School Roll Projections Methodology

How the GLA's school roll projections are produced

There is no single accepted method for projecting school numbers and London boroughs have recently faced major challenges in providing places to meet a growing child population. Harrow, along with the majority of other London boroughs, commission's school roll projections from the Greater London Authority's (GLA) School Roll Projections Service. The GLA provides the baseline projections to which local knowledge is applied to make reasonable adjustments in line with pressure at Reception, Year 7 and other school year groups.

Summary methodology

The school roll projection model creates a roll projection for each school based on the GLA population projections of the wards where its pupils live.

For each ward of residence in London, National Curriculum (NC) year (R to 11) and sex, the proportion of children of the corresponding age attending each mainstream state school is calculated. These proportions are carried forward as the pupils age through the school in the years being projected.

For new pupils entering a school in future years, for example at reception, there is currently no information on what proportions of the residents will attend the schools. In this case the proportions are calculated as averages over the latest years of actuals, with 4 being the default number of years used (2015, 2016, 2017 and 2018), but there is the option to use a different number. The same approach is used at years 7 and 12, even if the school is an all through school as it is assumed that there will be significant changes in the cohort at this point.

Where a school has opened recently, the proportions for its new intake are determined by averaging over all years used for calculating new intake (default of 4), even if it was only opened, for example, last year. This means that new schools will show lower projections going forward. The reverse is true for schools which closed within the back series – they will still provide a contribution to the projection going forward if they were open at any point during the back series. As a consequence, results for individual schools that have opened or closed within the back-series period may now appear counterintuitive, but the results are expected to be more robust at borough or planning area level as they take into account all of the pupils who were, or were not, present in schools in those areas during the back-year period.

For the current round year (2019), the actual school rolls submitted specify roll numbers but we have no information on wards of residence of the pupils. For this year, the number of pupils from the roll attributed to each ward are estimated by averaging over the previous years' patterns, with the default being 3 years (2016, 2017 and 2018), and scaling to ensure that the total numbers at each school for each age and sex match the submitted rolls.

The rolled forward and calculated new intake proportions for future years are then applied to the population projections to give projections of the number of children on roll by school by age and sex. Due to lower retention rates, sixth form projections are calculated using a survival ratio as the cohort ages through sixth form. School level projections are then aggregated to planning areas and borough totals.

Options for running the model

Appendix A describes the SRP model and the methodology behind the numbers of years of data used for **ward distribution of current roll** and **new intake** options in more detail.

The default is to use 3 years of back-data for both as well as the current 2019 roll for the new intake (known as the **3/4 option**).

The **3/4** option uses:

- **three** years of past detailed flow data (2016-2018) to define relationships between ward of residence and school attended for the 2019 roll
- **four** years of school-level rolls (2016-2019) to calculate the size of the new intake.

This gives a result that incorporates several years of past data to smooth out fluctuations in the data in terms of wards pupils come from and number of pupils in the new intake, giving more stable results than fewer years data.

The **3/1** option uses:

- **three** years of past detailed flow data (2016-2018) to define relationships between ward of residence and school attended for the 2019 roll
- **one** year of school-level rolls (2019) to calculate the size of the new intake.

This setup aims to smooth fluctuations in the underlying patterns of pupil movement, while reflecting only the most recent data in terms of overall number of pupils on roll.

In some situations, there may be a case for using just one year of data to estimate both the ward-school relationships and the new intake (a 1/1 option projection). This can be when patterns have changed in the most recent year and you believe they will continue into the future. However, the benefits of using only the most recent patterns can be outweighed by the issues of 'noisy' data. For example, in the latest year there may be a ward where there is no intake from that year but there is normally. This will result in the school roll projections not incorporating future population changes in that ward.

Which population projection to use?

School roll projections can be run based on the following population projection variants:

- **Borough Preferred Option (BPO).** This is the default option, but is dependent on boroughs providing development data to be incorporated into the projections.
- **The latest GLA ward level Strategic Housing Land Availability Assessment (SHLAA)** -capped AHS (average household size) population projections for all wards in your LA. This is the option if boroughs have not provided their development data.

Influences of population projections

The biggest driver of projected future school rolls is the population projections for schools' catchment areas. The underlying factors include:

- **Development**
  The amount of development projected in a LA will affect that authority's population projections and in turn its school roll projections. More development generally means that

the LA will attract more people and its population will therefore rise. If population increases, there will consequently be more children and so school roll projections will also rise.

