

Data Description Sheet for
“When is it Fairer to Allocate Resources for Training
to Lower-Performing versus Higher-Performing Employees?”

Eric W. Chan, *The University of Texas at Austin*

Mackenzie K. Feinberg, *Emory University*

Martin Wiernsperger, *Cornell University*

1. Data Handling and Analyses

Data handling and analyses were conducted by Eric Chan, Kenzie Feinberg, and Martin Wiernsperger.

2. Data Generation

See the article for a detailed description of the procedures used to generate the data. The main experiment was programmed using Lioness and housed on a private server. The experiment was run between 04/12/2024 and 04/19/2024 with undergraduate and graduate business school students from the McCombs Behavioral Lab participant pool at The University of Texas at Austin. The full data was downloaded from the private server on 04/22/2024.

The additional lab experiment was programmed using Lioness and housed on the same private server. For this additional lab experiment, we recruited participants from two different universities. The first run of the additional experiment took place between 03/06/2026 and 03/09/2026 with undergraduate and graduate business school students from the McCombs Behavioral Lab participant pool at The University of Texas at Austin. The data was downloaded from the private server on 3/10/2026. The second run of the additional lab experiment took place between 04/08/2026 to 04/10/2026 with undergraduate and graduate business school students from the participant pool of the Lab for Experimental Economics & Decision Research at Cornell University. The data was downloaded from the private server on 04/10/2026.

The supplemental experiment was run on 05/06/2026 with online Prolific participants. The full data was downloaded from the private server on 05/06/2026.

The main lab experimental data was collected by Eric Chan and Kenzie Feinberg. The additional lab experiment data was collected by Eric Chan and Kenzie Feinberg at UT Austin, and Martin Wiernsperger at Cornell University. The online experimental data from Prolific was collected by Martin Wiernsperger. The co-author team can vouch for the stated source of the raw data. The entire co-author team has access to the raw data.

3. Proprietary Nature of Data

Not applicable. Data is available from the authors upon request.

4. Data Collection and Processing

See Sections 3 and 4 of the article for a detailed description of the steps to collect and process the data. The full experimental instruments are available in the data replication package that can be assessed on the JAR Online Supplements website.

5. Raw Data Processing

All processing of the experimental data used in the reported analyses is performed through the STATA run code. The run code imports the raw experimental data, applies the sample inclusion and exclusion criteria described in the article, constructs and re-codes the variables used in the analyses, and creates the final analysis samples for each experiment. No manual changes to the analysis data are required outside of the provided code.

6. Provision of Code

The complete STATA run code is included in the data replication package available on the JAR Online Supplements website. The code is organized to reproduce the analyses for each experiment and includes the commands used to prepare the final analysis data, execute the statistical analyses reported in the article, and produce the output underlying the reported tables and results. The code is accompanied by the experimental instruments described above.

7. Comprehensive Log File

The data replication package includes comprehensive STATA log files showing execution of the run code from data processing through the statistical analyses and reported output. Because this study involves human participants rather than archival firm-level data, separate participant-level identifiers are not provided to protect participant privacy.

8. Assurance of Data Maintenance

We will ensure that at least one coauthor will maintain the data and programs according to the guidelines.