Notes regarding final replication.

Cleaning

1.a- Pilot

cleaning_pilot_rep.do - Main cleaning do file for pilot data and its surveys. Runs *name_cleaning_pilot_rep.do* for name cleaning.

1.b- Scaleup

cleaning_scaleup.do - Main cleanin do file for scaleup, surveys, and 2018 followup. Runs *name_cleaning_scaleup.do* for name cleaning.

1.c- Historical data

cleaning_hd_rep.do - Main cleaning do file for hd. Runs *name_cleaning_hd_rep.do* for name cleaning.

1.d- Time preferences

mxfls_cleaning.do appends and cleans answers to MxFLS data.

DB_time_pref.do Constructs a comparable database of frequencies of answers between MxFLS survey and P1 surveys.

1.e-P1 & P2 Followups

importAppend2020Followups.do Imports and cleans outcome data for both P1 and P2

cleaning_missingCasefiles.do Cleans initial casefilings that were recovered during 2019 - 2020

clean_missing_predictions.do cleans calculator predictions that were originally missing

cleanTerminaciones_v2.do cleans P1 & P2 2020 followups

1.f- Pilot 3

cleanPilot3.do cleans P3 database for followup

CleaningP3.do cleans survey (2m & 2w) data, as well as treatment data. It runs several auxiliay do files: *cleaning_2m_survey.do*, *cleaning_2w_survey.do*

1.g- ITT sample for P1 & P2

ITTSample.do selects casefiles that entered the ITT sample for the calculator experimental arm.

Main Results

Tables

T1: Pilots description

pilotsDescription.do Gets data for P1 & P2. May consider redoing the do file / table to include P3. Igual para incluir fechas etc.

T2: Summary statistics

SS_replicates.do Somehow the table got lost and this one does not replicate the one in the overleaf (we have more obs) I'll check the rest of the results to make sure this table is correct.

T3: main regression

ITTVersion.do This creates P1 & P2 columns except for CF
treatment_effects_IV_CF_noConciliator_rep_NoDuplicatesITT.do Performs the CF regression under different assumptions (presented also in A.C)
p3ResultMainTable.do Performs the regression with P3 outcomes

These results get merged in the excel file *TreatmentEffectsOnSettlement_main.xlsx*

T4: TE by lawyer type

teByLawyer.do Performs separate regressions for private and public lawyers. Result is formated in the te_bylawyer.xlsx file.

T5 Expectation updating

te_p3_updating.do Does the regression. Still misssing a xlsx to provide format.

T6: Outcomes by treatments

endMode.do Creates the table in a preformated xlsx

T7: Recovery at 42 months (1 & 2)

welfare42monthsP1P2.do Slight changes in the results.

T8: Welfare effects P3

P3welfare.do Missing an excel table that formats this.

Figures

F1: Claimed vs compensation

Fig1_amount_plots_rep.R (ahorita la checo)

F2: Time duration HD

caseending_overtime_rep.do

F3: Knowledge of the law and lawsuit

knowledge_allByType.do

F4: Treatment format

No do file

F5: Calculator prediction for settled casefiles

calculatorDistributionForSettled.do

F6: Calculator distribution for cases that won in court

caseValueKdensitiesdo.do does it. It also does it disaggregated by phase.

Appendix A

Tables

A1: Casefile variables Description

No do file
histograms_npv_hd.do <- quien sabe qué es esto

A2: Survey SS For P1.

surveyP1SS.do

A3: Experiment integrity

complianceTable.do does panel A using only P1 data
ComplianceP2Tables.do does panel B using only P2 data
showUp.do does panel C using P1 & P2 data

A4: Balance table

balance.do Me faltan observaciones del P1. Supongo que vienen de las neuvas iniciales? Está raro

A5: Fit assessment of the calculators (discrete)

calc_predictions.R (NO SE HA CORRIDO)

A6: Fit assessment of the calculators (continoous)

calc_predictions.R (NO SE HA CORRIDO)

Figures

FA1: Covariate distribution comparison HS vs P1

covariate_plots.R and *covariate_plots_hd.R*

FA3: Covariate distribution comparison HS vs P2

v_covariate_plots.R_ and *covariate_plots_hd.R*

Appendix C

Tables

TC.1 .- Amount asked (log), amount won (log), and probability of winning - Historic Data

reg_amount.do Output gets aggregated in the *TableReg1_log.xlsx* file.

TC.2.- Balance of casefiles having negative recovery amount.

negReturnHD.do. Output gets written in *negative_returners_balance.xlsx*

TC.3.- Balance regression on characteristics conditional on employee present

ep_balance.do. Output gets formatted in *ep_balance.xlsx*

TC.4.- Employee presence

pEmployeePresence.do.

TC.5.- Expectations Relative to Prediction

prediction_cases_pooled_MergeWithPilotOperation.do does Panel A w/P1 & P2
ExpectationsFromP3forPrediction_cases_pooles_panelB.do does Panel B with P3
Results are merged and formatted in table *prediction_cases_pooled.xlsx*

TC.6.- First stage and robustness for the control function regression

treatment_effects_IV_CF.do. Problem with column 1, gotta get back to it.

TC.7.- Heterogeneity in treatment effects

te_heterogeneity.do produces raw output, *te_heterogeneity.xlsx* formats it.

TC.8.- Treatment Effects with placebo arm - Phase 1

tePlacebos.do

TC.9.- Updating - Phase 1

update_reg_theta_rel_uc.do performs the regression with underconfidents, while *update_reg_theta_rel_oc.do* does it for over confidents. The table is actually composed of 2 different tables, both of which exist in the excel workbook *update_reg_theta_rel.xlsx*.

TC.10 - Duration of Cases by Treatment

welfare42monthsP1P2.do does, among other things, this regression. *duration.xlsx* adds format to the results.

Figures

F1.C.- Stylized Depiction of the Labor Justice Process

F2.C.-Calculator Treatment Format (example) - Phase 2

F3.C.- Calculator Treatment Format (example) - Phase 3

F4.C.- Distribution of Amount Collected, by Type of Lawyer

cdf_value_claims.do does both subgraphs and appends them

F5.C.-

oc_comparison_B.do does 6 subfigures, they get appended directly on latex.

F6.C.- Settlement Amount vs. Calculator

ratioGananciaConvenio.do

F7.C.- Outcomes when Plaintiff was Present, by Treatment

welfare42monthsp1p2.do does the 2 subfigures, latex compiled them as one

F8.C.- Calculator Predictions for Plaintiff Court Judgment, By Phase

caseValueKdensitiesdo.do

##Figures