AVANTI FELLOWS · RESEARCH & DATA

Non-Linear Returns

The Dramatic Impact of Test Preparation on College Access: A Comparative Study in the Navodaya Vidyalaya Schools

Created 7 July 2026 · revised 2 November 2026 (see the version notes)

> **About this paper.** This is an internal analysis by Avanti Fellows, shared for transparency. It has **not been peer-reviewed**. It was prepared with substantial help from AI tools (Anthropic's Claude), which wrote much of the code and drafted much of the text under the direction of Avanti staff. The numbers are produced by code from the underlying data, and the sources and methods are documented so they can be checked, but errors are possible. Read the findings as preliminary, and contact Avanti Fellows before citing them.

Executive Summary

Avanti runs intensive on-campus coaching, its Centres of Excellence (CoE), inside a subset of India's Jawahar Navodaya Vidyalayas (JNVs), the government residential schools for rural students. A lighter-touch programme, Nodal, runs in other JNVs. The two arms draw from the same rural, residential, largely reserved-category population, but they are not balanced: CoE students walk in with visibly stronger Class-10 board marks. This note asks a simple question with that imbalance built in: students at a CoE score much higher on the JEE engineering entrance exam than students at a Nodal centre, but they also start ahead. So is the CoE gap real teaching, or just better students?

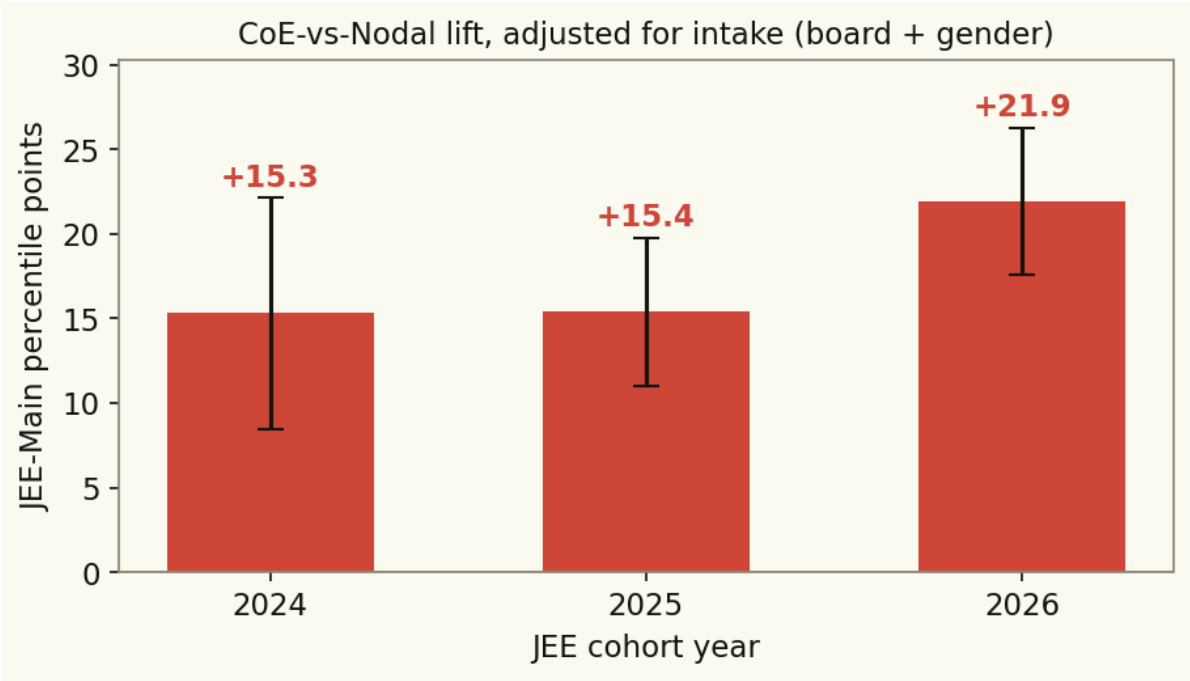
We measure the gap, then strip out the head start. In the two cohorts with reliable samples, 2025 and 2026, CoE students beat Nodal students by 25.7 and 34.4 percentile points on the raw average. Once we hold Class-10 board marks and gender fixed, **a lift of about 15.3 points (2025) and 20.5 points (2026) remains**. That figure is stable across *every* way we specify the intake control: a quadratic, an assumption-free board-decile design, and a common-support-only sample all land in the same place (§7). A simple *linear* intake control gives 15.4 and 21.9; in 2026 it under-adjusts where CoE students cluster near the board ceiling, so we read it as an upper bound. The lift is significant throughout (wild-cluster bootstrap $p \le 0.001$; n = 2,386 students with both a board mark and a JEE result), and **pooling the two years gives about +17.7 points** (linear +18.4 [15.3, 21.6]). The adjustment for intake removes about two-fifths of the raw gap (43%-47%); the majority is left over, large and stable as controls pile on. An earlier cohort, 2024, points the same way, but one Dakshana centre holds 64% of its 300 CoE students, so we report it as suggestive and do not lean on it.

Why the title says "non-linear returns." A percentile point is not a small thing. About 10 lakh students sit JEE Main, so a 15.3-percentile-point difference is roughly **1.5 lakh rank positions**, and a 20.5-point difference about 2.0 lakh. At the bottom of the range this is the difference between *not clearing the exam at all* and clearing it; higher up it is the difference between a rank that opens nothing and a rank that opens an **NIT or IIT** seat. The returns are non-linear in a second sense too: the same coaching yields a larger percentile jump for a mid-range student than for one already near the ceiling (Section 7). And the return does not stop at admission: the seat a student wins compounds into lifetime earnings. Avanti's alumni analysis puts the income and social return on exactly this movement in the tens of times the cost per student (see the Alumni RoI analysis); we cite it for scale, not as part of this paper's causal chain.

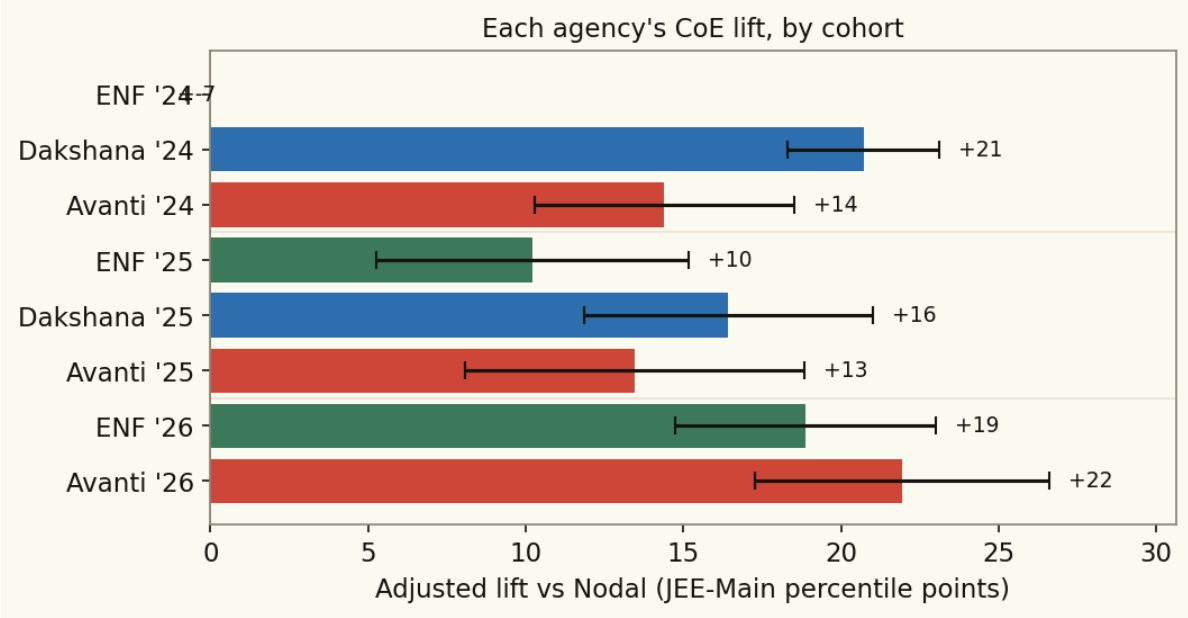
This is an observational comparison, not a randomised trial, so we write "is associated with", not "causes". CoE students are selected, and we can only adjust for the selection we observe. The residential setting holds family circumstances roughly fixed, the controls handle the one visible intake gap (board marks; the groups are comparable on category once a data-collection gap is corrected), and the result barely moves as controls pile on. A standard sensitivity bound (Oster) is

supportive but not decisive: the result survives the author's preferred assumption and weakens under the most conservative one. The honest reading: the CoE percentile advantage is much larger than the measured intake difference can account for, and the most plausible remaining objection, that CoE students are more motivated in ways we do not measure, is itself the behaviour an intensive programme is meant to produce.

The comparison is CoE against Avanti's lighter Nodal programme, so the lift measures what intensity adds, not the whole effect of a centre. Section 8 adds the step below it. There is no untreated group in the Navodaya system to compare with, since in 2024 every JEE taker outside a centre had online support, so the comparison is with Avanti's own online students. In 2025 a Nodal student was 23.7 points, and a CoE student 70.8 points, more likely to qualify in JEE Main than an Avanti online student with the same Class-10 marks. Online results are incompletely recorded in 2025, and the 2024 CoE arm rests on few centres, so we treat this as a supporting comparison (§8).



