

Applying Statistical Process Control Methods to School Performance Data

1. Introduction

The use of Statistical Process Control (SPC) methods in monitoring health outcomes is commonplace¹. This paper considers how such methods might be applied to measures of school performance.

We note that the presentation of statistical significance ('confidence intervals') of school outcome controlling for national variation in school intakes ('risk adjustment') is standard practice in education. Indicators of pupil/student progress and school/FE college value added have been published by DfE and associated bodies for many years.

However, procedures to formally define systems for setting "acceptable" limits of performance at school/college-level which reflect realistic variation around that performance, either by statistics or by decree, are not. DfE has set *de minimis* KS2/KS4 floor standards for schools, but not developed a structure for the bounds of performance or improvement to which it wants all schools (or colleges) to meet as a minimum. .

The use of SPC diagrams - here expressed in 'funnel charts' - may go some way to aid understanding of the significance of school performance generally and in relation to that of an actual - or hypothetical - group (or groups) of schools whose levels of pupil attainment or progress are considered to be of a minimum desired quality.

2. Funnel Charts

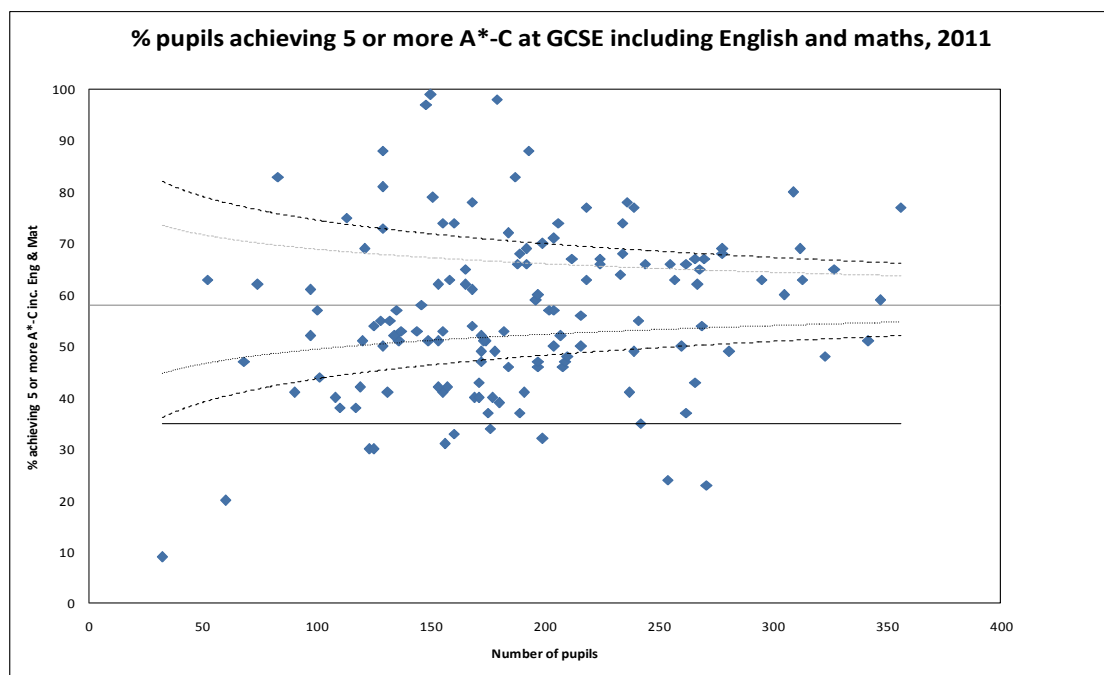
In this paper, we illustrate the use of SPC charts using the proportion of pupils achieving 5 or more GCSEs at grades A*-C (or equivalent) including English and mathematics (5ACEM), the most widely used measure of secondary school performance outturn. We focus our analyses on state-funded mainstream schools (including Academies): in other words, we exclude independent and special schools. In 2011, 59% of pupils in our defined set of schools achieved this standard.

¹ See <http://www.apho.org.uk/resource/view.aspx?RID=39445>

The Funnel Chart approach can be applied to any institutional indicator. In this paper we provide an example where KS4 pupil outcome (aggregated to school-level) has been adjusted for their (individual) KS2 prior attainment and context, and another where we consider (only at school-level) the rate of change in 5ACEM between two recent years.

