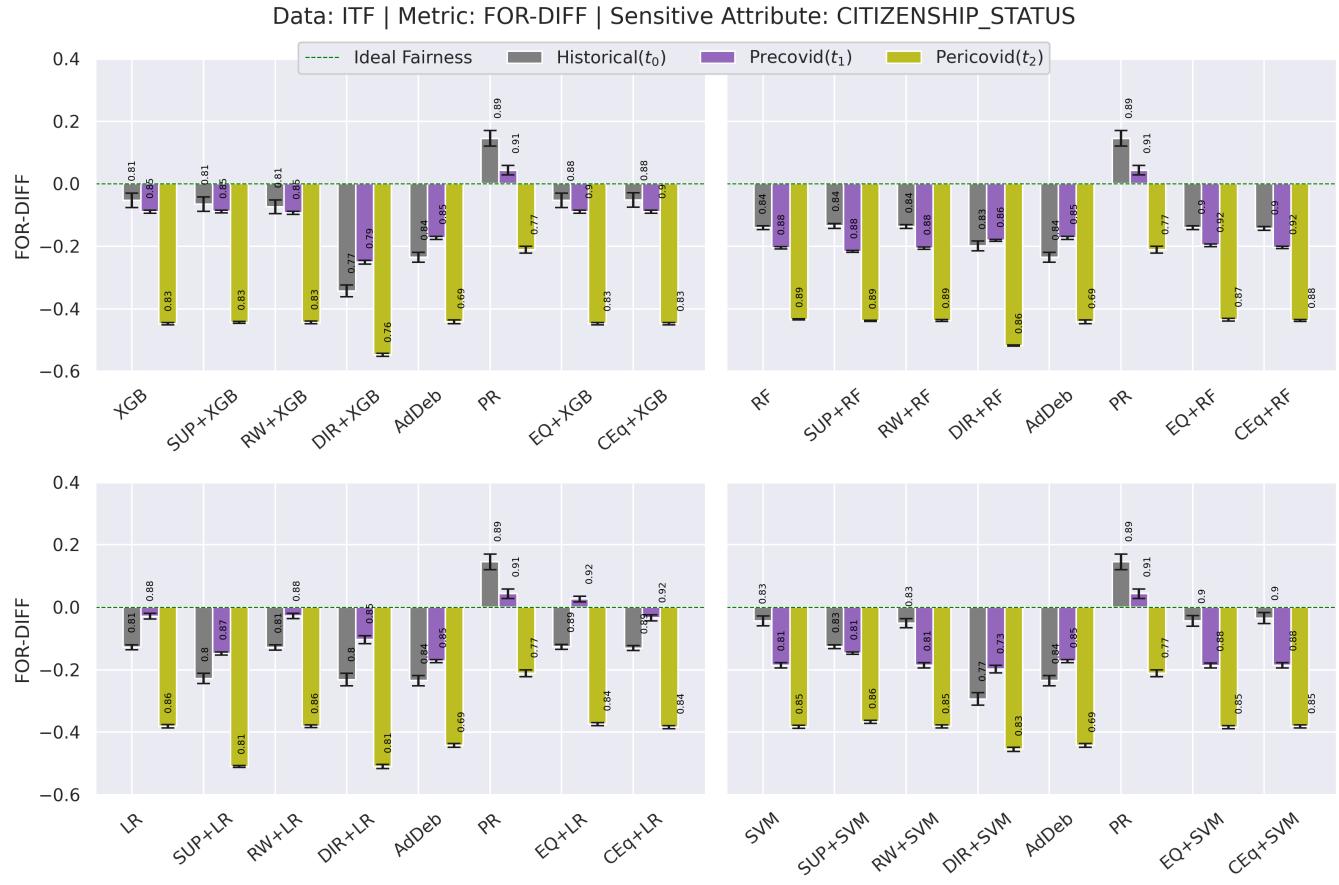


Figure 32: FOR-DIFF for ITF dataset.

