Privacy Considerations in Merging Sensitive Data: Case Study of the CARES Act Mortgage Forbearance Program

Larry Cordell

Federal Reserve Bank of Philadelphia

Presentation to the National Academies of Sciences, Engineering and Medicine

May 23, 2023

Disclaimer

- Any findings, views, or opinions in this presentation are those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

FEDERAL RESERVE BANK PHILADELPHIA

Federal Reserve Guiding Principals on Research

"Research and policy analysis can be data-intensive activities that involve analyzing detailed information. While some of this data include personal information, it is typically anonymized, or de-identified with strict prohibitions against re-identification. Some of this information is acquired from third-party data providers.... The safeguarding and responsible use of information for research is of the utmost importance to the Reserve Banks."

-- From Federal Reserve Bank of Philadelphia Privacy Notice (philadelphiafed.org)

Toll of COVID-19 on Mortgage Market Was Significant

- Pandemic market shutdown affected mortgage markets especially hard.
  - Unemployment peaked at 14.9% in April 2020.
  - By May, 6.6 million mortgages were past due on, effects of which were unequal across borrower race and income.

Many Black Homeowners Are Falling Further Behind on Their Mortgages

The share of Black homeowners in forbearance in mid-April was about 11%, more than double the overall rate

Wall Street Journal, May 4, 2021

- In April 2020, the CARES Act was passed, offering mortgage forbearance as an option to borrowers.
  - Forbearance allows borrowers to temporarily pause paying their mortgages without affecting their credit. Past due payments are expected to eventually be paid back.

Research Questions

- Was the CARES Act Forbearance Program addressing those most in need of assistance?

- What about longer-term assistance to help borrowers stay in their homes once forbearance ended and past due payments needed to be paid back?

- How did the effects differ by race/ethnicity and income?

- We study these questions through the mortgage market
  - We have novel data that show:
    - Every borrower who received assistance in the form of forbearance
    - What actual payments were made on their mortgages
    - Borrower race and income on mortgages
  - These data points were not well captured through credit bureau reporting or previously existing mortgage servicing databases.

Data Sources

1) McDash Flash
   - New database with servicer-provided loan payment, forbearance, and loss mitigation status
2) McDash Primary Data
   - Loan origination and performance history data from mortgage servicing records
3) Credit Risk Insights Servicing McDash (CRISM) Data
   - Information on borrower's other credit accounts from their Equifax credit reports
4) Confidential Home Mortgage Disclosure Act (HMDA) Data
   - Borrower and co-borrower race, age, sex, ethnicity, application income and additional loan origination information

- In this case, all data are already anonymized.
- In addition, we take the extra step of scrubbing database-specific loan identifiers before the data is used by researchers.

Matching Process

- Equifax uses a proprietary algorithm to match mortgage loans in McDash to their credit bureau files to create the CRISM database.

- We "fuzzy-match" CRISM to HMDA using
  - Geography of property: HMDA provides Census Tract and McDash provides zip codes
  - Loan characteristics: loan amount, lien status, occupancy, loan purpose, loan type
  - Closing date

- We end up with 47% of CRISM loans matched to both HMDA and McDash Flash.

- We then verify that characteristics of matched loans are representative of the overall CRISM sample.

Our Analysis Shows that Pandemic Exacerbated Inequality
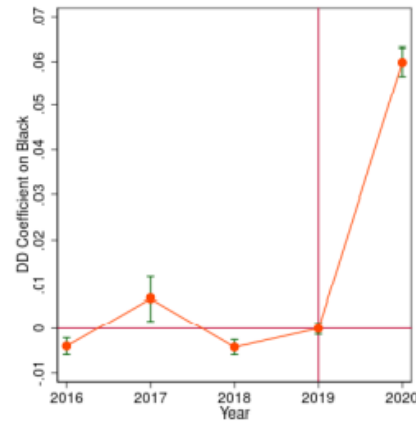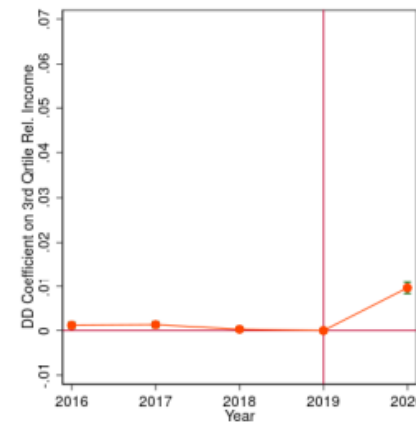

