Equity and Efficiency in Financial Aid Targeting*

Juliana Londoño-Vélez
Luis Esteban Álvarez
Catherine Rodríguez
Fabio Sánchez

This version: October 31, 2025.

Abstract

In public finance models, a planner allocates transfers to balance equity–efficiency trade-offs. These are central in higher education, where targeting combines SES and academic ability. We study how financial aid targeting affects equity and efficiency by comparing incidence, treatment effects, and MVPFs across SES and test scores using RD and boundary discontinuity designs. We find unambiguous welfare gains from targeting low-SES high-achievers, with substantial treatment effect heterogeneity. Targeting higher-scoring students reduces progressivity, enrollment effects, and the MVPF—especially when students respond to policy incentives. In contrast, targeting poorer high-achievers enhances both equity and efficiency, aligning redistributive and efficiency goals. **JEL:** H21, H23, I22, I24

*Corresponding author: Juliana Londoño-Vélez, Assistant Professor of Economics, Department of Economics, UCLA, Bunche Hall 8283, 350 Portola Plaza, Los Angeles, CA 90095, USA, (e-mail: j.londonovelez@econ.ucla.edu). Luis Esteban Álvarez is a Ph.D. student at UCLA. Catherine Rodríguez and Fabio Sánchez are researchers in the Department of Economics at Universidad de Los Andes. We have benefited from comments and suggestions from Susan Dynarski, Yotam Shem-Tov, Adriana Lleras-Muney, Damián Vergara, Natalie Bau, Felipe Goncalves, Daniel Haanwinckel, Pat Kennedy, Matt Notowidigdo, Jishnu Das, and numerous seminar participants at UCLA, the All-California Labor Economics Conference at Stanford University, Oxford University, Georgetown University, Texas A&M, the World Bank, and Banco de la República. Luis José Zapata Bobadilla provided outstanding research assistance. We are grateful to Patricia Moreno, Alejandro Corrales, and Yamit López from Colombia's Department of National Planning (DNP). We gratefully acknowledge financial support from the Spencer Foundation through Large Grant #10040436. This study was approved by UCLA's IRB (IRB#24-000220).

1 Introduction

In canonical public finance models, a social planner allocates transfers to maximize a social welfare function subject to informational and budget constraints (Mirrlees, 1971). Optimal transfers trade off *equity*—captured by social marginal welfare weights, which reflect how much society values transfers to different groups—and *efficiency*, captured by either individuals' behavioral response (Saez, 2002) or, in the marginal value of public funds (MVPF) framework, by the fiscal externalities that determine the net cost of transfers (Hendren and Sprung-Keyser, 2020).

These trade-offs are especially salient in higher education, where targeting typically combines socioeconomic status (SES) and academic ability. Targeting low-SES students promotes redistribution and can raise enrollment under liquidity constraints, but may generate limited fiscal externalities if few complete college. Conversely, targeting high-achieving students can yield large fiscal externalities if ability proxies productivity and may reflect social preferences for merit, but risks regressive incidence if ability is correlated with SES. Real-world policies—such as need- and merit-based financial aid or automatic admission for top students (e.g., Texas's top 10% rule, California's UC top 9% guarantee)—embody these tensions. Designing optimal policies requires understanding how targeting shapes incidence, treatment effects on planner-valued outcomes, and the MVPF. Yet rigorous evidence remains scarce.

This paper fills this gap by providing causal evidence on how financial aid targeting shapes equity and efficiency. We use rich administrative microdata and quasi-experimental variation from Colombia, where aid eligibility depends on both standardized test scores and a proxy-means test (PMT) for household poverty. We first benchmark the equity–efficiency trade-off by comparing the incidence of aid (who receives it), treatment effects, and MVPFs across test scores and poverty using regression discontinuity (RD) and boundary discontinuity designs applied to an initial cohort whose eligibility was determined before students knew of the policy. We then reexamine this trade-off when students can respond strategically, exploiting later reforms that shifted the eligibility thresholds over time.

We find unambiguous gains from targeting low-SES high achievers: financial aid substantially improves student outcomes and generates fiscal externalities. However, the design of aid targeting substantially shapes both equity and efficiency. First, the incidence of aid varies sharply across test scores and SES. Following Abadie (2003), we characterize *compliers*—students who receive aid when eligible—and show

that targeting those in the bottom half of the SES distribution channels resources toward students from poor neighborhoods, public high schools, and first-generation college backgrounds, along with other indicators of disadvantage. Conditional on SES, targeting students in the top 10% of test scores reduces progressivity because test scores are strongly correlated with baseline SES. Crucially, substantial heterogeneity exists even within the top decile: the share of first-generation college students among compliers falls by nearly 50% between the 91st and 99th percentiles of test scores, with similar declines across other SES indicators.

To assess efficiency, we estimate heterogeneous treatment effects (HTEs) on planner-valued outcomes such as college enrollment, college attainment, and expected post-college earnings. A canonical rationale for public investment in financial aid is that individuals may underinvest in higher education when credit constraints limit borrowing, when behavioral biases distort educational choices, or when education generates positive externalities (Dynarski et al., 2022; Lochner, 2011). These distortions justify a social welfare function that places positive value on expanding college access and attainment. We find that financial aid increases college enrollment across the distributions of test scores and SES, but the effects decline sharply with ability. The largest gains occur among low-SES students near the 91st percentile of test scores, while effects fall by roughly 10 percentage points (p.p.) between the 91st and 99th percentiles—producing a similarly steep drop in graduation impacts. Even among low-SES students, those with higher test scores are disproportionately inframarginal—that is, likely to enroll and graduate regardless of aid—so their relative treatment effects are significantly smaller. For these students, financial aid primarily affects the *intensive margin*, shifting the type or quality of institution attended rather than enrollment itself.

Another rationale for public investment in financial aid is the fiscal externality it creates through the tax-and-transfer system: by increasing lifetime earnings, college attendance and attainment raise individuals' lifetime tax contributions under a progressive income tax. We show that colleges differ enormously in their value added to future earnings. Financial aid substantially increases enrollment in higher–value-added institutions across the joint distributions of test scores and SES. Because high-scoring students use aid to access higher-return but more expensive colleges, expected earnings gains remain flat across test scores, while fiscal costs rise—implying that the MVPF likely declines with test scores.

In contrast, targeting high achievers from poorer backgrounds enhances redistribution

without sacrificing efficiency. Conditional on test scores, absolute increases in enrollment and projected earnings are similar across the SES distribution, while relative gains are larger among poorer students: those from lower-SES backgrounds start with much lower baseline enrollment and graduation rates—particularly at high–value-added institutions—yet experience substantial projected earnings gains when they do enroll. The MVPF exceeds one for all high achievers below the median SES level and remains stable or rises with poverty, indicating welfare gains from targeting this group. Under standard social welfare weights, these findings imply that optimal targeting of financial aid should prioritize socioeconomic disadvantage over academic rank—in this setting, equity and efficiency are mutually reinforcing.[1]

Finally, we reassess the equity–efficiency trade-off when students can respond strategically—that is, when test scores and measured SES become 'endogenous' to policy incentives. Between 2014 and 2019, the policy kept the *poverty* cutoff fixed but gradually raised the *test-score* threshold for eligibility from the 91st to the 98th percentile. Exploiting this variation, we find that higher test-score thresholds shift aid toward more advantaged students, reduce extensive-margin enrollment, and concentrate effects on the intensive margin, consistent with the baseline analysis. Strategic responses amplify these patterns because those who adjust are more advantaged and inframarginal to enrollment. As a result, enrollment effects decline sharply—from 29 p.p. in the first year (91st-percentile cutoff) to a statistically insignificant 3 p.p. by year five (98th-percentile cutoff). The impacts shifts toward college choice, particularly into higher–value-added institutions. Yet neither observed nor predicted earnings gains rise proportionally, indicating that tightening the test-score threshold does not enhance the policy's fiscal externality.

Our findings contribute to three strands of the literature. First, we build on extensive research on financial aid policy, as reviewed by Dynarski and Scott-Clayton (2013), Nguyen et al. (2019), and Dynarski et al. (2022). This work estimates the causal effects of need-based, merit-based, and hybrid aid programs, which target different populations and yield varied impacts on student outcomes. Program design matters: need-based aid typically increases enrollment and completion, while merit-based aid often benefits higher-income, inframarginal students. Hybrid programs combining SES and academic ability tend to produce large effects for certain groups (Bettinger et al., 2019), highlighting efficiency differences by targeting rule. Understanding

[1] We reach similar conclusions using the causal random forest (CRF) method by Athey et al. (2019) to estimate HTEs.

treatment effect heterogeneity is important for designing welfare-maximizing policies, yet evidence remains limited. We contribute by showing how eligibility design shapes recipients' sociodemographic composition, treatment effects, the behavioral margins of response, and MVPFs.[2]

Second, our results inform the optimal design of financial aid policy. They complement Colas et al. (2021), who use a structural model of selection into college with NLSY data to derive the optimal financial aid schedule. In contrast, we provide reduced-form causal estimates of treatment effects across the joint distribution of SES and academic ability. We evaluate targeting rules based on these two dimensions—common in real-world programs—and quantify the resulting equity–efficiency trade-off. We also account for heterogeneity in returns across postsecondary institutions, which is substantial in the U.S. and elsewhere (Lovenheim and Smith, 2023). Consistent with theory (Brown et al., 2011), there are sizable welfare gains from targeting financial aid to low-SES high achievers—not only because it improves redistribution but also because the higher-return institutions they can access maximize the fiscal externality of aid.

Third, we contribute to a growing literature that leverages treatment effect heterogeneity to improve the welfare impacts of policy (e.g., Aiken et al., 2022; Athey et al., 2025; Bertrand et al., 2024; Bhattacharya and Dupas, 2012; Björkegren et al., 2025; Haushofer et al., 2025). Most of this work uses randomized interventions—typically cash transfers—combined with machine learning to estimate HTEs and target programs based on predicted impacts rather than observable traits such as poverty. We extend this framework to higher-education policy, where targeting decisions are especially consequential. Unlike cash transfers—where recipients typically value each dollar similarly—education vouchers generate sharply different efficiency gains across recipients. Our results show that common eligibility rules based on standardized test scores can yield systematically suboptimal allocations because test scores are negatively correlated with both poverty and treatment effects. When targeting low-SES high achievers, prioritizing high scorers may inadvertently exclude the students who would benefit most—.

The remainder of this paper is organized as follows. Section 2 describes the background and data. Section 3 presents the methodology. Section 4 analyzes

[2] Recent work by Barahona et al. (2025) underscores the importance of targeting in student financial aid policy. They focus on supply-side responses—specifically, how targeting affects tuition—whereas we examine how targeting affects recipients' SES and treatment effects.

the baseline equity–efficiency trade-off, while Section 5 reassesses the trade-off once students endogenously respond to policy incentives. Finally, Section 6 offers a discussion.

2 Background and Data

Undergraduate admissions in Colombia are decentralized and major-specific: applicants apply directly to specific major–college pairs. Admission decisions primarily depend on performance in the national standardized high school exit exam, SABER 11, which assesses critical reading, mathematics, social and citizenship studies, natural sciences, and English. Nearly all high school seniors take this exam, regardless of whether they plan to pursue higher education.

To allocate social transfers to households, the Colombian government uses SISBEN, a PMT. During our study period, households received a continuous score ranging from 0 to 100 (poorest to richest) based on factors like major assets, household composition, education, and *estrato*, another proxy of SES.[3] We use the SISBEN score as our primary measure of poverty, classifying students as 'low-SES' if their score is below a cutoff near the median, and 'high-SES' otherwise. Students without a SISBEN score are assumed to be high-SES and are included in the analysis.[4]

Our empirical analysis draws on administrative data from several sources:

1. **Test scores.** We use the population of SABER 11 test takers from the *Instituto Colombiano para la Evaluación de la Educación* (Icfes), which oversees standardized testing in Colombia. These data include test scores and sociodemographic information (e.g., residential stratum, parental education, sex) and cover the fall semesters of 2012–2019.[5]

2. **Poverty.** We use information from DNP's SISBEN registry from 2012 to 2019, including all variables used in the PMT.

[3] The *estrato* is based on neighborhood and housing quality and ranges from 1 to 6 (poorest to richest).
[4] Because SISBEN is an application-based PMT that households typically complete only when they seek means-tested benefits, most non-applicants are in practice higher-SES. In 2014, roughly one-third of high-school test takers did not have a SISBEN score.
[5] From this sample, we exclude 2% of test takers who had previous college experience before (re)taking the exam. We also drop about 5% of individuals who were not registered high school students when taking SABER 11, since program beneficiaries had to be high school students.

3. **Policy take-up.** We use administrative records from ICETEX, the agency managing Colombia's financial aid programs. These include recipients of *Ser Pilo Paga* (SPP; 2014–2017) and its successor, *Generación E-Excelencia* (2018–2019). The data also report direct educational costs.

4. **College enrollment.** We use information from the Ministry of Education's *Sistema Nacional de Información de la Educación Superior* (SNIES), which tracks students in the postsecondary education system and provides panel information on enrollment status, institution, type of degree program (e.g., associate, bachelor's), and field of study (e.g., STEM) from 2013 to 2020.

5. **College completion.** We use administrative data from Icfes on all students who took Colombia's mandatory college graduation exams. Bachelor's degree students must take the SABER PRO exam to graduate. Prior to 2016, associate degree students were also required to take the SABER PRO exam; starting in 2016, they must take the SABER T&T exam instead. We observe SABER PRO test takers from 2013 to 2021 and SABER T&T test takers from 2016 to 2021. We use exam participation as a proxy for college graduation.

6. **Value-added and labor-market outcomes.** We construct measures of college-by-major "value-added" using data from Icfes and SNIES, linked to social security records from Colombia's Ministry of Health and Social Protection's *Planilla Integrada de Liquidación de Aportes* (PILA). PILA provides detailed payroll and earnings information for all formal workers in Colombia. Our main value-added measure is based on two pre-reform cohorts—those who took the SABER 11 exam in 2012 and 2013—and tracks their earnings nine years later, in 2021 and 2022, respectively. We also measure earnings through 2023.

7. **Educational expenditures.** We complement ICETEX data on direct program costs with expenditure information from financial accounts and balance sheets reported to Colombia's Ministry of Education. Together, these sources allow us to estimate the total cost of attendance, including government subsidies to public institutions per enrollee.

We now present four key background facts.

Fact #1: SES predicts test scores. Figure A.1 plots the distribution of high school

test scores by SES. High-SES students (top half of the wealth distribution) perform about 17 points, or over half of a standard deviation (σ) higher than low-SES students (bottom half of the wealth distribution) on the high school exit exam. They are also about three times more likely than low-SES students to score in the top decile of the test-score distribution.

Table A.1 presents descriptive statistics for student test takers, ranked by their test score and SES. Compared to the average student in Column (1), high-achievers—those scoring at least 310 out of 500 (top 9.3%) in Column (2)—tend to come from more advantaged backgrounds. They have higher SISBEN wealth scores and residential strata, both proxies of SES. They are nearly four times more likely to have college-educated parents and more than twice as likely to attend private high schools and study full-time. High achievers are also more likely to live in urban areas, particularly Bogotá, and less likely to reside in economically vulnerable regions.[6] Additionally, they are more often white and male.

Restricting the sample to low-SES students in Column (3) reduces some socioeconomic disparities. However, compared to the average student in Column (1), the 'low-SES high achievers' in Column (3) remain more likely to have college-educated parents, less likely to live in the poorest stratum, and more likely to be male. Importantly, this group is highly heterogeneous, and test scores remain strongly correlated with baseline SES. For instance, students scoring 359 or higher (top 2%) are over ten p.p. less likely to live in the poorest stratum than those scoring 310 or higher (top 10%). They are also nearly twice as likely to have college-educated parents, more likely to attend private high schools, and more often male. This persistent correlation between test scores and SES implies that financial aid policies based on test scores—even when conditioned on financial need—fail to reach the most vulnerable students.

Fact #2: Colleges vary substantially in quality and price. We use two different measures of educational quality. First, we consider a government-certified quality status known as "High-Quality Accreditation" (HQA). Colleges earn this status through a peer review process, and researchers and policymakers widely regard it as a proxy for educational quality (OECD, 2016). By 2014, about 12% of colleges had achieved HQA, enrolling about one-fifth of all students. HQA colleges outperform others in key metrics, including higher entry and exit test scores, graduation rates,

[6] Amazonas, Arauca, Caquetá, Casanare, Cesar, Chocó,, Guainía, Guaviare, La Guajira, San Andrés, Vaupés, and Vichada.

per-student spending, and a greater proportion of faculty with Ph.D. degrees (Table A.2).

Second, we assess college quality by estimating its contribution to students' future earnings. Following Londoño-Vélez et al. (forthcoming), we estimate earnings impacts nine years after high school for individuals who enrolled in a given college–major pair, regardless of completion, controlling for baseline individual, household, and high school characteristics. Figure A.2 shows substantial heterogeneity in value-added across college types—some institutions significantly raise earnings, while others yield little or even negative returns. Private HQA colleges exhibit the highest earnings value-added but are generally more expensive (Table A.2). Before the policy, these colleges rarely offered financial aid, resulting in a strong correlation between earnings value-added and tuition (Riehl et al., 2018).

Fact #3: Without financial aid, there are large socioeconomic gaps in college enrollment and quality. Access to college—especially to expensive, high–value-added colleges—depended heavily on a student's financial resources, and limited financial aid prevented many low-SES students from enrolling in college or accessing high-quality colleges. Table A.3 shows that over 80% of low-SES students do not enroll in college immediately after high school, and fewer than 1% attend private HQA colleges. Most who do enroll choose non-HQA institutions. In contrast, high-SES students are 11 p.p. more likely to enroll in any college, three times more likely to attend an HQA college, and six times more likely to enroll in a private HQA college.

Test scores strongly predict college enrollment: students in the top 10% are about 30 p.p. more likely to enroll than the average student. Yet even among these high achievers, low-SES students are roughly 10 p.p. less likely to attend college than their high-SES peers. These disparities persist over time. Six years later, high-achieving high-SES students are about 15 points more likely to have ever enrolled in an HQA college, nearly 20 points more likely to attend a private HQA college, and 4 points more likely to attend any college at all than equally high-achieving low-SES students.

Fact #4: Financial aid improves college enrollment and quality. To reduce these socioeconomic disparities, the national government launched a forgivable student loan program in 2014 called SPP. The program covered full tuition for any bachelor's degree, provided the student enrolled in an HQA college. In its first year, eligibility required (1) a SABER 11 exam score of 310 or higher—at or above the 90.7th

percentile—and (2) economic disadvantage, as determined by the student's SISBEN score (Figure 1).[7] Among 2014 high school test takers, 53% met the poverty threshold; as described previously, we refer to these students as 'low SES' and all others as 'high SES.'

Exploiting the policy's sharp eligibility thresholds, Londoño-Vélez et al. (2020, forthcoming) show that financial aid strongly influences college-going behavior. Students just above the cutoff are significantly more likely to enroll—particularly in HQA colleges—raising the earnings value-added of the college they attend. We replicate their findings in Column (1) of Table 1: in the policy's first year, eligibility increased overall enrollment by 29.4 p.p., HQA enrollment by 48.0 p.p., and the earnings value-added of the college and major attended. In short, financial aid not only expands access to college but also steers students toward higher-quality institutions with greater long-run returns, ultimately boosting post-college earnings.

3 Methodology

Our goal is to identify for whom financial aid is most effective—and, therefore, whom to target when resources are limited. This section will introduce three complementary econometric approaches: (i) a multidimensional RD design that exploits strict eligibility cutoffs and policy changes over time; (ii) a boundary discontinuity design that traces treatment effect heterogeneity along the full eligibility frontier in the baseline period; and (iii) causal forest models that uncover heterogeneity across the broader covariate space, beyond the eligibility thresholds.

We denote by $Y \in \{1, 0\}$ the random variable that expresses a binary outcome—such as whether a student enrolls in college immediately after high school ($Y = 1$) or not ($Y = 0$)—and $T \in \{1, 0\}$ for whether the student receives financial aid ($T = 1$) or not ($T = 0$).[8] We use standard potential-outcomes notation and write $Y(1)$ for the enrollment status a student would have if they had been assigned to the treatment, and $Y(0)$ for the enrollment status had they been assigned to control. The treatment of that student is then $Y(1) - Y(0)$, and we are interested in how this treatment effect varies with some baseline characteristics X (e.g., test score, poverty level). Specifically, we

[7] The applicant's SISBEN index had to fall below a cutoff that varied by geographic location: 57.21 in the 14 main metropolitan areas; 56.32 in other urban areas; and 40.75 in rural areas.

[8] We present the methodology using the outcome of whether a student enrolls in college (i.e., the extensive margin). However, the same approach applies when considering other outcomes, like whether a student enrolls in an HQA college.

aim to estimate the conditional average treatment effect

$$\tau(x) = E\left[Y(1) - Y(0)|X = x\right]$$

of students with characteristics $X = x$. When estimating $\tau(x)$, we face the challenge that for every student we only observe one of the potential outcomes $Y(1)$ and $Y(0)$, namely the realized enrollment decision $Y = Y(T)$ for their actual treatment status T. However, since treatment is as good as random close to the eligibility cutoffs, the realization of $Y(1)$, $Y(0)$, and X are independent of T in an infinitesimal neighborhood of the cutoff, so we can identify treatment effects from $\tau(x) = E\left[Y|T = 1, X = x\right] - E\left[Y|T = 0, X = x\right]$.

We use several empirical strategies to estimate the causal effects of financial aid and how treatment effects vary by test scores and poverty levels. First, we utilize an RD design exploiting the test score and poverty cutoffs determining financial aid eligibility. Specifically, we compare the effects of financial aid across two types of compliers identified by the multidimensional RD setting: (1) poverty-eligible students near the test score cutoff and (2) test-score–eligible students near the poverty cutoff. We report separate estimates by collapsing the discontinuity into a single dimension for each student and use the 'rdrobust' package by Cattaneo et al. (2014). Then, we leverage time variation in the eligibility cutoffs, comparing how these estimates vary as the targeting changes.[9]

Second, we analyze whether the treatment effects vary across the distribution of student test scores and poverty by estimating HTEs along the two boundaries—a boundary discontinuity design. Notably, the two-dimensional nature of the RD score creates a boundary of the treatment region $\mathcal{B}$ in the test score–poverty plane along which assignment to treatment changes discontinuously from zero to one (Figure 1). For each point $\mathbf{b} \in \mathcal{B}$, the treatment effect at that point is given by $\tau(\mathbf{b}) =$

[9] Appendix B provides evidence supporting the validity of the RD methodology. Each year, the government announced the policy's test-score cutoff only after students had taken the SABER 11 exam, preventing manipulation or sorting around the threshold. As expected, formal tests using the local polynomial density estimator of Cattaneo et al. (2018, 2020a) fail to reject the null of no manipulation (Figures B.1 and B.2). We also test for covariate imbalance around the test-score and poverty cutoffs each year (Tables B.1 and B.2). Although balance is rejected in a few bandwidth–year combinations, these rejections are not systematic, supporting the validity of the identifying assumption.

$E\left[Y_i(1) - Y_i(0) | \mathbf{X}_i = \mathbf{b}\right]$ and under regularity conditions,

$$\tau(\mathbf{b}) = \lim_{\substack{d(\mathbf{x},\mathbf{b})\to 0 \\ \mathbf{x}\in\mathcal{B}_t}} E\left[Y_i | \mathbf{X_i} = \mathbf{x}\right] - \lim_{\substack{d(\mathbf{x},\mathbf{b})\to 0 \\ \mathbf{x}\in\mathcal{B}_c}} E\left[Y_i | \mathbf{X_i} = \mathbf{x}\right]$$

where $\mathcal{B}_c$ and $\mathcal{B}_t$ denote the control and treatment areas, respectively, and $d(\cdot,\cdot)$ is a metric. Compared to the one-dimensional score case, this multi-score design generates a curve of treatment effects $\tau(\mathbf{b})$, one for each boundary point $\mathbf{b} \in \mathcal{B}$. We examine distance-based estimates using 'rdmulti' package by Cattaneo et al. (2020b), and show robustness with the recent location-based approach using the 'rd2d' package by Cattaneo et al. (2025).

Lastly, in Appendix C, we use the CRF algorithm, a machine learning method designed to estimate HTEs in observational studies (Athey et al., 2019). Its main goal is to estimate conditional average treatment effects (CATEs)—the effect of financial aid adjusting for a person's observed covariates at baseline. Compared to the RD and boundary discontinuity approaches, CRF enables estimating HTEs away from the eligibility cutoffs. Appendix C provides full implementation details; here, we summarize the causal-forest procedure to estimate CATEs as follows. First, the algorithm draws 3,000 bootstrap samples from the training data. For each sample, it builds a causal tree by recursively partitioning the covariate space to maximize heterogeneity in subgroup ATEs. Second, the causal forest aggregates the resulting 3,000 causal trees into weights, $\hat{w}_i(x)$, for each training observation (Y_i, T_i, X_i). These weights reflect how often a given covariate vector x falls into the same leaf as X_i across the different trees. Third, the CATE at covariate vector x is estimated as a weighted average treatment effect

$$\hat{\tau}(x) = \frac{\sum_{T_i=1} \hat{w}_i(x) Y_i}{\sum_{T_i=1} \hat{w}_i(x)} - \frac{\sum_{T_i=0} \hat{w}_i(x) Y_i}{\sum_{T_i=0} \hat{w}_i(x)}$$

of nearby observations (Y_i, T_i, X_i). To ensure valid inference, we adopt an "honest" forest: Y_i is not used when constructing $\hat{w}_i(x)$. This sample-splitting approach guarantees that the resulting estimates are consistent and asymptotically normal.

4 The Baseline Equity–Efficiency Trade-Off

This section benchmarks equity–efficiency trade-offs using a baseline cohort of students who took the exam before the policy was announced. We leverage cross-sectional variation in financial aid eligibility to study how targeting shapes equity, defined as the sociodemographic characteristics of compliers. To assess efficiency, we compare treatment effects on various outcomes (e.g., college enrollment, attainment, predicted post-college earnings) and the MVPF across test scores and poverty levels.

4.1 Equity

We characterize who is impacted by financial aid by estimating the average baseline characteristics of *compliers*—students who receive aid when eligible (Abadie, 2003).

In our multidimensional RD setting, we identify two types of compliers: (i) low-SES students near the test-score cutoff and (ii) high-scoring students near the poverty cutoff. Columns (1) and (2) of Table 2 report their respective characteristics. Compliers at the poverty cutoff—by definition, higher-scoring and less poor—are more socioeconomically advantaged than those at the test-score cutoff. They are less than half as likely to have low-educated mothers, nearly twice as likely to have mothers with university education, and less than one-third as likely to live in the poorest residential stratum. They are also more likely to attend full-day high school, twice as likely to attend a private high school, and more often reside in Bogotá.

To characterize compliers along the test-score–poverty plane, Figure 2 plots their characteristics using the boundary discontinuity design, which compares individuals on either side of the multidimensional eligibility frontier. The figure traces how these characteristics vary along the full boundary—across both the test-score and poverty dimensions. The x-axis represents test scores, and the y-axis represents the wealth distribution (where higher y values indicate greater poverty). Compliers just barely eligible—those near the 91st test-score percentile and 53rd wealth percentile—appear in the bottom left of each panel. At the 91st percentile, 16% live in the poorest residential stratum, compared with less than 11% above the 99th percentile (panel a). Over the same range, the share attending full-day high school rises from 28% to 35% (panel b), and the share from private high schools increases from 27% to 37% (panel c). In contrast, the share of first-generation college students (based on mother's education) declines by 22 p.p.—from 64% to below 42% (panel d).

As noted earlier, the SISBEN score is a PMT—a weighted index based on household characteristics such as residential stratum and education. Because residential stratum is one of the components, it is closely correlated with the SISBEN score: for instance, the share of compliers living in the lowest residential stratum rises fourfold, from 16% at the 50th percentile of the SISBEN distribution to 64% among the poorest 1% (i.e., a poverty percentile of 100 in panel a).[10] Similarly, higher poverty is associated with a lower share of students attending high school full time (panel b). However, other SES measures exhibit non-monotonic patterns: students in the poorest decile are more likely to attend private high schools (panel c) and less likely to be first-generation college students (panel d) than those in the second-poorest decile. These non-monotonicities likely reflect that SES indicators are imperfectly correlated and that PMT scores imperfectly capture underlying poverty (Banerjee et al., 2024; Hanna and Olken, 2018).

4.2 Efficiency

We now examine how treatment effects vary by test scores and poverty. We begin by reporting treatment effects for the two types of compliers, then trace heterogeneity along the eligibility boundaries. Finally, we apply the causal forest algorithm to estimate heterogeneity across the full distributions of test scores and poverty, as described in Appendix C. Across all three approaches, the results show that policy targeting influences not only the magnitude of treatment effects but also the margins along which students respond.

As discussed above, our multidimensional RD design allows comparison of treatment effects for compliers at the test-score and poverty cutoffs: the former have lower SES and test scores than the latter. Comparing Columns (1) of Tables 1 and A.4, students near the test-score cutoff are more marginal to college enrollment—their enrollment rises by 29.4 p.p. (70.5%, relative to a control mean of 41.7) versus 23.6 p.p. (43.8%) near the poverty cutoff. The impact on HQA enrollment is also larger for the test-score group (48.0 p.p. or 505.3%) than for the poverty group (42.6 p.p. or 174.6%). Among students near the poverty cutoff, financial aid more strongly affects the intensive margin, shifting enrollment from non-HQA and public HQA colleges toward private HQA institutions.

[10] The distance to the SISBEN cutoff does not perfectly map to students' poverty percentiles because rural students are subject to a separate cutoff (see Footnote 7). However, rural students represent only about 10% of test takers, and excluding them does not affect the results.

To illustrate treatment effect heterogeneity along the test-score–poverty plane, Figures 3 and 4 present reduced-form estimates across the test-score and poverty distributions, respectively, using the boundary discontinuity design. Percentage-point effects (black circles with 95% confidence intervals) are shown alongside proportional effects (gray triangles), which scale estimates by the control mean. Financial aid eligibility increases college enrollment, quality, and attainment across both distributions, with large and statistically significant effects throughout.

