Consumer Credit Reporting Data

Benedict Guttman-Kenney

University of Chicago, Booth School of Business

Today!

- Ask me questions
  - No such thing as a dumb question!

Today!

- Ask me questions
  - No such thing as a dumb question!

Aims for today:

- Learn about what consumer credit reporting data is.
  - Consider this a data asset you could use for your own research.
- Walk through some examples of papers using these data.
  - Hopefully help you generate your own research ideas.

Your Feedback Today Will Help You & Your Peers!

Writing Guide to Consumer Credit Reporting Data.

- Paper in-preparation.

- AEA Panel, January 2023
(NBER Household Finance, December 2021).

Co-authors (Credit Reporting Avengers):

- Christa Gibbs (CFPB)

- Donghoon Lee (FRBNY)

- Scott Nelson (Booth, ex-CFPB),

- Wilbert van der Klaauw (FRBNY)

- Jialan Wang (UIUC, ex-CFPB)

Brief Background

My Background

- Combine big data with tools from economics and psychology to understand consumer behavior - primarily researching household debt and credit information.
- PhD Economics Candidate at Booth: starting year 5, on market in year 6.
- Pre-PhD: 6+ years at UK financial regulator (Financial Conduct Authority, FCA)
  - At FCA mainly in Behavioral Economics & Data Science Unit.
  - Secondment to Federal Reserve Bank of New York (FRBNY).

My Experiences of Consumer Credit Reporting Data

- **UK:** Studied effects of payday loans and credit card nudges as well as variety of regulatory work understanding consumer credit markets, credit reporting agencies, and credit data
  - Worked with Equifax, Experian, & Transunion data.
  - Built UK Consumer Credit Reporting Panel.
  - Merged UK Household Consumer Panel (Understanding Society) with Credit Files.

My Experiences of Consumer Credit Reporting Data

- **UK:** Studied effects of payday loans and credit card nudges as well as variety of regulatory work understanding consumer credit markets, credit reporting agencies, and credit data
  - Worked with Equifax, Experian, & Transunion data.
  - Built UK Consumer Credit Reporting Panel.
  - Merged UK Household Consumer Panel (Understanding Society) with Credit Files.
- **CANADA:** Studying effects of increasing credit card minimum payments in Quebec
  - Bank of Canada's TransUnion data.

My Experiences of Consumer Credit Reporting Data

- **UK:** Studied effects of payday loans and credit card nudges as well as variety of regulatory work understanding consumer credit markets, credit reporting agencies, and credit data
  - Worked with Equifax, Experian, & Transunion data.
  - Built UK Consumer Credit Reporting Panel.
  - Merged UK Household Consumer Panel (Understanding Society) with Credit Files.
- **CANADA:** Studying effects of increasing credit card minimum payments in Quebec
  - Bank of Canada's TransUnion data.
- **US:** Studied US household leverage using Federal Reserve's Equifax data
  - FRBNY-CCP & CRISM.
  - Buy Now, Pay Later (BNPL) potential collaboration between bureau & lenders.

My Experiences of Consumer Credit Reporting Data

- **UK:** Studied effects of payday loans and credit card nudges as well as variety of regulatory work understanding consumer credit markets, credit reporting agencies, and credit data
  - Worked with Equifax, Experian, & Transunion data.
  - Built UK Consumer Credit Reporting Panel.
  - Merged UK Household Consumer Panel (Understanding Society) with Credit Files.
- **CANADA:** Studying effects of increasing credit card minimum payments in Quebec
  - Bank of Canada's TransUnion data.
- **US:** Studied US household leverage using Federal Reserve's Equifax data
  - FRBNY-CCP & CRISM.
  - Buy Now, Pay Later (BNPL) potential collaboration between bureau & lenders.
- **Booth** Currently most active user of Booth's TransUnion data.
  - Natural disasters, credit reporting, divorce & marriage, medical debt, & immigration.

What's a Consumer Credit Reporting Agency?

Terminology

- **These are all the same:** Consumer Credit Reporting Agencies / Credit Bureaus / Consumer Reporting Agencies / CRAs.