Data: ITF | Metric: IF | Sensitive Attribute: CITIZENSHIP_STATUS

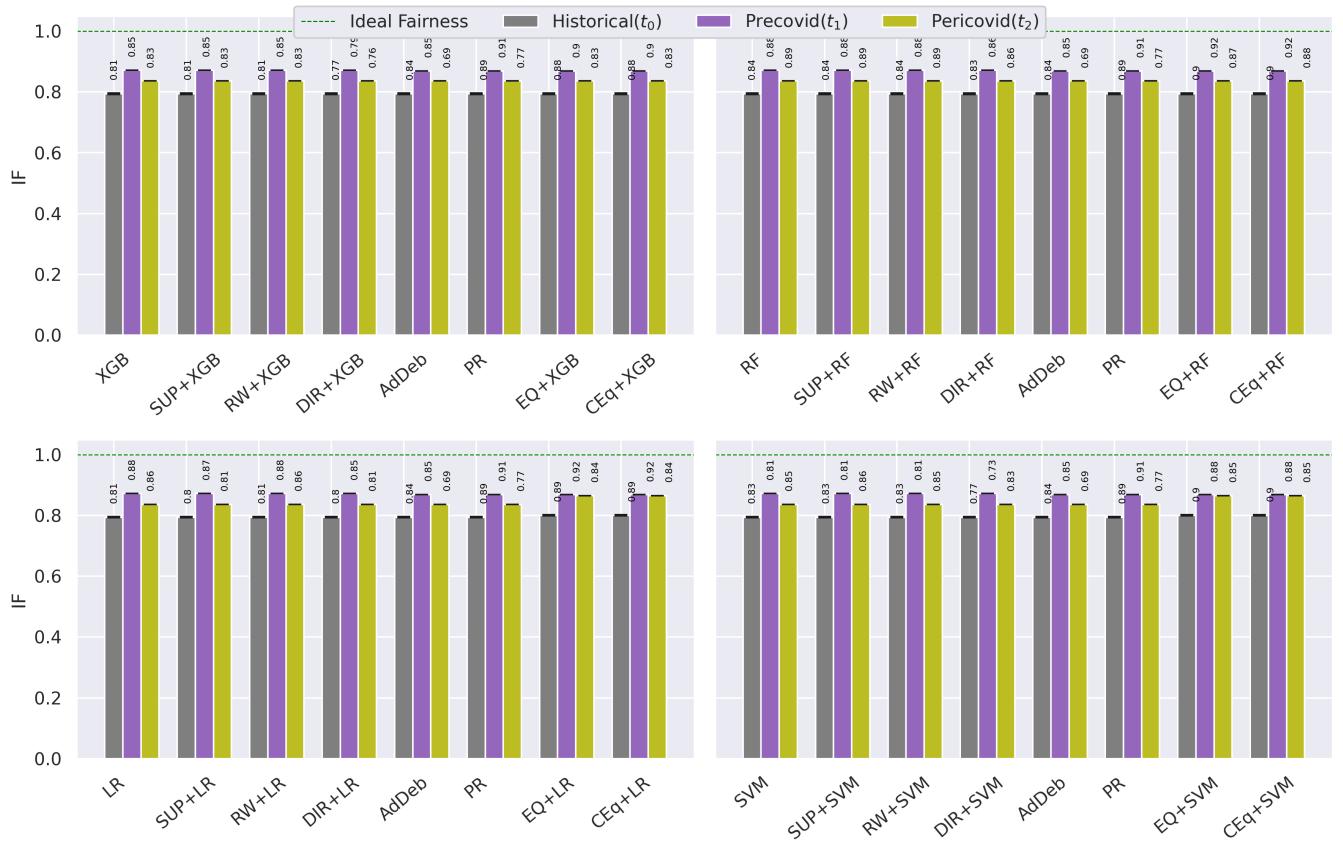


Figure 33: Consistency for ITF dataset.

