

# The CFPB Consumer Credit Panel: Direct Use and as a Sampling Frame

June 12, 2015



# About CFPB's CCP

---

- 5 million anonymized credit records
  - Excludes personally identifiable information
  - This is a 1-in-48 random sample of records in Dec 2012
  - Panel data at the individual level
- From one of the three large U.S. credit bureaus
- Extends from 2001 to 2015
- Credit scores from two providers
- For further info, see GAO's 2014 report

Differences from NY Fed's CCP:

- Different provider
- Data at the level of the individual tradeline

# Data Point: Medical Debt and Credit Scores

## May 20, 2014

Data point:  
Medical debt and credit  
scores

- Credit score is a tool to predict whether a borrower will default
- Traditionally, scoring models have not distinguished between medical and nonmedical debts
- Data Point showed that medical debt has less predictive power for future defaults than nonmedical debt
  - After release, credit rating agencies announced they would adjust scoring to account for these results
  - FICO 9, introduced a few months later, downweighted medical debt

# Scope: Credit Reporting and Collections

---

- Collections tradelines affect many consumers.
  - 31.6% of consumers with credit reports have one or more collections tradelines

Incidence of Collections tradelines by type	Percentage of consumer credit reports containing one or more collections tradelines originating from...
Medical or health care	19.4%
Cable, cellular, wireless, other telecommunications	8.7%
Utilities or energy	7.6%
Retail collections	6.9%
Banking	2.7%
Financial	1.5%

- Significant impact to Consumer Credit Scores
  - One medical collection causes FICO 8 credit score of 680 to drop by over 40 points; 780 score to drop by over 100 points

# Score Differentials, Medical v. Non-Medical

## Green Medical

Any: 8.4

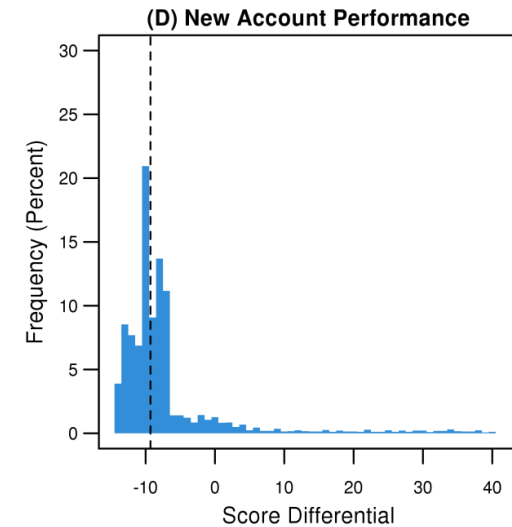
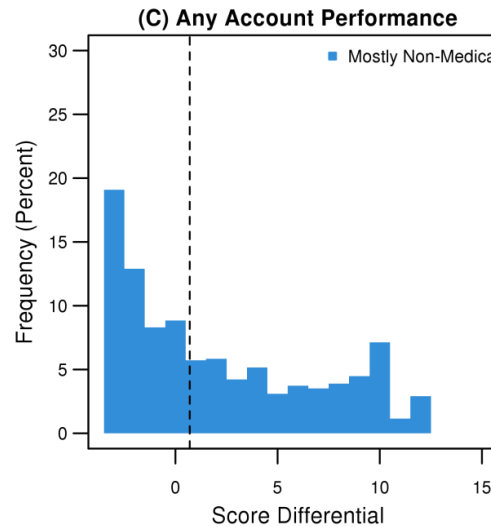
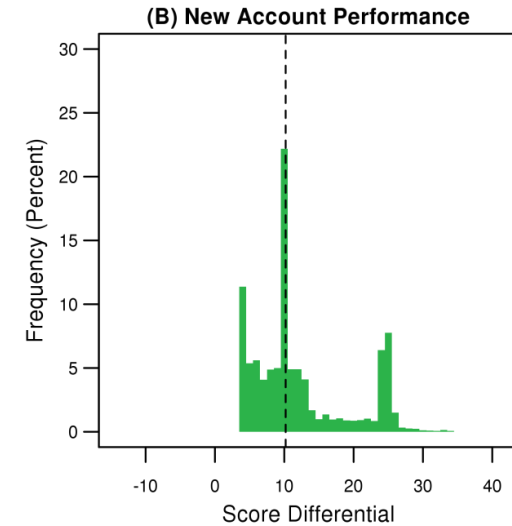
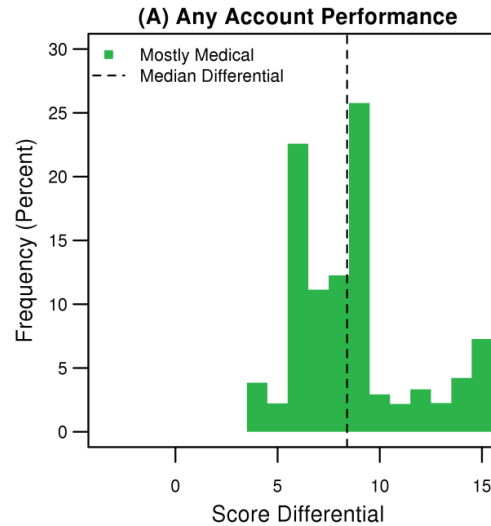
New: 10.2

