

Supplementary Material for *Implications of the school-household network structure on SARS-CoV-2 transmission under different school reopening strategies in England:*

Additional Figures:

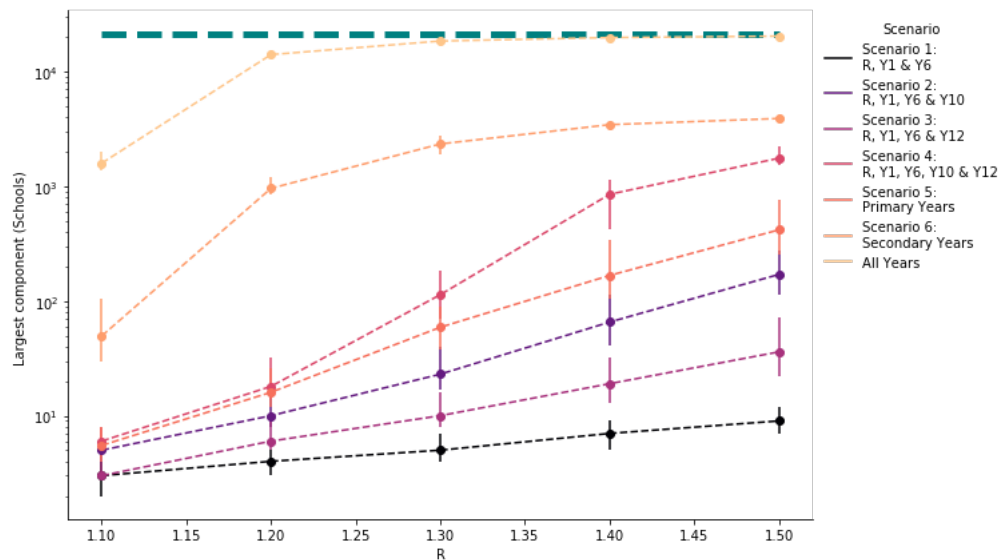


Figure S1 The number of schools in each largest connected component of the binary transmission networks generated from transmission probability networks for school reopening scenarios.

Verification of address matching:

Approximately half of students had a unique postal address code for their residence. We evaluated our address matching against this code for the students that had one. Table S1 shows the proportion of households that were correctly identified as a single household (99.78%) and the proportion of households that were mistakenly merged with either 1 or 2 others 0.22% and 0.001% respectively).

Table S1. Household addresses with a UK standard unique property reference number (UPRN), against households identified by postcode and address matching. 53% of households (2.6m of 4.9m) had a UPRN.

	Number of true unique households by UPRN (N=2627601)
1 household correctly identified	2621685 (99.78%)
2 households incorrectly merged	5902 (0.22%)
3 households incorrectly merged	14 (<0.001%)

Sensitivity analysis:

Analysis run with $q=0.3$ (double within-household transmission probability) and $q=0.08$ (half within-household transmission probability)