Chart 1 shows 5ACEM proportions for a random, approximate 5%, sample of schools (n=138). The horizontal axis is set at the 5ACEM national average (59%) and the dotted lines either side indicate 2 and 3 standard deviations (SD) of the *national* variation in outcome. An additional horizontal line denotes the Government’s notional floor performance standard² of 35%.

Chart 1: Funnel Chart of Secondary School Performance 2011, Sample of 5% of Schools



The chart shows large numbers of schools above and below the 3 SD control limits. In SPC jargon we observe a large number of schools that are “out-of-control”: that is, their 5ACEM outturn given the national variation in that measure can be considered an ‘outlier’.

² Schools are only considered to be below the floor target if rates of expected progress in both English and mathematics are also below average

As school outcomes are heavily correlated with prior attainment, schools with low or high ability intakes are more likely to be plotted below or above control limits, as indicated in Charts 2 and 3. These charts provide examples for ordered groups of schools. We return - in Chart 7 – to the impact in terms of ‘outliers’ when an adjustment is made to *pupil-level* KS4 outcome given their specific prior attainment and context aggregated to school-level. This ‘value added’ calculation is of the sort which DfE publish currently.

Chart 2: Schools in the random 5% sample and in the lowest 20% of schools nationally for prior attainment

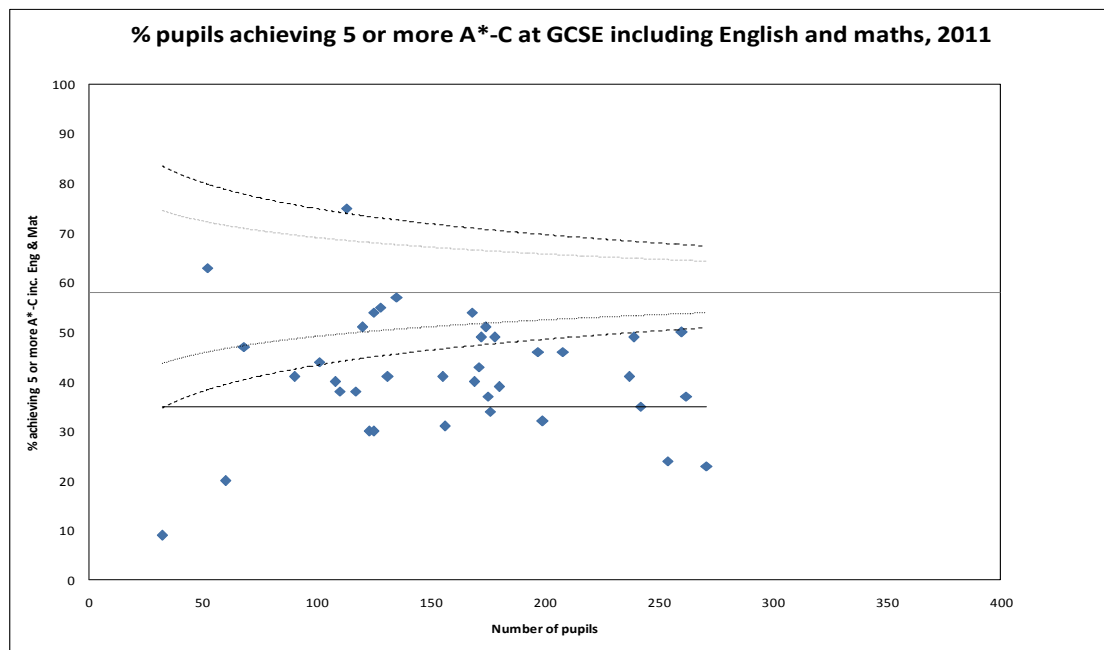
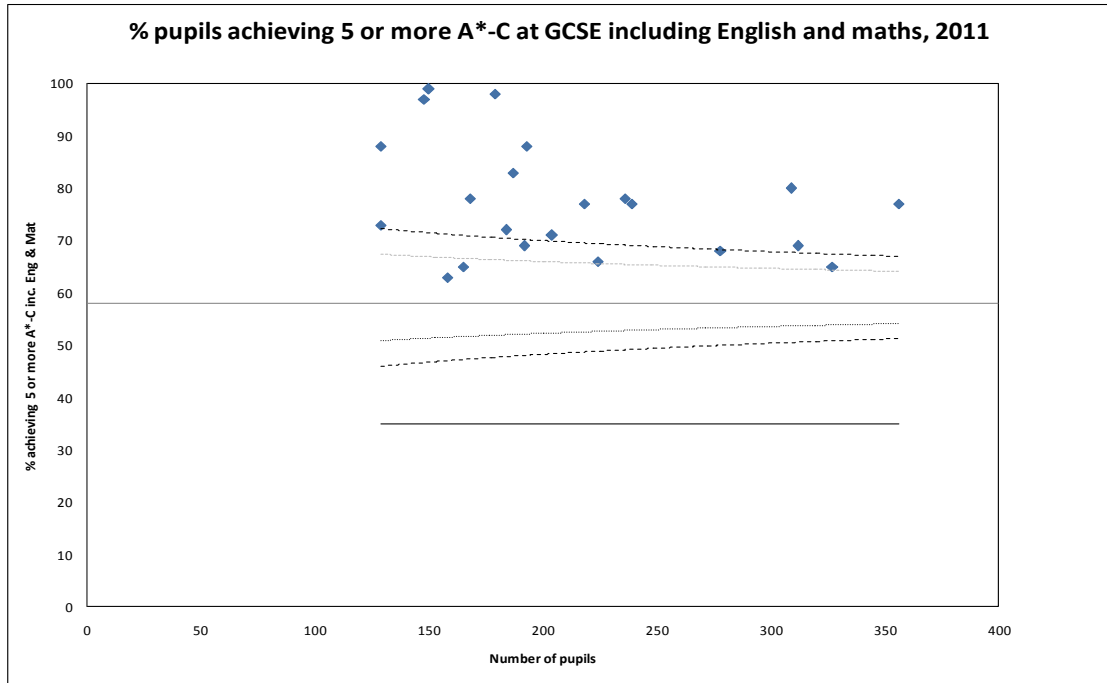


