DCC/Northern Alliance of Scotland Report

Stuart McIntyre*, Gennaro Rossi†

July 15, 2021

*Fraser of Allander Institute (FAI), Department of Economics, University of Strathclyde, Glasgow, UK.

1 Introduction

In this report we provide some data visualisation and descriptive statistics documenting educational inequalities across Scottish primary schools. In particular, we will closely follow the approach adopted by Chetty et al. (2020), who examine income segregation across US colleges by calculating income mobility rates for each college. In their work, colleges characterised by high intergenerational mobility where those with a higher share of high-income (top 20% of income distribution) alumni coming from low-income family (bottom 20% of income distribution). However, we will be focusing on "educational mobility", in other words how well students from low socio-economic background perform at school. This is motivated by the vast literature linking poverty or deprivation to low educational attainment (see Robertson and McHardy 2021 for a review) as well as the long-standing involvement of the Scottish Government in relation to the attainment gap.[1] For this reason, we will be using free school meals (henceforth, FSM) registration at the school level and provide a descriptive analysis of the link between socio-economic background and educational attainment in primary schools. A similar approach was followed by Blanden et al. (2007) and Jerrim and Macmillan (2015).

2 Background, Data and Sample Construction

We used data from the Scottish Pupils Census for all primary schools in Scotland between 2015 and 2018. In particular, we only included schools which are observable every year in the above-mentioned interval, and which had no "missing" stages, i.e. an enrollment count of zero for any of the stages from P1 to P7. The census includes all pupils enrolled in Local Authority-funded schools, and for each of them we observe their gender, ethnicity, stage, an identifier for the school they are enrolled in and, most importantly for this analysis, whether or not they are registered for FSM. We then match these data, by virtue of an anonymised candidate number, to *Curriculum for Excellence* teacher-based assessments in literacy and numeracy as well as literacy sub-categories such as reading, writing, listening & talking, for pupils in P1, P4 and P7. Therefore, for each pupil, in each school, we observe the above-mentioned demographic features, alongside whether or not they performed at/above the expected level for relevant primary school stage their in. Ultimately, we obtain the following information pooled across school years 2015/16-2018/19 at the stage- and school-level: *i)* the percentage of pupils registered for FSM; *ii)* the percentage of pupils who perform *at/above level* in both literacy and numeracy.

[1] Although the attainment gap refers to deprivation of the area a pupil lives in.

3 Results

3.1 Mobility Rates by Stage

In Figure 1 we report a set of schools, broken down by stage. On the horizontal axes, we measure the percentage of pupils registered for FSM within a specific school-stage (*Access*), whereas on the vertical axes we report the percentage of pupils on FSM who performed at level (*Success Rate*).