The lift is not one provider's doing. Each operating agency is compared with Nodal, pooled over the years it ran with a cohort fixed effect, so no agency is compared across exam years. Avanti's CoEs lift JEE percentile by +17.6 points and Dakshana's by +18.4, statistically indistinguishable ($p = 0.27$). ENF's come out lower at +10.5 ($p = 0.04$ against Avanti), on fewer students and centres, though in 2026 ENF (+18.9) and Avanti (+21.9) are close (§7). And the benefit reaches the disadvantaged rather than a creamy layer: the adjusted lift among reserved-category students (+12.7 in 2025, +22.6 in 2026) is at least as large as among non-reserved students. As a descriptive backdrop, top-tier seat rates (IIT, NIT/IIIT, medical, top engineering) are far higher at CoEs than in the online programmes, though that comparison is unadjusted and on a different population, so we keep it out of the evidentiary chain.



Avanti and Dakshana clear the same double-digit bar against Nodal in every year they ran, and ENF does in 2026: the intensive model, not one provider, is what moves the number. The sharpest sign of this: three centres Dakshana ran in 2025 that Avanti took over in 2026 (Bundi, Kottayam, Lucknow) held their near-ceiling clearance and percentiles straight through the operator change (Section 7.1). A note of caution the raw funnel makes explicit (Section 7.1): raw pass rates differ between the agencies in step with the Class-10 intake each started from, so raw rates alone cannot rank them; the *adjusted* lift is the like-for-like yardstick.

How sure are we: this is the strongest *observational* statement the data supports, and we have corrected it for the few-school inference problem and the most lenient sensitivity tuning. It is not proof of cause. A phased or randomised CoE rollout would settle it; until then, the size of the lift, its stability under controls, and its consistency across agencies and across reserved and non-reserved students make pure selection an unlikely full explanation.

Version note (2 November 2026). Three corrections to the student record, each found while asking how CoE and Nodal compare with the rest of the Navodaya cohort. 1. **The CoE arm was missing students.** Dakshana's 2024 centre and every ENF centre were filed outside it, because no attribution table covered them. The NTA results file's own programme tag now places them: the 2024 CoE arm grows from 71 to 300, and 2025's from 308 to 401. 2. **Retakes are excluded in every programme.** A student who passed Class 12 before the JEE year is out of the sample, whatever centre tags them. 3. **Standard errors cluster on the centre.** A student's cluster is the school where they sat the Class-12 board. Before, students with no school on record shared a single cluster.

The headline moves from 15.6 to 15.3 (2025), stays at 20.5 (2026), and moves from 17.9 to 17.7 pooled. The prior text is frozen in `analysis/avanti-impact-paper/v3/`. The same revision adds Section 8, which compares both centre arms with Avanti's own online students in 2024 and 2025; it replaces a short companion note published earlier that day.

Version note (7 October 2026). The first version of this paper reported 14.5 (2025) and 22.3 (2026) under the flexible intake controls, and 16.5 and 27.1 under a linear one. Since then the student record changed in ways that touch this sample: a student who sat JEE twice now appears in two cohorts, so the second, uncoached attempt of a 2025 student sat in the 2026 sample, mostly on the

Nodal side; and board-mark links and JEE qualification flags were corrected. This version counts each student once, at their **first attempt**, the rule the JEE forecast model (Model B) adopted on the same day. The headline moves to 15.3 and 20.5, the direction and size of the finding do not change, and the prior text is frozen in the analysis repository (`analysis/avanti-impact-paper/v2/`).

Non-Linear Returns: The Dramatic Impact of Test Preparation on College Access

1. Abstract

We examine whether Avanti's intensive on-campus programme, the Centre of Excellence (CoE), is associated with higher performance on the JEE Main engineering entrance examination, relative to Avanti's lighter-touch Nodal programme, among students in India's Jawahar Navodaya Vidyalaya residential schools. Linking Class-10 board records to JEE results through a person-level identity crosswalk, we study 2,386 students across the 2024, 2025 and 2026 cohorts who have both a board mark and a JEE score, each at their first attempt. CoE students outscore Nodal students by 25.7 to 34.4 JEE percentile points unadjusted in the two reliable cohorts. Controlling for Class-10 board marks and gender, with standard errors clustered by school, an adjusted lift of about 15.3 points (2025) and 20.5 points (2026) remains under a flexible intake control, and is stable across a ladder of intake controls (quadratic, board-decile fixed effects, and a common-support-only sample; §7); a linear control returns 15.4 / 21.9, which we treat as an upper bound. Because treatment is nested in school and the treated clusters are few, we confirm significance with a restricted wild-cluster bootstrap ($p \leq 0.001$ in both years) and report a pooled two-year estimate of about +17.7 points (linear +18.4 [15.3, 21.6]). A third cohort (2024) agrees in sign, but one centre holds 64% of its CoE students and its bootstrap interval is [1.6, 29.4], so we treat it as suggestive only. Intake differences explain about two-fifths of the raw gap; the coefficient is stable as controls are added. An Oster sensitivity bound supports the result under the author's baseline assumption but weakens under a stricter one, so we call it supportive, not decisive. The lift appears separately for the Avanti, Dakshana and ENF agencies, and is at least as large for reserved-category students as for non-reserved ones. Against Avanti's own online students, who were at no centre, Nodal students were 23.7 and CoE students 70.8 percentage points more likely to qualify in JEE Main in 2025, a comparison limited by incomplete recording of online results. We do not claim causality: CoE students are selected on characteristics we cannot fully observe. The size and stability of the adjusted lift make pure selection an unlikely complete explanation, and the residual threat, selection on motivation, is the margin an intensive programme is designed to move.

2. Introduction

India admits a few hundred thousand students each year to its selective engineering and medical colleges through the JEE and NEET examinations, and the competition rewards sustained, structured preparation that most rural students cannot access. The Jawahar Navodaya Vidyalayas (JNVs) are government residential schools built to carry talented rural children toward higher education; Avanti operates inside them. In some JNVs Avanti runs a Centre of Excellence (CoE): intensive, on-campus, teacher-led coaching for the exam. In others it runs a Nodal programme: lighter, less intensive support. Both sit inside the same residential system and draw the same rural, largely reserved-category, first-generation population.

This gives a natural comparison. If CoE students do better on the JEE than Nodal students, how much of that is the programme and how much is the students Avanti happens to place in a CoE? The question matters because the headline gap is large and the temptation to read it as pure impact is

strong, while the equal and opposite temptation, to dismiss it as selection, is just as available. We do neither. We measure the gap, decompose it into the part explained by measured intake and the part that is not, and test how fragile the residual is.

The result, previewed: the raw CoE-vs-Nodal gap is 25.7 to 34.4 percentile points in the two reliable cohorts; adjusting for Class-10 board marks and gender leaves about 15.3 (2025) and 20.5 (2026) points, about +17.7 pooled; that residual barely moves as controls are added. We do not claim a causal effect, because assignment to CoE is not random. We claim that the lift is far larger than the observable intake difference, that it survives the controls available and cluster-correct inference, and that the path to a causal estimate is a phased rollout, which we describe.

This note is a replication and consolidation: the regression follows the earlier `11_avanti_impact_regression` analysis, rebuilt on a single validated student record so every number traces to one source. A hypothesised mechanism, why intensive preparation might move the JEE, is documented separately in the companion practice-and-qualification study (the nvs-lakshya paper), which shows on a related (non-CoE) JNV population that doing full-length practice tests predicts qualification. We do not measure practice volume on this paper's sample, so that mechanism is a plausible explanation we import, not one we test here.

3. Setting

A JNV is a co-educational residential school, two students per district on average, selected by an entrance test at Grade 6 and serving predominantly rural, reserved-category households. Because students board, their home environment during the exam-preparation years is largely the school's, not the family's. This is the design lever the paper leans on: comparing two programmes inside the residential system holds many household confounders, income, parental schooling, local infrastructure, roughly fixed by construction, in a way a comparison across day schools could not.

Within that system, Avanti's **CoE** is the intensive arm: a dedicated on-campus coaching set-up with teachers, a structured JEE/NEET curriculum, and regular full-length testing. The **Nodal** programme is the lighter arm, less contact time and less intensive preparation. A separate, much larger online programme (PMU) reaches students with broadcast and self-paced material; we use it only as an outcome benchmark in Section 7, not in the regression. The CoE label spans three operating agencies, Avanti, Dakshana and ENF; Section 7 splits them.

The outcome is the **JEE Main percentile**, the candidate's national percentile rank (0 to 100), not a raw mark. A percentile of 88 means the student scored above 88% of all test-takers nationally.

4. Data

4.1 Sources and linkage

Three record systems are joined: Class-10 CBSE board results (the intake ability measure), the NTA JEE Main results (the outcome), and Avanti's enrolment table (the programme and demographics). These systems share no common key. We link them through a person-level identity crosswalk that resolves the same student across all three by exact id where one exists and by name-plus-date-of-birth matching otherwise, assigning each student one stable `person_id` (documented in `analysis/identity_crosswalk/`). Programme assignment (CoE vs Nodal) and the canonical bucket

come from the committed classification (`analysis/program-classification/`); both, plus the JEE outcome and board marks, are assembled in the golden student record (`analysis/golden-student-record/`), which this paper reads.