Chart 3: Schools in the random 5% sample and in the highest 20% of schools nationally for prior attainment



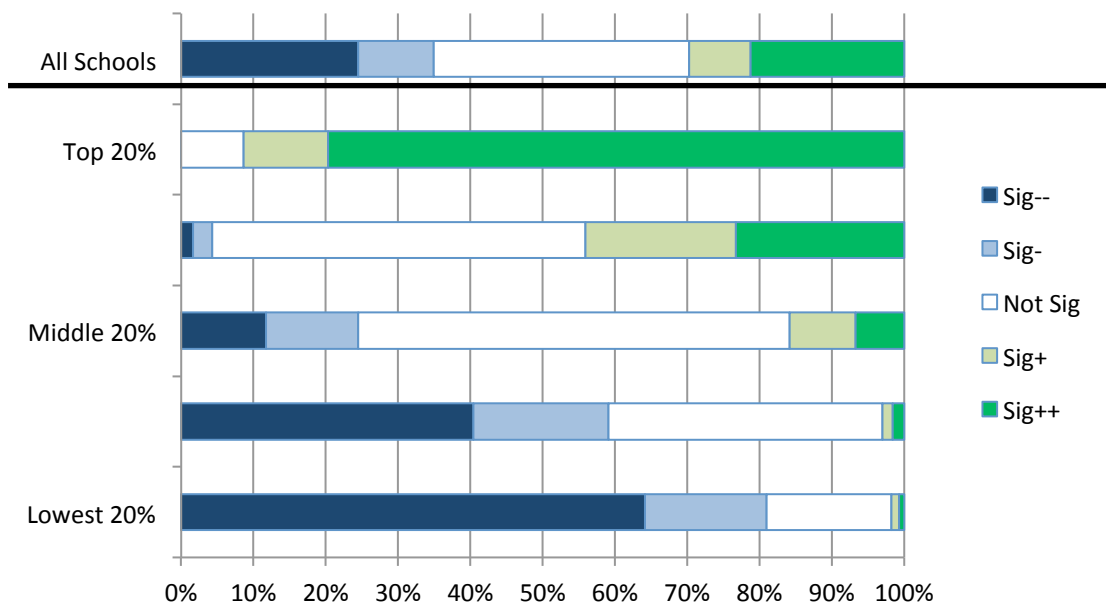
3. Displaying Statistical Significance

Although funnel charts (as a means of presentation) have not been widely used in education³, schools are used to tests of statistical significance through RAISEonline and FFT Live. Traditionally, pupil and school performance has been shown on a three point significance scale (significantly below, not significant, significantly above) with a 95% confidence level used to determine boundaries.

