Supplementary analysis of non-response and attrition

A simple way of looking at attrition by disability status is to calculate how the share of those who are disabled, according to the definition used in our paper, changes across the waves of the survey. If attrition were linked to disability status we would expect to see declining (unweighted) proportions of disabled young people in the sample over time. Table A1 shows that there is some differential attrition but this is not substantial in terms of the composition of the sample that is disabled across the waves (Column 1: Unweighted %). There is a slightly larger degree of differential attrition between waves 4 and 7 for disabled young people, but this may well be related to their patterns of school attendance after age 16 and therefore the extent to which they can be so easily tracked. Note that while Table A1 examines the distribution across the whole sample for each wave, our analysis at Waves 4 and 5 is on conditional samples. We discuss this issue further below.

Table A1: Shares of Wave1-7 samples made up of disabled young people, weighted and unweighted

Wave	Unweighted %, full sample	Unweighted %: those supplying information on KS2	Weighted %	Outcome measured at this point
1	13.3 (N=15,660)	12.9 (N=14,466)	15.3	N/A
2	12.8 (N=13,128)	12.6 (N=12,232)	15.3	University expectations
3	12.5 (N=12, 144)	12.3 (N=11,471)	15.3	Level 2 qualifications
4	12.4 (N=11,238)	12.1 (N=10,652)	15.2	Continued in full-time upper secondary education (on conditional sample)
7	11.3 (N=8,115)	11.3 (N=7,740)	14.8	Attend university (on conditional sample)

Note: Disability is measured using analytic measure drawing on wave 1 variables of school-limiting long-standing illness and current special educational needs (excluding gifted and talented and English as a second language).

Administrative data is linked to the survey responses, including Key Stage 2 (KS2) test scores from age 11. KS2 is available for the vast majority of young people in the survey. Column 2 of the table therefore provides the share of disabled young people at each wave, retaining only those for whom information on KS2 was supplied. We can see that those who provided KS2 information were slightly selected on disability status (i.e. a slightly smaller proportion of disabled people provided this information) but differential attrition across the waves did not substantially affect the composition of the sample.

The final column of Table A1 illustrates the sample composition, i.e. the proportions of disabled, once we adjust for complex survey design and weighting. Disability is under-represented in the overall sample but adjustments allow the proportions of disabled to remain constant throughout, apart from a very slight diminution in W7. The aim of weights is to account for differential non-response and they appear to be effective for disability. It is to be noted that in our paper, weights are applied in all analyses.

It is also possible to model response at later waves with disability as an explanatory variable. The degree of differential attrition among disabled young people illustrated in Table A1 may be driven by other characteristics that make young people less likely to respond or be traced and which are correlated with disability status. Since disabled young people are likely to have other individual and family characteristics that are independently associated with attrition, it is relevant to control for these measures in such models. Reflecting our rather parsimonious modelling strategy, we include only a simple set of covariates. More complex response models implemented by the data collection team, and using slightly different definitions of disability, failed to identify a significant effect of disability on response (Department for Education 2011).

Since retention in the sample at wave t+1 was conditional on participation at wave t, we model response at consecutive wages conditional on response at the previous wave. From our response models, we can identify the predicted percentages of those who continue to the next sweep, conditioning on selected relevant characteristics collected at wave 1, namely parental education, ethnic group, and child sex: see Table A2, column 1. The simple response models indicate that disability is significantly associated with a greater propensity not to respond at most sweeps, though the differences are mostly small.

In addition, we estimated another response model where we also condition on KS2 attainment: Table A2, column 3. Note that we lack KS2 information for around 7 percent of the initial sample and for around 10 per cent of disabled children. This means that controlling for KS2 provides us with a slightly selected KS2 sample, as we discussed in relation to Table A1. However, this model does suggest that non-response on disability is driven by KS2 score, that is, by differences in educational attainment, since disabled children are no less likely than their non-disabled counterparts to respond in this model, and, indeed are slightly more likely to respond at wave 7.

As there is some variation in disability by ethnicity, we also show estimated response probabilities for the different ethnic groups in our sample in Table A2. These show that the differences in response for some ethnic groups are much greater than for disability. Our final sample is therefore somewhat selected in terms of ethnic group. Moreover, controlling for KS2 score does not reduce the differences by ethnic group as much as it does for disability, suggesting that ethnic group differentials in response are not driven primarily by differences in educational attainment.

Table A2: Estimates from logistic regression of predicted response at waves 2, 3, 4 and 7 for disability and ethnic group

		Predicted response of those responding at prior wave		
		All, excl KS2 in model	Had made relevant prior transition, excl KS2	All, incl KS2 score in model
W2		All resp at W1	N/A	All resp at W1 & w. KS2 score
	Not disabled	85		85
	Disabled	82*		86
	Ethnic group			
	White	86		87
	Mixed	81*		83*
	Indian	85		86
	Pakistani	82*		84*
	Bangladeshi	84		85*
	Black Caribbean	77*		79*
	Black African	73*		76*
	Other	77*		81*
N		15,404		14,244
W3		All resp at W2	N/A	All resp at W2 & w. KS2 score
	Not disabled	93		93
	Disabled	91*		93
	Ethnic group			
	White	93		94
	Mixed	91*		92
	Indian	93		93
	Pakistani	93		94
	Bangladeshi	95*		95
	Black Caribbean	86*		87*
	Black African	84*		86*
	Other	91		93
N		13,036		12,247
		All resp at W3	W3 resps who attained Level 2	All resp at W3 & w. KS2 score
W4	Not disabled	93	94	92
	Disabled	91*	95	93
	Ethnic group			
	White	93	95	93
	Mixed	91	93	91
	Indian	95	95	94
	Pakistani	92	92*	93
	Bangladeshi	91*	94	92
	Black Caribbean	85*	89*	87*
	Black African	85*	90|*	87*
	Other	91	91*	91
N		12,067	7,318	11,399
W7		All resp at W4	W4 resps, stayed on in ed. & attained Level 2	All resp at W4 & w. KS2 score
	Not disabled	73	82	72

Disabled	68*	83	75*
Ethnic group			
White	71	82	71
Mixed	66*	77	67*
Indian	83*	87*	83*
Pakistani	76*	82	78*
Bangladeshi	77*	85	79*
Black Caribbean	65*	82	68
Black African	63*	74*	70
Other	74*	81	76
N	11,171	5,621	10,590

Note: All predictions adjust for sex, parental qualifications, ethnic group and disability.
*indicates a statistically significant difference at p<0.05.

It is worth noting that in our analysis of outcomes at waves 4 and 7, we are selecting intentionally on performance at Level 2 qualifications and continuing in full-time academic upper secondary education with Level 2 qualifications, respectively. We are thus estimating disability differentials on a constructed sample, who might be expected to have higher response. Column 2 of Table A2 therefore shows the response rates at Waves 4 and Waves 7 for those who are eligible to be included in our selected sample at those sweeps. We see that there are no differences in response by disability at Wave 4 for this group; and at Wave 7 disabled young people are somewhat more likely to respond. The differences in response by ethnic group are also somewhat attenuated for these samples.

This analysis shows that there is some loss of representativeness in our sample as disability is associated with slightly higher rates of attrition. However, there is no evidence that disability is a major driver of attrition. Despite the loss of cases and the move towards a more educationally selected sample, the composition remains largely consistent.

Reference

Department for Education [DfE]. 2011 *LSYPE User Guide to the Datasets: Wave 1 to Wave 7*, London: DfE.