



Technical Appendix to: *Revisiting Student Loan Forgiveness: An Update Based on the White House Plan*

Jacob Goss, Daniel Mangrum, Joelle Scally

Data

New York Fed/Equifax Consumer Credit Panel

For the main analysis, we use the New York Fed/Equifax Consumer Credit Panel (CCP) which is a 5% representative sample of credit reports from Equifax. The data contain detailed information on the debts of U.S. adults including loan types, origination dates, balances, and delinquencies. In addition to data on credit, the data includes other information on credit reports including age, credit score (Equifax Risk Score 3.0), and the Census place identifiers containing the home address. The 5% sample contains 14 million primary sample individuals, of which 2.2 million (15%) have a student loan on their credit report as of the second quarter of 2022.

To gather information about borrower income and demographics, we match each borrower in the CCP to information about their Census block group (CBG).¹ We chose the Census block group as the most granular unit that we can merge with data on neighborhood income and demographics. Block groups divide census tracts, are designed to follow neighborhoods, and generally contain between 600 and 3,000 people.

American Community Survey

We match borrowers' CBG to the 5-year American Community Survey (ACS) for information regarding neighborhood income and neighborhood demographics.

¹For those with missing CBG identifiers, we use zip code for merging to other data.

The 5-year ACS is a nationally representative survey that incorporates multi-year estimates to improve the reliability of statistics in areas with small populations. For current income, we use the 2014-2019 ACS, and for historical income, we use the 5-year samples from the 2005-2009 waves through the 2015-2019 ACS. We use IPUMS NHGIS to access each wave. We take median household income by Census block group to split borrowers into neighborhood income deciles², ending up with about 20,000 CBGs in each decile. We take populations by race to split borrowers into two groups, those in CBGs with at least 50% of residents identifying as white non-Hispanic (majority white), and those in CBGs with less than 50% of residents identifying this way (majority minority), classifying 148,707 CBGs as majority-white and 71,625 as majority-minority.

Federal Reserve Bank of New York's Survey of Consumer Expectations Credit Access Module

We use data from the Federal Reserve Bank of New York's Survey of Consumer Expectations (SCE) to estimate the probability each student loan borrower is under the income cap based on credit score and age. The SCE is a monthly, nationally representative, internet-based survey of around 1,300 heads of households on their expectations for economic outcomes. The SCE offers an additional module on credit access every 4 months. We take data from the 2017-2019 SCE and gather respondents' age and self-reported income from the main survey, and merge with the Credit Access Module for respondents' self-reported credit scores and student loan status. Income is reported in the first month surveyed, and we take age and survey weight from this observation as well. Respondents' credit scores are generally not available the first month that they are surveyed and scores may change over time, so we take the non-missing first credit score reported in order to most closely match the respondents' score when income was reported. The main sample has 4,086 respondents with no variables missing, and after restricting to student loan holders we're left with a sample of 1,201 respondents.

National Post-secondary Student Aid Study

We use data from several waves of the National Post-secondary Student Aid Study (NPSAS) from the National Center for Education Statistics (NCES) to establish the

²The bins are: less than \$33,044, \$33,044 to \$42,190, \$42,191 to \$49,706, \$49,707 to \$56,461, \$56,462 to \$63,922, \$63,923 to \$72,564, \$72,565 to \$83,133, \$83,134 to \$97,377, \$97,378 to \$121,032, and \$121,033 or more

probability of Pell grant receipt based on income, federal student loan borrowing status, and dependency status. The NPSAS is a nationally representative cross-sectional survey every 4 years that collects detailed information on college students including demographic, income, financial aid, and student loan details. We use the PowerStats tool to create tables detailing the probability of Pell grant receipt for college students by income bins and dependency status. We use the same income bins from the ACS 5-year surveys and we filter to include only those students who have borrowed student loans.³ We compute separate probabilities for dependent and independent students. For dependent students, we use parental income to create probabilities by income bin. For independent students, we use combined student and spouse (if any) income. As an example, one cell of this data reports that 85% of dependent students with parental income between \$30,000 and \$34,999 received a Pell grant in 2012. Another example cell reports that 21% of independent students with combined student and spousal income between \$45,000 and \$49,999 received a Pell grant in 2004.