4.2 Sample construction

The estimating sample is every student who (a) is in a JNV CoE or Nodal programme, (b) sat JEE Main in 2024, 2025 or 2026, (c) has both a Class-10 board mark and a JEE percentile (the regression needs both), and (d) is on their **first** JEE attempt. The golden record keeps one row per exam attempt, so a student who sat JEE in 2025 and again in 2026 has a second row in 2026; that retake is uncoached and typically scores higher, and it would otherwise sit in the 2026 sample as if it were a programme outcome. We drop it, using the same rule as the JEE forecast model (`repeat_attempts` in `../jee-main-prediction/model.py`). The same rule drops any attempt by a student who passed Class 12 before the JEE year (NTA's year of passing), whichever centre tags them: 88 of the 122 ENF centre students tagged in 2024 are such retakes. The board-mark requirement is the binding restriction: board linkage is incomplete in the earlier cohorts. Requiring a board mark may select on documentation as well as shrink the sample, a concern strongest in 2024; we report results per cohort rather than pooled so the 2024 cell is visible rather than hidden inside an average.

The binding number is treated clusters, not total clusters. The econometrics literature is explicit on this: when only a few groups are treated, standard inference is inappropriate even if the total number of groups is large (Conley & Taber 2011), and adding more control clusters does not fix it (MacKinnon, Nielsen & Webb 2023). Because a JNV is either a CoE school or a Nodal school, treatment is nested in school: clustering standard errors by school means the effective sample for inference is the number of *treated* (CoE) schools. A student's school here is the one where they sat the Class-12 board, which for a CoE student is the centre itself; 2026 has no Class-12 board record yet, so it uses the school on the student record. That gives **10 treated schools in 2024**, 15 in 2025 and 16 in 2026. Cluster-robust standard errors are unreliable below roughly 30 clusters, so for every cohort we report a restricted wild-cluster bootstrap (Section 6) as the primary inference. Two cautions from that literature apply. With very few treated clusters the restricted wild bootstrap tends to *under*-reject, so its p-values are conservative (MacKinnon & Webb 2018; MacKinnon, Nielsen & Webb 2023); in simulations it works well only from about seven treated clusters (MacKinnon & Webb 2017). And it is reliable only when clusters are of similar size (MacKinnon & Webb 2018). 2024 fails that test badly: one Dakshana centre holds 64% of its CoE students. So we treat the 2024 cohort as suggestive rather than as evidence, although it now has more treated schools. Randomization inference, which those authors recommend as a further check, is not done here.

Table 1. Samples ledger. One row per cohort; the estimand is the within-cohort CoE-vs-Nodal contrast. Comparisons are regression-adjusted (not matched or weighted); standard errors are clustered by school, with the wild-cluster bootstrap as primary because treated clusters are few.

JEE cohort	CoE students	Nodal students	schools (clusters)	CoE (treated) clusters
2024	300	465	31	**10**
2025	401	388	33	**15**
2026	507	325	27	**16**
Total	**1208**	**1178**	(overlap across cohorts)	

4.3 Variable definitions

- `jee_pct`: JEE Main national percentile (0 to 100), the outcome.
- `treat`: 1 if the student is in any CoE agency (Avanti / Dakshana / ENF), 0 if Nodal.
- `board`: the mean of the Class-10 mathematics and science percentages, the intake-ability control. Both subjects enter the regression separately (`math10`, `sci10`).
- `female`: 1 if female.
- `reserved`: 1 if the student's category is SC, ST, OBC or EWS, 0 if general, missing where category was not collected. One agency (Dakshana) supplied no category, so its `reserved` is missing, not 0; treating the gap as 0 would mis-read missing data as "general-category". A composition descriptor, not a control.

4.4 Balance: the selection we can see

Table 2. Balance, CoE vs Nodal. The reader's first identification test: do the groups differ on intake? On one axis they clearly do. CoE students walk in with markedly higher Class-10 board marks in every cohort, so part of the raw outcome gap is that head start, which Section 5 strips out. On category the groups are comparable: the reserved-category share, read on the data that was actually collected (Avanti-CoE only, since Dakshana supplied no category), is 83% for CoE against 82% for Nodal in 2025, and 84% against 89% in 2026. Gender is not balanced and moves from year to year (Section 9), which is why it is in the regression.

	2024 CoE	2024 Nodal	2025 CoE	2025 Nodal	2026 CoE	2026 Nodal
N	300	465	401	388	507	325
Board (math+sci)/2	92.6	85.6	90.8	82.8	91.9	82.0
% female	19.0	38.1	38.9	40.7	28.6	42.2
% reserved (collected only)	89.8	85.6	74.8	82.2	83.6	88.6
Mean JEE percentile	89.9	62.8	90.0	64.3	88.3	53.9
Raw gap (CoE − Nodal)	**+27.1**		**+25.7**		**+34.4**	

So the selection that matters most is on board marks (7.0 to 9.9 points higher for CoE). That is exactly what Section 5 adjusts for. An earlier draft of this analysis recorded a large CoE-vs-Nodal reserved-category gap, but that was a data artefact: Dakshana's missing category had been coerced to zero. With the missing data marked missing, category is balanced, and adding reserved as a control moves the lift by at most 0.10 points (Section 7), confirming it is not a hidden confounder. If adjusting for the one real imbalance, board marks, removed the gap, the CoE story would be intake. It does not.

5. Empirical strategy

For each cohort we estimate, by ordinary least squares,

```
jee_pct_i = β·treat_i + γ·math10_i + δ·sci10_i + ζ·female_i + ε_i
```

with standard errors clustered by school. The coefficient of interest is β, the CoE-vs-Nodal difference in JEE percentile after holding Class-10 board marks and gender fixed.

Identification, stated honestly. Assignment to CoE is not random, so β is not a causal effect in general. Two design features narrow the threat. First, both arms sit inside the JNV residential system, so household and local confounders that would dominate a cross-school comparison are roughly common to both groups by construction. Second, the controls absorb the one large, measured intake difference, prior academic ability (board marks); on category the groups are already comparable (Table 2), and gender is controlled. What we cannot do here is add school fixed effects: a CoE and a Nodal are different schools, so school effects would absorb the very contrast we want. We therefore cluster on school for inference but identify β off the cross-school CoE-vs-Nodal difference, conditional on intake.

Why Nodal is the control. β is the effect of *intensity*, what a CoE adds over a lighter centre, and it understates the whole effect of a centre by whatever Nodal itself adds. The natural alternative is not "no programme": the 2024 NTA file tags every Navodaya JEE taker outside a centre as an online student (`JNV ENABLE`, 6,665 first-attempt takers). Nodal stays the control for three reasons. It exists in every cohort. Both arms sit the exam at near-universal rates, while only some online students do. And Avanti records both arms' results, while most other JNV results in 2025 cannot be matched to a student. Section 8 compares both centre arms with Avanti's own online students in 2024 and 2025.

The named residual threat is selection on unobservables, most plausibly student motivation or effort that is correlated with being placed in (or choosing) a CoE and not captured by board marks. We cannot rule it out. We do three things against it: report the nested ladder (does β survive adding controls?), decompose the gap (how much is intake?), and compute the Oster bound (how strong would the unobservable have to be?).

6. Results

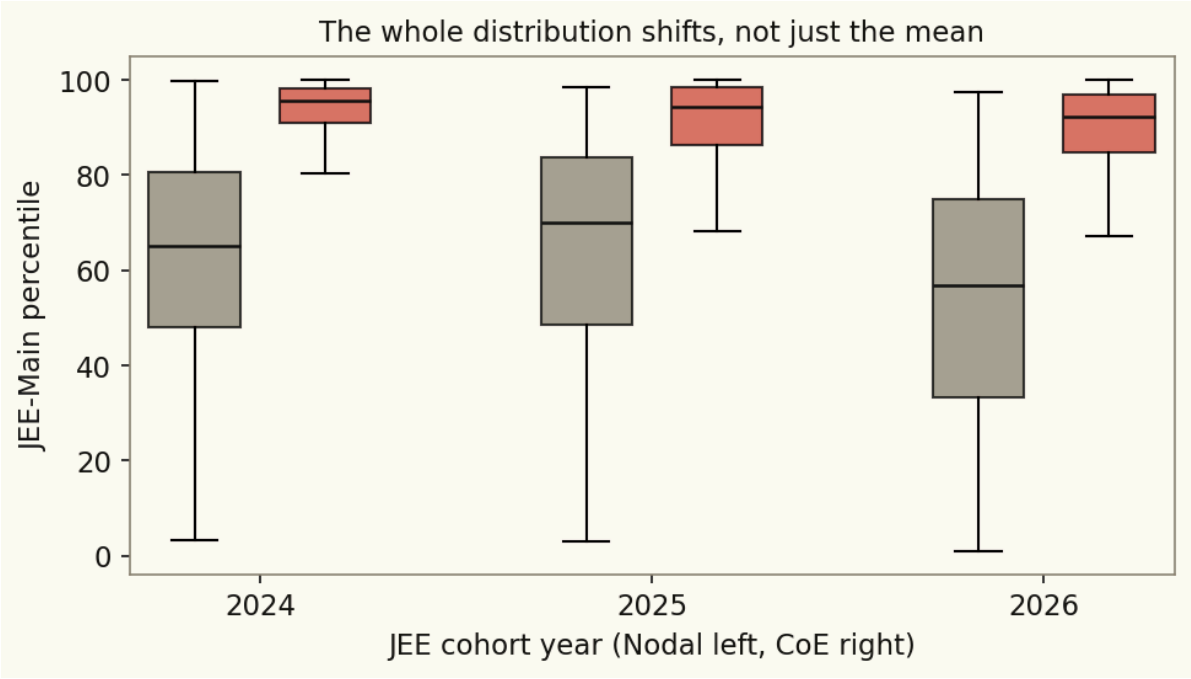
Table 3. The nested ladder: CoE-vs-Nodal coefficient on `jee_pct`. Each cell is β [95% CI]; M1 is the raw gap, M2 adds the two board marks, M3 adds gender. The 95% CIs shown are the asymptotic cluster-robust ones, but the *primary inference is the restricted wild-cluster bootstrap* (rightmost column), because treatment is nested in few schools.

