Predicting postsecondary attendance by family income in the United States using multilevel regression with poststratification

Benjamin T. Skinner*
University of Florida

William R. Doyle
Vanderbilt University

25 January 2024

Abstract

Despite billions of dollars spent yearly to fund higher education for low-income youth, no government agency tracks how many low-income young people attend college by state. Whereas proxy measures like Pell grant receipt address the number of already enrolled low-income students, direct estimates from U.S. Census surveys likely overestimate low-income youth enrollment due to their design. Using Bayesian multilevel regression with poststratification (MRP) to estimate postsecondary attendance rates by family income in each of the 50 states and the District of Columbia, we find substantial variation in attendance rates between income groups across the country.

Keywords: college access, low income, multilevel regression with poststratification, MRP

JEL: C11, C18, I22, I23, I24, I28

* *Corresponding author*: btskinner@coe.ufl.edu; (352) 273-4296. Benjamin Skinner is an assistant professor of higher education and policy at the University of Florida. Will Doyle is a professor of public policy and higher education at Vanderbilt University. All data used in the project are either publicly-available or can be requested from governmental agencies. For instructions on obtaining data as well as replication files, see https://github.com/btskinner/lowinc_enroll_mrp_rep. Neither author has any disclosures related to this project to report. We thank Richard Blissett, Taylor Burtch, Josh Clinton, Chris Fonnesbeck, Heather Jack, Daniel Klasik, Olivia Morales, and participants at both 2021 ASHE and 2016 AEFP conferences for their helpful comments on various iterations of this paper. All errors and mistakes in interpretation remain our own.

1 Introduction

In the United States, the vast majority of federal, state, and institutional financial aid money goes to individuals from low-income families in order to increase their college participation rates (Abraham & Clark, 2006; Baum & Payea, 2013). Thanks to large-scale longitudinal surveys of high school-aged young people regularly conducted since the late 1970s, we have reliable national and regional level estimates of postsecondary attendance by family income over time (Adelman et al., 2003; Berkner & Chavez, 1997; Bozick & Lauff, 2007). Despite decades of financial aid directed toward low-income youth, estimates from these surveys have consistently shown large gaps in postsecondary attendance by family income and wealth (Lovenheim & Reynolds, 2013). According to the most recent federal longitudinal survey of students transitioning from high school to young adulthood, 49 percent of low-income young people attend some form of postsecondary education compared to 74 percent of middle-income young people (Duprey et al., 2020).

In order to evaluate the efficacy of state-level financial aid policies, it would be beneficial to have state-specific estimates of college attendance by family income. By virtue of their design, however, nationally representative federal surveys like those noted above cannot be used to make inferences at the state level. Looking to other data sources, states with well-structured administrative data systems may be able to track the inverse, the proportion of college students from low-income families currently enrolled in college, $P(low\,income \mid college)$. If their systems span the P-20 pipeline, they may even be able to provide estimates of the proportion of interest, $P(college \mid low\,income)$, among those students who remain within the state system. Should these latter estimates exist, however, they are not widely available. Across the country, we find no systematic measures of postsecondary enrollment by family income at the state level.

Other means of estimating low-income youth enrollment in college have their own problems. To determine income-based eligibility, financial aid programs typically rely on the Free Application for Federal Student Aid (FAFSA), which asks questions about parental income in addition to student income (Dynarski & Scott-Clayton, 2013). Dependent students, even those living apart from the parents or guardians who financially support them, have their need categorized each year by their family's financial situation via the FAFSA. Not all students, however, complete the FAFSA. Whereas some low-income youth lack robust knowledge about the availability of need-based aid and the requirements of the FAFSA, others fail to complete the forms due to their complexity and the range of highly specific financial information they request (Dynarski et al., 2022). Other state-specific estimates of low-income youth college enrollment that come from U.S. Census Bureau surveys cannot be trusted due to their household-based sampling designs, which only capture incomes of those in a single household and are not the same as family or parental income for many young people. Because Census data often misclassify financially dependent young people from moderate to high-income families as low income (U.S. Census Bureau, 2009), estimates from these sources likely overestimate rates of college attendance among low-income young people.

Based on these difficulties, it is tempting to use estimates of $P(low\,income \mid college)$ taken from institutional sources such as the Integrated Postsecondary Education Data System (IPEDS) as proxies for $P(college \mid low\,income)$. Nevertheless, there is no necessary within-state correlation between the proportion of low-income college students and the proportion of youth from low-income families who enroll in higher education: $P(low\,income \mid college) \not\equiv P(college \mid low\,income)$. It is entirely possible that states with high proportions of low-income college students still have large populations of low-income young people who fail to enroll. Conversely, it is also possible that many low-income young people who enroll are not identified as such because they do not complete the FAFSA and therefore do not benefit from financial supports they might otherwise receive (Kofoed, 2017). A lack of clear estimates of enrollment by family income at the state level means that neither federal and state policymakers nor college administrators have clear evidence that their policies are effective in increasing college participation among low-income youth.

To fill this informational gap, we use multilevel regression with poststratification (MRP, Park et al., 2004) to estimate postsecondary attendance rates by family income in each of the 50 states and the District of Columbia. In this multi-step procedure, we first estimate the probability of postsecondary enrollment among high school graduates by family income among strata of a recent cohort of students surveyed in the High School Longitudinal Study of 2009 (Duprey et al., 2020). We next weight these predictions with matched state-level population characteristics in order to estimate college attendance rates by family income within each state. To validate our procedure, we compare MRP estimates to representative estimates that can be computed for ten states using HSLS09 survey-provided sampling weights. (We also validate the use of MRP for our research question using simulated data with known properties, which we present in the Appendix.) We conclude by comparing our MRP estimates to Census- and IPEDS-derived estimates of low-income youth enrollment to demonstrate their non-equivalence.

We contribute to the literature by providing the best possible state-specific estimates of the probability of college attendance among low-income youth. In addition to finding that low-income young persons are less likely than their higher-income peers to attend college overall, we find substantial variation in both the difference in attendance rates between these groups and their respective attendance rates across the states. Our estimates differ substantively from those taken directly from the U.S. Census as well as proxy measures such as the proportion of currently enrolled students who are Pell eligible and those who fall within the lowest family income band. While we believe our principled MRP procedure offers robust estimates of college enrollment among low-income youth across the United States, our ultimate hope is that our findings are obviated by better data collection and reporting in the future.

2 Background

It is well-established that young people from low-income families are less likely to attend higher education than their peers (Chen et al., 2017), even though higher education could serve as a means of increasing their opportunity (Corak, 2013). Many researchers have documented the struggles of young people from low-income families to attend higher education (Bell et al., 2009; Dowd, 2004; McDonough, 1997; Perna & Jones, 2013). In addition to poorer access to advising resources at their schools and often less accurate beliefs about the potential costs and benefits of college (Perna, 2006), one particular reason for comparatively lower rates of college attendance may be that young people from low-income families are more price responsive than their peers (Bartik et al., 2021; Carneiro & Heckman, 2002; Deming & Dynarski, 2009; Dynarski, 2002; Dynarski et al., 2021; Heller, 1997; Kane, 2006; Leslie & Brinkman, 1987), substantiating the need for income-based financial aid (Kane, 1999).

Given that young people from low-income families face the biggest barriers to attending college, the nation should keep track of attendance rates in higher education by family income (Deming & Dynarski, 2009). Furthermore, these data should be available at the state level, as states are the key players in setting higher education policy (Callan & Finney, 1997; Denning, 2019; Gurantz, 2022; Richardson et al., 1999; Scott-Clayton & Schudde, 2020; Zumeta et al., 2012). State leaders decide the amount of higher education that will be supplied, determine its price, and play a large role in determining how much financial aid will be provided to students (Zumeta et al., 2012). Yet state leaders make all of these decisions without clearly knowing how enrollment levels across different family income groups will be impacted. Federal data sets from the U.S. Census and Education Department cannot provide accurate estimates of college attendance rates by family income at the state level due to their design. If the states themselves are tracking enrollment by income, we have not been able to find any reports of their findings. A comprehensive review of state policy documents yielded no studies that track the proportion of individuals who attend college by family income.

Calculating postsecondary attendance rates by family income at the national level using federal surveys is a relatively simple task often done in the context of a larger project. Reports generated by the National Center for Education Statistics (NCES) find substantial gaps in postsecondary attendance by family income, even after taking into account the academic preparation of high school graduates (Berkner & Chavez, 1997; Bozick & Lauff, 2007). Duncan & Murnane (2011) find that gaps in postsecondary attendance by income did not close over the two decades prior to their study. NCES longitudinal surveys are designed to be generalizable to a national and often regional population of high school students (e.g., S. J. Ingels et al., 2014). In the case of the High School Longitudinal Study of 2009 (HSLS09), they are representative in a small subset of states (Duprey et al., 2020). These surveys are not designed, however, to support estimates at the state level across all states. Because their complex sampling procedures do not include stratification at the state level or cluster sampling within states, sampled students are not representative of the state in which they attended high school, even when restricted-use data files that indicate students' states of enrollment are used (S. J. Ingels et al., 2014).

The U.S. Census Bureau's American Community Survey (ACS) and Current Population Survey (CPS) represent another pair of data sources that one might consider using to estimate college attendance by family income level. Nevertheless, these surveys are plagued by a common problem that is particularly severe in the case of young people such as college students. Both Census surveys utilize a sampling procedure based on the households where people live (U.S. Census Bureau, 2012a, 2013). This design does not generally present a problem for estimating overall family income levels for older respondents, even at the state level. Estimating family income levels for young adults, however, represents a challenge. Survey questionnaires ask for information on all persons who are residents of the household at the time the forms are received. Residency follows the "two-month rule," which says that

> [i]n general, people who are away from the sample unit for two months or less are considered to be current residents, even though they are not staying there

> when the interview is conducted, while people who have been or will be away for more than two months are considered not to be current residents (U.S. Census Bureau, 2009, pp. 6–2).

While specific allowances are made for child dependents under 18 years of age who are away for boarding school, none are made for college-age dependents who are away for longer than the two month period.

The sampling design of Census surveys such as ACS means that when researchers limit analyses to observations of students enrolled in college, they are likely to miss a significant proportion of students whose families were interviewed while they were living apart for school. In addition, researchers are likely to receive limited family information on young people living in group housing. For this latter group, data limitations originate from the household-based survey procedure. Individual household members are asked to report their own income, which is later combined to create an estimate of total household income. While incomes of close relatives within a household are separately combined to form a family income, the income of family members outside of the household, which includes parents of a still dependent survey respondent, is not included in this estimate. Thus, if a surveyed young person lives in a household other than that of their parents or guardians for longer than the two-month window—in a household made up of other college students, for example—their reported family income will reflect only their individual earnings and will likely be lower than expected considering their true financial dependency. This would mean that dependent students from high income families could be classified as low income if they were living apart from their household for more than two months.

The end result is that family income levels reported by young people between 18- and 24-years-old are unreliable (U.S. Census Bureau, 2012b, 2012a, 2013). Because estimates of college attendance by family income taken from household-based Census data are likely to overstate the proportion of young people from low-income families in higher education, they should not be used to estimate college attendance rates by family income. Otherwise, such estimates, were they to be used in the evaluation of student aid policies, could lead researchers and policymakers alike to be more sanguine about the efficacy of family income-based access policies than is warranted.

What we do have at the state and institutional level are proxy measures of the proportion of students *already in college* who are from low-income families, $P(low\,income \mid college)$. Many states and institutions, for instance, report the proportion or number of students who are Pell eligible (Rosinger & Ford, 2019; Tebbs & Turner, 2005). However, indirect measures like the percentage of students who are Pell eligible do not directly measure the college-going rates of young people from low-income families. These measures ignore young persons who are not enrolled in higher education or, in the case of Pell grants, do not apply for federal student aid. As a result, researchers and policymakers do not currently know what percentage of young people from different family income levels transition from high school to college in each state.

Understanding the impact of state and federal policy on postsecondary attendance rates requires an inverse of the above estimate, that is, an estimate of the probability that a young person will enroll in college conditional on being from a low-income family, $P(college \mid low\,income)$. We have established that such estimates cannot be produced by a straightforward analysis of currently available federal longitudinal or cross-sectional data sets. Instead, we must utilize both types of data in order to come up with estimates of the likely level of college attendance by income in each state. Our approach, which we describe in more detail in the next section, begins with estimating the probability of enrollment in postsecondary education using data that covers the period from when students are first enrolled in high school to when they are 18- to 19-years-old, graduated from high school, and eligible to enroll in college. From this federal longitudinal survey, we use characteristics of high school students aged 14- to 15-years-old that include their family income, race/ethnicity, and gender to predict the probability that a student will attend postsecondary education in the year after high school graduation. We then turn to state-level Census data from the base year of the longitudinal survey to get counts of the number of 14- and 15-year-olds in each state with the same combinations of characteristics as those in our sample of high school students. Using these counts to weight our estimates, we produce representative state-level estimates of college attendance among low-income young persons.

3 Methodology

We use a Bayesian approach for estimation and inference. While a full explanation of Bayesian statistics is beyond the scope of this paper, the primary difference between Bayesian and frequentist inferential approaches lies in the treatment of the unknown parameter. Whereas Bayesian approaches treat the unknown parameter as a random variable, frequentist statistics treat the unknown parameter as a fixed value (Gelman et al., 2013). Bayesian approaches estimate the distribution of the unknown parameter, θ, using

$$P(\theta|X) \propto P(X|\theta) \times P(\theta),$$

where the posterior distribution, $P(\theta|X)$, is proportional to the likelihood of the data given the parameter, $P(X|\theta)$, times a prior, $P(\theta)$, which represents the initial state of the analyst's beliefs about the distribution of the population parameter. This posterior distribution represents an updated estimate of the likely distribution of the unknown parameter after factoring in the likelihood of the data. Our discussion will focus on summary measures—mostly quantiles—of the posterior distribution of various parameters. The summary measures in our work will discuss the probability that the population parameter is in a certain range—the credible interval. In contrast, frequentist statistics focus on the probability of observing values in an infinite number of repeated samples under assumptions about the sampling distribution (null hypothesis), which is generally not an estimate of interest in most policy applications (Gelman et al., 1995).

To recover state-level estimates of enrollment among low-income young persons in college, we use multilevel regression with poststratification (MRP), a statistical technique that has been widely used in the political science literature to estimate public opinion (Gao et al., 2019; Gelman et al., 2010; Gelman & Little, 1997; Howe et al., 2015; Kastellec et al., 2019; Kennedy & Gelman, 2019; Lax & Phillips, 2009; Lei et al., 2017; Lipps & Schraff, 2019; Little, 1993; Pacheco, 2011; Park et al., 2004; Wang et al., 2015; Warshaw & Rodden, 2012). Researchers in other disciplines such as public health (Downes et al., 2018; Eke et al., 2016; Zhang et al., 2014) and education policy (Ortagus et al., 2021) have also used MRP to produce representative estimates using non-representative data.

MRP works using two data sets and two primary analysis steps. First, a multilevel model of the form,

$$P(y_i = 1) = logit^{-1}(\beta_0 + \sum_{k=1}^{K} \alpha^k_{j[i]} + Z_i\gamma), \tag{1}$$

is fit to a binary outcome of interest, y_i, using observations, i, from the first data set that contains non-representative survey responses. The outcome could be voting for a particular candidate, supporting a policy position, or, in our case, enrolling in college. The right-hand side parameters in the model include a grand mean, β_0, and a suite of random intercepts, α^k, indicating demographic categories and geographic areas that separate each observation into a limited number of population cells, j. In addition, the right-hand side of the model includes second-level covariates and parameters, $Z_i\gamma$, that are associated with the area to which one wishes to poststratify (e.g., state). Once fit, predicted probabilities, π_j, are computed for each population cell in the data set. As one example, one would predict the likelihood that a low-income (α^1) white (α^2) male (α^3) high school graduate (α^4) from Kentucky (α^5) enrolls in college. The total number of population cells for which predictions are computed would equal the cross of all categories in α^k.

