Panel attrition in the General Population Sample and the Immigrant and Ethnic Minority Boost of *Understanding Society*

Pablo Cabrera-Álvarez[#], Nicole James and Peter Lynn

Institute for Social and Economic Research, University of Essex

Corresponding Author.

Panel attrition in the General Population Sample and the Immigrant and Ethnic Minority Boost of *Understanding Society*

Pablo Cabrera-Álvarez, Nicole James and Peter Lynn

Institute for Social and Economic Research, University of Essex

Non-Technical Summary

In longitudinal studies where panel members are asked to participate regularly, some can drop out of the panel while still eligible to be interviewed. Attrition poses a threat to data quality in longitudinal studies. First, the decrease in sample size might make the analysis of some sample subgroups difficult. Second, the propensity to drop from the study varies across sample subgroups, and the differential attrition might result in biased survey estimates.

The analysis reported in this paper aims to describe the evolution of panel attrition in two samples of *Understanding Society*, the General Population Sample (GPS), a sample representative of the Great Britain population, and the Immigrant and Ethnic Minority Boost (IEMB), which covers people from certain ethnic minorities and immigrants living in Great Britain households. This report investigates the evolution of panel attrition by focusing on the initial sample of respondents and exploring the drop in response rates from wave to wave. The change in wave response rates over time is analysed across sample subgroups based on the baseline characteristics of the panel members. Furthermore, we investigate the effectiveness of survey weights to tackle the impact of attrition in the GPS sample. The analysis of the GPS covers waves 1 to 11, while the IEMB covers waves 6 to 11.

The results show that the GPS lost 60.1% of the initial wave respondents after eleven waves. This drop in participation is particularly noticeable among younger people, panel members with an ethnic minority background, on lower income or with no qualifications. However, the survey weights were able to mitigate the impact of attrition. Regarding the IEMB, it lost 67.3% of the initial wave respondents in five waves, with Black people and those of Bangladeshi or Pakistani origin more likely to drop out.

Panel attrition in the General Population Sample and the Immigrant and Ethnic Minority Boost of Understanding Society

Pablo Cabrera-Álvarez, Nicole James and Peter Lynn

Institute for Social and Economic Research, University of Essex

Abstract: This working paper reports an analysis of panel attrition in the General Population Sample (GPS) and the Immigrant and Ethnic Minority Boost (IEMB) of *Understanding Society*. The analysis involves the comparison of the wave response rates for the initial samples of respondents overall and across a set of groups formed by baseline characteristics. In addition, we assess the effectiveness of the survey weights to tackle attrition. The analysis of the GPS covers waves 1 to 11, while the IEMB covers waves 6 to 11.

Corresponding Author: Pablo Cabrera-Álvarez, Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, United Kingdom, pcabre@essex.ac.uk.

Executive Summary

1. The General Population Sample (GPS) of *Understanding Society* has lost 60.1% of the first wave respondents between waves 2 and 11. This erosion is more pronounced than the one observed in the British Household Panel Survey (BHPS), which after 13 waves, retained 69.9% of the starting sample of respondents.

2. Some subgroups of the GPS have been more likely to abandon the panel, negatively impacting the sample balance. Based on their initial wave characteristics, the groups more likely to attrit have been youngsters, ethnic minorities, participants with poor health, those living in London, Scotland or Wales, on lower incomes, full-time students or unemployed, singles, participants with no qualifications, renting their houses and lone parents. In the period between waves 10 and 11, which covers the beginning of the covid-19 crisis, the level of erosion was higher among older respondents.

3. The survey weights effectively tackle panel attrition for the set of variables included in this analysis. The survey weights can restore the sample profile mitigating the impact of attrition.

4. The Immigrant and Ethnic Minority Boost (IEMB) sample has lost 67.3% of the initial wave respondents in five waves, from wave 6 to 11. Almost half of the initial sample of respondents (45.4%) abandoned the panel between waves 6 and 8. Afterwards, the attrition rate decelerated to 15.2% between waves 8 and 10.

5. In the IEMB, youngsters were more likely to abandon the sample, although the participants aged 60 and over had the highest attrition rate between waves 10 and 11 when the covid-19 pandemic was onset. In terms of ethnicity, Blacks, Bangladeshis and Pakistanis were more likely to drop over time, as well as those living in London. Then, panel members on lower incomes, unemployed, full-time students, without academic qualifications or renting their accommodation were more likely to attrit. The differences were less remarkable for the subgroups formed by the variables general health status and household type.