Cohort	M1 raw	M2 + board	M3 + board + gender (preferred)	Wild-cluster bootstrap (M3)
2024	+27.1 [20.0, 34.2]	+17.0 [8.2, 25.8]	**+15.3 [8.5, 22.1]**	**p = 0.033, CI [1.6, 29.4]**
2025	+25.7 [20.2, 31.1]	+15.2 [10.2, 20.2]	**+15.4 [11.0, 19.8]**	**p < 0.001, CI [10.5, 20.4]**
2026	+34.4 [29.9, 39.0]	+22.5 [18.0, 27.1]	**+21.9 [17.6, 26.3]**	**p < 0.001, CI [17.0, 26.8]**

Read the preferred (M3) estimate in words: in 2026, a CoE student scored on average 21.9 JEE percentile points higher than a Nodal student with the same Class-10 board marks and gender, and the wild-cluster bootstrap puts that in [17.0, 26.8]. The 2025 lift is 15.4 points [10.5, 20.4]. The 2024 cohort points the same way (+15.3) and clears the bootstrap at p = 0.033, but its interval [1.6, 29.4] is very wide because one centre holds 64% of its CoE students; we report it for completeness and do not rely on it. Pooling the two reliable cohorts with a cohort fixed effect gives **+18.4 points [15.3, 21.6]** across 56 clusters, the headline that does not lean on the thin year. These M3 figures use a *linear* board control; §7 shows a fuller, less-parametric intake control (board-decile fixed effects, common-

support trim) puts them at about **+15.3 (2025), +20.5 (2026) and +17.7 pooled**, which we adopt as the paper's defensible headline; the linear estimates here are the upper bound.

The β-stability argument. Adding the board controls pulls β down from the raw gap (the head start was real), but adding gender on top moves it very little: 15.2 to 15.4 in 2025, 22.5 to 21.9 in 2026. A coefficient that stops moving once the main confounder is in is a sign of limited remaining confounding from the observed variables. (Section 7 also shows the lift survives a more flexible, non-linear board control, a sterner test than adding gender.)



The shift is distributional, not driven by a few top scorers: the CoE box sits above the Nodal box across the whole range in every cohort.

7. Robustness

Decomposition (Oaxaca). Splitting the raw gap into the part explained by intake (board + gender, valued at the Nodal return) and the residual:

Cohort	Raw gap	Explained by intake	Unexplained (lift)	% unexplained
2024	27.1	11.6	15.6	57%
2025	25.7	12.0	13.7	53%
2026	34.4	14.7	19.7	57%

Measured intake accounts for about two-fifths of the gap (43%-47%). The rest, 53%-57%, is unexplained by the observables. The unexplained part is what a causal effect would have to live in, alongside any unobserved selection. Under ignorability the unexplained component can be read as an average effect on the treated, but the individual coefficients on board marks and gender are not causal (Fortin, Lemieux & Firpo 2011).

Selection-on-unobservables bound (Oster 2019; Altonji, Elder and Taber 2005). The Oster δ* asks how strong selection on unobservables would have to be, relative to selection on the

observed controls, to drive β to zero; δ* > 1 is the usual robustness threshold. δ* depends on an assumed maximum R-squared, Rmax, that a fully-controlled regression could reach, and it is sensitive to that choice, so we report it across the standard grid rather than at one favourable point.

Cohort	δ* at Rmax = 1.3·R1 (Oster's baseline)	δ* at Rmax = 2·R1 (conservative)	β* at δ = 1, Rmax = 2·R1
2025	2.37	1.00	+0.0
2026	1.64	0.78	−6.1
2024	1.94	0.71	−6.2

Read honestly: under Oster's baseline Rmax = 1.3·R1 the bound clears 1 for all cohorts (the lowest is 1.64), so the result is robust on the author's recommended assumption. Under the more conservative Rmax = 2·R1 it sits around 1 (1.00 in 2025, 0.78 in 2026). We therefore call the Oster evidence **supportive but not decisive**: it does not overturn the lift, but it does not place it beyond reach of an unusually strong unobservable either. Oster's own recommended Rmax of 1.3 times the controlled R-squared is calibrated so that 90% of randomized results in top journals survive it, and she treats Rmax = 1 as too strict (Oster 2019). A further limit: this kind of bound can reach the wrong conclusion when the included controls are themselves correlated with the omitted factors (Diegert, Masten & Poirier 2026), and Class-10 marks are almost certainly correlated with unmeasured ability and motivation. (We disclose this rather than report only the favourable 1.3·R1 figure. JEE percentile is a noisy single-day outcome, which is an argument for a lower Rmax, but we do not lean on that.)

Intake-control ladder (robustness to the *shape* of the control). Board marks are bounded at 100 and CoE students cluster near the ceiling (Table 2), so a *linear* board control can under-adjust intake exactly where the CoE mass sits, leaving part of the head start in the coefficient. We climb a ladder of progressively less-parametric intake controls and watch whether the lift moves: linear → quadratic → **board-decile fixed effects** (ten intake bins, no functional form imposed) → **common-support trim** (dropping the non-overlapping board range so the lift is not identified off extrapolation across the CoE–Nodal intake gap). A head-start artefact should fall as the control tightens; a real effect should stabilise.

Cohort	linear	quadratic	decile-FE	common-support
2025	+15.4	+15.4	+15.2	+15.4
2026	+21.9	+20.6	+20.7	+20.1
pooled	+18.4	+17.8	+17.8	+17.6

It stabilises. In 2025 every rung lands at 15.2-15.4: the linear control was already adequate. In 2026 the richer controls trim the linear figure by about 1.4 points, to 20.1-20.7, and then do *not* fall further as the control moves from a quadratic to an assumption-free decile design to a common-support-only sample. The lift is **not an artefact of the functional form of the intake control, nor of extrapolating across the board gap between the groups.** We read the richer-control figures (their mean: 15.3, 20.5, 17.7 pooled) as the defensible estimate and the linear estimate as an upper bound. (All rungs cluster-robust by school, p < 1e-10 throughout.)

Bounded outcome, and why the lift is not a constant. The percentile outcome is bounded at 100, so an additive OLS coefficient cannot literally apply to every student: a linear +15.4 on a student already at the 90th percentile would imply an impossible 105. Three checks show this does not undo

the result, but it does reshape how to read it. First, the misspecification is small in practice: refitting the model as a **fractional logit** (Papke & Wooldridge 1996: a bounded [0,1] specification that *cannot* predict past the ceiling) gives an average marginal effect of **+16.4 points (2025) and +21.8 points (2026)**, essentially the OLS estimate, from the model that respects the boundary. In-sample, fitted OLS values exceed 100 for 6.5% of students in 2025 and 0.0% in 2026. Second, the effect is genuinely **heterogeneous, and attenuates at high baseline** exactly as a ceiling forces: the raw CoE-Nodal gap is +22.8 (2025) / +30.2 (2026) among the lowest board quartile but compresses to +7.7 / +10.2 among the highest (the treat×board interaction is negative and significant in both years, $p \le 0.001$). So the headline is an **average** treatment effect, not a per-student constant: a low-starting student gains far more than the average, a student already near the ceiling gains less. Third, and this is the reading that matters for expansion: because high-board students are the ones a CoE most **selects**, the marginal effect on the *already-strong, already-selected* student is smaller than on the typical rural entrant with room to move. The programme's value is largest exactly where the need is greatest, which is the opposite of a creaming story.

Reserved-category as a control. Adding the reserved indicator (on the cohorts where it was collected) moves the lift by at most 0.10 points in 2025 and 2026, confirming that category composition is not a hidden confounder and justifying its use as a descriptor rather than a control (Section 4.4).

