README The Effects of Changes in Local Bank Health on Household Consumption

Daniel Cooper and Joe Peek

How to Access the Data

- **Panel Study of Income Dynamics (PSID)**
  - Public data can be downloaded from `https://psidonline.isr.umich.edu/`
  - Instructions for obtaining access to restricted geographic identifiers can be found at `https://simba.isr.umich.edu/restricted/ProcessReq.aspx`
- **Call Reports**
  - Public Option #1a
    - see `https://www.chicagofed.org/banking/financial-institution-reports/commercial-bank-data-complete-2001-2010`
    - Only Commercial Banks. 2001-2010
  - Public Option #1b
    - see `https://www.chicagofed.org/banking/financial-institution-reports/commercial-bank-data-complete-1976-2000`
    - Only Commercial Banks. 1976-2000
  - Public Option #2
    - see `https://cdr.ffiec.gov/public/PWS/DownloadBulkData.aspx`
    - Only Commercial Banks. Only back to 2001. Seems to be a subset of variables.
  - Structure Data - Public Option #1
    - see `https://www.chicagofed.org/banking/financial-institution-reports/commercial-bank-structure-data`
    - Structure Info - M&A's, transformations, etc.
  - Structure Data - Public Option #2
    - see `https://www.ffiec.gov/npw/FinancialReport/DataDownload`
    - Only current quarter, but contains structure info

- **FDIC Summary of Deposits (SOD)**
  - see `https://www7.fdic.gov/sod/dynaDownload.asp?barItem=6`
  - Data prior to 1994 are obtained from the Board of Governors / National Information Center Tables - SUMD
- **BLS Quarterly Census of Employment and Wages (QCEW)**
  - NAIC and SIC
  - Used to construct MSA-industry level employment data and then a Bartik indicator
  - Downloaded from `https://www.bls.gov/cew/datatoc.htm`
- **Census CBSA Codes**
  - from U.S. Census Bureau, Population Division; Office of Management and Budget, February 2013 delineations. Internet Release Date: March 2013
- **CoreLogic House Price Index**
  - HPI data can be purchased from CoreLogic, see `https://www.corelogic.com/products/corelogic-hpi.aspx`
- **FRBNY Consumer Credit Panel: Equifax**
  - Credit data can be purchased from Equifax
- **CAMELS *confidential supervisory bank ratings***
  (**c**apital adequacy, **a**sset quality, **m**anagement, **e**arnings, **l**iquidity, and **s**ensitivity to market risk)

Replication Files

Bank Health Data

See the supplemental README (a**Readme_bhmeasures**) for a detailed explanation of the suite of files needed to construct the bank health measures.

1. *bh2020-sub-00-project-structure-master.do*
2. *bh2020-sub-01a-createCallData.do*
3. *bh2020-sub-01b-prepCallData.do*
4. *bh2020-sub-01c-prepDepositsData.do*
5. *bh2020-sub-02a-mergeCallSOD.do*
6. *bh2020-sub-03a-splitDatasetsByLocation.do*
7. *bh2020-sub-04a-makeBhMeasures.do*
8. *bh2020-sub-05a-reCombineAndFinalCleaning.do*
9. ***bankhealth_lags.do*** simple cleaning of raw BH data, generating lags

PSID Data

10. ***psid_mnemonics_bh.do*** mnemonics needed to pull public PSID data

11. ***psid_agesex_samehead.do***

    (a) Adjusts for errors in the age and sex variables in the household level data with individual data.

    (b) Adjusts for issues with the variable that denotes whether the household head has changed between waves. This is important because we assign a 'new' household whenever the head of household changes.

12. ***psid_educ_data_cleaning.do*** cleans education variables, requires output from *psid_agesex_samehead.do*

13. ***psid_ageearnings.do***

    (a) sets up separate time series of demographic and employment information for head/wife using individual and household level data

    (b) constructs age earnings profiles and income deviations for head/wife separately then re-forms households

    (c) see comments within the code for detailed steps, even more extensive detail on sample restrictions is available upon request

14. ***psid_taxsim.dta*** Output from TAXSIM program which calculates federal and state income tax liabilities from survey data. See `http://users.nber.org/~taxsim/`

BPP

See the supplemental README (a**Readme_BPP**) for a detailed explanation of the suite of files needed to construct the imputed nondurables consumption measure following Blundell, Pistaferi and Preston (2006).

15. *BPP_cons_impute_cex.do*
16. *merge_jan20.dta*
17. *pricescex2020.dta*
18. *imputebetas_bpp.dta*
19. *BPP_cons_impute_psid.do*
20. *psiddata_bpp.dta*

Other Data

21. ***pce_bh_def.do*** Creates a PCE defaltor (base year 2000).

22. ***hpi.do*** CoreLogic House Price Index for states and MSAs, calculates lagged growth rate, deflates. (Data must be obtained directly from CoreLogic, see above.)

23. ***emp_bartik_naic_sic.do*** Creates lagged Bartik employment growth for states and MSAs.

Merging & Analysis

Tables 2 - A10

24. ***maindata.do*** merges together all non-equifax data and creates the main dataset(s) for the regressions.

25. ***tables.do*** Creates all tables for the paper except Table 1 (see *equifax.do*).

Table 1

26. ***equifax.do*** Uses individual-level Equifax data to get MSA-level measures of real per capita loans and regresses with BH, employment growth, HP growth. (Equifax data must be obtained separately, see above.)