LAs should assume that significant changes in assumed development will be accompanied by corresponding changes in projected rolls. If LAs are unsure what development assumptions have been used in the past, the GLA is able to provide this information.

- **Births**
  The number of births in an area will have a direct effect on the number of children on roll four years later. 2012 saw the highest number of births in London with these children starting school in either academic year 2016/17 or 2017/18 depending on when in the year they were born. Many areas have seen a fall in birth numbers since and this has led to subsequent projections of future births and therefore rolls, being correspondingly lower.

- **Migration**
  Migration, both from other areas within the UK and internationally, can significantly influence population projections. However, the most recent migration data available for the 2018 model run covers the period mid-2015 to mid-2016. Subsequent changes to patterns are not accounted for in the projections.

  The GLA has created an Excel based dashboard that allows boroughs to see in-, out- and net flow of children to/from their LA from elsewhere in London. It is available to download from the London Datastore and will be updated annually:
  http://data.london.gov.uk/dataset/internal-migration-flows-school-age-children-visualisation

  ONS releases both mid-year international and internal migration data by single year of age and sex at the end of June each year. The former is released as part of the mid-year components of change and the latter as part of the internal migration estimates series.

  Analysis of trends in the mid-year estimate series, and comparison to administrative sources, suggests that there has been an over estimation of the number of 0-4 year olds in London as a whole since 2011 in the official data. We believe that this is the result of underestimation of international out migration flows in the young population. Following this analysis, the GLA has taken the decision to revise the estimates of migration and population used as the basis for projections. See Appedix B for details.

What the School Roll Projection Model does and does not take into account

School closures
There is currently no provision in the model to take account of planned school closures.

New schools
There is currently no provision in the model to include planned new schools that have yet to open. Where a school has recently opened, it will not have existed at the 2017 January census so we have no information on the wards from which the school draws its pupils. In this case, it is assumed to draw its intake from across the local authority as a whole.

A new school is assumed if the DfE number given in the actual rolls (or its corresponding 'previous DfE number') cannot be matched to a DfE number in the national pupil database extract that the model uses.

Children who live outside London
The base population projections for areas outside of London are at local authority level. Therefore, pupil flow data for children resident outside of London is aggregated to LA level rather than ward level.

The City of London
The City of London is treated by the model as one entity to match the population projections used by the model.

Age to NC year
Boroughs should provide all data by national curriculum year. However, the population projections refer to children's age instead of year group. To line up the population projections with the school roll data, the model converts age to year group (Reception <-> age four, Year 1 <-> age five, etc.).

Cross border mobility
The model takes cross border mobility into account explicitly as it uses information about pupils' home wards from the national pupil database (NPD). The detailed flow data (i.e. assumed flows for the whole projection period for an authority's schools) can be provided on request.

Child yield
Child yields are not incorporated into the model in the same sense that many people think about them. The borough projection models contain assumptions about the age and gender characteristics of migration flows between locations. The difference in characteristics between in- and out-flows defines the resulting population age structure. These migration flows are influenced by assumed development in the model and new development tends to be associated with increased numbers of children in the population.

Limitations and considerations

The models are simplifications of complex real-world processes. They project forwards relationships taken from past data, so they are projections and not predictions. Many factors are not explicitly taken into account and LAs should be aware of the limitations of the models when interpreting results.

Among the factors that the current models do not account for are:

- Changes to future patterns of migration;
- Changes to future planned development;
- Changes to parental preferences for schools;
- Constraints to the capacity of schools;
- Schools opening/closing in neighbouring boroughs
- Future changes to provision, e.g. schools opening or closing, or changes to the characteristics of schools; and
- Future changes to the character of local areas, e.g. gentrification or the impact of welfare reform

Appendix A. School Roll Projection Model

Introduction
Not all children attend school in their borough of residence. This is particularly the case in London where the geographic size of local authorities is relatively small and transport networks enable children to travel beyond their borough boundary. Additionally, for children who live close to a borough boundary, their closest school may be in a neighbouring authority.

To create school roll projections based on the ward level population projections it is necessary to know where pupils come from. The National Pupils Database (NPD), based on the School Census, provides home ward and school attended for all pupils attending state funded schools.

As can be seen from the summary diagram at Appendix B, there are four key stages to the projections.
- For the wards that the school draws pupils from, estimate the proportion of the ward attending the school in the current academic year by NC year and sex. (**ward distribution of current roll)**.
- By NC year and sex estimate the proportions from each ward moving forward **(aging).**
- Estimate the proportions from each ward for future new intake **(new intake)**.
- Aggregate to obtain school, planning area and Borough projections **(aggregating projections)**.