Terminology

- **These are all the same:** Consumer Credit Reporting Agencies / Credit Bureaus / Consumer Reporting Agencies / CRAs.
- Credit Data / Credit Files / Credit Reports / Credit Bureau Data.

Terminology

- **These are all the same:** Consumer Credit Reporting Agencies / Credit Bureaus / Consumer Reporting Agencies / CRAs.
- Credit Data / Credit Files / Credit Reports / Credit Bureau Data.
- Public Credit Registries vs. Private Credit Bureau
  - **(Public) Credit Registries:** Lenders legally required to report debts and administered by public organization (e.g. central bank, government department).
  - Common in South America mainland Europe.

Terminology

- **These are all the same:** Consumer Credit Reporting Agencies / Credit Bureaus / Consumer Reporting Agencies / CRAs.
- Credit Data / Credit Files / Credit Reports / Credit Bureau Data.
- Public Credit Registries vs. Private Credit Bureau
  - **(Public) Credit Registries:** Lenders legally required to report debts and administered by public organization (e.g. central bank, government department).
  - Common in South America mainland Europe.
  - **(Private) Credit Reporting Agencies:** Lenders report debts (generally voluntarily but may be required by law) and administered by private organization.
  - US, UK, Canada, Mexico.
  - Equifax, Experian & TransUnion are the main providers.
- Today I'm only speaking about **Consumer Credit Reporting Agencies.**
  - N.b. may co-exist with credit registries.

Terminology

- **These are all the same:** Consumer Credit Reporting Agencies / Credit Bureaus / Consumer Reporting Agencies / CRAs.
- Credit Data / Credit Files / Credit Reports / Credit Bureau Data.
- Public Credit Registries vs. Private Credit Bureau
  - **(Public) Credit Registries:** Lenders legally required to report debts and administered by public organization (e.g. central bank, government department).
  - Common in South America mainland Europe.
  - **(Private) Credit Reporting Agencies:** Lenders report debts (generally voluntarily but may be required by law) and administered by private organization.
  - US, UK, Canada, Mexico.
  - Equifax, Experian & TransUnion are the main providers.
- Today I'm only speaking about **Consumer Credit Reporting Agencies.**
  - N.b. may co-exist with credit registries.
- **Business credit reports** out-of-scope.
  - Much less known about these data. Only a couple of 2022 papers.

5

Why Do Credit Bureaus Exist? The Problem Lenders Face

- Lending is risky.
- If consumer doesn't pay you back (default) you lose money.

Why Do Credit Bureaus Exist? The Problem Lenders Face

- Lending is risky.
- If consumer doesn't pay you back (default) you lose money.
- **Information asymmetry** between borrowers and lenders. Jaffee & Russell 1976, Stiglitz & Weiss 1981, Pagano & Jappelli 1993, Jappelli & Pagano 2002)
- Those who you least want to lend to want to borrow most (adverse selection).

Why Do Credit Bureaus Exist? The Problem Lenders Face

- Lending is risky.
- If consumer doesn't pay you back (default) you lose money.
- **Information asymmetry** between borrowers and lenders. Jaffee & Russell 1976, Stiglitz & Weiss 1981, Pagano & Jappelli 1993, Jappelli & Pagano 2002)
- Those who you least want to lend to want to borrow most (adverse selection).
- Lenders predict expected risk of default. Determines (i) whether to accept credit application and (ii) on what terms (e.g. interest rate, amount lent).

Why Do Credit Bureaus Exist? The Problem Lenders Face

- Lending is risky.
- If consumer doesn't pay you back (default) you lose money.
- **Information asymmetry** between borrowers and lenders. Jaffee & Russell 1976, Stiglitz & Weiss 1981, Pagano & Jappelli 1993, Jappelli & Pagano 2002)
- Those who you least want to lend to want to borrow most (adverse selection).
- Lenders predict expected risk of default. Determines (i) whether to accept credit application and (ii) on what terms (e.g. interest rate, amount lent).
- Has someone paid back in past strong predictor of future behavior.