In the second step, values of π_j are aggregated to the area of interest, θ_S, using

$$\theta_S = \frac{\sum_{j \in S} N_j \pi_j}{\sum_{j \in S} N_j} \tag{2}$$

which reweights each demographic cell's predicted probability using corresponding population counts, N_j, from the second data set. Population cell counts in the second poststratification data set are often constructed using Census data. In general, the area of interest, θ_S, could be any geographic or institutional (see Ortagus et al., 2021) level for which population cell counts can be computed. The geographic area of interest in our study is the state.

Demographic cells indicated in the first data set and multilevel regression must match those available in the second poststratification matrix.[1] This often limits the unit-level information that can be used to predict the outcome. For example, high school GPA, which would be positively correlated with college enrollment, cannot be included in the multilevel model because high school GPA is not included in Census data.[2] For this reason, second-level covariates are important for improving fit and producing better area-specific poststratified predictions (Park et al., 2004).

With this approach, we use nationally-representative data to first estimate the probability of college attendance by individual characteristics, including race/ethnicity, gender and age. We then use state-level estimates of the numbers of individuals in each of those categories to predict enrollment rates by income at the state level.

4 Estimating college enrollment among low-income youth

To estimate the proportion of low-income youth who attend college, we follow the same MRP procedure outlined above. In the first step, we use student-level data from the most recent NCES survey, HSLS09, to estimate the likelihood of enrolling in college. Our multilevel

[1] Kastellec et al. (2015) propose a method for using a poststratification matrix comprised of non-Census variables (in their case, partisanship) and that incorporates uncertainty in the poststratification matrix. It remains true, however, that demographic cells, j, in the primary data set and poststratification matrix must correspond.

[2] This was not the case for Ortagus et al. (2021), who used administrative data and poststratified to colleges in the original sample frame rather than a geographic area. Nevertheless, it holds true in our study as we use publicly available Census data to construct the poststratification matrix.

logistic regression takes the form,

$$P(y_i = 1) = logit^{-1}(\beta_0 + \beta^{lowinc} * lowinc_i + \beta^{female} * female_i + \alpha^{race/ethnicity}_{re[i]} \\ + \alpha^{region}_{r[i]} + \alpha^{state}_{s[i]} + \alpha^{state.lowinc}_{sl[i]} + Z_i\gamma), \tag{3}$$

in which we use random effects parameters for student characteristics that include gender, race/ethnicity, and income status as well as state and regional indicators. We fully interact income status, our binary covariate of interest, with state indicators to account for potential differences in low-income college enrollment across the states. Second-level covariates in Z include state-level percentages of adults with a Bachelor's degree or higher, the proportion of college students who attend public two-year institutions, the average tuition at four-year public institutions, a measure of average distance to public two-year institutions, and county population-weighted state average unemployment rate for the years aligning with the year of on-time college enrollment for students surveyed in HSLS09. We give all regression parameters weakly informative normal priors: $\alpha, \beta, \gamma \sim N(0, \sigma); \sigma \sim N_+(0, 1)$.

Our second level estimates come from the literature on predictors of low-income college enrollment. Based on national surveys, analysts have shown that students who live in states with higher educational attainment are more likely to attend postsecondary education, as are students who live in states with more community colleges (Bozick, 2009; Doyle & Skinner, 2016). Similarly, the "sticker price" of four-year colleges has been shown to be a reliable predictor of on-time college enrollment, as colleges with lower tuition are perceived to be more affordable, regardless of actual net price (Hemelt & Marcotte, 2011). We include unemployment rate due to its inverse relationship with college attendance, particularly among low-income youth who forgo earnings when they choose continued schooling over immediate employment.

After fitting equation 3, we create predicted probabilities of enrollment for each demographic cell in our data set. Once we have calculated enrollment probabilities for each demographic cell, $N_J = 1{,}224$ unique cells, we compute the weighted probability of enrollment for each state, θ_S, with equation 2 using population counts as weights. Within each state, we estimate two values of θ_S, θ_S^{hi} and θ_S^{lo}, representing the likelihood of college enrollment among middle/high-income and low-income youth, respectively.

5 Data

Data for our study come from two primary sources. Unit-level data come from a student longitudinal survey conducted by the National Center for Education Statistics, the High School Longitudinal Study of 2009 (HSLS09, Duprey et al., 2020). As with prior NCES longitudinal surveys, HSLS09 tracked a nationally-representative sample of ninth-graders starting in 2009 as they moved through high school and either enrolled in postsecondary institutions or entered the workforce. Unlike prior NCES surveys, HSLS09 is representative in 10 states—California, Florida, Georgia, Michigan, North Carolina, Ohio, Pennsylvania, Tennessee, Texas, and Washington—though it cannot be used to construct state-level estimates across the entire United States.

We use information from the first and fourth wave of the survey. Data on the state of residence for all students in the base year of the survey come from the restricted-use data files. We use the variable `x4fb16enrstat` to determine those students who enrolled within one-year of earning a high school diploma or GED. Using this variable allows us to differentiate between pre-graduation dual enrollments and postsecondary matriculation—the outcome of interest. We use the base-year family income variable, `x1famincome`, which discretizes incomes into 13 bins of non-equal size, to construct a binary variable of income status. We assign low-income status for all students with family incomes below $35,000, which equals between 150% and 185% of the federal poverty line for a family of four in 2009. Because federal loan policy does not make state-specific cost of living adjustments in determination of need, we do not adjust the poverty indicator by state. Our primary unit-level data set has $N = 13{,}020$ observations, 26% of whom are coded as low-income.

Second-level state characteristics are taken from various sources covering the year of on-time college enrollment for students in the HSLS09 sample. These sources include the College Scorecard, the American Community Survey, and the Bureau of Labor Statistics. We follow Doyle & Skinner (2016) in computing a measure of distance that takes into account all two-year colleges in the state rather than just the closest. Specifically, our measure is the population-weighted average of $\sum \frac{1}{distance^2}$ between each census block group and all public two-year institutions in the state.

Finally, to construct our poststratification matrix of demographic cell sizes, we use population data from the American Community Survey. We select data from the period that best aligns with the year of the HSLS09 sample, which are single-year population estimates for 14 to 15-year-olds in 2009.[3] Using HSLS09 and these population weights, our results are most applicable to the population of young persons who graduated high school in the early 2010s.[4]

6 Results

Results from our regression model are shown in Figure 1 (see Table A1 and Table A2 for summary values). Center dots in each figure represent median posterior ($\hat{\theta}_{q50}$) values, with lines representing 95% credible intervals. Values are unadjusted and on the logit scale. Holding all else equal, first level parameters show that low-income youth are less likely to

[3] In addition to single-year estimates, we also poststratified results using 3- and 5-year estimates to provide more stable population estimates, particularly for small subpopulations. For each multi-year poststratification matrix, we used Census estimates in which 2009 was the central year: 2008-2010 3-year estimates and 2007-2011 5-year estimates. Our results did not differ within meaniful degrees of rounding across these different estimates, likely due to the large size of our subnational areas. All results in the paper use single year ACS estimates for poststratification. Other estimates are available via the replication code.

[4] Ideally, we would not condition college enrollment on college graduation since we cannot do the same for our poststratification matrix. However, because almost all college enrollees have a high school diploma or GED, there is little effective difference between $\mathbb{1}(college)$ and $\mathbb{1}(college \mid HS)$. Out of the full analysis sample, fewer than 25 observations report non-dual enrollment postsecondary enrollment without having earning a high school diploma or GED.

enroll in college than their middle-high-income peers. Women and those who identify as Asian / Pacific Islander are more likely to attend than men and other racial/ethnic groups. Both sets of findings are consistent with the literature on college enrollment. While we find less evidence of differences in enrollment across other racial/ethnic groups when controlling for income, we note that our model does not account for the role of racial discrimination on income and its downstream relationship to enrollment. At the second level, the slightly positive shift in the posteriors for Bachelor's degree, two-year institutions, and unemployment rate alongside the slightly negative shift in the posteriors for four-year tuition and distance to two-year institutions are similarly in line with the access literature, though all parameters are generally close to zero. Comparatively, we estimate more differences in attendance across states, particularly among middle-high income young persons. Though many state-specific medians are close to zero, the widths of their credible intervals suggest heterogeneity in enrollments across the sample within subgroups.

Figure 2 presents our primary results, showing the probability of college enrollment across the 50 states and the District of Columbia for both low-income (red) and middle-to-high-income (blue green) high school graduates in the top panel (A). The figure is ordered by low-income enrollment, median values ($\hat{\theta}^{state}_{q50}$) which range from 34.8% [21.0,49.1] of recent high school graduates in Utah to 68.1% [57.2,79.1] in Mississippi, a range of 33.3 p.p. with a median of 50.7% [44.8,56.4] in Illinois. Because we interact the indicator of low-income status with state-specific random intercepts, we allow for state-specific differences in the relationship between income status and college enrollment. Within-state differences between low- and middle-to-high-income enrollment are shown in panel B of Figure 2. These are ordered from the lowest to highest difference between the two populations. Differences range between 19.7 p.p. [12.2,25.9] in Mississippi and 30.3 p.p. [23.2,35.9] in Wyoming with a median of 27.3 p.p. [22.9,31.3] in California.

While the size of the 95% credible intervals demonstrate some uncertainty in our estimates, which at the most ranges 31.2 p.p. [40.0,71.2] for low-income enrollment in Hawaii, we make two notes. First, that the credible intervals for low- and middle-to-high-income estimates within 47 of 50 states and D.C. do not overlap provides strong evidence that the on-time enrollment rates of low-income students are lower than that of their higher income peers across the country. Second, we are able to provide estimates in low population states and, in the case of the District of Columbia, a location entirely unsampled by HSLS09.

In the next section, we more formally validate our estimates for which we have other state-representative results. But first, we provide some context as to the face validity to a particular aspect of our findings. One estimate that stands out is Mississippi. Whereas Mississippi has historically ranked comparatively poorly among states on various education-related measures such as rates of children receiving free-and-reduced-price lunch, students performing at or above basic levels of reading achievement, teacher salaries, high school graduation rates, and overall educational attainment (Snyder & Dillow, 2013), our estimates put it at the top of the rankings for low-income youth college enrollment as well as first among states for the lowest gap in enrollment rates between low- and middle-high income students. Despite this apparent contradiction, we find a number of reasons to trust our estimate for Mississippi.

First, we note that our estimates are comprised of those young persons who graduated high school or earned a GED. Because Mississippi had a comparatively lower high school graduation rate than the rest of the nation during the period of our study, an averaged freshman graduation rate [AFGR] of 63.8 compared to the national rate of 78.6 (Table 219.40, Snyder & Dillow, 2013), we would expect the college matriculation rate among high school graduates in the state to be higher if, all else equal, students on the margins are more likely to drop out before earning a high school diploma as it would mechanically change the denominator used to compute the rate. In addition, we again note that our indicator for low-income status does not take into account regional differences in purchasing power or cost of living. We choose not to adjust our poverty estimates by state since federal financial aid funding formulas do not adjust for state-specific cost of living differences. This choice means that a comparatively higher proportion of young persons in Mississippi are categorized as low income who, were they to have a similar but cost-adjusted financial station in another state, who would not be considered low income. Again for mechanical reasons, Mississippi might be expected to have a greater rate of low-income youth attendance as well as smaller gaps between low and middle-high income enrollment rates. Finally, we note that Mississippi has some of the lowest costs among four-year public schools in the nation, even compared to its neighboring states. Using data from the College Scorecard, average net tuition in 2013 among public four-year institutions in Mississippi was $6,517, which is at the 25th percentile across all states, and lower than the averages of bordering states: Alabama ($8,579), Arkansas ($6,970), Louisiana ($6,826), and Tennessee ($8,220). Combining these various aspects of Mississippi's unique context, we find it plausible that it would rank first among the states in low-income youth college attendance. We are not alone in finding this result. The National Center for Education Statistics reports Mississippi as having the highest rate of high school graduates attending college in the nation, while at the same time ranking Mississippi well below the median in the proportion of 18-24 year-olds attending higher education (De Brey et al., 2021). Of course, initial enrollment is not the same as persistence and eventual degree attainment, a point we return to in the paper's conclusion.

7 Validation of MRP

Having presented our state-level MRP estimates of low-income versus mid/high-income college enrollment and the differences between them, we now turn to validating. We begin by first checking our estimates against those produced when using survey weights for the limited number of representative states.

7.1 Comparison to 10 representative states

While not representative for all states, the primary data set we use to produce these estimates, HSLS09, is representative for a subset of states when survey weights are used. To validate our MRP procedure, we use these weights to produce state-level estimates in the ten states for which HSLS09 is representative—California, Florida, Georgia, Michigan, North Carolina, Ohio, Pennsylvania, Tennessee, Texas, and Washington. In this section, we compare survey-weighted estimates to those we derived using MRP.

For each state in this representative subsample, we estimate the survey-weighted mean of respondents who attended college by our indicator of low-income status using two weights: the base year survey weight, `W1STUDENT`, and the first year-fourth year longitudinal weight, `W4W1STU`. Whereas the latter longitudinal weight is the more correct survey weight to use since it covers both the initial survey wave, when covariates were collected, and the fourth wave, when the outcome measure of college enrollment was collected, we also provide estimates using the initial base year weights. In practice, the resulting differences between the two survey-weighted estimates are negligible.

We fit the same model,

$$P(y_i = 1) = logit^{-1}(\beta_0 + \beta^{lowinc} * lowinc_i), \tag{4}$$

for each state twice, using each weight. So that we can recover a full posterior distribution of estimates and thereby make a more proper comparison to our MRP estimates, we use the `brms` R package (Bürkner, 2021) to fit equation 4. By design, brms treats weights as frequency weights, multiplying the likelihood by the weight as if the observation were repeated X times. For example, if an observation had the average `W1STUDENT` weight of 191.9 (S. J. Ingels et al., 2011), it would be treated as if it were effectively 192 observations. Multiplying each observation by a positive amount would imply we observe more data than we do, artificially increasing our certainty, and resulting in inappropriately small credible intervals. To prevent this, we adjust both sets of weights using,

$$\tilde{w}_i = N_{obs} \frac{w_i}{\sum w}, \tag{5}$$

in which the adjusted observation weight, $\tilde{w}_i$, is the provided sample weight, w_i, divided by the sum of all weights and multiplied by the analytic sample size, N_{obs}. This procedure recenters the weights relative to their respective contributions in the population while preserving the same overall information for the Bayesian model.[5] After fitting equation 4 with adjusted weights, we use estimated model parameters to compute posterior distributions of the predicted probability of college enrollment for each state.

Figure 3 presents the results for this validation exercises (see appendix Table A4 for summary values). Three estimates of low-income college enrollment (top section) and middle-high income enrollment (bottom section) by state are shown in each plot area: base year weight weighted estimates (red dash), longitudinal weight weighted estimates (green dot), and MRP estimates (blue solid). Visually, our MRP estimates generally align with survey weight-derived estimates in most facets, differing most notably in estimates for low-income enrollment in Washington (MRP lower), middle-high income enrollment in Florida (MRP higher), and both income levels for Ohio (MRP again higher).