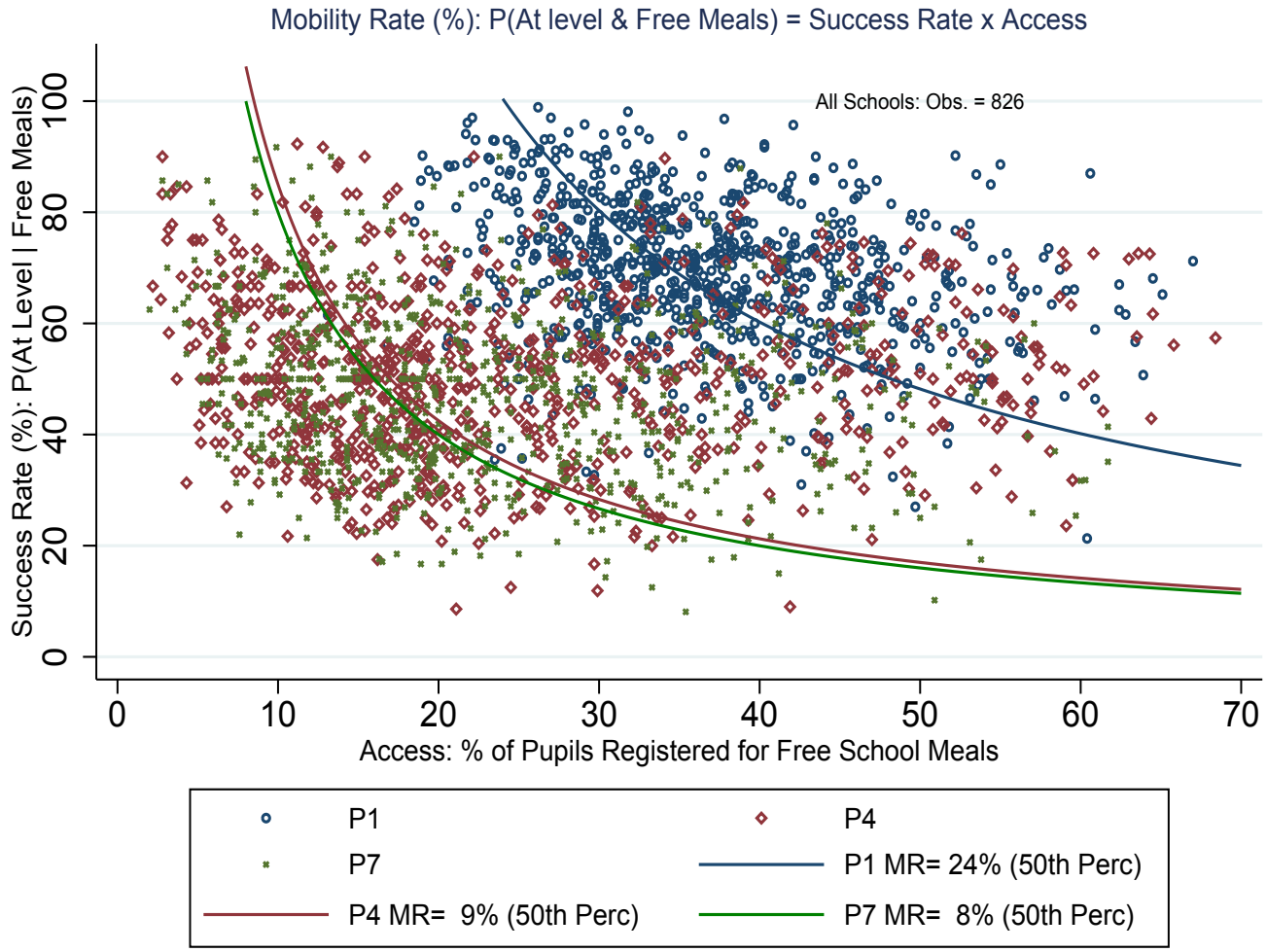

Figure 1: Mobility Rates By Stage - All Schools

By multiplying these two measures, we obtain the *Mobility Rate*, or simply put, the percentage of all pupils in a specific stage-school who are registered for FSM and performed at/above level.[2] If a school-stage has a

[2]The subtle difference between *Mobility Rate* and *Success Rate* is that the previous refers to all the pupils in stage-school, whereas the latter refers only to those in school-stage who are on FSM. See Technical Appendix (section 4) for more details.

high percentage of pupils on FSM (moving to the rightward on the horizontal axis) and a high percentage of pupils on FSM who performed at level (moving upward the vertical axis) this school will have a high *Mobility Rate*. Therefore, as we move from the bottom-left to the top-right of the chart, we go from low-mobility to high-mobility school-stages. The same level of mobility can be achieved with different combinations of *Access* and *Success Rate*. This is the idea underpinning the three downward-slope curves presented in the chart. Let us focus on the navy curve. Along this curve are located all the schools in the sample, whose P1 cohorts recorded a mobility rate of 24%, namely 24% of students are FSM-registered and performed at/above level. This value corresponds to the 50th percentile, or median value, within P1 school-cohorts. What this means is that 50% of P1 school-cohorts have a mobility rate of 24% or more, and 50% of P1 school-cohorts have a mobility rate below 24%. By focusing on the top-end of this curve we can see that there are P1 school-cohorts whose percentage of FSM-registered pupils is just below 30%, and nearly 90% of these performed at/above level. Likewise, P1 school-cohorts whose percentage of FSM-registered pupils is around 50%, and nearly half of which performed at/above level, record the same ($0.5 \times 0.48 = 24\%$) mobility rate. A similar rationale applies to the maroon and green lines, which represents the median mobility rates among P4 and P7 school-cohorts respectively. What emerges from this chart is that P1 school-cohorts seem to be characterised by a mobility rate which is larger (by a factor of 3, approximately) than those of their P4 and P7 counterparts. The reason is twofold: *i)* the share of pupils performing at level is much larger among P1, than it is for P4 and P7 cohorts;[3] *ii)* as a result of the extension of FSM eligibility to all P1-P3 pupils, regardless of their household income, there are many more FSM-registered pupils among P1 than P4 and P7 cohorts.[4] In other words, if both these measures are larger, there will be a higher chance that a randomly picked pupil from a P1 cohort will be FSM-registered and performed at/above compared to one from a P4/P7 cohort. Moreover, another aspect to be noted is the "small" number of schools present in this sample (826 against around 2,000 in total). This is the result of statistical disclosure control measures, i.e. some school-stages had a count of students on FSM (or on FSM and performing at level) below 5, hence this implies a risk of identification of the subjects and the relevant schools have been omitted from the sample.[5]