Why Do Credit Bureaus Exist? The Problem Lenders Face

- Lending is risky.
- If consumer doesn't pay you back (default) you lose money.
- **Information asymmetry** between borrowers and lenders. Jaffee & Russell 1976, Stiglitz & Weiss 1981, Pagano & Jappelli 1993, Jappelli & Pagano 2002)
- Those who you least want to lend to want to borrow most (adverse selection).
- Lenders predict expected risk of default. Determines (i) whether to accept credit application and (ii) on what terms (e.g. interest rate, amount lent).
- Has someone paid back in past strong predictor of future behavior.
- Many credit agreements. e.g. # credit card affects likelihood of repaying debt.

Why Do Credit Bureaus Exist? How Credit Bureaus Help Lenders

- Bureaus solve a co-ordination problem.
- Positive externalities from sharing data.
- Limit free-riding.

Why Do Credit Bureaus Exist? How Credit Bureaus Help Lenders

- Bureaus solve a co-ordination problem.
- Positive externalities from sharing data.
- Limit free-riding.
- On consumer side mean non-lender specific reputational impact to non-payment.
- May increase incentive to repay (i.e. reduce moral hazard of strategic default)

What Do Credit Bureaus Do?

- Collate information provided by lenders about consumers.
- Sell information back to lenders (/other firms) about consumers.
  - For evaluating credit applications
  - For targeting marketing
- Being able to track consumers e.g. across addresses, name changes, lenders.
- Standardization in how firms report (not completely though)

Matching Technology

- Large part of credit agencies' business is tracking a consumer over time e.g. name & address changes.

Matching Technology

- Large part of credit agencies' business is tracking a consumer over time
  e.g. name & address changes.
- Important as consumers with poor credit have incentive to 'hide' bad credit.

Matching Technology

- Large part of credit agencies' business is tracking a consumer over time
  e.g. name & address changes.
- Important as consumers with poor credit have incentive to 'hide' bad credit.
- Keep history of names, addresses, social security numbers (SSNs), dates of birth.
  - Proprietary algorithms for matching data.

Matching Technology

- Large part of credit agencies' business is tracking a consumer over time
  e.g. name & address changes.
- Important as consumers with poor credit have incentive to 'hide' bad credit.
- Keep history of names, addresses, social security numbers (SSNs), dates of birth.
  - Proprietary algorithms for matching data.
- NOT EVERYONE has a credit file (approx. 11%).

Matching Technology

- Large part of credit agencies' business is tracking a consumer over time
  e.g. name & address changes.
- Important as consumers with poor credit have incentive to 'hide' bad credit.
- Keep history of names, addresses, social security numbers (SSNs), dates of birth.
  - Proprietary algorithms for matching data.
- NOT EVERYONE has a credit file (approx. 11%).
- SOME PEOPLE have multiple credit files ('fragmented') where can't consolidate.
- Need to be aware of this and can introduce criteria to restrict to higher-quality sample
  (e.g. dates of birth, social security numbers, present in tradeline data).

Credit Data

Credit Data

- Observe most household debts.

Credit Data

- Observe most household debts.
- Do not observe assets, incomes (beyond crude estimates), expenditures (beyond credit repayments), most demographics.

Credit Data

- Observe most household debts.
- Do not observe assets, incomes (beyond crude estimates), expenditures (beyond credit repayments), most demographics.
- Good coverage of main credit, bad coverage of subprime credit
  e.g. payday loans (though some 'alternative data' on this).

Credit Data

- Observe most household debts.
- Do not observe assets, incomes (beyond crude estimates), expenditures (beyond credit repayments), most demographics.
- Good coverage of main credit, bad coverage of subprime credit e.g. payday loans (though some 'alternative data' on this).
- Observe date of birth, geography.

Credit Data

- Observe most household debts.
- Do not observe assets, incomes (beyond crude estimates), expenditures (beyond credit repayments), most demographics.
- Good coverage of main credit, bad coverage of subprime credit
  e.g. payday loans (though some 'alternative data' on this).
- Observe date of birth, geography.
- Can construct variety of economically-useful measures.