More formally, we offer two test statistics comparing MRP estimates to each of the survey weight-derived estimates. The first statistic, OVL, shows the proportion of overlap between the two distributions. The second, $|\Delta\theta|$, provides the absolute percentage point difference in the median values (θ_{q50}) of the two distributions. Whereas policy-makers, stakeholders, and other researchers may be more interested in single point estimate values of state-level college enrollment by state, our Bayesian framework supports more direct comparisons of estimate uncertainty. For these reasons, we believe a qualitative combination

[5] In frequentist procedure, we would use the corresponding balanced repeated replicate weights provided by HSLS09 to calculate appropriately adjusted errors for the weighted point estimates. Because estimate errors are directly computed as a function of posterior distributions in a Bayesian framework, we do not (nor can we) use post-estimation BRR weighting adjustments.

of the two test statistics provides the best approach to model comparison.

Across all states and both weighted samples, the mean overlap between MRP estimates and their corresponding weighted estimates is 0.56 (median: 0.63). These values are approximately the same when split by low-income and middle-high-income estimates. On average, MRP estimates have a higher degree of overlap with base year-weighted estimates (0.63) than longitudinal-weighted estimates (0.56). Across all estimates, the greatest degree of overlap is between MRP and longitudinal-weighted estimates of low-income enrollment in Texas (0.92); the lowest is between MRP and longitudinal-weighted estimates of middle-high-income enrollments in Ohio (0.04).

Turning to absolute difference in median posterior values, the average difference is 2.94 percentage points (p.p.; median: 2.17 p.p.). Unlike with the overlap comparison, the average absolute median difference does not change greatly by weight type—base: 2.92 p.p.; long: 2.97 p.p. On the other hand, the mean difference for low-income estimates at 3.6 p.p. is larger than that for middle-high-income estimates at 2.25 p.p. The lowest estimated difference in median posterior values is between MRP and middle-high-income estimates using base-year weights (0.13 p.p.); the largest difference is in Ohio for low-income estimates using longitudinal weights (9.72).

Considered together, we find solid evidence of commensurability between MRP estimates and those obtained with survey-provided analytic weights. On average, the full distribution of MRP posterior estimates has greater than 50% overlap and has a median not more than 3 percentage points away from weighted estimates. In the worse case example of Ohio, median MRP estimates remained within 10 percentage points of the weighted estimate. Though a nearly double-digit median difference in estimates of college attendance is meaningfully different, we make one observation. Due to missing data (mostly surrounding family income values used to calculate the low-income indicator), it is unlikely that either set of survey weights retains its same degree of representation as we make no adjustments to account for this missingness. Therefore, while we compare MRP estimates to weighted estimates, there is no external metric by which we can conclusively adjudicate in favor of one over the other as being closer to the true population value. This is not to say we blindly accept the MRP estimate over the weighted estimate when they differ. We do, however, argue in favor of the comparative transparency of the MRP procedure over the use of weights of unclear applicability.

7.2 Simulation

Additionally, we offer another validation exercise in which we follow the MRP procedure on synthetic data structured to mimic NCES surveys like HSLS09 but with known properties and relationships between covariates. In this simulation, we follow the example of Park et al. (2004), but with modifications that make it more pertinent to our use case. Briefly, we (1) generate a known national population, from which we (2) take a number of nationally- but not state-representative samples, and, (3) following our MRP procedure, are able to recover reasonable state-level estimates of a parameter of interest. See Appendix B for a more complete description of this simulation exercise.

8 Comparison to proxy estimates

In this final section, we compare our estimates of low-income student enrollment with publicly available, potentially cognate estimates that researchers and policymakers might be tempted to use in place of the actual statistic of interest. We consider three measures: (1) a direct measure of low-income student enrollment, (2) the percentage of students enrolled in college who receive Pell grant funding, and (3) the percentage of students enrolled in college whose families earn less than $30,000 a year.

The first cognate measure, which comes from IPUMS microcensus data for the American Community Survey, is the state-level percentages of 18 and 19-year-olds in 2013 who have earned a high school credential (diploma or GED) but not a postsecondary credential and who are currently enrolled in college as undergraduates. We further filter this group to those young people with family incomes less than $35,000 per year—the cut-off we use for the HSLS09 sample, adjusted for inflation.[6] Panel A of Figure 4 compares these ACS estimates to our MRP estimates. For each state, the median posterior value is shown by state abbreviation and vertical lines show 95% credible intervals. The diagonal dashed line represents the point at which the two estimates are the same. Values below the dashed line indicate that direct Census estimates are higher than those obtained using MRP. As expected based on the way Census family income does not account for financial dependency across households, Census estimates skew higher than MRP estimates, with a median difference of 22.5 p.p. [14.9,32.1]. Only two states—Alaska and Hawaii—slightly underestimate low-income college student enrollment compared to MRP estimates. In practical terms, this means that using direct estimates from the Census are likely to overstate the participation on low-income young persons in college by a large margin in most states.

Panel B of Figure 4 compares MRP estimates to the percentage of first-time in college undergraduates who participated in the Pell grant program in the 2013-2014 academic year. The percentage of students who are Pell eligible is computed using institution-level data from IPEDS, aggregated to the state level using undergraduate full-time equivalent (FTE) counts as weights. Unlike the first cognate measure, which represents the probability of enrolling in college conditional on being low income, $P(college \mid low\,income)$, the relation of interest and what we estimate using MRP, the Pell estimate measures the probability of being low-income conditional on being enrolled in college, $P(low\,income \mid college)$. Nevertheless, the percentage of students who either use Pell or are Pell eligible is often used as a proxy for low-income enrollment. Converse to ACS estimates, Pell proxy estimates tend to understate college enrollment among low-income youth.

In panel C of Figure 4, we compare MRP estimates to the percentage of first-time in college undergraduates with family incomes less than $30,000 in the 2013-2014 academic

[6]We use an inflation adjustment pegged to 2009 real dollar values as given by the Federal Reserve Economic Data file, USACPIALLAINMEI (https://fred.stlouisfed.org/series/USACPIALLAINMEI).

year, which comes from the student financial aid component of IPEDS. This income category represents the closest approximation to the low-income category we use for our primary MRP estimates. We are limited to reporting only this grouping of low-income students given IPEDS' reporting standards for enrollment by income. As with the Pell measure, this financial aid cohort measure is averaged from the institution level to the state level using undergraduate FTE enrollments as weights. It also represents the inverse probability of the MRP estimates: $P(\mathit{low\,income} \mid \mathit{college})$ rather than $P(\mathit{college} \mid \mathit{low\,income})$. Panel C shows that as with the Pell-eligible measures, the proportion of the financial aid cohort that is low-income used as a proxy for low-income enrollment tends to underestimate compared to MRP.

In addition to the visual comparison between our MRP estimates and these proxies, we provide two correlation statistics in the open panel of Figure 4. These include both Pearson (ρ_p)and Spearman (ρ_s) rank-order correlation coefficients. We include Spearman correlations to account for monotonic relationships that may exist between MRP and proxy values even in cases where they are not linearly related. The highest levels of correlation we observe are between MRP and ACS estimates (Panel A), with $\rho_p = 0.33$ and $\rho_s = 0.4$, which are not particularly strong. For the other proxies, the correlations are weaker—Pell: $\rho_p = -0.03$ and $\rho_s = -0.13$; financial aid cohort $< \$30,000$: $\rho_p = -0.15$ and $\rho_s = -0.10$.

Together, the comparisons presented in Figure 4 demonstrate that three common cognate measures of low-income youth college enrollment are not strongly aligned with our more principled MRP estimates. Using U.S. Census-reported information on enrollment by household income substantially over-reports the proportion of low-income young people who enroll in higher education. As we find no observable relationship between our MRP estimates and either measure provided by IPEDS, we provide evidence that using a measure of $P(\mathit{low\,income} \mid \mathit{college})$ is a poor substitute for the actual measure of interest, $P(\mathit{college} \mid \mathit{low\,income})$. While these proxy measures give us useful information — for example, differences in Pell recipient attendance across sectors — they cannot give reliable information about the rates at which low-income youth attend college. This finding should not be unexpected since measures based on the probability of being low-income given that a person is enrolled in college, $P(low\,income \mid college)$, have no necessary reason to correlate with the more policy relevant measure of the probability of enrolling in college given that a person is low-income, $P(college \mid low\,income)$, despite an understandable desire to substitute the former for the latter as a matter of convenience. Metrics and prior research that use these proxies for low-income youth enrollment should be interpreted with caution. Based on these results, we conclude that while existing measures of low-income youth enrollment in college do not capture the policy-relevant outcome, our MRP-based estimates do.

9 Conclusion

The significance of this study is two-fold. First, we offer state-level estimates of low-income college enrollment for a recent cohort of young persons. These estimates provide better evidence of the efficacy of policies meant to support low-income youth enrollment in college at the state level, where many aid policies are set and funded. Second, we generate our estimates using a statistical procedure, MRP, that we believe can be usefully applied to other education policy questions for which data at the proper level of inference is otherwise limited (e.g., Ortagus et al., 2021). While researchers may find our estimates useful in future analyses, we believe that our findings are most immediately useful to policymakers who currently lack robust state-level estimates of college-going by low-income youth. With these estimates, policymakers have a baseline upon which to evaluate the effectiveness of college access policies such as need- and merit-based aid, promise programs, and free college plans.

The goal of federal grant aid in the form of the Pell Grant program has been to increase attendance rates in postsecondary education (Kane, 1999). The primary focus of efforts to increase participation in the Pell Grant program has been young people (Deming & Dynarski, 2009). Similarly, the primary goal of many state policymakers has been to increase college attendance rates among low-income young people through lower tuition and financial aid programs (Zumeta et al., 2012). Based on our findings here, considerable work remains to accomplish these goals.

Our results highlight several salient facts regarding attendance in higher education by income across states. First, fewer low-income young people attend college in every state than do their middle-income or high-income peers. The median gap between low-income postsecondary attendance and high-income attendance credibly ranges 33.8 p.p. across states. This means that despite 50 years of sustained effort to close the college attendance gap by income, low-income young persons still face considerable barriers to attending college, even in the states with the lowest net prices. Second, the probability of attending college for low-income young people varies dramatically across states. In the lowest performing states, about 34 percent of low-income young people attend higher education, while in the highest performing states, about 67.8 percent of low-income young people attend higher education. This means that a low-income young person's chance for attending college depends crucially on their state of residence.

We do not argue that initial attendance is the only relevant goal of state-level higher education policy. Students must persist and eventually earn a postsecondary credential in order for them to reap the full benefits of their college education (Bird et al., 2022). Mississippi again serves as an example. Despite its high rate in our estimates of college enrollment among low-income youth, it remains lower ranked among the states in terms of overall degree attainment. Considering the increasing costs of a college credential for which many students take on student loan debt (Ma & Pender, 2021), it is important that students do not merely attend some college before leaving without a credential. Nonetheless, access and initial enrollment is an important first step in the process. Future research could use similar MRP methodology to produce state-level estimates of rates of college persistence and graduation among low-income youth as direct extensions of this paper. MRP could be also be used to estimate within-state college access — by county or metropolitan statistical area (MSA), for example — among low-income youth in order to evaluate state-specific policies or simply better describe disparities across regions within the state.

Other future research might use these estimated postsecondary attendance rates to help find patterns regarding the effectiveness of various state-level policies. While it is well-established that lowering the price of higher education increases college attendance rates, particularly among young people, more research can be done to establish the effectiveness of other policies to encourage more young people to enroll. Technical issues aside, the failure of current data sources to provide reliable estimates on low-income postsecondary enrollment by state has substantial implications for education policy and cannot be overstated. Even with a combined $40 billion in spending from public sources on grant aid for low-income students, we know next to nothing about college attendance patterns by income at the state level (Baum & Payea, 2013; Ma & Pender, 2021). It is our hope that the methods proposed here will no longer be necessary in the years to come, as this data will be collected and reported.

References

Abraham, K. G., & Clark, M. A. (2006). Financial aid and students' college decisions: Evidence from the District of Columbia tuition assistance grant program. *Journal of Human Resources*, *41*(3), 578–610. https://doi.org/10.3368/jhr.XLI.3.578

Adelman, C., Daniel, B., & Berkovits, I. (2003). *Postsecondary attainment, attendance, curriculum, and performance: Selected results from the NELS:88/2000 postsecondary education transcript study (PETS), 2000* (NCES No. 2003-394). National Center for Education Statistics. http://nces.ed.gov/pubs2003/2003394.pdf

Bartik, T. J., Hershbein, B., & Lachowska, M. (2021). The effects of the Kalamazoo promise scholarship on college enrollment and completion. *Journal of Human Resources*, *56*(1), 269–310. https://doi.org/10.3368/jhr.56.1.0416-7824R4

Baum, S., & Payea, K. (2013). *Trends in student aid 2013* [Trends in Higher Education Series]. The College Board.

Bell, A. D., Rowan-Kenyon, H. T., & Perna, L. W. (2009). College knowledge of 9th and 11th grade students: Variation by school and state context. *The Journal of Higher Education*, *80*(6), 663–685. https://doi.org/10.1353/jhe.0.0074

Berkner, L., & Chavez, L. (1997). *Access to postsecondary education for the 1992 high school graduate* (No. 98-105). National Center for Education Statistics. http://nces.ed.gov/pubs98/98105.pdf

Bird, K. A., Castleman, B. L., Fischer, B., & Skinner, B. T. (2022). Unfinished business? Academic and labor market profile of adults with substantial college credits but no degree. *Educational Evaluation and Policy Analysis*, *44*(3), 404–428.

Blair, G., Cooper, J., Coppock, A., Humphreys, M., Rudkin, A., & Fultz, N. (2021). *Fabricatr: Imagine your data before you collect it*. https://CRAN.R-project.org/package=fabricatr

Bozick, R. (2009). Job opportunities, economic resources, and the postsecondary destinations of American youth. *Demography*, *46*(3), 493–512.

Bozick, R., & Lauff, E. (2007). *Education longitudinal study of 2002 (ELS:2002): A first look at the initial postsecondary experiences of the sophomore class of 2002* (NCES No. 2008-308). National Center for Education Statistics. http://nces.ed.gov/pubs2008/2008308.pdf

Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, *100*(5), 1–54. https://doi.org/10.18637/jss.v100.i05

Callan, P., & Finney, J. (1997). *Public and private financing of higher education*. ACE/Oryx Press.

Carneiro, P., & Heckman, J. J. (2002). The evidence on credit constraints in post-secondary schooling. *The Economic Journal*, *112*(482), 705–734. https://doi.org/10.1111/1468-0297.00075

Chen, X., Lauff, E., Arbeit, C. A., Henke, R., Skomsvold, P., & Hufford, J. (2017). *Early millennials: The sophomore class of 2002 a decade later* (NCES No. 2017-437). U.S. Department of Education. National Center for Education Statistics.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *The Journal of Economic Perspectives*, *27*(3), 79–102.

De Brey, C., Snyder, T. D., Zhang, A., & Dillow, S. A. (2021). *Digest of education statistics 2019.* (NCES No. 2021-009). National Center for Education Statistics.

Deming, D., & Dynarski, S. (2009). *Into college, out of poverty? Policies to increase the postsecondary attainment of the poor* (NBER Working Paper No. 15387). National Bureau of Economic Research. http://www.nber.org/papers/w15387

Denning, J. T. (2019). Born under a lucky star: Financial aid, college completion, labor supply, and credit constraints. *Journal of Human Resources*, *54*(3), 760–784.

Dowd, A. C. (2004). Income and financial aid effects on persistence and degree attainment in public colleges. *Education Policy Analysis Archives*, *12*(21).

Downes, M., Gurrin, L. C., English, D. R., Pirkis, J., Currier, D., Spittal, M. J., & Carlin, J. B. (2018). Multilevel regression and poststratification: A modelling approach to estimating population quantities from highly selected survey samples. *American Journal of Epidemiology*, *187*(8), 1780–1790.

