Readme File

Replication material for "A Quantitative Model of Banking Industry Dynamics"

by Dean Corbae and Pablo D'Erasmo

The accompanying replication.zip file contains two folders.

1. "Data"
   - This folder contains the programs (Stata and Matlab) and available data used for the analysis.
2. "Model"
   - This folder contains the Fortran codes used to compute the model as well the Matlab codes and model output to create Figures and Tables.

Data Availability and Provenance Statements

We use data from four sources:

1. *Consolidated Report of Condition and Income (known as Call Reports)* that banks submit to the Federal Reserve each quarter. Report of Condition and Income data are available for all banks regulated by the Federal Reserve System, Federal Deposit Insurance Corporation FDIC, and the Comptroller of the Currency. All financial data is on an individual bank basis and we consolidate the data at the Bank Holding Company Level.

   Data is subject to distribution restrictions. We access the data through the Federal Reserve Board of Governors. A public version of the data can be accessed through the Federal Financial Institutions Examination Council Central Data Repository's Public Data Distribution https://cdr.ffiec.gov/public/ManageFacsimiles.aspx (link last accessed July 20, 2025) or the Federal Reserve Bank of Chicago https://www.chicagofed.org/banking/financial-institution-reports/commercial-bank-data (link last accessed July 20, 2025).

   The list of all variables used and definitions can be found in the online Appendix of the paper as well as in the file "data_appendix.pdf" included with the replication materials (folder "\Data").

2. Federal Financial Institutions Examination Council, National Information Center (NIC): Relationship data that describes ownership between bank holding companies and Transformation table that contains details on mergers and acquisitions https://www.ffiec.gov/npw/FinancialReport/DataDownload (link last accessed April 20, 2025).
3. *Federal Reserve Bank of St Louis Economic Data (FRED)*

   We retrieve the data from FRED, Federal Reserve Bank of St Luis: https://fred.stlouisfed.org/ We use the following variables

- Consumer Price Index from the U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCSL] retrieved from https://fred.stlouisfed.org/series/CPIAUCSL
- Real gross domestic product [GDPC1] from the U.S. Bureau of Economic Analysis retrieved from at https://fred.stlouisfed.org/series/GDPC1

- Nominal gross domestic product from the U.S. Bureau of Economic Analysis retrieved from https://fred.stlouisfed.org/series/GDP
- Loans and Leases in Bank Credit, All Commercial Banks (LOANS) from the Board of Governors of the Federal Reserve System (US) retrieved from https://fred.stlouisfed.org/series/LOANS
- Deposits, All Commercial Banks (DPSACBM027NBOG) from the Board of Governors of the Federal Reserve System (US) retrieved from https://fred.stlouisfed.org/series/DPSACBM027NBOG

Dataset list

Data File	Source	Notes	Provided
\Data\call_reports_q_01.dta	Federal Reserve Board of Governors	Call Reports data appendix_data.pdf presents the list of variables used	No
\Data\topholder_table_q_01.dta	Federal Reserve Board of Governors	Table that identifies the top holder of banks in the sample	No
\Data\CSV_TRANSFORMATIONS.csv	National Information Center (NIC)	Data on commercial bank failures, mergers, and acquisitions	Yes
\Data\CPIAUCSL.csv	FRED	CPI quarterly frequency (end of period)	Yes
\Data\GDPC1csv	FRED	Real GDP	Yes
\Data\GDPDEF.csv	FRED	GDP Deflator	Yes
\Data\GDP	FRED	Nominal GDP	Yes
\Data\LOANINV	FRED	Total Loans	Yes
\Data\DPSACBM027SBOG	FRED	Deposits	Yes
\Data\mydata_bhcb_allq.dta	Federal Reserve Board of Governors FRED NIC	Derived sample. Combines multiple sources, provides the base for all the moments and analysis at bank/bank holding company level	No
\Data\data_concentration_dep_national_level.mat	Derived sample (bank sample)	Measures of concentration	Yes
\Data\ sod_1994-2023.dta	Summary of Deposits (SOD)	Sample with deposits at branch level (yearly frequency)	No
\Data\ data_hhi_dep_state_level.mat	Derived sample (HHI state level)	HHI at state level	Yes