Geographic Definitions and Crosswalks

The following excel files contain geographic definitions and crosswalks that we use in various parts of the analysis.

- *census_cbsa_2013_codes* CBSA codes from U.S. Census Bureau, Population Division; Office of Management and Budget, February 2013 delineations
- *us_postal_codes* complete list of US zip codes for *equifax.do*
- *ZIP_CBSA_032014* zip code to CBSA crosswalk for *equifax.do*

Data Dictionary

This dictionary describes the variables that contribute to the tables in the paper in their final forms.

Main Data (output from *maindata.do*)

- $unique = ER30001 * 1000 + ER30002$, PSID unique individual identifier (combination of 1968 ID and person number)
- $location_id$, location group for clustering standard errors and fixed effects, constructed from randomized dummies for MSAs and states' non-MSA areas such that it **cannot** be used to identify households' locations, simply used to group them.
- ln_psidc_food: Log real household food consumption, sum of food at home, food away from home, and food delivered. see *psid_mnemonics_bh.do*
- bpp: Imputed nondurables consumption, see *aReadme_BPP*
- ln_psidc_altc1: Alternative Consumption Measure, 2005-2015, log of the sum of $hfurn$, $hrec$, $vaca$, $school$, $cloth$, $food$. see *psid_mnemonics_bh.do*

- *ln_famincrat*: Log Income, log of real after tax (using TAXSIM, see *psid_taxsim.dta*) total family income
- *hp_ev_w*1: House Price Growth, see *hpi.do*
- *bartik_ebev_w*1: Bartik Empl. Growth, see *emp_bartik_naic_sic.do*
- The **BH variables** have complex names as there are many variations of them. However, they follow a clear structure:
  - **General structure**:
    $$(re/npl)_a_50_ev$$
    * re/npl: real estate non-performing loans or all non-performing loans
    * a: stands for 'total assets', the denominator of the BH measure
    * 50: percentile of BH we are looking at in the location
    * *ev*: (stands for 'everywhere') means we use the MSA measure where available, and the state's rural (non-MSA) area measure if the household resides outside an MSA (or lives in an MSA that has too few banks)
  - **Multi-locational BH structure**:
    $$(re/npl)_a_50_(5/10)m_bhc_ev$$
    * $5/10$: maximum percentage of location deposits a bank may have to be considered a multi-locational bank. In other words, percentage of location deposits that we use to define multi-locational banks i.e. multi-locational bank if less than 10%/5% of its deposits are from that location. See the paper for more details.
    * m: multi-locational (rather than local) measure
    * *bhc*: multi-locational measure based on Bank Holding Companies (BHC)
  - **Constraint designation structure**:
    $$BH_(c/u)(l/h)(/_30)(/_5)$$
    * c/u: constrained/unconstrained
    * l/h: low/high liquid asset holdings
    * $/_30$: negative income deviation/neg inc dev or age under 30
      "Under 30":
      · if single, head under 30
      · if single earner, earner under 30
      · if dual earner, both under 30
    * $/_5$: negative income deviation/5% neg inc dev
  - Baseline BH measure: $npl_a_50_5m_bhc_ev$, meaning the median NPL ratio for multi-locational banks in the local area where ML banks are the local branches of parent bank holding companies (BHCs) that have less than 5 percent of overall deposits in a given location, but where ML BHCs account for at least a 5 percent share of all local deposits

- *vl_BH*: Income Volatility x BH measure. Income volatility is the mean squared deviation of actual household income from predicted household income.
- *ln_l4_LocShrM05bhc_100_ev* is the log percent of local deposits held by mutli-locational banks in an area. We use this measure to drop locations that don't have at least 5/10 percent of their local deposits held in multi-locational banks.
  - *every_dum_npl_a_bhc*$(5/10)rf$: dummy for regression sample that performs the locations restriction described above.
- *totwr_100*: Wealth/Income, the ratio of households' assets (bonds, stocks, cash, and IRA holdings along with the net value of any businesses, vehicles, or non-primary-residence real estate) less any nonhousing debt plus housing equity (house value less outstanding mortgage debt), if any, relative to lagged income. (Income is multiplied by 100 for ease of presentation.)
- *totar_100*: Financial Assets/Income, bonds, stocks, cash, and IRA holdings along with the net value of any businesses, vehicles, or non-primary-residence real estate relative to lagged income x100.
- *totdr_100*: Debt/Income, non-housing debt relative to lagged income x100.
- *ageh*, *age2*: Household head's age, age$^2/1000$
- $famsize$, $famsizessq$: Family size, family size squared
- *m_se*, *m_de*: Married single earner, married dual earner (omitted category is single)

Equifax Data (output from *equifax.do*)

- *_w* means windsorized at 1% (by year)
  - *nonmort_cap_real_ch_w*, Real per capita non-mortgage loan growth excluding home equity. See *equifax.do* for detail on how these loan growth measures are constructed from the CCP Equifax data.
  - *nonmort_he_cap_real_ch_w*, Real per captia non-mortgage loan growth including home equity
  - *l_npl_a_50_bhc_w*, ML NPL50, baseline BH measure (see above)
  - *l_hp_gr_w*, House price growth, see *hpi.do*
  - *l_emp_gr_w*, Bartik employment growth, see *emp_bartik_naic_sic.do*

Software

- Stata15/Stata16 MP
- LaTeX (TeX distribution software: MiKTeX 2.9; editor/compiler: WinEdt 10.3)