Doyle, W. R., & Skinner, B. T. (2016). Estimating the education-earnings equation using geographic variation. *Economics of Education Review*, *53*, 254–267.

Duncan, G. J., & Murnane, R. J. (2011). *Whither opportunity?: Rising inequality, schools, and children's life chances*. Russell Sage Foundation.

Duprey, P., M. D., Wilson, D. H., Jewell, D. M., Brown, D. S., Caves, L. R., Kinney, S. K., Mattox, T. L., Smith Ritchie, N., Rogers, J. E., Spagnardi, C. M., & Wescott, J. D. (2020). *High school longitudinal study of 2009 (HSLS:09) postsecondary education transcript study and student financial aid records collection data file documentation* (NCES No. 2020-004). U.S. Department of Education. Washington, DC: National Center for Education Statistics, Institute of Education Sciences. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020004

Dynarski, S. (2002). The behavioral and distributional implications of aid for college. *The American Economic Review*, *92*(2), 279–285. http://www.jstor.org/stable/3083417

Dynarski, S., Libassi, C., Michelmore, K., & Owen, S. (2021). Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. *American Economic Review*, *111*(6), 1721–1756. https://doi.org/10.1257/aer.20200451

Dynarski, S., Page, L. C., & Scott-Clayton, J. (2022). *College costs, financial aid, and student decisions* (Working Paper No. 30275). National Bureau of Economic Research. https://doi.org/10.3386/w30275

Dynarski, S., & Scott-Clayton, J. (2013). *Financial aid policy: Lessons from research* (Working Paper No. 18710). National Bureau of Economic Research. https://doi.org/10.3386/w18710

Eke, P. I., Zhang, X., Lu, H., Wei, L., Thornton-Evans, G., Greenlund, K. J., Holt, J. B., & Croft, J. B. (2016). Predicting periodontitis at state and local levels in the United States. *Journal of Dental Research*, *95*(5), 515–522.

Gao, Y., Kennedy, L., Simpson, D., & Gelman, A. (2019). Improving multilevel regression and poststratification with structured priors. *arXiv Preprint arXiv:1908.06716*.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Chapman; Hall/CRC.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman; Hall/CRC.

Gelman, A., Lee, D., & Ghitza, Y. (2010). Public opinion on health care reform. *The Forum*, *8*(1).

Gelman, A., & Little, T. C. (1997). Poststratification into many categories using hierarchical logistic regression. *Survey Methodology*, *23*, 127–135.

Gurantz, O. (2022). Impacts of state aid for nontraditional students on educational and labor market outcomes. *Journal of Human Resources*, *57*(1), 1–32.

Heller, D. E. (1997). Student price response in higher education: An update to Leslie and Brinkman. *Journal of Higher Education*, *68*(6), 624–659.

Hemelt, S. W., & Marcotte, D. E. (2011). The impact of tuition increases on enrollment at public colleges and universities. *Educational Evaluation and Policy Analysis*, *33*(4), 435–457.

Howe, P. D., Mildenberger, M., Marlon, J. R., & Leiserowitz, A. (2015). Geographic variation in opinions on climate change at state and local scales in the USA. *Nature Climate Change*, *5*, 596–603.

Ingels, S. J., Pratt, D. J., Alexander, C. P., Jewell, D. M., Lauff, E., Mattox, T. L., & Wilson, D. (2014). *Education longitudinal study of 2002 third follow-up data file documentation* (NCES No. 2014-364). National Center for Education Statistics.

Ingels, S. J., Pratt, D. J., Herget, D. R., Burns, L. J., Dever, J. A., Ottem, R., Rogers, J. E., Jin, Y., & Leinwand, S. (2011). *High school longitudinal study of 2009 (HSLS:09). Base-year data file documentation* (NCES No. 2011-328). U.S. Department of Education. Washington, DC: National Center for Education Statistics. https://nces.ed.gov/surveys/hsls09/pdf/2011328_1.pdf

Kane, T. J. (1999). Has financial aid policy succeeded in ensuring access to college? In *The price of admission*. Brookings.

Kane, T. J. (2006). Public intervention in post-secondary education. *Handbook of the Economics of Education*, *2*, 1369–1401.

Kastellec, J. P., Lax, J. R., Malecki, M., & Phillips, J. H. (2015). Polarizing the electoral connection: Partisan representation in Supreme Court confirmation politics. *The Journal of Politics*, *77*(3), 787–804. https://doi.org/0.1086/681261

Kastellec, J. P., Lax, J. R., & Phillips, J. H. (2019). *Estimating state public opinion with multi-level regression and poststratification using R.*

Kennedy, L., & Gelman, A. (2019). *Know your population and know your model: Using model-based regression and poststratification to generalize findings beyond the observed sample* (arXiv Preprint No. 1906.11323v1).

Kofoed, M. S. (2017). To apply or not to apply: FAFSA completion and financial aid gaps. *Research in Higher Education*, *58*(1), 1–39. https://doi.org/10.1007/s11162-016-9418-y

Lax, J. R., & Phillips, J. H. (2009). How should we estimate public opinion in the states? *American Journal of Political Science*, *53*(1), 107–121.

Lei, R., Gelman, A., & Ghitza, Y. (2017). The 2008 election: A preregistered replication analysis. *Statistics and Public Policy*, *4*(1), 1–8.

Leslie, L. L., & Brinkman, P. T. (1987). Student price response in higher education: The student demand studies. *Journal of Higher Education*, *58*(2), 181–204.

Lipps, J., & Schraff, D. (2019). Estimating subnational preferences across the European Union. *Political Science Research and Methods*, 1—9.

Little, R. J. A. (1993). Post-stratification: A modeler's perspective. *Journal of the American*

Statistical Association, *88*(423), 1001–1012. https://doi.org/10.1080/01621459.1993.10476368
Lovenheim, M. F., & Reynolds, C. L. (2013). The effect of housing wealth on college choice: Evidence from the housing boom. *Journal of Human Resources*, *48*(1), 1–35. http://jhr.uwpress.org/content/48/1/1
Ma, J., & Pender, M. (2021). *Trends in student aid 2021* [Trends in Higher Education Series]. The College Board.
McDonough, P. M. (1997). *Choosing colleges: How social class and schools structure opportunity*. SUNY Press.
Ortagus, J. C., Skinner, B. T., & Tanner, M. J. (2021). Investigating why academically successful community college students leave college without a degree. *AERA Open*, *7*(1), 1–17. https://doi.org/10.1177/23328584211065724
Pacheco, J. (2011). Using national surveys to measure dynamic U.S. State public opinion. *State Politics & Policy Quarterly*, *11*(4), 415–439.
Park, D. K., Gelman, A., & Bafumi, J. (2004). Bayesian multilevel estimation with poststratification: State-level estimates from national polls. *Political Analysis*, *12*(4), 375–385.
Perna, L. W. (2006). Studying college access and choice: A proposed conceptual model. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 21, pp. 99–157). Springer Netherlands. https://doi.org/10.1007/1-4020-4512-3_3
Perna, L. W., & Jones, A. (2013). *The state of college access and completion: Improving college success for students from underrepresented groups*. Routledge.
R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/
Richardson, R., Bracco, K. R., Callan, P., & Finney, J. (1999). *Designing state higher education systems for a new century*. ACE/Oryx Press.
Rosinger, K. O., & Ford, K. S. (2019). Pell grant versus income data in postsecondary research. *Educational Researcher*, *48*(5), 309–315. https://doi.org/10.3102/0013189X19852102
Scott-Clayton, J., & Schudde, L. (2020). The consequences of performance standards in need-based aid: Evidence from community colleges. *Journal of Human Resources*, *55*(4), N.PAG.
Snyder, T. D., & Dillow, S. A. (2013). *Digest of education statistics 2013* (NCES No. 2015-011). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. https://nces.ed.gov/pubs2015/2015011.pdf
Stan Development Team. (2021). *Stan modeling language users guide and reference manual, 2.28*. https://mc-stan.org
Tebbs, J., & Turner, S. (2005). Low-income students a caution about using data on pell grant recipients. *Change: The Magazine of Higher Learning*, *37*(4), 34–43.
U.S. Census Bureau. (2009). *Design and methodology American Community Survey* [ACS-DM1]. United States Census Bureau.
U.S. Census Bureau. (2012a). *American community survey and Puerto Rico community survey 2012 subject definitions*. United States Census Bureau. https://www2.census.gov/programs-surveys/acs/tech_docs/subject_definitions/2012_ACSSubjectDefinitions.pdf
U.S. Census Bureau. (2012b). *Current population survey, October 2012 school enrollment and internet use supplement file* [CPS-12]. United States Census Bureau. http://www.

census.gov/prod/techdoc/cps/cpsoct12.pdf

U.S. Census Bureau. (2013). *Instructions for applying statistical testing to the 2010-2012 ACS 3-year data and the 2008-2012 ACS 5-year data.* United States Census Bureau.

Wang, W., Rothschild, D., Goel, S., & Gelman, A. (2015). Forecasting elections with non-representative polls. *International Journal of Forecasting*, *31*(3), 980–991.

Warshaw, C., & Rodden, J. (2012). How should we measure district-level public opinion on individual issues? *The Journal of Politics*, *74*(1), 203–219.

Zhang, X., Holt, J. B., Lu, H., Wheaton, A. G., Ford, E. S., Greenlund, K. J., & Croft, J. B. (2014). Multilevel regression and poststratification for small-area estimation of population health outcomes: A case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system. *American Journal of Epidemiology*, *179*(8), 1025–1033.

Zumeta, W., Breneman, D. W., Callan, P. M., & Finney, J. E. (2012). *Financing American higher education in the era of globalization.* Harvard Education Press.

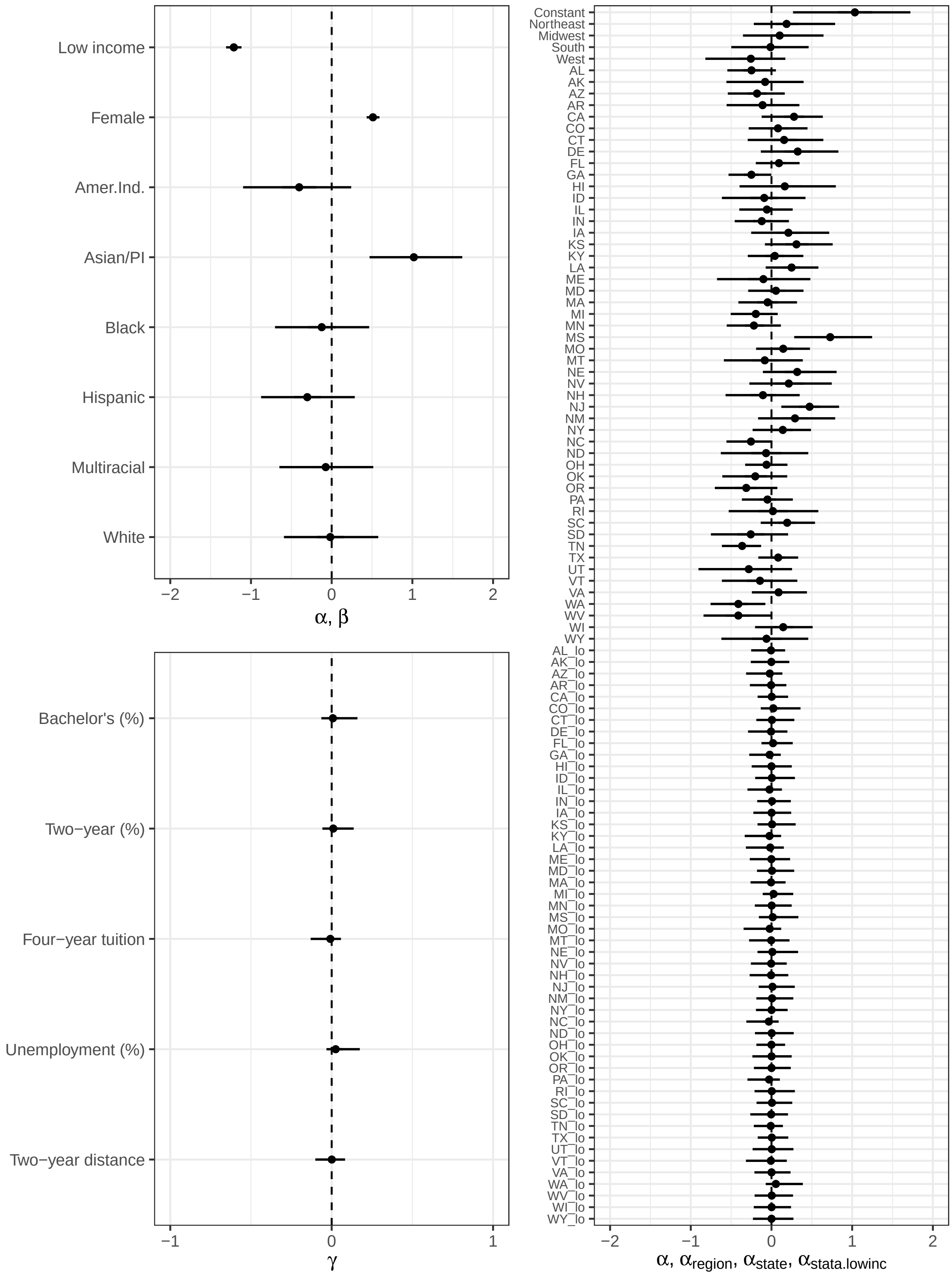

Figure 1: Regression parameters for the full model. Center dots represent $\hat{\theta}_{q50}$, with lines representing 95% credible intervals: $\hat{\theta}_{q2.5}$ and $\hat{\theta}_{q97.5}$.