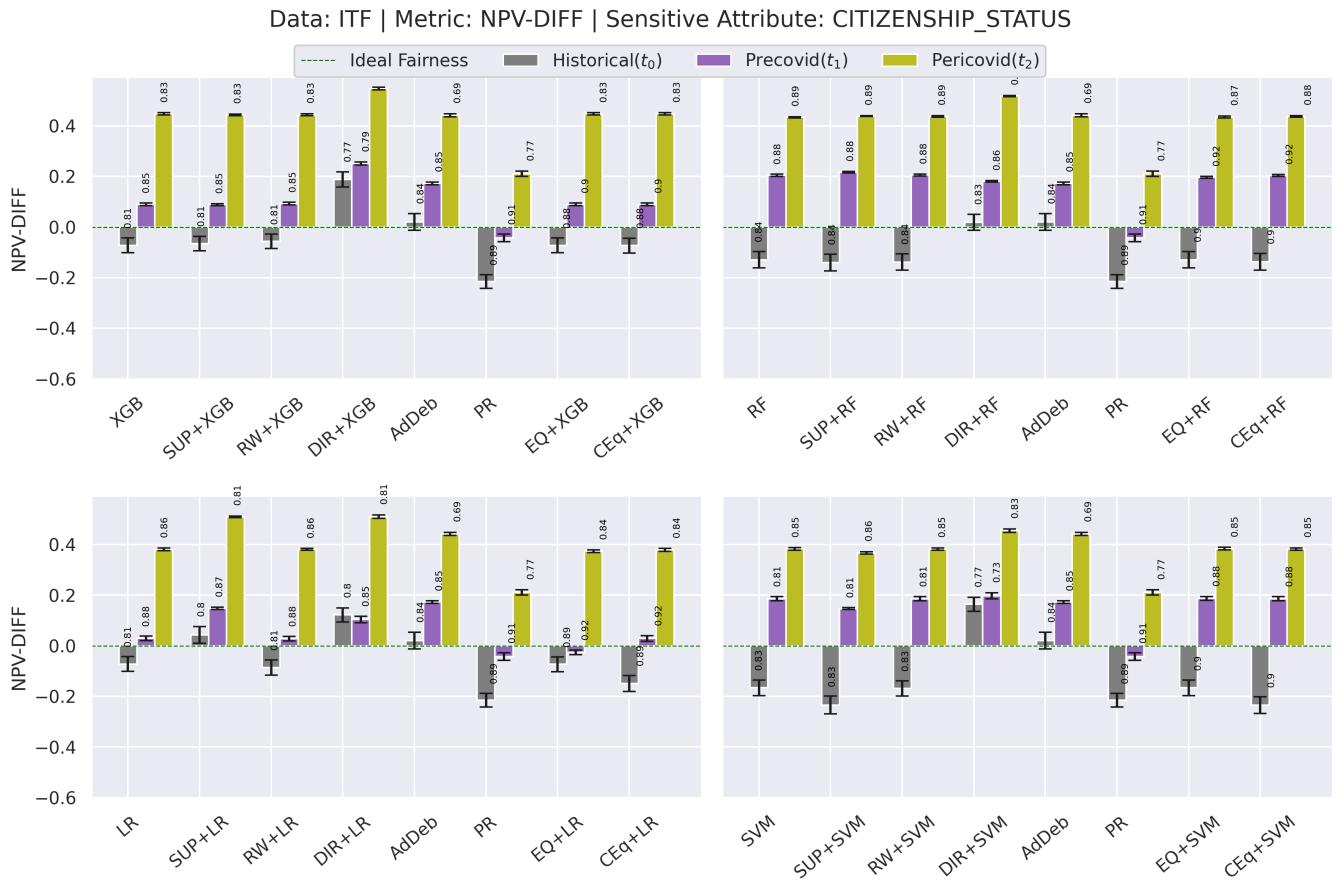


Figure 34: NPV-DIFF for ITF dataset.