Introduction

The main objective of this paper is to assess the impact of panel attrition on the sample profile of the General Population Sample (GPS) and the Immigrant and Ethnic Minority Boost (IEMB) of *Understanding Society*. This analysis addresses the evolution of attrition over time, focusing on the variation between the data collected before and during the covid-19 pandemic. In addition, this paper evaluates the effectiveness of survey weights to mitigate the effect of panel attrition on the sample profile of the GPS.

This report continues a series of publications about the representativity of the sample of *Understanding Society*. In 2018, Lynn and Borkowska evaluated the representativity of the British Household Panel Study (BHPS) and the GPS of *Understanding Society*. They compared the initial samples to the population figures from the census and investigated the impact of panel attrition over time using the respondents to the initial wave as the reference point. The analysis presented in this paper relies on upon and extends that work on the sample representativeness of *Understanding Society*, adding a specific analysis of the IEMB. These analyses use a set of eleven variables collected for the respondents at the initial wave of the study.

This analysis covers the GPS (2009), the main sample of *Understanding Society*, representative of the Great Britain population. The impact of panel attrition on this sample was evaluated by Lynn and Borkowska (2018), covering the period between waves 2 and 7. This analysis extends that work by including waves 10, whose data collection occurred almost entirely before the beginning of the COVID-19 pandemic in the United Kingdom in March 2020, and 11, which took place between January 2019 and June 2021. Furthermore, we present a set of measures, R-indicators and coefficients of variation (CV), to help assess the ability of survey weights to correct the deviations generated by the erosion of the panel over time.

This report also includes an analysis of the attrition in the IEMB sample. This sample was first interviewed in wave 6 (2014-16) and is a probability sample of households where at least one person had been born outside of the UK or considered themselves or parents or grandparents to be of one of the main ethnic minority groups in the UK (i.e., Indian, Pakistani, Bangladeshi, Black Caribbean or Black African) (Lynn et al., 2018). This is the first analysis of the impact of panel attrition in this sample covering waves 7 to 11.

Methods

This paper comprises two analyses: 1) an attrition analysis of the GPS and the IEMB sample; 2) an analysis of the effectiveness of survey weights in reducing the effects of attrition in the GPS.

The attrition analysis compares the response rates for different sample subgroups over time, covering waves 2 to 11 in the GPS and 7 to 11 in the IEMB. Then, for the GPS, we compare the evolution of the R-indicators and coefficients of variation (CV), two measures that assess the level of representativeness of the sample with respect to a set of auxiliary variables. Comparing the R-indicators and CVs for the unweighted and weighted GPS allows us to evaluate the ability of the survey weights to mitigate the impact of panel attrition on the sample profile.

Attrition analysis

This analysis compares the response rates for different sample subgroups over time. In this section, we explain the difficulties we found and the methodological decisions we made in order to compute the rates. These decisions cover four aspects: 1) the reference sample, 2) the definition of survey respondents, 3) the assumptions about the eligibility of the sample members to compute the response rates, and 4) the interpretation of the response rates.

The attrition analysis requires a reference sample to observe the effect of dropouts over time. In this analysis, the reference point is formed by the respondents to the initial wave, which was different for the GPS and IEMB. The GPS was recruited at wave 1 of *Understanding Society* (2009-11), while the initial wave for the IEMB sample corresponds to wave 6 of *Understanding Society* (2014-16). The definition of respondents, the numerator in calculating the response rates, covers the panel members completing the adult questionnaire and the proxy respondents, for whom another household member responded to some questions on their behalf.

