



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 1% random sample of individuals. Where applicable, we multiply the results by 100 to obtain the full-population estimate.

Open accounts. With the exception of collections, we count balances for only “open” accounts, so 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 2010Q2, 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 VantageScores® that are outside the possible ranges, or people that 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.

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.

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 now 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.

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 backend 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 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 Figures 1, 3, and 4, we use averages weighted by population. For the other figures we use simple averages.

Nominal dollars. We currently make no adjustments for inflation.

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 (e.g., foreclosures) 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

This research was made possible through support from Arnold Ventures, The James Irvine Foundation, the University of California Office of the President Multicampus Research Programs and Initiatives, MRP-19-600774 and M21PR3278, and the Bylo Chacon Foundation. The views expressed are those of the authors and do not necessarily reflect the views of our funders. All opinions and errors should be attributed entirely to the authors.

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.