AVANTI FELLOWS · RESEARCH & DATA

JEE qualification forecast: methodology

The model, its assumptions, and a like-for-like comparison with the linear regression it replaces

Created 21 July 2026 · updated 7 October 2026 · technical note

In plain terms

What the model does. For each CoE student it estimates a range of JEE-Mains percentiles they are likely to score, not a single number. From that range it reads off the chance that they clear their category's cutoff. The forecast count is the sum of those chances: ten students who each have a 70% chance count as seven qualifiers. Nobody is labelled "will" or "won't".

What goes into a student's estimate. Seven things, each learned from 2,063 past CoE and Nodal students (2024 to 2026), each at their first JEE attempt. The details are in Section 3.7.

- **Grade-10 board marks** (maths and science average). Higher marks, higher percentile. This is the strongest steady signal.
- **Being at a CoE rather than a Nodal centre.** The largest single factor. For an OBC boy with an 85 board average, it moves the expected percentile from about the 68th to the 87th.
- **Being a girl.** Compare a girl and a boy with the same marks, category, stream and centre type: the girl scores lower on JEE, on average. At a CoE that is about 6 percentile points at a board of 85. This is the gap that remains **after** marks are accounted for, so it is not explained by girls arriving with weaker marks. Girls in fact enter with board marks within about 2 points of boys'. The model measures this gap; it cannot tell us why it exists.
- **Category.** EWS students score highest for given marks and SC and ST students lowest. Category matters twice: once here, and again because each category has its own cutoff.
- **The student's school**, through a small, shrunk-to-average adjustment for each centre we have seen before.
- **The year**, which shifts every student up or down together. Years vary, so the forecast carries that as uncertainty.
- **Stream** (PCM vs PCMB), which makes no difference.

What none of this explains. Two students who match on every factor above still land far apart. For an OBC boy at a CoE with an 85 board average, the typical estimate is the 87th percentile. But two-thirds of students like him land anywhere between the 72nd and the 95th. That scatter is the "residual". It is why the model gives each student a probability, not a verdict, and why a single student's estimate is a weak guide while a count across a cohort is a good one.

What this means for girls.

1. **A girl's forecast is lower than a boy's with the same marks**, because past girls with those marks scored lower. The model is reporting what has happened, not passing a judgement on what girls can do.
2. **A CoE does not close the gap, and may not lift girls quite as much as boys.** We tested whether the CoE effect differs by gender. The estimate is that the gap is a little wider at a CoE than at a Nodal centre, but the evidence is weak (p = 0.08), so read it as "no sign that a CoE narrows the gap", not as a measured difference.
3. **The gap is wider for girls with high board marks**, and it was wider in 2025 than in 2024. Model B uses one girls' adjustment for every student, so at the very top it is slightly generous to girls. We tested adding a term for this. It made the high-board girls' estimates more accurate at the M1 bar, but less accurate for boys and no better at the qualifying bar the forecast counts, so it is not adopted (Section 3.7).

1. Purpose and framing

The task is a forecast. For the students enrolled in CoEs today, how many will clear the JEE-Mains cutoff for their category. It is not a causal question. We are not estimating what the CoE programme adds over some counterfactual; we are estimating where these particular students are likely to land, given what we observe about them now. That framing sets the requirements. A forecast of a count near a threshold needs three things a plain regression does not provide: a prediction that respects the 0 to 100 bound of a percentile, a per-student probability rather than a hard verdict, and an honest decomposition of what we do not know into the reported uncertainty.

The statistical character of the problem is prediction under distribution shift. We have three past cohorts (2024 to 2026) with known outcomes, and we predict two future cohorts (2027, 2028) whose composition differs: lower board marks and a heavier weight on OBC and General students, who face the highest cutoffs. A method that merely fits the past well is not enough; it has to transport to a differently-composed, as-yet-unobserved cohort, and it has to say how uncertain that transport is.