The denominator of the response rate is the number of initial wave respondents eligible for an adult interview. The eligibility situation may change over time: participants who have moved out of the country, been institutionalised, or died before the next wave of data collection are ineligible and, therefore, should be excluded from the response rates calculations. In some instances, the fieldwork force learns about the participants' circumstances that lead to a change in their eligibility status. However, some participants become unavailable from one wave to another, and the fieldwork force never learns whether this is a case of genuine nonresponse or

they became ineligible (i.e., moved out of the UK, got institutionalised or died). Ignoring these changes in the eligibility of the sample would lead to underestimating response rates. This underestimation is likely to increase over waves as more participants become ineligible, especially for some subgroups such as elderlies. In this analysis, apart from excluding from the calculations the ineligible cases identified during the data collection, we have implemented a mortality adjustment that removes part of the underestimation due to undetected deceased participants. These adjustments are different for the GPS and IEMB analyses:

1. For the GPS, for waves 2 to 8, the adjustment takes the form of a propensity coefficient representing the probability of being deceased and not observed at each wave. The calculation of this propensity relies on a survival model that uses data from the official mortality registers, census, and the data collected during the fieldwork (Kamisnka, 2021). This adjustment is unavailable for waves 9 onwards at the moment. For the attrition analysis of waves 10 and 11, we used the wave 8 propensity adjustment and removed the cases identified to have died in this period from the mortality registers and during the data collection.

2. The propensity adjustment was not available for the IEMB sample, given the lack of official mortality data by ethnic background. Instead, to control part of the unobserved mortality, we removed from the calculations the participants known to have died between waves 8 and 11 identified from the mortality registers.

Finally, we should mention that response rates are sample-based estimates and therefore are subject to sampling error. Therefore, minor differences between the rates should not lead us to conclude that these rates are different in the population.

Evaluation of the survey weights in tackling panel attrition

The second objective of the analysis is to evaluate the performance of survey weights to adjust the sample profile in the GPS. R-indicators and coefficients of variation (CV) were used to show the ability of the weights to correct the effect of panel attrition (see Schouten et al., 2016).

R-indicators measure the representativeness of a sample with respect to a set of auxiliary variables known for respondents and nonrespondents. This indicator is based on the variability of the response propensities, which are unknown and need to be estimated using a propensity model fitted with the auxiliary information. Auxiliary variables play a vital role in the

calculation and interpretation of the R-indicators, and the representativeness diagnostics based on them are limited to the auxiliary variables included in the model. In this analysis, the variables used to compute the response propensities that lead to the R-indicator are gender, age, ethnic background, Government Office Region, personal income, general health, employment status, marital status, education, household type, and household tenure.

R-indicators can take values between 0 and 1, where values close to 0 indicate that response propensities vary across the groups defined by the auxiliary variables, meaning that some groups are overrepresented while others are underrepresented in the sample of respondents. Therefore, values close to 0 indicate that the sample has representativity issues with regard to the auxiliary variables included in the analysis. In contrast, values close to 1 indicate a slight variation of the response propensities, meaning that the responses are representative in terms of the auxiliary variables.

The coefficient of variation is a measure related to the R-indicator representing the estimated maximal absolute standardised bias assuming that nonresponse correlates maximally with the auxiliary variables. This measure shows a worst-case scenario in terms of bias. When the sample is balanced in terms of the auxiliary variables used in the propensity model, the CV would be close to zero, indicating an absence of potential bias.

Unconditional partial CVs were derived to evaluate the contribution of single auxiliary variables to the lack of representativeness of the responses. Note that these measures are not controlled for the effect of the rest of the auxiliary variables. For instance, the unconditional partial indicator of the auxiliary variable marital status is correlated with age since older people tend to be widowed or younger adults are more likely to be single. Thus, when interpreting the CV, we are looking at the contribution of marital status without controlling for the rest of the auxiliary variables. In the analysis, we compare the weighted and the unweighted versions of these representativeness indicators to evaluate the ability of survey weights to mitigate the effects of panel attrition.

Note that the R-indicators and CVs are sample-based estimates and therefore are subject to error. This implies that minor differences could be due to the random variability introduced at the sampling stage.

Results

In this section, we present the results of the attrition analysis of the GPS and the IEMB sample. The table with the results for the overall samples and the crosstabulations by gender, age and ethnic background are presented in this section. The tables for the rest of the variables can be found in the annexes A (GPS) and B (IEMB).