Focusing on the extensive margin—the likelihood of enrolling in *any* college—Figure 3 shows a sharp decline in effects with test scores: from 27.1 p.p. at the 91st percentile to 12.5 p.p. at the top (panel a). Because higher-scoring students are more likely to enroll even without aid, proportional effects also fall, from 60% at the 91st percentile to 18.9% at the top. Since the policy aimed to promote enrollment in HQA colleges, treatment effects on HQA enrollment (panel b) are larger than those on overall enrollment, reflecting both extensive- and intensive-margin impacts, especially from non-HQA institutions. These effects also decline with test scores, though more gradually: from 49.5 p.p. (515%) at the 91st percentile of test scores to 37.9 p.p. (142%) at the top. This pattern suggests that higher-scoring students tend to reallocate within the postsecondary sector—along the *intensive* margin—rather than move from non-enrollment to enrollment. The gradient flattens further when focusing on private HQA enrollment (panel c), where effects are 51.5 p.p. (1456%) at the 91st percentile of test scores and 44.8 p.p. (521%) at the top. Top-scoring students thus are more likely to substitute across institution types, especially from public to private HQA colleges.

Together, these patterns suggest that students at the top of the test-score distribution are more inframarginal with respect to the college enrollment decision: they would enroll somewhere even without financial aid. For these students, financial aid primarily shifts the *type* of institution they attend—inducing reallocation toward more selective and better-resourced colleges—rather than affecting *whether* they enroll at all.

Turning to longer-term outcomes, Figure 3 shows that the impact of financial aid on B.A. *attainment* is smaller than its impact on *enrollment* because some students drop out before completing their degrees. Nevertheless, treatment effects on this outcome are large and highly statistically significant across all college types—any college (panel d), HQA colleges (panel e), and private HQA colleges (panel f). Since test scores predict persistence in college, treatment effects decline monotonically with test scores, as higher-scoring students are more inframarginal to the completion decision.

Figure 3 shows reduced-form effects of financial aid eligibility, but only a subset

of eligible students take up the program. Because higher-scoring students are more likely to be admitted at HQA colleges, take-up rises sharply with test scores—from about 60% near the 91st percentile to nearly 80% at the top (Figure A.3). Consequently, the instrumental variable (IV) estimates, which scale reduced-form effects by take-up, exhibit an even steeper gradient. For instance, the IV estimate for B.A. attainment at private HQA colleges falls by nearly 20 p.p., from 64.6 p.p. at the 91st percentile to 45.6 p.p. at the top (Figure A.4). The proportional decline—from 1717% to 377%—suggests diminishing marginal returns to targeting increasingly high-ability students.

If returns to postsecondary education were homogeneous, targeting low-SES students near the 91st percentile of the test-score distribution—where enrollment and attainment effects are largest—would maximize expected future earnings and, *ceteris paribus*, the fiscal externality of the policy. In practice, however, returns vary widely across colleges, and financial aid moves students into higher–value-added institutions. Even among top scorers, aid substantially upgrades college quality, yet the value-added gradient with respect to test scores is flat (panel g): value-added gains do not increase with test scores. If value-added predicts compliers' realized earnings, earnings gains from financial aid are positive but not increasing across the test-score distribution.[11] Together with the earlier equity results, this implies that targeting *higher-scoring* low-SES students weakens progressivity without generating offsetting efficiency gains.

We now examine treatment effect heterogeneity by poverty using the boundary discontinuity design. Figure 4 reports percentage-point and proportional effects across the poverty distribution. Financial aid eligibility significantly increases college enrollment, HQA enrollment, and private HQA enrollment at all poverty levels (panels a–c). While absolute effects remain fairly constant, proportional effects rise sharply with poverty because poorer students have lower baseline enrollment rates—especially in private HQA colleges. For example, the proportional effect on private

[11] Table V in Londoño-Vélez et al. (forthcoming) compares realized earnings nine years after high school with value-added predictions and find them closely aligned. At the test-score cutoff, realized earnings increase by 0.231 (s.e. 0.041) versus 0.203 (s.e. 0.014) based on value-added; at the poverty cutoff, the corresponding estimates are 0.152 (s.e. 0.089) and 0.179 (s.e. 0.029). The smaller gain for students at the poverty cutoff—who have higher test scores—again suggests diminishing returns to targeting higher-scoring students. Unfortunately, the boundary discontinuity design yields relatively imprecise earnings estimates, limiting interpretability. Figure A.5 plots effects by test score and poverty using 2023 earnings. For the 2014 cohort (nine years after high school), earnings effects are roughly flat across test scores and rise modestly with poverty, though confidence intervals are wide. For the 2016 cohort (eight years post–high school), effects appear to increase with both test scores and poverty but remain imprecisely estimated.

HQA enrollment increases nearly fivefold, from 1456% at the 47th percentile of poverty to almost 7000% for the poorest students.

Because low-SES students face higher dropout risk, the policy's effect on college *completion* is smaller than its effect on initial *enrollment* (panel d). Moreover, poverty is a strong predictor of dropout, so the impact of aid on B.A. attainment declines with poverty: the IV estimate falls from 30.0 p.p. at the 47th percentile to 13.1 p.p. among the poorest 1% (panel d, Figure A.6). Still, the effects on HQA and private HQA attainment remain large and positive across the poverty distribution (panels e and f). For instance, the IV estimate for private HQA attainment is 64.6 p.p. at the 47th percentile and remains 57.4 p.p. among the poorest students. (Again, these effects exceed those on overall B.A. attainment at comparable poverty levels, since they capture both extensive- and intensive-margin effects.) Because few poor students attend HQA or private HQA colleges without aid, the proportional effects rise sharply with poverty. And since these institutions add more value, the impact on value-added remains relatively stable across the poverty distribution (panel g).

Figure 5 (Figure A.7) summarizes the reduced-form (IV) estimates along both the test-score and poverty eligibility boundaries. A social planner aiming to maximize impacts on enrollment, attainment, or expected future earnings would avoid targeting low-SES students above the 93rd percentile of test scores. Instead, the optimal focus would be on low-SES students near the 91st percentile and those with similar test scores but greater poverty.[12] In this context, policies that prioritize equity—by targeting *poorer* high-achieving students—also enhance efficiency.

These results are robust to alternative HTE estimators. Cattaneo et al. (2025) show that conventional univariate distance-to-boundary estimators can suffer from misspecification bias when the assignment frontier includes kinks. To address this, they propose a bivariate location-based estimator that accounts for the frontier's geometry. Figures A.8 and A.9 compare the distance-based and location-based specifications in our boundary discontinuity design, using test scores and the poverty score as the running variables, respectively. The point estimates and confidence intervals are similar across methods. Importantly, the pattern of heterogeneity remains: treatment effects on college enrollment and attainment decline with test

[12] Our design does not allow estimation below the 91st test-score percentile, but targeting students far below this range is unlikely to increase impacts on HQA enrollment or expected earnings. Admissions to HQA institutions are highly test-score selective: the average test score at HQA private colleges is 1.77 standard deviations above the median (Table A.2), and even higher for competitive STEM majors, which yield the highest labor-market returns.

scores but are stable across the poverty distribution, and effects on predicted post-college earnings are flat across both dimensions.

Moreover, Appendix C estimates HTEs using the CRF algorithm (Athey et al., 2019). The results show that treatment effects are consistently positive and precisely estimated across all outcome domains. Financial aid does not discourage any subgroup from enrolling in college, attending HQA institutions, attending higher-return colleges, completing degrees, or improving labor-market outcomes. Despite these uniformly positive effects, their magnitudes vary considerably across individuals. A variable-importance analysis identifies the *test score* as the strongest predictor of treatment effects on enrollment and attainment. For both predicted and realized earnings, test scores remain the top predictor, but their influence diminishes—especially for realized earnings—while SES becomes much more salient. Consistent with the RD and boundary discontinuity results, the CRF estimates reveal that treatment effects on educational outcomes decline sharply with test scores yet remain relatively flat across the poverty distribution. In contrast, effects on predicted and realized earnings are largely invariant to both test scores and poverty. Overall, these findings confirm that targeting low-SES high achievers is efficient, though efficiency falls with test scores.

4.3 Social Welfare

We conduct a prospective cost-benefit analysis using the concept of the MVPF, defined as the ratio of the willingness-to-pay (WTP) for the policy to its net fiscal costs.

WTP. In the MVPF framework, marginal recipients—those whose behavior changes in response to the policy—derive welfare gains from the resulting increase in their net-of-tax lifetime income.[13] Inframarginal individuals, who would have attended an HQA college even without the policy, value it by the transfer they receive. Accordingly, total WTP equals the present discounted value of the net-of-tax lifetime earnings gains for marginal beneficiaries, plus the discounted value of transfers for inframarginal individuals.

We make several simplifying assumptions to project lifetime earnings impacts. As

[13] This requires assuming that marginal individuals were not perfectly optimizing when the policy was introduced—under perfect optimization, their WTP would be zero by the envelope condition. However, if individuals face liquidity or credit constraints, it is reasonable to approximate their WTP using the policy's estimated effect on lifetime earnings rather than assuming it is zero (Hendren and Sprung-Keyser, 2020).

noted in Footnote 11, because the boundary discontinuity design yields imprecise earnings estimates, we use value-added effects as proxies for long-run earnings gains. These estimates, derived from pre-policy cohorts, reflect the average return for a typical student in each college–major pair. We assume that students induced to enroll in or switch colleges or majors due to the policy experience similar returns. Earnings effects are assumed to materialize nine years after high school (around age 26). Lifetime earnings trajectories are based on population-level profiles from the 2019 *Gran Encuesta Integrada de Hogares*, Colombia's main labor force survey. We project percentage earnings gains forward under constant proportional effects (Hendren and Sprung-Keyser, 2020), assuming retirement at age 60, a 19% income tax rate, and a 3% discount rate.

For inframarginal individuals—those who would have attended an HQA college even without the policy—we assume their WTP equals the value the transfers they receive, including tuition subsidies and living stipends recorded in ICETEX administrative data over eight years. We calculate the present discounted value of these transfers by summing annual disbursements and discounting them to the first year at a 3% annual rate. The share of inframarginal individuals is inferred from the proportion of barely ineligible students who enroll in an HQA college—about 7% near the test-score cutoff and up to 34% among those with higher test scores.

Figure 6 shows that WTP is substantial across the test-score and poverty distributions, ranging from approximately USD 18,000 to USD 26,500 (panel a). Mirroring the heterogeneity in college value-added, WTP is relatively flat across test scores and highest among poorer students.

Net Fiscal Costs. The net fiscal cost to the government equals the present value of additional educational expenditures minus projected increases in future income tax revenue. Incremental government outlays are estimated as the additional public subsidy per student, obtained from boundary discontinuity estimates. This includes direct transfers to financial aid recipients (from ICETEX data) and average per-student subsidies for all students—recipients and non-recipients—enrolled in public institutions, based on financial data reported to Colombia's Ministry of Education. We assume private college students without financial aid generate no additional fiscal

cost, as their tuition covers all expenses.[14] Net costs are then adjusted by subtracting projected lifetime income tax revenue, applying a 19% tax rate to the estimated earnings gains.

Figure 6 shows that the net fiscal costs vary considerably (panel b): they are lowest for the poorest students and highest for the highest-achieving students, who disproportionately use aid to enroll in more expensive colleges and majors—particularly at high-cost private HQA colleges.

The MVPF. Figure 6 shows that the MVPF exceeds one across the entire distribution (panel c), indicating that targeting low-SES high achievers generates welfare gains. Estimates range from roughly 1.7 to 4.9, suggesting that the government's return per dollar spent can nearly triple depending on how aid is targeted—even within the group of low-SES high achievers. Although fiscal costs increase with test scores, the associated WTP (benefits) do not, resulting in a sharp decline in the MVPF at higher achievement levels. For instance, the MVPF falls from 2.2 at the 91st percentile to 1.8 at the very top of the test-score distribution. In contrast, the MVPF is relatively flat across the poverty distribution—it does not decline with poverty. If anything, it increases, reaching 4.9 for students with the highest poverty levels, indicating that financial aid yields the highest social return for these students.

In the MVPF framework, social welfare increases when public funds are reallocated toward groups with higher MVPF values, as these groups yield greater welfare gains per dollar of government spending. For example, reallocating financial aid from students near the 91st test-score percentile with median poverty levels ($MVPF_A = 2.2$) or from the extremely poor ($MVPF_C = 4.9$) to median-poverty students with the very highest test-scores ($MVPF_B = 1.8$) would improve welfare only if society values one dollar to the latter group more than \$1.2 (= 2.2 / 1.8) to the first group and more than \$2.7 (= 4.9 / 1.8) to the second. Under standard welfare weights that do not strongly favor test-score achievement, such reallocation would reduce the program's progressivity without meaningful efficiency gains, thereby lowering overall social welfare.

[14] To clarify, our cost measure—the additional government outlay estimated from the boundary discontinuity design—includes HQA tuition subsidies, living stipends, changes in public operating subsidies, and the higher costs of shifting students from associate to bachelor's degree programs. It excludes other expenses such as books, supplies, housing, and transportation.

5 The Equity–Efficiency Trade-Off When Students Respond

We now reassess the equity–efficiency trade-off when students respond strategically to policy incentives.

Awareness of the policy rose quickly after its announcement, with two key consequences. First, demand for it grew markedly, increasing take-up: at the test score cutoff, the share of eligible students receiving aid rose from about 60% in 2014 to over 88% three years later (Figure B.3 and Table A.5). Second, the policy incentivized younger low-SES cohorts to improve exam performance (Bernal and Penney, 2019; Laajaj et al., 2022). The number of low-SES students scoring at least 318 out of 500 rose by more than 60%, from roughly 10,000 in 2014 to nearly 16,200 in 2015. By 2016, the number of low-SES students scoring 342 or above had *tripled* relative to 2014 (Londoño-Vélez et al., 2025). Combined with the expanded SISBEN survey rollout, these trends substantially increased the pool of low-SES students attaining top-decile test scores.

With demand outstripping the policy's fixed budget, the policy progressively tightened test-score eligibility while keeping the poverty threshold constant. The cutoff rose from 310 in 2014 (top 9.3%) to 318 in 2015 (top 8.6%), 342 in 2016 (top 4.6%), and 348 in 2017 (top 3.5%). In 2018, the new administration increased it further to 359 (top 1.9%). When actual demand later fell short of expectations—partly due to uncertainty about the program's continuation (Bernal et al., 2024)—the government extended benefits to the next three highest-scoring low-SES students in each department. In 2019, the cutoff was slightly relaxed to 350 (top 2.6%).[15] In what follows, we exploit this panel variation in eligibility thresholds to reassess how changes in financial-aid targeting shape equity and efficiency when students can respond to policy incentives, which themselves evolve over time.[16]

[15] The policy's shift toward progressively higher test-score thresholds could be interpreted as reflecting implicit welfare weights that place greater value on test scores than on financial need (Björkegren et al., 2025)—a point we revisit in Section 6.

[16] As noted in Footnote 9, eligibility thresholds were announced each year only *after* students had taken the SABER 11 exam, precluding any strategic manipulation or sorting around the test-score cutoff (Figure B.1). Furthermore, because the PMT formula is not publicly disclosed, there is no evidence of bunching below the poverty cutoffs (Figure B.2).

5.1 Equity

Both students' behavioral responses to policy incentives and the rising test-score cutoff shaped who was targeted and who took up aid. Figure 7 shows that as the test-score cutoff increased, compliers became progressively more socioeconomically advantaged. When the threshold was at the 91st percentile, 69% of compliers had mothers without postsecondary education, compared to 42% at the 98th percentile; the share with university-educated mothers rose from 15% to 36%. Similarly, the proportion from the poorest socioeconomic stratum fell from 33% to 23%, while the share from private high schools doubled from 17% to 36%. Table 2 shows similar trends near the poverty cutoff: as the policy became more test-score selective, compliers increasingly came from more educated families, wealthier households, and private-school backgrounds. These shifts mirror the baseline patterns in Section 4 but grew stronger over time, suggesting that younger low-SES students aware of the policy and striving to qualify tended to come from relatively more advantaged socioeconomic backgrounds.

As a result, the policy became even less progressive over time, shifting benefits toward students from more advantaged backgrounds. Table A.6, which presents descriptive statistics on financial aid recipients, highlights the broader implications of these shifts. Between 2014 and 2018, the share of recipients whose mothers had no postsecondary education dropped by 41%, from 64.6% to 38.3%. Meanwhile, the share of beneficiaries from private high schools nearly doubled, and the proportion from the wealthiest three residential strata also rose meaningfully.

5.2 Efficiency

To assess efficiency, Figure 8 plots the likelihood that low-SES students enroll in any college (panel a), HQA colleges (panel b), and private HQA colleges (panel c) immediately after high school, by test score, for six cohorts graduating between 2014 and 2019. As described in Section 2, enrollment rises with test scores, and financial aid eligibility substantially increases enrollment—especially in private HQA colleges.

As predicted at baseline, the enrollment effect of financial aid declines as the test-score cutoff rises, both in magnitude and significance. Figure 9 shows RD estimates with confidence intervals: the reduced-form effect falls from 29.4 p.p. at the 91st percentile (p-value = 0.00) to 3.0 p.p. at the 98th percentile (p-value = 0.26). In relative terms (blue triangles), the effect drops from a significant 70% to a non-significant

5%. This attenuation reflects higher-scoring students being inframarginal to college enrollment. Later cohorts—more aware of and responsive to policy incentives—are also more likely inframarginal. As take-up increased over time, the IV coefficient declined even more sharply, from 49.5 p.p. (119%; p-value = 0.00) to 5.1 p.p. (8%; p-value = 0.33). Thus, along the extensive margin, greater emphasis on test scores appears both less equitable and less efficient.[17]

However, the policy had large effects on the *intensive* margin, altering the type of college attended. Figure 10 shows sizable increases in HQA and private HQA enrollment across all test-score cutoffs. Although RD and boundary discontinuity estimates are not directly comparable—since they identify effects for different local populations—later cohorts of compliers, increasingly aware of the policy incentives, also appear more inframarginal on the HQA margin: the HQA effect declined from 48.0 p.p. (504%) at the 91st percentile to 15.4 p.p. (42%) at the 98th percentile. At the 91st percentile, extensive-margin responses dominate—roughly two-thirds of the HQA enrollment gain reflects overall college attendance, with the remainder due to shifts from non-HQA colleges. By contrast, at the 98th percentile, the policy operates mainly on the intensive margin, reallocating students from non-HQA and public HQA institutions to HQA colleges—specifically, private ones.[18]

Indeed, higher-scoring students primarily use financial aid to attend private institutions. Figure A.10 shows that as the test-score cutoff rises, enrollment shifted almost one-for-one from public to private colleges. This pattern persisted even after 2017 (96.5th percentile), when the policy added financial incentives favoring public institutions—first a lump-sum payment equal to half a monthly minimum wage, and from 2018 (98.1th percentile), larger grants of US$927—US$3,379. Despite these incentives, students continued to prefer private higher education, revealing strong demand for private HQA colleges.

The shift toward private HQA institutions likely influences long-run labor-market outcomes, as these colleges exhibit stronger job-placement records. Figure 11 shows that access to high–value-added colleges rises with test scores and that financial aid

[17] Table 1 reports the policy's effects on additional outcomes, including enrollment in bachelor's degree programs and STEM fields. Higher-scoring students are more likely to enroll in bachelor's degrees and select STEM majors. As a result, increasing test-score selectivity also reduces the policy's impacts on both outcomes, since these students are largely inframarginal on these margins as well.

[18] The policy generated incentives for institutions to obtain HQA status: while 33 colleges had HQA by October 2014, this number had nearly doubled six years later (Table A.2). Nevertheless, the finding that financial aid expanded HQA enrollment is robust to considering colleges with HQA status defined annually (Table A.7).

significantly increases value added across the entire test-score distribution. Figure 12 further indicates that these gains remain substantial even at the top, where students move from public to private HQA colleges offering better career prospects. Overall, value-added effects are relatively stable across test scores. If value-added continues to predict future earnings for aid recipients, tighter test-score targeting may reduce enrollment impacts without diminishing long-run gains.

Similar patterns appear at the poverty cutoff. Extensive-margin impacts fell sharply over time—by roughly 20 p.p. in reduced-form estimates (Tables A.4) and 40 p.p. in IV estimates (Table A.8)—becoming small and statistically insignificant once the policy targeted the 97th test-score percentile or higher. Effects on HQA enrollment followed the same pattern. Among these higher-achieving, higher-SES compliers, financial aid still yields positive impacts on value-added when targeted to the very highest achievers, though the smaller sample affects precision.

Finally, our labor-market data allow us to examine realized earnings eight years after high school—for the 2014 (measured in 2022) and 2015 (measured in 2023) cohorts—when financial aid's impacts seem to stabilize (Londoño-Vélez et al., forthcoming). Figure A.11 shows IV estimates of financial aid's effect on formal monthly earnings, using the test-score (panel a) and poverty-index (panel b) running variables. Across cohorts and cutoffs, the impacts are positive, sizable, and statistically significant. Consistent with the value-added results, earnings gains remain roughly stable as the test-score cutoff increases—*they do not rise*. Although we do not compare MVPF estimates across cohorts, the lack of rising returns to test-score selectivity suggests that the MVPF is also unlikely to increase.[19]

6 Discussion

We studied how financial aid targeting based on standardized test scores and household SES affects program incidence, treatment effects, and the MVPF. Using both cross-sectional and panel variation in eligibility criteria, we found that directing aid to low-

[19] We do not compare MVPF estimates across cohorts for several reasons. First, we lack post-college earnings for the most recent cohorts to estimate WTP, and imputing them from pre-reform value-added would require assuming that the policy did not affect value-added itself. Second, the policy's direct cost evolved over time as subsidy designs changed. Additional modifications—such as the 2018 mandate requiring private colleges to grant a 25% tuition discount and the reduction of relocation stipends from four to two monthly minimum wages (three if no HQA colleges were available in the applicant's department)—further complicate isolating the fiscal impact of targeting changes. Finally, cost data for recent cohorts remain incomplete because degree completion is not yet observed.

SES high achievers enhances social welfare. Yet equity and efficiency vary markedly within this group. Because test scores are highly correlated with SES, progressivity declines as test score rise. Treatment effects on college enrollment also fall with test scores, shifting impacts toward the intensive margin. Despite substantial return heterogeneity across colleges, predicted earnings gains remain large and positive throughout the test-score distribution. However, rising fiscal costs at higher scores lower the MVPF. These patterns intensify when students respond strategically to policy incentives, as responders are more advantaged and inframarginal on both enrollment and quality margins. In contrast, treatment effects and MVPFs are highest among *poorer* high achievers—indicating that targeting this group strengthens both equity and efficiency, aligning the two goals.

Interestingly, the policy's increasing test-score selectivity over time deviates from the standard predictions of a welfare-maximizing planner. Although the government never articulated a rationale for these shifts, one might be tempted—following the logic of Björkegren et al. (2025)—to infer implicit social preferences, suggesting that the planner assigned higher welfare weights to higher-scoring students. However, such interpretations risk conflating normative objectives with political constraints. In practice, policymakers' reluctance to tighten eligibility by lowering the poverty threshold, and their choice to emphasize test scores instead, likely reflected political feasibility: as demand from middle-income households grew, excluding them became politically costly. Facing a fixed budget and mounting political pressure, policymakers may have found adjusting the test-score cutoff the more politically viable lever. Unlike in the standard utilitarian framework of optimal tax theory, real-world policy design is often shaped by political economy forces (Persson and Tabellini, 2002; Saez and Piketty, 2013), which may prevent governments from targeting policies according to their intended impacts, or which populations are most valued.

We conclude by outlining several avenues for future research. First, our analysis focused on students in the bottom half of the wealth distribution and the top decile of the test-score distribution. We restrict attention to this group because students with lower test scores are unlikely to gain admission to HQA colleges, and because the Colombian policy—as occurs with many higher education policies worldwide—specifically targets low-SES high achievers. Still, if the planner does not weight college quality heavily, it could be valuable to examine the equity and efficiency consequences of lowering the test-score cutoff below the top decile, where the policy may be likelier to influence students' decisions over whether to enroll and what degree programs to

attend. Second, our value-added estimates are derived from pre-reform cohorts and capture the average return for the typical student in each college–major pair. However, the students who are induced to enroll in or switch colleges and majors because of the policy may differ in their characteristics and, consequently, may experience different returns. Third, the policy itself may have affected the value added by colleges. Incorporating these dimensions could yield further insights.

References

Abadie, Alberto, "Semiparametric Instrumental Variable Estimation of Treatment Response Models," *Journal of Econometrics*, 2003, *113* (2), 231–263.

Aiken, Emily, Simone Bellue, Dean Karlan, Christopher Udry, and Joshua E. Blumenstock, "Machine learning and phone data can improve targeting of humanitarian aid," *Nature*, March 2022, *603* (7903), 864–870. Epub 2022 Mar 16.

Athey, Susan, Julie Tibshirani, and Stefan Wager, "Generalized Random Forests," *Annals of Statistics*, April 2019, *47* (2), 1148–1178.

_ **, Niall Keleher, and Jann Spiess**, "Machine learning who to nudge: Causal vs predictive targeting in a field experiment on student financial aid renewal," *Journal of Econometrics*, 2025, p. 105945.

Banerjee, Abhijit, Rema Hanna, Benjamin A. Olken, and Diana Sverdlin Lisker, "Social Protection in the Developing World," *Journal of Economic Literature*, December 2024, *62* (4), 1349–1421.

Barahona, Nano, Cauê Dobbin, and Sebastián Otero, "Equilibrium Price Responses to Targeted Student Financial Aid," Working Paper 33833, National Bureau of Economic Research May 2025.

Bernal, Gloria L. and Jeffrey Penney, "Scholarships and student effort: Evidence from Colombia's Ser Pilo Paga program," *Economics of Education Review*, 2019, *72*, 121–130.

_ **, Luz K. Abadía, Luis E. Álvarez Arango, and Kristof De Witte**, "Financial Aid Uncertainty and Low-Income Students' Higher Education Preferences," *Higher Education*, 2024, *87*, 1845–1863.

Bertrand, Marianne, Bruno Crépon, Alicia Marguerie, and Patrick Premand, "Do Workfare Programs Live Up to Their Promises? Experimental Evidence from Côte d'Ivoire," 2024. mimeo.

Bettinger, Eric, Oded Gurantz, Laura Kawano, Bruce Sacerdote, and Michael Stevens, "The Long Run Impacts of Merit Aid: Evidence from California's Cal Grant," *American Economic Journal: Economic Policy*, 2019, *11* (1), 64–94.

Bhattacharya, Debopam and Pascaline Dupas, "Inferring welfare maximizing treatment assignment under budget constraints," *Journal of Econometrics*, 2012, *167* (1), 168–196.

Björkegren, Daniel, Joshua E. Blumenstock, and Samsun Knight, "What do Policies Value?," *Review of Economic Studies*, 10 2025, p. rdaf089.

Brown, Meta, John Karl Scholz, and Ananth Seshadri, "A New Test of Borrowing Constraints for Education," *The Review of Economic Studies*, 11 2011, *79* (2), 511–538.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma, "Manipulation Testing Based on Density Discontinuity," *The Stata Journal*, 2018, *18* (1), 234–261.

_ , _ , and _ , "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

_ , Rocio Titiunik, and Ruiqi Yu, "Estimation and Inference in Boundary Discontinuity Designs," May 2025.

_ , Rocío Titiunik, and Gonzalo Vazquez-Bare, "Analysis of regression-discontinuity designs with multiple cutoffs or multiple scores," *The Stata Journal*, 2020, *20* (4), 866–891.

_ , Sebastian Calonico, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Colas, Mark, Sebastian Findeisen, and Dominik Sachs, "Optimal Need-Based Financial Aid," *Journal of Political Economy*, 2021, *129* (2), 492–533.

Dynarski, Susan and Judith Scott-Clayton, "Financial Aid Policy: Lessons from Research," *The Future of Children*, 2013, *23* (1), 67–91.

_ , **Lindsay C Page, and Judith Scott-Clayton**, "College Costs, Financial Aid, and Student Decisions," Working Paper 30275, National Bureau of Economic Research July 2022.

Hanna, Rema and Benjamin A. Olken, "Universal Basic Incomes versus Targeted Transfers: Anti-Poverty Programs in Developing Countries," *Journal of Economic Perspectives*, November 2018, *32* (4), 201–26.

Haushofer, Johannes, Paul Neihaus, Carlos Paramo, Edward Miguel, and Michael Walker, "Targeting Impact versus Deprivation," June 2025.

Hendren, Nathaniel and Ben Sprung-Keyser, "A Unified Welfare Analysis of Government Policies," *The Quarterly Journal of Economics*, 03 2020, *135* (3), 1209–1318.

Laajaj, Rachid, Andrés Moya, and Fabio Sánchez, "Equality of opportunity and human capital accumulation: Motivational effect of a nationwide scholarship in Colombia," *Journal of Development Economics*, 2022, *154*, 102754.

Lochner, Lance, "Chapter 2 - Nonproduction Benefits of Education: Crime, Health, and Good Citizenship," in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of The Economics of Education*, Vol. 4 of *Handbook of the Economics of Education*, Elsevier, 2011, pp. 183–282.

Londoño-Vélez, Juliana, Catherine Rodríguez, and Fabio Sánchez, "Upstream and Downstream Impacts of College Merit-Based Financial Aid for Low-Income Students: Ser Pilo Paga in Colombia," *American Economic Journal: Economic Policy*, May 2020, *12* (2), 193–227.

_ , _ , _ , **and Luis E. Alvarez**, "Financial Aid and Upward Mobility: Evidence from Colombia's Ser Pilo Paga," *Journal of Political Economy*, forthcoming.

