Read Me file for JPE Replication Package — The Effect of Gender Discrimination on Labor Supply

Nickolas Gagnon, Kristof Bosmans, Arno Riedl

July 2024

Summary. The package contains the following files: (i) datasets for Studies 1–6, (ii) statistical analysis codes for the article and for the online appendices, and (iii) original instructions and experiment/survey codes for Studies 1–6[1].

Software and formats. We provide Stata codes for statistical analyses. We use Stata SE 18 and assume the code and data files are in the same directory and the paths of all .do files are changed before any is run since several .do files run another .do file as part of their code. For Stata, you need to install the $st0381_1$ package for Dunn's tests and the *estout* package for some tables. We provide Qualtrics codes for conducting the studies. In terms of formats, analysis codes are in .do, datasets are in .dta, instructions are in .docx, and codes for conducting the studies are in .qsv.

Analysis codes. There are 13 analysis codes (13 separate .do files). There are six .do files for preparing data, one per study (GBR24CleaningStudy1, GBR24CleaningStudy2, ... GBR24CleaningStudy6). For the article, there are four .do files to analyze data (GBR24AnalysisStudies1-2-3, GBR24AnalysisStudy4, GBR24AnalysisStudy5, and GBR24AdditionalInfoWomExperienceDiscrimination[2]). For the appendices, there are three .do files to analyze data (GBR24AppendicesABFiguresTables, GBR24AppendixCParameters, and GBR24AppendixDAnalysisStudy6). In terms of ***running order***, the first set of .do files that can be run are the six .do files for preparing data (GBR24CleaningStudy1,...), but running them is not necessary since the other .do files automatically run them whenever they are required. Then, the four .do files to analyze data for the article (GBR24AnalysisStudies1-2-3, GBR24AnalysisStudy4, GBR24AnalysisStudy5, and GBR24AdditionalInfoWomExperienceDiscrimination) can be run separately and will automatically run the necessary .do files for preparing data. Finally, the three files for the appendices can be run separately, but will run the six .do files for preparing data and the four .do files for the main text whenever they are required. Note that because of dependencies between.do files, all (loading and saving) paths should be changed to the same directory before running .do files.

Anonymized datasets. We provide six datasets, one per study (GBR24DataStudy1, GBR24DataStudy2, ... GBR24DataStudy6), after pre-cleaning them for public use. The six .do files for data cleaning are

[1] We generally only provide the code/instructions for men or women instead of both as these are identical except for gender.

[2] This .do file is to analyze experiences of gender discrimination by women across our studies and a survey of women we used to recruit the participants in Studies 4 and 5.

then used to create six additional .dta files (GBR24DataCleanStudy1.dta, GBR24DataCleanStudy2.dta, ... GBR24DataCleanStudy6.dta). We also provide a dataset containing the discrimination experience of women in the initial survey we used to recruit women for Studies 4 and 5 (GBR24CleanDataInitialSurveyStudies4-5). In the pre-cleaning, we do not remove variables used in data analyses. We removed variables with potential for personal identification, such as Prolific IDs (PIDs), participation time (replaced with period indicators), and questions with free-form answers. We also removed variables automatically created by Qualtrics with no use for analysis. We used PIDs to merge Qualtrics answers and socio-demographic information provided by Prolific, and to remove invalid observations[3]. We also removed the handful of participants whose self-reported gender in the experiment did not match the gender they reported to Prolific. Finally, we removed the data for the Study 1 schemes that are not part of the article (see schemes in Gagnon, Bosmans, and Riedl, SSRN Working Paper 3519540, 2020).

[3]Software issues lead to more than one observation from the same participant for $< 2\%$ of participants (concentrated in Study 1). These were typically two participations, with one participation stopping around the consent form or the instructions and the other participation passing the instructions and the comprehension questions and going through to the allocation to a payment scheme (i.e., allocation to a treatment), so that only one participation was allocated to a treatment. We kept the participation going to the treatment allocation for the typical case where only one participation reached a treatment. For the rare cases where participants reached the treatment allocation twice, we kept the first participation (and for three participants who reached the treatment allocation twice at around the same time, we deleted both participations).