Charts 1-3 above show a five point banding of school performance, additionally indicating significance at the 99.8% level. This is summarised for all schools nationally in Chart 4, according to quintile KS2 prior attainment band.

Overall, 21% of schools are in the Sig⁺⁺ category (significantly above, 99.8% confidence level) and 24% in the Sig⁻⁻ category (significantly below, 99.8% confidence level). Almost half of all state-funded schools are, in terms of the outcome measure and significance categorisation used here, either above or below “extreme limits”.

Chart 4: Significance States for the Percentage of Pupils achieving 5 or more A*-C Grades at GCSE (or Equivalent) 2011, All State-Funded Mainstream Schools in England by KS2 Prior Attainment Band



³ Other forms of SPC chart have been used, for instance in measures of value added from GCSE to A level (e.g. ALIS, LAT)

Typically, tests of statistical significance used in the education sector are based on comparison to the *national* average and the distribution of performance of *all* state-funded mainstream schools.

The SPC approach, however, does not use the performance and its distribution of all appropriate schools but instead defines a group of schools that are deemed *a priori* (by some agreed means) to be “in control” (as then defined) and which taken as a group provide a benchmark range (a ‘target’). In Section 4 we consider two approaches to defining such a set of schools.

4. Overdispersion

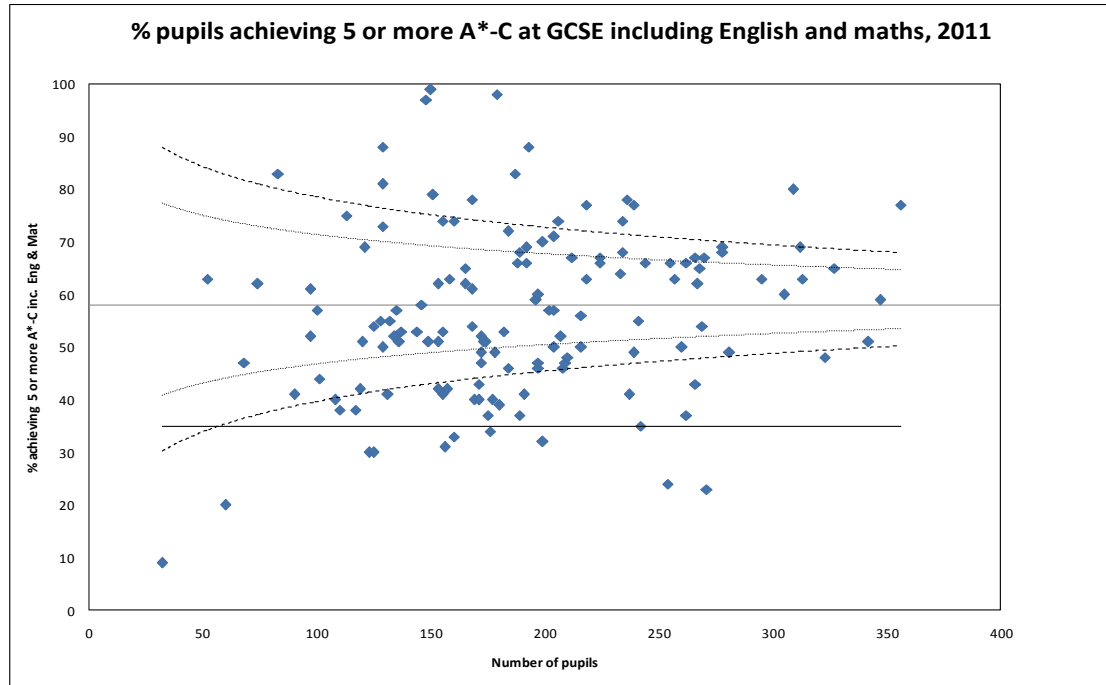
The above charts show large numbers of schools at which the observed process - the proportion of pupils achieving 5 or more GCSE A*-C including English and maths - is, superficially at least, ‘out-of-control’. This observation is referred to as ‘overdispersion’ and is often a consequence, for example, to reflect adequately the factors which affect pupil progress (referred to as “casemix” in the health literature) that vary between pupils within schools (for example, prior attainment). We consider the impact of such ‘risk adjustment’ in Section 5.