Londoño-Vélez, Juliana, Catherine Rodriguez, Fabio Sánchez, and Luis Esteban Álvarez-Arango, "Targeting Social Assistance: The Evolution of College Financial Aid in Colombia," *AEA Papers and Proceedings*, May 2025, *115*, 340–44.

Lovenheim, Michael and Jonathan Smith, "Chapter 4 - Returns to different postsecondary investments: Institution type, academic programs, and credentials," in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Eric A.*

Hanushek, Stephen Machin, and Ludger Woessmann, eds., Vol. 6 of *Handbook of the Economics of Education*, Elsevier, 2023, pp. 187–318.

Mirrlees, James A., "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 1971, *38* (2), 175–208.

Nguyen, Tuan D., Jenna W. Kramer, and Brent J. Evans, "The Effects of Grant Aid on Student Persistence and Degree Attainment: A Systematic Review and Meta-Analysis of the Causal Evidence," *Review of Educational Research*, 2019, *89* (6), 831–874.

OECD, *Education in Colombia*, OECD Publishing, 2016.

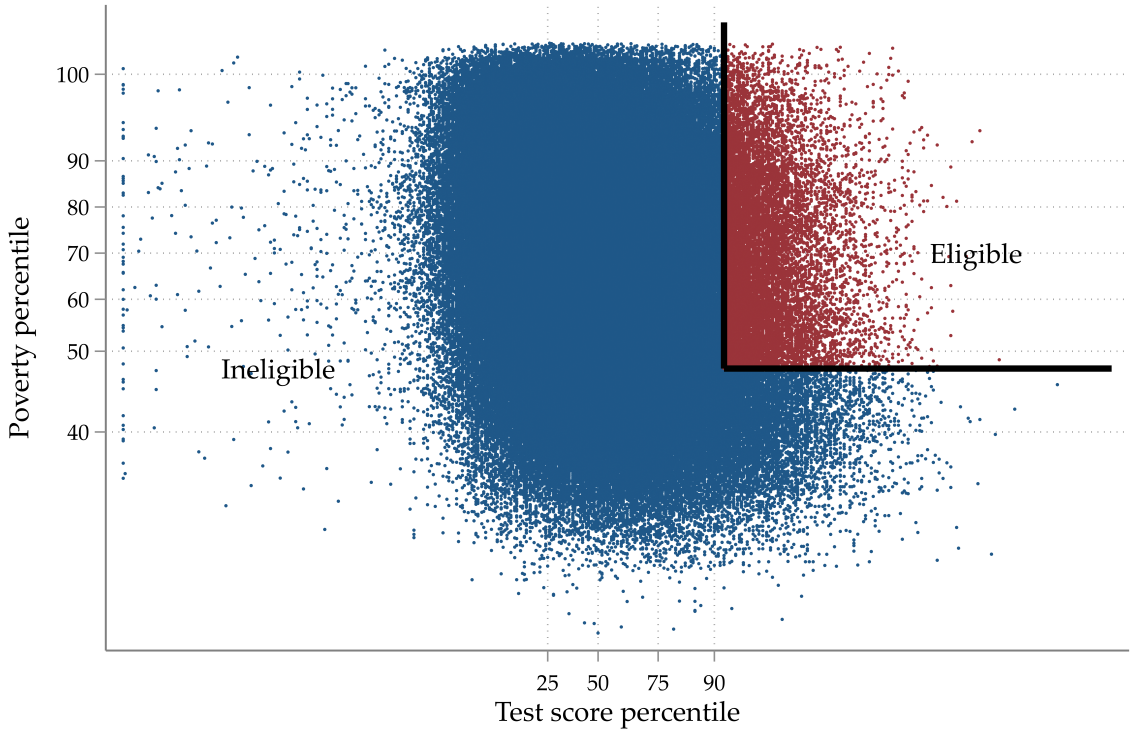
Persson, Torsten and Guido Tabellini, "Political Economics and Public Finance," in Alan J. Auerbach and Martin Feldstein, eds., *Handbook of Public Economics*, Vol. 3, Amsterdam: North-Holland, 2002, chapter 24, pp. 991–1042.

Riehl, Evan, Juan E. Saavedra, and Miguel Urquiola, "Learning and Earning: An Approximation to College Value Added in Two Dimensions," in "Productivity in Higher Education" NBER Chapters, National Bureau of Economic Research, Inc, December 2018, pp. 105–132.

Saez, Emmanuel, "Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 2002, *117* (3), 1039–1073.

_ **and Thomas Piketty**, "Optimal Labor Income Taxation," in "Handbook of Public Economics," Vol. 5, Elsevier, 2013, pp. 391–474.

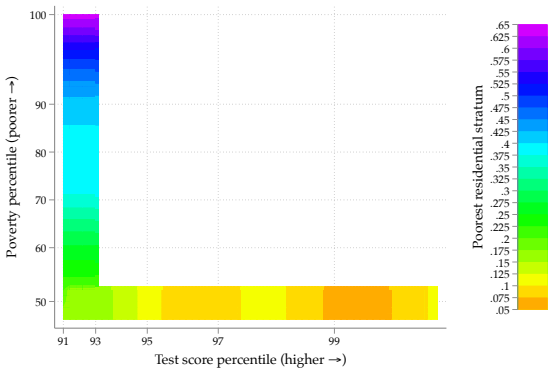
Figure 1: Targeting Financial Aid Based on Test Scores and SES



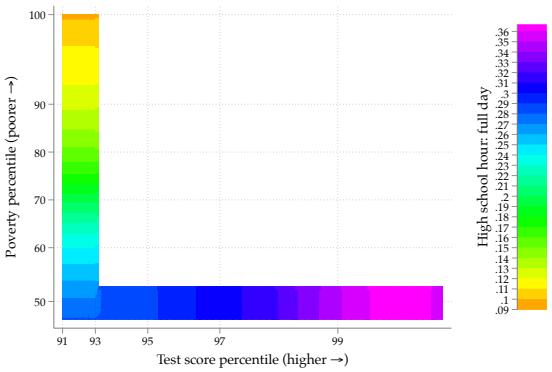
Notes: This figure illustrates how financial aid is targeted in Colombia. Each point represents a high school test taker, positioned according to their test score (x-axis) and poverty score (y-axis, where high y values indicate greater poverty). The red region marks students who qualified for financial aid in 2014—those in the poorest half of the SISBEN wealth distribution with test scores of at least 310 out of 500, placing them in the top decile nationwide. We refer to this group as low-SES high achievers. All other students, shown in blue, were ineligible for financial aid.

Figure 2: Complier Characteristics Along the Boundaries of Financial Aid Eligibility

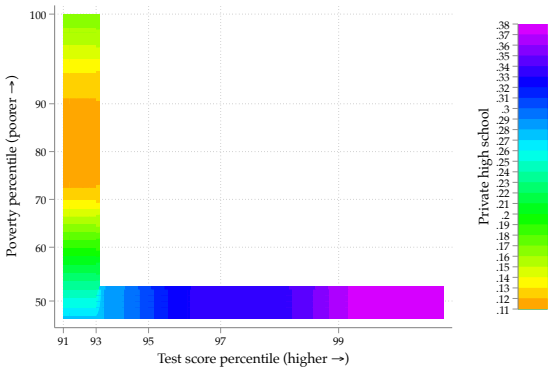
(a) Poorest residential stratum



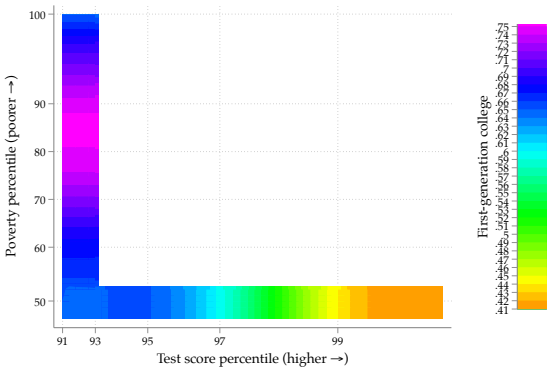
(b) High school hours: full day



(c) High school: private



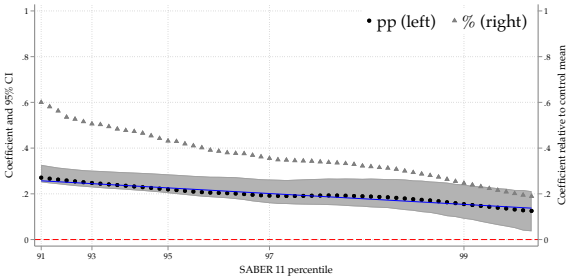
(d) First-generation college student



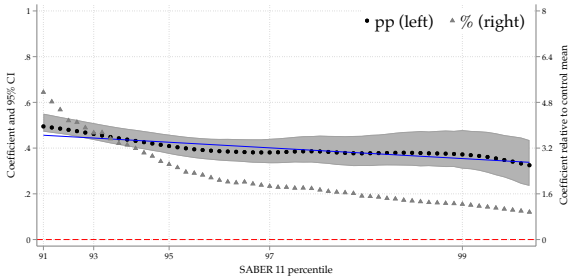
Notes: The figure shows how complier characteristics vary across test scores and poverty, estimated using the rdmulti package (Cattaneo et al., 2020b). Following Abadie (2003), who demonstrates that in an IV (like a fuzzy RD) the mean of a characteristic among compliers can be expressed as a Wald ratio, we compute these characteristics by regressing $Beneficiary_i \cdot X_i$ on $Beneficiary_i$, using the eligibility indicator as an instrument for $Beneficiary_i$. Intuitively, this identifies the average value of characteristic X_i for individuals whose treatment status changes at the strict eligibility cutoff. The sample is restricted to students who took the exam in the fall of 2014. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), and ICETEX.

Figure 3: Heterogeneity by Test Score: Reduced-Form Estimates

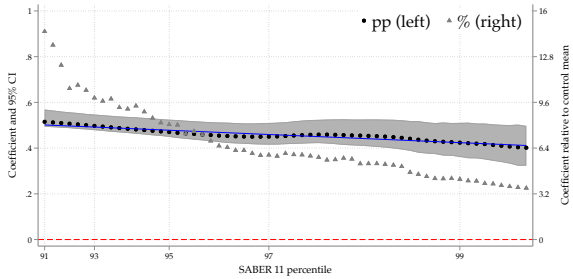
(a) Any college enrollment



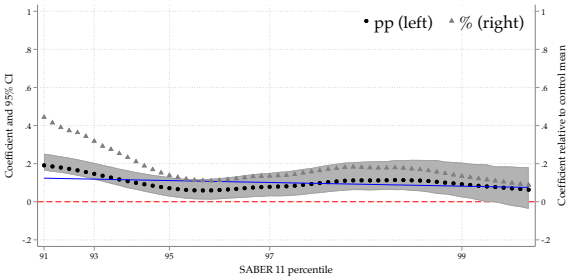
(b) HQA enrollment



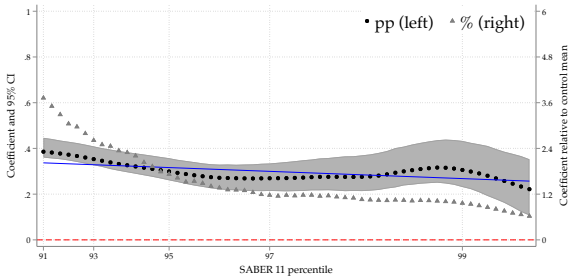
(c) Private HQA enrollment



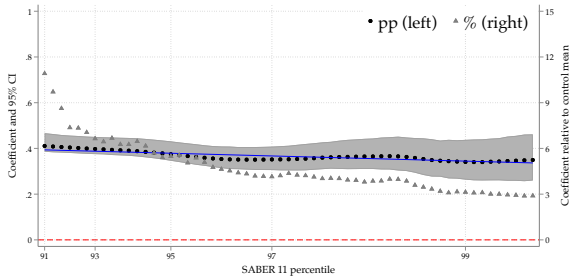
(d) Any B.A.



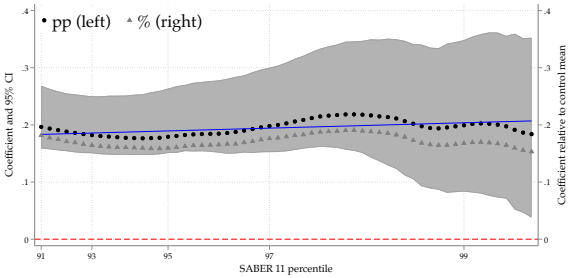
(e) HQA B.A.



(f) Private HQA B.A.



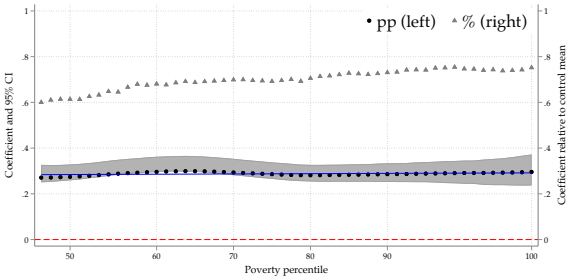
(g) Value-added



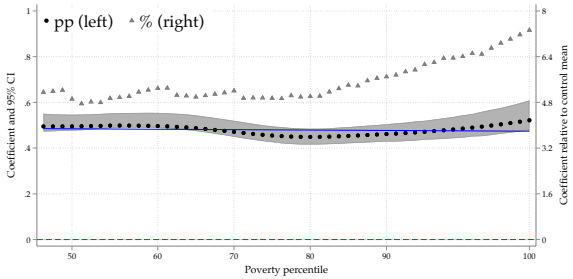
Notes: The figure compares the reduced-form RD estimates across test scores for students at the poverty threshold. The gray area represents the robust bias-corrected 95% confidence intervals. The blue line is a linear fit prediction for the RD estimate on the test score. Conventional local linear RD estimates and robust bias-corrected standard errors are estimated with package rdmulti (Cattaneo et al., 2020b). Value-added is based on the average monthly earnings nine years after high school for student who ever attended a given college–major combination, relative to students who never attended any college, after controling for baseline covariates, as detailed in the main text. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), SABER PRO (Icfes), SABER T&T (Icfes), and PILA (MinSalud).

Figure 4: Heterogeneity by Level of Poverty: Reduced-Form Estimates

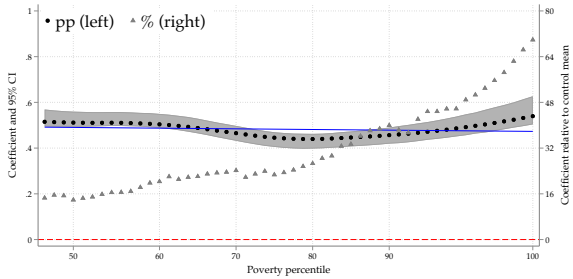
(a) Any enrollment



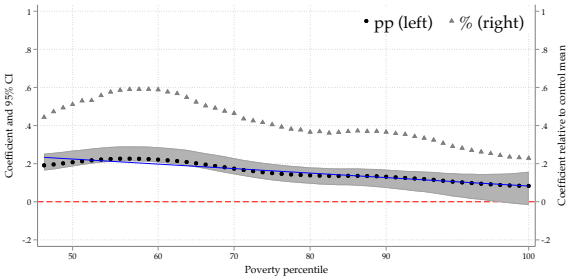
(b) HQA enrollment



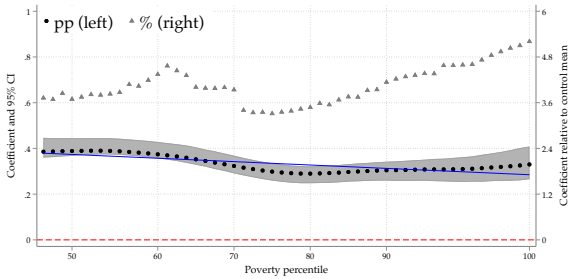
(c) Private HQA enrollment



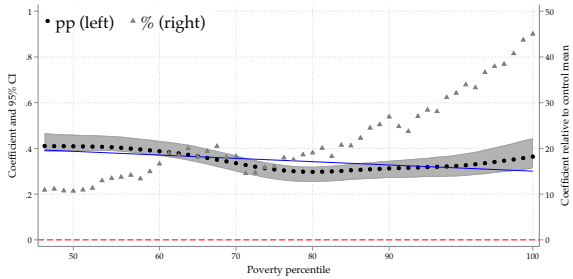
(d) Any B.A.



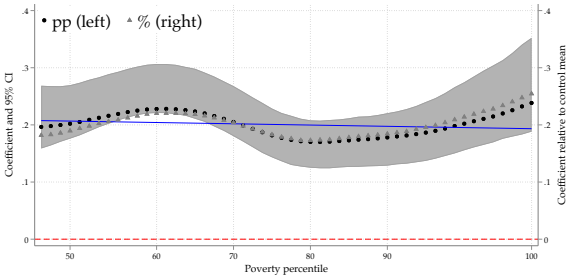
(e) HQA B.A.



(f) Private HQA B.A.



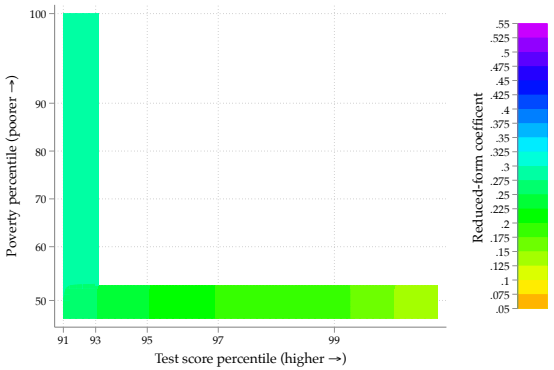
(g) Earnings value-added



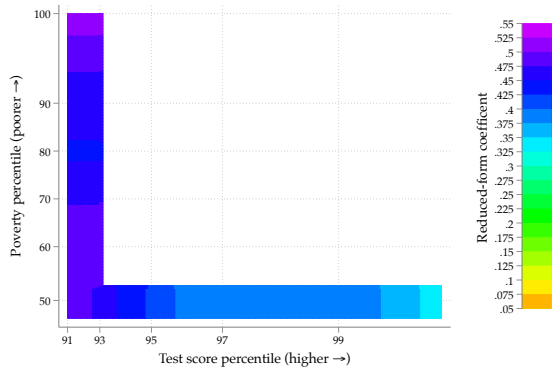
Notes: The figure compares the reduced-form RD estimates across poverty levels (the negation of the distance to the SISBEN cutoff) for students at the test-score threshold. The gray area represents the robust bias-corrected 95% confidence intervals. The blue line is a linear fit prediction for the RD estimate on the poverty score. Conventional local linear RD estimates and robust bias-corrected standard errors are estimated with package rdmulti (Cattaneo et al., 2020b). Earnings value-added are based on average monthly earnings 9 years after high school, calculated using pre-reform data and detailed in the main text. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), SABER PRO (Icfes), SABER T&T (Icfes), and PILA (MinSalud).

Figure 5: Heterogeneous Treatment Effects: Reduced-Form Estimates

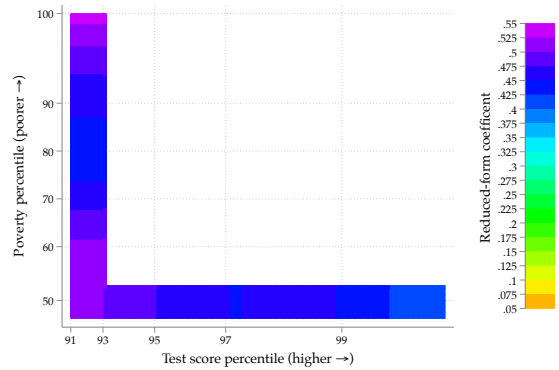
(a) Any college enrollment



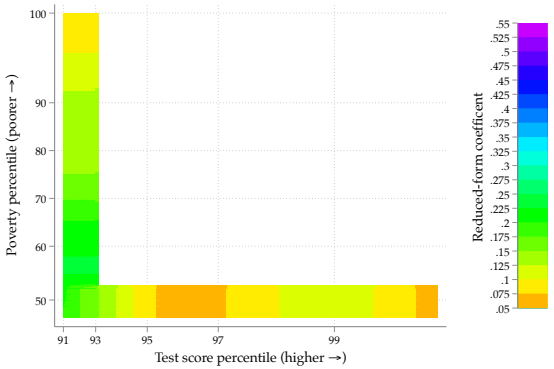
(b) HQA enrollment



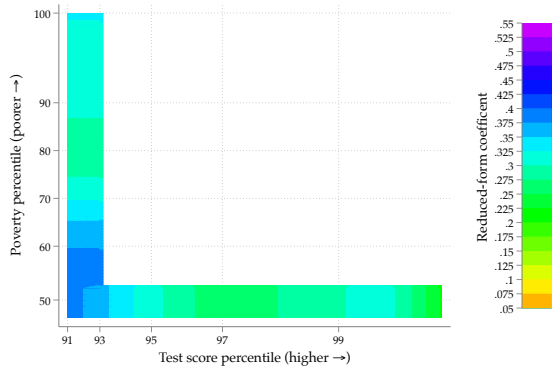
(c) Private HQA enrollment



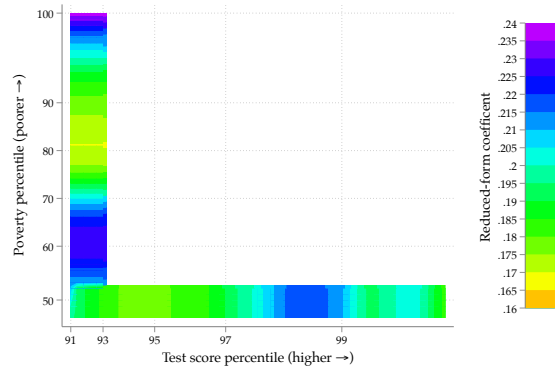
(d) Any B.A.



(e) HQA B.A.

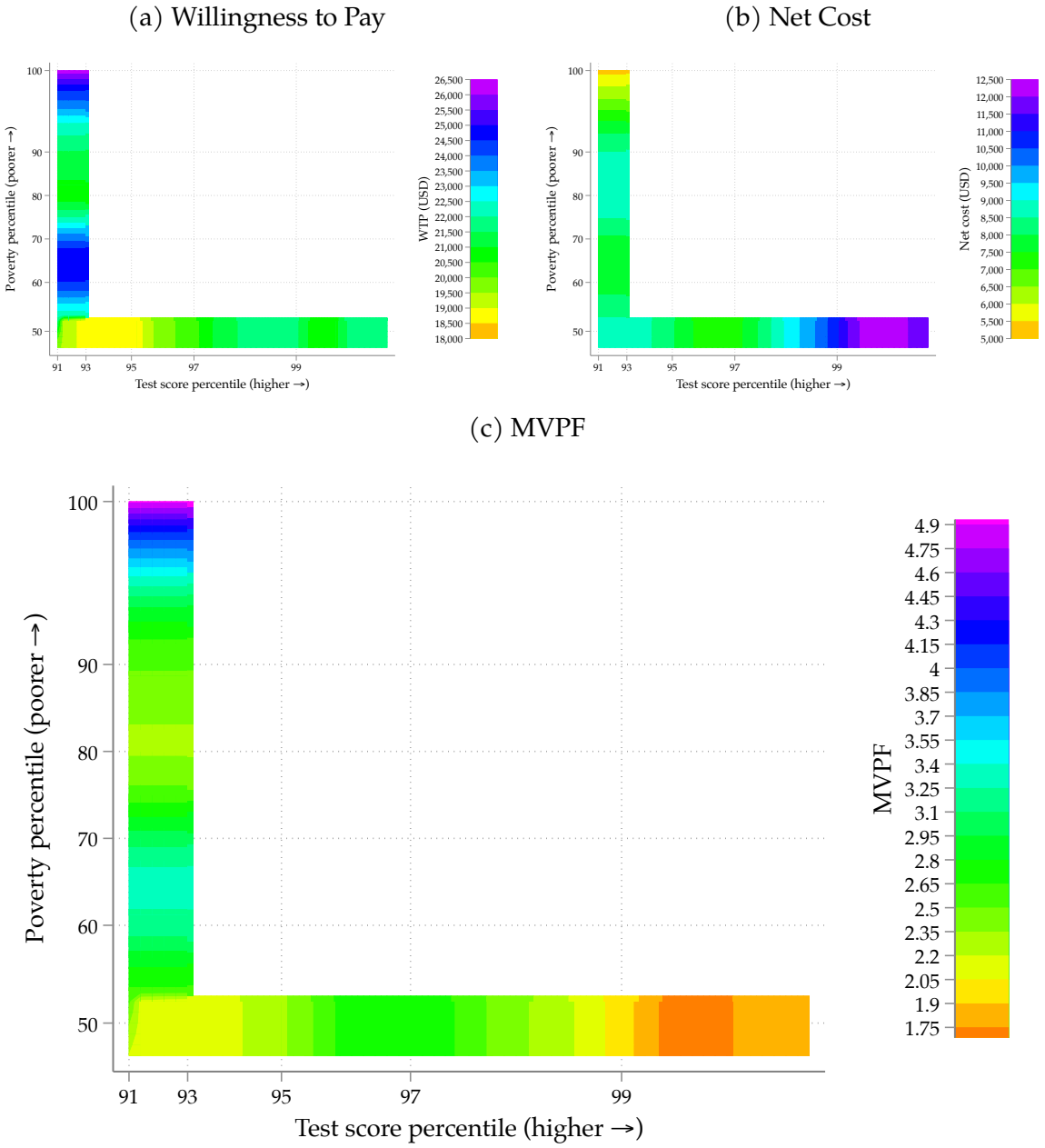


(f) Value-added



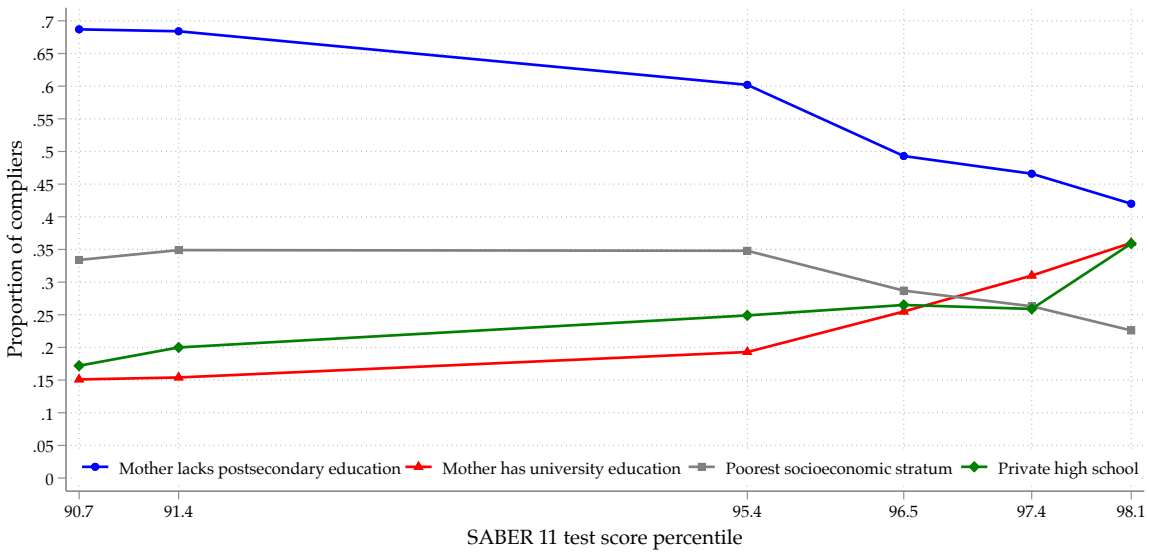
Notes: The figure compares the reduced-form RD estimates across the distribution of test scores and poverty, estimated with package rdmulti (Cattaneo et al., 2020b). Value-added is based on average monthly earnings 9 years after high school for students who ever attended a given college–major pair, calculated using pre-reform data and detailed in the main text. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), SABER PRO (Icfes), and PILA (MinSalud).

Figure 6: The MVPF Across Test Scores and Poverty



Notes: The figure compares the WTP (panel a), net government cost (panel b), and the MVPF (panel c) across the distribution of test scores and poverty, estimated with package rdmulti (Cattaneo et al., 2020b). WTP equals the projected earnings gains based on college value-added for marginal recipients and the financial aid transfer for inframarginal recipients. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), ICETEX, and PILA (MinSalud).

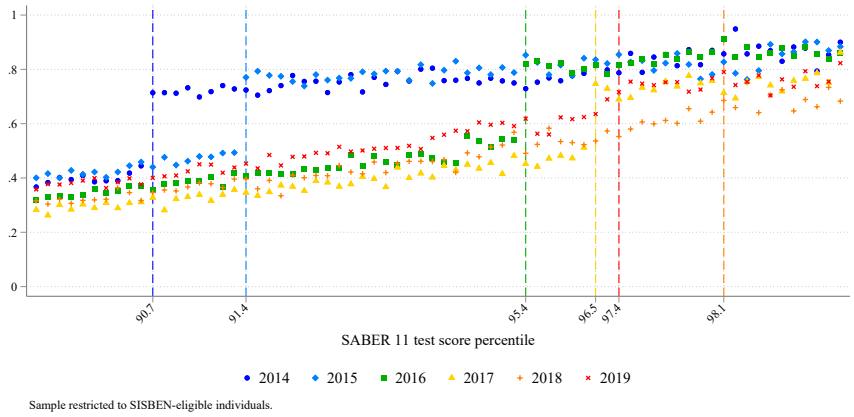
Figure 7: Complier Characteristics Across Test-Score Cutoffs



Notes: This figure plots the characteristics of compliers as a function of the equivalent SABER 11 percentile cutoff used to determine financial aid eligibility in years 2014 to 2019. Following the approach of Abadie (2003), we estimate the mean characteristic X_i among compliers by regressing $Beneficiary_i \cdot X_i$ on $Beneficiary_i$, using the eligibility indicator as an instrument for $Beneficiary_i$. The sample is restricted to SISBEN-eligible individuals. *Source:* Authors' calculations based on Table 2.