GPS

Eleven waves after the recruitment of the GPS, 39.9% of those responding to an adult interview in wave 1 participated (Table 1). The attrition analysis of the BHPS might serve as a reference point to evaluate the magnitude of the panel erosion in the GPS. At wave 13 of the BHPS, 69.9% of the sample of wave 1 respondents still participated, 30 p.p. higher than the GPS. In line with Lynn and Borkowska's (2018) conclusions, the GPS erodes at a higher rate compared to the BHPS. Most dropouts occurred in the first four waves (35.9%); afterwards, the attrition rate significantly decreased. Between waves 10 and 11, the response rate decreased by 2.8 p.p., from 42.7% to 39.9%.

The response rates were similar among men and women over time. Regarding participants' age (in 2009), the attrition rates were significantly higher among the younger sample members during the first four waves. However, after that, the oldest group of the sample registered a higher attrition rate compared to the rest of the age groups. For example, while the average reduction in response rates between waves 7 and 10 was 8.3 p.p. for the overall sample, this difference was 14.2 p.p. for participants aged 70 and older. Something similar happened between waves 10 and 11, where the decrease in the response rate of the 70 and older group was 5.4 p.p. compared to the average of 2.8 p.p. However, this increase in the attrition rate could be overestimated due to undetected mortality in the panel, which predominantly affects the oldest group of participants. The non-white participants, especially Blacks and Bangladeshis, were more likely to abandon the study than British whites between waves 2 and 4. These differences among groups softened over time and, between waves 10 and 11, all groups presented attrition rates relatively close to the average of 2.8 p.p.

Table 1. GPS Attrition: Sex, Age and Ethnic Background

		Wave 1 (2009-11)	Wave 4 (2012-14)	Wave 7 (2015-17)	Wave 10 (2018-20)	Wave 11 (2019-21)
Full sample		43,673	64.1	51.0	42.7	39.9
Sex						
	Male	19,773	63.5	50.5	41.5	38.8
	Female	23,900	64.6	51.5	43.6	40.7
Age in 2009	**16-19**	2,700	51.2	35.2	26.3	24.0
	20-29	6,389	51.7	38.7	31.4	28.7
	30-39	7,408	64.1	49.7	40.9	38.4
	40-49	8,266	65.9	53.9	46.1	43.8
	50-59	6,891	70.6	59.2	52.5	50.0
	60-69	6,287	71.3	59.9	52.9	49.8
	70+	5,732	65.9	50.4	36.2	30.8
Ethnic group	**White**	39,761	65.4	52.4	44.2	41.3
	Black	961	47.5	34.3	23.2	20.0
	Indian	885	53.7	40.0	30.0	28.1
	Pakistani	551	53.9	40.7	32.6	30.8
	Bangladeshi	197	42.8	35.8	26.6	27.9
	Other Asian	510	49.7	38.8	27.3	24.3
	Mixed	456	57.1	41.4	33.2	30.6
	Other	304	46.8	31.7	21.5	21.5

The response rates have been similar for the different levels of general health status, although those declaring poor health at wave 1 presented a slightly higher attrition rate over time (Table 3). At wave 11, 32.1% of the participants with poor health responded, while the response rate among those with excellent, very good or good health was above 40%. In terms of the Government Office Region (GOR) of residence, attrition was higher in Greater London and Scotland between waves 1 and 4, but afterwards, the attrition rates tended to homogenise, except for Wales, which experienced a drop of 21 p.p. in the response rates between waves 4 and 7 (Table 4).

The attrition also affected, to a greater extent, those on lower incomes at wave 1 (Table 5). At wave 11, the response rate for the bottom quintile of income was 32.2%, compared to 49.7% of the top quintile. Regarding the employment status, participants in full-time education or unemployed in 2009-10 presented higher attrition rates in the first four waves; afterwards, the attrition rates decelerated (Table 6). For instance, in wave 4, the response rate of retired

participants was 69.0%, while those in full-time education had a response rate 18.7 p.p. lower. At wave 11, the difference between these two groups had slightly eroded to 14.5 p.p.

Marital status is also related to panel attrition (Table 7). Single participants in wave 1 were more likely to abandon the panel, especially in the first four waves. The attrition rate was also above average for widowed participants after wave 4. These differences can also be partially explained by age, which is closely related to marital status. Regarding education, panel members with no qualifications were more likely to drop out from the study between waves 2 and 7, while the graduated ones presented the lowest attrition rates over the entire period under study (Table 8).

