CALIFORNIA POLICY LAB

Technical Appendix: California Credit Dashboard

SARAH HOOVER, STEVE RAMOS, EVAN WHITE

The California Credit Dashboard provides detailed information about household finances and debt, information that has not historically been available to the public at the state-level. Our data come from the University of California Consumer Credit Panel, a dataset of credit-bureau data provided by Experian, one of the three nationwide credit reporting agencies. This technical appendix provides more information about the methodology and data used in the dashboard.

Sample. All figures are based on a 2% random sample of individuals. Where applicable, we multiply the results by 50 to obtain the full-population estimate. Not all Californians have access to credit, so one might expect the number of consumers in the UC-CCP sample to be less than the state population. However, when comparing credit bureau data population counts to population estimates from the American Community Survey (ACS), the New York Federal Reserve Bank reported finding the opposite in some cases. Thus, when comparing the n values from our figures to other data sources, like the Census, there will not be perfect alignment. It is important to note that while our reported n values may be overestimating the true n value, the aggregate balances are not impacted.

Open accounts. With the exception of collections, we count balances only for "open" accounts, and do not count any paid, charged off, or otherwise closed accounts.

Other exclusions.

- We exclude three categories of records that follow exclusions explained in the New York Federal Reserve Bank's analyses. First, we drop fragmented (aka "frag") files, which include consumers marked as deceased or who, after Q2 2010, are missing a birth date on their file. Second, we also exclude accounts that have not been furnished by lenders within the prior quarter.

Third, we exclude open balances that are either zero or missing. In both of the latter cases, these accounts are unlikely to still be open.

- We also drop any observations with ages or credit scores that are outside the possible ranges, or people whose available geographic data can't be placed in a region. We check to ensure that each of these exclusions is justified, minor, and does not skew our results.

- We also drop tradelines that are exact duplicates on all relevant measures. This drops approximately 0.1% of tradelines in a given quarter.

Loan types. We construct simplified loan types from a combination of the "account type" and the "kind of business" codes. "Auto" includes both loans and leases. "Other" includes tradelines that are not easily classified into the other categories, and is largely composed of miscellaneous revolving accounts and unsecured installment loans, largely under $15,000.

Credit scores: There are many different types of credit scores. We use one version provided by the credit bureau that ranges from 300-850 (the same as FICO® scores), with five scoring buckets roughly defined as follows: deep subprime (300-580), subprime (580-619), near-prime (620-659), prime (660-719), and super-prime (720-850). The data show some volatility in credit scores between 2006 and 2010

that may reflect actual volatility or may be caused by unreliable scoring data.

Regions. Economic regions are from the [California Economic Strategy Panel Regions](). For most individuals (81% in Q3 2022), we observe both county FIPS code and ZIP code, while for some observations (19%), we only observe a ZIP code. ZIP codes can be split into more than one county, so for those with only a ZIP code on file, we use the individual's FIPS code in prior and subsequent quarters to interpolate their missing FIPS information, provided their ZIP code at the time matches the prior/subsequent ZIP code. We give precedence to the prior location in cases where the two conflict. After doing so, only 11% of observations (down from 19%) are missing a county FIPS code. In these cases, we use HUD's ZIP-county crosswalks to place individuals into counties, based on the county containing the majority of that ZIP code's population. Counties then nest into the regions linked.

Generations. Generations are defined in accordance with the [Pew Research Center.]() The credit bureau did not record dates of birth before June 2010. For consumers that appear both before and after June 2010, we backwards infer their age prior to 2010. All others are listed as "no birthdate." We use birth year from the most recent quarter's data. Consumers are grouped into generations by their year of birth, and the dashboard reports ages for each generation based on a January 1st birthdate.

Weighting accounts. In the data, we are able to observe both primary account holders as well as co-signers, authorized users, and joint accounts. To avoid double-counting certain accounts, we weight the tradeline information for all primary account holders (ECOA codes 1 and 7) as 1, all joint account holders (ECOA codes 2 and 4) as 0.5, and all authorized users and co-signers (ECOA codes 3, 5, and 6) as 0. We apply these weights to total balances and total origination amounts. But we do not apply them to average balances and average origination amounts so that these charts show the actual loan amounts, rather than the value split between joint borrowers. We also do not apply the weights to the N values.

Cell suppression. Cells with less than 10 observations have been suppressed. This is one of many measures we take to protect these anonymized data from potential reidentification.

Data anomalies. Credit-bureau data contain various anomalies, including errors introduced by data furnishers (e.g., lenders). In addition, the credit bureau from which we receive the data underwent systems changes around June 2010 and in some years prior to that, and the data may show discontinuities around those times due to back-end changes in how data were collected and processed. Discontinuities after June 2010 may be due to furnishing errors (e.g., [this]()) or due to changes in how the bureaus report data (e.g., starting in 2023, the bureaus will [stop reporting medical debts of less than $500]()), as well as true discontinuities, such as those caused by [forbearance]() during the pandemic.

Origination projections. Months often elapse before a new loan starts being reported in the credit data. Thus, for recent quarters, observed originations will be underestimates of the true volume. For this reason, we project the most recent 4 quarters of originations data using information about underestimates in that loan type over the prior 3 years.

Seasonal adjustments. The seasonal adjustment is a 4-quarter rolling average of the unadjusted data. For the charts showing averages and rates we use rolling averages weighted by population. For the other figures we use simple averages.

Inflation adjustment. The inflation adjustment uses CPI-U (series CUUR0000SA0) and is normalized to the most recent quarter.

FUTURE ADDITIONS

Credit-bureau data provide a unique and near-real-time perspective on the financial health of Californians. Pending funding availability, we plan to update the California Credit Dashboard on a quarterly basis, as well as provide more in-depth analyses on specific credit issues. We also hope to build new features into the Dashboard, such as tracking new outcomes and expanding the ability to break out the data into smaller geographies and new demographic groupings.

Want to support this effort? The California Policy Lab is seeking philanthropic support to maintain and improve this Dashboard as a tool to inform policymakers and the general public about the financial health of Californians and our state's economy. Please reach out for more information.

ACKNOWLEDGMENTS

The California Policy Lab generates research insights for government impact. We are an independent, nonpartisan research institute at the University of California with sites at the Berkeley and Los Angeles campuses.

This research publication reflects the views of the authors and not necessarily the views of our funders, our staff, our advisory board, Experian, or the Regents of the University of California.