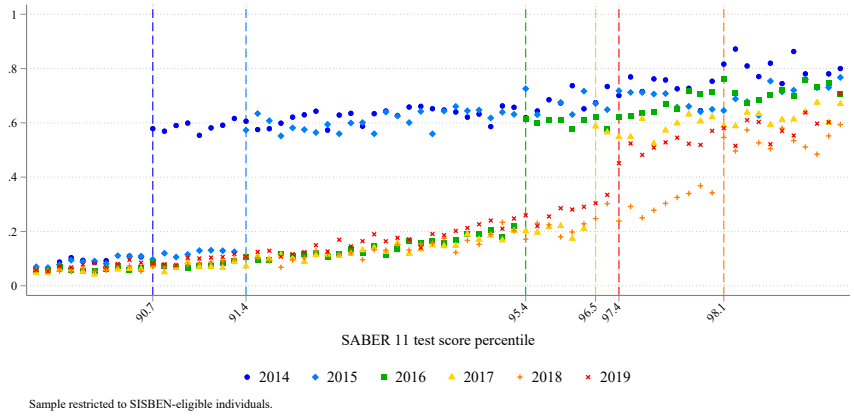
Figure 8: Enrollment as a Function of Test Scores for Low-SES Students

(a) Any College



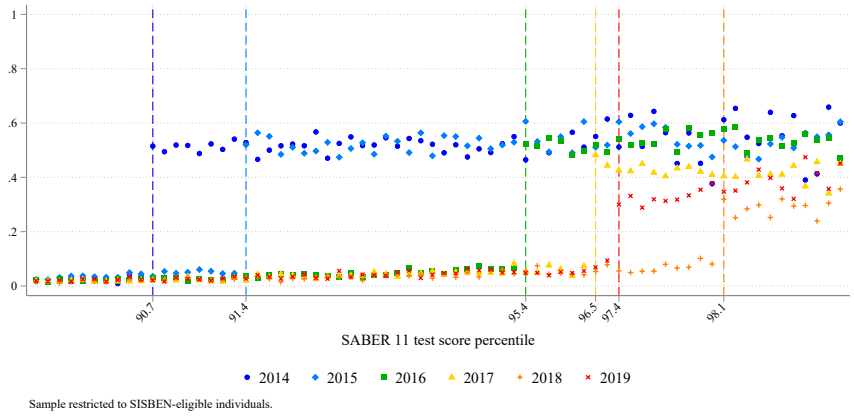
Sample restricted to SISBEN-eligible individuals.

(b) HQA College



Sample restricted to SISBEN-eligible individuals.

(c) Private HQA College

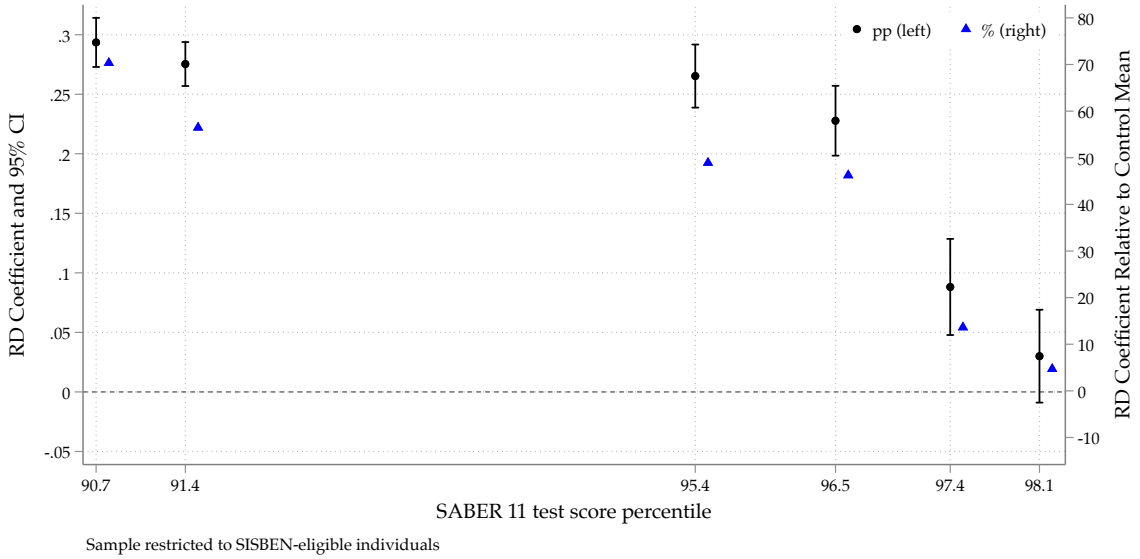


Sample restricted to SISBEN-eligible individuals.

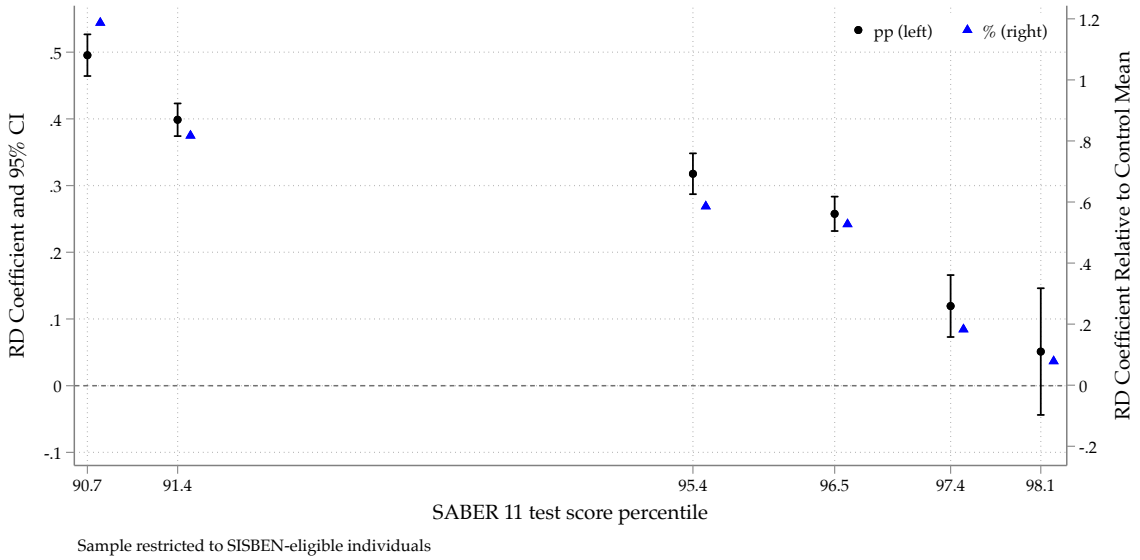
Notes: This figure plots the likelihood of enrolling in any college (panel a), an HQA college (panel b), or a private HQA college (panel c) immediately after completing high school as a function of the test score in years 2014 to 2019. The vertical lines illustrate the test-score cutoff determining financial aid eligibility each year. The sample is restricted to low-SES individuals. HQA status is defined at baseline (October 2014). *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), and SNIES (MEN).

Figure 9: The Enrollment Effect is Decreasing in Test Scores

(a) Reduced Form



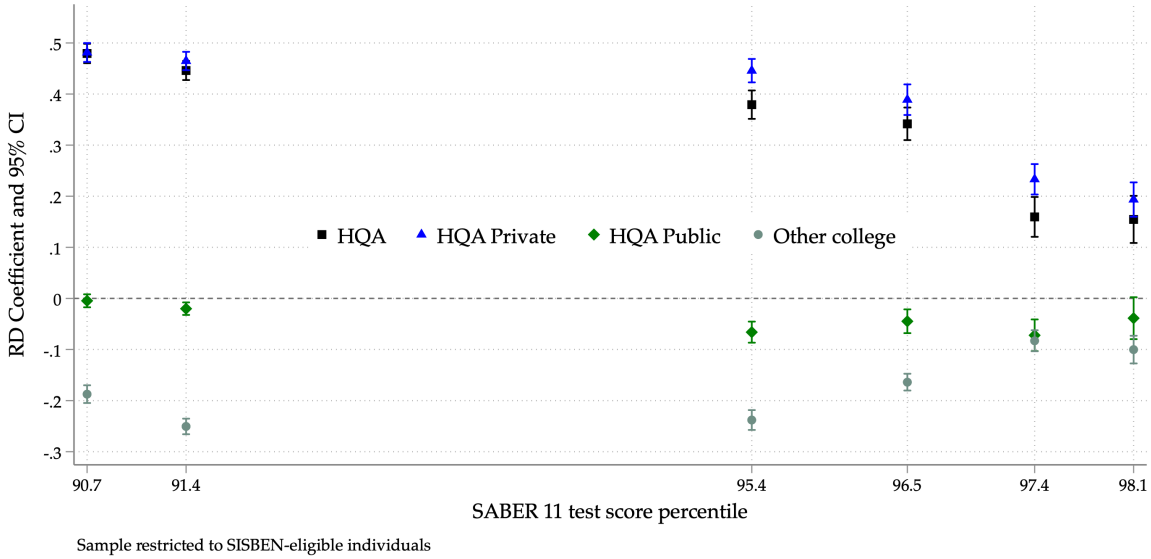
(b) IV



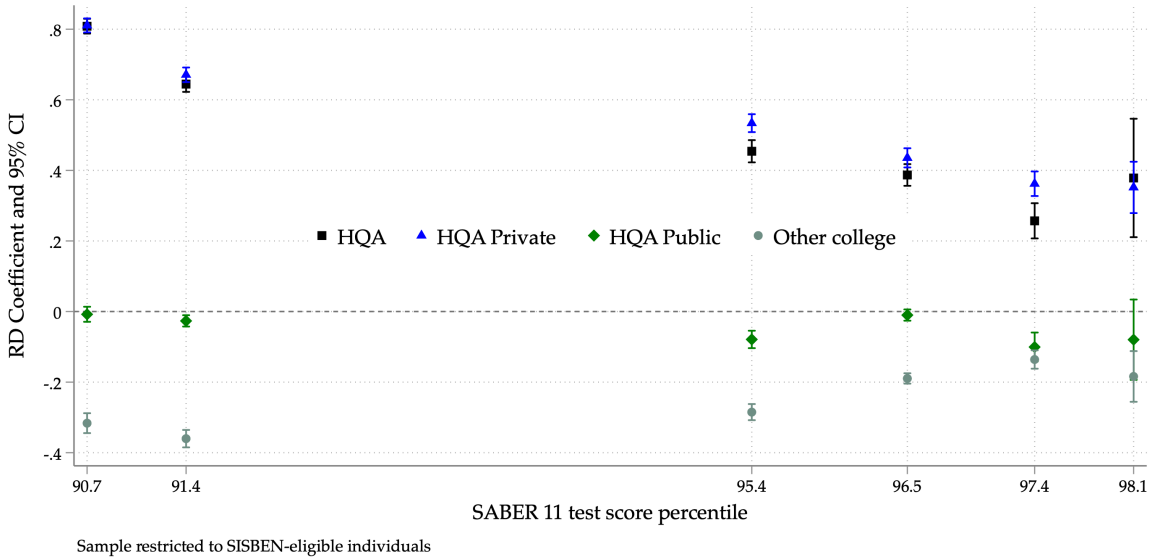
Notes: The figures show how financial aid impacted the likelihood of college enrollment immediately after high school by test-score cutoff for each year from 2014 to 2019. Panel (a) presents the reduced-form effect, while Panel (b) scales this effect by the likelihood of receiving financial aid. The sample is restricted to SISBEN-eligible individuals. Conventional local linear RD estimates and 95% confidence intervals, computed with rdrobust package (Cattaneo et al., 2014), are shown in black (left y-axis). Blue markers divide the RD coefficient by the comparison group's mean to express the effect in percentage terms (right y-axis). *Sources:* Tables 1 and 3.

Figure 10: The Intensive-Margin Effects Dominate as Test Scores Increase

(a) Reduced Form

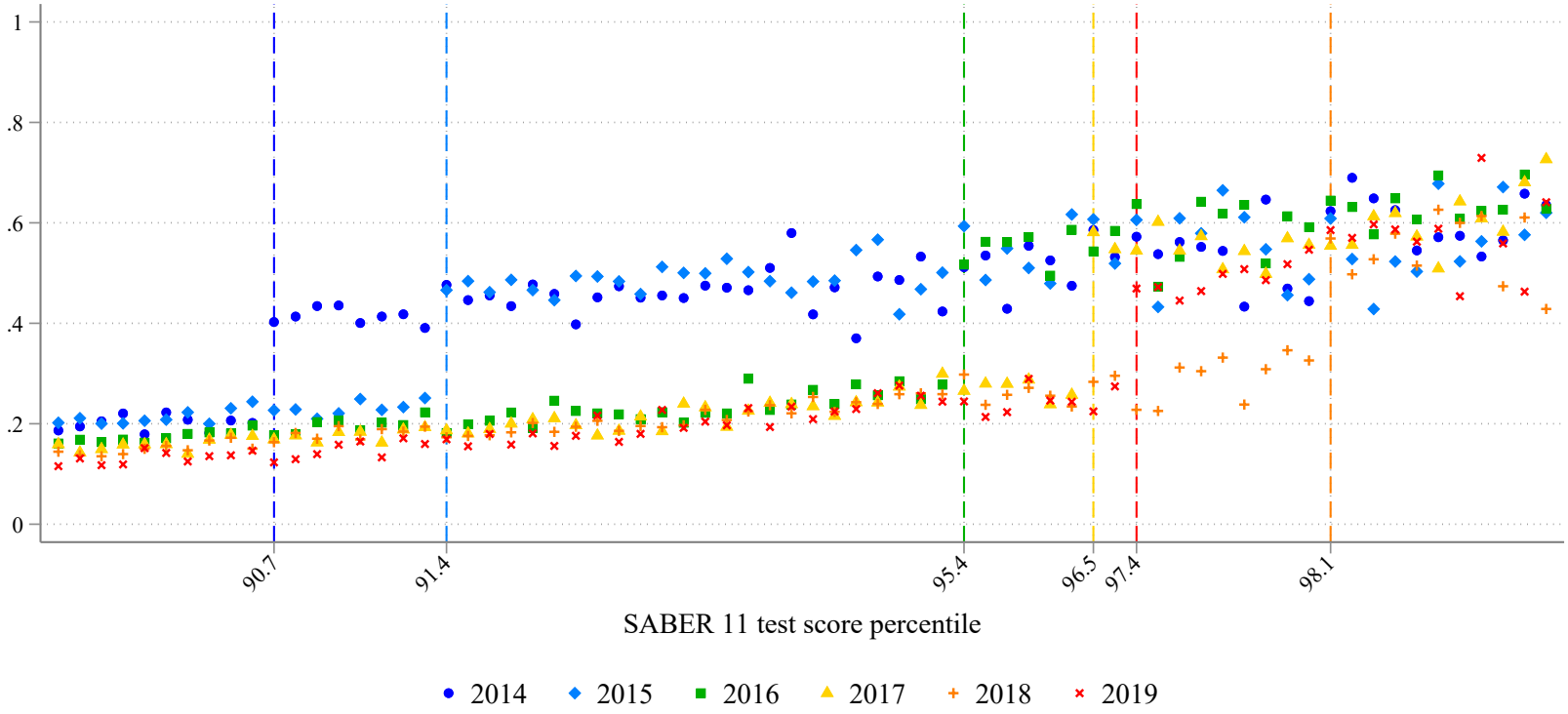


(b) IV



Notes: The figures show how financial aid impacted the likelihood of HQA enrollment, private HQA enrollment, public HQA enrollment, or any other enrollment immediately after high school by test-score cutoff for each year from 2014 to 2019. Panel (a) plots the reduced-form effect, while Panel (b) scales it by the likelihood of receiving financial aid. The sample is restricted to SISBEN-eligible individuals. Conventional local linear RD estimates and 95% confidence intervals estimated with package rdrobust (Cattaneo et al., 2014). HQA status is defined at baseline (October 2014). *Sources:* Tables 1 and 3.

Figure 11: Value-Added Increases Across the Test-Score Distribution

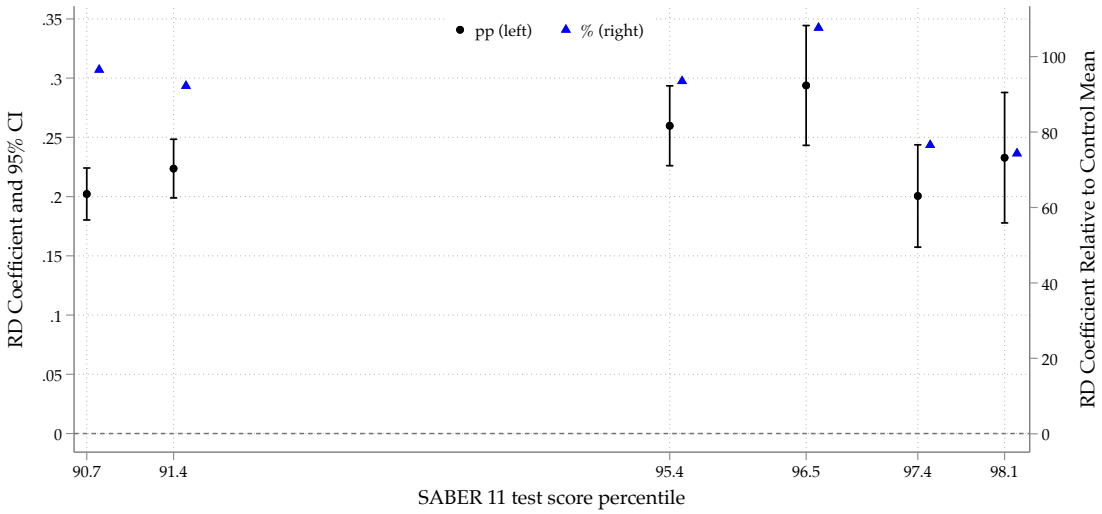


Sample restricted to SISBEN-eligible individuals.

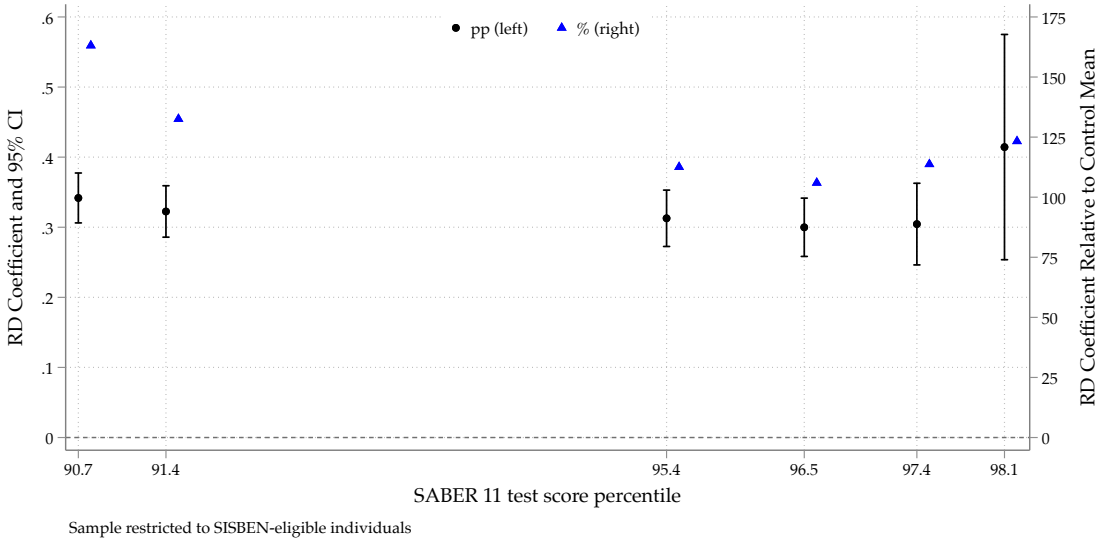
Notes: This figure plots the earnings value-added of each college–major pair for low-SES students who enrolled in college immediately after completing high school, as a function of the test score between 2014 and 2019. Value-added is based on formal monthly earnings nine years after high school— expressed in multiples of the monthly minimum wage for full-time workers—relative to students who did not enroll in any postsecondary institution within six years of high school completion. Vertical lines indicate the SABER 11 eligibility cutoff for financial aid in each year. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), and PILA (MinSalud).

Figure 12: The Impact on Value-Added is Roughly Constant in Test Scores

(a) Reduced Form



(b) IV



Notes: The figures plot the impact of financial aid on earnings value-added of the college–major pair as a function of the test-score cutoff for each year from 2014 to 2019. Formal monthly earnings are measured nine years after high school and the college–major fixed effects are expressed relative to students who did not access any postsecondary institution within six years from high school. Panel (a) plots the reduced-form effect, while Panel (b) scales it by the likelihood of receiving financial aid. The sample is restricted to SISBEN-eligible individuals. Conventional local linear RD estimates and 95% confidence intervals estimated with package rdrobust (Cattaneo et al., 2014) displayed in black. The blue markers divide the RD coefficient by the control mean to express this effect in percentage terms. *Sources:* Tables 1 and 3.

Table 1: Reduced-Form Estimates by Test-Score Selectivity

	Test-Score Percentile / Year					
	90.7 2014 (1)	91.4 2015 (2)	95.4 2016 (3)	96.5 2017 (4)	97.4 2019 (5)	98.1 2018 (6)
Any college	0.294	0.275	0.265	0.228	0.088	0.030
	(0.013)	(0.011)	(0.016)	(0.018)	(0.025)	(0.023)
	[0.417]	[0.488]	[0.542]	[0.493]	[0.647]	[0.640]
HQA college	0.480	0.446	0.379	0.343	0.161	0.154
	(0.011)	(0.011)	(0.017)	(0.019)	(0.023)	(0.028)
	[0.095]	[0.131]	[0.215]	[0.220]	[0.320]	[0.363]
Private HQA college	0.479	0.464	0.447	0.389	0.234	0.191
	(0.011)	(0.011)	(0.013)	(0.018)	(0.018)	(0.021)
	[0.028]	[0.053]	[0.068]	[0.066]	[0.064]	[0.082]
Public HQA college	-0.034	-0.043	-0.114	-0.085	-0.088	-0.079
	(0.009)	(0.009)	(0.013)	(0.015)	(0.022)	(0.026)
	[0.125]	[0.132]	[0.217]	[0.227]	[0.339]	[0.370]
Other college	-0.184	-0.170	-0.113	-0.114	-0.070	-0.128
	(0.011)	(0.011)	(0.015)	(0.016)	(0.024)	(0.023)
	[0.321]	[0.359]	[0.330]	[0.273]	[0.327]	[0.284]
Bachelor's degree program	0.355	0.341	0.306	0.258	0.107	0.036
	(0.013)	(0.013)	(0.016)	(0.019)	(0.024)	(0.022)
	[0.307]	[0.405]	[0.490]	[0.460]	[0.617]	[0.621]
Associate degree program	-0.065	-0.068	-0.042	-0.031	-0.025	-0.005
	(0.007)	(0.006)	(0.006)	(0.005)	(0.007)	(0.007)
	[0.114]	[0.088]	[0.055]	[0.034]	[0.036]	[0.018]
STEM	0.138	0.146	0.109	0.126	0.062	0.017
	(0.012)	(0.012)	(0.017)	(0.019)	(0.025)	(0.025)
	[0.192]	[0.227]	[0.284]	[0.264]	[0.380]	[0.376]
STEM Plus	0.218	0.207	0.202	0.157	0.064	0.047
	(0.014)	(0.012)	(0.019)	(0.019)	(0.025)	(0.029)
	[0.366]	[0.414]	[0.456]	[0.417]	[0.545]	[0.534]
SSH	0.057	0.052	0.050	0.065	0.022	-0.030
	(0.007)	(0.009)	(0.012)	(0.011)	(0.012)	(0.019)
	[0.042]	[0.065]	[0.076]	[0.063]	[0.089]	[0.109]
Arts	0.018	0.016	0.013	0.009	-0.002	0.007
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
	[0.010]	[0.012]	[0.013]	[0.011]	[0.017]	[0.008]
Value-added	0.202	0.224	0.260	0.294	0.201	0.233
	(0.013)	(0.015)	(0.020)	(0.029)	(0.026)	(0.033)
	[0.210]	[0.242]	[0.278]	[0.273]	[0.262]	[0.313]

Notes: This table reports reduced-form RD estimates of the effect of financial aid eligibility on college enrollment, type, degree program, and major immediately after high school. The last row reports effects on the earnings value-added of the first college–major pair attended. Columns are ordered by test-score cutoff; the running variable is the distance to the test-score cutoff, and the sample includes low-SES students. Because each outcome uses its own optimal RD bandwidth, total enrollment effects may not equal the sum across categories. Local linear RD estimates, robust standard errors (in parentheses), and control means (in brackets) are computed using rdrobust (Cattaneo et al., 2014). *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), and PILA (MinSalud).

Table 2: Complier Characteristics by Test-Score Selectivity

Test score percentile	90.7		91.4		95.4		96.5		97.4		98.1	
Year	2014		2015		2016		2017		2019		2018	
Instrument	Test score	Poverty	Test score	Poverty	Test score	Poverty	Test score	Poverty	Test score	Poverty	Test score	Poverty
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Female	0.460	0.464	0.480	0.453	0.469	0.507	0.464	0.414	0.415	0.433	0.372	0.319
Age	16.384	16.256	16.368	16.300	16.293	16.294	16.252	16.340	16.223	16.209	16.266	16.211
Ethnic minority	0.032	0.038	0.030	0.004	0.027	0.016	0.029	0.003	0.039	0.001	0.015	0.021
Mother's education: primary	0.203	0.096	0.213	0.127	0.166	0.103	0.125	0.086	0.093	0.052	0.064	-0.010
Mother's education: secondary	0.488	0.440	0.479	0.430	0.445	0.309	0.384	0.320	0.387	0.529	0.372	0.306
Mother's education: T&T	0.157	0.191	0.155	0.192	0.195	0.237	0.236	0.261	0.220	0.086	0.204	0.182
Mother's education: professional	0.151	0.278	0.154	0.253	0.193	0.346	0.255	0.354	0.310	0.353	0.360	0.504
Father's education: primary	0.315	0.179	0.291	0.174	0.248	0.141	0.170	0.119	0.168	0.166	0.122	0.043
Father's education: secondary	0.436	0.390	0.437	0.400	0.428	0.397	0.406	0.383	0.381	0.324	0.329	0.304
Father's education: T&T	0.118	0.181	0.121	0.172	0.141	0.152	0.153	0.199	0.141	0.174	0.180	0.223
Father's education: professional	0.131	0.252	0.151	0.260	0.182	0.318	0.273	0.315	0.307	0.304	0.351	0.435
Household SES: Stratum 1	0.334	0.097	0.349	0.094	0.348	0.128	0.287	0.084	0.263	0.144	0.226	0.071
Household SES: Stratum 2	0.441	0.514	0.451	0.545	0.454	0.455	0.468	0.453	0.454	0.370	0.397	0.408
Household SES: Stratum 3	0.212	0.356	0.186	0.308	0.186	0.380	0.193	0.395	0.229	0.377	0.274	0.423
Household SES: Strata 4-6	0.021	0.033	0.013	0.043	0.015	0.056	0.037	0.070	0.036	0.108	0.050	0.089
School hours: Full day	0.201	0.325	0.207	0.374	0.223	0.433	0.226	0.408	0.213	0.506	0.297	0.474
Private school	0.172	0.379	0.200	0.344	0.249	0.430	0.265	0.428	0.259	0.486	0.359	0.463
Family has internet	0.630	0.787	0.618	0.831	0.628	0.808	0.692	0.883	0.773	0.813	0.807	0.921
Family has a laptop	0.752	0.923	0.732	0.872	0.699	0.857	0.735	0.922	0.710	0.803	0.788	0.950
Family has a car	0.182	0.312	0.171	0.296	0.167	0.310	0.236	0.341	0.267	0.459	0.293	0.426
School location: Urban	0.919	0.960	0.907	0.950	0.941	0.961	0.945	0.976	0.955	1.017	0.934	0.981
Vulnerable regions	0.053	0.040	0.059	0.024	0.056	0.034	0.067	0.023	0.062	0.057	0.059	0.023
Not Bogota	0.816	0.632	0.827	0.639	0.870	0.641	0.891	0.662	0.938	0.796	0.895	0.614

Notes: This table reports complier characteristics at the test-score and poverty eligibility cutoffs over time. Following Abadie (2003), mean characteristics X_i among compliers are estimated by regressing $Beneficiary_i \cdot X_i$ on $Beneficiary_i$, instrumented by the eligibility indicator. Local linear RD estimates use the rdrobust package (Cattaneo et al., 2014). Columns alternate by cutoff and test-score selectivity: odd columns use the test-score threshold (low-SES sample) and even columns use the poverty threshold (test-score–eligible sample). Each entry is a Wald estimator—the ratio of two local discontinuities in a fuzzy RD design—and may lie outside the unit interval in finite samples. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), and ICETEX.