We present a *general* method of handling ‘over dispersion’ in Chart 5, where we show a revised version of Chart 1 in which we have applied a Variance Inflation Factor (VIF) to the control limits.

For the purposes of illustration only, we have defined a set of “in-control” schools as those in the second and third quartiles of the distribution of schools on our measure of the proportion of pupils achieving 5 or more GCSEs at grades A*-C (or equivalent) including English and mathematics to which the VIF has been applied.

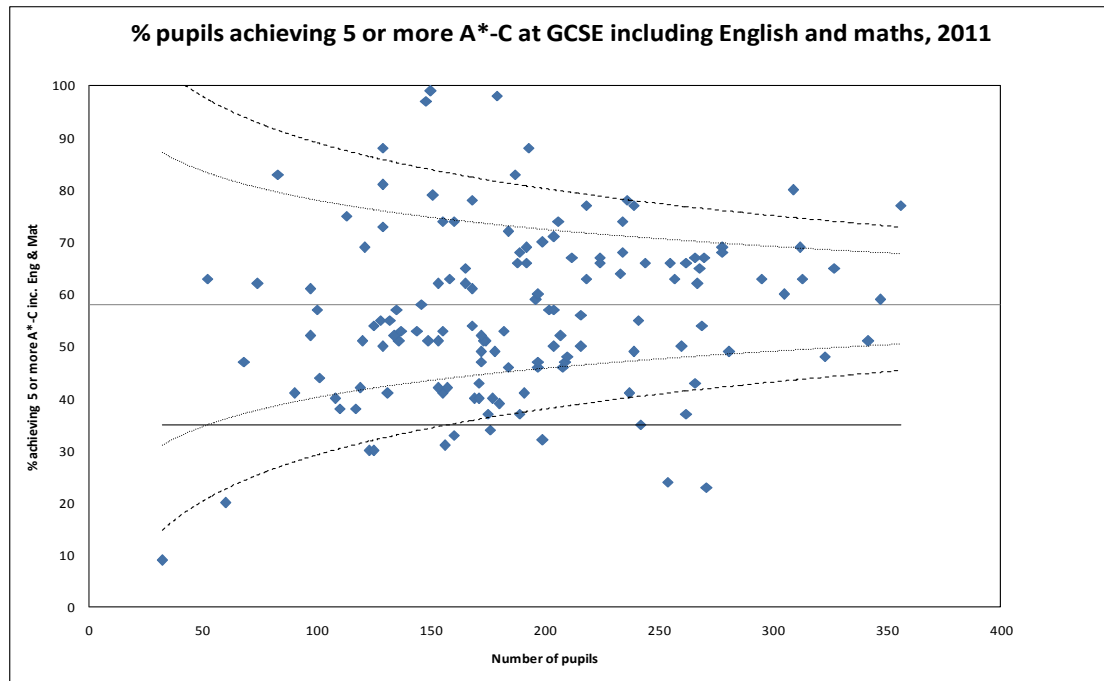
The VIF ‘stretches’ the distribution of the outcomes *of all* schools such that, compared to Chart 1, the control limits have been moved further away from the *national* mean performance leading to a reduction in “out-of-control” schools.

Chart 5: Funnel Chart of Secondary School Performance 2011 Accounting for Over-Dispersion through a Variance Inflation Factor, Sample of 5% of Schools



An alternative approach, again for illustration only, is to use the DfE-defined floor target to 'fix' the lower control limit and adjust the other control limits relative to it. Around 20% of schools have cohorts of 150 or less at Key Stage 4. In Chart 6 we have set the lower control limit to 35% for a school with 150 pupils and, as in Chart 5, we observe far fewer "out-of-control" schools.