\Data\data_di_index_bysize.mat	Derived sample (DI index)	Diversification index by bank size	Yes
\Data\data_Hstat01.mat \Data\reg_Hstat.csv	Derived sample (Hstat)	H-stat by bank size and year	Yes
\Data\all_banks_y1988.mat \Data\all_banks_y2016.mat	Derived sample (deposit ranking)	Deposits and Deposit ranking for years 1988 and 2016	Yes
\Data\reg_adj_zipfs_assets_top3000.csv \Data\reg_adj_zipfs_loans_top3000.csv \Data\reg_adj_zipfs_dep_top3000.csv	Derived sample (bank sample)	Estimated power law coefficients	Yes
\Data\stats_markup_top35.mat \Data\stats_markup_dist.mat	Derived sample (markups)	Estimates of median markup (all banks) and average for Top 35	Yes
\Data\data_exit_rates01.mat	Derived sample (exit rates)	Estimates of exit and failure rates by quarter	Yes
\Model\data_reg_crisis01.mat	Model Output	Output files from the baseline model to estimate regressions in Table 7	Yes
\Model\data_baseline_granular01.mat	Model Output	Output files from the baseline model to estimate regressions in Table 9	Yes
\Model\data_spillovers01.mat	Model Output	Output files from the baseline model to create Figures 8, 9, 10, 11	Yes

Description of Programs

Data Programs (folder "\Data")

The following is a list of the Stata and Matlab codes we provide. These codes were used to compute all data moments and create the corresponding figures.

1. cpi_quarterly_aux01.do: takes CPI data from FRED (CPIAUCSL.csv), creates inflation index, and makes it into a Stata sample: quarterly_data_cpi.dta
2. det_rgdp_aux01.do: takes real GDP and GDP Deflator data from FRED (GDPC1.csv and GDPDEF.csv), creates series for log-real gdp as well as series for GDP deflator in Stata. Uses

tsfilter hp from Stata. Samples created are quarterly_data_real_gdp_6023.dta and quarterly_data_gdp_def.dta.

3. aux_quarterly01.do: creates and renames some variables from the sample of banks in call_reports_q_01.dta. Creates an intermediate sample mydata_allq.dta
4. aux_quarterly_top_holders.do: loads table with top holders of different types (topholder_rable_q_01.dta) and identifies the regulatory top holder for each bank. Creates an intermediate sample banks_with_topholders_qtr.dta
5. aux_quarterly_transf_table.do: creates sample that identifies targets and acquirers in mergers. It also identifies bank failures. Uses CSV_TRANSFORMATIONS.csv and banks_with_topholders_qtr.dta and creates transf_table_targets.dta and transf_table_acquirers.dta
6. aux_quarterly02.do: incorporates information on entry, exit, merger, and top holders into the sample of banks. Creates intermediate sample mydata_allq_with_toph.dta
7. bhcb_quarterly01.do: loads the bank level data, combines with information on banks' top holders and generates the sample at the Bank Holding Company Level for those banks with top holders and keeps individual banks for banks with no top holders. Sample generated is mydata_bhcb_allq.dta which is the main sample for analysis.
8. work_quarterly_concentration01.do: this files loads bank sample and creates measures of concentration. Resulting sample in data_concentration_dep_national_level.xls/.mat and market_shares_by_size01.log
9. fig_q_concentration01.m: loads data on concentration at national level and state level and creates plots (Figure 1 and Figure 2).
10. sod_work_01_bhc.do: Takes data on deposits at branch level and generates HHI index at State level (yearly frequency). It also uses information on branches to create the diversification index.
11. work_rose_panzar01.do: takes bank level data and estimates H-stat by bank size. Generates sample reg_Hstat.csv (data_Hstat01.mat).
12. fig_Hstat.m: Loads data with Hstat by bank size and creates figure (Figure 3).
13. work_zipfs01.do: uses bank level data to estimate Deviations from Zipf's Law (presented in Figure 4 and Table A.2
14. fig_zipfs01.m: loads data on distribution of deposits and ranking to plot in Zipfs distribution (Figure 4)
15. work_quarterly_AR1_deposits01.do: estimates AR(1) deposit process by bank size. Estimated results in results_dep_AR1_bysize.log
16. work_quarterly_costs01.do: estimates bank trans-log equations for marginal non-interest expenses and non-interest income. Provides estimates of marginal net non-interest income, markups and other measures by bank size.
17. fig_markups01.m: this file uses data on the markup distribution to plot the median across all banks and the average markup for the Top 35
18. work_quarterly_exit_entry01.do: this file creates exit/failure rates time series and exports the data for the chart
19. fig_q_exit_rates.m: loads the data and makes the plot (Figure 7)
20. work_balance_sheet01.do: creates balance sheet ratios makes the explained normalization and presents the log file with results.