The household characteristics can also be related to response propensities. Lone parents in the initial wave and households with more than two adults and children have been more affected by panel attrition over time (Table 9). Finally, in terms of tenure status, panel members in rented houses have been more likely to drop from the study compared to those in owned properties at the initial wave of the study (Table 10).

The R-indicator and CV for the whole sample are presented in

Figure 1. The plot on the left (a) represents the loss in representativity due to panel attrition for the auxiliary variables included in the analysis. The R-indicator corroborates the findings from the attrition rate analysis in the previous paragraphs: most of the variation in the attrition rates occurred in the first waves of the fieldwork. However, the evolution of the CV tells us a slightly different story: the deterioration of the sample profile occurred linearly. This difference between the two indicators should be explained by the assumption underlying the CV, which represents a worst-case scenario where there is a maximum correlation between nonresponse and the auxiliary variables. Plot b) shows the ability of weights to mitigate the accumulated imbalance of the sample profile. The weights have effectively tackled the effects of attrition with respect to the auxiliary variables included in the analysis. This finding is, to some extent, paradoxical. The auxiliary variables included in the analysis are the same predictors that can be used in the weighting to adjust the deviations since we have the information for respondents and nonrespondents. The good news is that we have auxiliary variables correlated with response which might be correlated to other target variables of the study.
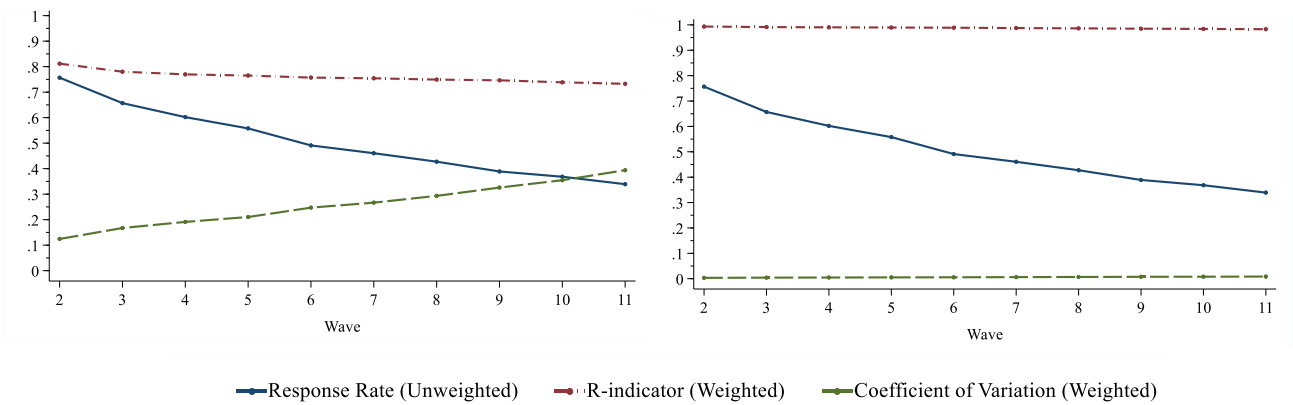
b) Weighted

Figure 1. R-indicators and coefficient of variation for the a) unweighted and b) weighted full sample

The unconditional partial CVs measure the individual contributions of the auxiliary variables to the bias (
Figure 2). The results are in line with the previous graph. In plot a), the unweighted CVs of the individual auxiliary variables linearly increase over time, with age, housing tenure and education contributing the most to the lack of representativeness at wave 11. However, in plot b), the weighted unconditional partial CVs are close to zero and constant over time, meaning that the survey weights are effective in rebalancing the sample profile.