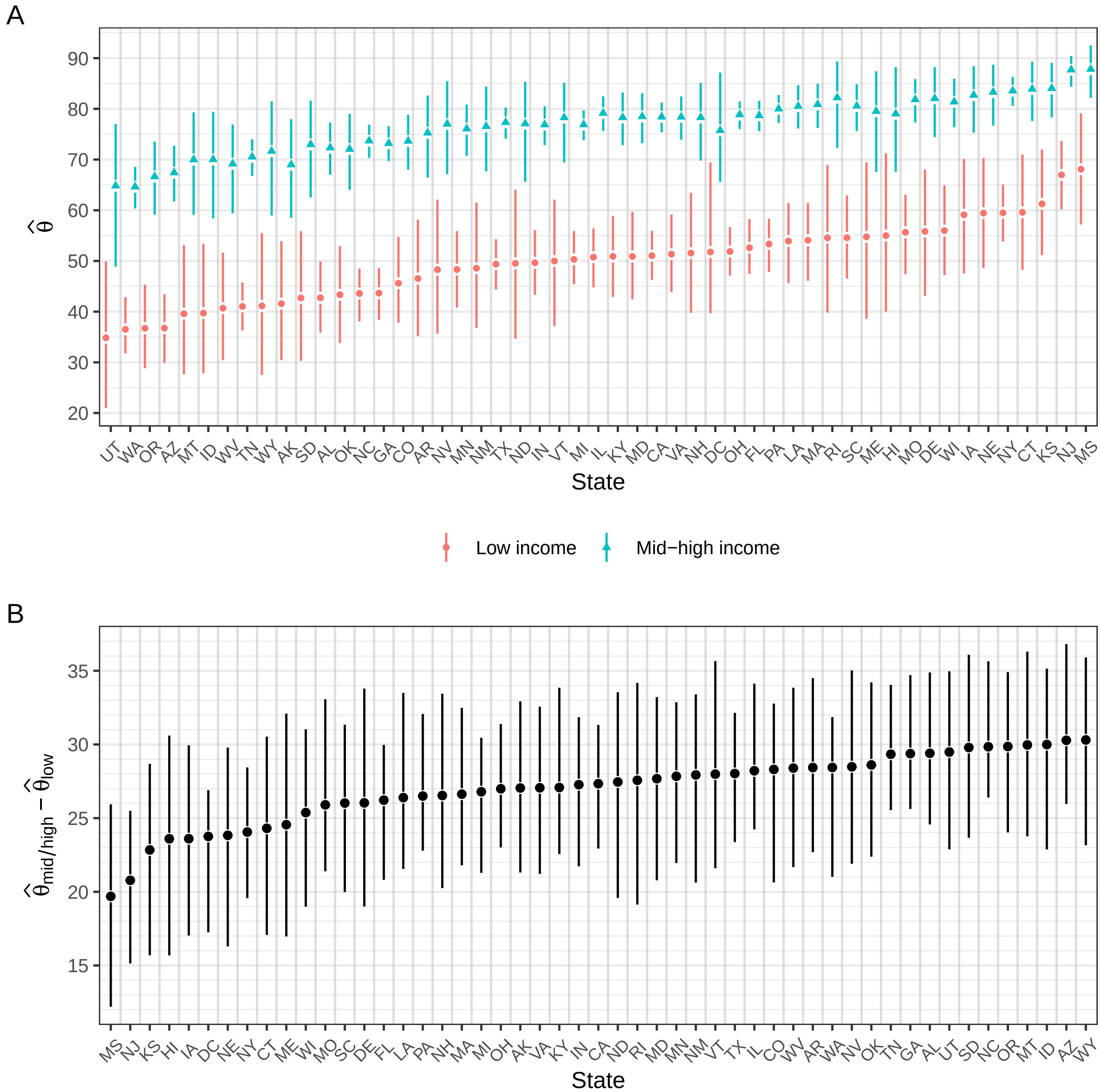

Figure 2: State-level posterior estimates of college enrollment by income. Center dots are $\hat{\theta}_{q50}$ and vertical lines 95% credible intervals. Panel A compares estimates for low-income vs middle-high income students; panel B presents the within-state difference in enrollment by income group.

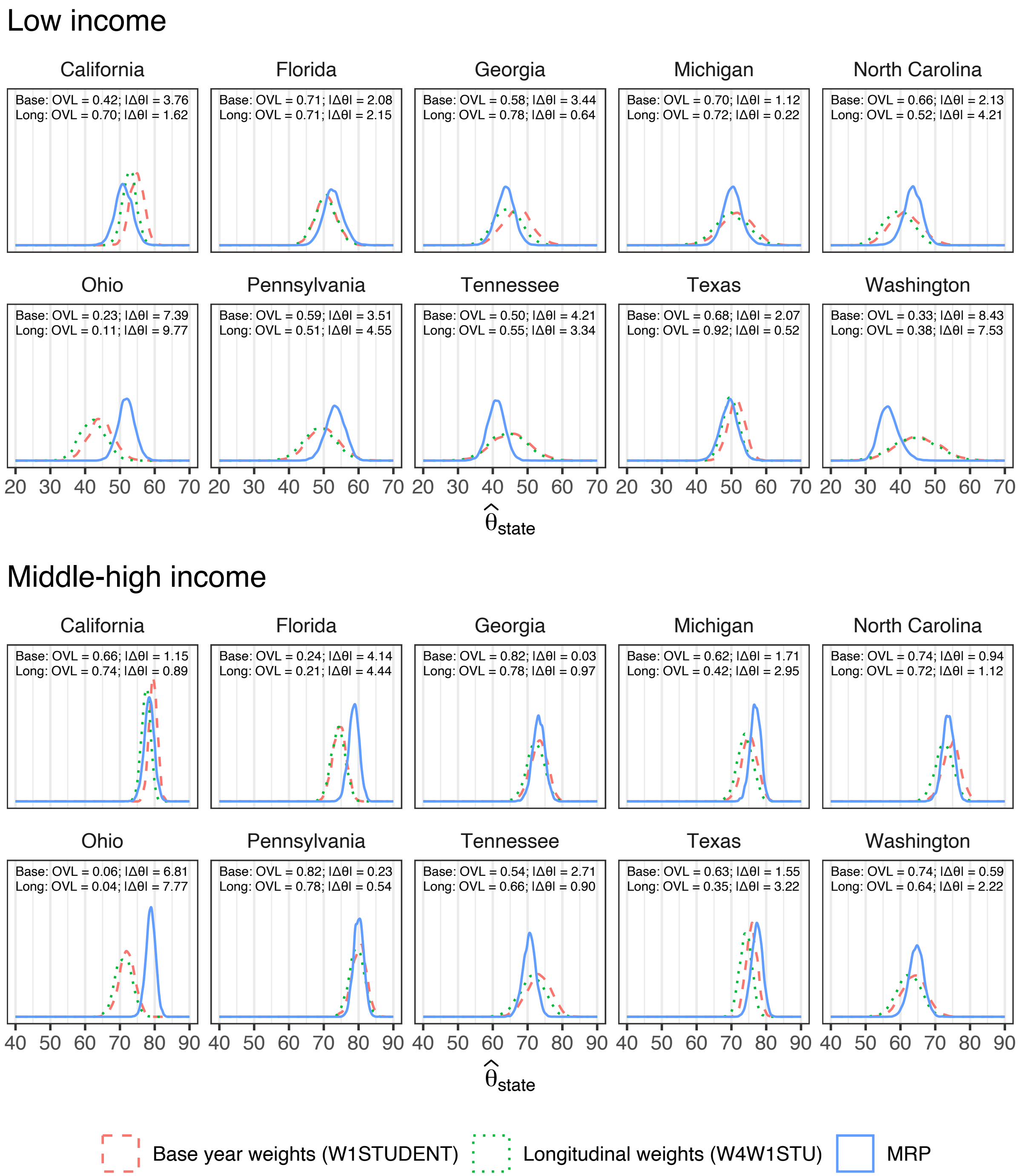

Figure 3: State-level validation comparing posterior estimates using two sampling weights—base year (W1STUDENT, red dashed) and longitudinal (W4W1STU, green dots)—to MRP estimates (solid blue) in 10 states for which HSLS09 is representative. Within each facet, the overlap (OVL) proportion between the MRP posterior and that of each weighted estimate posterior is given. In addition, the absolute percentage point difference between the MRP median and corresponding weighted posterior median is given by $|\Delta\theta|$.

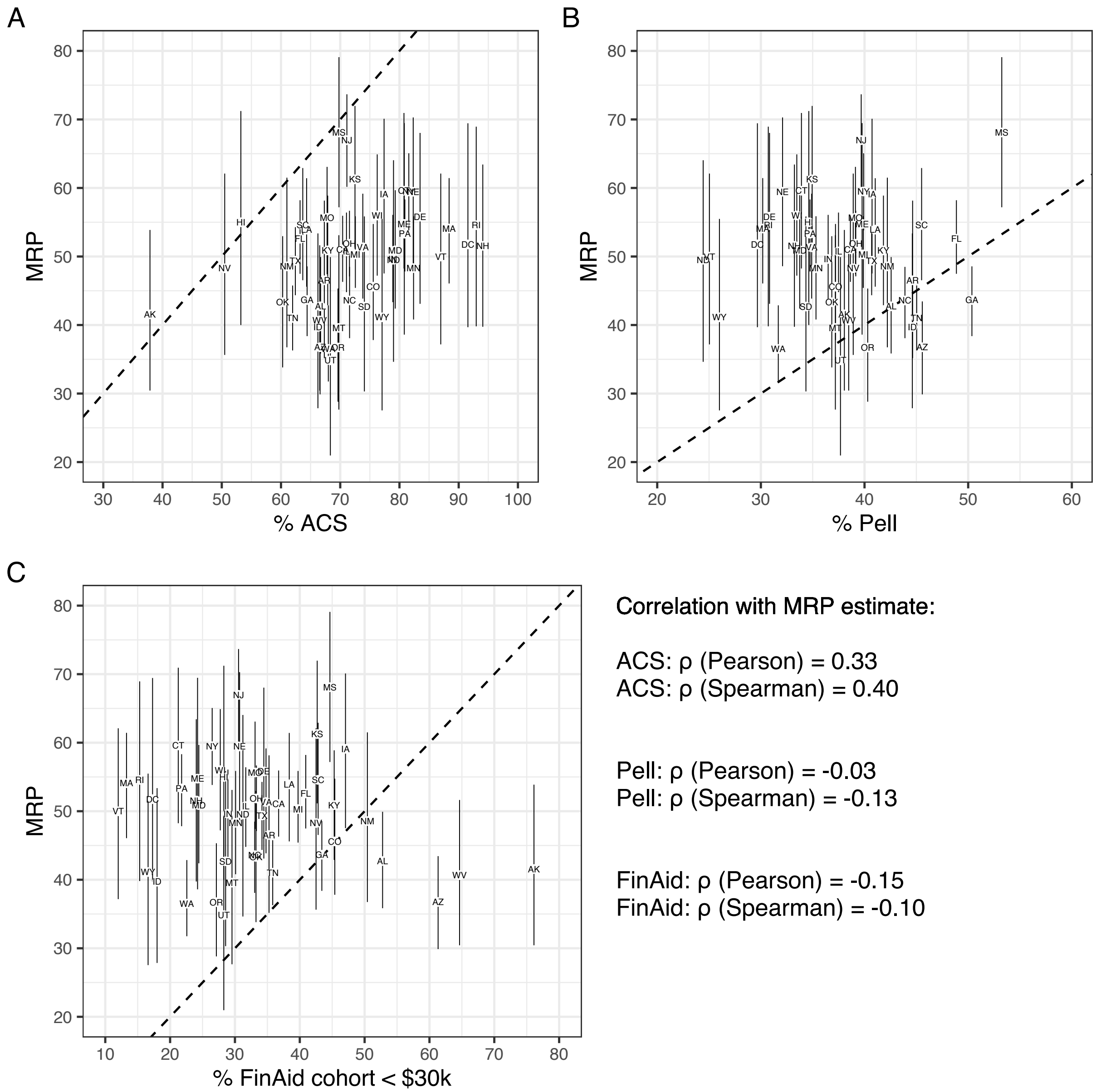

Figure 4: Comparison of MRP estimates to proxy measures of low-income student enrollment. Panel A comparison estimates come from the American Community Survey (ACS) and both institution-level measures in panels B and C come from the Integrated Postsecondary Education Data System (IPEDS); all are aggregated to the state level. Panel A compares MRP estimates of low-income student enrollment to those taken directly from ACS estimates; panel B compares the average student population that receives Pell grant funding; and panel C compares the percentage of the financial aid cohort in the $< \$30,000$ category out of the full time, first time in college cohort. Dashed lines in each figure represent 45 degree lines of parity between measures. Pearson and Spearman rank-order correlation coefficients between MRP estimates and each proxy measure are shown in the bottom right quandrant of the figure.

Table A1: Summary of first and second level regression posterior distributions

	$\hat{\theta}_{q50}$	95% C.I.
First level		
Low income	-1.21	[-1.31,-1.12]
Female	0.51	[0.43,0.59]
American Indian	-0.40	[-1.10,0.24]
Asian / Pacific Islander	1.02	[0.47,1.64]
Black	-0.12	[-0.70,0.46]
Hispanic	-0.30	[-0.88,0.29]
Multiracial	-0.07	[-0.65,0.52]
White	-0.02	[-0.59,0.58]
Second level		
Bachelor's (%)	0.01	[-0.06,0.16]
Four-year tuition	-0.01	[-0.13,0.06]
Two-year (%)	0.01	[-0.06,0.14]
Two-year distance	0.00	[-0.10,0.08]
Unemployment (%)	0.02	[-0.03,0.17]

Notes. 95% credible intervals (*C.I.*) are computed at $\hat{\theta}_{q2.5}$ and $\hat{\theta}_{q97.5}$.

Table A2: Regional and state-specific regression posterior distributions

	$\hat{\theta}_{q50}$	95% C.I.		$\hat{\theta}_{q50}$	95% C.I.
Constant	1.04	[0.27,1.78]			
Region					
Northeast	0.19	[-0.22,0.80]			
Midwest	0.10	[-0.35,0.65]			
South	-0.01	[-0.50,0.46]			
West	-0.26	[-0.82,0.17]			
State			**State × Low-income**		
Alabama	-0.25	[-0.55,0.06]	... × low income	-0.01	[-0.25,0.17]
Alaska	-0.08	[-0.56,0.40]	... × low income	-0.00	[-0.26,0.22]
Arizona	-0.18	[-0.54,0.16]	... × low income	-0.02	[-0.32,0.14]
Arkansas	-0.11	[-0.56,0.35]	... × low income	-0.00	[-0.27,0.18]
California	0.28	[-0.12,0.64]	... × low income	0.00	[-0.17,0.21]
Colorado	0.08	[-0.28,0.45]	... × low income	0.02	[-0.13,0.36]
Connecticut	0.16	[-0.29,0.64]	... × low income	0.01	[-0.19,0.28]
Delaware	0.32	[-0.13,0.83]	... × low income	-0.01	[-0.29,0.20]
Florida	0.09	[-0.19,0.35]	... × low income	0.02	[-0.12,0.26]
Georgia	-0.25	[-0.53,-0.00]	... × low income	-0.02	[-0.28,0.12]
Hawaii	0.16	[-0.40,0.80]	... × low income	0.00	[-0.24,0.25]
Idaho	-0.09	[-0.61,0.42]	... × low income	0.00	[-0.20,0.29]
Illinois	-0.06	[-0.40,0.26]	... × low income	-0.02	[-0.30,0.13]
Indiana	-0.12	[-0.46,0.22]	... × low income	0.01	[-0.18,0.24]
Iowa	0.21	[-0.25,0.71]	... × low income	0.00	[-0.23,0.24]
Kansas	0.31	[-0.08,0.76]	... × low income	0.01	[-0.17,0.30]
Kentucky	0.04	[-0.29,0.40]	... × low income	-0.02	[-0.33,0.12]
Louisiana	0.25	[-0.07,0.58]	... × low income	-0.02	[-0.32,0.15]
Maine	-0.10	[-0.68,0.48]	... × low income	-0.00	[-0.27,0.23]
Maryland	0.06	[-0.29,0.40]	... × low income	0.01	[-0.18,0.28]
Massachusetts	-0.05	[-0.41,0.32]	... × low income	-0.01	[-0.26,0.17]
Michigan	-0.19	[-0.51,0.08]	... × low income	0.03	[-0.11,0.27]
Minnesota	-0.22	[-0.56,0.12]	... × low income	0.00	[-0.21,0.25]
Mississippi	0.73	[0.28,1.25]	... × low income	0.02	[-0.16,0.33]
Missouri	0.15	[-0.19,0.48]	... × low income	-0.02	[-0.35,0.12]
Montana	-0.08	[-0.59,0.39]	... × low income	-0.00	[-0.28,0.22]
Nebraska	0.32	[-0.11,0.81]	... × low income	0.01	[-0.17,0.33]
Nevada	0.21	[-0.27,0.75]	... × low income	-0.00	[-0.26,0.19]
New Hampshire	-0.11	[-0.57,0.35]	... × low income	-0.00	[-0.27,0.21]
New Jersey	0.47	[0.12,0.84]	... × low income	0.01	[-0.16,0.29]
New Mexico	0.29	[-0.17,0.79]	... × low income	0.01	[-0.19,0.27]
New York	0.14	[-0.23,0.49]	... × low income	0.00	[-0.19,0.20]
North Carolina	-0.25	[-0.56,0.00]	... × low income	-0.03	[-0.31,0.09]
North Dakota	-0.07	[-0.63,0.46]	... × low income	0.00	[-0.20,0.28]
Ohio	-0.06	[-0.33,0.20]	... × low income	-0.00	[-0.19,0.17]
Oklahoma	-0.20	[-0.61,0.20]	... × low income	0.00	[-0.24,0.25]
Oregon	-0.31	[-0.70,0.07]	... × low income	0.00	[-0.22,0.24]
Pennsylvania	-0.05	[-0.37,0.27]	... × low income	-0.03	[-0.30,0.10]
Rhode Island	0.02	[-0.53,0.58]	... × low income	0.00	[-0.21,0.29]
South Carolina	0.19	[-0.13,0.54]	... × low income	0.01	[-0.18,0.26]
South Dakota	-0.26	[-0.75,0.21]	... × low income	-0.00	[-0.26,0.20]
Tennessee	-0.36	[-0.61,-0.13]	... × low income	-0.01	[-0.22,0.14]
Texas	0.08	[-0.16,0.33]	... × low income	0.00	[-0.17,0.21]
Utah	-0.28	[-0.91,0.26]	... × low income	0.00	[-0.23,0.27]
Vermont	-0.14	[-0.61,0.32]	... × low income	-0.01	[-0.32,0.19]
Virginia	0.09	[-0.24,0.44]	... × low income	0.00	[-0.21,0.24]
Washington	-0.41	[-0.76,-0.08]	... × low income	0.05	[-0.07,0.39]
West Virginia	-0.41	[-0.84,0.01]	... × low income	0.00	[-0.21,0.27]
Wisconsin	0.15	[-0.20,0.51]	... × low income	0.00	[-0.22,0.24]
Wyoming	-0.06	[-0.62,0.46]	... × low income	0.00	[-0.23,0.27]

Notes. 95% credible intervals (*C.I.*) are computed at $\hat{\theta}_{q2.5}$ and $\hat{\theta}_{q97.5}$.