Credit Data Shows a Limited History of Consumers

Credit File Information	Reporting Duration
Hard Credit Inquiry	2 years from inquiry date
Open Credit Agreement	Indefinitely
Closed, Non-Delinquent Credit Agreement	10 years from agreement's last activity
Delinquent Credit Agreement	7 years from payment first 30 days past due
Foreclosure Public Record	7 years
Bankruptcy (Chapter 13)	7 years
Bankruptcy (Chapters 7, 11, 12)	10 years

Credit Data

Most researchers use data 2000s onwards.

- Earlier available but less accessible / lower quality.

- Public Records
- Collections
- Tradelines (this is rawest data and so most useful IMHO)
- Customs/roll-ups including Credit Scores
- (Inquiries)

Many papers written only using quarterly or annual customs data.

Public Records

- **Bankruptcy**
  - Gross, Kluender, Liu, Notowidigdo, & Wang (2021) *American Economic Review* ('AER') study effects of bankruptcy reform.
  - Gross, Notowidigdo, & Wang (2020) *American Economic Journal: Macro* and Dobbie, Goldsmith-Pinkham, & Song (2020) *Journal of Finance* ('JF') study credit, consumption, & labor market effects of removal of bankruptcy flags.

Public Records

- **Bankruptcy**
  - Gross, Kluender, Liu, Notowidigdo, & Wang (2021) *American Economic Review* ('AER') study effects of bankruptcy reform.
  - Gross, Notowidigdo, & Wang (2020) *American Economic Journal: Macro* and Dobbie, Goldsmith-Pinkham, & Song (2020) *Journal of Finance* ('JF') study credit, consumption, & labor market effects of removal of bankruptcy flags.
- **Other Public Records:** Foreclosures, Tax Liens.
  - Fulford & Nagypal (2022) exploit 2017 law change removing public records and find it ↑ credit access due to redistribution.

- **Medical Debt**
  - Dobkin, Finklestein, Kluender, & Notowidigdo (2018) *American Economic Review* event study of effects of hospital admissions.
  - Kluender, Mahoney, Wong, & Yin (2021) *JAMA* Guttman-Kenney, Kluender, Mahoney, Wong, Xia, & Yin (2022) *JAMA Health Forum* analyze medical debt with COVID update.

- **Medical Debt**
  - Dobkin, Finklestein, Kluender, & Notowidigdo (2018) *American Economic Review* event study of effects of hospital admissions.
  - Kluender, Mahoney, Wong, & Yin (2021) *JAMA* Guttman-Kenney, Kluender, Mahoney, Wong, Xia, & Yin (2022) *JAMA Health Forum* analyze medical debt with COVID update.

- **Non-Medical Debt**
  - Romeo & Sandler (2022) *Journal of Public Economics*, Fonseca (2022) Journal of Finance) exploit state debt collection law changes to evaluate effects on credit access.

What Determines Consumer Financial Distress?
Place- and Person-Based Factors

Use credit file data to (1) measure mobility for causal analysis (movers) and (2) outcomes.

Large geographic variation in outcomes.

(A) Debt in Collections in Past 3 Years (%)
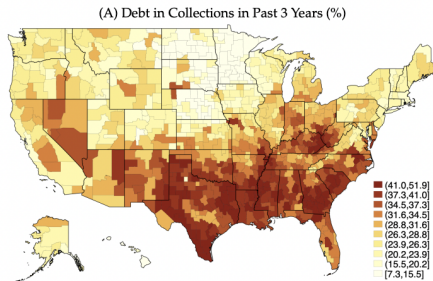
After moving small change in collections (instead person-based), large change in bankruptcy (place-based).
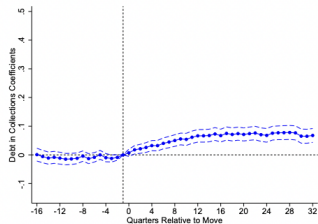
Keys, Mahoney, & Yang (2022) *Review of Financial Studies ('RFS')*.