Table 3: IV Estimates by Test-Score Selectivity

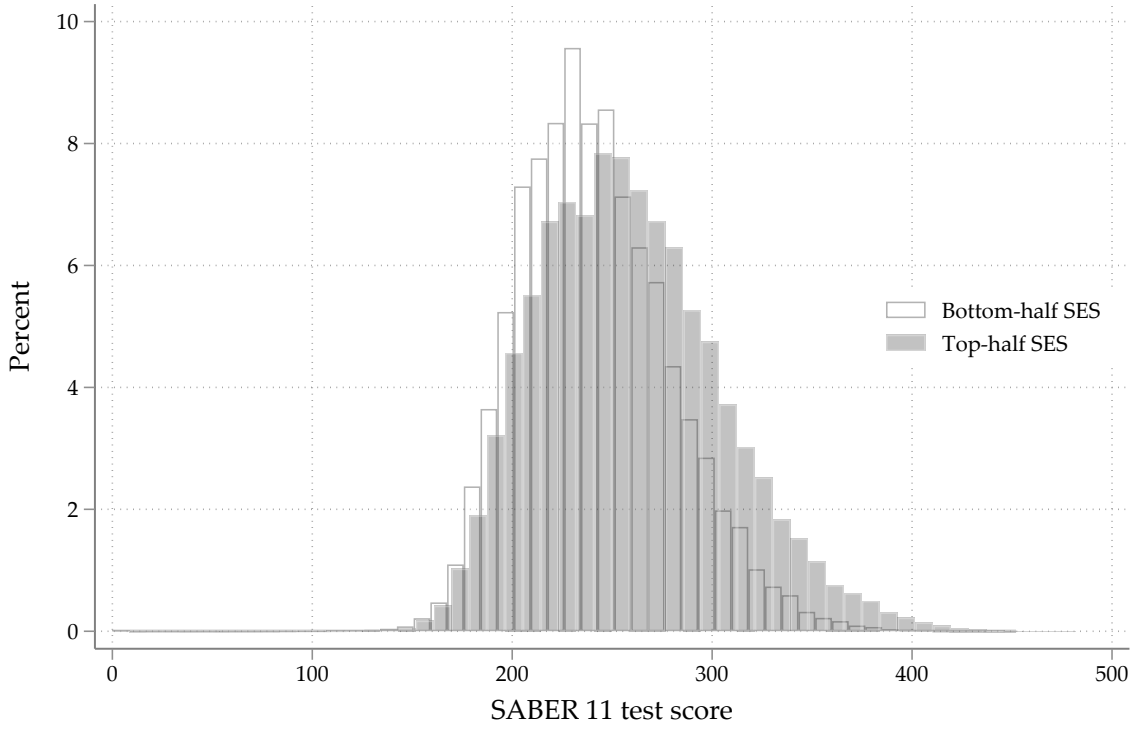
	Test-Score Percentile / Year					
	90.7 2014 (1)	91.4 2015 (2)	95.4 2016 (3)	96.5 2017 (4)	97.4 2019 (5)	98.1 2018 (6)
Any college	0.495	0.399	0.318	0.258	0.119	0.051
	(0.019)	(0.013)	(0.018)	(0.016)	(0.026)	(0.051)
	[0.417]	[0.488]	[0.542]	[0.489]	[0.653]	[0.651]
HQA college	0.810	0.643	0.454	0.388	0.258	0.375
	(0.012)	(0.012)	(0.019)	(0.017)	(0.028)	(0.090)
	[0.095]	[0.131]	[0.215]	[0.220]	[0.312]	[0.378]
Private HQA college	0.808	0.670	0.535	0.436	0.363	0.342
	(0.012)	(0.011)	(0.015)	(0.016)	(0.020)	(0.039)
	[0.028]	[0.053]	[0.068]	[0.065]	[0.063]	[0.087]
Public HQA college	-0.007	-0.027	-0.080	-0.010	-0.095	-0.083
	(0.012)	(0.009)	(0.016)	(0.010)	(0.022)	(0.066)
	[0.071]	[0.078]	[0.146]	[0.133]	[0.241]	[0.283]
Other college	-0.311	-0.246	-0.136	-0.131	-0.143	-0.279
	(0.018)	(0.014)	(0.018)	(0.013)	(0.026)	(0.078)
	[0.321]	[0.358]	[0.330]	[0.277]	[0.344]	[0.264]
Bachelor's degree program	0.597	0.497	0.366	0.292	0.165	0.072
	(0.020)	(0.014)	(0.019)	(0.016)	(0.026)	(0.060)
	[0.307]	[0.400]	[0.490]	[0.455]	[0.612]	[0.632]
Associate degree program	-0.110	-0.098	-0.050	-0.035	-0.049	-0.007
	(0.012)	(0.009)	(0.007)	(0.005)	(0.008)	(0.013)
	[0.114]	[0.088]	[0.055]	[0.035]	[0.042]	[0.017]
STEM	0.233	0.212	0.131	0.139	0.103	0.058
	(0.020)	(0.017)	(0.021)	(0.017)	(0.029)	(0.058)
	[0.192]	[0.228]	[0.284]	[0.262]	[0.371]	[0.374]
STEM Plus	0.367	0.303	0.242	0.174	0.089	0.143
	(0.023)	(0.014)	(0.023)	(0.018)	(0.028)	(0.088)
	[0.366]	[0.412]	[0.456]	[0.416]	[0.547]	[0.529]
SSH	0.096	0.076	0.060	0.070	0.037	-0.065
	(0.012)	(0.011)	(0.014)	(0.010)	(0.014)	(0.044)
	[0.042]	[0.064]	[0.076]	[0.068]	[0.087]	[0.109]
Arts	0.030	0.023	0.015	0.012	0.000	0.012
	(0.006)	(0.005)	(0.005)	(0.004)	(0.006)	(0.011)
	[0.010]	[0.012]	[0.013]	[0.010]	[0.015]	[0.008]
Value-added	0.342	0.323	0.313	0.300	0.305	0.414
	(0.021)	(0.020)	(0.024)	(0.024)	(0.033)	(0.087)
	[0.210]	[0.243]	[0.278]	[0.283]	[0.268]	[0.336]

Notes: This table reports IV estimates—scaling the reduced-form RD estimates by program take-up—on college enrollment, college type, degree program, and major immediately after high school. The last row reports effects on the earnings value-added of the first college–major pair attended. Columns correspond to different test-score cutoffs; the running variable is the distance to the cutoff, and the sample includes low-SES students. Because each outcome uses its own optimal RD bandwidth, total enrollment effects may not equal the sum across categories. Local linear RD estimates, robust standard errors (in parentheses), and control means (in brackets) are computed using rdrobust (Cattaneo et al., 2014). *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), and PILA (MinSalud).

Online Appendix

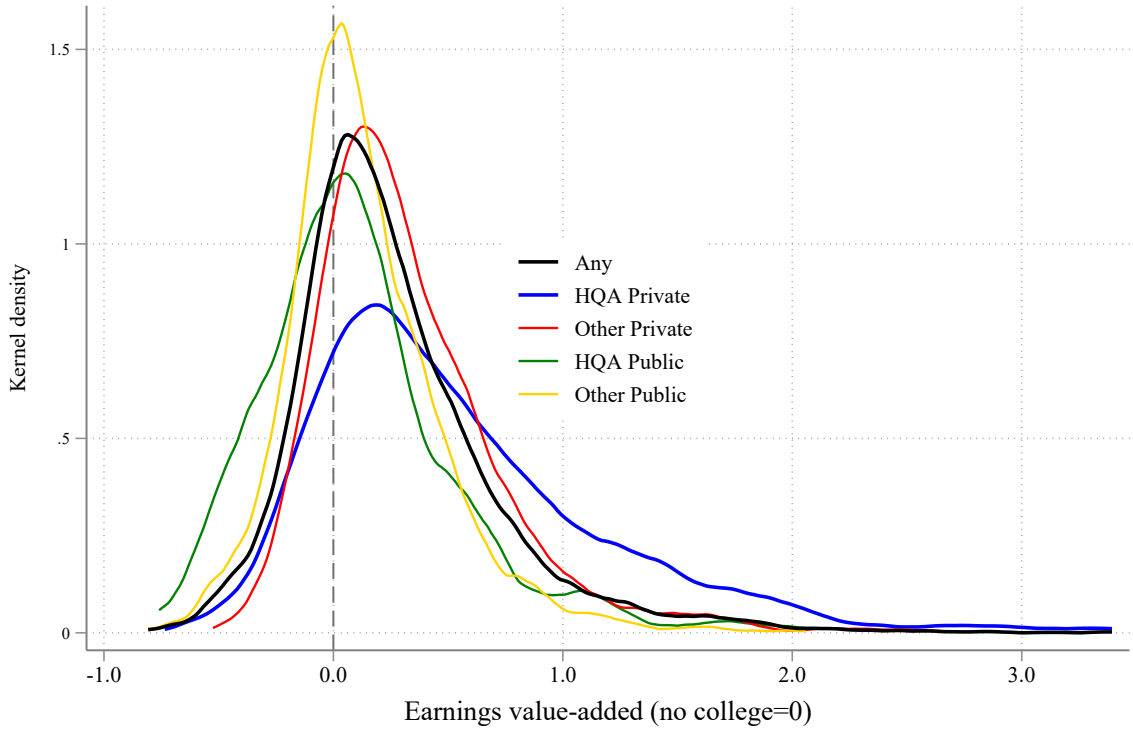
Appendix A Additional Figures and Tables

Figure A.1: The Distribution of High School Test Scores by SES



Notes: The figure plots the distribution of fall 2014 test scores for students in the bottom and top half of the wealth distribution. Students without a SISBEN score are assumed to be high-SES. *Sources:* Authors' calculations based on SABER 11 (Icfes) and SISBEN (DNP).

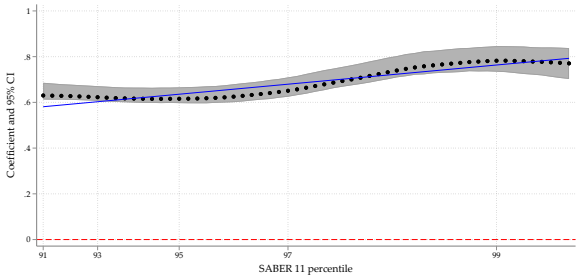
Figure A.2: The Distribution of Earnings Value-Added



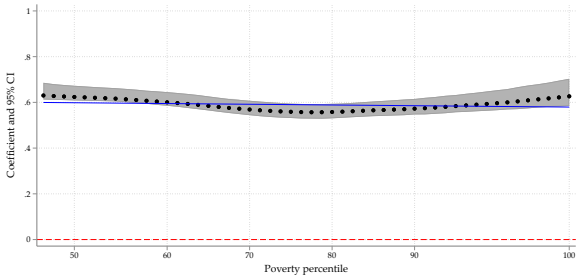
Notes: The figure shows the distribution of earnings value-added. Earnings are measured nine years after high school for college graduates, dropouts, and individuals who never attended college. College–major fixed effects are expressed relative to those with no college experience and plotted separately by college type. Controls include age, test scores (third-degree polynomials), sex, ethnicity, household size, socioeconomic stratum, SISBEN score and missing indicator, parental education, cohort, high school schedule, private-school status, urban location, and high school–by–cohort leave-one-out means of key characteristics. The sample includes students who took the exam in fall 2012 and 2013. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), and PILA (MinSalud).

Figure A.3: Heterogeneous Treatment Effects: The Take-Up Rate

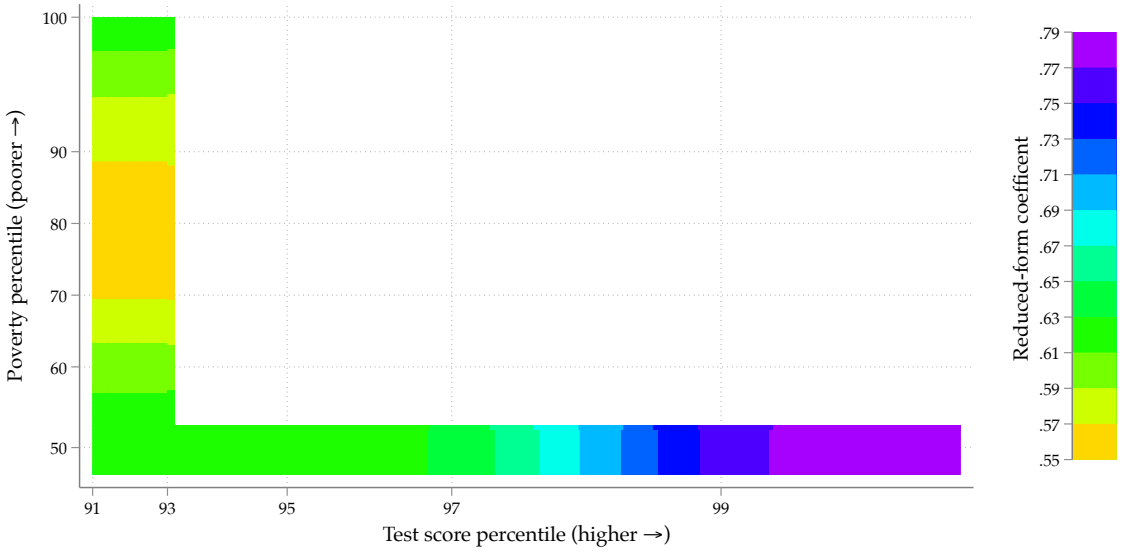
(a) By Test Score



(b) By Level of Poverty



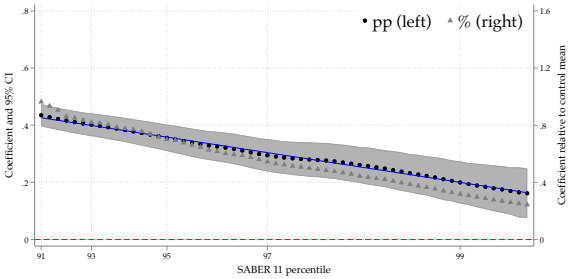
(c) Heatplot



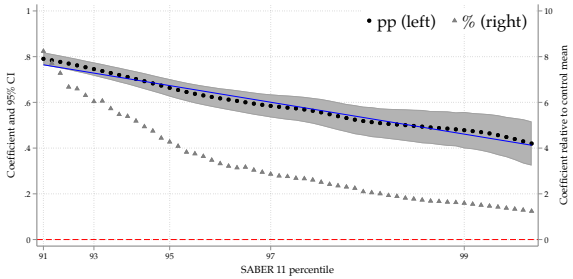
Notes: The figure presents the RD estimates of the likelihood of receiving SPP—the first-stage effect—across the distributions of test scores and poverty, estimated using the rdmulti package (Cattaneo et al., 2020b). Panels (a) and (b) display the corresponding point estimates along with 95% confidence intervals, plotted separately over the distributions of test scores and poverty, respectively, while Panel (c) compares the point estimates at the two boundaries. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), and ICETEX.

Figure A.4: Heterogeneity by Test Score: IV Estimates

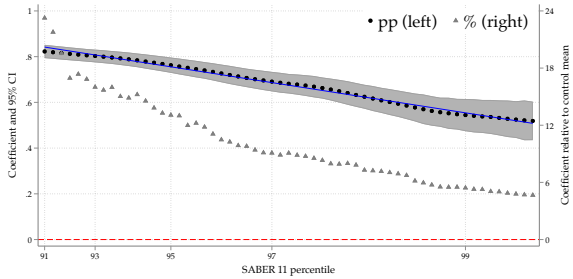
(a) Any enrollment



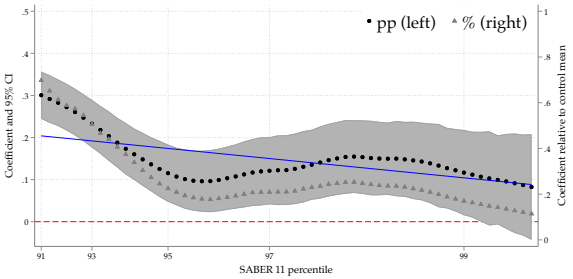
(b) HQA enrollment



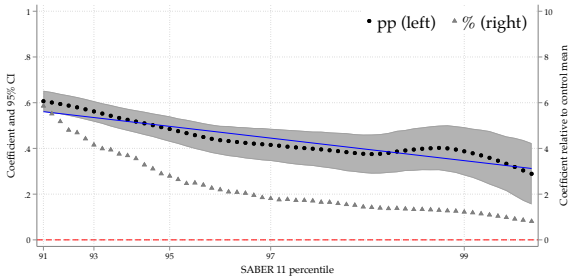
(c) Private HQA enrollment



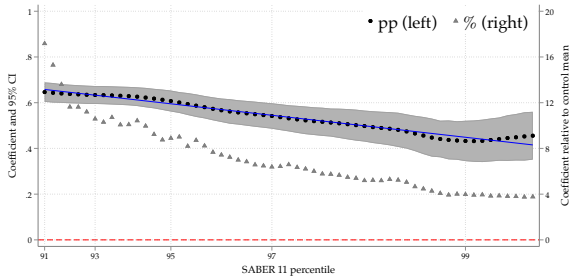
(d) Any B.A.



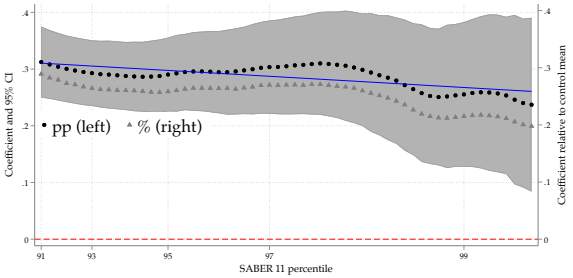
(e) HQA B.A.



(f) Private HQA B.A.



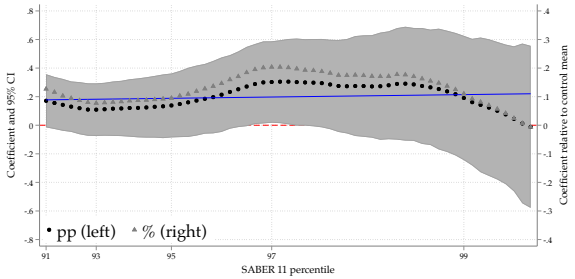
(g) Value-added



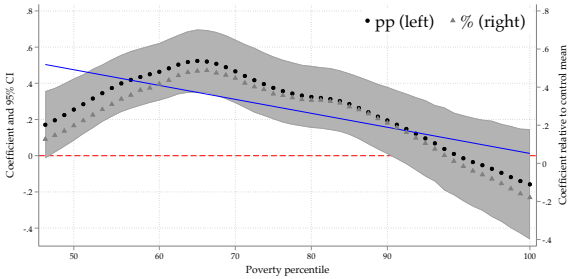
Notes: The figure compares the RD-IV estimates across the distribution of test scores for students at the poverty threshold. The gray area represents the robust bias-corrected 95% confidence intervals. The blue line is a linear fit prediction for the RD estimate on the test score. Conventional local linear RD estimates and robust bias-corrected standard errors are estimated with package rdmulti (Cattaneo et al., 2020b). Value-added is based on the average monthly earnings nine years after high school for student who ever attended a given college–major combination, relative to students who never attended any college, after controling for baseline covariates, as detailed in the main text. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), ICETEX, SNIES (MEN), SABER PRO (Icfes), and PILA (MinSalud).

Figure A.5: Heterogeneous Treatment Effects: Realized Earnings

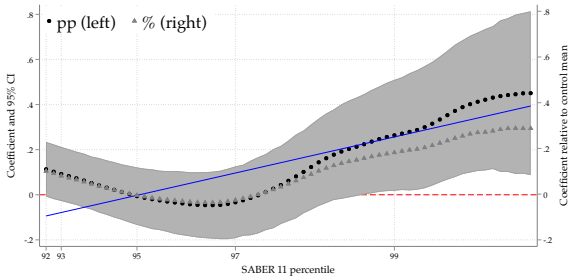
(a) By Test Score: 2014 cohort



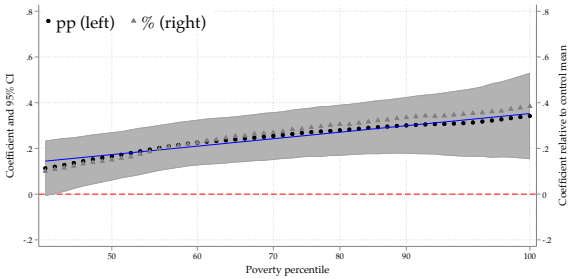
(b) By Level of Poverty: 2014 cohort



(c) By Test Score: 2015 cohort



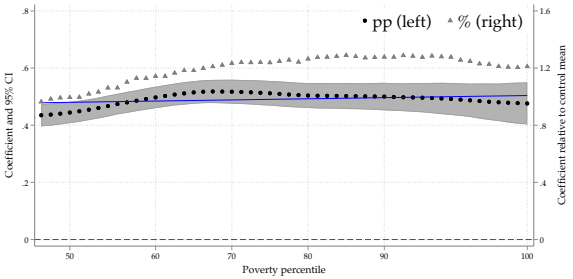
(d) By Level of Poverty: 2015 cohort



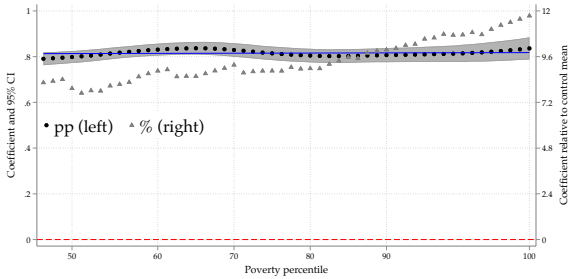
Notes: The figure reports RD-IV estimates of the impact of financial aid on formal monthly earnings in 2023, expressed in multiples of the statutory minimum wage for full-time workers, across the distributions of test scores and poverty using the rdmulti package (Cattaneo et al., 2020b). Panels (a) and (b) present results for the 2014 cohort (nine years post-high school). Panels (c) and (d) correspond to the 2015 cohort (eight years post-high school). *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), PILA (MinSalud), and ICETEX.

Figure A.6: Heterogeneity by Level of Poverty: IV Estimates

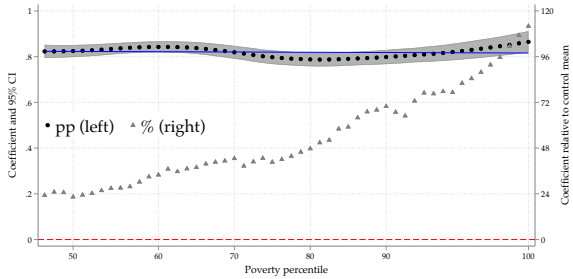
(a) Any enrollment



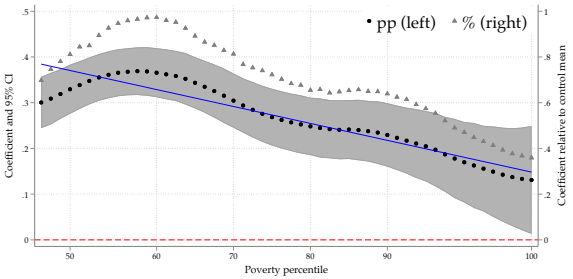
(b) HQA enrollment



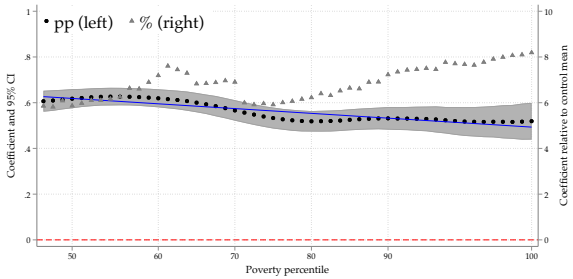
(c) Private HQA enrollment



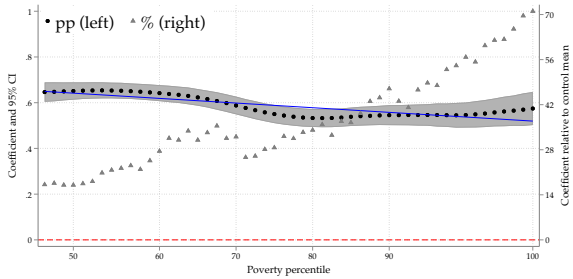
(d) Any B.A.



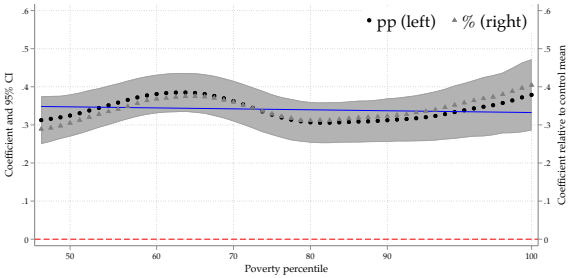
(e) HQA B.A.



(f) Private HQA B.A.

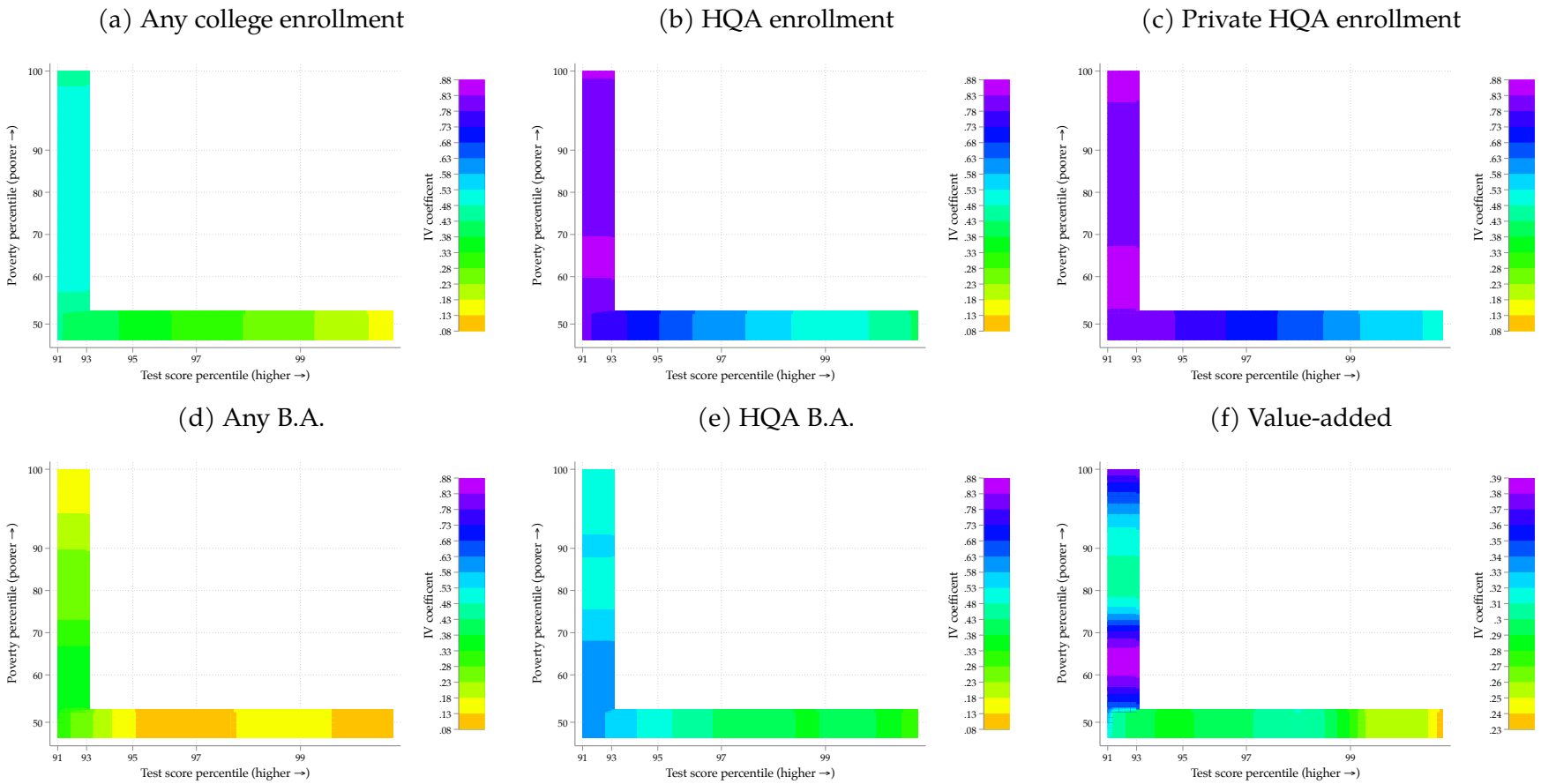


(g) Value-added



Notes: The figure compares the RD-IV estimates across the distribution of poverty (the negation of the household SISBEN score) for students at the test score threshold. The gray area represents the robust bias-corrected 95% confidence intervals. The blue line is a linear fit prediction for the RD estimate on the poverty score. Conventional local linear RD estimates and robust bias-corrected standard errors are estimated with package rdmulti (Cattaneo et al., 2020b). Value-added is based on the average monthly earnings nine years after high school for student who ever attended a given college–major combination, relative to students who never attended any college, after controling for baseline covariates, as detailed in the main text. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), ICETEX, SABER PRO (Icfes), and PILA (MinSalud).

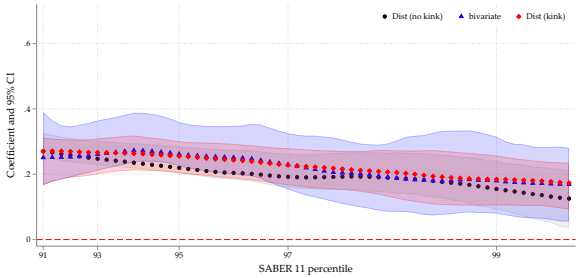
Figure A.7: Heterogeneous Treatment Effects: IV Estimates



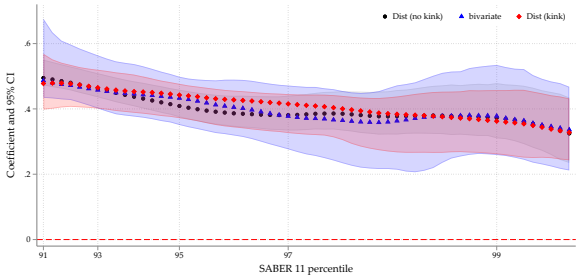
Notes: The figure compares the RD IV estimates across the distribution of test scores and poverty, estimated with package rdmulti (Cattaneo et al., 2020b). Value-added is based on average monthly earnings 9 years after high school for students who ever attended a given college–major pair, calculated using pre-reform data and detailed in the main text. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), SABER PRO (Icfes), and PILA (MinSalud).

