2020 Grades – Problems with the model

Ofqual have published their statistical model for larger cohorts. In brief, the boards will:

- Calculate a grade distribution for each subject in a given centre as a percentage at each grade (actually, cumulative percentage)
- Impose this on the actual number of candidates. Some rounding is then done to get whole numbers at each grade.
- This is then used to assign decimalised "scores". For example, if 4 students are to get grade 8 under this model then they will get decimalised scores of 8.00, 8.25, 8.50 and 8.75 in rank order. (The actual formula is a little more complicated than this).
- Combine the decimalised scores from all centres taking this subject into one list and rank in order.
- Impose a national distribution of grades on this national data set, making sure percentages at each grade are not too high (so now our student who had 8.00 may get moved down into the 7s).
- Assign these nationally adjusted grades as the actual grades for each centre, overriding the calculated centre distribution.
- Go back and adjust grades for smaller groups so that CAG grades come into play for groups under 15 and play a major role for groups of 5 or fewer. This will almost certainly raise some grades for smaller groups, but definitely not for groups of more than 15.

The decimalised scores and national standardisation is not used with subjects where fewer than 500 enter nationally.

The predicted grade distribution is made up of several parts:

c = centre average from the last 3 years (or fewer for some subjects according to the rules)
r = percentage of the 2020 cohort with KS4 data (e.g. 95% gives r = 0.95)
p = percentage that students would have achieved at this centre in the previous 3 years if their grade distribution followed the national value-added picture (this is calculated on a subject level basis, using the prior attainment of the students doing that subject at that centre.)
q = percentage that students at this centre would achieve in 2020 if their grade distribution followed the national value-added picture from the last 3 years.
P = actual percentage assigned to the centre for 2020

The equation is $P = c + r(q - p)$. This can be understood as taking the average from the last 3 years, c, and then adding on a correction factor, $q - p$, which takes into account the GCSE grades of the 2020 cohort in this subject at this centre, compared to the previous 3 years.

I see four obvious problems:

1) The problem of applying a statistical prediction to a relatively small cohort. It is statistically valid to estimate things like averages for cohorts of 30 or so, it is less valid to argue that a cohort of this size should follow a fixed distribution exactly. Most of the individual group sizes are too small for statistical validity (generally, fewer than 5 students for an individual grade category is statistically uncertain).
2) Problem 1 is then compounded by the number of rounding decisions that must be made. Nationally, this will average out with some winners and losers. The problem is that a student who is on a grade border for one subject will usually be on the border for others. They will tend to have all three grades rounded down instead of one or two.
3) The third problem applies to centres like Durham Johnston where value added is historically much better than national. Nationally, weaker students do very badly at A level, here at DJCS weaker students who succeed in Y12 go on to get large value added figures at A level. By introducing a correction factor that uses the national VA figures, our predicted percentages have been lowered at grades A*/A and increased at grades D/E/U.
4) The final problem is that some students will now be moved down a grade to make the national percentages OK, despite perhaps some already having been moved down a little from their CAG grade. The following example illustrates: Suppose we submitted 5 students for grade 8 in a particular subject. The calculated distribution says 4.4, which is rounded down to 4, so one student is moved down into the 7s immediately. The second lowest now gets a decimalised score of 8.00 and then gets moved down into the 7s by the national standardisation. The upshot is that we wanted 5 grade 8s, the initial model said 4.4, but we got 3!

Problem 4 is a particular issue with subjects where most centres are small, examples include further maths and most MFL at A level. Candidates in the small centres will have decimalised scores calculated, based on the formulas specified. These will be combined with the large centres to potentially pull down all grades to fit the national distribution required. The small centres then get their CAGs back at the final stage (probably higher than the statistical distribution would give if the CAGs are "optimistic") but the large centres are now stuck with lowered grades.

Dr Mark Thornber

13 August 2020