Chart 6: Funnel Chart of Secondary School Performance 2011 Accounting for Over-Dispersion through a Policy Threshold, Sample of 5% of Schools

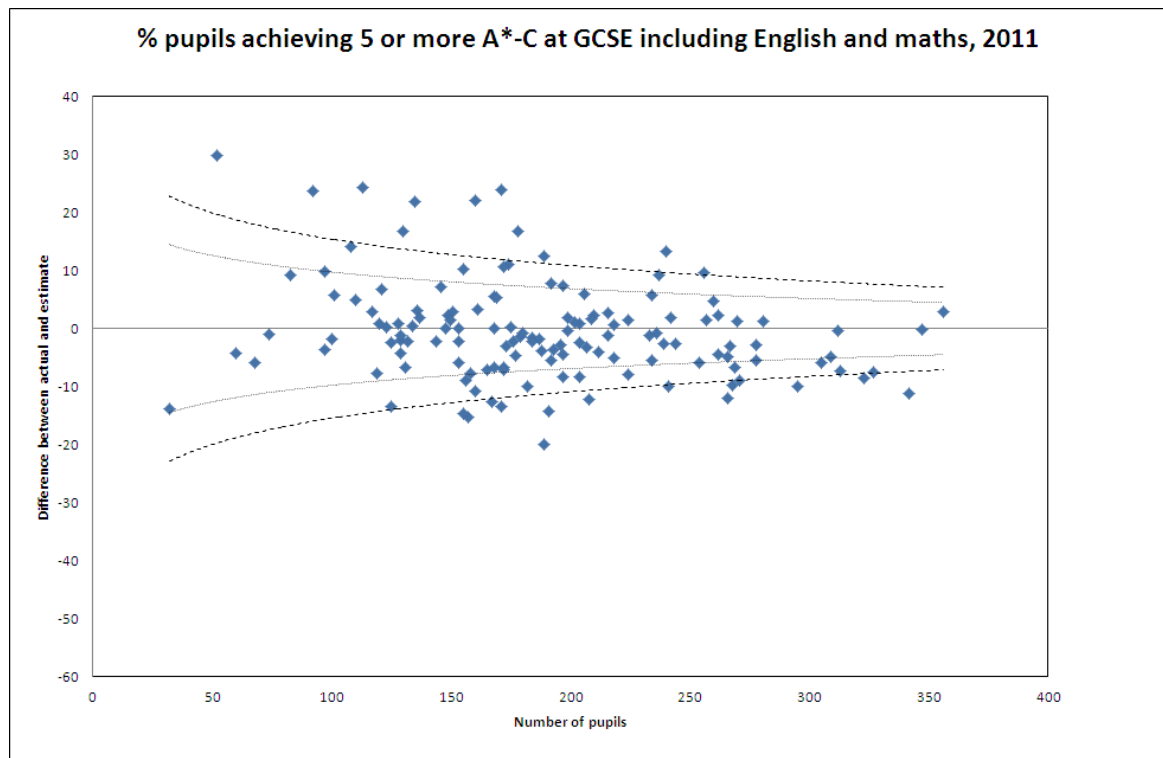


5. Controlling for Intake (Risk Adjustment)

As school-level outcomes are heavily correlated with pupil prior attainment and contexts (for example, socio-economic deprivation, ethnicity, special educational needs), we show in Chart 7 the consequences of controlling for variation between schools in pupil intakes.

In this chart, instead of displaying the proportion of pupils achieving 5 or more A*-C grades at GCSE including English and maths, we display the differences between this proportion and an estimate based on pupil prior attainment and context. We apply *national* control limits as we did in Chart 1 but we note that Chart 7 is centred at zero, indicating a school that has performed in line with expectation.

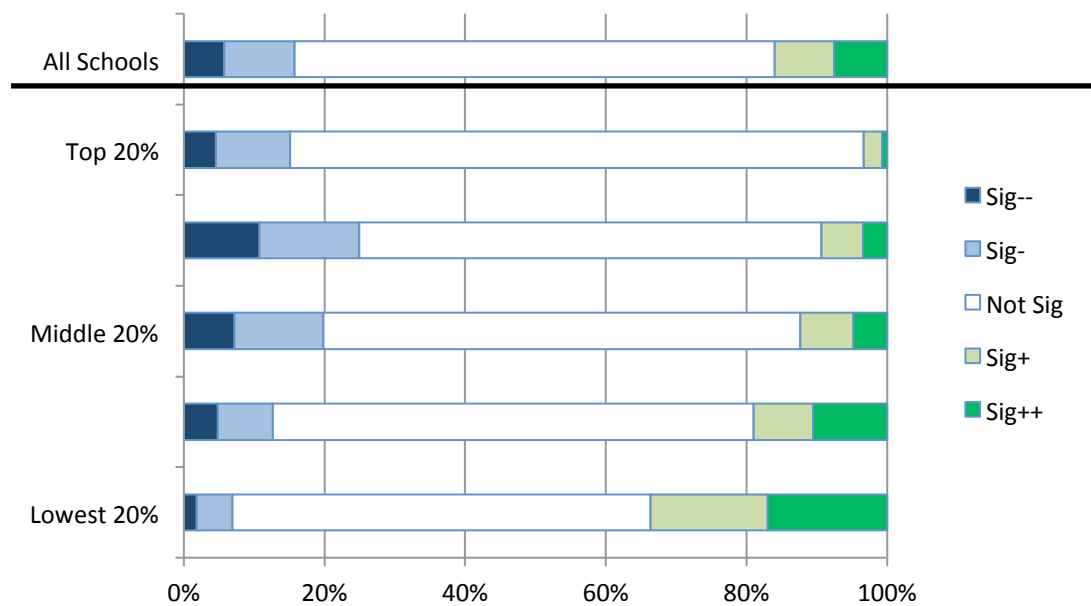
Chart 7: Funnel Chart of Secondary School Performance 2011 Adjusted for Pupil Prior Attainment and Contexts, Sample of 5% of Schools