Figure A.8: Robustness: Heterogeneous Treatment Effects by Test Score

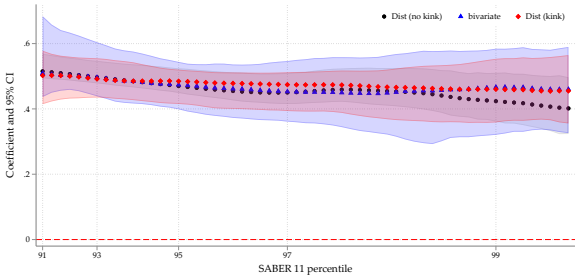
(a) Any college enrollment



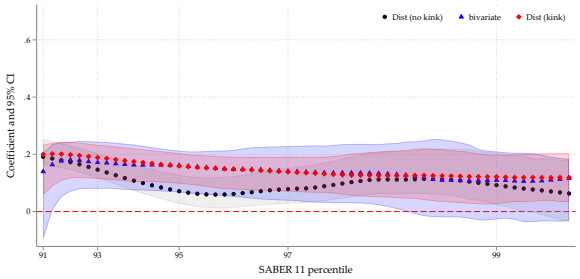
(b) HQA enrollment



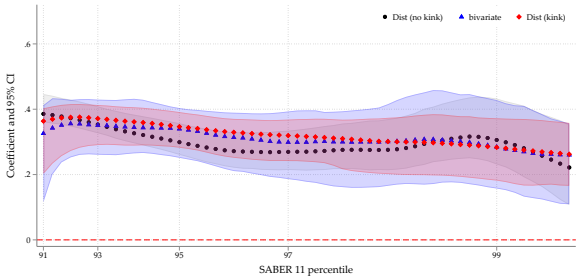
(c) Private HQA enrollment



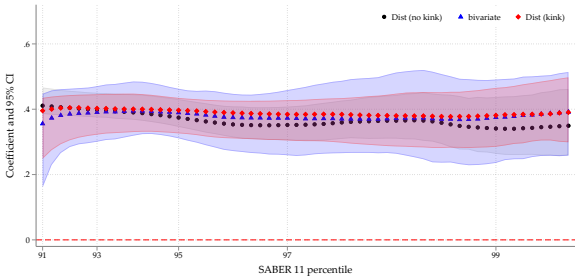
(d) Any B.A.



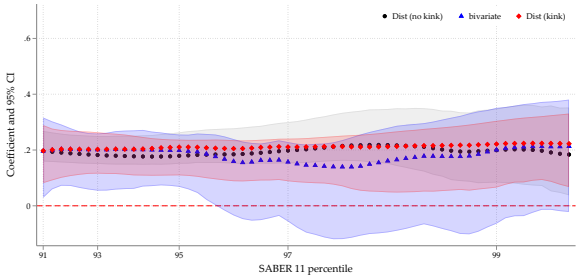
(e) HQA B.A.



(f) Private HQA B.A.



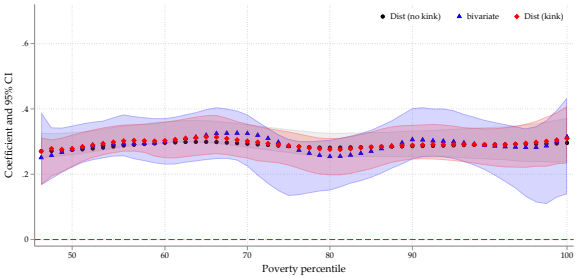
(g) Value-added



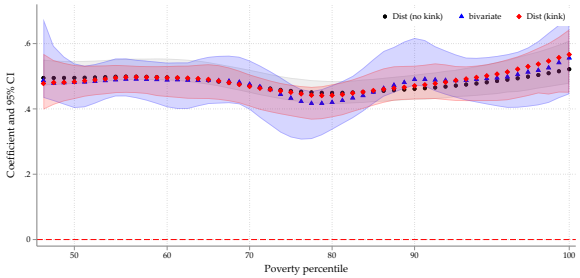
Notes: The figure compares the reduced-form local polynomial RD estimates and 95% confidence intervals across the distribution of test scores for students at the poverty threshold. Distance-based estimates without kink adjustment (in black) are estimated with package rdmulti (Cattaneo et al., 2020b). Distance-based estimates with kink adjustment (in red) and estimates based on bivariate coordinates (in blue) are estimated with package rd2d (Cattaneo et al., 2025). *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), SABER PRO (Icfes), SABER T&T (Icfes), and PILA (MinSalud).

Figure A.9: Robustness: Heterogeneous Treatment Effects by Poverty

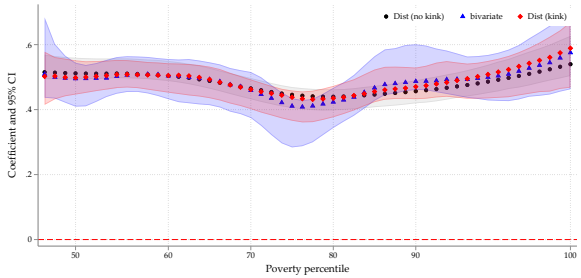
(a) Any college enrollment



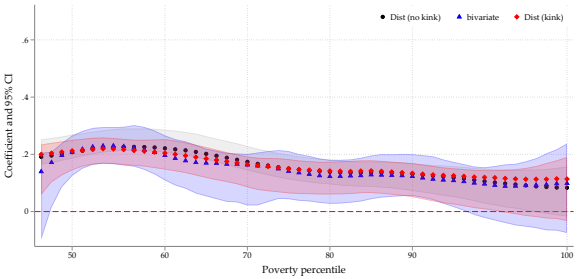
(b) HQA enrollment



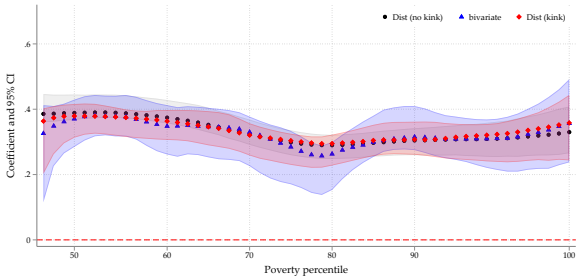
(c) Private HQA enrollment



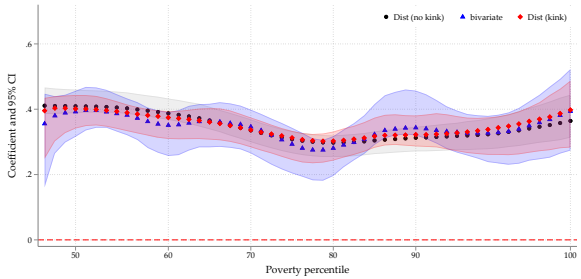
(d) Any B.A.



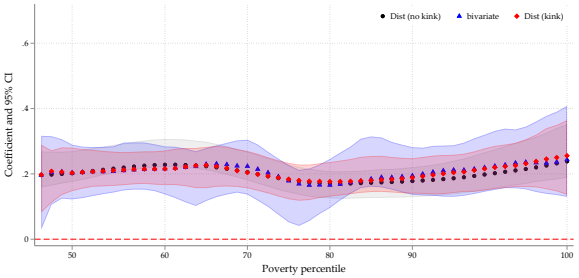
(e) HQA B.A.



(f) Private HQA B.A.



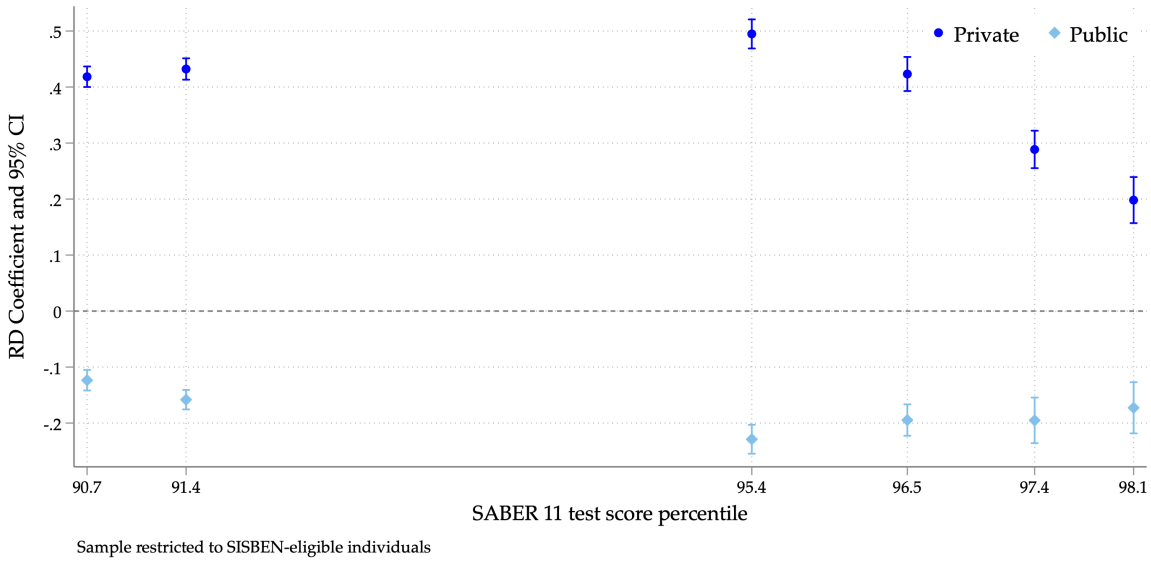
(g) Value-added



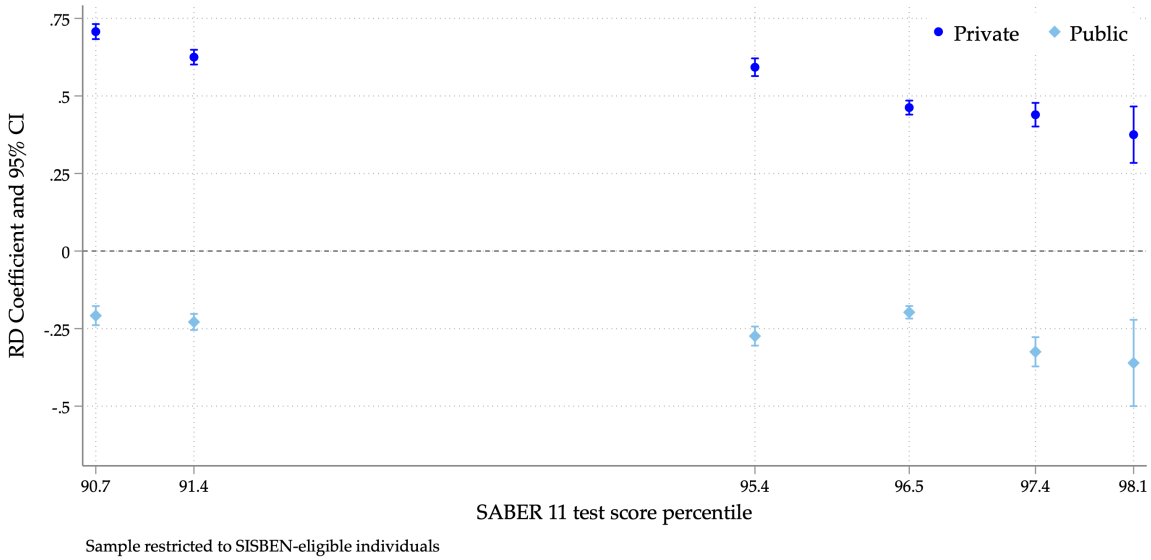
Notes: The figure compares the reduced-form local polynomial RD estimates and 95% confidence intervals across the distribution of poverty for students at the test-score threshold. Distance-based estimates without kink adjustment (in black) are estimated with package rdmulti (Cattaneo et al., 2020b). Distance-based estimates with kink adjustment (in red) and estimates based on bivariate coordinates (in blue) are estimated with package rd2d (Cattaneo et al., 2025). *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), SABER PRO (Icfes), SABER T&T (Icfes), and PILA (MinSalud).

Figure A.10: Students Shift from Public to Private Institutions

(a) Reduced Form



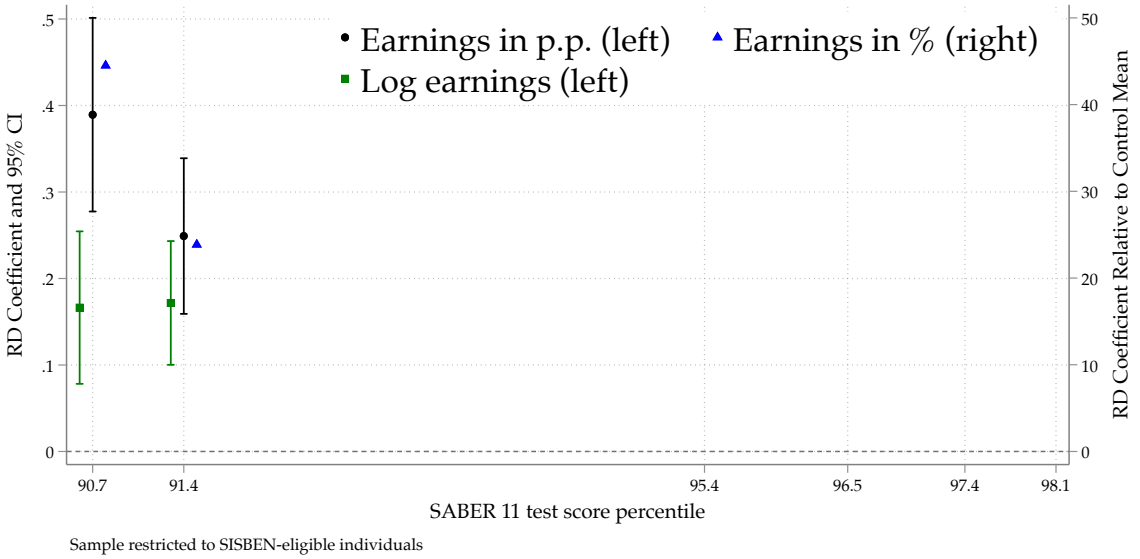
(b) IV



Notes: The figures plot the impact of financial aid on the likelihood of enrolling in a public vs. private institution immediately after completing high school as a function of the test-score cutoff. Panel (a) plots the reduced-form effect, while Panel (b) scales it by the likelihood of receiving financial aid. The sample is restricted to SISBEN-eligible individuals. Conventional local linear RD estimates and 95% confidence intervals estimated with package rdrobust (Cattaneo et al., 2014). *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), and ICETEX.

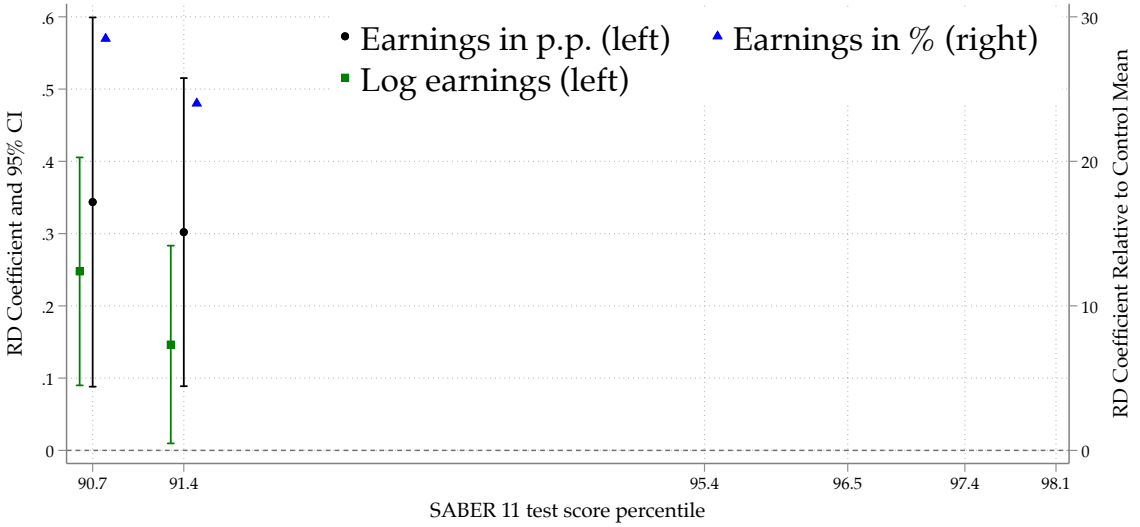
Figure A.11: Earnings Impacts Do Not Rise With Test Scores

(a) Test score is the running variable



Sample restricted to SISBEN-eligible individuals

(b) Poverty is the running variable



Sample restricted to test-score-eligible individuals

Notes: The figures show 2SLS estimates of the impact of financial aid on formal monthly earnings eight years after high school by the test-score cutoff. Earnings include zeros for individuals not formally employed and are expressed in multiples of the monthly minimum wage. They are measured in 2022 for the 2014 cohort (90.7 percentile) and 2023 for the 2015 cohort (91.4th percentile). Panel (a) uses the test score as the running variable among SISBEN-eligible individuals; panel (b) uses the poverty index among test-score–eligible individuals. Black markers report conventional local linear RD estimates with 95% confidence intervals from rdrobust package (Cattaneo et al., 2014); blue markers scale effects by the control mean; and green markers show impacts on log earnings. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), PILA (MinSalud), and ICETEX.

Table A.1: Student Characteristics By Baseline Test Score and Poverty (2014 Cohort)

	All	Min. score	Restricted to Low-SES Students					
	students	310	310	318	342	348	350	359
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SISBEN Score	36.32	45.49	34.99	35.32	36.21	36.25	36.15	36.22
Female	0.551	0.45	0.43	0.418	0.367	0.358	0.353	0.335
Age	17.66	16.40	16.37	16.33	16.24	16.22	16.22	16.18
Ethnic minority	0.047	0.033	0.041	0.042	0.043	0.041	0.041	0.035
Mother's education: primary	0.366	0.1	0.209	0.197	0.148	0.14	0.137	0.126
Mother's education: secondary	0.432	0.331	0.458	0.444	0.413	0.395	0.396	0.384
Mother's education: T&T	0.091	0.174	0.16	0.168	0.179	0.178	0.176	0.181
Mother's education: professional	0.112	0.395	0.173	0.191	0.26	0.287	0.291	0.309
Father's education: primary	0.43	0.142	0.283	0.265	0.216	0.206	0.204	0.189
Father's education: secondary	0.386	0.319	0.417	0.411	0.362	0.345	0.344	0.334
Father's education: T&T	0.073	0.14	0.121	0.126	0.151	0.159	0.16	0.165
Father's education: professional	0.111	0.399	0.179	0.198	0.271	0.29	0.292	0.313
Household SES: Stratum 1	0.425	0.12	0.298	0.279	0.226	0.214	0.209	0.191
Household SES: Stratum 2	0.346	0.302	0.45	0.452	0.45	0.453	0.447	0.456
Household SES: Stratum 3	0.176	0.349	0.221	0.233	0.276	0.282	0.29	0.295
Household SES: Strata 4-6	0.050	0.228	0.028	0.032	0.046	0.049	0.052	0.057
School hours: Full day	0.196	0.454	0.226	0.24	0.282	0.294	0.294	0.31
Private school	0.255	0.527	0.218	0.237	0.307	0.327	0.333	0.355
Family has internet	0.506	0.827	0.648	0.666	0.721	0.729	0.734	0.742
Family has a laptop	0.616	0.897	0.776	0.79	0.831	0.837	0.838	0.845
Family has a car	0.206	0.466	0.21	0.219	0.245	0.254	0.257	0.271
School location: Urban	0.862	0.941	0.929	0.936	0.958	0.959	0.961	0.964
Vulnerable regions	0.071	0.036	0.053	0.052	0.045	0.044	0.042	0.046
Not Bogota	0.831	0.685	0.825	0.822	0.813	0.812	0.812	0.822
N	542,341	50,547	14,487	10,025	3,099	2,298	2,060	1,273

Notes: This table presents baseline characteristics fpr students who took the SABER 11 exam in fall 2014. Column (1) reports statistics for all test takers registered as students, i.e., those with school codes. Column (2) restricts the sample to students scoring at least 310 out of 500, corresponding to the top 9.3% of the score distribution. Columns (3)–(8) further limit the sample to students eligible for financial aid based on their household SISBEN score. Column (3) includes all low-SES students within the top 9.3% of scores, while Columns (4)–(8) progressively focus on students with increasingly higher SABER 11 scores. *Sources:* Authors' calculations based on SABER 11 (Icfes) and SISBEN (DNP).

Table A.2: Baseline Characteristics by College Type and Date of HQA

	Date by which HQA was obtained							
	Before October 2014				After October 2014			
	HQA		No HQA		HQA		No HQA	
	Private (1)	Public (2)	Private (3)	Public (4)	Private (5)	Public (6)	Private (7)	Public (8)
SABER 11 score	303.10	293.32	258.94	262.19	277.59	289.66	255.78	255.22
(Standardized) SABER 11 score	1.77	1.44	0.30	0.41	0.92	1.32	0.19	0.17
(Standardized) SABER PRO score	0.50	0.27	-0.18	-0.16	0.05	0.12	-0.22	-0.24
Graduation rate (%)	75.08	66.61	59.30	59.58	71.23	67.59	57.28	57.55
Faculty with Ph.D. (%)	11.86	15.16	1.91	2.43	5.22	4.72	1.29	1.83
Program length (semesters)	9.18	9.59	9.16	9.44	9.29	9.38	9.14	9.45
Sticker price of tuition (in min. wages)	17.65	6.39	8.06	3.72	12.18	5.77	7.36	2.96
Public spending per student (in min. wages)		14.49		8.53		6.58		9.01
Earnings value-added (9 years after high school)	0.67	0.05	0.33	0.12	0.42	0.01	0.31	0.15
N students	50,642	60,704	214,203	267,636	45,618	41,262	168,585	226,374
N college campuses	21	16	207	79	30	16	177	63
N colleges	20	13	181	66	16	13	165	53

Notes: This table reports descriptive statistics by college type and by the date on which each institution obtained HQA status. Columns (3)–(4) and (7)–(8) include institutions that had not obtained HQA by October 2014 or that never obtained it, respectively. The sample consists of students who took the high school exit exam in fall 2012 and 2013 and enrolled in college within six years of graduation. Earnings value-added is measured nine years after high school and expressed relative to individuals who did not enroll in any college within six years of high school completion. *Sources:* Authors' calculations based on SABER 11 (Icfes), SABER PRO (Icfes), SNIES (MEN), institutional financial accounts and balance sheets (MEN), CNA (MEN), PILA (MinSalud).

Table A.3: College Enrollment by Household Wealth, Test Scores, and College Type

	None	Any college	HQA College			Other college
			Any	Private	Public	
	%	%	%	%	%	%
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: *Enrollment immediately after high school*					
All students						
Low SES	82.6	17.4	2.4	0.8	1.6	15.0
High SES	71.7	28.3	7.8	5.0	2.7	20.5
Top 10% test scores						
Low SES	52.2	47.9	17.8	4.8	13.1	30.0
High SES	42.9	57.1	30.8	18.7	12.1	26.3
	Panel B: *Enrollment within six years from high school*					
All students						
Low SES	54.66	45.34	6.17	1.64	4.53	39.18
High SES	41.19	58.81	13.61	7.77	5.84	45.2
Top 10% test scores						
Low SES	17.28	82.72	29.96	6.75	23.21	52.76
High SES	13.2	86.8	44.77	25.74	19.03	42.03

Notes: This table reports descriptive statistics on college enrollment by household wealth, SABER 11 test scores, and college type, prior to the expansion of financial aid. Panel (a) shows enrollment immediately after high school graduation, and Panel (b) shows enrollment within six years. The sample consists of fall 2013 high school exit test takers. Students are classified into two SES groups—above- and below-median wealth—based on their household's SISBEN wealth score, and further by whether they scored in the top 10% of the SABER 11 distribution. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), and SNIES (MEN).

Table A.4: Reduced Form Estimates Using SISBEN

	Test-Score Percentile / Year					
	90.7 2014 (1)	91.4 2015 (2)	95.4 2016 (3)	96.5 2017 (4)	97.4 2019 (5)	98.1 2018 (6)
Any college	0.236	0.220	0.264	0.178	0.026	0.023
	(0.025)	(0.029)	(0.047)	(0.060)	(0.062)	(0.067)
	[0.539]	[0.609]	[0.585]	[0.543]	[0.700]	[0.681]
HQA college	0.426	0.403	0.399	0.225	0.050	0.082
	(0.023)	(0.030)	(0.048)	(0.055)	(0.060)	(0.075)
	[0.244]	[0.276]	[0.308]	[0.380]	[0.497]	[0.534]
Private HQA college	0.485	0.466	0.394	0.221	0.118	0.101
	(0.019)	(0.025)	(0.039)	(0.050)	(0.055)	(0.067)
	[0.075]	[0.102]	[0.137]	[0.149]	[0.174]	[0.202]
Public HQA college	-0.090	-0.108	-0.079	-0.027	-0.021	-0.080
	(0.022)	(0.024)	(0.044)	(0.061)	(0.065)	(0.071)
	[0.237]	[0.253]	[0.273]	[0.284]	[0.366]	[0.430]
Other college	-0.184	-0.193	-0.129	-0.064	-0.029	-0.054
	(0.021)	(0.032)	(0.035)	(0.041)	(0.057)	(0.056)
	[0.292]	[0.346]	[0.280]	[0.174]	[0.211]	[0.135]
Bachelor's degree program	0.296	0.296	0.346	0.192	0.020	0.028
	(0.026)	(0.026)	(0.049)	(0.061)	(0.066)	(0.067)
	[0.445]	[0.513]	[0.503]	[0.509]	[0.698]	[0.680]
Associate degree program	-0.061	-0.083	-0.066	-0.015	-0.011	0.001
	(0.012)	(0.017)	(0.018)	(0.016)	(0.013)	(0.004)
	[0.094]	[0.099]	[0.071]	[0.028]	[0.015]	[-0.001]
STEM	0.089	0.114	0.073	0.017	0.007	0.080
	(0.027)	(0.032)	(0.044)	(0.054)	(0.059)	(0.084)
	[0.284]	[0.325]	[0.367]	[0.363]	[0.406]	[0.396]
STEM Plus	0.136	0.138	0.118	0.072	-0.030	0.040
	(0.030)	(0.032)	(0.046)	(0.052)	(0.067)	(0.079)
	[0.461]	[0.517]	[0.518]	[0.489]	[0.581]	[0.536]
SSH	0.102	0.077	0.112	0.065	0.061	-0.037
	(0.021)	(0.018)	(0.032)	(0.033)	(0.047)	(0.053)
	[0.049]	[0.067]	[0.069]	[0.056]	[0.100]	[0.143]
Arts	0.018	0.001	0.030	0.026	0.005	0.021
	(0.009)	(0.010)	(0.015)	(0.014)	(0.014)	(0.016)
	[0.014]	[0.023]	[0.007]	[0.007]	[0.003]	[-0.001]
Value-added	0.202	0.130	0.103	0.147	0.099	0.160
	(0.027)	(0.040)	(0.056)	(0.080)	(0.099)	(0.123)
	[0.318]	[0.397]	[0.458]	[0.412]	[0.356]	[0.415]

Notes: This table presents the reduced-form effect of financial aid eligibility on the likelihood of accessing college immediately after high school completion using an RDD. The running variable is the distance to the SISBEN score cutoff and the sample is restricted to SABER 11-eligible students. Because each outcome uses a distinct optimal RD bandwidth, the effect on overall enrollment need not equal the sum across mutually exclusive categories. The last row presents effects on earnings "value-added" of the attended college–major pair, as described in the main text. Conventional local linear RD estimates, robust standard errors (in parentheses), and conventional control means (in brackets) are estimated with package rdrobust (Cattaneo et al., 2014). *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), and PILA (MinSalud)

Table A.5: First Stage Estimates Over Time

	Test-Score Percentile Cutoff / Year					
	90.7	91.4	95.4	96.5	97.4	98.1
	2014	2015	2016	2017	2019	2018
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: *SABER 11 is the running variable*						
RD Coefficient	0.596	0.692	0.835	0.882	0.649	0.385
	(0.011)	(0.009)	(0.010)	(0.009)	(0.015)	(0.032)
Control mean	[0.000]	[0.000]	[0.000]	[0.010]	[0.001]	[0.450]
Panel B: *SISBEN is the running variable*						
RD Coefficient	0.659	0.745	0.723	0.829	0.469	0.749
	(0.018)	(0.016)	(0.024)	(0.023)	(0.043)	(0.052)
Control mean	[0.005]	[0.012]	[0.064]	[0.038]	[0.105]	[0.049]

Notes: This table reports RD estimates of the effect of financial aid eligibility on the probability of receiving SPP (Columns 1–4) or GE (Columns 5–6). Panel (a) uses distance to the SABER 11 cutoff as the running variable, restricting to SISBEN-eligible students; Panel (b) uses distance to the SISBEN score, restricting to SABER 11-eligible students. Estimates are based on local linear regressions using rdrobust package (Cattaneo et al., 2014), with robust standard errors in parentheses and control means in brackets. The effect of eligibility on aid receipt rises over time, except in 2018, when an unusually high test-score cutoff and uncertainty after the change in administration reduced take-up. To meet the 4,000 beneficiary target, the government extended awards to additional SISBEN-eligible students. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), and ICETEX.

