

Replication Instructions for

“Mortgage Refinancing, Consumer Spending, and Competition: Evidence from the Home Affordable Refinance Program”

by Sumit Agarwal, Gene Amromin, Souphala Chomsisengphet, Tim Landvoigt, Tomasz Piskorski, Amit Seru & Vincent Yao

Data Availability Statement

Summary of Data Availability: No data can be made publicly available.

Details on Each Data Source

We use several confidential datasets in our paper. If you need additional guidance beyond the one provided in this document, please contact the authors directly.

The *Equifax-Black Knight Financial Services Credit Risk Insight Servicing McDash* data (Equifax-BKFS CRISM, 2008-2013), is currently commercially offered by Black Knight and Equifax (<https://www.blackknightinc.com/>). It covers the majority of the US mortgage market during our sample period, and most loans guaranteed by the Government Sponsored Enterprises (GSEs). The data also contains information on interest rates and borrower and loan-specific characteristics, including FICO score at origination, loan-to-value ratio, five-digit zip code of origination, loan purpose, and whether the loan is fixed or adjustable-rate. It also includes dynamic data on monthly payments, outstanding mortgage balances, delinquency status, and prepayment. The data provider also provides us each borrower’s credit bureau records, merged using unique consumer identifiers. We exploit this data to track a borrower across time to study her refinancing history, including mortgage terms across loans. It allows to account for a host of loan, property, and borrower characteristics. The data also provides us with a borrower’s monthly credit history, including auto debt balance information.

This data has been used by researchers in academic institutions as well as governmental and financial intuitions (we gained access to our data through the Federal Reserve Bank of Chicago where one of our co-authors works).¹ The other researchers could purchase this data subject to Data Use Agreements and conduct similar analysis to ours whenever we use this data. Requests for complimentary demonstrations of the Equifax-BKFS CRISM can be sent to BlackKnightDNA@bkfs.com.

The second main dataset we use in our paper is a proprietary internal database of conforming mortgages securitized by a Large Secondary Market Participant (from the Federal National Mortgage Association (Fannie Mae) where one of our co-authors used to work). This loan-level

¹ See for example Martin Beraja, Andreas Fuster, Erik Hurst, Joseph Vavra, Regional Heterogeneity and the Refinancing Channel of Monetary Policy, *The Quarterly Journal of Economics*, Volume 134, Issue 1, February 2019, Pages 109–183.

monthly panel data has detailed dynamic information on a rich array of loan, property, and borrower characteristics (e.g., interest rates, location of the property and current borrower credit scores and LTV ratios) and monthly payment history. Importantly, this data contains unique Social Security Numbers (SSN) for each borrower, allowing us to track the refinancing history of each borrower, the servicer responsible for prior and current mortgage of the borrower, and whether refinancing was done under HARP. It allows us to obtain all the present and prior mortgage terms including all relevant information on fees applied during the refinancing process. Most importantly for our purposes this data *includes* administratively set *GSE g-fees* charged for the insurance of default risk and the lender identity field that are not included in the Equifax-BKFS CRISM data. Researchers interested in obtaining the data should contact the authors.

In our regional analysis we use additional proprietary databases. The first dataset collected by the Office of the Comptroller of Currency allows us to measure the aggregate quarterly credit card spending of borrowers in a particular zip code. Currently, this data cannot be accessed by the general public and is not commercially available. The researchers could try to gain access to similar data with Data Use Agreements from the credit bureau agencies such as Equifax or Experian (e.g., <https://www.experian.com/consumer-information/consumer-spending-data>; <https://datadrivenmarketing.equifax.com/digital-targeting-segments/credit-card-spending-propensities/>). The second database comprises the auto sales data from R. L. Polk & Company, which allows us to directly measure the car purchases in a zip code. The researchers could gain access to this data with Data Use Agreements with R. L. Polk & Company (<https://ihsmarkit.com/btp/polk.html>). Finally, we also use zip code level house price indices from CoreLogic. The researchers could gain access to this data with Data Use Agreements with CoreLogic (<https://www.corelogic.com/industry/real-estate-solutions.aspx>).

Computational Requirements

The replicationProgram.do file can be run in Stata 15 (<https://www.stata.com/stata15/>). The program can be run on the PC with the equivalent or higher technical specifications: 3.2GHz 6-core 8th-generation Intel Core i7, 16GB 2666MHz DDR4 Onboard Memory, 1TB SSD storage. The run time for each table should take between several minutes to several hours (less than eight hours).

Analysis Code

The ReplicationProgram.do file contains the programs and directions needed to replicate the results from “Mortgage Refinancing, Consumer Spending, and Competition: Evidence from the Home Affordable Refinance Program.” The replicationProgram.do file can be run in Stata 15.