## Blue

Non-Medical

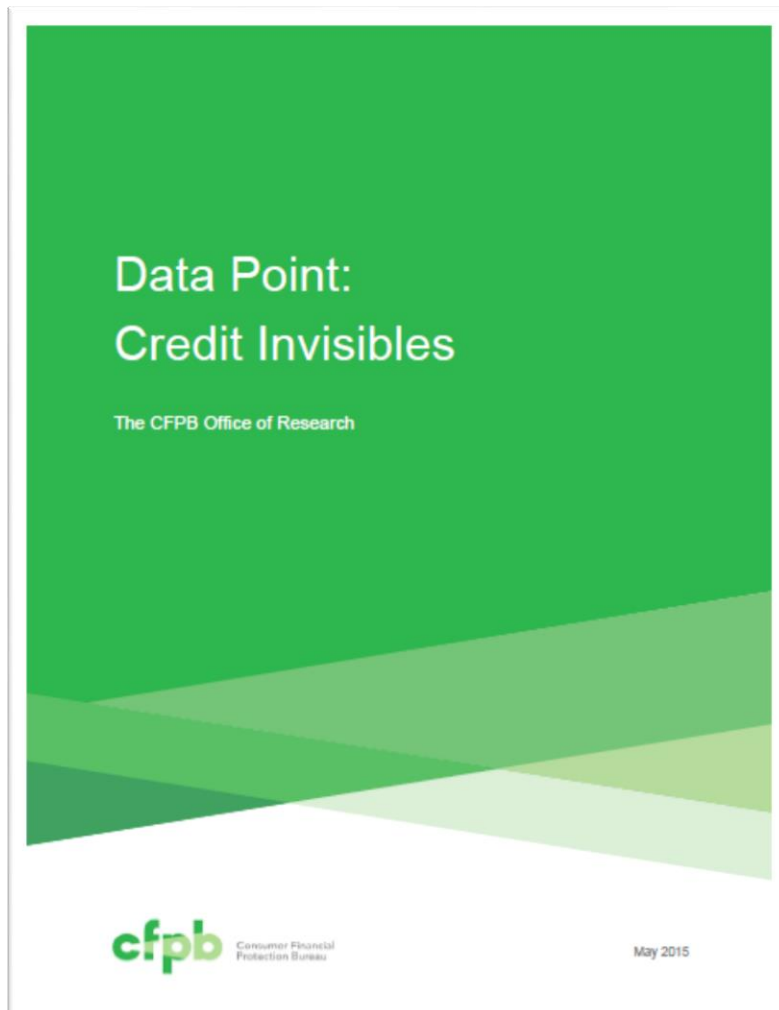
Any: 0.7

New: -9.3



# Data Point: Credit Invisibles

## May 5, 2015



Looks at the number and characteristics of consumers who:

- Have no credit record (“credit invisibles”)
- Have a credit record that cannot be scored
  - Insufficient-unscored
  - Stale-unscored

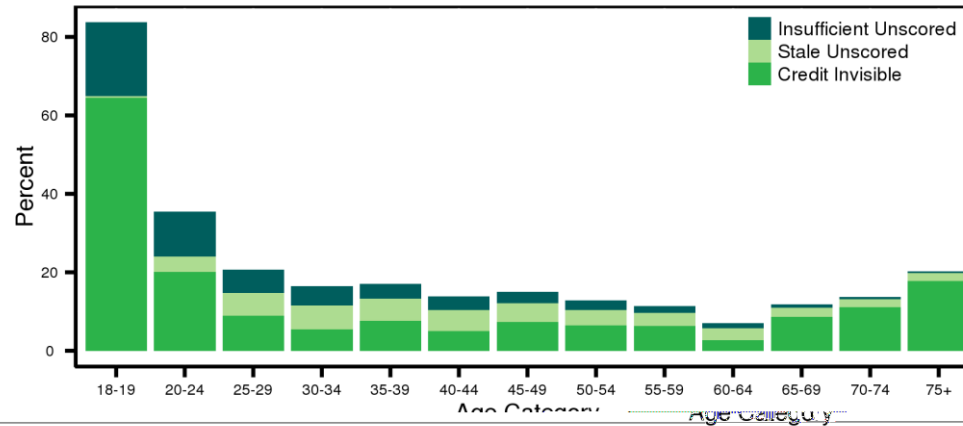
# How Can We See People Who Are Invisible?