Table A3: Poststratified estimates of college attendance by state

	Low income		Mid-high income		$\theta_{Mid/high} - \theta_{Low}$	
	$\hat{\theta}_{q50}$	95% C.I.	$\hat{\theta}_{q50}$	95% C.I.	$\hat{\theta}_{q50}$	95% C.I.
Alabama	42.73	[35.85,49.92]	72.34	[67.00,77.31]	29.40	[24.57,34.89]
Alaska	41.56	[30.44,53.88]	68.99	[58.50,77.96]	27.04	[21.31,32.93]
Arizona	36.76	[29.88,43.44]	67.42	[61.72,72.71]	30.29	[25.96,36.82]
Arkansas	46.53	[35.16,58.15]	75.27	[66.43,82.62]	28.43	[22.69,34.51]
California	51.04	[46.30,55.94]	78.42	[75.36,81.25]	27.34	[22.94,31.33]
Colorado	45.60	[37.81,54.75]	73.66	[67.99,78.82]	28.30	[20.64,32.78]
Connecticut	59.60	[48.25,70.94]	83.92	[77.59,89.33]	24.31	[17.07,30.53]
Delaware	55.81	[43.10,68.03]	82.06	[74.43,88.25]	26.04	[19.00,33.79]
District of Columbia	51.76	[39.68,69.44]	75.75	[65.58,87.14]	23.76	[17.25,26.90]
Florida	52.59	[47.48,58.21]	78.69	[75.61,81.57]	26.22	[20.81,29.97]
Georgia	43.65	[38.38,48.56]	73.20	[69.67,76.60]	29.38	[25.62,34.71]
Hawaii	55.00	[40.00,71.23]	79.03	[67.58,88.22]	23.60	[15.68,30.61]
Idaho	39.72	[27.86,53.37]	70.02	[58.38,79.42]	30.00	[22.88,35.15]
Illinois	50.74	[44.80,56.41]	79.17	[75.66,82.47]	28.21	[24.23,34.13]
Indiana	49.64	[43.34,56.07]	76.89	[72.84,80.50]	27.27	[21.73,31.85]
Iowa	59.10	[47.53,70.10]	82.71	[75.31,88.41]	23.60	[17.03,29.95]
Kansas	61.26	[51.15,71.97]	84.05	[78.32,89.08]	22.84	[15.69,28.69]
Kentucky	50.89	[42.89,58.89]	78.31	[72.84,83.23]	27.08	[22.56,33.85]
Louisiana	53.92	[45.60,61.42]	80.51	[76.13,84.65]	26.39	[21.55,33.50]
Maine	54.75	[38.60,69.47]	79.54	[67.54,87.42]	24.55	[16.97,32.09]
Maryland	50.90	[42.40,59.67]	78.49	[73.22,83.06]	27.68	[20.78,33.22]
Massachusetts	54.08	[46.07,61.44]	80.88	[76.22,84.96]	26.63	[21.80,32.49]
Michigan	50.31	[45.42,55.87]	76.90	[73.82,79.69]	26.79	[21.29,30.45]
Minnesota	48.30	[40.80,55.87]	76.09	[70.73,80.84]	27.84	[21.96,32.87]
Mississippi	68.10	[57.18,79.09]	87.83	[82.17,92.50]	19.69	[12.20,25.94]
Missouri	55.65	[47.40,63.08]	81.85	[77.31,85.87]	25.90	[21.39,33.07]
Montana	39.58	[27.66,53.11]	70.00	[59.08,79.31]	29.97	[23.76,36.30]
Nebraska	59.45	[48.58,70.29]	83.30	[76.69,88.72]	23.83	[16.29,29.79]
Nevada	48.28	[35.64,62.10]	77.02	[67.13,85.48]	28.49	[21.90,35.03]
New Hampshire	51.56	[39.74,63.42]	78.31	[69.79,85.12]	26.54	[20.25,33.45]
New Jersey	66.98	[60.17,73.67]	87.71	[84.39,90.43]	20.78	[15.13,25.49]
New Mexico	48.56	[36.75,61.50]	76.56	[67.66,84.42]	27.93	[20.63,33.40]
New York	59.49	[53.83,65.06]	83.57	[80.58,86.30]	24.05	[19.57,28.44]
North Carolina	43.58	[38.08,48.49]	73.70	[70.33,76.85]	29.85	[26.40,35.65]
North Dakota	49.53	[34.63,64.05]	77.09	[65.60,85.35]	27.45	[19.58,33.55]
Ohio	51.86	[47.10,56.67]	78.87	[75.99,81.44]	26.99	[23.01,31.38]
Oklahoma	43.33	[33.81,52.96]	72.07	[64.03,79.02]	28.61	[22.38,34.22]
Oregon	36.73	[28.81,45.32]	66.65	[59.13,73.53]	29.86	[24.03,34.91]
Pennsylvania	53.33	[47.81,58.32]	80.01	[77.22,82.71]	26.50	[22.79,32.07]
Rhode Island	54.57	[39.81,68.94]	82.22	[72.29,89.38]	27.57	[19.13,34.18]
South Carolina	54.58	[46.54,62.91]	80.57	[75.62,84.85]	26.02	[19.98,31.34]
South Dakota	42.69	[30.31,55.85]	72.98	[62.52,81.61]	29.79	[23.66,36.09]
Tennessee	41.03	[36.29,45.76]	70.52	[66.79,74.01]	29.34	[25.54,34.04]
Texas	49.38	[44.35,54.28]	77.34	[74.14,80.25]	28.03	[23.37,32.15]
Utah	34.84	[20.96,49.91]	64.83	[48.88,77.02]	29.49	[22.88,34.97]
Vermont	49.99	[37.17,62.10]	78.30	[69.42,85.15]	27.99	[21.60,35.66]
Virginia	51.34	[43.86,59.17]	78.40	[73.94,82.45]	27.06	[21.21,32.56]
Washington	36.51	[31.76,42.87]	64.63	[60.36,68.55]	28.44	[21.01,31.86]
West Virginia	40.67	[30.44,51.64]	69.16	[59.38,76.92]	28.40	[21.68,33.86]
Wisconsin	55.98	[47.21,64.90]	81.40	[76.37,85.95]	25.38	[18.99,31.03]
Wyoming	41.13	[27.54,55.48]	71.67	[58.96,81.50]	30.31	[23.16,35.91]

Notes. 95% credible intervals (*C.I.*) are computed at $\hat{\theta}_{q2.5}$ and $\hat{\theta}_{q97.5}$.

Table A4: Validation for 10 representative states

			$\hat{\theta}_{q50}$	95% C.I.	Overlap	$\lvert\Delta\hat{\theta}_{q50}\rvert$
California	Low income	MRP	51.04	[46.30,55.94]	-	-
		Base (W1STUDENT)	54.87	[50.72,58.83]	0.42	3.74
		Longitudinal (W4W1STU)	52.75	[48.71,56.56]	0.7	1.66
	Mid-high income	MRP	78.42	[75.36,81.25]	-	-
		Base (W1STUDENT)	79.54	[77.10,81.75]	0.66	1.12
		Longitudinal (W4W1STU)	77.52	[74.91,79.93]	0.74	0.9
Florida	Low income	MRP	52.59	[47.48,58.21]	-	-
		Base (W1STUDENT)	50.56	[44.69,56.55]	0.71	2.13
		Longitudinal (W4W1STU)	50.53	[44.35,56.66]	0.71	2.14
	Mid-high income	MRP	78.69	[75.61,81.58]	-	-
		Base (W1STUDENT)	74.52	[70.68,78.07]	0.24	4.14
		Longitudinal (W4W1STU)	74.17	[70.01,77.98]	0.21	4.48
Georgia	Low income	MRP	43.65	[38.38,48.56]	-	-
		Base (W1STUDENT)	47.03	[38.87,55.19]	0.58	3.45
		Longitudinal (W4W1STU)	44.21	[36.62,52.06]	0.78	0.63
	Mid-high income	MRP	73.20	[69.67,76.60]	-	-
		Base (W1STUDENT)	73.31	[68.38,77.73]	0.82	0.13
		Longitudinal (W4W1STU)	72.26	[67.28,76.99]	0.78	0.89
Michigan	Low income	MRP	50.31	[45.42,55.87]	-	-
		Base (W1STUDENT)	51.52	[42.15,60.16]	0.7	1.13
		Longitudinal (W4W1STU)	50.03	[41.05,58.99]	0.72	0.29
	Mid-high income	MRP	76.90	[73.82,79.69]	-	-
		Base (W1STUDENT)	75.19	[70.61,79.50]	0.62	1.71
		Longitudinal (W4W1STU)	73.92	[69.44,78.27]	0.42	2.94
North Carolina	Low income	MRP	43.58	[38.08,48.49]	-	-
		Base (W1STUDENT)	41.40	[33.16,50.39]	0.66	2.26
		Longitudinal (W4W1STU)	39.38	[31.18,47.36]	0.52	4.15
	Mid-high income	MRP	73.70	[70.33,76.86]	-	-
		Base (W1STUDENT)	74.76	[69.56,79.55]	0.74	1.02
		Longitudinal (W4W1STU)	72.60	[67.48,77.57]	0.72	1.13
Ohio	Low income	MRP	51.86	[47.10,56.67]	-	-
		Base (W1STUDENT)	44.42	[37.48,51.74]	0.23	7.42
		Longitudinal (W4W1STU)	42.17	[35.29,49.25]	0.11	9.72
	Mid-high income	MRP	78.87	[75.99,81.44]	-	-
		Base (W1STUDENT)	72.00	[67.40,76.33]	0.06	6.84
		Longitudinal (W4W1STU)	71.06	[66.41,75.73]	0.04	7.75
Pennsylvania	Low income	MRP	53.33	[47.81,58.32]	-	-
		Base (W1STUDENT)	49.84	[41.04,58.74]	0.59	3.55
		Longitudinal (W4W1STU)	48.80	[39.84,57.64]	0.51	4.56
	Mid-high income	MRP	80.01	[77.22,82.71]	-	-
		Base (W1STUDENT)	80.23	[75.95,84.07]	0.82	0.22
		Longitudinal (W4W1STU)	79.44	[75.26,83.51]	0.78	0.55
Tennessee	Low income	MRP	41.03	[36.29,45.76]	-	-
		Base (W1STUDENT)	45.25	[34.90,56.01]	0.5	4.25
		Longitudinal (W4W1STU)	44.37	[34.34,54.50]	0.55	3.3
	Mid-high income	MRP	70.52	[66.79,74.01]	-	-
		Base (W1STUDENT)	73.20	[65.86,79.41]	0.54	2.69
		Longitudinal (W4W1STU)	71.36	[63.85,77.96]	0.66	0.9
Texas	Low income	MRP	49.38	[44.35,54.28]	-	-
		Base (W1STUDENT)	51.43	[46.61,55.98]	0.68	2.06
		Longitudinal (W4W1STU)	49.81	[45.27,54.48]	0.92	0.53
	Mid-high income	MRP	77.34	[74.14,80.25]	-	-
		Base (W1STUDENT)	75.81	[72.52,78.97]	0.63	1.54
		Longitudinal (W4W1STU)	74.11	[70.76,77.45]	0.35	3.22
Washington	Low income	MRP	36.51	[31.76,42.87]	-	-
		Base (W1STUDENT)	45.04	[33.33,57.50]	0.33	8.37
		Longitudinal (W4W1STU)	44.18	[31.79,56.33]	0.38	7.45
	Mid-high income	MRP	64.63	[60.36,68.55]	-	-
		Base (W1STUDENT)	63.99	[56.93,70.60]	0.74	0.63
		Longitudinal (W4W1STU)	62.44	[55.55,68.92]	0.64	2.2

Notes. Overlap values represent the proportion of the weighted posterior distribution that overlaps with MRP estimates for the state. $\lvert\Delta\hat{\theta}_{q50}\rvert$: the absolute distance in percentage points between the median of the weighted posterior distribution and the median of the MRP posterior distribution.

B.1 Simulation

To demonstrate MRP's validity, we follow the example of Park et al. (2004) and follow the MRP procedure on a synthetic data. Simulated data used in this exercise represent a population of students across the 50 states plus the District of Columbia and has been constructed to resemble the HSLS09 data set we use in our primary analyses of low-income youth enrollment in college.[7] Because we construct these data, however, we know the ground truth against which we can compare estimates recovered using MRP.

The total population size in these simulated data is N = 1,000,027 and is spread out across the states in rough correspondence to the relative population size of each state. Every unit, i, has three observed characteristics: a continuous value of x between 10,000 and 50,000, a categorical value of $v \in \{1, 2, 3, 4, 5, 6\}$, and a categorical value of $w \in \{0, 1\}$. These covariates — x, v, and w — may be thought of as roughly corresponding to a student's income, race/ethnicity, and gender, respectively, as they might appear in an administrative data set or longitudinal survey (hence the limited value ranges). The distributions of these covariates do not correspond to particular groups, however, and we do not place any specific meaning to them beyond their use to separate observations into various subgroups. For each observation, we next construct a dummy variable, D, in which $D = 1$ when $x <$ 22,000 and $D = 0$ otherwise. $D = 1$ for approximately 23% of the population. This covariate corresponds to our indicator for low-income status in terms of its national distribution, though again, we do not give it this particular interpretation in our synthetic data.

We generate a binary outcome measure, y, using a random draw from a Bernoulli distribution in which the latent probability is a linear combination of various individual- and state-level characteristics (normally distributed variables, z, that differ across states with varying degrees of within-region correlation) plus noise. $P(y = 1) = \Theta$ is the average of y across the data set and represents the ground truth population value. When creating values of y across the simulated data set, the following rules applied:

[7] Data are simulated using the R package, `fabricatr` (Blair et al., 2021).