The code uses the following datasets (that should be stored in “C:\Data\” folder):

borrower_panel.dta: The quarterly panel data on credit outcomes and characteristics of borrowers with treatment and control group loans. The treatment group consists of borrowers with HARP eligible loans (that is GSE loans with current LTV greater than 80%) and the control group consists of borrowers with prime FRM loans not guaranteed by the GSEs (the non-agency loans) with current LTV greater than 80% as of the beginning of 2008. Data is extracted from the Equifax-BKFS CRISM, 2008-2013 data.

borrower_panel_HARPrefi.dta: The data covering quarterly credit outcomes of borrowers with program eligible treatment loans as described above that refinanced during our sample period. The data is extracted from the Equifax-BKFS CRISM, 2008-2013 data. For each borrower the data covers the period from four quarters prior to the refinancing date of the loan and the eight quarters after (including the refinancing quarter).

regional.dta: The data on regional outcome variables covering 2008-2013 period. The data on quarterly consumer card spending growth was collected by the Office of the Comptroller of Currency, 2008-2013. The data on auto spending at the zipcode level is from Polk. The data on foreclosures at the zip code level is from the Equifax-BKFS CRISM, 2008-2013 data, while the data on house prices is from CoreLogic. The data for % *Below CLL* is from Di Maggio et al. (2017). The pre-program program eligible share, *HARP Eligible Share*, is the fraction of outstanding first-lien GSE mortgage loans in a zip code that have current LTV ratios greater than 80 prior to the program implementation in the Equifax-BKFS CRISM, 2008-2013 data. The zip code ARM share, zip code average interest rate, and zip code LTV is also from the Equifax-BKFS CRISM, 2008-2013 data.

harp_refinance_0912.dta: Data containing characteristics of loans that refinanced under the HARP program during 2009-2012 period extracted from the Large Secondary Market Participant, 2005-2013

conform_refinance_0912.dta: Data containing characteristics of conforming loans that refinanced under regular GSE program guidelines during 2009-2012 period extracted from the Large Secondary Market Participant, 2005-2013 data.

conform_refinance_0509.dta: Data containing characteristics of conforming loans that refinanced under regular GSE program guidelines during 2005-2009 extracted from the Large Secondary Market Participant, 2005-2013 data.

HarpSpread_PrgChange.dta: Data containing characteristics of loans that refinanced under the HARP program during July 2012 till June 2013 extracted from the Large Secondary Market Participant, 2005-2013 data.

HARPsVsConfrmRefi_PrgChange.dta: Data tracking refinancing outcomes on HARP eligible loans and loans that are eligible for regular confirming refinancing from July 2012 till June 2013 extracted from the Large Secondary Market Participant, 2005-2013 data.

Output Files

The output files will be stored in “C:\Output\” folder.

Tables:

- | | | |
|---------------|---------------|---------------|
| • Table1.csv | • Table2C.csv | • Table3B.csv |
| • Table2A.csv | • Table2D.csv | • Table4A.csv |
| • Table2B.csv | • Table3A.csv | • Table4B.csv |

- Table5.csv
- Table6.csv

Figures:

- Figure1a.png
- Figure1b.png
- Figure2a.png
- Figure2b.png
- Figure3a.png
- Figure3b.png
- Figure4a.png
- Figure4b.png
- Figure5a.png
- Figure5b.png

References

CoreLogic, 2008-2013, “Zip-code level CoreLogic Home Price Index (HPI).” CoreLogic. (Accessed January 12, 2015).

Di Maggio, M., Kermani, A., Keys, B., Piskorski, T., Ramcharan, R., Seru, A., Yao, V., 2017. Interest rate pass-through: Mortgage rates, household consumption and voluntary deleveraging. *American Economic Review* 107, 3550–3588.

Equifax-BKSF CRISM, 2008-2013, “Credit Risk Insight Servicing McDash (CRISM) dataset produced by merging mortgage servicing records from Black Knight Financial Services (BKFS) with credit bureau data from Equifax.” Black Knight Financial Services. (Accessed August 16, 2016).

Large Secondary Market Participant, 2005-2013, “Loan origination and monthly performance records of residential mortgages from mortgage pools guaranteed by the Federal National Mortgage Association.” Federal National Mortgage Association (Accessed December 10, 2014).

Office of the Comptroller of the Currency, 2008-2013, “Zip-code level quarterly credit card spending data collected by the Office of the Comptroller of Currency.” Office of the Comptroller of Currency (Accessed May 28, 2015).

R. L. Polk & Company, 2008-2013, “Zip-code level auto sales index.” R. L. Polk & Company (Accessed January 21, 2015).