Tradeline

Account-level data e.g. one observation per credit card held by a consumer per month.

Tradeline

Account-level data e.g. one observation per credit card held by a consumer per month.

- Account opening ('origination') details: Date, amount lent or credit line, duration, anonymized lender (tricky!).
- Account performance: Balance, credit limit, default status, payment amount (sometimes!), closed date.

Tradeline

Account-level data e.g. one observation per credit card held by a consumer per month.

- Account opening ('origination') details: Date, amount lent or credit line, duration, anonymized lender (tricky!).
- Account performance: Balance, credit limit, default status, payment amount (sometimes!), closed date.

What products?

- Auto loans, credit cards (& retail cards) , mortgages (& home equity lines of credit HELOCs), student loans & other unsecured loans.

Tradeline

Account-level data e.g. one observation per credit card held by a consumer per month.

- Account opening ('origination') details: Date, amount lent or credit line, duration, anonymized lender (tricky!).
- Account performance: Balance, credit limit, default status, payment amount (sometimes!), closed date.

What products?

- Auto loans, credit cards (& retail cards) , mortgages (& home equity lines of credit HELOCs), student loans & other unsecured loans.

What's missing?

- Utilities, payday loans (though some alternative credit file data available) and Buy Now Pay Later (BNPL).

16

How Do Individuals Repay Their Debt? The Balance-Matching Heuristic

Use credit file data to explore behavioral economics of heuristics.

How consumers allocate payments across multiple credit cards?
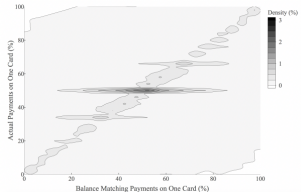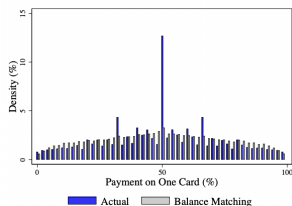
Horse-race comparing models
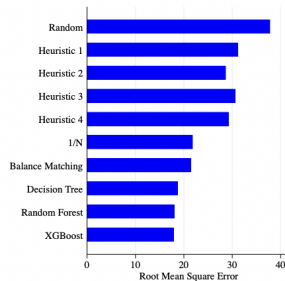
Gathergood, Mahoney, Stewart, & Weber (2019) *Economics Bulletin*
replicates their 2019 *American Economic Review ('AER')*.

Liquidity versus Wealth in Household Debt Obligations: Evidence from Housing Policy in the Great Recession

Use credit file data as measures of consumption (credit card spending, new autos) as outcomes for effects of mortgage relief during Great Recession.
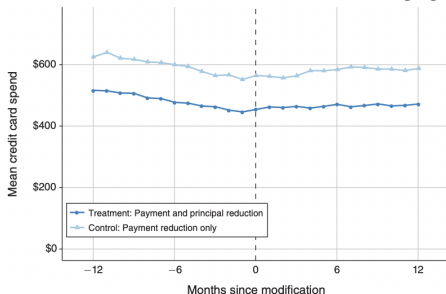

FIGURE 3. EFFECT OF PRINCIPAL REDUCTION ON CONSUMPTION

N.b. Credit card measure becomes highly selected post-2015 due to firms stopping reporting payments data (Guttman-Kenney & Shahidinejad, 2022)
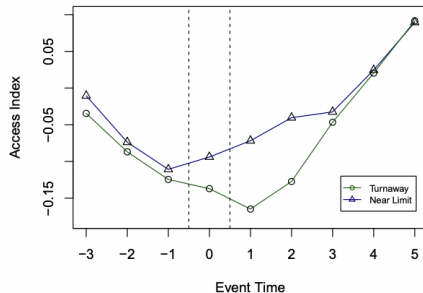
Ganong & Noel (2020) *American Economic Review ('AER')*.