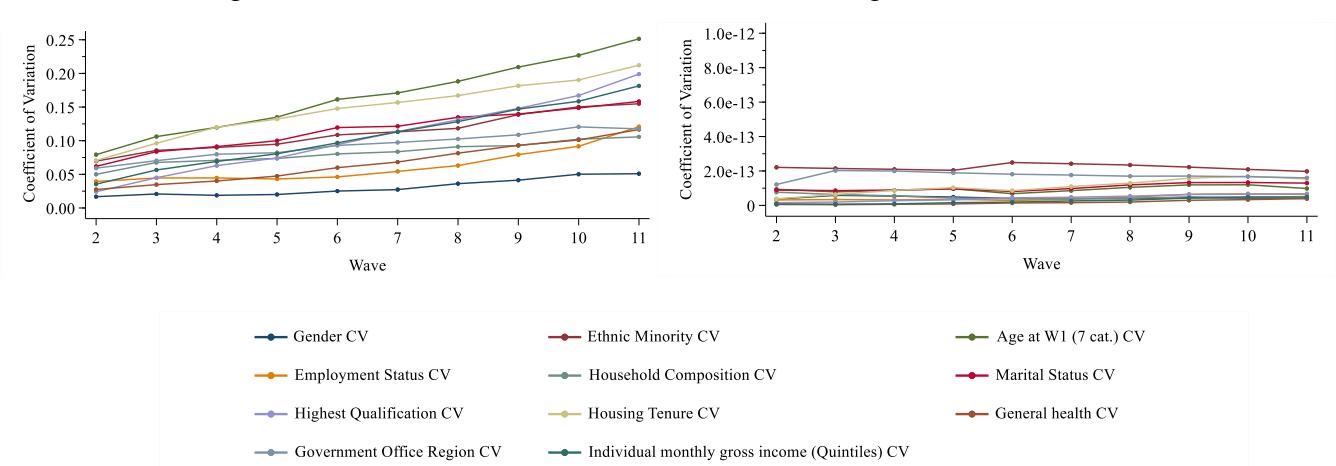a)   Unweighted   b) Weighted

Figure 2. Evolution of partial unconditional coefficients of variation over time for a set of auxiliary variables

IEMB

The IEMB was recruited at wave 6 of *Understanding Society*, in 2014-16; five waves later, 32.7% of respondents to the initial survey completed the adult interview (Table 2). The dropouts were more numerous during the first waves of the study. For instance, the response rate dropped 45.4 p.p. between waves 6 and 8 and 15.2 p.p. between waves 8 and 10.

Table 2. IEMB Attrition: Sex, Age and Ethnic Background

		Wave 6 (2014-16)	Wave 8 (2016-18)	Wave 10 (2018-20)	Wave 11 (2019-21)
Total		4,301	54.6	39.4	32.7
Sex					
	Male	1,964	51.6	35.6	29.6
	Female	2,337	57.2	42.4	35.3
Age in 2014	**16-19**	345	48.8	32.6	27.7
	20-29	957	45.2	30.5	26.4
	30-39	1,140	57.3	39.6	31.7
	40-49	866	58.0	43.3	36.8
	50-59	507	56.7	43.4	37.8
	60-69	241	69.4	57.3	44.9
	70+	220	62.7	48.4	36.2
Ethnic group	**White**	972	53.1	38.8	31.7
	Black	1,028	50.7	30.9	26.4
	Indian	720	58.6	48.5	46.2
	Pakistani	664	64.3	48.5	36.3
	Bangladeshi	211	54.8	42.0	28.8
	Other Asian	301	46.6	32.1	28.1
	Mixed	217	56.0	41.4	33.3
	Other	185	44.3	25.7	18.3

Men show a higher attrition rate than women. At wave 11, 35.6% of females were still responding to the survey compared to 29.6% of males. Younger participants were more likely to drop in the first waves. However, between waves 10 and 11, which include part of the covid-19 pandemic, the attrition rates for the oldest panel members doubled the average. While the average reduction in response rates was 6.7 p.p. in this period, this reduction was above 12 p.p. for participants aged 60 and over. As mentioned in the previous section, this attrition rate might be overestimated due to undetected mortality, which is more pronounced in these age groups. It could also be that the change in survey modes to web and telephone during the pandemic

affected this group to a greater extent (Cabrera-Álvarez et al., 2021). Regarding their ethnic background, participants with Indian, Pakistani and Mixed backgrounds were less likely to abandon the sample, while blacks were more prone to leave the panel. In the transition from wave 10 to 11, Pakistani and Bangladeshi participants presented a more significant attrition rate than the rest groups, 12.2 p.p. and 13.2 p.p., respectively.

