Replication code for Centralizing Over-The-Counter Markets?

By Jason Allen and Milena Wittwer, 03/2023

Remark: The data is proprietary.

DATA CREATION AND CLEANING

mainfile_datacreation.do

reads in all the other do.files
INPUT: many input files with raw data
OUTPUT: data_base_bonds_bills.dta

public_auctions.do

brings in the public auction information from
https://www.bankofcanada.ca/markets/government-securities-auctions/

cleaning.do

data cleaning, including isin names and settlement dates

create_variables.do

creates variables, including counterpart-type and client-type, institutional versus retail

bloomberg_candeal.do

merges in quote data from bloomberg (bloomberg_2020, bloomberg_2020_2, bloomberg_2020_3)
and candeal (avg_hourly_candeal_bonds_bills)
runs: **bloomberg_duplicates.do** to eliminate duplicates

double_counting.do

expands the inter-dealer market so that each trade has a leg. This is marked with the variable Dleg. Dleg=0 is only the original data set

data_yas_v2.do and data_yas_bl_v2.do and data_yas_can.do

in-sheets the yields in YAS from .cvs files

DESCRIPTIVE EVIDENCE - TABLES AND FIGURES

DescriptiveEvidence.do

- provides the descriptive evidence, including figures and tables, in the draft (everything that isn't related to the estimated model or counterfactual)

EventStudy.do

- runs the events study of investors who lose platform access

Homedealer.do
- assigns a homedealer to all investors

Figure3.do
- creates Figure 3 using quote data from CanDeal

ESTIMATION OF THE STRUCTURAL MODEL

Mainfile_ModelEstimation_STATA.do

need to choose specification: homedealer (0 or 1) and week (-1,0, 1) and largetrade (0 or 1)
- estimates sigma and backs out xij
- outsheets the data that is used in the Matlab code for the second step of the estimation

BaseData.do

uses *data_base_bonds_bills.dta* created in mainfile_datacreation.do and restricts the sample as described in the paper
- eliminates outliers (upper and lower 1% of distribution)
- fills in missing quotes
- keeps only hours where CanDeal is open
- normalizes all yields

creates *data_base_bills_clients_matlab_2022b.dta*, which is the main data sample used throughout the paper with dealer-to-investor trades (only including in-house trades)

ModelData.do

uses *data_base_bonds_bills.dta* to add in auction data, *data_base_bills_clients_matlab_2022b.dta* and *data_auction_expectations.dta* (with the supply shock estimates)
- excludes in-house trades, data after July 30 2019, by non-primary dealers, drop one dealer due to misreporting + some extra cleaning (empty yields, 2 holidays)
- creates useful variables for the regressions

Marketshares.do
- Is used in ModelData.do to create market shares

Regressions.do, RegressionHomedealer.do, RegressionHomedealer_SL.do
- each runs the IV market share regression for the respective model specification: model without loyalty benefit, with loyalty benefit, with loyalty benefit and different trade sizes

outsheets *"MatlabEstimationData_$week _$homedealer _$keepall _$quantityrobust _$largetrade _$fej.csv"* for Matlab Code

Instrument_do

- Combines the auction estimates, obtained by running the code from Allen, Kastl, and Wittwer (2022) with the full data
- Output: *data_auction_expecations.dta* to be used in the estimation of the main model