We thank the reviewers for their thoughtful and useful feedback. Here, we address the main concerns raised by the reviewers. We will also fix minor typos in the final version of the paper.

Adding comparison with other uncoupled regression: Since Reviewers 1 and 4 share the concern on the SVMRank baseline in the experiments, we answer it here. Strictly speaking, there is no uncoupled regression method that can be used in our setting to the best of our knowledge. As we stated in the paper, other uncoupled regression methods require external contextual data or strong assumptions that do not hold for the benchmark datasets that we used in our experiments. However, we may obtain some linear model by using the methods discussed in Hsu et al. [1] or Pananjady et al. [2], though their linear assumptions on the target function does not hold in our settings. We will add empirical comparison with these methods in the final version.

To Reviewer 1

The discussion about the comparison between the TT and RA approaches in Figure 1: Thank you for pointing it out. As you suspect, the discussion should have gone for Figure 2. We will fix it in the final version.

The error bars seem strange: It is because we used log-plot in figures. Since the standard deviation of the TT method is large, the mean accuracy minus the standard deviation goes negative, which yields the strange shape in the log-plot. (Note that log of negative is treated as minus infinity.)

To Reviewer 3

Practical use of algorithm: To the best of our knowledge, there is no uncoupled regression method can be used in our setting. Hence, we compare our methods to the SVMRank benchmark, which is the closest to our setting. To make such a comparison meaningful, we decided to focus on the linear kernel. We would try different kernels in SVMRank or more complex models in the final version.

Computational requirements: The loss function in the RA approach is just the sum of empirical means of the loss function, which does not take extra computation compared to ordinal empirical risk estimation methods. Moreover, by using the approximation described in Appendix A.1, the TT approach can be computed similarly to the logistic regression.

References

[1] D. J. Hsu, K. Shi, and X. Sun. Linear regression without correspondence. In *Proceedings of the 30th Advances in Neural Information Processing Systems*, pages 1531–1540, 2017.

[2] A. Pananjady, M. J. Wainwright, and T. A. Courtade. Linear regression with shuffled data: Statistical and computational limits of permutation recovery. *IEEE Transactions on Information Theory*, 64(5):3286–3300, 2018.