Methods

Sample Construction

We begin our analysis by restricting the CCP sample to only federal student loans owned by the federal government since these are the only loans eligible for forgiveness through this plan. To select such loans, we identify loan servicer sub-portfolios in the CCP data that shifted to automatic forbearance, shifting all loans to current status between February 2020 and May 2020. These servicer portfolios each had delinquency rates above 0% prior to the administrative forbearance event then subsequently had 0% delinquency rates after forbearance. We also include servicers with 100% delinquency rates since these are defaulted portfolios which were not automatically marked current in March 2020. This selection process excludes private student loans and FFEL student loans owned by commercial banks but includes FFEL loans owned by the federal government. Through this selection, we arrive at an implied 38 million unique borrowers with federal student loans totaling \$1.42 trillion as of the second quarter of 2022. This aligns closely

³The bins are: 0 to \$9,999, \$10,000 to \$14,999, \$15,000 to \$19,999, \$20,000 to \$24,999, \$25,000 to \$29,999, \$30,000 to \$34,999, \$35,000 to \$39,999, \$40,000 to \$44,999, \$45,000 to \$49,999, \$50,000 to \$59,999, \$60,000 to \$74,999, \$75,000 to \$99,999, \$100,000 to \$124,999, \$125,000 to \$149,999, \$150,000 to \$199,999, and \$200,000 or more

to the total outstanding for the “federally-managed” portfolio from the Federal Student Aid portfolio, which total \$1.476.

Estimating Income Eligibility

The White House plan allows for loan forgiveness for borrowers whose 2020 or 2021 income was less than \$125,000 for individuals or \$250,000 for households. Since borrower income and tax filing status are not available in the CCP, we estimate the probability each borrower will qualify for forgiveness using details about their location, age, and credit score. We derive these probabilities from two sources. First, we use the household income distribution at the Census block group level from the ACS. We approximate the probability of income eligibility by taking the share of individuals in each Census block group whose household income is less than \$200,000. We choose the \$200,000 threshold because this is the nearest threshold available in the data to the weighted average of \$125,000 and \$250,000 based on the share of tax-filers filing individual versus joint or head-of-household returns.

Next, we use the SCE Credit Access Module 2017 to 2019 waves to compute a similar probability according to age, self-reported credit score, and student loan borrowing status. Credit scores are binned into five mutually exclusive categories: less than 620, 620 to 659, 660 to 719, 720 to 759, and greater than 760. We also categorize borrowers into five bins by age: less than 30, 30 to 39, 40 to 49, 50 to 59, and greater than 60. We restrict the sample to only those respondents with outstanding student loans. We then compute the share of respondents with student loans whose household income is below \$200,000 within each age by credit score bin. For example, the estimates suggest that 97% of student loan borrowers under 30 with a credit score between 620 and 659 are income eligible for forgiveness while 84% of student loan borrowers between 40 and 49 with credit scores above 760 would be income eligible.

Lastly, we create an overall probability of being income eligible for forgiveness by averaging the probability based on Census block group income distributions and the probability based on age and credit score. We use these probabilities to compute the expected total amount of forgiveness under the White House plan. We estimate that 5.1% of federal student loan borrowers would be excluded from forgiveness due to income.

Estimating Pell Grant Eligibility

For each borrower in our sample, we estimate the probability of having ever received a Pell grant using the income distribution of each borrower's Census block group around the time of their first federal student loan disbursement combined with the probability of Pell receipt by income and dependency status. This method has the advantage of using information on income during college rather than using current income as other studies use. We begin by identifying the first instance of federal student loan borrowing for each borrower and collect the Census block group identifier during that year and the year prior (typically when students apply for federal aid). If the borrower did not have a credit report in the year prior, we use the neighborhood identifier for the year following. We then merge the household income distribution for each Census block group from the ACS 5-year for the nearest year available. Since the ACS 5-year is only available for 2009 and later, we use the income distribution for 2005-2009 for borrowing events before 2009. Similarly, we use the income distribution from the 2016-2020 ACS for borrowing events after 2020.

Next, we use data from NPSAS survey waves on the probability of Pell receipt by income bins and dependency status to estimate the probability a borrower received a Pell grant during a borrowing event. We use the 2000 wave for borrowing events in 2001 and before, the 2004 wave for 2002 through 2005, the 2008 wave for 2006 through 2009, the 2012 wave for 2010 through 2013, and the 2016 wave for events in 2014 and later. We assign borrowers in the CCP as dependent students if their borrowing event occurred at age 23 or younger and independent status otherwise.