1. Observations with $D = 1$ are less likely to have $y = 1$ than those with $D = 0$, meaning that $\Theta_{D=1} < \Theta_{D=0}$.
2. Values of Θ vary across observed characteristics, w and v, both through main effects and interaction with D.
3. Values of Θ_{state} also vary across states as a function of three state-level covariates, Z, region, and random noise.
4. The relationship between D and Θ_{state} varies across states due to different distributions of individual characteristics, w and v, across the states as well as interactions between D and random state-level noise.

In our analyses, we are interested in obtaining unbiased state-level estimates of $P(y = 1)$ by values of D.

B.1.1 Samples from the simulated population

For our analyses, we draw four samples from the population. Two are samples of $N =$ 10,000 (1% of the population) and two are smaller samples of $N =$ 1,000 (0.1% of the population). Within each sample size, we draw a simple random sample where all population units have an equal probability of selection. For the other two samples, we use weights such that observations with the following characteristics are more likely to be selected: (1) $D = 1$; (2) small v groups (those representing less than 10% of population); and (3) from the Northeastern region of the (simulated) United States. Samples are drawn from the full population of observations with no sub-sampling within region or state. Throughout the next section, we distinguish between true population values (Θ), observed sample values (θ), and estimates ($\hat{\theta}$) of $P(y = 1)$.

B.1.2 Comparison of samples and estimates to population values

Figure B1 compares population differences in Θ by $D = 0$ (left panel) and $D = 1$ (right panel) to estimates using each sample. Population values are presented to the left of the black line within each facet. Except for the population values, all points represent observed values, θ. Though there is some bias in that estimates $\theta_{D=0}$ are a little high whereas estimates $\theta_{D=1}$ are low, they are within ± 5 percentage points (p.p.) of the true value (see Table B1 for precise numbers). As with NCES data sets like HSLS09, national estimates provide a reasonable approximation of the truth.[8]

State-level population differences in Θ_{state} are presented in the top panel of Figure B2. This panel shows that $\Theta_{state|D=1} < \Theta_{state|D=0}$ across all states, with variation in both values of Θ_{state} across the states. Observed values of θ_{state} across the four samples are presented in the four bottom panels of Figure B2 (see Table B2 for precise values). As seen in the figure, state-level values of θ_{state} are generally highly biased. In some states, $\theta_{state} = 100$, meaning that values of D are perfectly collinear with outcomes y. In other states, only one level of D is represented, that is, some conditions are unobserved in that state. Because none of the four samples were stratified by state, it is only within some of the larger states that an estimate $\hat{\theta}_{state}$ (with accompanying estimates of uncertainty, not presented) somewhat approximates the true value, Θ_{state}. Using inverse weights with the weighted samples does not improve state-level estimates, which is expected since the weights apply to a national-level sampling procedure. Figure B2 demonstrates the problem with using national longitudinal surveys to estimate state-level characteristics: given that the sampling procedure is not designed to provide state-level estimates, those estimates will be highly biased in many states.

Figure B3 directly compares observed state-level values of θ_{state} and the true population values, Θ_{state}. Each facet represents a different sample. On the x-axis is the sample value of θ_{state} and on the y-axis is the population (true) value of Θ_{state}. States are plotted by their abbreviation, with red values representing when $D = 1$ and blue green values when $D = 0$. The 45-degree line in each facet represents equality between sample estimate and the true value. Across the four samples, values of θ are on average lower when $D = 1$, shown by their

[8]When including standard errors and inverse sampling weights for the weighted sample, the true population value is contained in the 95% confidence intervals in 10/12 estimates (using weighted samples both with and without weighting adjustment).

comparatively lower position on the 45-degree line. In the larger 1% samples, estimates of θ when $D = 0$ are closer to the line, likely due to their representing about 77% of the simulated population, with correlations between Θ_{state} and θ_{state} between 0.6 and 0.82. In the smaller 0.1% samples, however, correlations when $D = 0$ drop to 0.32.

In all state-specific samples, values of θ among $D = 1$ observations show greater bias from the population truth, particularly for some of the smaller states. In some cases, there are estimates of 0 or 100, suggesting very few observed values in the state. Some states do not have estimates in a particular sample. For example, no observations from Alaska were sampled in the 0.1% weighted sample so an estimate for the state does not show up at all in bottom right facet. Across states that were sampled, bias is even greater among both D conditions in the smaller 0.1% samples, with even more extreme values of θ_{state}. Average correlations between true and sampled values of θ when $D = 1$ range from 0.21 and 0.57 in the 1% sample to 0.12 and 0.42 in the 0.1% sample.

Though these data are simulated, the samples from the population behave as we might expect national surveys designed like HSLS09 to behave. While we can recover reasonable national-level estimates from sample data sets, small area estimates at the state level are highly biased and cannot be recovered with supplied survey weights. The simulation study at this point reflects the reality of the situation for analysts. Using standard techniques, we cannot obtain reliable estimates of the parameter of interest—the probability that a young person from a low-income family will enroll in college. We next turn to our proposed solution to see how well it can provide estimates with credible intervals that cover the known simulated population parameters.

B.1.3 Using MRP on simulated data

To recover state-level estimates, we fit the following multilevel regression to each sample type:

$$P(y_i = 1) = logit^{-1}(\beta_0 + \beta_1 * D_i + \alpha^{state}_{s[i]} + \alpha^{state.D}_{s[i]} + Z_i\gamma). \quad \text{(B1)}$$

In addition to a grand mean parameter, β_0, each state is given a random intercept, α^{state}, and the key covariate of interest, D, is estimated as both a main effect, β_1, and in interaction with each state, $\alpha^{state.D}$. We also include state-level covariates, Z, which we consider to be known from the population data set. Even though we have further observation level covariates, v and w, that we know are part of the data generating process, we fit a simplified model that only separates observations within each state into two demographic categories: $D = 0$ and $D = 1$.

After fitting equation B1 and producing cell-specific predicted probabilities, $\hat{\pi}_j$, we poststratify our estimates using population counts from the original population data set.[9] These counts come from collapsing the population data set (N = 1,000,027) by summing matching demographic cells ($D \in \{0, 1\}$) within each state. This is the only way that we use the population data set in our MRP procedure for the simulation. Figure B4 compares true values of Θ_{state} to the median of our poststratified estimates, $\hat{\theta}_{state}$. As with Figure B3, each facet represents a unique sample. Across all samples, the poststratified estimates are much closer to the true population values. This is most apparent in the $D = 1$ values, which are shrunk toward the 45 degree line (signaling less bias). The $D = 0$ values are even more tightly bunched on the 45 degree line and no state-specific estimate sits at an extreme of 0 or 100. Across all samples, the lowest correlation between the known value of Θ and our MRP estimate is 0.67; the highest is 0.92. Another important result is that we are able to provide estimates for states which were not included in the weighted sample, such as Alaska.

To make the comparison clearer, Figure B5 plots the population truth (Θ_{state}, red circle), observed sample mean (θ_{state}, green triangle), and poststratified estimate ($\hat{\theta}_{state}$, blue square as median of the poststratified posterior distribution with 95% credible intervals) across both

[9]We use the Stan probabilistic programming language (Stan Development Team, 2021) in conjunction with the R programming language (R Core Team, 2021) to fit all Bayesian models in our paper. For equation B1, we assign all regression parameters weakly informative normal priors: $\alpha, \beta, \gamma \sim N(0, \sigma); \sigma \sim N_+(0, 1)$.

0.1% samples across all states. Except in a few cases, the 95% CIs of the poststratified estimates contain the true value across the states, even when the state sample average is either very different from the ground truth or unobserved because the state was not included in the sampling—again, see Alaska in both weighted sample panels, which does not contain an observed sample mean estimate. This is true in both the simple random sample as well as the weighted random sample, the latter of which oversampled units based on values of D as well as characteristics we did not include in the model. Overall, the 95% credible intervals of our MRP estimates contain the true state population value 88 to 98 percent of the time. In addition, the average absolute difference between median MRP posterior values and the true population values range from 3.11 to 6.38 percentage points. This is in comparison to average absolute differences of 14.61 to 25.75 percentage points between true and directly sampled values (ignoring those states for which direct samples were unavailable).

We conclude from this simulation exercise that MRP is a viable solution for recovering state-level estimates of college enrollment among low-income young persons. With data generated to replicate conditions very similar to the ones that apply in our actual data, these simulation results show that MRP generates posterior distributions that overwhelmingly include the true value of the population parameter even when (1) we use a simplified model of the data generating process and (2) the data for a given state is minimal or nonexistent. These results do not mechanically follow from the construction of the sample: while each sampling procedure produces naive estimates that are quite different from the actual values, MRP estimates, which rely on the types of data available to us in our empirical approach, reliably include known true values.

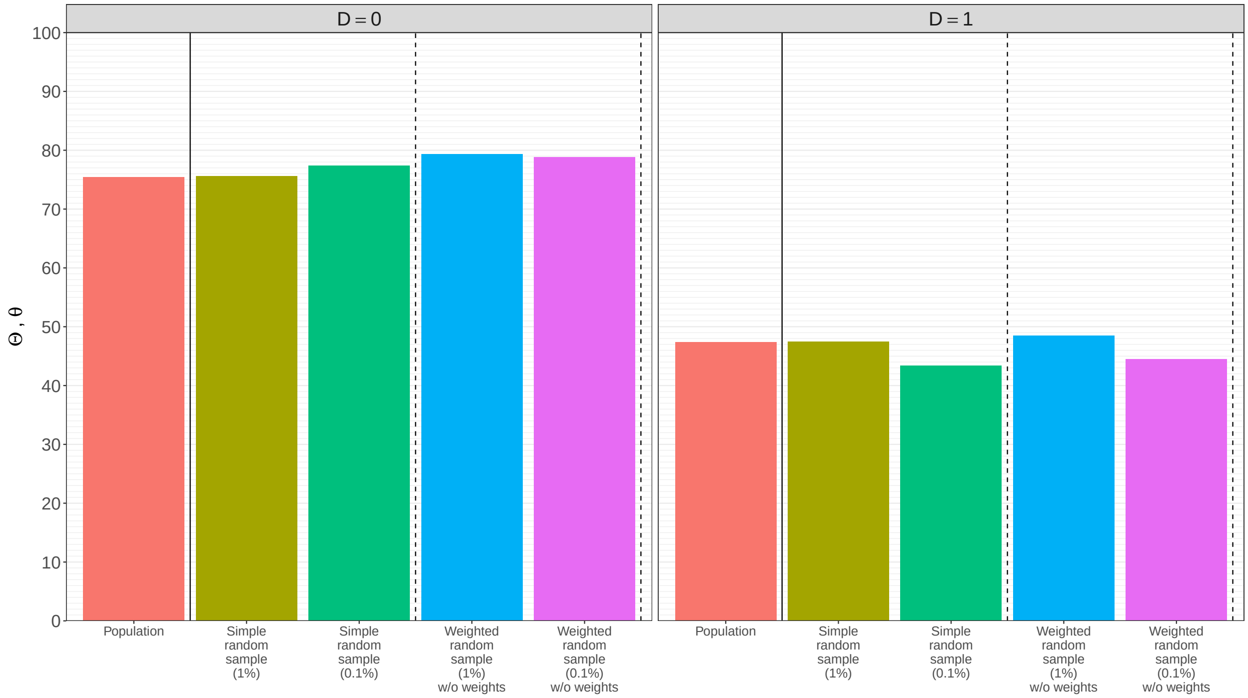

Figure B1: From simulated data, national values of Θ compared to observed values of θ by $D \in \{0, 1\}$. Samples of 1% and 0.1% have $N = 10{,}000$ and $N = 1{,}000$ observations, respectively. Weighted random samples oversample some subpopulations and estimates are presented without sampling weights. All estimates were computed as simple mean statistics.

Figure B2: From simulated data, state-specific values of Θ_{state} (Population) compared to observed values of θ_{state} (Samples) by $D \in \{0, 1\}$. Samples of 1% and 0.1% have $N = 10{,}000$ and $N = 1{,}000$ observations, respectively. Weighted random samples are presented without weighting adjustment.

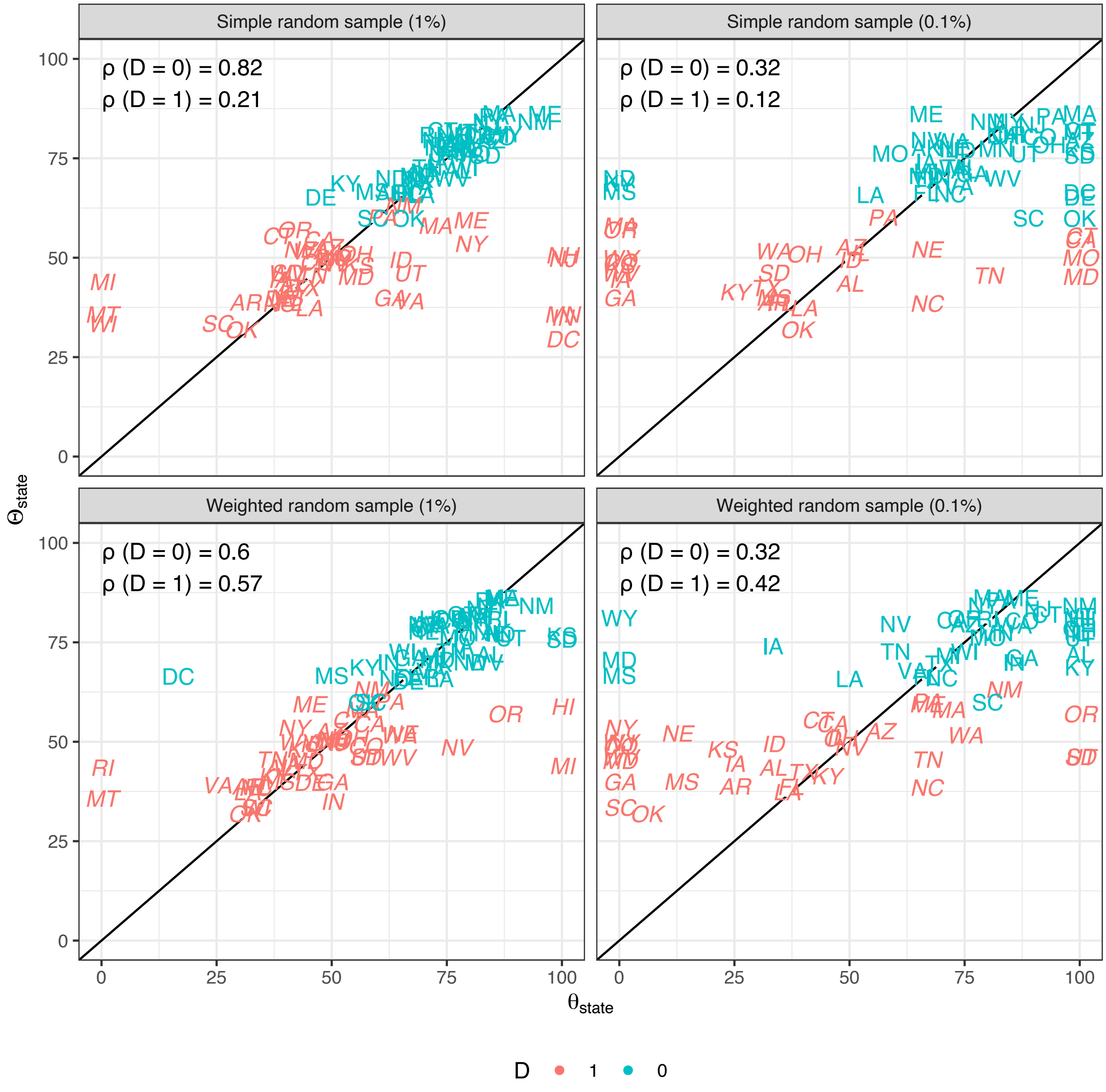

Figure B3: From simulated data, a comparison of observed values of θ_{state} across samples to true values of Θ_{state}. States are plotted twice in each facet, once for each value of D, except in the case when the sample does not have observations for a particular state by D combination. The 45 degree line on each plot represents equality between unobserved true (Θ_{state}) and observed (θ_{state}) values. Weighted samples are presented without weighted adjustments.