[3]For more details on Achievement of Curriculum of Excellence, see here.
[4]For more details on free school meals registration and take-up, see here.
[5]See Figure A1, for a version which includes only "large" schools, as in excluding schools from the bottom 25% of enrollment count.

3.2 Mobility Rates by Area

The reduced number of observations due to statistical disclosure control, alongside the "inflated" number of FSM-registration in P1 cohorts implied we decided to pursue a different strategy, namely pooling together P4 and P7 cohorts and omitting P1 ones. Hence, not only are we able to present larger counts – by summing the number of pupils in P4 and P7– and therefore need not omit a relatively large amount of schools, but we are also able to use FSM-registration as a more representative measure of socio-economic status, being FSM-registration directly linked to income in these school stage.[6] An output similar to Figure 1 is presented in Figure 2. In this instance, *Access* measures the percentage of P4 and P7 pupils who are FSM-registered, whereas *Success Rate* is the share of these who also performed at/above level. Therefore, each dot represents an entire school. Instead of distinguishing school-stages, here we separate schools in urban areas from those in rural areas.[7] What can be noticed is that the 197 rural/small town schools present in this sample[8] are characterised by a smaller median level of mobility (5%) than their urban counterparts (7%). This means that among rural and small town schools in this sample, the portion of P4 and P7 pupils who are FSM-registered AND performed at/above level in literacy and numeracy is marginally smaller than in urban schools. This result is not necessary driven by a (yet marginal) discrepancy in attainment between rural and urban schools, and it might well be the result of a smaller number of FSM-registered pupils among rural schools in the first place. Figure A2 reports the same chart, but only for larger schools, i.e. above the 25th percentile.

3.3 Mobility Rates Among Northern Alliance Local Authorities

Figure 3 presents a similar chart to Figure 2, but here we highlight 180 schools within Local Authorities which are part of the Northern Alliance of Scotland. Similarly to Figure 2, we find a slightly different pattern in mobility rates. For schools in Local Authorities which are part of the Northern Alliance half of the schools have a mobility rate above 4%, whereas the other half is below this value. This measure is slightly higher for schools belonging

[6]FSM eligibility was made universal in schools of Glasgow City among P4 students as well, starting from August 2018. However, we do not think this will affect our analysis by a significant extent.

[7]Here we use a 6-categories classification made by the Scottish government based on population size: Large urban areas (settlements with population greater than 125,000); Other urban (settlements with population between 10,000 and 124,999); Accessible small town (settlements with population between 3,000 and 9,999 and within 30 minutes drive of a settlement with a population of 10,000 or more); Remote small town (settlements with population between 3,000 and 9,999 and more than 30 minutes drive from a settlement with a population of 10,000 or more); Accessible rural (areas with a population of less than 3,000 and within 30 minutes drive of a settlement with a population of 10,00 or more); Remote rural (areas with a population of less than 3,000 and more than 30 minutes drive from a settlement with a population of 10,00 or more). We regroup in the following way: 1) Urban = Large and Other Urban; Rural = Accessible and Remote Rural plus Small Towns, whether accessible or remote.

[8]Smaller sample size with respect to what originally can be expected is due to suppression of observations due to statistical disclosure control.