2. Data and estimand

We fit on past JNV CoE and Nodal students, cohorts 2024 to 2026, from the golden student record, restricted to those who sat JEE Mains and have a Grade-10 board mark, at their **first** JEE attempt (n = 2,063; 886 CoE, 1,177 Nodal; 794 girls; golden record as of 7 October 2026 with the NTA-flag fix). The unit is a student. Since 22 July 2026 the golden record holds one row per student per exam attempt, so a student who sat JEE in 2025 and again in 2026 has a second row in 2026. That second attempt came after the programme and is not the outcome forecast here, so `model.first_attempts()` keeps each student's earliest exam year, linking two golden ids that share an Avanti student id. It drops 93 rows from this sample; left in, they cut the CoE coefficient from 1.13 to 0.86 and the FY28 forecast from 73% to 68%. The outcome is the JEE-Mains national percentile. The estimand is the conditional distribution of that percentile given a student's intake, for a CoE-type student, from which the qualification probability and the expected count follow. Full provenance and coverage funnels are in the README. This note assumes the golden record after its category-recovery fix, without which category was missing for the 2024 and 2025 CoE cohorts.

3. The model

3.1 Outcome transform

Let `pct` be the JEE percentile in (0, 100). We model its logit,

```
y = logit(pct / 100) = log( p / (1 − p) ),   p = pct / 100,
```

with `pct` clipped to [0.5, 99.5] so the transform is finite at the boundary. Modelling `y` rather than `pct` has two consequences that matter. First, any prediction, when mapped back through the inverse logit, lies strictly in (0, 100); the model cannot emit an impossible percentile. Second, the scale is multiplicative in the odds of out-ranking a random test-taker, so a fixed additive effect on `y` is a large percentile move in the middle of the distribution and a small one near the ceiling, which is the correct behaviour: moving from the 94th to the 99th percentile is far harder than from the 50th to the 55th.

3.2 Estimating equation

We fit a linear mixed model on `y`:

```
y_ic = x_ic' β + u_s(i) + τ_c + ε_ic
u_s ~ N(0, σ²_school),   ε_ic ~ N(0, σ²_resid)
```

for student `i` in school `s(i)` and cohort `c`. The fixed-effect design `x'β` contains board marks entered as a quadratic (`b`, `b²`, centred at 85), category (SC, ST, OBC, EWS, GEN), gender, stream, a CoE indicator, and the CoE indicator interacted with the board terms (`coe·b`, `coe·b²`). The interaction lets the board-to-outcome relationship differ in shape between CoE and Nodal, so the CoE contribution is heterogeneous in board rather than a single constant. Cohort enters as fixed effects `τ_c`; their dispersion is used in Section 3.4 as the transfer variance. Estimation is by restricted maximum likelihood (statsmodels `MixedLM`, L-BFGS).

Fitted variance components on the full sample: residual SD 0.96, school SD 0.42, and a between-cohort SD of 0.24 (all on the logit scale; Section 3.7 explains each). The school grouping uses a canonicalised school key (upper-cased, whitespace-collapsed): raw school strings drift in case between sources, which silently split one school's random intercept into two half-weight groups (10 phantom clusters before the fix).

Board floor at prediction time. The fitted `b²` coefficient is positive (about 0.0021), so the quadratic reaches its minimum near a board of 62 and rises again below it. Taken at face value, an OBC male CoE student with a board of 40 would be predicted at the 88th JEE percentile, against the 66th at a board of 60. The training data cannot support that upturn: its 5th-percentile board is 69 and only 20 rows fall below 60. So when scoring, any student below the training 5th percentile is scored as if at it (`BOARD_FLOOR_QUANTILE` in `model.py`, the same rule as `analysis/jnv_potential`). The prediction is then flat below 69 and rising above it, so a lower board can never earn a higher forecast. The fit itself is unchanged. On the October 2026 roster the floor touches 16 FY27 and 17 FY28 students and moves the expected counts by 0.1 and 0.3 qualifiers.

3.3 School random intercept and Nodal borrow-strength

The school term `u_s` is a random intercept, estimated with partial pooling. A centre with few students is shrunk toward the overall mean in proportion to how little it carries, so small centres do not swing on a handful of results, and a centre we have never seen contributes no idiosyncratic offset (its intercept is set to its prior mean of zero, and the school variance is added to that student's predictive spread instead). This is what makes the per-centre numbers stable and what lets the model score a brand-new centre honestly.

Nodal students enter the fit with the CoE indicator off. This borrows strength: the board-to-outcome shape is estimated on the larger pooled sample, which stabilises the sparse low-board CoE cells (few high-scoring-board General students exist in past CoE data), while the CoE indicator and its board interaction carry the CoE-specific level and slope. The CoE prediction uses the indicator on.