Proportion of ward population attending a school
For each ward in London, national curriculum (NC) year, and sex, the proportion of children attending each mainstream state school is calculated as follows: Divide the number of pupils of that sex who attend the school in that NC year who live in the ward by the total number of children of the equivalent age group and sex who live in the ward (the base population). NC year is matched to age at the beginning of the school year. For example, reception pupils are matched to children from the population projection who were 4 years old at 1st September 2017.

$$Pupil\ ward\ to\ school\ flow\ proportion = \frac{number\ from\ home\ ward\ attending\ that\ school}{home\ ward\ base\ population}$$

Ward distribution of current roll
For the 2019 cohort, boroughs provide the numbers on roll for each school by age and sex. As the NPD data is not available yet to obtain the home ward information for the new intake years in 2018, the average home ward patterns over a number of previous years are taken as a proxy. The default being three years (2016–2018) with options for a different number e.g. only the most recent year (2018). These averaged patterns are scaled to ensure that the numbers across all wards equal the number on roll for each school, NC year and sex as submitted in the 2019 rolls.

Aging
There is no information on the proportion of pupils from each ward beyond the years for which we hold NPD and pupil on roll data. Beyond this point the proportion of pupils from a ward is carried forward as children age. Therefore, the proportion of year 3 pupils living in ward *a* and attending school *z* in 2019 is the same as the proportion of year 4 pupils living in ward *a* and attending school *z* in 2020. Figure 1 shows the aging of proportions through the projection period. It can also be seen from Figure 1, as projections move further forward, proportions for new intake cohorts need to be estimated. In the special case of aging from the NPD 2018 to the current roll year of 2019, the proportions are scaled after aging to ensure that the total numbers at each school for each age and sex match the submitted 2019 rolls.

Figure 2 shows aging of proportions, with the proportion of the latest intake of reception pupils carried forward until, in this example, 2024 and beyond when this proportion is applied to all year groups from that ward.

Figure 1. Aging of primary school pupils' resident in one ward

NC year	NPD 2015	NPD 2016	NPD 2017	estimated 2018	Aging 2019	Aging 2020	Aging 2021	Aging 2022	Aging 2022
R	3/5	2/5	1/3	1/4	?	?	?	?	?
1	3/5	3/5	2/3	1/2	1/4	?	?	?	?
2	2/5	4/5	4/5	3/4	1/2	1/4	?	?	?
3					3/4	1/2	1/4	?	?
4						3/4	1/2	1/4	?

Figure 2. Aging of primary school pupils' resident in one ward

NC year	NPD 2015	NPD 2016	NPD 2017	estimated 2018	Aging 2019	Aging 2020	Aging 2021	Aging 2022	Aging 2022
R	3/5	2/5	1/3	1/4	1/4	1/4	1/4	1/4	1/4
1	3/5	3/5	2/3	1/2	1/4	1/4	1/4	1/4	1/4
2	2/5	4/5	4/5	3/4	1/2	1/4	1/4	1/4	1/4
3					3/4	1/2	1/4	1/4	1/4
4						3/4	1/2	1/4	1/4

New intake

The GLA do not have information on which wards pupils entering a school in its intake NC year (reception year in the above example) will come from in the future, and what proportion of each ward's population will attend. To estimate this, the proportions of pupils in that NC year attending the school from each ward in previous years are averaged to give an estimated proportion to use for calculation of new intake in future years. The default number of years to average over is 4 (2016 - 2019); 3 years from the NPD and the estimated current year proportions. As noted below there is the option to use a different number of years. The same approach is used at NC years 7 and 12, even if the school is an all-through school as it is assumed that there will be significant changes in the cohort at this point.

This year there has been a change in the way that the model treats schools which have opened or closed within the years used to calculate future new intake. Where a school has opened recently, the proportion for its new intake is calculated by averaging over all years used for calculating new intake (default of four), even if it was only opened, for example, last year. This means that new schools will show lower projections going forward. The reverse is true for schools which closed within the back series – they will still provide a contribution to the projection going forward if they were open at any point during the back series. As a consequence, results for individual schools that have opened or closed within the back-series period may now appear counterintuitive, but the results are expected to be more robust at borough or planning area level as they take into account all of the pupils who were, or were not, present in schools in those areas during the back-year period.