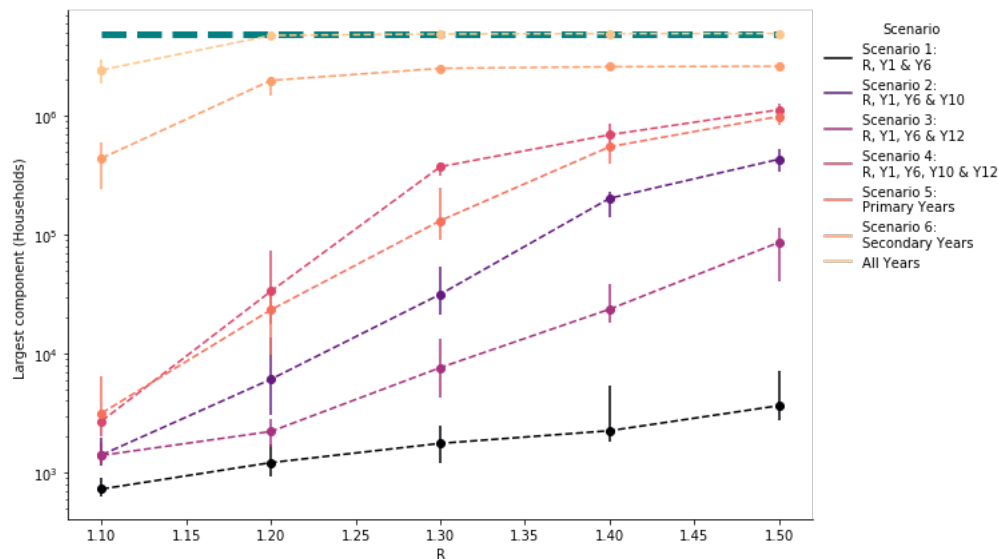


Figure S2 (Sensitivity analysis with $q = 0.3$) The number of households with children attending a school in each largest connected component of the binary transmission networks (estimated potential outbreak cluster size) generated from transmission probability networks for school reopening scenarios.

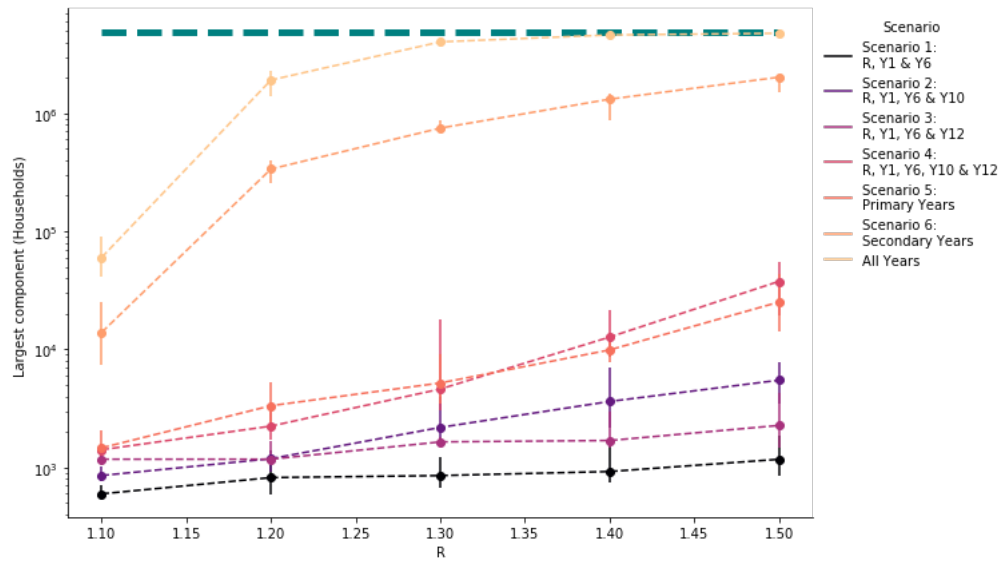


Figure S3 (Sensitivity analysis with $q = 0.08$) The number of households with children attending a school in each largest connected component of the binary transmission networks (estimated potential outbreak cluster size) generated from transmission probability networks for school reopening scenarios.