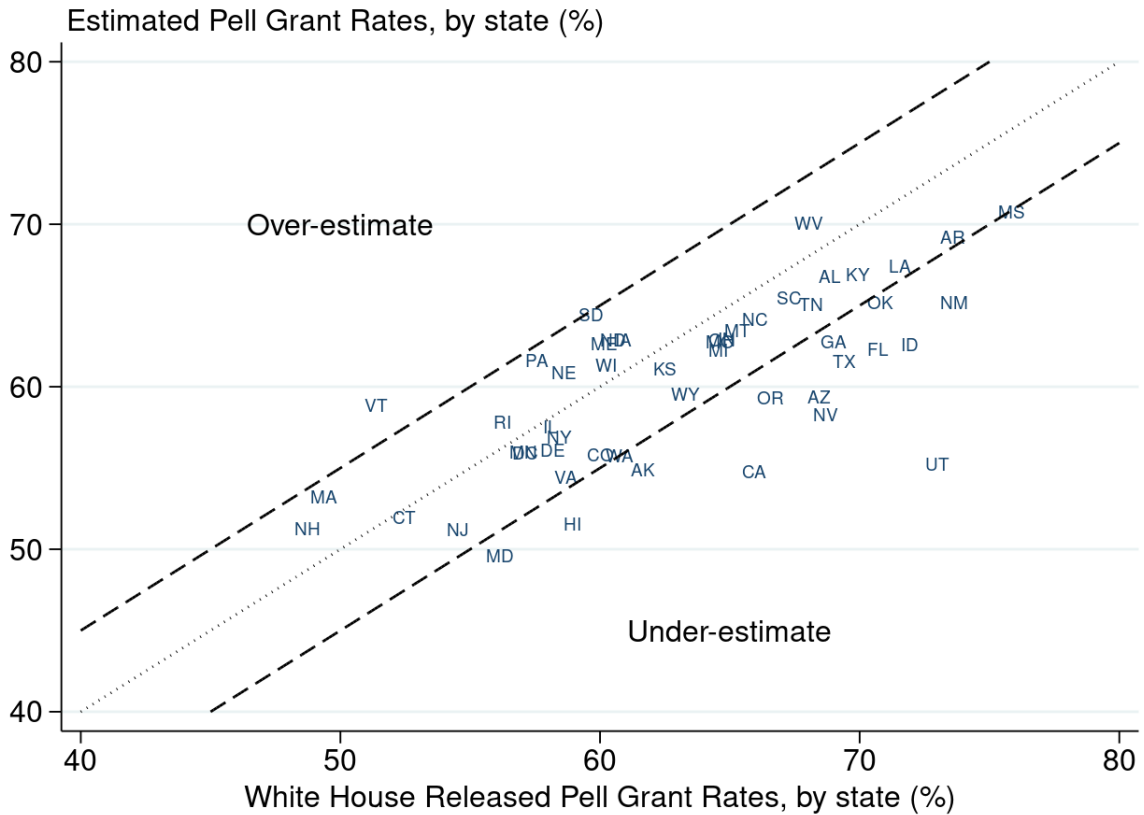
We use simulations to combine data on neighborhood household income distributions with data on the probability of Pell grant receipt by income and dependency status to arrive at a probability that each borrower ever received a Pell grant. For each simulation, the algorithm proceeds as follows:

1. Each borrower has two opportunities to receive a Pell grant for each simulation. The year of the first federal student loan disbursement is the first. The second is the year prior if a credit report exists, but the year after otherwise.
2. Two random numbers are drawn from a uniform distribution between 0 and 1. Each draw assigns a borrower into an income bin according to the distribution of household income for their Census block group for the corresponding year.

3. The assigned income bin, year of borrowing, and age at borrowing then maps to a probability of Pell grant receipt from the NPSAS survey.
4. A second random draw decides whether the borrower receives a Pell grant for that borrowing year according to the matched probability of Pell receipt.
5. If either of the two opportunities results in Pell grant receipt, this borrower is marked as a Pell recipient for the simulation.

We run 200 simulations of this algorithm for our 1.9 million observations and estimate the probability of having ever received a Pell grant by computing the share of the simulations that result in at least one Pell grant. The White House released a statistic that 60% of federal student loan borrowers ever received a Pell grant. Our algorithm performs remarkably well to match this statistic as we arrive at an implied cumulative Pell grant rate of 59.6%. Additionally, we can compare Pell grant rates across states using newly available data from the White House that reports the estimated number of eligible borrowers and the estimated number of eligible Pell grant recipients. In the figure below, we plot the White House eligible Pell grant rate for each state on the x-axis and our estimated eligible Pell grant rate for each state on the y-axis. 34 of the states (including DC) have predicted Pell grant rates within 5 percentage points of the White House rates. At least one contributing factor to our error in the Pell rate is variation in household size. Pell grant rates are higher for larger households due to the Pell grant formula. We underestimate Pell rates in the states with the largest households (Utah) and overestimate in the states with the smallest households (Vermont). The scale of our error does not drastically change the aggregate cost estimate or the distribution of beneficiaries. If we adjust each borrower's probability of Pell grant by the state level difference between our estimate and the White House Pell grant rate, the total value of cancelled loans increases from \$441 billion to \$449 billion.

Figure 1: State comparison of White House reported Pell grant rates (x-axis) against estimated Pell grant rates (y-axis)



Notes: White House Pell grant rates come from a recent release that details the total number of federal student loan borrowers and the total number of federal student loan borrowers who ever received a Pell grant. We plot these rates against our own estimates of Pell grant rates as detailed above. The lower-right area denotes states in which we under-estimate Pell grant rates while the upper-left area denotes states that we over-estimate Pell grant rates. The dotted line denotes 45-degree line and the two dashed lines denote a 5 percentage point window above and below an exact match.