3.4 From a fitted model to a predictive distribution

For a student in a new cohort (2027 or 2028) the predicted logit outcome is Normal,

```
y* ~ N( μ*, σ*² )
μ*  = x*' β̂  +  û_s (if the school was seen, else 0)  +  τ̄
σ*² = σ²_resid  +  Var(τ)  +  σ²_school · 1{school unseen}  +  x*' Σ̂_β x*
```

The mean adds the average cohort effect $\bar{\tau}$ because 2027 and 2028 are new draws, not any observed year. The variance is the sum of four sources, each with a clear meaning: residual student-level variation; the between-cohort transfer $\mathrm{Var}(\tau)$, which is the price of predicting an unobserved year; the school variance for centres with no history; and the parameter uncertainty in $\hat{\beta}$. The transfer term is estimated from three cohorts, so it is itself imprecise; we treat it as a floor on year-to-year uncertainty, not a tight quantity.

3.5 From a distribution to a count

Qualification is $\mathrm{pct} \geq c$ for the student's category cutoff c. On the logit scale this is $y^* \geq \mathrm{logit}(c/100)$, so the probability is

$$P_i = P(y^* \geq \mathrm{logit}(c/100)) = \Phi\left((\mu^* - \mathrm{logit}(c/100)) / \sigma^* \right).$$

The expected number of qualifiers is the sum of these probabilities, $E[\#] = \Sigma_i P_i$ (the mean of a Poisson-binomial). This is the correct expectation of the count, $E[\Sigma\ 1\{\mathrm{pct}_i \geq c\}] = \Sigma\ P(\mathrm{pct}_i \geq c)$. It is not the same as thresholding a point prediction, $\Sigma\ 1\{\hat{y}_i \geq c\}$, which is what the linear approach does; the two differ by Jensen's inequality and, more importantly, the threshold discards all information about students near the line, who are precisely where a count is decided.

3.6 Uncertainty: the reported band

The band is a Monte-Carlo simulation. In each of 2,000 iterations we draw a new-cohort shift from the transfer distribution, $\delta \sim N(0, \mathrm{Var}(\tau))$, and a cutoff for each category from the recent-year set

(the 2027 and 2028 cutoffs are unknown), then recompute the realized count as a Poisson-binomial draw. The 10th, 50th, and 90th percentiles of the simulated counts form the band. By construction the band is dominated by the two quantities that actually move a count of students sitting near the line, the cutoff and the year-to-year shift, and not by sampling noise in the coefficients, which is small at this sample size.

One honest note on what the band contains: because each iteration realizes the count as a Poisson-binomial draw, this is a *realized-count* band, not a band on the expectation — it deliberately includes the irreducible student-level realization noise (roughly ±1 point on a cohort this size) on top of the cutoff and transfer uncertainty. A band on the expected count alone would be slightly narrower.

3.7 What each factor does, in plain language

How to read a logit effect. The model adds up its factors on the logit scale (Section 3.1) and only then converts the total back into a percentile. A fixed step on the logit scale is a large percentile move in the middle of the range and a small one near the top: +0.1 is about +2.5 percentile points around the 50th percentile, but under +1 point around the 90th. So every effect below is constant on the logit scale, and its size in percentile points depends on where the student starts.

The fitted effects. From `model.fit()` on the golden record of 7 October 2026 (first attempts): 2,063 CoE and Nodal students, 2024 to 2026, 794 of them girls. Every number in this section is printed by `stats.py`. The reference student is a boy, EWS, PCMB, at a Nodal centre, in the 2024 cohort.

Factor	Effect (logit)	s.e.	What it means
Board average, per mark (around 85)	+0.10	0.004	A steady climb with marks, slightly curved (`b²` = +0.002)
Studying at a CoE (vs Nodal)	**+1.13**	0.11	The largest single effect. It grows a little with board marks (`coe:b` +0.014 per mark, p = 0.02): about +0.2 logits more at a board of 95 than at 80
Girl (vs a boy with the same marks, category, stream, centre type)	**−0.48**	0.05	Clearly below zero (p < 0.001)
Category vs EWS: General	−0.22	0.10	Effect on the percentile. Each category's own cutoff is applied afterwards, so General faces the highest bar as well
Category vs EWS: OBC	−0.30	0.09	
Category vs EWS: SC	−0.51	0.09	
Category vs EWS: ST	−0.49	0.10	
PCM vs PCMB stream	+0.13	0.07	Small and not clearly different from zero (p = 0.07)
Cohort 2025 / 2026, vs 2024	+0.27 / −0.22	0.07 / 0.09	Year-to-year level. A new year gets the average, and the spread becomes the transfer uncertainty