---

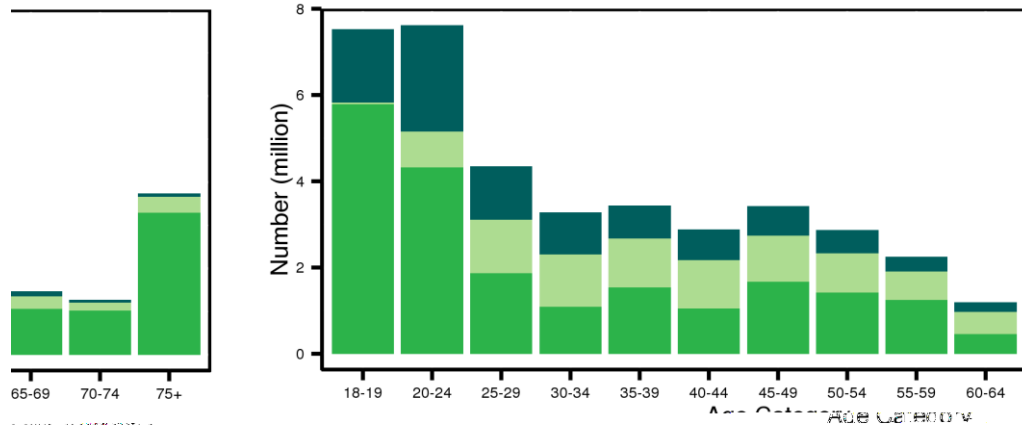
- By comparison with “truth” as measured by the Census
  - Our measure is inaccurate in exact proportion to Census
- Allows breakdowns by categories as fine as (but no finer than) those reported by Census