Sixth form

It was found that projecting the proportions forward in the sixth form years over-projected the numbers of pupils in NC years 13 and 14. For this reason sixth form projections are calculated using a survival ratio as the cohort ages through sixth form. For example, for each of the projection years, the number of year 13 pupils in a school in that year is a fixed proportion of year 12 pupils at the school the year before. Year 12 pupils are always treated as new entry even if the school also includes younger years. Projections for year 12 pupils are calculated using the methodology outlined in the New Intake passage above.

Aggregating proportions

For each NC year and sex, the proportions of each ward attending a school is projected, then multiplied by the ward population projection to obtain the number from that ward attending the school. This is summed across all wards that pupils are resident in to obtain the school projection for a particular NC year and sex. Schools are summed to planning area and Borough totals.

Appendix B. School Roll Projection Model summary diagram
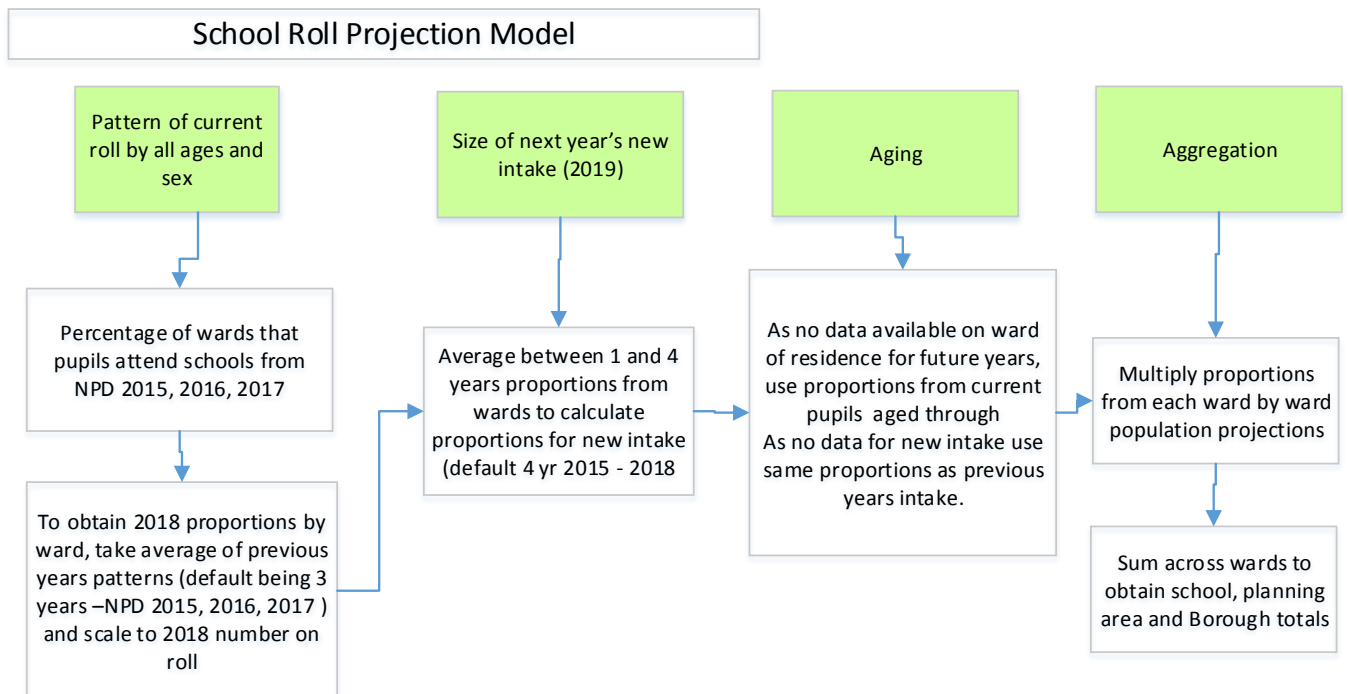
Appendix C. Population projection model

As noted above, the GLA's ward level population projections drive the school roll projection model.

The population projections incorporate annual birth, death and migration data to mid-2017. Future birth trends in fertility and mortality are based on the principal assumptions from ONS's 2016-based National Population Projections (NPP) for England. Household formation rates are taken from the 2014-based DCLG (Department for Communities and Local Government) subnational household projections. Past dwelling completions are taken from the London Development Database. Assumed future housing trajectories are derived from the 2016 Strategic Housing Land Availability Assessment.