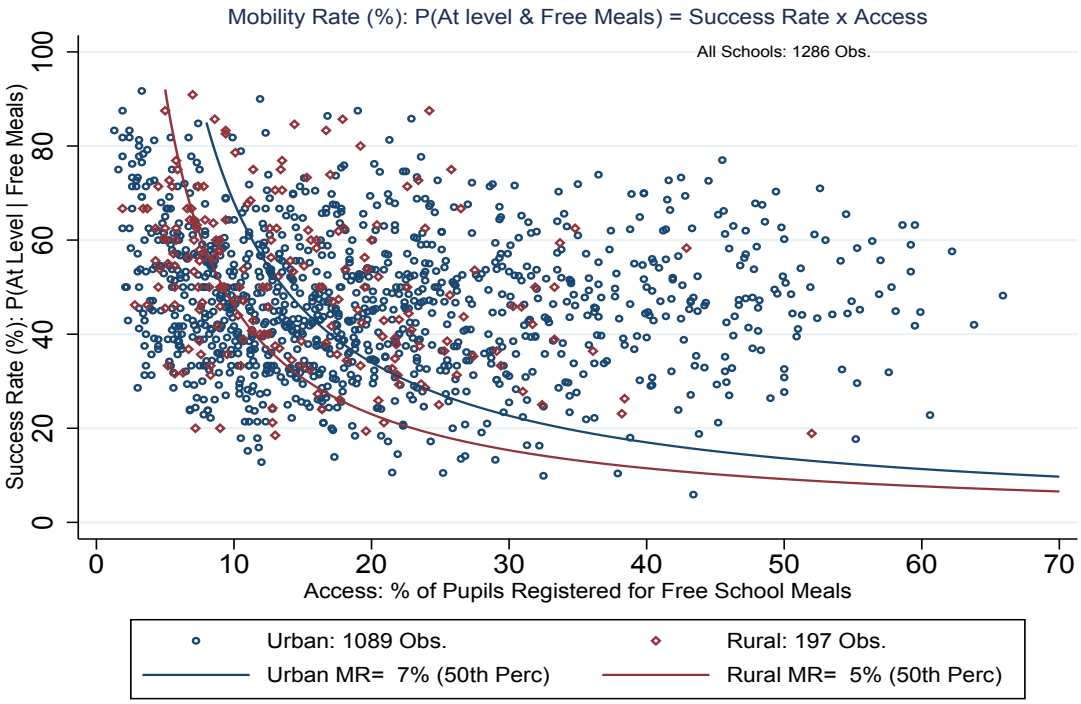

Figure 2: Mobility Rates By Area - All Schools

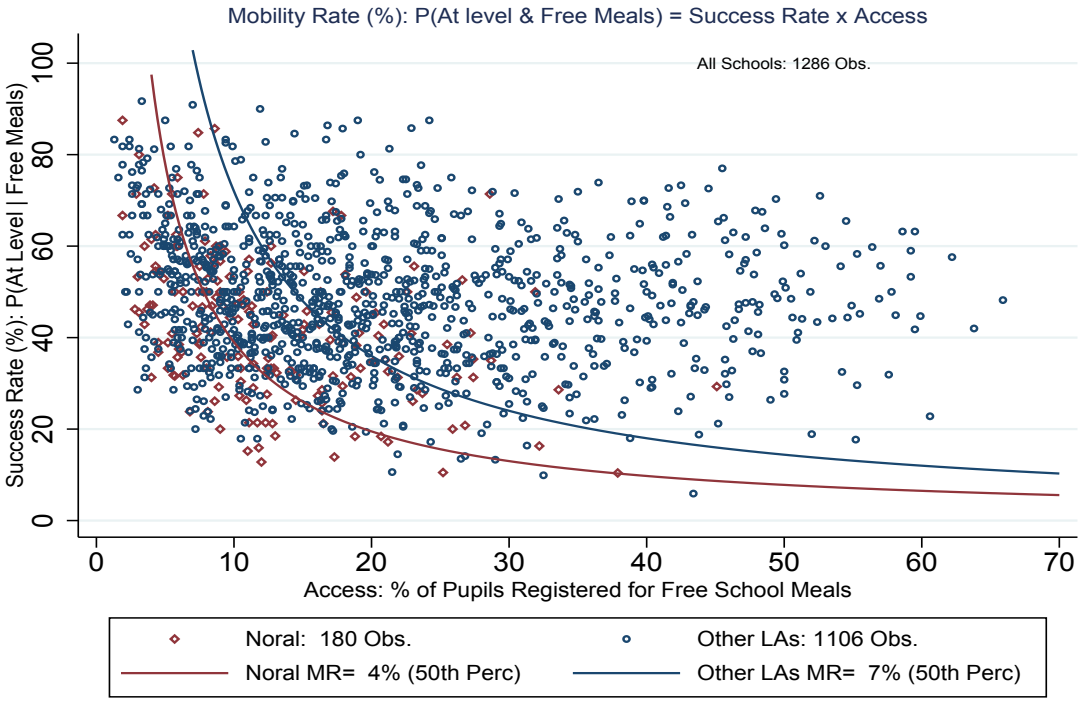

Figure 3: Mobility Rates By Area (Northern Alliance of Scotland) - All Schools

to other local authorities.