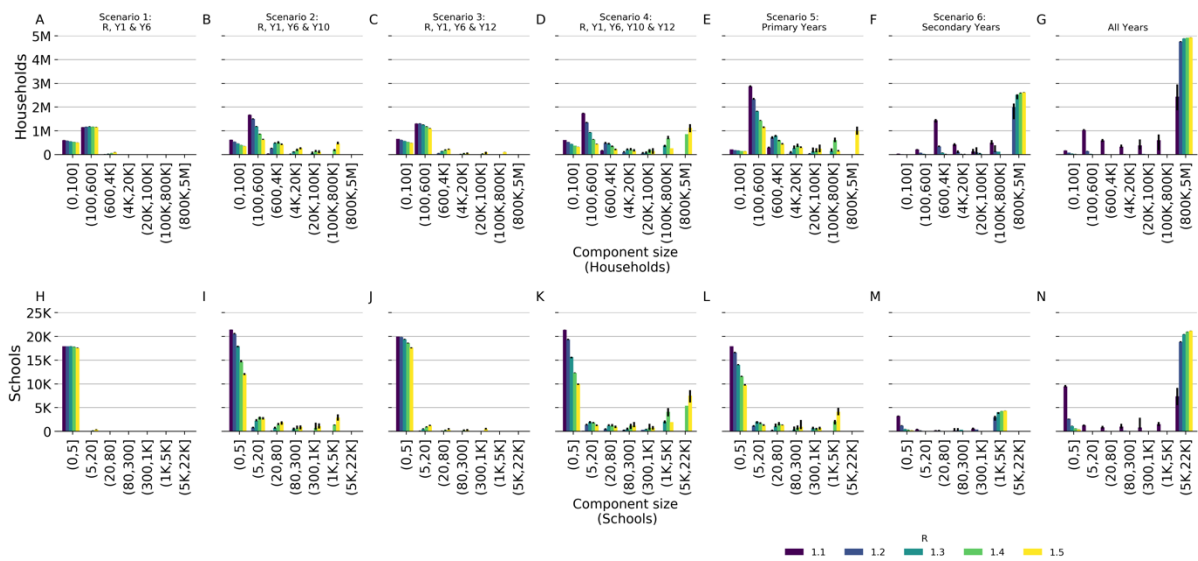


Figure S4 (Sensitivity analysis with $q=0.3$) Connected component distribution of the binary transmission networks generated for school reopening scenarios and R values of 1.1 to 1.5 (indicated by colour). By school (A-G), i.e. the number of schools in a component size in each bin and by household (H-N)

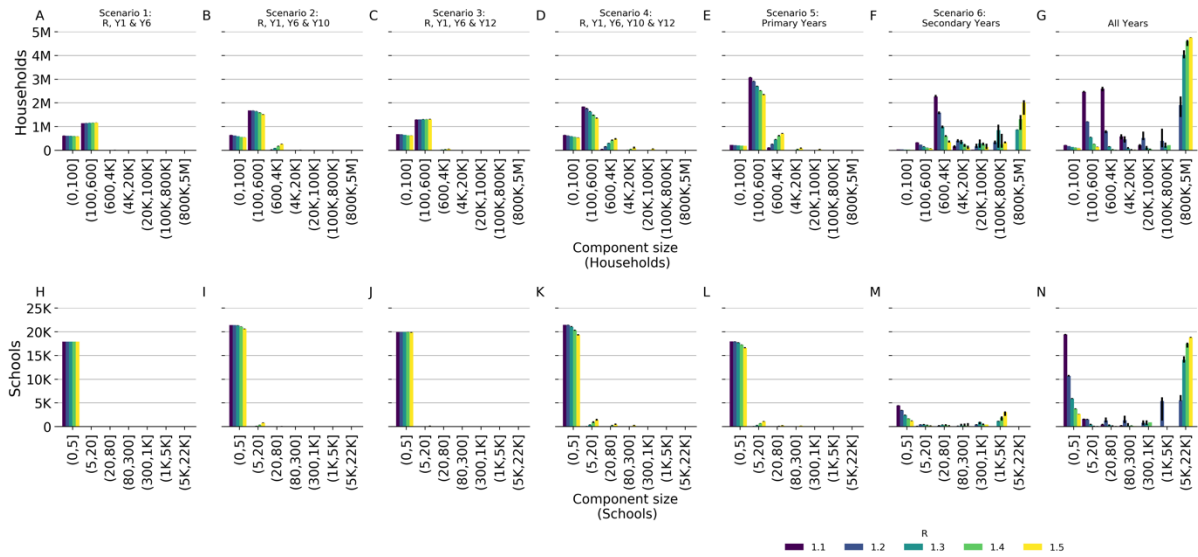


Figure S5 (Sensitivity analysis with $q=0.08$) Connected component distribution of the binary transmission networks generated for school reopening scenarios and R values of 1.1 to 1.5 (indicated by colour). By school (A-G), i.e. the number of schools in a component size in each bin and by household (H-N)