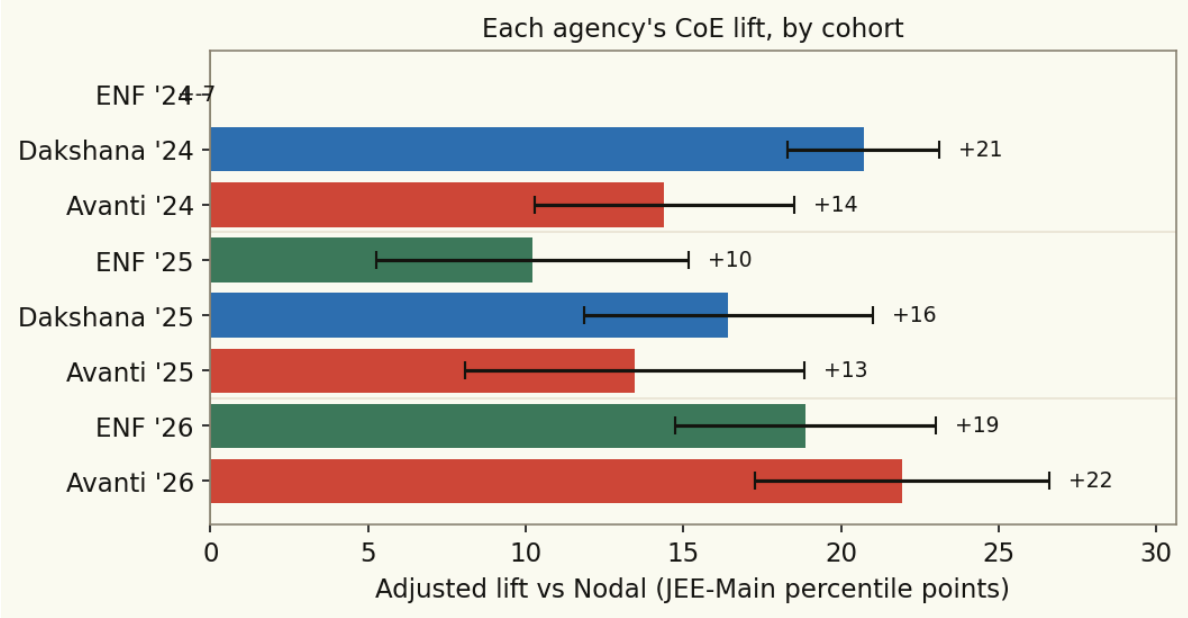
Agency split. The CoE bucket pools three operating agencies, Avanti, Dakshana and ENF. The question is whether the lift is one agency carrying the others, or a property of the intensive model itself. Per cohort, against Nodal ($n \ge 15$ per arm; cluster-robust 95% CIs):

Cohort	Avanti CoE	Dakshana CoE	ENF CoE
2024	+14.4 [10.3, 18.5] (n=75)	+20.7 [18.3, 23.1] (n=192)	−6.6 [−11.5, −1.7] (n=33)
2025	+13.4 [8.1, 18.8] (n=108)	+16.4 [11.8, 21.0] (n=242)	+10.2 [5.3, 15.2] (n=51)
2026	+21.9 [17.3, 26.6] (n=403)	not estimable	+18.9 [14.8, 23.0] (n=104)

A single cell rests on very few centres: Dakshana's 2024 cell is one centre, and ENF's are three to five. So the comparison that carries weight pools each agency over the cohorts it ran, with a cohort fixed effect, so that no agency is compared across exam years. Inference is a wild-cluster bootstrap over centres:

Agency	Cohorts	CoE students	School clusters	Lift vs Nodal [wild-bootstrap 95% CI]	p
Avanti CoE	2024, 2025, 2026	586	19	**+17.6** [13.2, 21.5]	$p < 0.001$
Dakshana CoE	2024, 2025	434	8	**+18.4** [10.9, 26.0]	$p < 0.001$
ENF CoE	2024, 2025, 2026	188	10	**+10.5** [2.9, 18.4]	$p = 0.017$

Avanti (+17.6) and Dakshana (+18.4) are **statistically indistinguishable** ($p = 0.27$ for the difference, from one model with all three agencies). ENF's pooled lift is about 5.5 points lower than Avanti's ($p = 0.04$), but that rests on 2024 and 2025. ENF's 2024 cell is −6.6, on 33 first-attempt students. By 2026 ENF (+18.9) and Avanti (+21.9) are close. With this few ENF centres we read the ENF gap as unsettled, not as a ranking.



7.1 Raw outcomes behind the agency lifts

The adjusted lifts above are residuals, so it is worth seeing the raw funnel each agency actually produced, and the intake it started from. Two things stand out: every CoE agency clears the exam at far higher rates than Nodal, and the differences that remain between the CoE agencies track the Class-10 intake each began with. (Split by exam; percentiles are JEE-Main percentiles; the Advanced columns use a category rank, lower = better. JEE Advanced ranks for 2026 are not yet loaded, so the Advanced table stops at 2025. These raw tables count every JEE-Main appearer with a percentile, board mark or not, so their n differs from Table 1.)

Table A. Raw JEE Main, of students who appeared.

Cohort	Agency	n appeared	% cleared Main	Median percentile	Class-10 board intake
2024	Avanti CoE	77	93.5	92.8	93.7
2024	Dakshana CoE	198	99.0	96.4	92.6
2024	ENF CoE	119	75.6	89.0	90.0
2024	Nodal	495	36.4	64.3	85.6
2025	Avanti CoE	117	76.1	84.3	87.0
2025	Dakshana CoE	252	96.8	96.8	92.0
2025	ENF CoE	58	87.9	90.6	93.0
2025	Nodal	438	40.0	67.2	82.8
2026	Avanti CoE	414	90.3	92.3	91.3
2026	ENF CoE	104	100.0	92.4	93.8
2026	Nodal	424	40.6	64.0	83.2

Table B. Raw JEE Advanced, of students who appeared (2024–2025).

Cohort	Agency	% of appearers who qualified JEE Advanced	Qualifiers with a category rank (n)	Median category rank
2024	Avanti CoE	36.4	22	2,792

Cohort	Agency	% of appearers who qualified JEE Advanced	Qualifiers with a category rank (n)	Median category rank
2024	Dakshana CoE	87.9	168	862
2024	ENF CoE	26.9	32	3,015
2024	Nodal	1.4	5	3,298†
2025	Avanti CoE	17.1	20	2,512
2025	Dakshana CoE	69.8	161	1,839
2025	ENF CoE	34.5	18	3,717
2025	Nodal	4.1	16	5,587

† Fewer than 10 ranked Advanced qualifiers, so the median rank is indicative, not reliable. The share of appearers who qualified Advanced counts a missing flag as not qualified; an earlier version averaged over flagged rows only, which overstated it (for example Nodal 2025 read 54.8% against the correct 4.5%).

Different starting lines, which is why the lift, not the rate, is the yardstick. Read against Nodal, every CoE agency is in a different league: Main clear rates of 75.6–100.0% against Nodal's 36.4–40.6%, and a far higher share qualifying JEE Advanced. Across the CoE agencies the raw rates are not identical, and the gaps line up with intake. Dakshana's stronger 2025 raw numbers (96.8% cleared Main, median percentile 96.8, 69.8% qualifying Advanced, median Advanced rank 1,839) rest on a Class-10 board intake of 92.0, above Avanti's 87.0 (76.1% cleared Main, median 84.3, 17.1% qualifying Advanced). The Advanced gap is wide, and Advanced is where a few board points matter most, since only the very top of the Main distribution qualifies. That is precisely why a raw pass rate cannot rank the agencies: it rewards whoever admitted the stronger cohort. Once board marks are held fixed, the two agencies' adjusted Main lifts are statistically indistinguishable (Avanti +13.4, Dakshana +16.4 in 2025). This paper does not adjust the Advanced outcome, so it makes no claim about the agencies' relative Advanced performance.

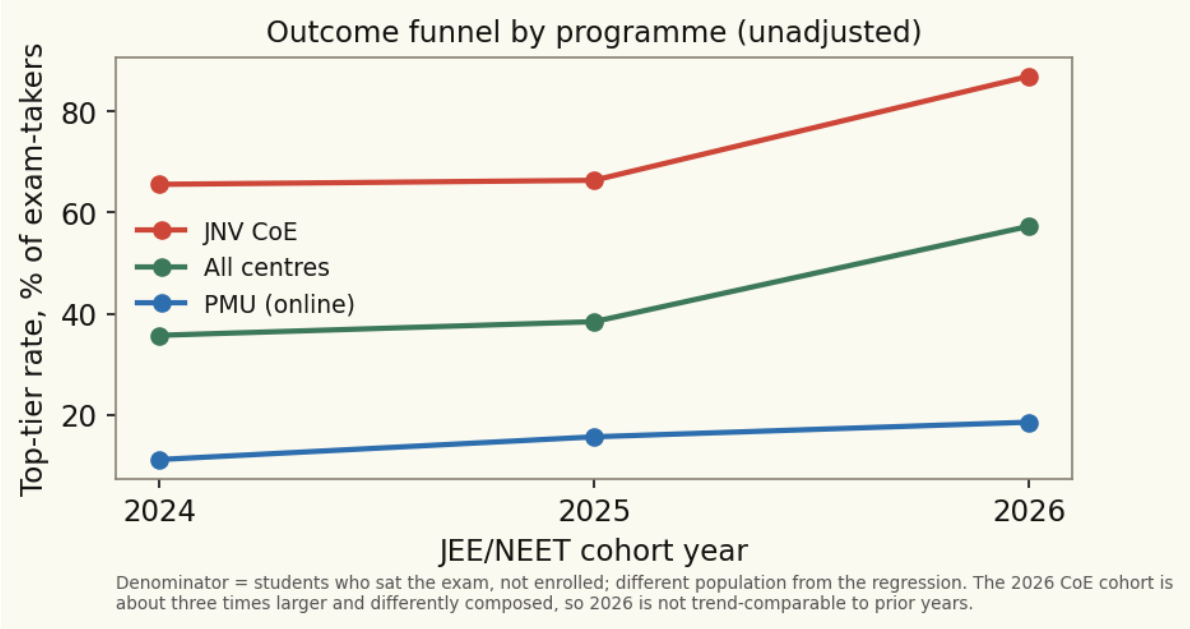
The three handover centres: the model, not the operator. The cleanest test of "is it the agency or the model?" comes from three centres Dakshana ran in 2025 that **Avanti took over in 2026**: JNV Bundi, JNV Kottayam and JNV Lucknow. The school, the catchment and the residential setting are held fixed; only the operator changed between consecutive cohorts. If the CoE result were a Dakshana-specific effect, it should drop when Avanti takes the wheel.