4 School Communities

In this section we provide some insights on how heterogeneous are schools in terms of their intakes. In the Scottish primary school system, school populations are largely contingent on catchment areas, with relatively few pupils attending a different school from the designated one, on the basis of a "placing request". Scotland is split in 6,976 data zone,[9] i.e. statistical units which include between 500 and 1,000 inhabitants. Figure 4 shows that nearly 30% of Scottish primary schools catchment areas stretch across seven or more data zones. This might suggest a fairly high degree of heterogeneity in school intakes. However, we need to look closer at how heterogeneous are these data zones in terms of Scottish Index of Multiple Deprivation. Each data zone is

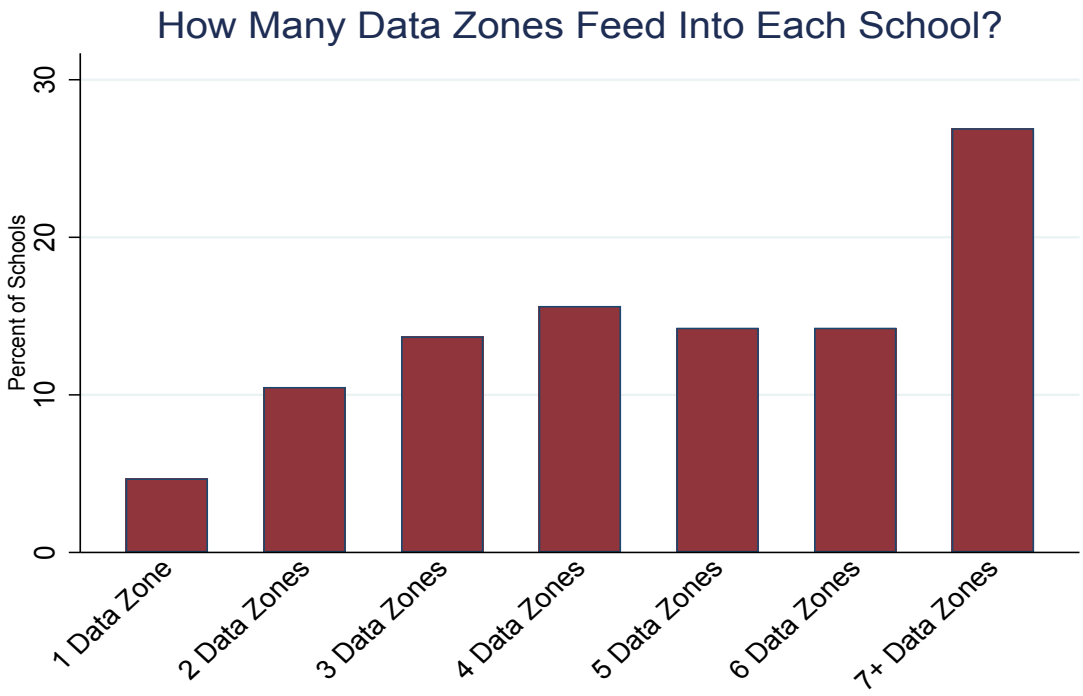

Figure 4: School Composition - Data Zones

ranked from 1 to 6,976, with 1 being the least deprived and 6,976 the most deprived. Figure 5 provides a similar picture to Figure 4, but instead looks precisely at how many deciles are in each school. It can be seen that a relative majority of schools (25%) "contains" five SIMD deciles.

[9]Following the re-definition in 2011.