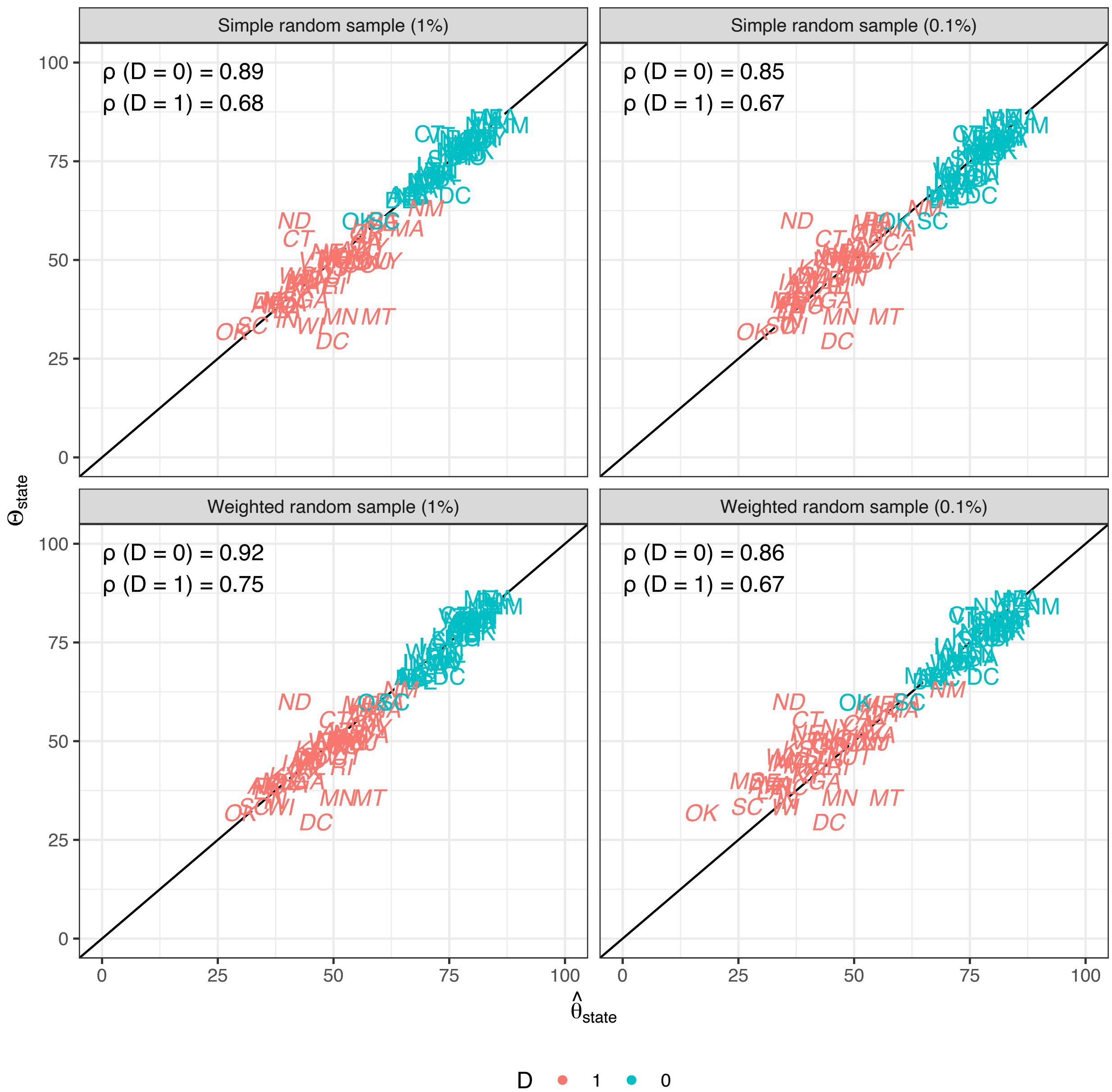

Figure B4: From simulated data, a comparison of poststratified values of $\hat{\theta}_{state}$ across samples to true values Θ_{state}. Again, each state is plotted twice, once for each D condition. Because poststratification estimates can be produced for states with no observations in the sample, there are no missing states in any facet.

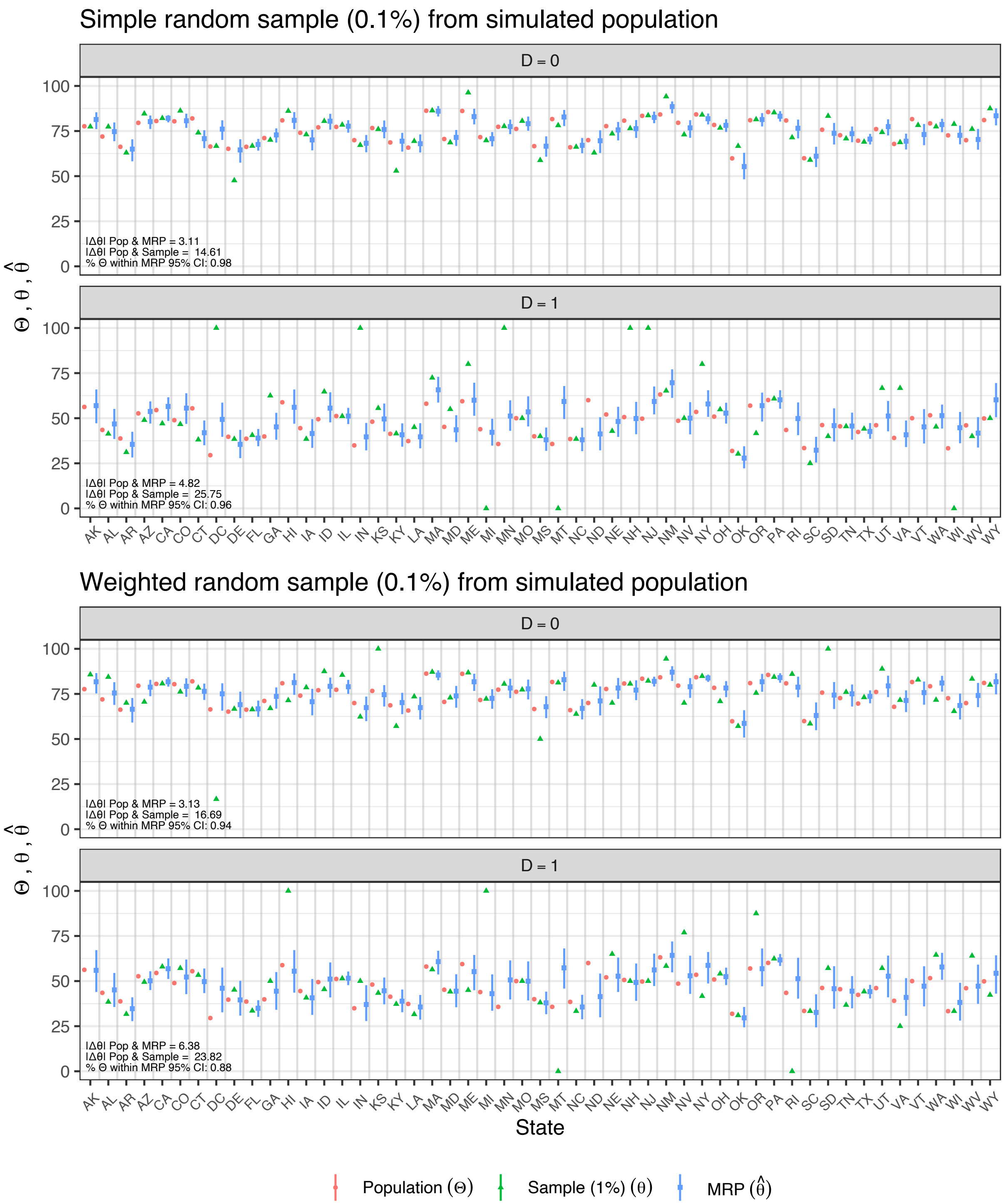

Figure B5: Comparison of true (Θ_{state}), observed (θ_{state}), and poststratified estimates ($\hat{\theta}_{state}$) for the 0.1% simple random sample (top panels) and 0.1% weighted random sample (bottom panels). Lines on the poststratified estimates represent 95% credible intervals. The top and bottom facets for each sample group represent values when D is equal to 0 and 1, respectively.

Table B1: National values of Θ from simulated data

	$D = 0$	$D = 1$
Population	75.41	47.40
	(0.05)	(0.11)
Simple random sample (1%)	75.63	47.50
	(0.49)	(1.07)
Simple random sample (0.1%)	77.42	43.35
	(1.48)	(3.49)
Weighted random sample (1%) w/o weights	79.32	48.48
	(0.56)	(0.73)
Weighted random sample (0.1%) w/o weights	78.83	44.52
	(1.72)	(2.38)
Weighted random sample (1%) w/ weights	76.32	47.46
	(0.59)	(0.73)
Weighted random sample (0.1%) w/ weights	79.65	42.64
	(1.70)	(2.37)

Notes. From simulated data, national values of Θ compared to observed values of θ by $D \in 0, 1$. Samples of 1% and 0.1% have $N = 10{,}000$ and $N = 1{,}000$ observations, respectively. Weighted random samples oversample some subpopulations and estimates are presented without sampling weights. All estimates were computed as simple mean statistics.

Table B2: State-specific samples from simulated population

	Pop		SRS (1%)		SRS (0.1%)		WRS (1%)		WRS (0.1%)	
	$D=0$	$D=1$	$D=0$	$D=1$	$D=0$	$D=1$	$D=0$	$D=1$	$D=0$	$D=1$
AK	77.64	56.25	77.42	-	66.67	-	85.71	-	-	-
AL	71.96	43.50	77.39	41.38	75.00	50.00	84.38	38.46	100.00	33.33
AR	66.26	38.77	62.96	31.15	-	33.33	70.00	31.71	-	25.00
AZ	79.54	52.67	84.62	48.95	100.00	50.00	70.59	49.47	75.00	56.52
CA	80.56	54.51	82.11	47.06	83.33	100.00	80.72	57.99	72.34	46.15
CO	80.39	48.91	86.23	46.67	91.30	0.00	76.19	57.14	87.50	0.00
CT	81.99	55.47	74.07	38.10	100.00	100.00	78.40	53.33	92.86	42.86
DC	66.41	29.51	66.67	100.00	100.00	-	16.67	-	-	-
DE	65.17	39.71	47.62	38.46	100.00	-	66.67	45.16	-	-
FL	66.25	38.64	66.73	40.59	66.67	34.78	66.36	33.54	66.67	37.14
GA	71.11	39.88	70.06	62.50	76.67	0.00	67.01	50.00	87.50	0.00
HI	80.92	58.82	86.11	-	85.71	-	71.43	100.00	100.00	-
IA	74.04	44.47	73.08	38.46	66.67	0.00	78.57	40.74	33.33	25.00
ID	77.04	49.45	80.49	64.71	75.00	50.00	87.50	45.45	-	33.33
IL	77.29	51.19	78.43	51.08	71.43	52.17	85.48	51.43	100.00	46.94
IN	69.94	34.93	67.20	100.00	69.57	-	62.34	50.00	85.71	-
KS	76.66	48.06	76.00	55.56	100.00	0.00	100.00	43.40	-	22.22
KY	68.68	41.34	52.94	41.49	0.00	25.00	57.14	37.27	100.00	45.00
LA	65.79	37.33	69.39	45.00	54.55	40.00	73.53	31.58	50.00	36.36
MA	86.25	58.04	86.36	72.41	100.00	0.00	87.16	56.41	80.56	71.43
MD	70.58	45.20	68.64	55.00	66.67	100.00	72.92	44.12	0.00	0.00
ME	86.13	59.42	96.30	80.00	66.67	-	86.79	45.00	87.50	66.67
MI	71.63	43.90	69.81	0.00	66.67	-	72.22	100.00	71.43	-
MN	77.35	35.71	77.64	100.00	81.82	-	80.52	-	81.82	-
MO	76.19	50.00	80.56	50.00	58.82	100.00	77.33	50.00	80.00	-
MS	66.62	39.96	58.82	40.00	0.00	33.33	50.00	38.10	0.00	13.33

	Pop		SRS (1%)		SRS (0.1%)		WRS (1%)		WRS (0.1%)	
	$D=0$	$D=1$	$D=0$	$D=1$	$D=0$	$D=1$	$D=0$	$D=1$	$D=0$	$D=1$
MT	81.63	35.71	78.12	0	100	-	81.25	0	-	-
NC	66	38.46	66.19	38.46	71.88	66.67	63.86	33.33	70	66.67
ND	69.97	60	62.96	-	0	-	80	-	-	-
NE	77.73	52.07	73.53	42.86	72.73	66.67	70	65	100	12.5
NH	80.76	50.61	76.47	100	83.33	-	80.6	50	100	-
NJ	83.42	49.75	83.66	100	89.66	-	82.22	50	90.91	0
NM	84.19	63.14	94.12	65.22	80	-	94.44	58.33	100	83.33
NV	79.58	48.57	73.03	50	66.67	-	70	76.92	60	50
NY	84.22	53.49	83.99	80	84.48	-	84.85	41.67	79.1	0
OH	78.37	50.88	76.74	54.86	93.33	40	70.89	54.08	100	47.83
OK	59.85	31.87	66.67	30.23	100	38.46	57.14	31.02	-	5.88
OR	80.99	56.98	81.43	41.67	84.62	0	75.51	87.5	75	100
PA	85.54	60.09	85.25	60.82	93.75	57.14	84.36	62.39	82.86	66.67
RI	80.86	43.48	71.43	-	-	-	86	0	80	-
SC	59.93	33.45	58.96	25	88.89	-	58.49	33.33	80	0
SD	75.71	46.19	83.33	40	100	33.33	100	57.14	-	100
TN	72.67	45.51	70.81	45.45	73.33	80	76	36.73	60	66.67
TX	69.64	42.37	69	44.12	66.67	31.58	73.09	44.16	69.57	39.58
UT	76.08	46.11	74.34	66.67	88.24	-	88.89	57.14	100	100
VA	67.85	39.07	68.75	66.67	73.68	-	71.57	25	63.64	-
VT	81.6	50	78.26	-	100	-	82.86	-	100	-
WA	79.25	51.63	77.46	45.24	72.22	33.33	71.67	64.49	85.71	75
WI	72.61	33.33	78.85	0	72.41	-	65.38	33.33	75	-
WV	69.92	46.05	76.09	40	83.33	0	83.33	64	-	0
WY	81.08	49.86	87.5	50	-	0	80	42.31	0	0

Notes. *SRS*: simple random sample; *WRS*: weighted random sample. From simulated data, state-specific values of Θ_{state} (Population) compared to observed values of θ_{state} (Samples) by $D \in 0,1$. Samples of 1% and 0.1% have N = 10,000 and N = 1,000 observations, respectively. Weighted random samples are presented without weighting adjustment. Values with hyphen represent samples in which no observations were observed.