Centre	2025, Dakshana (n / % Main / median pct / board)	2026, Avanti (n / % Main / median pct / board)
Bundi	48 / 100% / 97.8 / 90.5	50 / 100% / 97.6 / 90.2
Kottayam	49 / 95.9% / 96.7 / 91.8	50 / 98% / 93.3 / 89.3
Lucknow	36 / 100% / 96.9 / 92.9	38 / 100% / 94.6 / 93.0

It does not drop. Across all three, Main clearance stays at the ceiling (98–100%) and the median percentile holds within 3.4 points through the operator change, while the Class-10 board intake under Avanti is equal or slightly lower (Bundi 90.5 → 90.2, Kottayam 91.8 → 89.3, Lucknow 92.9 → 93.0). The intensive model's output survived the handover intact, which is what "the lift attaches to the model, not the provider" predicts. Caveats: these are consecutive *cohorts*, not the same students;

2025 and 2026 face different exam difficulty; and JEE Advanced 2026 ranks are not yet loaded, so the swap is compared on JEE Main only. It is three centres, so we read it as corroboration of the agency split, not a standalone result.

Outcome backdrop (the funnel), read with care. As a descriptive backdrop, the share of each programme's *exam-takers* who reached a top-tier outcome (IIT, NIT/IIIT, MBBS/BDS, or a top engineering seat) is far higher at JNV CoEs (66%, 66%, 87% across 2024 to 2026) than at the online PMU programme (11 to 18%), with pooled physical centres in between (36 to 57%). Three cautions keep this out of the evidentiary chain. The denominator is students who *sat the exam*, not who enrolled, and exam-sitters are tracked more completely at physical centres than online, so coverage differs across the groups being compared. It is not adjusted for intake. And the 2026 CoE cohort is about three times larger and differently composed than earlier years, so the jump in 2026 is not a trend. (NEET 2026 is now included: 2026 CoE counts 53 MBBS/BDS outcomes.) We report the funnel as context for the regression, not as confirmation of it.



8. A second comparison: CoE and Nodal against Avanti's online students

The comparison so far measures what a CoE adds over Nodal. This section adds the step below it: how both centre arms compare with Navodaya students in Avanti's online programme, who were not at a centre. It answers a question the main contrast cannot, which is how much of a centre student's advantage over the rest of the cohort the lighter Nodal centre already delivers. It is a supporting comparison, for the reasons in Section 8.4, and it does not replace the CoE-versus-Nodal estimate.

8.1 Who is compared

There is no untreated group to compare with. The 2024 NTA results file tags every Navodaya JEE taker with the programme they were in at Class 12, and every first-attempt taker outside a centre is tagged as online:

NTA programme tag, 2024 (first attempts)	JEE takers
JNV ENABLE	6,665
Non JNV	4,397
Avanti Nodal	479
Dakshana CoE	192
Inconclusive	149
Avanti CoE	74
ENF CoE	34

`JNV ENABLE` is online support inside the school, run by Avanti or by the Ex-Navodayan Foundation (ENF). `Non JNV` students were in a JNV at Class 10 and left before Class 12. The comparison group is therefore an online one, and we use **Avanti's own online students**: the online test-prep programme in 2023-24, then JNV Enable and Enable Plus in 2024-25. The golden student record gives 3,587 of the 2024 students in this population an Avanti id but no programme. We place them from their own 2023-24 enrolment record, which puts 3,476 of them in a JNV online programme, and leave the rest out. The reason for using Avanti's own online students is measurement. Avanti records these students' JEE results in the same system as its CoE and Nodal results. Every other JNV student appears only in the NTA file, and in 2025 8,066 of its 12,103 rows carry no name, so they cannot be matched to a student or a board record. Section 8.4 widens the online arm to every non-centre JNV student in 2024, the one year the file allows.

The population in each cohort is every student who sat the JNV Class-12 board in maths and physics (PCM) and has Class-10 maths and science marks, at their first JEE attempt. A student who left the JNV after Class 10 did not sit that board, so drops out by construction.

Table 8. CoE, Nodal and Avanti online, by cohort. Class-12 PCM students with Class-10 marks. "Sat" is the share with a JEE Main result. "Qualified" counts a student with no result as not qualifying. Reserved share is over students whose category was recorded (Dakshana supplied none). The 2024 CoE arm's low female share reflects Dakshana's 2024 centre, whose students are all boys; gender is a control in every model.

Cohort	Arm	Students	Schools	Class-10 board, mean (SD)	Female	Reserved	Sat JEE Main	Qualified
2024	CoE	286	9	92.5 (6.2)	15%	94%	100%	92.3%
2024	Nodal	482	16	85.5 (8.6)	41%	86%	94%	33.4%
2024	Avanti online	4,434	443	83.2 (9.3)	37%	79%	67%	16.9%
2025	CoE	388	12	90.8 (7.5)	37%	73%	99%	93.8%
2025	Nodal	409	16	82.0 (9.9)	44%	82%	89%	37.9%
2025	Avanti online	2,078	192	78.9 (10.3)	41%	73%	48%	12.1%

The online arm starts below Nodal on Class-10 marks, and further below CoE, so intake again needs holding fixed. The arms also differ on something board marks cannot fix: who sits the exam. Almost every CoE and Nodal student has a JEE result. Among Avanti's online students, 67% do in 2024 and 48% in 2025.

8.2 Estimating equation

For each cohort we estimate, by ordinary least squares,

```
y_i = α + β_C·CoE_i + β_N·Nodal_i + γ·math10_i + δ·sci10_i + ζ·female_i + ε_i
```

with the online arm as the reference group, so β_C and β_N are the CoE and Nodal gaps against an Avanti online student with the same Class-10 marks and gender. There are two outcomes. The first is JEE Main qualification (0 or 100) for every PCM student, with a student who has no result counted as not qualifying. It does not depend on who chose to sit, so we lead with it. The second is the paper's own outcome, the JEE Main percentile, among those who sat. Inference is the restricted wild-cluster bootstrap over Class-12 schools, as in Section 6.

8.3 Results

Table 9. Gaps against Avanti's online students. β [wild-cluster bootstrap 95% CI]; n and school counts per cohort and outcome.

Cohort	Qualified: CoE	Qualified: Nodal	n / schools	Percentile (sat): CoE	Percentile (sat): Nodal	n sat
2024	+**63.6** [10.9, 117.5]	+**14.4** [7.4, 21.3]	5,202 / 458	+**21.7** [−3.0, 46.4]	+**4.2** [−0.2, 8.7]	3,706
2025	+**70.8** [64.0, 78.3]	+**23.7** [15.0, 33.3]	2,875 / 214	+**25.5** [21.1, 29.5]	+**9.0** [4.5, 13.5]	1,753

In 2025, when the CoE arm spans 12 centres, a Nodal student was **23.7 points [15.0, 33.3]** more likely to qualify in JEE Main than an Avanti online student with the same Class-10 marks and gender, and a CoE student **70.8 points [64.0, 78.3]** more likely. Among those who sat, the gaps are 9.0 and 25.5 percentile points. 2024 points the same way on qualification (Nodal 14.4, CoE 63.6) with wider intervals. Its CoE arm is a handful of centres, one of which holds 67% of the students, so the CoE interval even runs past 100 points, which no share can do, and among 2024 sitters the Nodal gap (4.2, [−0.2, 8.7]) cannot be told apart from zero.

Read as a ladder, the centre gap has two steps. In 2025 the Nodal step is 33% of a CoE student's qualification advantage over an online student, and the CoE adds the rest. On the percentile scale, the CoE-minus-Nodal step in this population is 16.5 points, close to the paper's main 2025 estimate of 15.4. That is a consistency check between the two samples, not new evidence.

8.4 What limits this comparison

- **Sitting is shaped by the programme.** The percentile comparison sets each centre's whole class against the online students who chose to sit, a stronger group than the online cohort as a whole. It understates the centres, which is why qualification among every PCM student leads.
- **2025 result coverage.** Only 48% of the 2025 online students have a JEE result, against 67% in 2024. Either fewer sat, or Avanti collected fewer results, and the NTA rows that would settle it are the unnamed ones. If online qualifiers went unreported, the 2025 qualification gaps are overstated. The percentile gaps are biased in an unknown direction.
- **Avanti's online students are only part of the online cohort.** Widening the 2024 online arm to every non-centre JNV student (the 2024 file lists them all, with names) gives Nodal

+16.4 and CoE +67.1 points on qualification. Avanti's online students do better than the wider online cohort, so the 2024 gaps reported above are the smaller ones.
- **Selection into each arm** is held fixed only on Class-10 marks and gender. Section 5's residential argument still applies within the JNV system, but the 2025 online students are spread over 192 schools, not placed in a few centres.
- **Two cohorts only.** The 2026 NTA file lists qualifiers only, and there is no 2026 Class-12 board record yet.