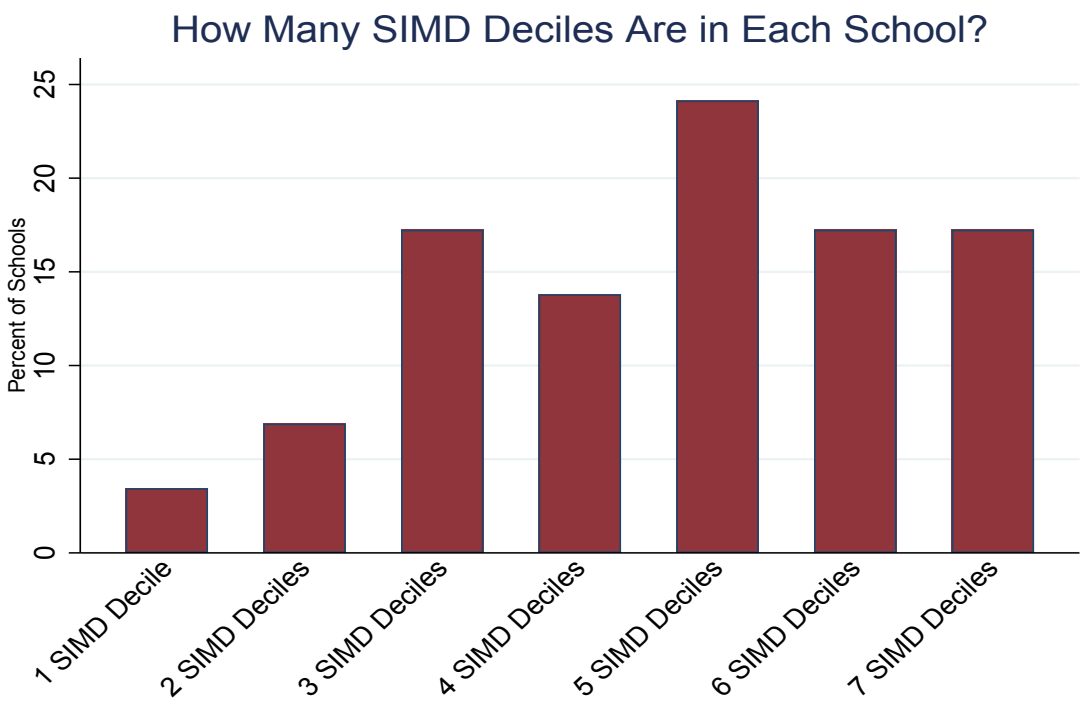

Figure 5: School Composition - SIMD Deciles

Another question that arises is, how well does the physical location of a school predicts its composition. Figure 6 classifies groups of school based on the deprivation decile of the data zone the school is actually in (horizontal axes), with respect to the (weighted) average of its composition (vertical axes). The size of each circle/diamond represents the number of schools in that specific point (at least 20). For example, schools whose location is in the first decile of deprivation, gather pupils coming from the first and second deciles. One thing that can be noticed is that schools located in "central" deciles, namely around a median level of deprivation, seem to be more heterogeneous than schools which instead are located at the two ends of the deprivation distribution. This pattern seem to be pretty consistent both for Northern Alliance's as well as for the other Local Authorities.