Three spreads complete the model. These describe variation rather than effects:

Spread (SD, logit)	Value	In plain terms
School	0.42	Centres differ. Two centres one SD apart differ by about 10 percentile points for a student near the 50th percentile. Small centres are pulled toward the average, so a centre with ten students is not judged on its ten results
Cohort transfer	0.24	How much a whole year moves up or down together. A new year is a fresh draw, which widens the forecast band
Residual	0.96	What none of the factors explain. An OBC boy at a CoE with an 85 board average has a typical estimate of the 87th percentile, but two-thirds of students like him land between the 72nd and the 95th. That scatter is why each student gets a probability

What the effects mean in percentiles. The table shows the median predicted percentile for an OBC, PCM student at a centre the model has not seen before, in a new cohort. Board marks are in the rows, and each column pair is a boy and a girl with everything else equal.

Board average	CoE: boy	CoE: girl	Gap	Nodal: boy	Nodal: girl	Gap
80	80.2	71.5	−8.6	58.0	46.2	−11.8
85	86.9	80.4	−6.4	68.1	57.0	−11.1
90	92.5	88.4	−4.1	78.5	69.4	−9.1
95	96.3	94.1	−2.1	87.4	81.2	−6.2

The CoE lift is visible in every row: at a board of 85 it takes a boy from the 68th to the 87th percentile. The girls' gap **narrows toward the top only because percentiles crowd together near 100**. The underlying effect is the same −0.48 in every row, and the checks below show that it actually grows with board marks.

Do girls and boys differ systematically? Yes. A girl with the same marks, category, stream and centre type as a boy scores lower in JEE, by −0.48 on the logit scale. This is not a gap in intake: across 2024 to 2026, girls entered with board averages between 0.6 points above and 2.8 points below boys', yet trailed them by 4 to 13 JEE percentile points.

Cohort	Boys: board / JEE pct	Girls: board / JEE pct	JEE gap
2024	86.4 / 67.7 (n = 312)	87.0 / 63.9 (n = 222)	−3.8
2025	86.7 / 79.5 (n = 406)	85.1 / 69.6 (n = 290)	−9.8
2026	89.0 / 79.2 (n = 551)	86.2 / 66.2 (n = 282)	−13.0

The data establish the size of the gap. They do not establish its cause, and this note does not offer one.

Three checks, each adding a single term to `FORMULA` in the same school-grouped MixedLM:

1. **CoE vs Nodal: no sign that a CoE closes the gap.** `female:coe` = −0.18 (s.e. 0.10, p = 0.08). The point estimate says the gap is a little wider at a CoE, but it is not clearly different from zero. (Before the repeat attempts were removed this read −0.08, p = 0.45.)
2. **Higher board marks: the gap is wider.** `female:b` = −0.019 per mark (s.e. 0.005, p < 0.001). The girls' effect is about −0.35 at a board of 80 and −0.64 at 95. Model B's single `female` term averages these, so it is slightly generous to high-board girls and slightly harsh to low-board girls.

3. **By cohort: the gap was wider in 2025.** With `female:C(cohort)` the 2024 girls' effect is −0.36, and 2025 and 2026 are a further −0.22 and −0.11 (p = 0.06 and 0.34). The model's single term pools all three years.

Should Model B carry `female:b`? We tested it the same way the model is validated (Section 7): fit on two cohorts, predict the held-out cohort's CoE students as if unseen, and compare against what happened. The table gives the count error (predicted minus actual qualifiers, as percentage points of the group) with and without the term.

Held-out CoE students	n	Qualifying bar: B	+ `female:b`	M1 bar: B	+ `female:b`
All	886	+0.3	+0.2	+3.9	+3.9
Boys	564	+0.7	+0.8	+4.7	+5.3
Girls	322	−0.6	−0.8	+2.5	+1.4
Girls, board 92 and above	181	−0.0	−1.2	+5.2	+2.9

(The M1 bar is the floor of that year's M1b band in `dim_nta_cutoffs`. The log score is the mean held-out log density of the logit outcome.) The term does what the check predicts at the M1 bar: high-board girls were over-predicted by 5.2 points and that falls to 2.9. It buys this by making boys worse, and the overall log score of the held-out percentiles falls (−1.628 to −1.637). At the qualifying bar, which is what the forecast counts, it changes nothing. **So Model B keeps the single girls' term.** Anyone using Model B's M1 probabilities for girls, as the JNV potential analysis does for its M1 layer, should read high-board girls' M1 numbers as slightly high. That analysis reports the same effects and checks in its Section 3.