(a) Black **(b)** Hispanic **(c)** Asian
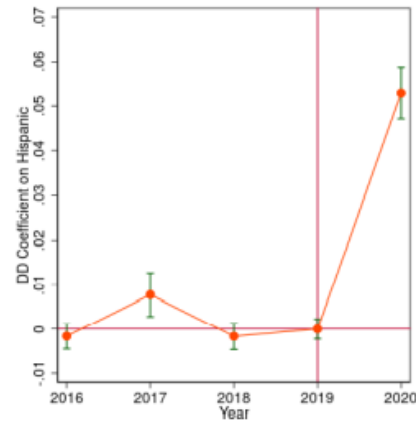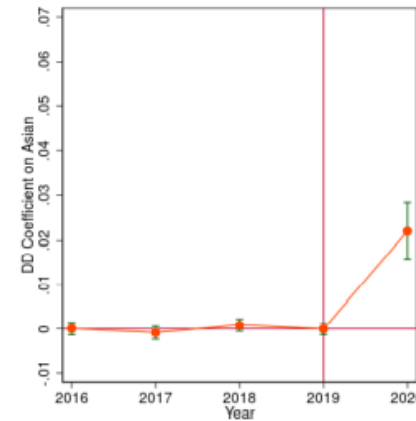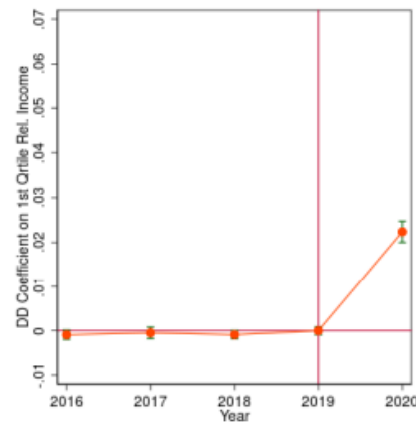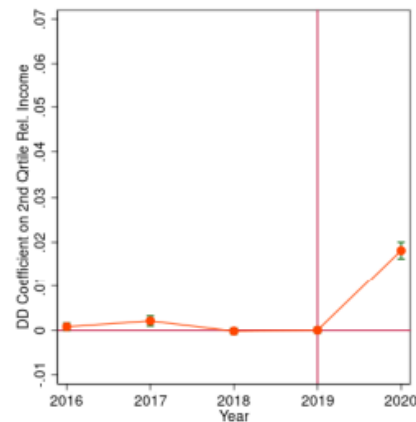
(d) 1st Qrtile Rel. Income **(e)** 2nd Qrtile Rel. Income **(f)** 3rd Qrtile Rel. Income

Source: An, Cordell, Geng and Lee (2022) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3789349