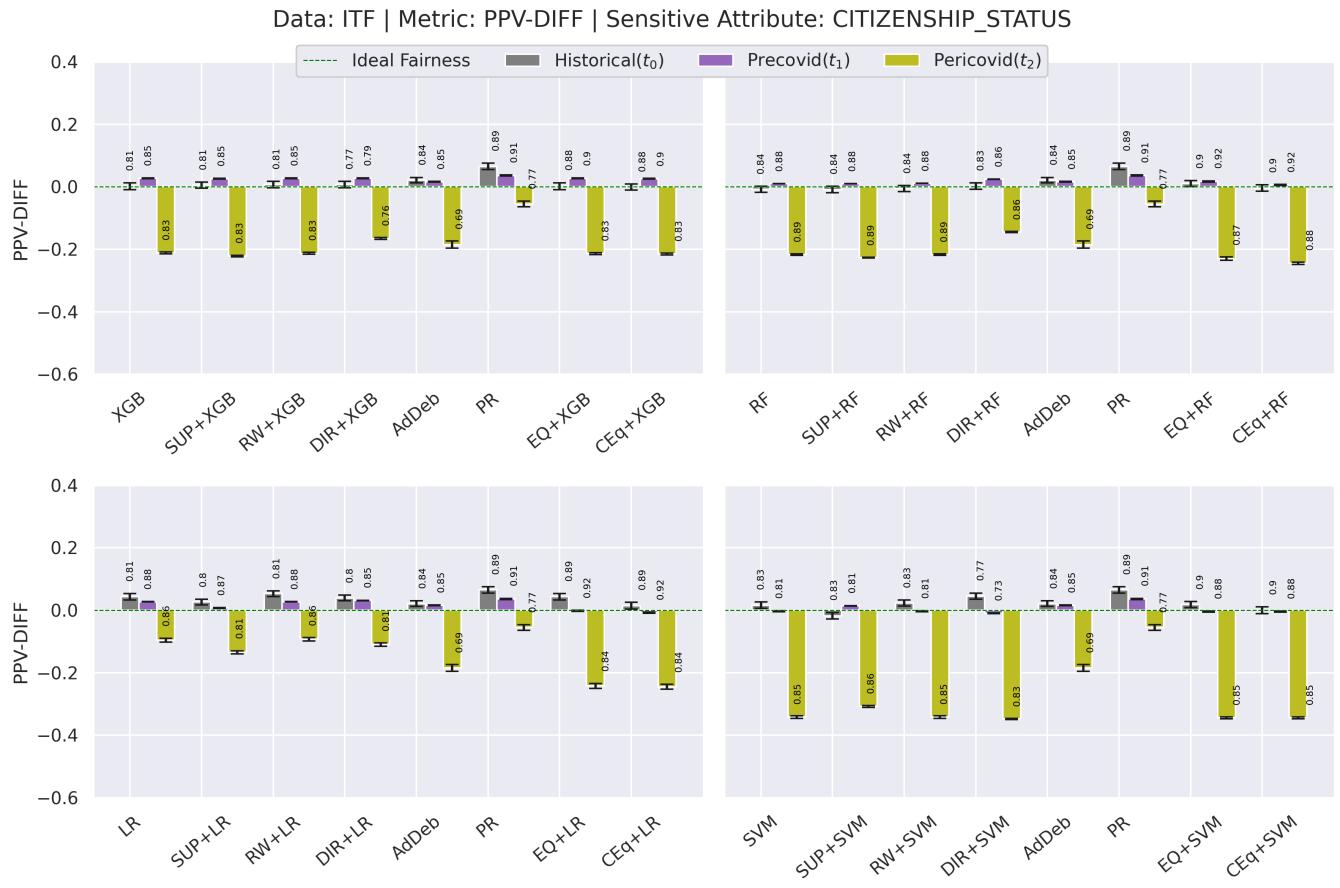


Figure 35: PPV-DIFF for ITF dataset.



Figure 36: WGEI for ITF dataset.