# By Age

(A) Share of Consumers that are Credit Invisible or Unscored

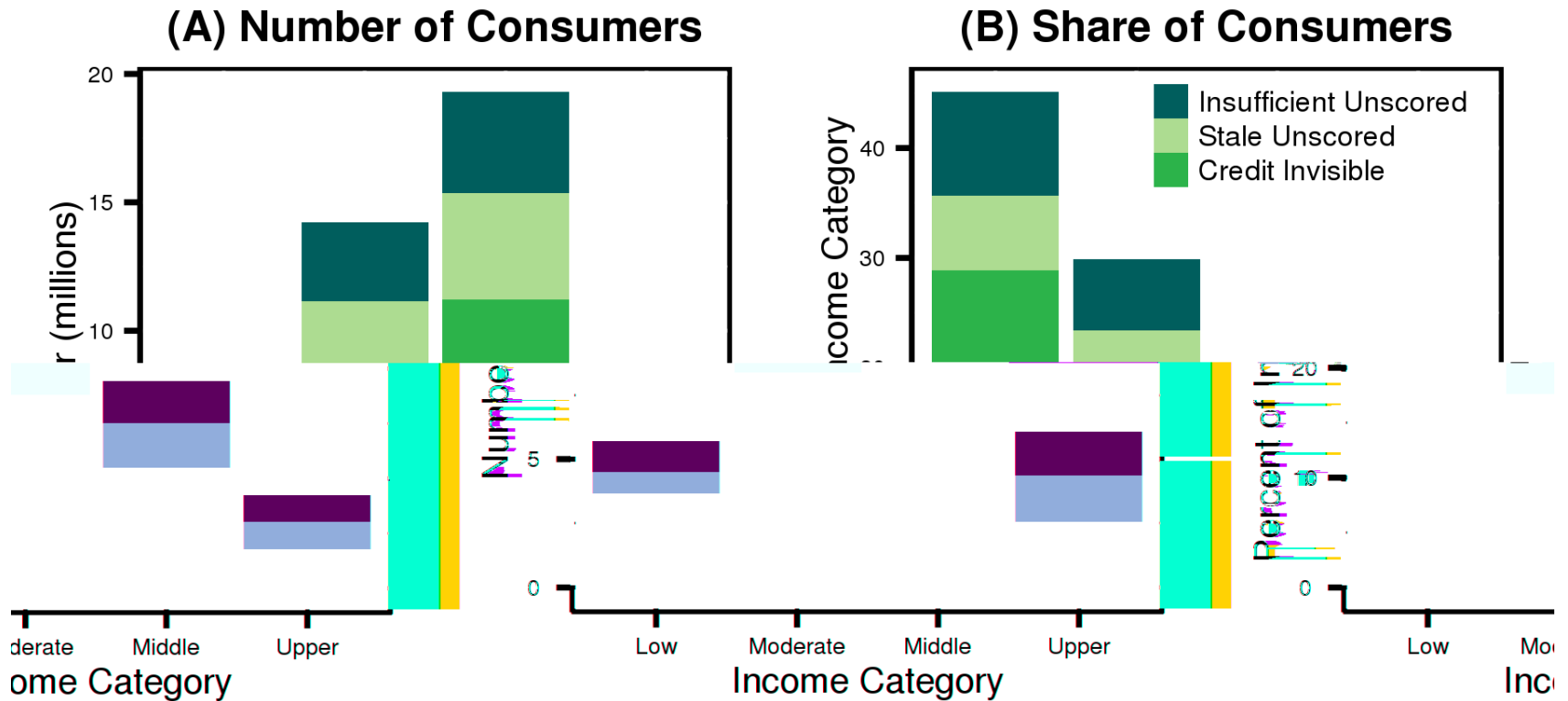


(B) Credit Invisibles and Unscored



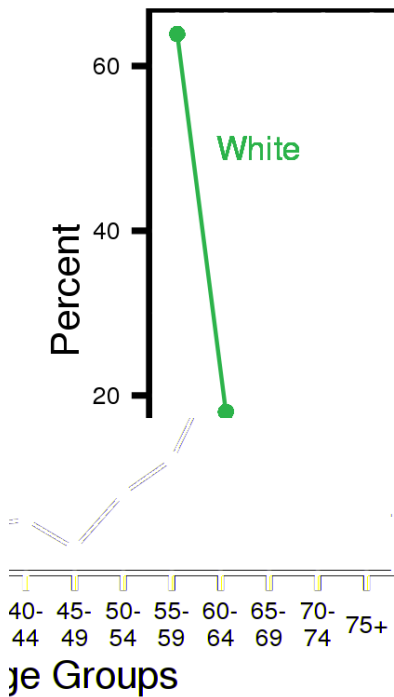


# By Income Level

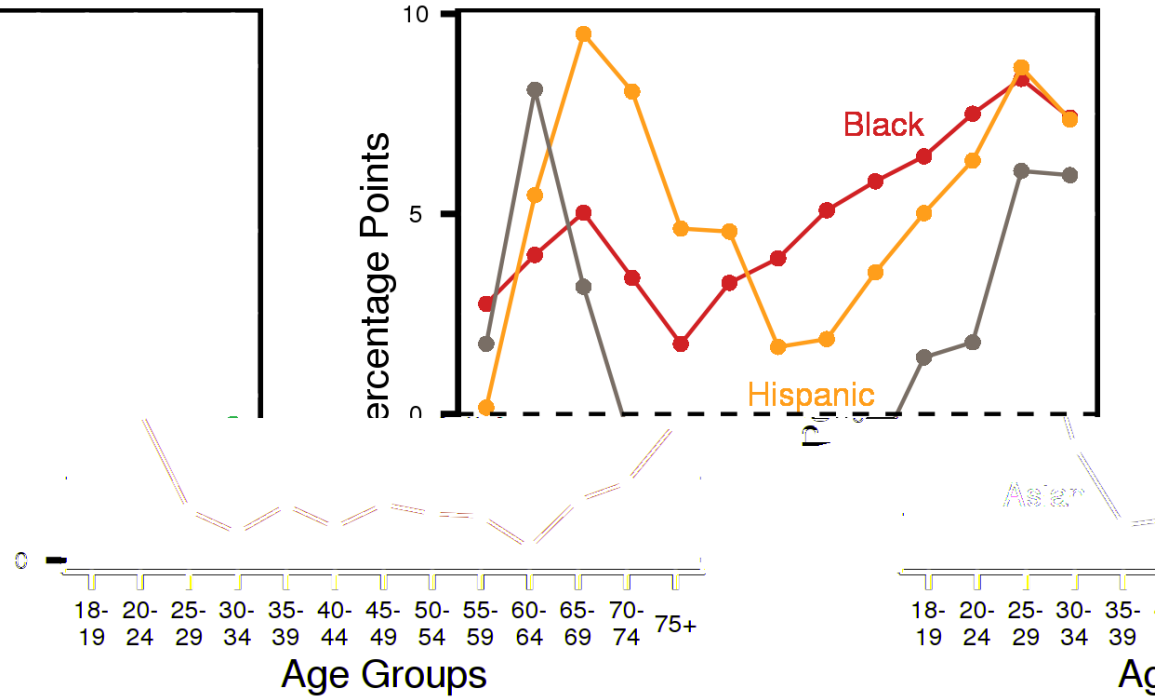


# By Race/Ethnicity & Age

**(A) Credit Invisibles**



**(B) Differentials**



# Surveys Using CCP as Sampling Frame

---

- Groves (2006)
  - Surveys based on sampling frame from administrative data permit modeling of nonresponse patterns
  - This may be able to help understand (and possibly correct for) NR bias
  - Permits combination of rich quantitative data with qualitative/perceptual data
- Broad coverage, rich info => CCP a natural frame for many questions relating to consumer financial circumstances

# Mechanics of Surveys from the CCP

---

1. CFPB Designs Survey (writes questions, etc)
2. Design sent to OMB for review
3. OMB vets
  1. Reviews survey design for quality and legal issues
  2. Publishes notice in Federal Register soliciting public comments
  3. After comment period closes, comments go to CFPB
4. CFPB revises survey in light of public comments
5. Resubmits to OMB for second round review
6. If all goes well, OMB approves revised survey
7. CFPB identifies sample to be surveyed from anonymized records
8. Sends sample info to CCP provider
9. Provider subcontracts with surveying firm to administer survey
10. When survey is done, anonymized results are returned to CFPB

# CCP's Virtues as a Sampling Frame

---

- Stratification
  - Can oversample populations of particular interest
  - Results made representative of population by weighting
- Analysis of nonresponse bias
  - Surveys often know little about characteristics of NR's
  - Permits comparison of R's to NR's in many dimensions
  - Can construct rich model of selection on observables
- No need to ask R's questions for which we already have data
  - For many such Q's, admin data more reliable anyway
  - Expands time for qualitative Q's not captured by admin data
  - R's may be more open to answering other questions if not asked the Q's whose answers we know from admin data
- Can relate survey responses to subsequent outcomes

# CCP's Limits as a Sampling Frame

---

- Initial contact must be by mail
  - Though survey can be taken by web using link on initial mailing, if R prefers that