Which Reversed Once Forbearances Gets Accounted For

(a) Black

(b) Hispanic

(c) Asian
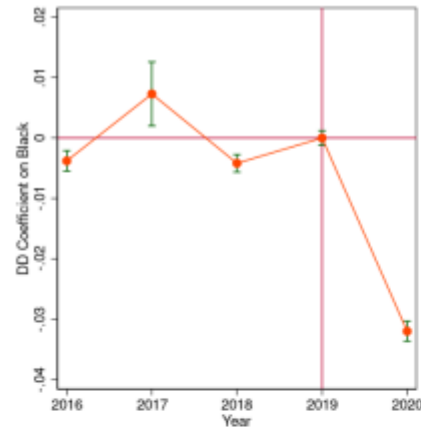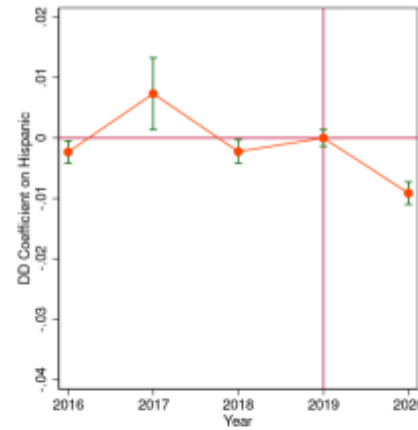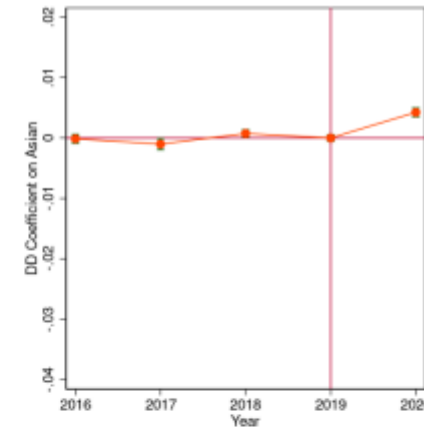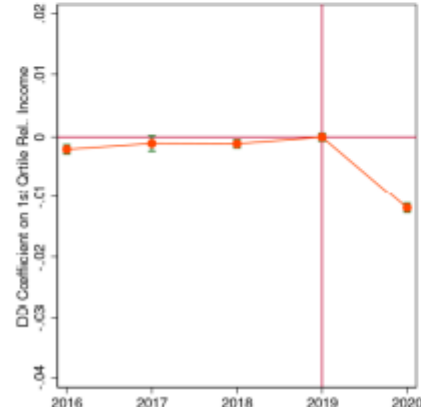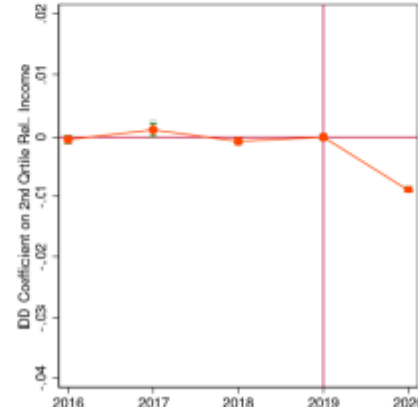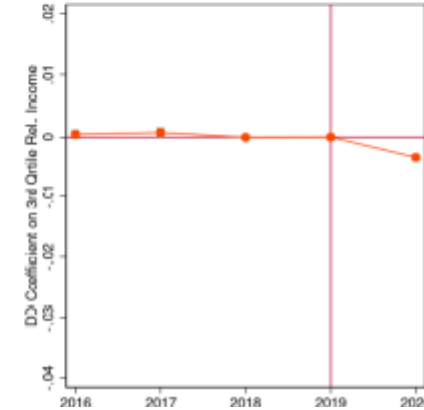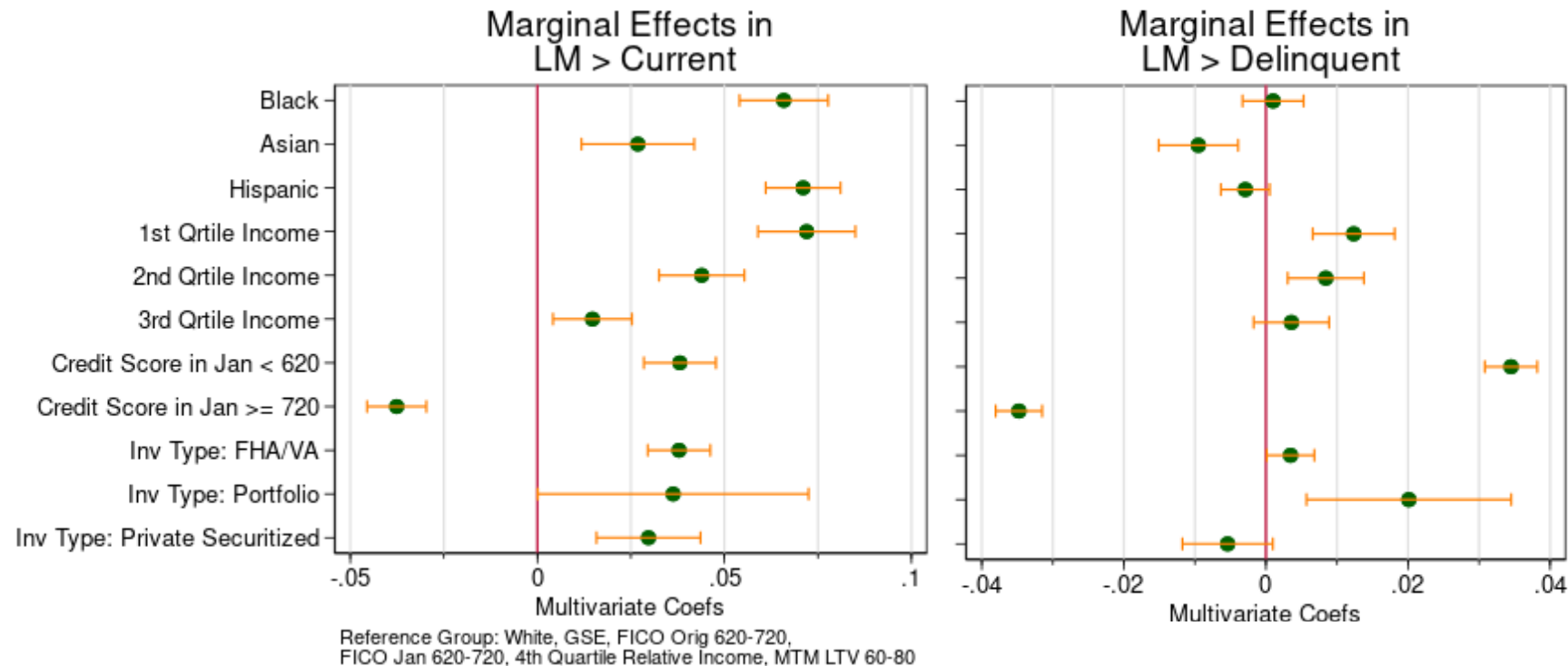
(d) 1st Qrtile Rel. Income

(e) 2nd Qrtile Rel. Income

(f) 3rd Qrtile Rel. Income

Source: An, Cordell, Geng and Lee (2022) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3789349

Post-Modification, Most Performing Well

Multivariate Regression Coefficients For Post-Loan Mod Exit Statuses As Of 2022 May

Reference Group: White, GSE, FICO Orig 620-720,
FICO Jan 620-720, 4th Quartile Relative Income, MTM LTV 60-80

Notes: Marginal effects calculated from multivariate multinomial logic regression with three exit statuses. These are calculated using loans that have ever entered forbearance during April 2020 – December 2020 and have ever entered Loss Mitigation. Mean rates for the three outcomes are: Current (55.9%), Delinquent (4.9%)

Source: An, Cordell, Geng and Lee (2022) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3789349

Conclusions

- A rich array of data exists to study important policy questions. Combining data from different sources can provide valuable insights.

- In combining these data, extra care must be taken to maintain the privacy of individuals.