Table A.6: Characteristics of Aid Recipients (Fall Cohorts of 2014–2019)

	SPP				GE	
	2014	2015	2016	2017	2019	2018
	(1)	(2)	(3)	(4)	(5)	(6)
SABER 11 percentile	94.8	95.3	97.5	98.0	98.6	98.9
Female	0.433	0.423	0.437	0.426	0.409	0.409
Age	16.314	16.285	16.242	16.217	16.183	16.192
Ethnic minority	0.041	0.026	0.027	0.040	0.030	0.024
Mother's education: primary	0.192	0.174	0.127	0.102	0.084	0.070
Mother's education: secondary	0.455	0.456	0.413	0.342	0.324	0.313
Mother's education: T&T	0.173	0.169	0.199	0.216	0.219	0.225
Mother's education: professional	0.181	0.201	0.261	0.340	0.373	0.393
Father's education: primary	0.268	0.240	0.198	0.154	0.139	0.121
Father's education: secondary	0.417	0.422	0.400	0.366	0.338	0.337
Father's education: T&T	0.128	0.139	0.143	0.164	0.179	0.160
Father's education: professional	0.188	0.198	0.259	0.316	0.344	0.382
Household SES: Stratum 1	0.285	0.321	0.314	0.282	0.240	0.253
Household SES: Stratum 2	0.453	0.454	0.450	0.448	0.431	0.433
Household SES: Stratum 3	0.230	0.202	0.205	0.212	0.253	0.242
Household SES: Strata 4-6	0.029	0.023	0.030	0.037	0.046	0.046
School hours: Full day	0.229	0.236	0.256	0.247	0.261	0.301
Private school	0.225	0.241	0.309	0.328	0.355	0.402
Family has internet	0.666	0.655	0.665	0.720	0.811	0.801
Family has a laptop	0.793	0.764	0.732	0.754	0.763	0.781
Family has a car	0.205	0.203	0.206	0.249	0.316	0.310
School location: Urban	0.930	0.933	0.949	0.950	0.962	0.955
Vulnerable regions	0.059	0.060	0.057	0.075	0.079	0.067
Not Bogota	0.817	0.852	0.887	0.906	0.943	0.918
N	9,373	11,929	8,618	7,180	4,126	3,461

Notes: This table reports baseline characteristics of financial aid recipients for cohorts taking the SABER 11 exam in the fall semesters of 2014 through 2019. Columns (1)–(4) present statistics for SPP recipients, while Columns (5) and (6) present statistics for GE recipients. *Sources:* Authors' calculations based on SABER 11 (Icfes) and ICETEX.

Table A.7: HQA Enrollment Impact when HQA Status is Defined Annually

	Test-Score Percentile / Year					
	90.7 2014 (1)	91.4 2015 (2)	95.4 2016 (3)	96.5 2017 (4)	97.4 2019 (5)	98.1 2018 (6)
Panel A: *SABER 11 is the running variable*						
HQA college (annual)	0.480	0.523	0.500	0.396	0.168	0.130
	(0.011)	(0.011)	(0.015)	(0.018)	(0.026)	(0.023)
	[0.095]	[0.149]	[0.258]	[0.299]	[0.520]	[0.476]
Other college (annual)	-0.184	-0.250	-0.238	-0.166	-0.084	-0.099
	(0.011)	(0.009)	(0.011)	(0.010)	(0.012)	(0.016)
	[0.321]	[0.345]	[0.287]	[0.194]	[0.132]	[0.163]
Panel B: *SISBEN is the running variable*						
HQA college (annual)	0.426	0.449	0.458	0.251	0.099	0.081
	(0.023)	(0.029)	(0.046)	(0.063)	(0.066)	(0.070)
	[0.244]	[0.305]	[0.367]	[0.433]	[0.611]	[0.622]
Other college (annual)	-0.184	-0.249	-0.189	-0.054	-0.060	-0.053
	(0.021)	(0.028)	(0.031)	(0.028)	(0.027)	(0.030)
	[0.292]	[0.318]	[0.222]	[0.081]	[0.083]	[0.052]

Notes: This table reports reduced-form RD estimates of the effect of financial aid eligibility on enrollment in an HQA college, defined annually, immediately after high school. In Panel (a), the running variable is the distance to the SABER 11 cutoff, restricting the sample to SISBEN-eligible students. In Panel (b), the running variable is the distance to the SISBEN cutoff, restricting the sample to SABER 11-eligible students. Local linear RD estimates, robust standard errors (in parentheses), and control means (in brackets) are computed using rdrobust package (Cattaneo et al., 2014).
Sources: Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), and SNIES (MEN).

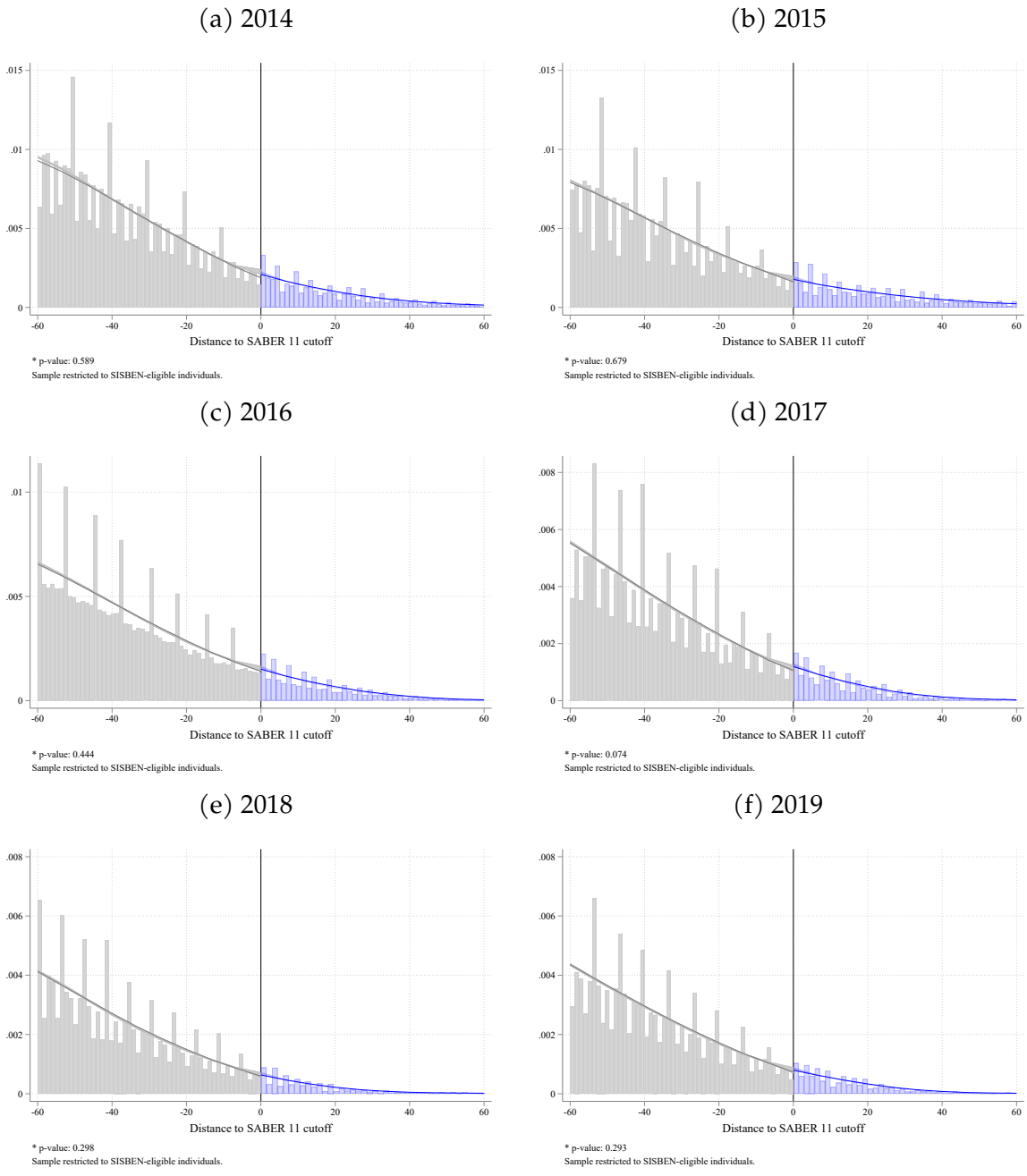
Table A.8: IV Estimates Using SISBEN

	Test-Score Percentile / Year					
	90.7 2014 (1)	91.4 2015 (2)	95.4 2016 (3)	96.5 2017 (4)	97.4 2019 (5)	98.1 2018 (6)
Any college	0.400	0.300	0.339	0.172	0.102	0.028
	(0.044)	(0.034)	(0.053)	(0.054)	(0.124)	(0.102)
	[0.518]	[0.603]	[0.601]	[0.573]	[0.686]	[0.675]
HQA college	0.661	0.545	0.526	0.249	0.109	0.091
	(0.037)	(0.039)	(0.056)	(0.056)	(0.125)	(0.116)
	[0.239]	[0.275]	[0.318]	[0.401]	[0.499]	[0.544]
Private HQA college	0.747	0.633	0.541	0.285	0.261	0.119
	(0.032)	(0.031)	(0.049)	(0.053)	(0.116)	(0.094)
	[0.069]	[0.101]	[0.137]	[0.148]	[0.173]	[0.212]
Public HQA college	-0.088	-0.089	-0.022	0.000	-0.147	-0.007
	(0.034)	(0.033)	(0.046)	(0.057)	(0.128)	(0.112)
	[0.171]	[0.173]	[0.185]	[0.227]	[0.320]	[0.318]
Other college	-0.271	-0.249	-0.178	-0.085	-0.024	-0.084
	(0.036)	(0.040)	(0.048)	(0.036)	(0.119)	(0.086)
	[0.284]	[0.336]	[0.280]	[0.177]	[0.196]	[0.135]
Bachelor's degree program	0.483	0.401	0.425	0.190	0.118	0.046
	(0.044)	(0.036)	(0.053)	(0.056)	(0.126)	(0.096)
	[0.430]	[0.513]	[0.532]	[0.547]	[0.671]	[0.674]
Associate degree program	-0.087	-0.093	-0.085	-0.022	-0.023	-0.001
	(0.023)	(0.017)	(0.023)	(0.015)	(0.027)	(0.006)
	[0.090]	[0.087]	[0.069]	[0.026]	[0.016]	[-0.000]
STEM	0.136	0.154	0.099	0.021	0.018	0.115
	(0.044)	(0.040)	(0.058)	(0.053)	(0.126)	(0.102)
	[0.284]	[0.323]	[0.373]	[0.349]	[0.405]	[0.388]
STEM Plus	0.210	0.190	0.161	0.091	-0.057	0.070
	(0.046)	(0.041)	(0.058)	(0.055)	(0.144)	(0.115)
	[0.460]	[0.516]	[0.527]	[0.484]	[0.577]	[0.526]
SSH	0.159	0.110	0.148	0.064	0.142	-0.066
	(0.032)	(0.027)	(0.042)	(0.034)	(0.103)	(0.075)
	[0.049]	[0.066]	[0.070]	[0.068]	[0.098]	[0.148]
Arts	0.029	0.000	0.041	0.024	0.014	0.031
	(0.014)	(0.014)	(0.020)	(0.016)	(0.031)	(0.025)
	[0.013]	[0.024]	[0.007]	[0.012]	[0.001]	[-0.001]
Value-added	0.278	0.180	0.165	0.240	0.221	0.229
	(0.054)	(0.055)	(0.065)	(0.080)	(0.223)	(0.145)
	[0.316]	[0.396]	[0.445]	[0.394]	[0.353]	[0.410]

Notes: This table presents the IV effect of financial aid, which scales the reduced-form effect by the likelihood of receiving financial aid, using an RDD. The running variable is the distance to the SISBEN score cutoff and the sample is restricted to SABER 11-eligible students. Because each outcome uses a distinct optimal RD bandwidth, the effect on overall enrollment need not equal the sum across mutually exclusive categories. Conventional local linear RD estimates, robust standard errors (in parentheses), and conventional control means (in brackets) are estimated with package rdrobust (Cattaneo et al., 2014). *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), ICETEX, and PILA (MinSalud).

Appendix B RD Validity

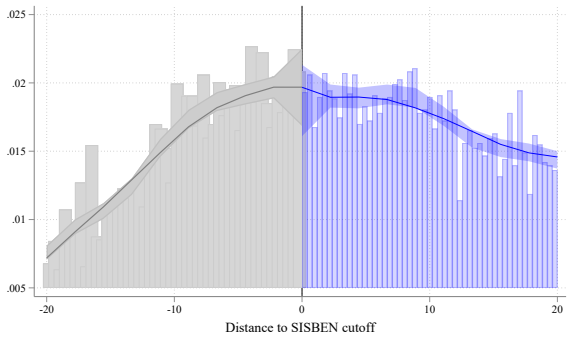
Figure B.1: Manipulation Testing based on Density Discontinuity of Test Scores



Notes: This figure tests for manipulation test scores based on density discontinuity, estimated with package rddensity (Cattaneo et al., 2018). The sample is restricted to low-SES individuals. The p-values suggest we cannot statistically detect manipulation. *Sources:* Authors' calculations based on SABER 11 (Icfes) and SISBEN (DNP).

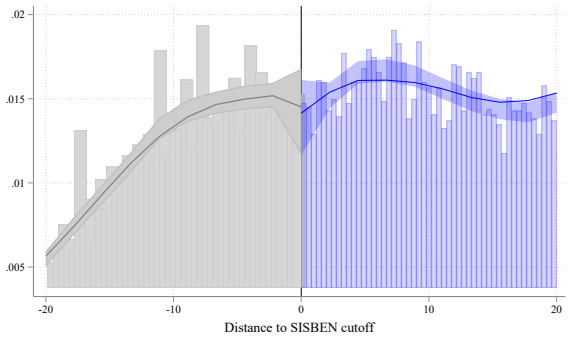
Figure B.2: Manipulation Testing based on Density Discontinuity of SISBEN Scores

(a) 2014



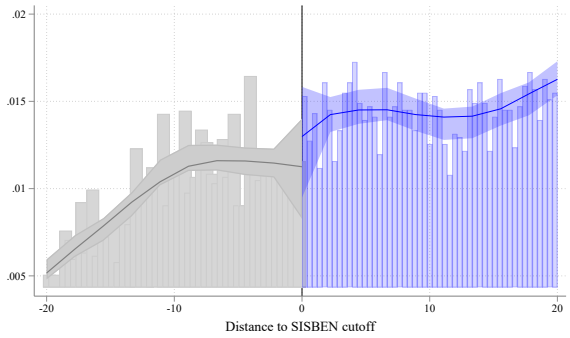
* p-value: 0.624
Sample restricted to SABER 11-eligible individuals.

(b) 2015



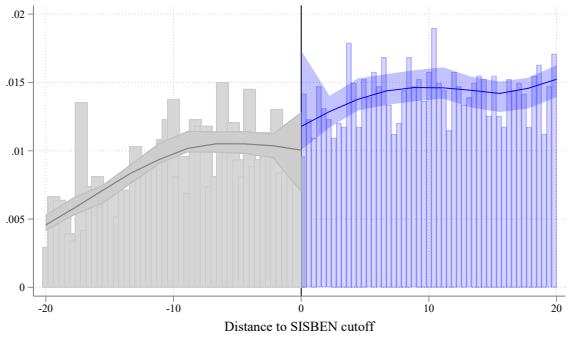
* p-value: 0.837
Sample restricted to SABER 11-eligible individuals.

(c) 2016



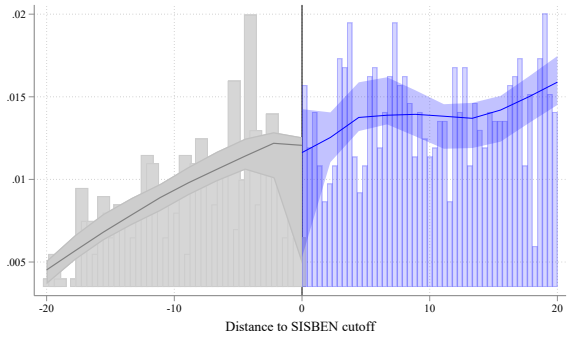
* p-value: 0.474
Sample restricted to SABER 11-eligible individuals.

(d) 2017



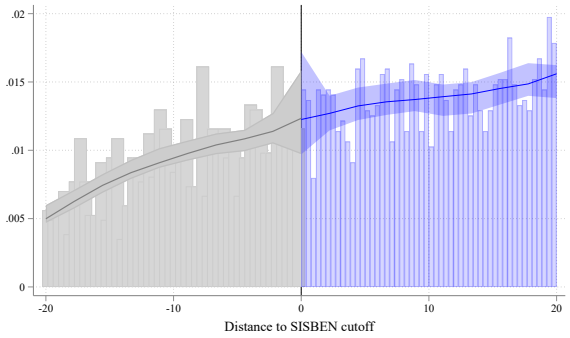
* p-value: 0.129
Sample restricted to SABER 11-eligible individuals.

(e) 2018



* p-value: 0.6980000000000001
Sample restricted to SABER 11-eligible individuals.

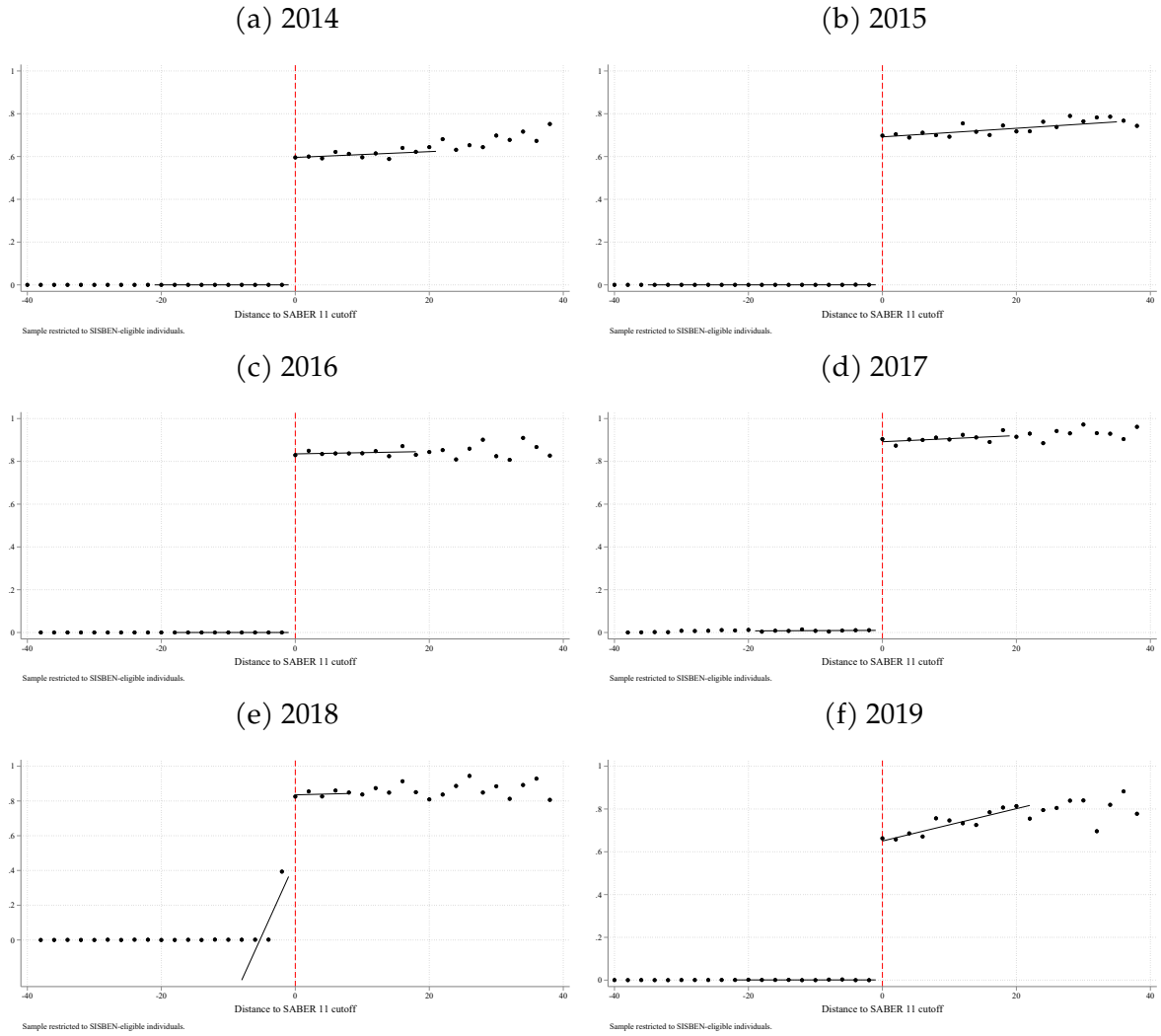
(f) 2019



* p-value: 0.8
Sample restricted to SABER 11-eligible individuals.

Notes: This figure tests for manipulation of the poverty score based on density discontinuity, estimated with package rddensity (Cattaneo et al., 2018). The sample is restricted to SABER 11-eligible test takers. The p-values suggest we cannot statistically detect manipulation. *Sources:* Authors' calculations based on SABER 11 (Icfes) and SISBEN (DNP).

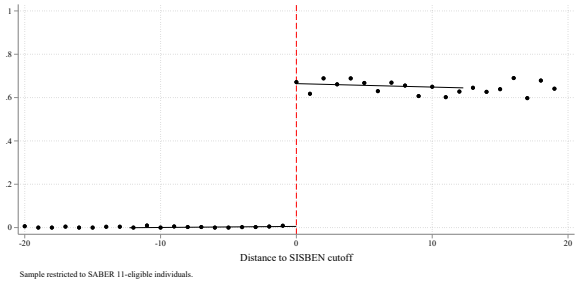
Figure B.3: The Probability of Receiving Aid: SABER 11 is the Running Variable



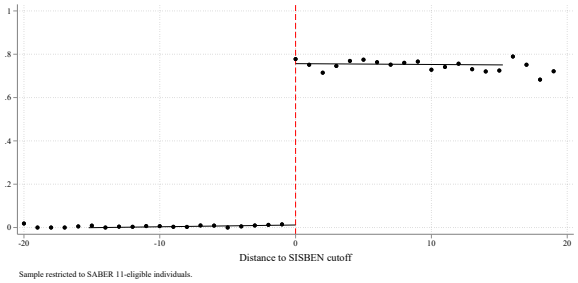
Notes: The figure plots the probability of receiving SPP or GE financial aid between 2014 and 2019 as a function of the distance to the SABER 11 eligibility cutoff for SISBEN-eligible students. Each marker shows the sample average within bins, and the fitted line is plotted using the optimal bandwidth (Cattaneo et al., 2014). Table A.5 reports the corresponding RD estimates and robust standard errors. The effect eligibility on aid receipt rises over time, except in 2018, when an unusually high test-score cutoff and uncertainty following the change in administration reduced take-up. To meet the target of 4,000 beneficiaries, the government extended awards to additional SISBEN-eligible students. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), and ICETEX.

Figure B.4: The Probability of Receiving Aid: SISBEN is the Running Variable

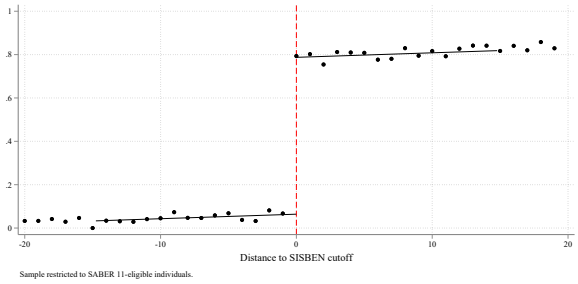
(a) 2014



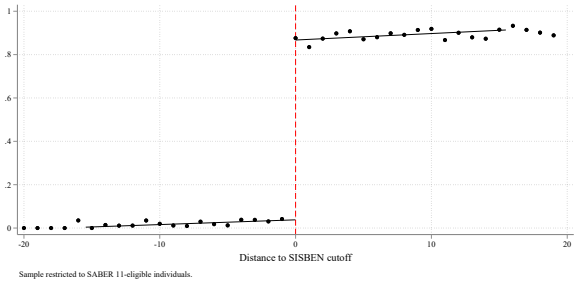
(b) 2015



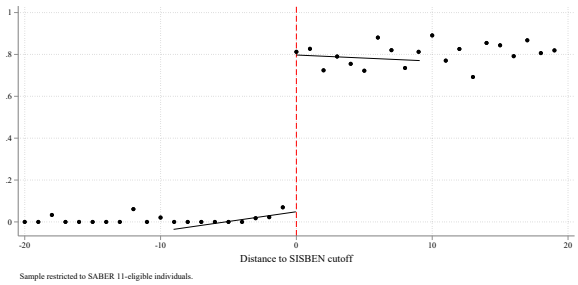
(c) 2016



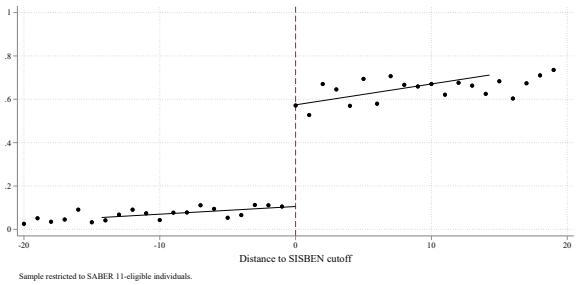
(d) 2017



(e) 2018



(f) 2019



Notes: The figures plot the probability of receiving SPP or GE financial aid programs between 2014 and 2019 as a function of the distance to the SISBEN eligibility cutoffs, restricting the sample to SABER 11-eligible students. The impact of eligibility on the probability of receiving financial aid increases over time. Sample average within bin. The line is plotted for the optimal bandwidth (Cattaneo et al., 2014). Table A.5 presents the RD estimates and robust standard errors. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), and ICETEX.

Table B.1: Covariate Balance Test around Policy's Test-Score Threshold

	2014		2015		2016		2017		2018		2019	
	RD Coeff. (1)	p-value (2)	RD Coeff. (3)	p-value (4)	RD Coeff. (5)	p-value (6)	RD Coeff. (7)	p-value (8)	RD Coeff. (9)	p-value (10)	RD Coeff. (11)	p-value (12)
Wealth percentile (including missing SISBEN)	-0.579	0.138	-0.534	0.198	0.334	0.522	0.636	0.213	0.478	0.876	0.747	0.406
Female	-0.015	0.217	0.009	0.372	0.004	0.786	0.007	0.707	-0.009	0.742	-0.027	0.232
Age	0.051	0.288	0.006	0.896	0.035	0.213	-0.016	0.428	0.014	0.644	-0.055	0.088
Ethnic minority	0.001	0.906	0.001	0.976	-0.001	0.701	0.005	0.368	-0.015	0.039	0.012	0.050
Employed	0.002	0.739	0.006	0.249	0.002	0.752	-0.016	0.175	-0.001	0.920	-0.007	0.707
Mother's education: primary	-0.010	0.278	-0.001	0.802	0.011	0.289	-0.005	0.794	-0.030	0.010	-0.001	0.811
Mother's education: secondary	-0.015	0.248	-0.009	0.635	-0.001	0.823	-0.002	0.761	0.022	0.254	0.031	0.109
Mother's education: T&T	0.002	0.745	0.005	0.598	-0.009	0.389	0.018	0.296	-0.014	0.476	-0.026	0.116
Mother's education: professional	0.021	0.007	0.008	0.327	-0.011	0.321	-0.014	0.327	0.033	0.198	-0.007	0.858
Father's education: primary	-0.003	0.745	0.004	0.748	0.027	0.056	-0.016	0.412	-0.025	0.129	0.009	0.538
Father's education: secondary	-0.004	0.558	-0.012	0.404	0.008	0.518	0.034	0.077	-0.019	0.463	-0.018	0.387
Father's education: T&T	0.003	0.583	0.002	0.925	-0.019	0.151	-0.044	0.005	0.001	0.903	-0.019	0.196
Father's education: professional	0.007	0.382	0.013	0.160	0.004	0.925	0.006	0.774	0.044	0.072	0.034	0.104
Household SES: Stratum 1	0.002	0.945	-0.011	0.249	0.006	0.600	0.012	0.386	-0.007	0.742	-0.018	0.443
Household SES: Stratum 2	-0.019	0.140	0.004	0.582	-0.001	0.893	-0.003	0.995	-0.022	0.413	0.012	0.534
Household SES: Stratum 3	0.012	0.238	0.008	0.400	0.005	0.636	-0.016	0.222	0.019	0.402	0.007	0.789
Household SES: Stratum 4	0.006	0.027	0.000	0.929	-0.005	0.217	0.001	0.852	0.010	0.266	0.010	0.220
Household SES: Stratum 5	0.001	0.524	-0.002	0.080	-0.001	0.508	0.004	0.111	0.001	0.776	-0.003	0.506
Household SES: Stratum 6	-0.001	0.204	-0.001	0.197	0.000	0.451	0.001	0.311	-0.002	0.231	-0.001	0.791
School hours: Full day	-0.004	0.702	-0.008	0.413	0.002	0.985	-0.004	0.863	0.028	0.179	0.010	0.640
School hours: Morning	0.000	0.955	0.000	0.817	-0.009	0.530	0.032	0.074	-0.037	0.155	-0.004	0.862
School hours: Evening	0.002	0.596	0.000	0.888	0.003	0.134	0.000	0.780	0.000	0.854	0.001	0.786
School hours: Afternoon	0.000	0.925	0.007	0.285	0.010	0.313	-0.027	0.038	0.011	0.431	-0.003	0.863
School hours: Weekends	0.003	0.444	0.000	0.667	-0.001	0.596	0.001	0.865	-0.001	0.642	-0.003	0.169
Private school	0.001	0.934	-0.002	0.621	0.013	0.298	0.016	0.369	0.018	0.514	0.005	0.811
Family has internet	0.019	0.144	0.012	0.236	0.004	0.915	-0.023	0.157	0.003	0.839	-0.004	0.856
Family has a laptop	0.004	0.771	0.008	0.312	-0.010	0.365	0.005	0.835	0.007	0.727	0.012	0.509
Family has a car	0.013	0.285	-0.007	0.456	0.008	0.603	0.033	0.040	0.029	0.248	0.012	0.390
School location: Urban	-0.006	0.540	-0.007	0.284	0.001	0.861	0.001	0.978	-0.011	0.300	0.005	0.583
Joint F-Stat (p-value, LB on bandwidth)	0.402		0.622		0.954		0.087		0.214		0.620	
Joint F-Stat (p-value, UB on bandwidth)	0.463		0.768		0.693		0.107		0.060		0.316	

Notes: This table plots the reduced-form RD coefficient where the outcome is a baseline characteristic and the running variable is the distance to the test-score cutoff. The sample is restricted to low-SES students. Odd columns present conventional coefficients, and even columns present p-values based on conventional standard errors. The last two rows report the p-value from a joint significance test using all baseline characteristics and small (± 20) or large (± 40) bandwidths. All results are estimated with package rdrobust (Cattaneo et al., 2014). *Sources:* Authors' calculations based on SABER 11 (Icfes) and SISBEN (DNP).