BPOs will be produced on two different models. The first set will follow the same basic methodology to that used to produce small area population projections in past years. However, the housing trajectory input for this borough is replaced by bespoke ward level housing

development data provided by the borough. Full methodology papers for the housing-led and small area projection models used to produce these outputs can be found on the GLA's London Datastore: https://data.london.gov.uk/dataset/projections-documentation. This set of projections will provide continuity with previous years.

The second set will use a methodology which has been updated to better account for housing characteristics and local context. The GLA has now improved the detail and range of data that can be included in the model, such as: number of bedrooms and tenure of existing and future development, house prices, and transport accessibility.

Domestic migration variants

This year, the GLA produced projections based on three different scenarios of domestic migration patterns. The goal in doing so was to try and provide users with suitable variants for use in school place planning work. The GLA chose different levels of domestic migration as the basis for the variant projections as this has the most potential to significantly change the number of children requiring places in the short term.

Outmigration rates of children from London fell dramatically following the financial crisis, as the effects on the housing market prevented many families who would otherwise have left London from doing so. As the housing market recovered, so too did the outflow of families and children. The latest estimates indicate that the proportion of children leaving the city each year is now back at levels we saw before the crisis.

The financial crisis showed that domestic migration of families could change significantly in a short period of time and therefore represents an important source of uncertainty for planners. To try and capture a realistic range of possible outcomes the created three scenarios based on previously observed patterns of migration.

- **High** – these use domestic migration rates based on the last few years only. These rates reflecta relatively high level of movement of families. The same rates are used for the whole projection period. For most areas, this will give a lower number of children in the population as families tend to move radially outwards from Inner London to Outer, and Outer London to the home counties and beyond. The effect is variable between areas though – as for example boroughs in Outer London will tend to gain additional inflows from Inner London that may offset their own outflows.
  This scenario can be thought of assuming that current levels of outmigration are the norm and will continue.

- **Medium** – these start off using the same rates as for the high variant – i.e. based on the very recent past. However, rates gradually move back towards the long-term average over the first ten years of the projection.
  Basing near-term behaviour on the recent past and moving back towards longer term trends is a common forecasting approach and this variant can be considered the 'standard' variant.

- **Low** – these assume an immediate drop in outmigration– using rates based on the few years following the financial crisis. As with the medium variant, the rates return to the long-term average over the first ten years.
  This scenario can be considered as representing a shock event and sudden slowing of the housing market followed by a gradual recovery. As it assumes a significant departure from recent patterns, it is primarily suitable for use in 'stress-testing' plans.

The periods on which these rates are based had both different levels of absolute migration and different age structures to the flows, with the movement of families being disproportionately affected by the financial crisis.

When applied within the housing-led model, users should be aware that the total population may not differ greatly between the variants. This is because the assumed housing stock is the same in each case and this determines the size of the population and therefore total net migration. While the total population may be similar in each case, the proportion of children in the population will still vary due to the different age structures of the flows, as well as different balance of in/out & international/domestic flows.

Higher projected population

For many areas, the latest projections give higher populations than previous rounds of outputs. Given the potential implications for planners, we felt it important to provide some additional information about this change so that users can appropriately interpret the results.

This year's housing-led projections are the first to incorporate household formation rates from ONS's 2016-based household projections. These replace rates from DCLG's 2014-based household projections as the default choice in the models. The methodology ONS used to project household formation gives higher average household sizes than the equivalent DCLG approach. When used in the population models, the result is typically to increase the number of people assumed to fit within a given number of homes.

The effect of this change in the household formation rates on long term population size might typically be of the order of five percent. However, it is important to note that where a large amount of new housing is expected, the difference in the *additional growth* in population is likely to be much larger. In this example, a five percent difference in the projected *population* is equivalent to a twenty-five percent difference in the population *growth*.

Choice of household formation rates

That the level of growth is sensitive to the household formation assumptions, makes the choice of rates more critical than in previous rounds of projections. This choice is complicated by the fact that future household formation trends are inherently uncertainty.

The move to using the ONS rather than DCLG rates in the projections was informed in part by how well each to have aligned with estimated change since 2011. These comparisons are made as part of the standard set-up of the model and through this process it was determined that the ONS rates have been, overall, more consistent with recent data on population and housing stock change. Analysis by ONS comparing different household projections with estimated numbers of households arrived at a similar conclusion.

Though the use of the ONS rates have been used as the default in the model, it is important to emphasise that there is no definitive 'correct' choice. The BPO projections are offered as a service to the boroughs and there is some flexibility in how the models are configured.