21. work_quarterly_returns01.do: computes loan returns and other measures affecting bank profits and prints averages, computes standard deviations, business cycle correlations (all banks and by bank size).
22. graphs_deposits_top35.do: loads data and create charts with distribution of banks for Top 35 banks identifying the Top 4 (and related charters).
23. work_quarterly_costs02.do: estimates elasticity of net marginal expenses across bank size.
24. work_dep_to_output01.do: estimates the ratio of loans to GDP and deposits to GDP using FRED Data.

Model Programs (folder "\Model")

These codes can be used to run all cases of the model. This includes the baseline parameterization, the Bank Run Granularity experiment (Section 7), and the Too-Big-to-Fail experiments presented in Section 8.

1. globals02.f90: this file sets parameters, options for policy counterfactuals and creates global functions. Relevant Options for experiments:
    - estimate = 0 (line 38) to run code for given parameters, set =1 to iterate on parameters define in subroutine myfun in subroutines.f90
    - baseline_case=0 to run with pre-determined entry costs. If set =1, entry costs will be set during the simulation routine to match corresponding moment. Note that if baseline_case=1 measure of household consumption is not valid.
    - compute_sp = 1 runs the experiment to compute granularity measures or spillovers.
    - Key parameters for Too-Big-To-Fail policies in lines 95-99
2. Functions02.f90: this file contains set of functions used during optimization routines.
3. Subroutines02.f90: this file contains the routines that sets the grids, solves the banks problem, aggregate functions, and simulates the model to generate moments.
   Note: subroutines from "Numerical Recipes in FORTRAN; The Art of Scientific Computing" not included as they are proprietary. See: http://www.nrbook.com/a/bookf90pdf.html to find the book and obtain a license
4. mysimplex02.f90: this file contains auxiliary routines to estimate the parameters of the model.
5. main02.f90: this code makes the call to all other routines and saves the functions.
6. simulations_reg02.m: this file loads model output and runs the regressions in Table 7
7. simulations_granular02.m: this file loads the model output and runs the regressions in Table 8 (input the index value so code selects the desired combination for dependent and independent variable)
8. simulations_spillovers02.m: this file loads the model output and creates Figures 8, 9, 10, 11.

Instructions to Replicators

Details Data

- The file \Data\data_appendix.pdf provides a list of variables downloaded from Consolidated Report of Condition and Income.