The Economic Consequences of Being Denied an Abortion

Use credit file data as measures of economic outcomes
for people with abortions vs. control group turned away due to gestational limits
(differences-in-difference).
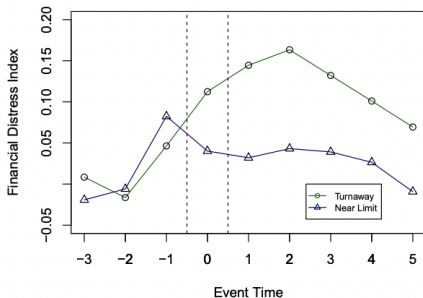
(a) Financial Distress Index   (b) Credit Access Index

Miller, Wherry, & Forster (2019) *American Economic Journal: Economic Policy ('AEJ: EP')*.

Customs / Roll-ups

Consumer-level e.g. one observation per credit card held by a consumer per month.

- Geography
- Credit scores
- Aggregates of other datasets (e.g. number of open tradelines, number of open tradelines in default)

Credit Scores

- Credit score takes raw data and predicts default.

Credit Scores

- Credit score takes raw data and predicts default.
- Less likely to repay (higher credit risk), lower credit score, less credit access.

Credit Scores

- Credit score takes raw data and predicts default.
- Less likely to repay (higher credit risk), lower credit score, less credit access.
- Designed to be a relative risk of default based on logit models (now using ML techniques).

Credit Scores

- Credit score takes raw data and predicts default.
- Less likely to repay (higher credit risk), lower credit score, less credit access.
- Designed to be a relative risk of default based on logit models (now using ML techniques).
- There are many types of credit scores.
- FICO, VantageScore most common proprietary algorithms used bylenders.
  - Different versions of these e.g. FICO 5, 8; VantageScore 3.0, 4.0 etc.
  - Versions targeted to particular markets e.g. autos, mortgages.
- Scales differ but all highly correlated.
- Can use any and often using ventiles (or similar) or groups (subprime, prime etc) based on thresholds of credit scoring company. Calculating your own credit score is fine (sometimes better), hacking FICO is not!

Credit Scores

- Credit score takes raw data and predicts default.
- Less likely to repay (higher credit risk), lower credit score, less credit access.
- Designed to be a relative risk of default based on logit models (now using ML techniques).
- There are many types of credit scores.
- FICO, VantageScore most common proprietary algorithms used bylenders.
  - Different versions of these e.g. FICO 5, 8; VantageScore 3.0, 4.0 etc.
  - Versions targeted to particular markets e.g. autos, mortgages.
- Scales differ but all highly correlated.
- Can use any and often using ventiles (or similar) or groups (subprime, prime etc) based on thresholds of credit scoring company. Calculating your own credit score is fine (sometimes better), hacking FICO is not!
- Some cutoffs are particularly important e.g. FICO 660 determines if US mortgage is eligible for government securitization.

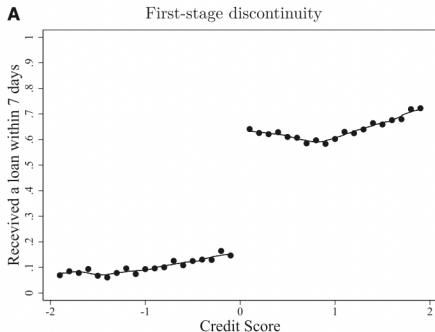## How Do Payday Loans Affect Borrowers? Evidence from the U.K. Market

1. Using **credit score** discontinuities as a source of variation for causal analysis (regression discontinuity design)

2. Using **credit bureau** data to measure consumer outcomes.

- Data: Payday loan transactions (90% applications, 99% approvals) & credit bureau.

Gathergood, Guttman-Kenney, & Hunt (2019) *Review of Financial Studies ('RFS')*.