Compared to raw outcomes (Chart 4), a much lower proportion of schools are above or below extreme limits (Sig-- or Sig++) when prior attainment and pupil contexts are taken into account (Chart 8).

Although there is minimal bias at pupil-level in outcome with respect to prior attainment and context (though the evidence is not shown here), we observe that schools with the lowest levels of pupil prior attainment are proportionately more likely to be found in the Sig++ category, and that schools with pupils in the third and fourth quintiles of prior attainment are more likely to be below the lower control limits. We do not speculate in this paper why such findings may occur.

Chart 8: Significance States for the Percentage of Pupils achieving 5 or more A*-C Grades at GCSE (or Equivalent) 2011, All State-Funded Mainstream Schools in England by Prior Attainment Band

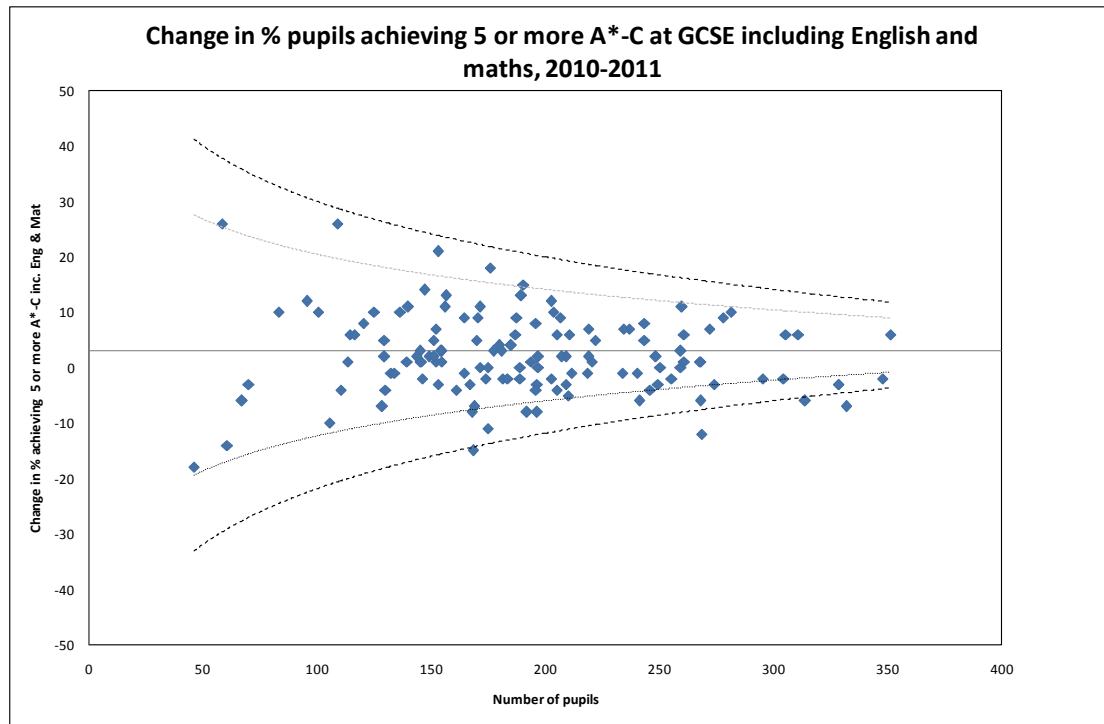


6. Year-on-Year Change

The SPC approach could be used to monitor year-on-year changes in school performance (Chart 9). We observe that, on average, state-funded mainstream secondary schools improved the percentage of pupils achieving 5 or more A*-C grades at GCSE (or equivalent) including English and mathematics by 3 percentage points between 2010 and 2011.