4. Identification and assumptions for a forecast

A forecast rests on assumptions of transport, not of causal exchangeability. We state them plainly.

- **In-population fit.** We estimate `E[y | x]` on the same programme (JNV CoE, plus Nodal for shape) we predict on, so there is no assumption that a national relationship transfers to our students. This is the central correction to the earlier model.
- **Conditional transport across cohorts.** We assume the mapping from intake to outcome, `E[y | x]`, is stable across cohorts up to the cohort-level shift `τ_c`, whose variance we carry into the band. If the mapping itself changed shape in 2028 (for reasons beyond level and composition), the point estimate would be off.
- **Common support.** The 2028 cohort is lower-board and more OBC than past CoE cohorts. Where a 2028 intake cell is thinly represented in the training data, the prediction leans on the pooled shape (including Nodal) and on students like them elsewhere; the real uncertainty there is wider than the band alone conveys. The category-recovery fix to the golden record materially improved this, since 2024 and 2025 CoE category is now complete.
- **Cutoff exogeneity.** The category cutoff enters only as the threshold, drawn from its recent distribution. It is not a regressor. (The earlier model used it as both, which is incoherent; see Section 5.)

5. Comparison with the linear regression it replaces

The point of this section is to show, on the same data, how a linear regression and our method differ. We hold the training sample and the predictors fixed and vary only the method:

- **Linear.** Ordinary least squares on the raw percentile, `pct ~ b + b² + coe + coe·b + coe·b² + category + female + stream`, giving one predicted percentile per student. A student is counted as a qualifier if that point prediction is at or above the category cutoff (a hard threshold). The count is the number above the line.
- **Ours.** The bounded hierarchical model above, giving a predictive distribution per student, a probability of clearing the cutoff, and a count equal to the sum of probabilities.

5.1 On held-out cohorts (leave-one-cohort-out)

Both methods are fit on two cohorts and asked to predict the third as if unseen. The truth is known.

Held-out cohort	n	Actual	Linear (threshold count)	Ours (sum of P)
2024	71	65	65	63.4
2025	308	274	241	252.7
2026	507	443	478	468.3
Pooled	**886**	**782**	**784 (error +0.2 pts)**	**784 (error +0.3 pts)**

The pooled totals tie, and in each of 2025 and 2026 ours misses by less than the linear count does (21 and 25 students against 33 and 35). Both methods miss the same way in each year, low in 2025 and high in 2026, because the year itself moved (Section 7). The pooled totals are close partly because those two misses cancel, and partly because these held-out cohorts qualify at 88 to 92%, so the region near the cutoff, where the two methods diverge, is thinly populated. The linear method also reaches its total in a way that does not survive inspection.

5.2 What the linear total hides

Over the same held-out predictions, the linear model produced **139 predicted percentiles above 100**, an impossible value, because OLS on an unbounded scale extrapolates through the ceiling. Our model produced **zero** such values, by construction. The linear ceiling behaviour is visible in a single profile (a General, male, PCM, CoE student):

Grade-10 board	Linear predicted percentile	Ours (median)
70	70.6	70.5
80	78.6	81.5
90	88.7	93.0
100	**100.9**	98.5

5.3 The near-cutoff student, where 2028 lives

The 2028 cohort clusters near the cutoffs, so the treatment of borderline students is decisive. Take an OBC, male, PCM, CoE student with a board of 78 and a cutoff of 72:

- **Linear:** predicts 76.9, which is above 72, so it counts the student as a full qualifier (1), with no hint that it is a close call.
- **Ours:** estimates a 0.61 probability of clearing 72, so the student contributes 0.61 to the expected count.

Across a cohort sitting on the line, the linear method's hard rounding of every near-cutoff student to 0 or 1 introduces error that cancels only by luck. On a batch where most students clear comfortably (the held-out years) it cancels well; on an OBC-heavy cohort clustered at the bar (2028) it need not, which is exactly the regime the forecast is for and the regime we cannot yet validate.