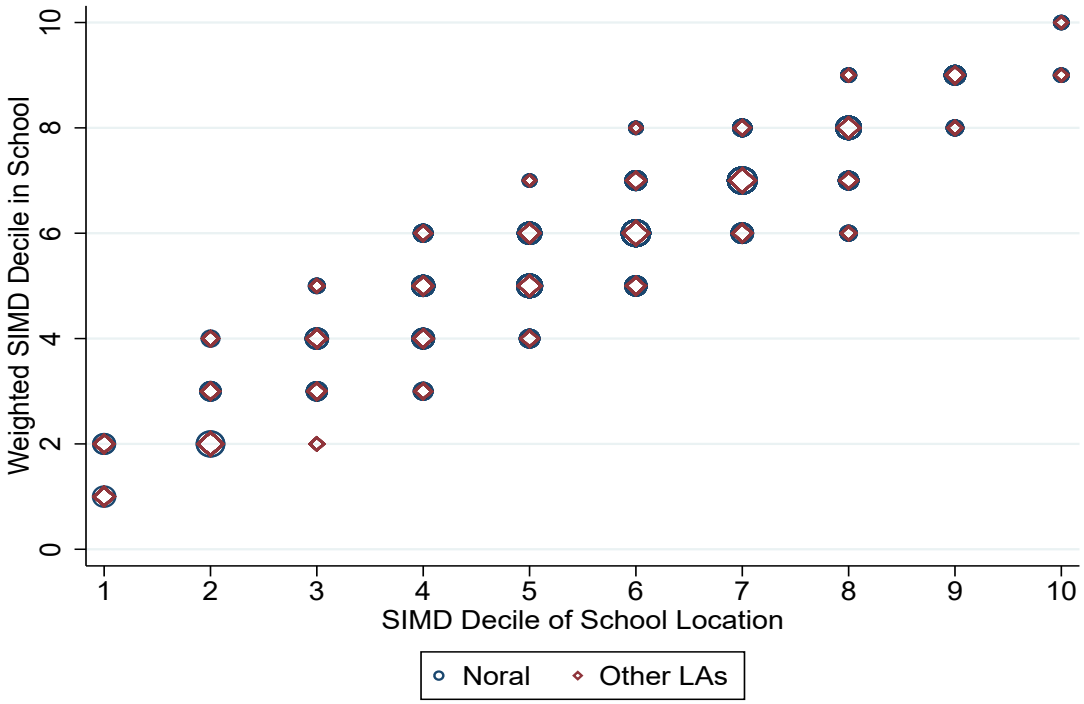

Figure 6: School Composition - SIMD Deciles

References

Blanden, Jo, Paul Gregg, and Lindsey Macmillan, "Accounting for intergenerational income persistence: noncognitive skills, ability and education," *The Economic Journal*, 2007, *117* (519), C43–C60.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan, "Income segregation and intergenerational mobility across colleges in the United States," *The Quarterly Journal of Economics*, 2020, *135* (3), 1567–1633.

Jerrim, John and Lindsey Macmillan, "Income inequality, intergenerational mobility, and the Great Gatsby Curve: Is education the key?," *Social Forces*, 2015, *94* (2), 505–533.

Robertson, Laura and Fiona McHardy, "The Poverty-related Attainment Gap: A Review of the Evidence," *The Poverty Alliance*, 2021.