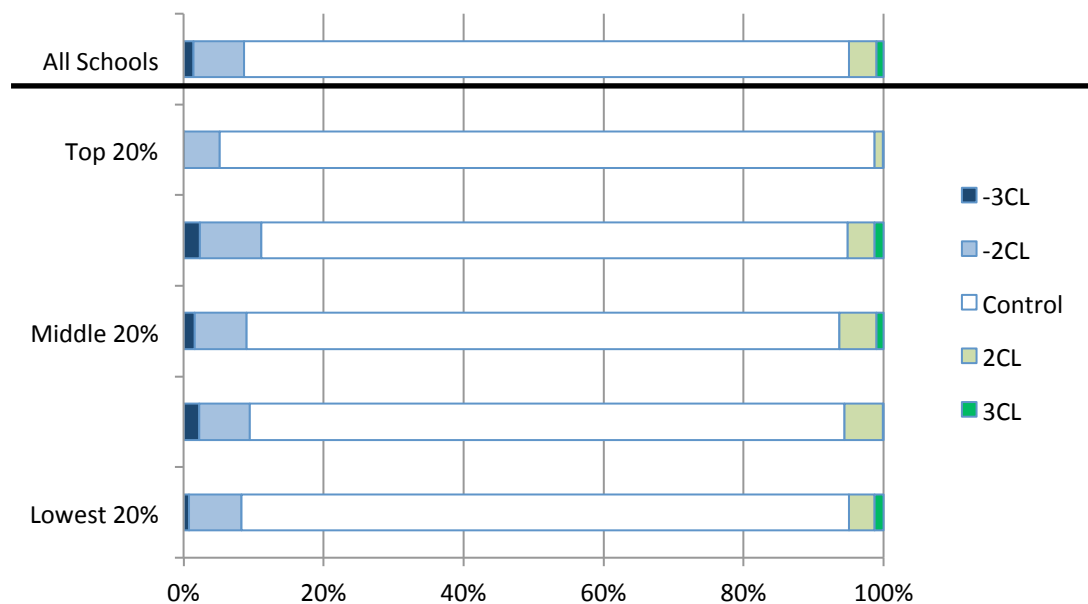
For illustration, control limits have been established based on the change amongst a group of schools deemed to be “in-control” - those in the 2nd quartile for improvement 2010-2011 with a distribution ranging from 2 to 6 percentage points.

Chart 9: Funnel Chart of the Change in Secondary School Performance 2010-11 for a Random Sample of 5% of Schools



We observe few “out-of-control” schools, which implies that the rate of change (at school-level) of pupils’ achieving the DfE headline measure of performance is not unacceptably different across schools even though many will have very different *absolute* levels of outcome. Chart 10 shows that just 1% of all state-funded mainstream secondary schools were above or below extreme limits in 2011 using this method.

Chart 10: Control Limit States for the Change in Percentage of Pupils achieving 5 or more A*-C Grades at GCSE (or Equivalent) Between 2010 and 2011, All State-Funded Mainstream Schools in England by Prior Attainment Band



7. Conclusions

Funnel charts are an attractive and simple way of presenting comparisons of school performance. These can show any school indicator whether this relates to measures of raw outcomes, measures of outcomes adjusted for pupil intake (prior attainment and contexts), or to monitor year-on-year change in outcomes.

The 'control limits' which indicate levels of unacceptably different levels of performance can be based either on national distributions (as current significance tests currently are), or on those of schools where performance (or progress) is deemed worthy of having 'control' status.

Indeed, the 'control performance distribution' which sets the control limits need not be based on *any* specific group of schools, in which case they can be said to be 'pre-defined'. However, there is no precedent within the education sector of defining such a distribution, and a process to determine such a group or groups (and changes to them over time) could well be involved and lengthy, requiring consensus between schools, politicians and inspectors.