5.4 Summary of differences

Dimension	Linear regression	Our method
Outcome scale	raw percentile, unbounded	logit of percentile, bounded to 0 to 100
Impossible predictions	139 of 886 above 100 (held-out)	0
CoE effect	fixed additive term	varies with board (interaction)
School structure	none	random intercept, partial pooling
Year-to-year change	none	between-cohort variance in the band
Per student	one number	a probability
The count	number above the line	sum of probabilities (expected value)
Uncertainty	none	band over cutoff and cohort transfer
Cutoff usage	(in the earlier national build) a regressor and the threshold	the threshold only, drawn as a distribution
Near-cutoff students	rounded to 0 or 1	counted in proportion to their odds

Two points specific to the earlier national build (v1, in the `godrej` folder) rather than to a well-specified in-population OLS. That build fit on the 2024 national cohort and added a flat +18 percentile CoE dummy, so it both transported a national relationship to our students and forced the CoE effect to be constant. It also used the category cutoff as a predictor and again as the threshold, which penalised General students twice and pushed their qualification rate to about 19%; removing that puts General near 46%. Its reported 91% accuracy was on a batch where 90% qualified, so a model predicting "everyone qualifies" would score about 90%; the number was not evidence the method worked.

6. Why the impact paper correctly used OLS

Our CoE-vs-Nodal impact paper reports an average lift and uses ordinary least squares, and that is the right choice there. An average treatment effect sits in the interior of the range, where the 0 to 100 bound does not bite, and an OLS coefficient is a directly interpretable average difference. That paper checked the bound: it re-estimated the lift with a fractional-logit specification and recovered

essentially the same average (about +18 and +27 percentile points against the OLS +16.5 and +27.1), with fitted values exceeding 100 for only 0.3% of students. The bounded machinery was confirmed unnecessary for an average and would only have made the headline unreadable.

The rule is to match the tool to the question. Estimating one average effect calls for OLS with clustered errors and the usual robustness battery. Predicting individuals and counting who crosses a line near the ceiling calls for the bounded, hierarchical, probability-based model here. The mismatch this rebuild corrects is using an OLS levels regression to classify individuals against a threshold.

7. Validation

We validate by leave-one-cohort-out, the same posture as forecasting an unseen year: fit on two cohorts, predict the third, and score the predicted expected count against the actual. Pooled, the model predicts 784 against an actual 782, an error of +0.3 percentage points on 886 students. **The individual years are not that close:**

Held-out cohort	n	Predicted	Actual	Error
2024	71	63 (89%)	65 (92%)	−2 pts
2025	308	253 (82%)	274 (89%)	−7 pts
2026	507	468 (92%)	443 (87%)	+5 pts

2025 was a stronger year than the other two and 2026 a weaker one, and a model that has not seen a year cannot know which kind it will be. That is exactly the cohort-transfer uncertainty the band carries (Section 3.6): a single year's count can miss by several points, so read a forecast for one year as its band, not its centre. (On the July 2026 golden record, with 786 held-out students, every year was within 2 points; the frozen note in `v2/` records those numbers.)

The same run reports calibration by category within each held-out year (validate_loco.py); the one systematic miss is General students in 2026 (76% predicted vs 51% actual, n = 68). The model runs optimistic precisely where the bar is highest, which is why the General projection is flagged as an upper read. Calibration of the count, not classification accuracy, is the proper score for a count deliverable; accuracy on a batch where most students qualify is inflated by the base rate and uninformative.

A sensitivity script (`robustness.py`) re-runs this validation under alternative specifications (linear rather than quadratic board, CoE-only without Nodal borrow-strength, and no school random effect) for anyone who wants to confirm the headline is not an artefact of one modelling choice.

8. Limitations

- **Board-only prior.** There is no Class-11 or Class-12 signal yet. The mock-test refresh will be materially more accurate and will replace these numbers, not merely confirm them.
- **Transfer variance from three cohorts.** The between-cohort term that widens the forecast band is estimated from three points, so it is itself uncertain; read it as a floor.
- **New centres.** With no history they carry no school effect and are scored from the population mean plus the school variance; their real uncertainty is wider than the band conveys.
- **Cutoffs.** The 2027 and 2028 cutoffs are unknown; we carry recent-year variation into the band, but a surprise move would shift the count.

- **M1 (NIT-calibre) is not forecast** from board data; it depends on mock-test performance and is deferred.

9. Reproducibility

Every number here is reproducible against BigQuery and the golden record: `model.py` (the model), `predict.py` (the FY27 and FY28 forecast and the centre panel), `validate_loco.py` (calibration), `compare_linear.py` (the Section 5 comparison), and `robustness.py` (the sensitivity battery). The de-identified per-student extract and the centre panel are published to Cloud Storage for review without warehouse access; see the folder README.