Table B.2: Covariate Balance Test around Policy's Poverty Thresholds

	2014		2015		2016		2017		2018		2019	
	RD Coeff.	p-value	RD Coeff.	p-value	RD Coeff.	p-value	RD Coeff.	p-value	RD Coeff.	p-value	RD Coeff.	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SABER 11 percentile	0.111	0.508	-0.060	0.588	0.092	0.322	-0.082	0.364	0.049	0.513	-0.075	0.281
Female	0.015	0.728	0.072	0.012	0.156	0.001	0.030	0.523	-0.013	0.714	-0.057	0.371
Age	0.001	0.922	0.034	0.547	0.023	0.741	0.063	0.312	-0.116	0.224	-0.035	0.810
Ethnic minority	0.017	0.051	-0.008	0.145	0.006	0.257	-0.001	0.897	0.014	0.458	-0.006	0.569
Employed	-0.008	0.349	-0.024	0.020	-0.002	0.893	-0.001	0.939	-0.031	0.554	0.013	0.792
Mother's education: primary	0.005	0.900	-0.024	0.210	-0.007	0.667	0.015	0.510	-0.025	0.653	-0.030	0.256
Mother's education: secondary	-0.059	0.038	0.028	0.295	-0.072	0.040	-0.060	0.152	0.044	0.652	0.085	0.070
Mother's education: T&T	-0.005	0.774	0.010	0.691	0.052	0.111	0.067	0.129	-0.061	0.356	-0.093	0.044
Mother's education: professional	0.056	0.005	-0.005	0.829	0.047	0.167	-0.015	0.807	0.045	0.505	0.032	0.437
Father's education: primary	0.010	0.907	0.004	0.903	0.014	0.713	-0.011	0.894	-0.018	0.696	0.030	0.716
Father's education: secondary	-0.061	0.018	-0.060	0.025	0.004	0.894	-0.021	0.545	-0.053	0.405	-0.024	0.678
Father's education: T&T	-0.006	0.597	0.015	0.359	-0.062	0.057	0.026	0.635	0.037	0.473	-0.024	0.599
Father's education: professional	0.057	0.015	0.036	0.086	0.044	0.187	0.001	0.792	0.033	0.628	0.061	0.216
Household SES: Stratum 1	-0.011	0.360	-0.008	0.674	0.002	0.983	0.004	0.836	-0.043	0.420	-0.002	0.936
Household SES: Stratum 2	0.000	0.926	0.045	0.189	0.020	0.805	0.029	0.685	0.077	0.336	0.073	0.199
Household SES: Stratum 3	0.006	0.788	-0.042	0.198	0.006	0.682	0.011	0.709	0.032	0.573	-0.080	0.198
Household SES: Stratum 4	0.010	0.276	0.008	0.418	-0.022	0.311	-0.038	0.172	-0.030	0.404	0.011	0.628
Household SES: Stratum 5	-0.002	0.565	-0.002	0.683	0.000	0.898	0.012	0.220	-0.016	0.374	0.006	0.633
Household SES: Stratum 6	0.000	0.745	0.000	0.351	0.001	0.850	-0.011	0.186	-0.016	0.262	-0.012	0.245
School hours: Full day	0.025	0.207	0.016	0.388	0.082	0.057	0.048	0.269	-0.016	0.970	0.089	0.093
School hours: Morning	-0.033	0.180	-0.004	0.702	-0.064	0.101	-0.025	0.469	0.015	0.936	-0.052	0.267
School hours: Evening	0.001	0.815	0.003	0.461	-0.001	0.667	-0.004	0.316	0.000	0.334	-0.005	0.416
School hours: Afternoon	0.016	0.342	-0.023	0.213	0.010	0.785	0.014	0.853	0.005	0.972	-0.028	0.437
School hours: Weekends	-0.007	0.041	0.000	0.955	0.003	0.276	0.000	0.192	0.000	0.000	0.000	0.589
Private school	0.058	0.012	-0.004	0.922	0.041	0.245	-0.024	0.866	-0.018	0.950	0.108	0.060
Family has internet	0.008	0.652	0.034	0.112	0.003	0.861	-0.023	0.381	0.067	0.157	0.055	0.191
Family has a laptop	0.031	0.038	0.001	0.825	-0.019	0.545	0.001	0.890	0.083	0.058	-0.016	0.700
Family has a car	0.060	0.014	-0.007	0.960	-0.052	0.167	0.007	0.833	0.045	0.397	-0.030	0.630
School location: Urban	-0.005	0.554	0.008	0.469	0.019	0.255	-0.008	0.652	0.074	0.098	0.038	0.059
Joint F-Stat (p-value, LB on bandwidth)	0.018		0.000		0.492		0.000		0.524		0.000	
Joint F-Stat (p-value, UB on bandwidth)	0.004		0.021		0.017		0.487		0.212		0.148	

Notes: This table plots the reduced-form RD coefficient where the outcome is a baseline characteristic and the running variable is the distance to the SISBEN cutoffs. The sample is restricted to SABER 11-eligible individuals. Odd columns present conventional coefficients, and even columns present p-values based on conventional standard errors. The last two rows report the p-value from a joint significance test using all baseline characteristics and small (± 7) or large (± 15) bandwidths. All results are estimated with package rdrobust (Cattaneo et al., 2014).
Sources: Authors' calculations based on SABER 11 (Icfes) and SISBEN (DNP).

Appendix C HTEs using Causal Random Forests

This section estimates the HTEs of financial aid on college enrollment, quality, graduation, and earnings using CRFs. We estimate separate CRFs for each outcome using the "causal_forest" function from the GRF package (Athey et al., 2019).

CRFs are nonparametric machine learning algorithms tailored to estimating CATEs. They extend the standard random forest framework to the causal setting by constructing trees that split the covariate space to maximize heterogeneity in subgroup ATEs. Formally, the CRF estimates the CATE, $\tau(x) = \mathbb{E}[Y(1) - Y(0) \mid X = x]$, where $Y(1)$ and $Y(0)$ denote potential outcomes under treatment and control, respectively, and X is the covariate vector. The algorithm recursively partitions the feature space to isolate subpopulations with differential responses to treatment, and averages across an ensemble of such trees to produce stabilized CATE predictions.

We estimate separate HTEs for ten outcomes: (1) immediate enrollment in any college; (2) immediate enrollment in an HQA college; (3) immediate enrollment in a private HQA college; (4) B.A. attainment from any college; (5) B.A. attainment from an HQA college; (6) B.A. attainment from a private HQA college; (7) earnings value-added; (8) earnings, including zeros for those not formally employed; (9) log earnings; and (10) formal employment. All labor-market outcomes are measured in 2023—nine years after high school for the 2014 cohort.

Sample Selection. Our analysis for the 2014 cohort uses the 2013 cohort of SABER 11 test takers as the pre-treatment sample. Specifically, the causal forest is estimated on (a) 2014 students who satisfy both the test-score and poverty cutoffs (eligibles), (b) 2013 students who would have satisfied those cutoffs had the policy existed (controls), and (c) 2014 students who miss one of the cutoffs within the optimal RD bandwidth (ineligibles), obtained via the “rdrobust" command (Cattaneo et al., 2014). The assumption that justifies pooling pre-policy eligibles (2013) with post-policy ineligibles (2014) is a “parallel counterfactual" or “no differential cohort trend" assumption conditional on covariates.

Covariates. The model includes a rich set of more than 30 baseline characteristics at the individual, household, and high school levels. These include: the student’s SABER 11 test score, sex, age, ethnic minority status, employment status at registration; household variables such as the SISBEN score and area, parental education (for both parents), socioeconomic stratum, family size, and indicators for home floor materials, internet access, laptop ownership, and car ownership; and school-level

variables including whether the high school is private, its class schedule, academic calendar, and urban location. Categorical covariates were one-hot encoded without an intercept, in line with best practices.

Estimation. Each CRF was trained using 3,000 trees to improve the stability and precision of the estimated treatment effects and their associated confidence intervals. We use “honest" estimation, which involves splitting data between training and estimation sets to avoid overfitting. Hyperparameters (e.g., sample fraction, minimum node size) were selected via the built-in cross-validation routine (tune.parameters = "all"). The out-of-bag (OOB) error metric serves as an internal cross-validation mechanism: each tree is trained on a bootstrap sample, and prediction error is computed using the holdout (out-of-bag) observations. Averaging these OOB errors across all trees yields a robust estimate of the model’s performance.

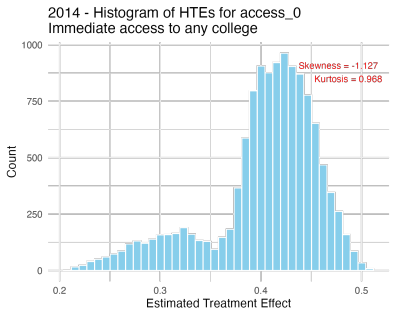
Results. Figure C.1 shows the distribution of estimated treatment effects for each outcome. For example, the treatment effects on enrollment in any college range from around 20 to 50 p.p. We do not find any evidence for negative treatment effects, so the policy is unlikely to have discouraged students from attending college. Figure C.2, which displays the point estimates (in blue) and associated 95% condidence intervals (in black), shows that these effects are precisely estimated. Estimates for all educational outcomes are precise, with narrow CIs. Estimates for labor-market outcomes exhibit slightly wider CIs but remain precise.

We next inspect which variables drive treatment effect heterogeneity. To do so, we report a simple variable-importance measure based on the causal forest used for estimating HTEs. This variable-importance measure counts how often each variable is used in a split within the causal forest, giving greater weight to splits occurring higher in the tree. Figure C.3 shows that the overall variable importance is high for all educational outcomes, with the test score being the strongest predictor of treatment effects. In contrast, for labor-market outcomes, overall variable importance is lower, and the poverty score can be more predictive of treatment effects than the test score.

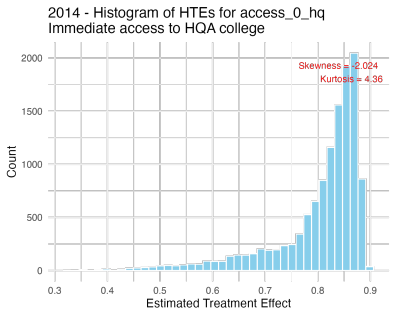
Figure C.4 displays 3D plots of HTEs by test score and poverty level. For educational outcomes, treatment effects decline sharply with higher test scores and remain relatively stable across poverty levels. In contrast, for labor-market outcomes, there is no clear linear pattern: treatment effects on employment appear to rise slightly with test scores, and effects on log earnings increase modestly with poverty, but both relationships are weak.

Figure C.1: Histograms of Estimated Treatment Effects across Outcomes, using Honest Estimates from the CRF Method: 2014 Cohort

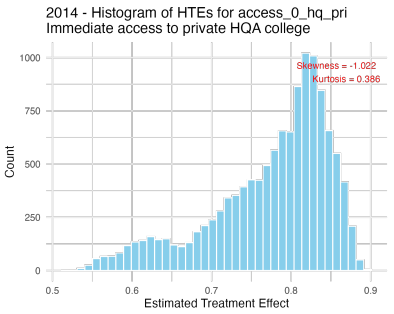
(a) Any enrollment



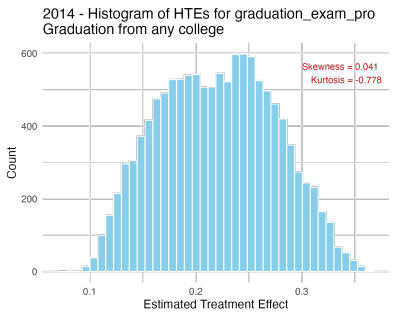
(b) HQA enrollment



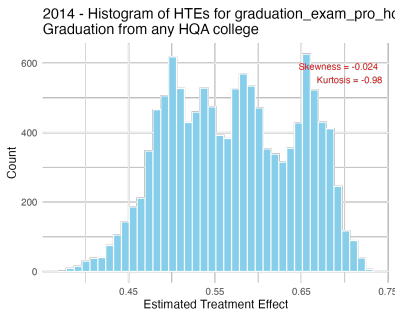
(c) Private HQA enrollment



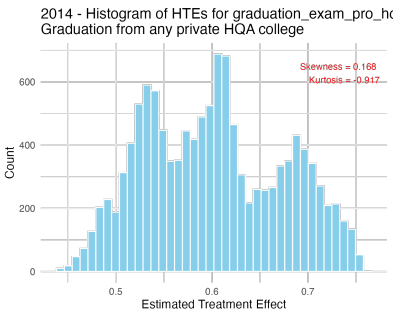
(d) Any B.A.



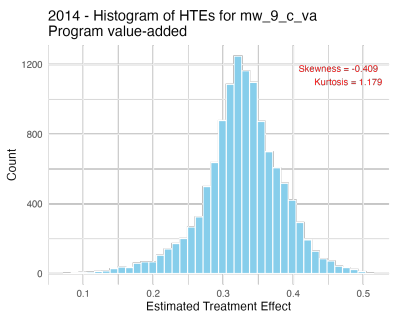
(e) HQA B.A.



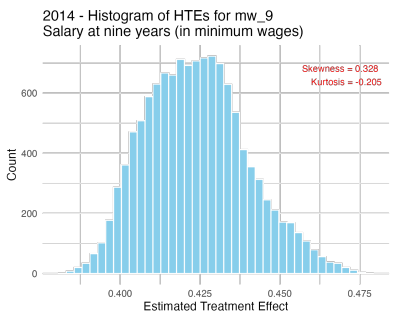
(f) Private HQA B.A.



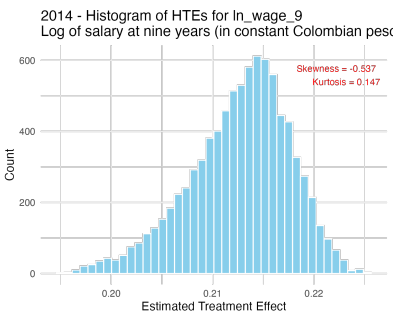
(g) Value-added



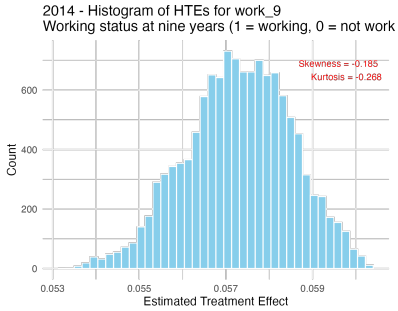
(h) Earnings (with zeros)



(i) Log Earnings



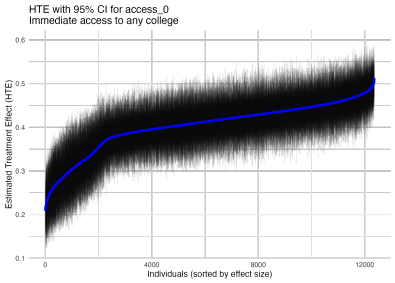
(j) Formal Employment



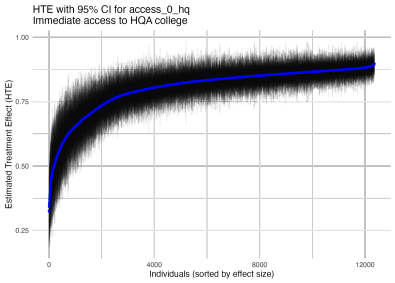
Notes: This figure shows the distribution of estimated treatment effects based on the 2014 cohort of test takers. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), SABER PRO (Icfes), ICETEX, and PILA (MinSalud).

Figure C.2: Confidence Intervals of Estimated Treatment Effects across Outcomes, using Honest Estimates from the CRF Method: 2014 Cohort

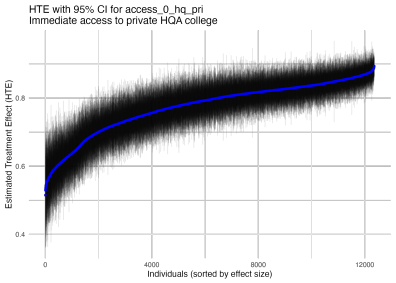
(a) Any enrollment



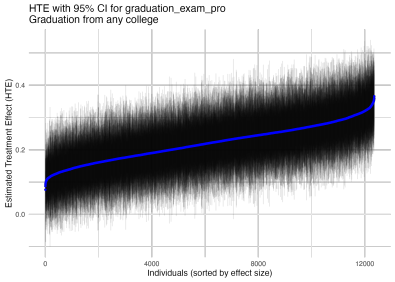
(b) HQA enrollment



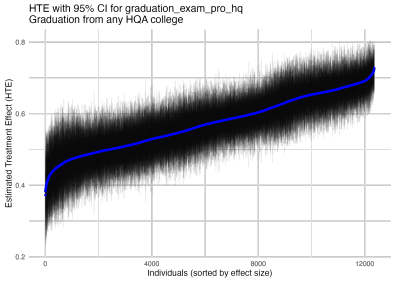
(c) Private HQA enrollment



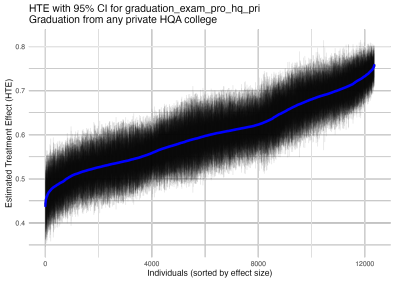
(d) Any B.A.



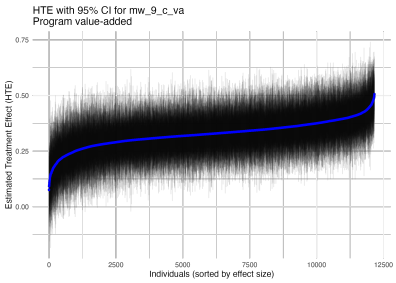
(e) HQA B.A.



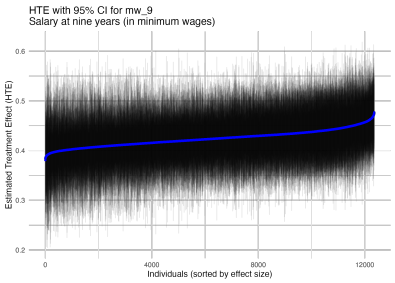
(f) Private HQA B.A.



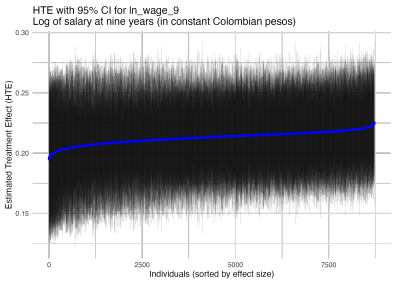
(g) Value-added



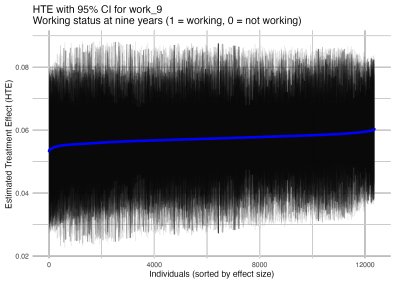
(h) Earnings (with zeros)



(i) Log Earnings



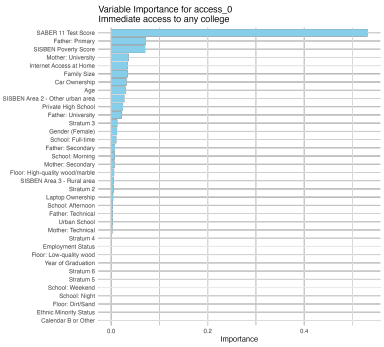
(j) Formal Employment



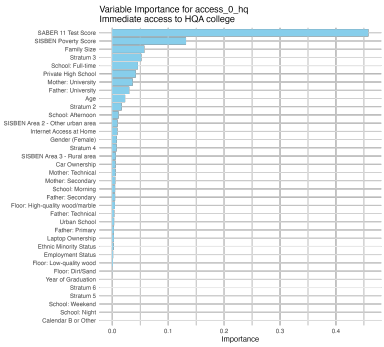
Notes: This figure shows the distribution of estimated treatment effects and associated 95% CIs based on the 2014 cohort of test takers. The blue line depicts the point estimates of the HTEs, while the black vertical lines denote the corresponding CIs. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), SABER PRO (Icfes), ICETEX, and PILA (MinSalud).

Figure C.3: Variable Importance Plots from the CRF Method: 2014 Cohort

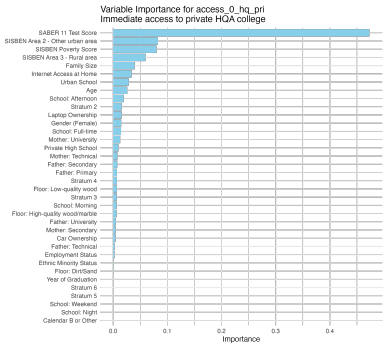
(a) Any enrollment



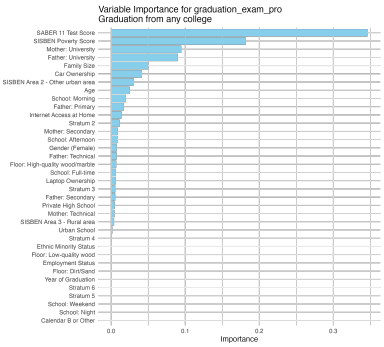
(b) HQA enrollment



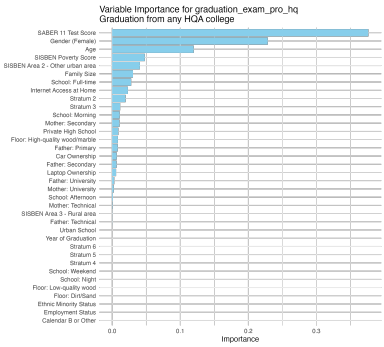
(c) Private HQA enrollment



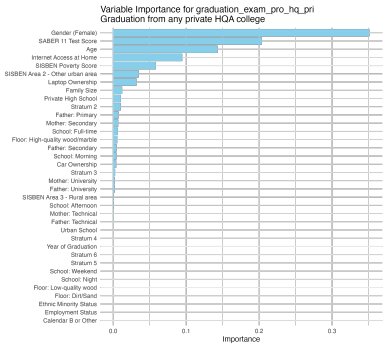
(d) Any B.A.



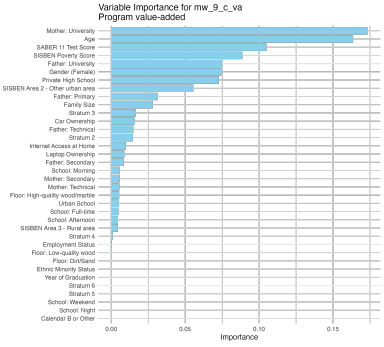
(e) HQA B.A.



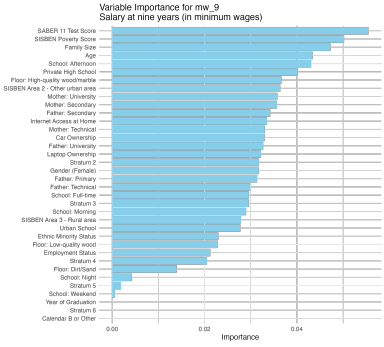
(f) Private HQA B.A.



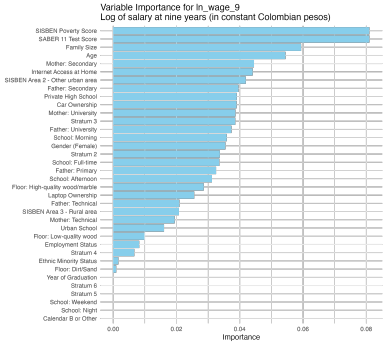
(g) Value-added



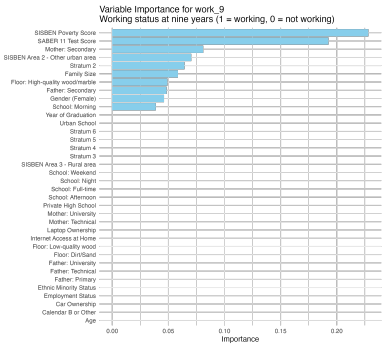
(h) Earnings (with zeros)



(i) Log Earnings



(j) Formal Employment



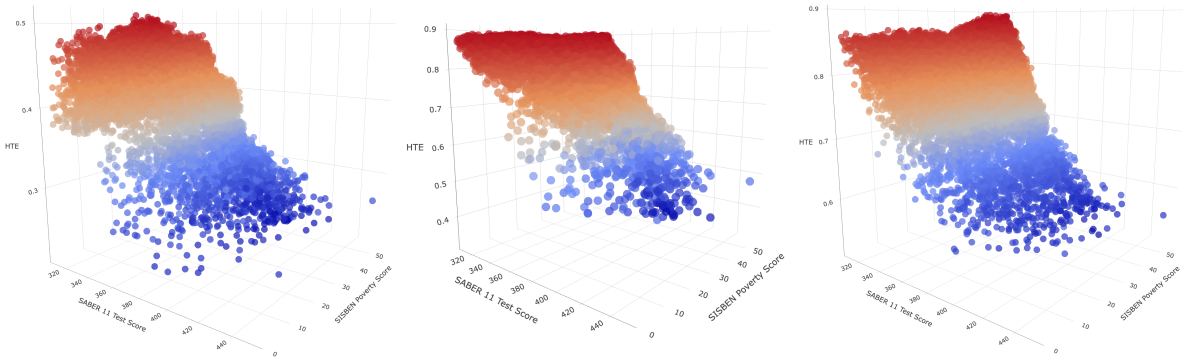
Notes: This figure reports the variable importance plot based on the 2014 cohort of test takers. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), SABER PRO (Icfes), ICETEX, and PILA (MinSalud).

Figure C.4: 3D HTE Plots by Test Score and Poverty: 2014 Cohort

(a) Any enrollment

(b) HQA enrollment

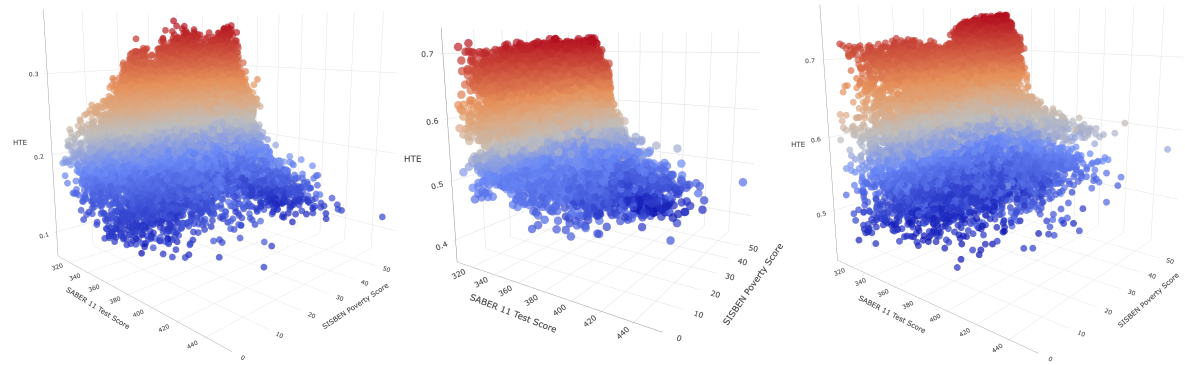
(c) Private HQA enrollment



(d) Any B.A.

(e) HQA B.A.

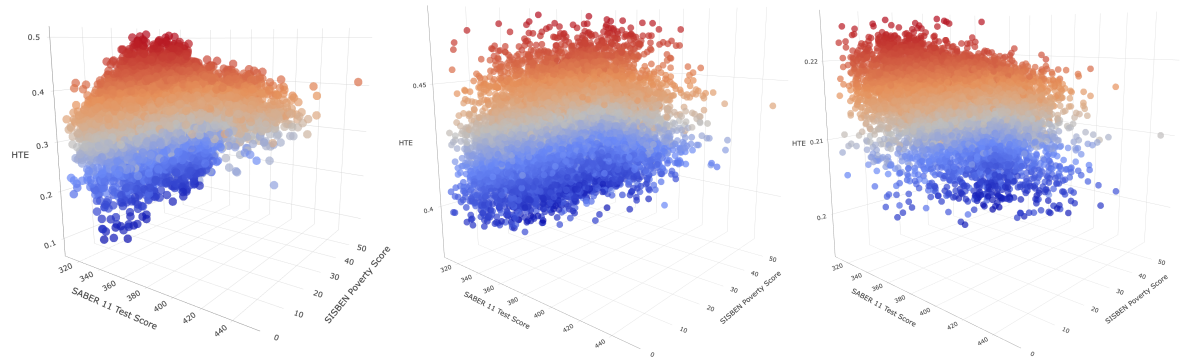
(f) Private HQA B.A.



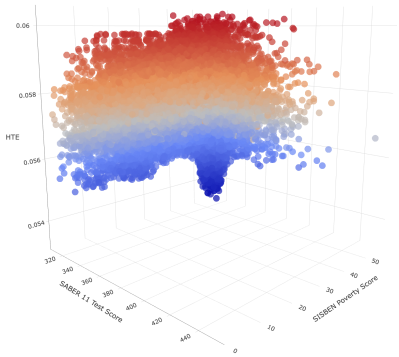
(g) Value-added

(h) Earnings (with zeros)

(i) Log Earnings



(j) Formal Employment



Notes: This figure provides 3D plots of the HTEs by test score and poverty based on the 2014 cohort of test takers. *Sources:* Authors' calculations based on SABER 11 (Icfes), SISBEN (DNP), SNIES (MEN), SABER PRO (Icfes), ICETEX, and PILA (MinSalud).