REPLICATION APPENDIX FOR

Mentoring, Educational Preferences and Career Choice: Evidence from Two Field Experiments in Bhutan

Ryotaro Hayashi
Asian Development Bank

Hyuncheol Bryant Kim
Yonsei University and Hong Kong University of Science and Technology

Norihiko Matsuda
Florida International University

Trinh Pham
KDI School of Public Policy and Management

INTRODUCTION

This appendix outlines the data and code necessary to replicate the study "Mentoring, Educational Preferences and Career Choice: Evidence from Two Field Experiments in Bhutan," authored by Hayashi, Kim, Matsuda and Pham (2025), published in the *Journal of Policy Analysis and Management*.

The replication package contains the dofiles that reproduce the results presented in the paper. Due to the use of confidential data in this study, the data cannot be made available in the replication package. However, the clean anonymized data that support the findings of this study are available from the corresponding author upon request. Access to the raw data is subject to approval by the Asian Development Bank (ADB) and the Bhutanese Ministry of Education and Skills Development, due to privacy and ethical restrictions.

DATA SETS

In this paper, we draw from eight sets of primary and administrative databases:

1. The baseline surveys of grade 9/10 and grade 11/12 students.
2. The follow-up surveys of grade 10 and grade 12 students.
3. The administrative database on list of students in Grade 11.
4. The administrative database on TVET trainees, exam test scores, student repeaters.
5. The baseline survey attendance sheets provided by local schools. This dataset is to identify relevant grade 10 and grade 12 in our final study sample.
6. The surveys containing mentor characteristics.
7. The surveys containing mentor report after each meeting session.
8. The Bhutanese administrative shapefiles

   The files are made available at https://data.humdata.org/dataset/cod-ab-btn. A copy of these files is provided in the replication package.

FILE STRUCTURE

The replication package is divided between three folders, structured as follows.

data: there are three subfolders.

- **raw:** this subfolder contains six smaller subfolders.
  - The subfolder **student** includes raw datasets from source (1), (2), and (3), as described above
  - The subfolder **confidential** includes raw datasets from source (4)
  - The subfolder **manual** includes raw datasets from source (5)
  - The subfolder **mentor** includes raw datasets from source (6)
  - The subfolder **report** includes raw datasets from source (7)
  - The subfolder **map** includes shapefiles of Bhutan, from source (8)
- **temp:** any temporary datasets that are generated during the data construction process will be temporarily saved in this subfolder and deleted afterwards
- **clean:** this subfolder contains the anonymized datasets generated after running the codes and they are ready for final analysis. These final datasets will be provided from the corresponding author upon request.
  - comparison_all_grade910_baseline.dta
  - comparison_grade910_baseline_reachable_wgrade.dta
  - comparison_all_grade1112_baseline.dta
  - comparison_grade1112_baseline_reachable_wgrade.dta
  - STEM_mentor.dta
  - TVET_mentor.dta
  - mentor_report_discussion.dta
  - grade10_analysis.dta
  - grade12_analysis.dta

code: there is one master do file that replicates all figures and tables in the text, using the aforementioned datasets

- 0.0.master_dofile.do

output: all figures and tables generated by the do file are stored within this folder

RUNNING THE CODE

To replicate the results, users should maintain the directory structure provided in the replication package and follow these steps:

1. **Obtain either raw or clean data**: Users must acquire either the raw data files from sources (1)-(7), or the clean data files specified above and place them in their respective subfolders within the replication package.
2. **Install required packages**: Ensure all Stata packages listed at the top of the 0.0.master_dofile.do file are installed.
3. **Run the code**: After completing the above steps, execute the 0.0.master_dofile.do file. The generated tables (.tex files) and figures (png) will be saved in the **Output** folder.

Technical Specifications:
The code was executed using Stata MP version 17 on a 64-bit Windows machine with an Intel® Core™ i5-13500 CPU (2.5 GHz) and 64 GB of RAM. The runtime was approximately 2 hours.