Appendix

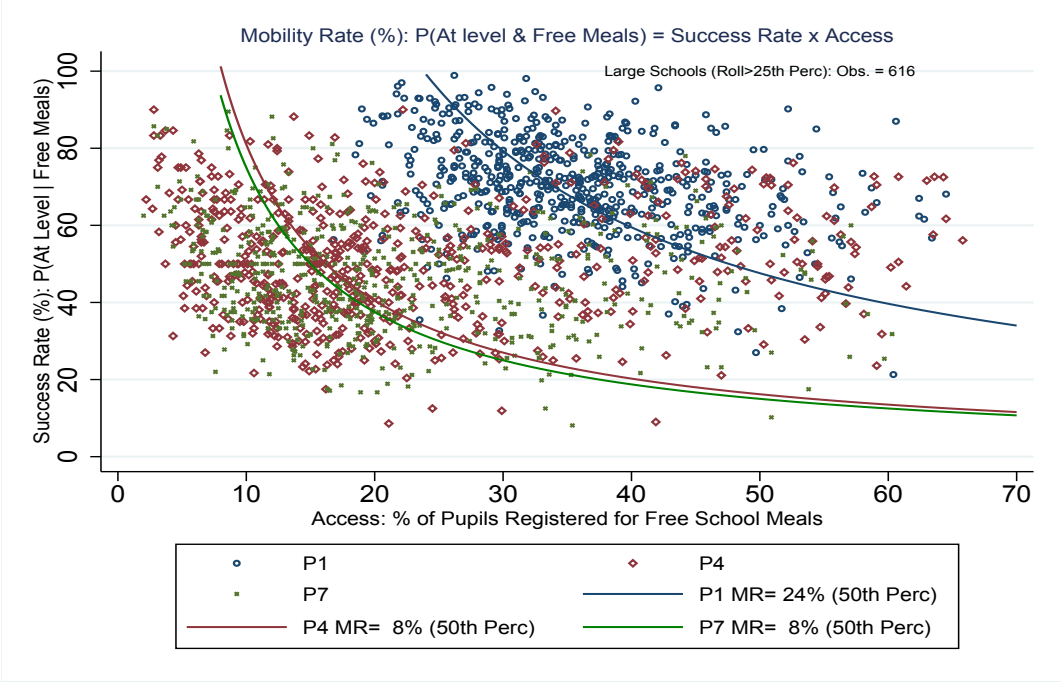

Figure A1: Mobility Rates By Stage - Large Schools

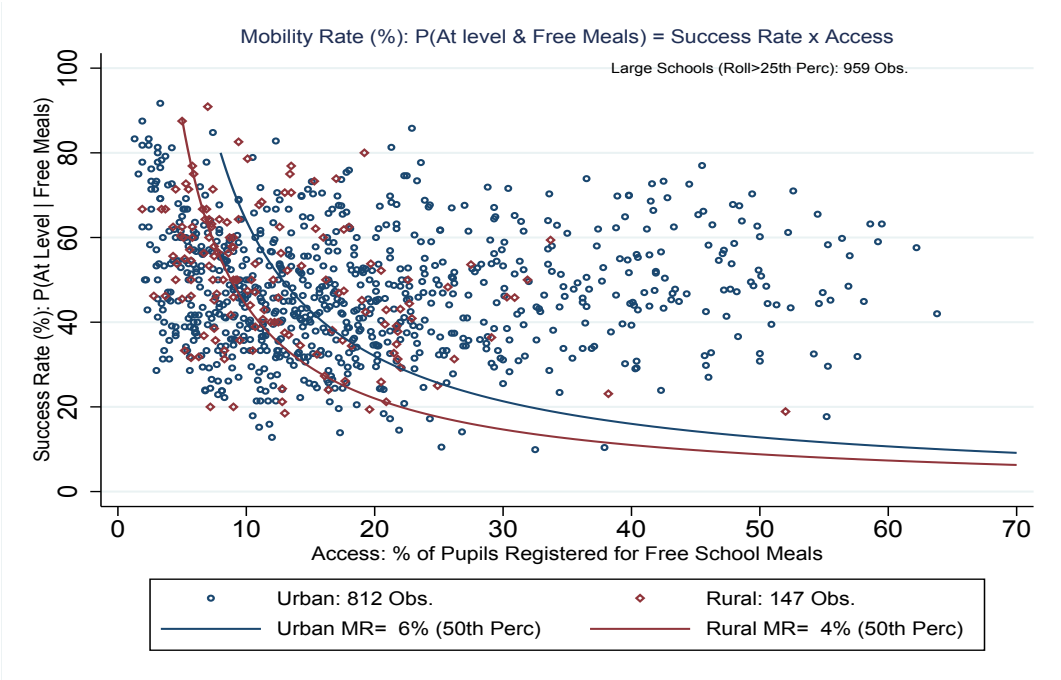

Figure A2: Mobility Rates By Area - Large Schools

Technical Appendix

The measure of mobility which we have employed in this report can be summarised by a standard formula for conditional probability, as illustrated below:

$$P(At\ level \mid Free\ Meals) = \frac{P(At\ level\ \&\ Free\ Meals)}{P(Free\ Meals)} \tag{1}$$

This formula represents the probability that a pupil randomly drawn from a certain population, whether a stage/school/Local Authority, performed at/above level in literacy and numeracy, given that this same pupils is FSM-registered (we also referred to this as *Success Rate*). In other words, if our reference population is school A in Local Authority X, the above formula helps me answering the following question: by picking a student at random within this school/LA, knowing that she is FSM-registered, what is the probability that she has also performed at/above level in that specific year? Given that FSM registration and whether or not the student performs at level are not two mutually exclusive events, the above-mentioned conditional probability can be illustrated as the ratio between

$$P(At\ level\ \&\ Free\ Meals) = \frac{No.\ of\ Pupils\ who\ Performed\ at\ Level\ \&\ on\ Free\ Meals}{Total\ No.\ of\ Pupils\ in\ School}$$

Namely the share of pupils in school A who are FSM-registered pupils AND performing at/above level, also referred to as *Mobility Rate* and

$$P(Free\ Meals) = \frac{No.\ of\ Pupils\ on\ Free\ Meals}{Total\ No.\ of\ Pupils\ in\ School}$$

Which is simply the share of pupils in school who are FSM-registered (*Access*). It is easy to work out from that the ratio between these two elements is

$$P(At\ level \mid Free\ Meals) = \frac{No.\ of\ Pupils\ who\ Performed\ at\ Level\ \&\ on\ Free\ Meals}{No.\ of\ Pupils\ on\ Free\ Meals}$$

namely the share of FSM-registered pupils in school A, who have also performed at/above level. By re-arranging Equation 1, it is easy to see how the *Mobility Rate* is just the product of *Access* and *Success Rate*

$$\underbrace{P(At\ level\ \&\ Free\ Meals)}_{\text{Mobility Rate}} = \underbrace{P(Free\ Meals)}_{\text{Access}} \times \underbrace{P(At\ level \mid Free\ Meals)}_{\text{Success Rate}}$$

Therefore, the same level of mobility (in our example above, the median level) can be achieved with different combinations of access and success rate. For example, a school whose 60% of pupils are FSM-registered, and 40% of these performed at level (.6×.4=.24, hence 24%) will be just as "mobile" as a school in which 30% of its 80% of FSM-registered pupils will have also passed the level (.8×.3=.24, hence 24%).