Details Fortran Codes

1) To obtain the output for the baseline model with imperfect competition, using the Fortran codes in folder "\Model\" on a machine/cluster running Intel Fortran and MPI,
   a) Compiled the codes using file compile_d01_f02.sh
      - The file b02 (to be created when compiling the files) is the executable file that is run to make the computation
      - Run the file loading the vector of shocks in file uvec
2) To obtain the output for Tables 8 and 9 "Bank Run Granularity", in file globals02.f90 set baseline_case=0 and compute_sp = 1 with no further changes in other files.
3) To obtain the results for Figure 8, 9, 10, and 11 "Aggregate Spillovers" in file globals02.f90 set baseline_case=0 and compute_sp = 1 plus adjust subroutines02.f90 as follows:
   (1) subroutine sub_spillovers: comment lines 4133-4134, uncomment lines 4137-4138 (periods 50-100 correspond to the baseline, 101 to 150 to the experiment with transitory national banks shocks, 201-250 to the experiment with transitory regional bank shocks).
   (2) Subroutine simu_spillovers:
      - Regional bank 1: uncomment lines 4467-4468, 4473-4480, comment lines 4954-4962, 4976-4984, 4993-5001 (this prevents writing over specified sequence of deposit shocks).
      - Regional bank 2: uncomment lines 4494-4495, comment lines 5033-5041, 5055-5063, 5072-5080 (this prevents writing over specified sequence of deposit shocks).
      - National bank: uncomment lines 4520-4546, comment lines 4877-4885, 4898-4905, 4915-4923 (this prevents writing over specified sequence of deposit shocks).
4) To run the Too-big-to-Fail counterfactual (no bailouts), using the provided version of the codes, only adjust globals02.f90 as follows:
   i) Set phibail_n =0.0_DP (line 95)
   ii) Set pbail = 0.0_DP (line 99)
      (1) To obtain spillovers effect, in addition to these changes set compute_sp=1.
5) To run the National and Regional banks bailouts experiment, using the provided version of the codes, only adjust globals02.f90 as follows:
   i) Set phibail_r = 0.90; (phibail_n=0.90 in provided code).

Computational Requirements

Software Requirements

- Stata (code was last run with version MP 17.0 for Windows)
    - estout: Additional Package: package st0085_2 Installed from http://www.stata-journal.com/software/sj14-2 “Making regression tables from stored estimates”
- Linux: CentOS 7.4 distribution
- Intel Fortran Compiler
    - Intel(R) MPI Library for Linux* OS, Version 2021
- Matlab (code was run with Matlab Release 2022a)
    - Symbolic Math Toolbox
    - Optimization Toolbox
    - Global Optimization toolbox

Memory and Runtime Requirements

The Stata code was last run on a 4-core Intel based laptop with Windows 11.

The Matlab code was last run on a 4-core Intel based laptop with Windows 11.

The Fortran Codes were run on a Linux cluster running Intel Fortran and MPI (as described in Software Requirements) with 148-core Intel server with 1024 GB ram.

List of Tables and Programs

Figure/Table #	Program	Output File	Note
Figure 1	\Data\work_quarterly_concentration01.do \Data\fig_q_concentration01.m	\Data\fig_c4c35_nbanks01q.eps	Stata code computes and exports needed information to create figure in Matlab
Figure 2	\Data\work_quarterly_concentration01.do \Data\sod_work01_bhc.do \Data\fig_q_concentration01.m	\Data\fig_hhi01.eps	Stata codes compute and exports needed information to create figure in Matlab
Figure 3	\Data\work_rose_panzar01.do \Data\fig_Hstat.m	\Data\fig_Hstat_year_bysize.eps	Stata code computes and exports needed information to create figure in Matlab
Figure 4	\Data\work_zips01.do \Data\fig_zips01.m	\Data\fig_zipf_dep_reg_2panels.eps	Stata code computes and exports needed information to create figure in Matlab
Figure 5	\Data\sod_work01.do \Data\fig_q_concentration01.m	\Data\fig_dep_state_di_all.eps	Stata code computes and exports needed information to create figure in Matlab
Table 1	\Data\work_quarterly_AR1_deposits01.do	\Data\results_dep_AR1_bysize.log	Estimates of deposit process by bank size
Table 2	\Data\work_quarterly_costs01.do	\Data\avg_costs_markups01.log	Estimates of bank cost structure by bank size and for all banks
Figure 6	\Data\work_quarterly_costs01.do \Data\fig_markups01.m	\Data\fig_markup_t35_median.eps	Stata file creates markups and series. Matlab file loads data and creates figure with median and Top 35 markup
Figure 7	\Data\work_quarterly_exit_entry01.do \Data\fig_q_exit_rates01.m	\Data\fig_fail_rate02.eps \Data\quarterly_exit_entry_stats.log	Stata file computes exit rates and matlab file plots the series
Table 3	\Data\work_balance_sheet01.do	\Data\Balance_sheet_stats.log	File computes balance sheet ratios displayed in log file
Table 5	\Data\work_quarterly_returns01.do	\Data\BC_corr_returns01_allbanks.log	Stata codes create corresponding measures and