  - Slower than electronic (internet; phone) surveys
- Use of credit record data imposes some release limitations
  - Nothing that compromises identity
  - Some other contractual restrictions (so far, not a problem)
- Persons without CR's are missing from the sampling frame
  - “Credit Invisibles”

# Examples of Use of CCP as Sampling Frame

---

## National Survey of Mortgage Borrowers

- Quarterly survey of borrowers with new mortgages
  - Near universal reporting of mortgages to CRA's
  - No need to massively oversample to find target surveyees
- Credit history prior to mortgage may be useful for many questions
- Of 15,000 recipients of survey, over 5000 responded
  - Very high response rate

# National Survey of Mortgage Borrowers

---

## Initial Results:

- Half of borrowers consider only a single lender or mortgage broker
  - Interesting given results in Hall and Woodward (2012)
- R's say the lender or broker is their primary source of information
- Proximity and reputation are important in choosing lender
- Unexplored so far: Information about expectations, life events, demographic characteristics
  - These are not in the CCP

## Future Plans:

- NSMB to be an ongoing effort
- One motivation:
  - Ability to head off crises like 2008-2009 by being able to see troubling developments as they occur



# Example 2: Debt Collection Survey

---

- Complaints about collection activities are regularly the biggest category in CFPB Consumer Complaints data
- But complaints only reflect experiences of persons motivated strongly enough to complain; cannot be used to get a rounded picture of the typical experience
- We surveyed more than 10,000 consumers
- Sampling frame chosen to obtain dense information on consumers likely to be in collections, while permitting estimates of the prevalence of the experiences surfaced in complaints data
  - No need to massively oversample to get sufficient sample size
- Results will be used for any future CFPB policymaking exercise affecting debt collection activities

# The CFPB's CCP: A Powerful Tool for "Evidence Based Policymaking"

---

- Direct Uses
  - Rich data on liabilities side of balance sheet, credit scoring, much else
- As a Sampling Frame
  - Generic PRA approvals are done
  - Comparatively rapid production of new survey results
  - Panel history of respondents may provide important insights