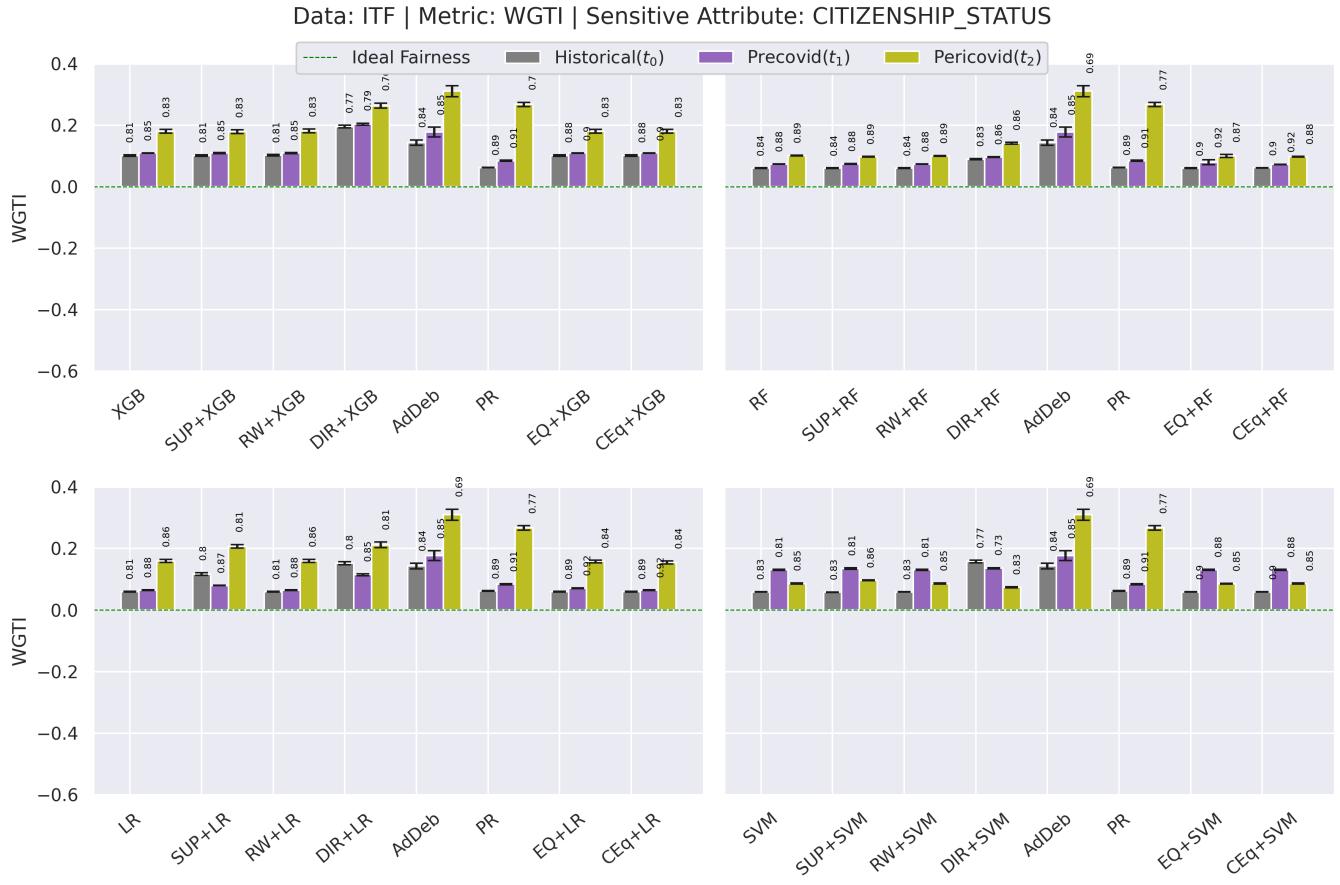


Figure 37: WGTI for ITF dataset.