	\Data\work_quarterly_costs01.do \Data\work_quarterly_exit_entry01.do \Model\globals02.f90 \Model\functions02.f90 \Model\subroutines02.f90 \Model\main02.f90	\Data\costs_markups_BC_corr01_allbanks.log \Data\exfail_BC_corr01_allbanks.log \Model\output_baseline01.txt	compute correlations in the data (presented in log files). Fortran codes solve the model and print result in output_baseline01.txt
Table 6	\Data\work_quarterly_returns01.do \Data\work_quarterly_costs01.do \Model\globals02.f90 \Model\functions02.f90 \Model\subroutines02.f90 \Model\main02.f90	\Data\avg_returns01_allbanks.log \Data\avg_costs_markups01.log \Model\output_baseline01.txt	Stata codes create corresponding measures and compute correlations in the data (presented in log files). Fortran codes solve the model and print result in output_baseline01.txt
Table 7	\Model\globals02.f90 \Model\functions02.f90 \Model\subroutines02.f90 \Model\main02.f90 \Model\simulation_reg02.m	\Model\reg_crisis.out	Fortran codes used to solve the model. Matlab file loads data from model (data_reg_crisis01.mat) and run the regressions. Output printed in reg_crisis.out
Table 8	\Model\globals02.f90 \Model\functions02.f90 \Model\subroutines02.f90 \Model\main02.f90	\Model\output_baseline_granular01.txt	Fortran codes used to solve the model. Output printed in txt file. Moments in the Table under “Some Additional Moments” (lines 304-313 for baseline, lines 431-440 for only dominant bank shock)
Table 9	\Model\globals02.f90 \Model\functions02.f90 \Model\subroutines02.f90 \Model\main02.f90 \Model\simulation_granular02.m	\Model\reg_granular_baseline01.out	Fortran codes used to solve the model. Matlab file loads data from model (data_baseline_granular01.mat) and run the regressions. Output printed in reg_granular_baseline01.out

Figures 8, 9, 10, 11	\Model\globals02.f90 \Model\functions02.f90 \Model\subroutines02.f90 \Model\main02.f90 \Model\sim_spillovers02.m	\Model\ fig_spillovers_all01.eps \Model\ fig_spillovers_all02_dn.eps \Model\ fig_spillovers_all02_dr.eps \Model\ fig_spillovers_all03.eps	Fortran codes used to solve the model. Matlab file loads data from model (data_spillovers01.mat) and creates figures.
Tables 10, 11	\Model\globals02.f90 \Model\functions02.f90 \Model\subroutines02.f90 \Model\main02.f90	\Model\output_baseline01.txt \Model\output_notbtf01.txt	Fortran codes used to solve the model (baseline case and no too-big-to-fail policies). Output files contain model moments used in the table.
Table 12	\Model\globals02.f90 \Model\functions02.f90 \Model\subroutines02.f90 \Model\main02.f90	\Model\output_baseline_granular01.txt \Model\output_notbtf_granular01.txt	Fortran codes used to solve the model (baseline case and no too-big-to-fail policies). Setting to run spillovers experiment.
Figure A.1	\Model\graphs_deposits_top35.do	\Data\dep_84_14.eps	Stata file loads data and creates charts
Table A.2	\Data\work_zipfs01.do	reg_adj_zipfs_assets_top3000.cvs reg_adj_zipfs_loans_top3000.cvs reg_adj_zipfs_dep_top3000.cvs	Stata file loads data, run the regressions and store results in csv files.
Table A.3	\Data\work_quarterly_exit_entry02.do	quarterly_exit_entry_bysize.log	Stata file loads data, creates statistics and prints results in log file
Table A.4	\Data\work_quarterly_returns01.do	\Data\avg_returns01_allbanks.log \Data\avg_returns01_top4.log \Data\avg_returns01_top5_35.log \Data\avg_returns01_fringe.log \Data\test_size_diff_avg01.log \Data\results_stddev_returns_8419_bysize.log	Stata files load data and compute moments. Averages reported in avg_returns01 files (all banks, top 4, top 5-35, fringe), tests for size differences in test_size_diff_avg01 files, and standard deviations in results_stddev_returns_8419_bysize.log