9. Heterogeneity

The equity cut: the lift reaches the disadvantaged. The central question for a programme serving rural, reserved-category students is whether the benefit concentrates in a better-off minority. It does not. Splitting the preferred regression by reserved-category status (Avanti and ENF CoEs against Nodal, since Dakshana supplied no category), the adjusted lift among reserved-category students is +12.7 points [7.6, 17.8] in 2025 (CoE n = 119) and +22.6 points [17.8, 27.4] in 2026 (CoE n = 424), at least as large as the lift among non-reserved students (+12.7 and +19.3; the 2025 non-reserved cell has only 40 CoE students). The CoE advantage is not concentrated in a better-off few. This is the paper's strongest equity statement and it is well powered on the reserved side; the female-interaction is not, so we do not report it.

Girls' share is a flag, not a result. The girls' share of the CoE arm against Nodal moves a great deal from year to year: 19.0% against 38.1% in 2024, 38.9% against 40.7% in 2025, and 28.6% against 42.2% in 2026, the year the CoE arm grew most. The 2026 gap sits awkwardly against Avanti's stated girls-and-engineering priority. The lift still holds for the girls who are enrolled, but the enrolment pattern itself deserves attention. The companion gender-stream study (the Lakshya analysis) treats the girls-and-engineering question directly on the larger non-CoE population.

10. Related work

Selective and residential schools. The closest evidence uses admission cut-offs to compare students just above and just below the line. In high-income settings, exam schools add little: offers of a seat at Boston and New York exam schools have little effect on most students' achievement (Abdulkadiroglu, Angrist & Pathak 2014), and New York exam schools raise course rigour but have little impact on SAT scores or college outcomes (Dobbie & Fryer 2014). In Kenya, elite national boarding schools show little evidence of raising learning; their reputations reflect who they select, and the authors warn that observers are effectively running the naive regression (Lucas & Mbiti 2014). Other settings find real gains: better schools in Trinidad and Tobago raise end-of-secondary exam performance, with effects about twice as large for girls (Jackson 2010); the SEED urban boarding schools in the United States raise reading and maths by about 0.2 SD a year, driven almost entirely by girls (Curto & Fryer 2014); and Karnataka's selective government model schools raise maths by 0.38 SD and pre-university entry by 11.5 points at comparable cost (Kumar 2019). The Romanian selective-school study (Pop-Eleches & Urquiola 2013) sits closest in design. A CoE differs from all of these in one respect: the treatment is targeted entrance-exam instruction within the same residential system, not a change of school or peer group. The selection warning applies with full force, which is why this paper leans on the intake ladder and the Oster bound.

Coaching access in India. Private tutoring in India is more common among urban, private-school and better-off students, and households are less likely to buy it for girls: at secondary level a girl is about 4 points less likely than a boy to receive it (Azam 2016). A free CoE place is, in part, a substitute for coaching these students' families would not or could not buy.

Gender and high stakes. The girls' penalty in this population matches a large literature. In China, girls' relative performance falls by 0.15 SD from a low-stakes mock to the Gaokao, cutting their chance of a Tier-1 university place by 15% (Cai, Lu, Pan & Zhong 2019). In Spain, girls' 0.18 SD advantage on low-stakes tests shrinks to 0.11 SD on high-stakes ones and reverses in science (Azmat, Calsamiglia & Iriberri 2016), and raising the weight of the entry exam cut women's admission-score advantage and their share of the most selective programmes (Arenas & Calsamiglia 2022). Czech women perform as well as men in less competitive admissions but are less likely to be admitted to the most selective universities (Jurajda & Münich 2011), and in the HEC Paris entrance exams women do worse than men in the most competitive stage but better in the less competitive ones (Ors, Palomino & Peyrache 2013). Responses to competition are one candidate explanation (Niederle & Vesterlund 2010). Upstream, Indian girls are about 20 points less likely than their brothers to take a science or commerce stream, a gap not explained by prior scores (Sahoo & Klasen 2021). At the IITs, women were about a fifth of JEE registrants but under 10% of batches before the 2018 supernumerary-seat scheme, which raised their share and saw them catch up academically by graduation (Sharma & Sengupta 2024).

The causal benchmark. The one randomised evaluation of an Avanti programme, a Grade-9 blended remedial course, found large gains in foundational maths and reading but no effect on the Class-10 board exam (Chiplunkar, Dhar & Nagesh 2020). It is the standard this observational paper falls short of, and a reminder that a programme's effect on one outcome does not carry to another.

11. Discussion and limitations

What the paper shows: among JNV residential-school students, being in a CoE is associated with a JEE percentile about 15 to 22 points higher than being in a Nodal programme, after adjusting for the main observable intake difference, and this association is stable under controls and survives a standard selection-on-unobservables bound under its baseline assumption.

What it does not show: that the CoE *causes* the lift. Assignment is not random. CoE students are positively selected on board marks (Table 2), and plausibly on unmeasured traits such as motivation. We adjust for the former and bound the latter, but we cannot eliminate it. Four further limits. The estimand is intensity, not a centre against online support: Nodal is itself a centre, so β understates the whole effect of a centre. Section 8 estimates the step below it, against Avanti's own online students in 2024 and 2025, with the caveats set out there. We cannot use school fixed effects (CoE and Nodal are different schools), so school-level confounders are handled only by the residential design argument, not absorbed. In the 2024 cohort one centre holds 64% of the CoE students, and its interval is very wide once inference is done correctly, so the evidence is really the 2025 and 2026 cohorts (and their pooled +17.7). And external validity is limited: the in-sample CoEs are a particular set of sites, each agency runs only a handful of centres (Dakshana's 2024 cohort is one centre), and a marginally-expanded CoE at a less-ready JNV could deliver less than these established sites. That last point is itself the argument for the staggered rollout below, which would estimate the effect precisely at expansion sites.

The most plausible surviving objection, selection on motivation, has a useful property: motivation and effort are precisely what an intensive coaching programme is built to raise. So even the conservative reading, that some of the lift is motivated students sorting into CoEs, returns a policy question rather than a dismissal: does placing a motivated student in a CoE, or building that motivation, produce the score? The clean way to separate the two is a **phased or randomised rollout**: assign the next tranche of CoE expansion across eligible JNVs in a staggered or lottery design, and the CoE-vs-Nodal contrast becomes causal. That is the recommended next step, and it is feasible because CoE expansion happens in tranches already.

A mechanism is hypothesised, not tested here. The companion practice-and-qualification study shows, on the larger non-CoE JNV population, that doing full-length evaluative practice tests predicts JEE qualification (each additional serious test is associated with about 1.6 times the odds of qualifying; predicted qualification rises from roughly 8% at zero tests to 20% at three). Intensive structured testing is the most visible thing a CoE does that a Nodal does less of, which makes it a plausible candidate for what the lift reflects. But that estimate is on a different population and a different outcome, and this paper carries no practice-volume measure for its 2,386 students, so we import the mechanism as a hypothesis, not as a decomposition of the CoE effect.

12. Conclusion

The CoE-vs-Nodal JEE gap is large, and most of it is not intake. After holding Class-10 board marks and gender fixed, with the intake control specified flexibly enough not to lean on functional form (§7), an advantage of about 15.3 points in 2025 and 20.5 in 2026 remains (about +17.7 pooled; a linear control gives 15.4 / 21.9 / 18.4, the upper bound), stable across every intake control we try, significant under cluster-correct inference, and of similar size across operating agencies and across reserved and non-reserved students. This is the strongest statement an observational comparison supports; it stops short of causal, and the honest next step is a staggered rollout that would make the contrast experimental. The policy margin is concrete: the intensive on-campus model is associated with outcomes the lighter-touch and online models are not, on the same rural residential population, and the benefit reaches the disadvantaged majority rather than concentrating in a better-off few.

13. Appendix: reproducibility

Every figure and number is reproducible from the golden student record, and every number in this paper is filled in from the scripts' outputs by `paper_render.py` (the prose lives in `paper.template.md`):

Artifact	Produced by
Regression, 3-way, funnel (the headline numbers + gates)	`../core-impact-analyses/analyze.py`, validated by `check.py`
Figures, balance table, nested ladder, validation CSV	`../core-impact-analyses/make_paper_assets.py`
Wild-cluster bootstrap, Oster grid, pooled spec, equity cut, intake ladder, bounded-outcome + ceiling heterogeneity (`robustness.json`)	`../core-impact-analyses/robustness.py`
Raw agency funnel and handover centres	`../core-impact-analyses/analyze_agency_raw.py`

Artifact	Produced by
CoE and Nodal against Avanti's online students, 2024 and 2025 (§8)	`../core-impact-analyses/online_compare.py`
The student record it reads	`../golden-student-record/` (identity crosswalk + classification + BQ)

This manuscript lives in git; the **figures**, the **PDF**, and the row-level **validation dataset** (`output/validation_impact.csv`, 2,386 rows, PII-stripped: no name or date of birth, and the student id and school each replaced by a stable hash, so the clustering reproduces without naming a student or a school) are hosted in Cloud Storage and linked below. A skeptical reader can re-run every table and regression, including the wild-cluster bootstrap and Oster grid, from the validation CSV without warehouse access.

Caveat on inputs. The JEE outcome comes from `jnv_fact_jee_results` through the golden record, not from the vendored NTA parquets the original `11_avanti_impact_regression` used, so the numbers drift from that analysis; the sign, significance and rough magnitude match (`check.py`).