Payday lenders have credit score cutoffs for approving applications:

Use Regression Discontinuity Design (RDD) to estimate causal effects:

Probability of any payday loan from any lender after first payday loan application

A — First-stage discontinuity

Probability of any missed payment 6-12 months after first payday loan application

C — 6–12 months after
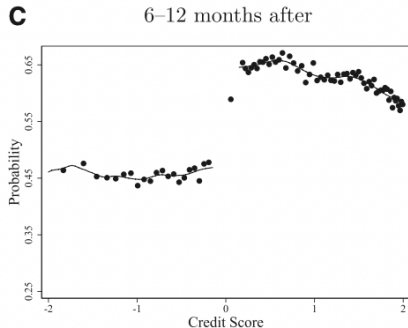
Gathergood, Guttman-Kenney, & Hunt (2019) *Review of Financial Studies.*
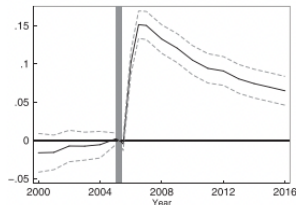
Long-run Net Distributionary Effects of Federal Disaster Insurance: The Case of Hurricane Katrina

Use credit file data to (1) measure mobility and (2) credit outcomes.
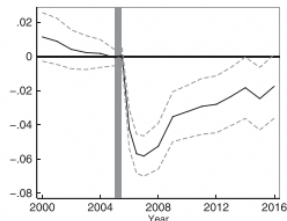
Hurricane Katrina caused exodus from New Orleans.

(a) Pre-Katrina Address

Reduced mortgage access for those most affected.

(a) Has Mortgage

Bleemer & van der Klauuw (2019) *Journal of Urban Economics ('JUE')*.

Using Credit Data for Your Research

Lots of Other Cool Household Finance Datasets

- JP Morgan Chase Institute - e.g. see Peter Ganong & Pascal Noel. Olivia Kim for recent JMP using it.
- Housing & Mortgage Datasets (HMDA, LPS, CRISM, CoreLogic, ATTOM)
- FinTech data (e.g. Mint, Yodlee)
- Spending data (e.g. MasterCard, VISA, Affinity, Earnest Research, Fable Data, SafeGraph)
- Credit offers (e.g. Mintel Comperemedia, Competiscan)
- 'Alternative' credit data covering payday loans (e.g. Clarity, FactorTrust, Veritec) or (e.g. Esusu, Experian Boost, Experian RentalBureau, TurboTenant).

At U.Chicago / Booth check out Fama-Miller Centre (FMC) but also other centres e.g. BFI, Kilts, Center for Applied AI.

Recent Examples of Job Market Papers (JMPs) Using Credit Data

- Nelson (2017) credit cards.
- Mello (2018) traffic fines.
- Buchak (2019) gig economy.
- Shahidinejad (2022) credit unions.
- (My JMP, credit reporting).

How Can Credit Data Help You?

- Sources of variation (e.g. removal of bankruptcy flags)
- Credit outcomes (e.g. credit access, delinquency)
- Non-credit outcomes (e.g. mobility, consumption)
- Matching technology (e.g. 2+ datasets with names & addresses how to match)
- Particular value merging credit files with other datasets (individual or regional levels)
  - hard work but high returns if pull off!

Where to access?

- CFPB and Federal Reserve have long-running panels.
- California Policy Lab has panel available to researchers at Universities of California researchers.
- Ad-hoc purchases or some historical panels (e.g. Stanford, Wharton, UIUC)

TransUnion Data at Booth

- Approx. 25 active users (last I heard).
- A great resource!
- 10% sample of US consumers with monthly credit reports July 2000 - 2022 (updated).
- Booth researchers can:
  - Write papers with these without cost.
  - Merge other datasets at cost.
- Access managed by Kilts Centre for Marketing. Submit research proposals (PhDs need faculty supervisor) for TU to approve and then TU reviews finished paper for public release.
- Vulcan hadoop cluster for analyzing.

GitHub

'Corporate Memory' / Institutional Knowledge

- Documentation
- Shared code
- Troubleshooting

(if you are approved by Kilts email me for access)

> **Advice**
> Become an institutional expert! If you are the person who knows most about a dataset people want to collaborate with you.

Thank you!

Please do reach out! Happy to skim your PhD research statements, discuss my PhD experiences, household finance research etc.

✉ **benedict@chicagobooth.edu**

🐦 @gk_ben

🌐 **www.benedictgk.com**