	\Data\work_quarterly_cost01.do	\Data\avg_costs_markups01.log \Data\test_size_diff_avg02.log \Data\results_stddev_costs_8419_bysize.log	Averages markups and costs (all banks and by size) in avg_costs_markups01.log, tests size differences in test_size_diff_avg02, and std dev in results_stddev_costs_8419_bysize.log
Table A.5	\Data\work_quarterly_returns01.do \Data\work_quarterly_cost01.do	\Data\BC_corr_returns01_allbanks.log \Data\BC_corr_returns01_top4.log \Data\BC_corr_returns01_top5_35.log \Data\BC_corr_returns01_fringe.log \Data\costs_markups_BC_corr01_allbanks.log \Data\costs_markups_BC_corr01_top4.log \Data\costs_markups_BC_corr01_top5_35.log \Data\costs_markups_BC_corr01_fringe.log	Stata file computes business cycle correlations and prints in log file
Table A.7/A.8	\Data\work_quarterly_return01.do \Data\work_quarterly_AR1_deposits01.do \Data\work_quarterly_costs01.do \Data\work_quarterly_costs02.do	\Data\avg_returns01_allbanks.log \Data\avg_returns01_top4.log \Data\avg_returns01_top5_35.log \Data\avg_returns01_fringe.log \Data\results_stddev_returns_8419_bysize.log \Data\results_dep_AR1_bysize.log \Data\avg_costs_markups01.log \Data\test_size_diff_avg02.log \Data\results_stddev_costs_8419_bysize.log \Data\estimates_cost_elasticity.log	Stata files compute moments from data and print in log files. Fortran codes compute model and print output in txt file.

	\Data\work_quarterly_concentration01.do \Data\work_balance_sheet01.do \Data\work_dep_to_output_do \Data\work_quarterly_entry_exit_entry01.do \Data\work_quarterly_entry_exit_entry02.do \Model\globals02.f90 \Model\functions02.f90 \Model\subroutines02.f90 \Model\main02.f90	\Data\market_shares_by_size01.log \Data\balance_sheet_stats.log \Data\dep_to_output01.log \Data\quarterly_exit_stats.log \Data\quarterly_exit_entry_by size.log \Model\output_baseline01.txt	
Table A.9	\Model\globals02.f90 \Model\functions02.f90 \Model\subroutines02.f90 \Model\main02.f90 \Model\simulation_granular02.m	\Model\reg_granular_baseline01.out	Fortran codes used to solve the model. Matlab file loads data from model (data_baseline_granular01.mat) and run the regressions. Output printed in reg_granular_baseline01.out
Tables A.10/A.11	\Model\globals02.f90 \Model\functions02.f90 \Model\subroutines02.f90 \Model\main02.f90	\Model\output_baseline01.txt \Model\output_reg_nat_tbtf01.txt	Fortran codes used to solve the model baseline specification and case with bailouts for regional and national banks. Baseline model results in output_baseline01.txt National and Regional Bank bailout case in output_reg_nat_tbtf01.txt

References

Vetterling, William T., and William H. Press. *Numerical recipes in Fortran: the art of scientific computing*. Vol. 1. Cambridge University Press, 1992.