

Readme

This file describes the replication package for Bruhn, Miriam and Caio Piza. 2022. "Missing Information: Why Don't More Firms Participate in Business Training?"

For this study, we collaborated with the Parana office of the Brazilian Micro and Small Business Support Service (SEBRAE-PR), a non-profit that provides a range of free business training services to small firms. SEBRAE-PR shared with us data on business practices they collected during their Business-to-Business (NaN, short for *Negocio-a-Negocio*) program in 2018. SEBRAE-PR also shared administrative data on the use of SEBRAE services. The study also uses data from the Relação Anual de Informações Sociais (RAIS) database from the Ministry of Labor. All of the data used is restricted-access and cannot be shared by the authors of the study, but it could be requested directly to SEBRAE-PR and the Ministry of Labor of Brazil.

Folder structure

The parent folder includes three subfolders.

1. Data: This folder should includes three files.
 - a. Baseline_NaN_RAIS.dta includes baseline business practices from SEBRAE-PR data, as well as RAIS data for 2015. RAIS is a dataset on employees of registered firms from the Brazilian Ministry of Labor and Employment.
 - b. Followup_Survey_RAIS.dta includes data from the follow-up survey, RAIS data for 2019 and RFB data for 2019. We use data from the Federal Tax Authority (RFB) to measure if a firm has officially closed.
 - c. SEBRAE_Services.dta includes daily administrative data from SEBRAE-PR on use of SEBRAE services. The do file Analysis_final.do, described below, codes outcome variables from this data and converts it to a monthly frequency.
2. Do Files: This folder includes Analysis_final.do. This code uses the three data files listed above to reproduce figure 1, partially reproduce table 1, and reproduce tables 2 through 4 in the main section of the paper.
3. Results: This folder contains the output files generated from running Analysis_final.do. These are the files generated when running the code that correspond to the exhibits in the main section of the paper:
 - a. Table_1.xlsx contains the summary statistics shown in Table 1 of the paper, except for the variables Average employee age (2015), Share of male employees (2015), Share of employees with college or higher (2015), Commerce sector, Service sector, and Starting capital (USD), which were added with a later version of the code not included in this reproducibility package
 - b. tables_2_3_4_5.xml contains the results shown in Table 2 of the paper (tab d_service_c) and Table 3 (tab Followup_Survey_RAIS)
 - c. Ftest_d_service_c.xls contains F-test p-values and control group means for table 2.
 - d. Ftest_d_material_c.xls contains F-test p-values and control group means for table 3.
 - e. Table_6.xlsx contains the results shown in Table 4 of the paper

Other outputs generated correspond to exhibits in the Appendix or to additional outputs not included in the paper.