Replication data

Re-run every table and regression here **without warehouse access**. All artifacts are hosted in Google Cloud Storage (aggregate / PII-stripped — no name or date of birth; internal id kept and school hashed so the fixed effects reproduce):

- Rendered PDF
- validation_impact.csv — 2,386 rows, the row-level regression input
- balance_table.csv · ladder.csv · robustness.json
- agency_raw_outcomes.csv — the §7.1 raw agency×year funnel
- handover_centres.csv — the §7.1 Dakshana→Avanti handover centres
- Figures: CoE lift · agency lift · funnel · JEE-percentile box

Sources

- Avanti Fellows Alumni RoI analysis (2026). Cohort earnings, unit cost, and social/dollar return per student, the income lift that a JEE-rank movement compounds into. Replication data: https://storage.googleapis.com/avantifellows-bq-assistant/analysis/alumni-roi/roi.csv (cited in the Executive Summary for scale; not part of this paper's causal chain).
- Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics* 37(2), 187-204. (The δ* bound used in Section 7.) download
- Altonji, J., Elder, T. and Taber, C. (2005). Selection on Observed and Unobserved Variables. *Journal of Political Economy* 113(1), 151-184. (The origin of the selection-proportionality idea Oster formalises.) download
- Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review* 14(3), 693-709. https://doi.org/10.2307/2525981 download; and Blinder, A. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *Journal of Human Resources* 8(4), 436-455. https://doi.org/10.2307/144855 download. (The decomposition in Section 7, valued at the Nodal return.)

- Ors, E., Palomino, F. and Peyrache, E. (2013). Performance Gender Gap: Does Competition Matter? *Journal of Labor Economics* 31(3), 443-499. https://doi.org/10.1086/669331 download
- Cameron, A., Gelbach, J. and Miller, D. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics* 90(3), 414-427. (The restricted wild-cluster bootstrap used as primary inference, appropriate for few treated clusters.) download
- Pop-Eleches, C. and Urquiola, M. (2013). Going to a Better School: Effects and Behavioral Responses. *American Economic Review* 103(4), 1289-1324. (The selective-school-to-high-stakes-exam design this paper's observational comparison sits closest to.) download
- Chiplunkar, G., Dhar, D. and Nagesh, R. (2020). Too Little, Too Late: Improving Post-Primary Learning Outcomes in India. Working paper, 29 June 2020. (A randomised trial of Avanti's Grade-9 blended remedial programme: large gains in foundational maths and reading, no effect on the Class-10 board exam. The house standard for causal identification that this observational paper falls short of.) https://riseprogramme.org/sites/default/files/inline-files/Dhar.pdf download
- Companion Avanti work: the regular-practice study (nvs-lakshya, *Showing Up Regularly*).

Added 7 October 2026. Each work below was downloaded, read, and its claims checked against the archived copy before being cited (archived copies open to avantifellows.org accounts; checksums in `SOURCES.md`). Where the copy read is a working-paper version of the journal article, the reference says so.

- Abdulkadiroglu, A., Angrist, J. D., & Pathak, P. A. (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. Econometrica, 82(1), 137-196. (Copy read: NBER Working Paper 17264, July 2011, revised September 2011.) https://doi.org/10.3982/ECTA10266 download
- Arenas, A. and Calsamiglia, C. (2022). Gender Differences in High-Stakes Performance and College Admission Policies. IZA Discussion Paper No. 15550. https://ideas.repec.org/p/iza/izadps/dp15550.html download
- Azam, M. (2016). Private Tutoring: Evidence from India. Review of Development Economics, 20(4), 739-761. (Copy read: IZA Discussion Paper No. 8770, January 2015.) https://doi.org/10.1111/rode.12195 download
- Azam, M. and Kingdon, G. G. (2013). Are Girls the Fairer Sex in India? Revisiting Intra-Household Allocation of Education Expenditure. World Development, 42, 143-164. Copy read: IZA Discussion Paper No. 5706 (2011). https://doi.org/10.1016/j.worlddev.2012.09.003 download
- Azmat, G., Calsamiglia, C. and Iriberri, N. (2016). Gender Differences in Response to Big Stakes. Journal of the European Economic Association, 14(6), 1372-1400. Copy read: CEP Discussion Paper No. 1314, revised October 2015. https://doi.org/10.1111/jeea.12180 download
- Borker, G. (2021). Safety First: Perceived Risk of Street Harassment and Educational Choices of Women. World Bank Policy Research Working Paper 9731. https://documents.worldbank.org/curated/en/723631626710146405 download
- Cai, X., Lu, Y., Pan, J., & Zhong, S. (2019). Gender Gap under Pressure: Evidence from China's National College Entrance Examination. Review of Economics and Statistics, 101(2), 249-263. (Copy read: authors' working-paper version, October 2016.) https://doi.org/10.1162/rest_a_00749 download

- Conley, T. G. and Taber, C. R. (2011). Inference with 'difference in differences' with a small number of policy changes. Review of Economics and Statistics, 93(1), 113-125. Read as NBER Technical Working Paper No. 312 (2005). https://doi.org/10.1162/REST_a_00049 download
- Curto, V. E., & Fryer, R. G. Jr. (2014). The Potential of Urban Boarding Schools for the Poor: Evidence from SEED. Journal of Labor Economics, 32(1), 65-93. (Copy read: NBER Working Paper 16746, 'Estimating the Returns to Urban Boarding Schools: Evidence from SEED', January 2011.) https://doi.org/10.1086/671798 download
- Diegert, P., Masten, M. A. and Poirier, A. (2026). Assessing omitted variable bias when the controls are endogenous. arXiv:2206.02303v6 [econ.EM], 4 Feb 2026. https://arxiv.org/abs/2206.02303 download
- Dobbie, W., & Fryer, R. G. Jr. (2014). The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools. American Economic Journal: Applied Economics, 6(3), 58-75. (Copy read: NBER Working Paper 17286, 'Exam High Schools and Academic Achievement: Evidence from New York City', August 2011.) https://doi.org/10.1257/app.6.3.58 download
- Fortin, N., Lemieux, T. and Firpo, S. (2011). Decomposition methods in economics. In O. Ashenfelter and D. Card (eds), Handbook of Labor Economics, Vol. 4A, 1-102. Read as NBER Working Paper No. 16045 (June 2010). https://doi.org/10.1016/S0169-7218(11)00407-2 download
- Jackson, C. K. (2010). Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago. The Economic Journal, 120(549), 1399-1429. https://doi.org/10.1111/j.1468-0297.2010.02371.x download
- Jurajda, S. and Munich, D. (2011). Gender Gap in Performance under Competitive Pressure: Admissions to Czech Universities. American Economic Review: Papers & Proceedings, 101(3), 514-518. Copy read: CERGE-EI Working Paper 371, 'Gender Gap in Admission Performance under Competitive Pressure', November 2008. https://doi.org/10.1257/aer.101.3.514 download
- Kumar, N. (2019). Public Schools Can Improve Student Outcomes: Evidence from a Natural Experiment in India. Job market paper, University of Illinois at Chicago, November 15, 2019 (unpublished). https://ideas.repec.org/p/jmp/jm2019/pku634.html download
- Lucas, A. M., & Mbiti, I. M. (2014). Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya. American Economic Journal: Applied Economics, 6(3), 234-263. https://doi.org/10.1257/app.6.3.234 download
- MacKinnon, J. G. and Webb, M. D. (2017). Wild bootstrap inference for wildly different cluster sizes. Journal of Applied Econometrics, 32(2), 233-254. Read as Queen's Economics Department Working Paper No. 1314 (corrected 2019-08). https://doi.org/10.1002/jae.2508 download
- MacKinnon, J. G. and Webb, M. D. (2018). The wild bootstrap for few (treated) clusters. Econometrics Journal, 21(2), 114-135. Read as Queen's Economics Department Working Paper No. 1364 (Nov 2017). https://doi.org/10.1111/ectj.12107 download
- MacKinnon, J. G., Nielsen, M. O. and Webb, M. D. (2023). Cluster-robust inference: A guide to empirical practice. Journal of Econometrics, 232(2), 272-299. Read as Queen's Economics Department Working Paper No. 1456 (rev. Nov 2022). https://doi.org/10.1016/j.jeconom.2022.04.001 download
- Niederle, M. and Vesterlund, L. (2010). Explaining the Gender Gap in Math Test Scores: The Role of Competition. Journal of Economic Perspectives, 24(2), 129-144. https://doi.org/10.1257/jep.24.2.129 download

- Papke, L. E. and Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. Journal of Applied Econometrics, 11(6), 619-632. Read as NBER Technical Working Paper No. 147 (November 1993). https://doi.org/10.1002/(SICI)1099-1255(199611)11:6<619::AID-JAE418>3.0.CO;2-1 download
- Sahoo, S. and Klasen, S. (2021). Gender Segregation in Education: Evidence from Higher-Secondary Stream Choice in India. Demography, 58(3), 987-1010. Copy read: IZA Discussion Paper No. 11660 (2018), 'Gender Segregation in Education and Its Implications for Labour Market Outcomes: Evidence from India'. https://doi.org/10.1215/00703370-9101395 download
- Sharma, S. and Sengupta, N. (2024). Debunking myths associated with affirmative action: using evidence from Supernumerary Seat Scheme. Working paper, IIT Delhi (presented at ACEGD 2024, ISI Delhi), 13 September 2024. https://www.isid.ac.in/~acegd/acegd2024/papers/SwatiSharma.pdf download