The participants' subjective health status is not closely related to the response rates in the IEMB (Table 11). Attrition has affected the extreme groups most, i.e. those with excellent or poor health. At wave 11, the response rates for these groups were very similar, 31.0% and 31.5%, respectively. However, between the last two waves, attrition has eroded the fair and poor health groups to a greater extent than the others. The response rate of participants with fair health dropped by 12.1 p.p., and the poor group suffered a similar change of 13.0 p.p. Panel members from London, the North of England, Scotland, and the Midlands were more likely to leave the study, and in the case of the latter three, the response rate dropped above average between waves 10 and 11 (Table 12).

Participants on lower incomes at the initial wave were more likely to attrit over time. At wave 11, only 1-in-4 participated from the lowest income quintile, while in the highest income group, the response rate was 11 p.p. higher, 36.2% (Table 13). Full-time students and unemployed at the initial wave present higher attrition rates compared to the other groups (Table 14). In the last two waves, a more considerable drop in response rates affected participants who were retired and family carers in wave 6.

Single participants have been more prone to refuse to participate in the study over time: 1-in-2 initial wave respondents did not participate at wave 8, and, at wave 11, the response rate had eroded to 29.2% (Table 15). However, between waves 10 and 11, the most significant drop in response rate was for the widowed participants at the initial wave, 10.0 p.p. Note that this group includes many elderly participants. Regarding education, participants with primary studies or less have suffered the highest levels of attrition, a loss of 72.7 p.p. between the initial wave and wave 11. In contrast, those with a university degree exhibit a lower attrition rate, 63.8 p.p. between waves 6 and 11.

Panel members living on their own and in large households – more than two adults – without children were less likely to stay in the panel (Table 17). After three waves, their response rates were 51.5% and 47.3%, respectively, descending to 29.5% and 28.7% in wave 11. Regarding

household tenure status, panel members in rented private houses or other – neither owned nor rented – were more likely to drop over time (Table 18).

References

Cabrera-Álvarez, P., Burton, J., & Lynn, P. (2021). COVID-19 and Mode Selection Effects in Understanding Society. *Understanding Society Working Papers*, *2021*(10), 32.

Kamisnka, O. (2021). *Weighting for mortality in a longitudinal study*. European Survey Research Association Conference, Online.

Lynn, P., & Borkowska, M. (2018). Some Indicators of Sample Representativeness and Attrition Bias for BHPS and Understanding Society. *Understanding Society Working Papers*, *2018*(01), 19.

Lynn, P., Nandi, A., Parutis, V., & Platt, L. (2018). Design and implementation of a high-quality probability sample of immigrants and ethnic minorities: Lessons learnt. *Demographic Research*, *38*, 513–548. https://doi.org/10.4054/DemRes.2018.38.21

Schouten, B., Cobben, F., Lundquist, P., & Wagner, J. (2016). Does more balanced survey response imply less nonresponse bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *179*(3), 727–748. https://doi.org/10.1111/rssa.12152

Annex A: Attrition in the General Population Sample

Notes to Annex A tables: Cells entries for wave 1 indicate the number of respondents to the adult interview in wave 1 (personal or proxy). The rest of the cells contain the response rate for the subgroup as the percentage of wave 1 respondents who completed the interview in that wave. Ineligible cases were removed from the response rates calculations and, as explained in the methods sections, further adjustments were implemented to deal with under-identified mortality. However, it is likely that remains some undetected ineligibility that might cause the underestimation of the response rates. The undetected ineligibility is likely to increase over time, especially in the oldest age groups.

Table 3. GPS Attrition: General Health Status

	Wave 1 (2009-11)	Wave 4 (2012-14)	Wave 7 (2015-17)	Wave 10 (2018-20)	Wave 11 (2019-21)
Excellent	8,022	62.8	50.4	43.5	41.1
Very good	14,015	65.0	52.5	44.6	41.8
Good	12,068	65.1	51.8	42.9	40.2
Fair	6,355	63.9	49.7	39.7	36.4
Poor	3,150	60.5	45.7	36.0	32.1

Note: General health status was not included in the proxy questionnaire, so analysis for this variable is restricted to sample members who completed the personal interview at wave 1.

Table 4. GPS Attrition: General Office Region (GOR)

	Wave 1 (2009-11)	Wave 4 (2012-14)	Wave 7 (2015-17)	Wave 10 (2018-20)	Wave 11 (2019-21)
North East	1,990	63.3	52.3	42.9	40.8
North West	4,975	64.1	49.2	42.0	38.4
Yorks & Humber	3,774	64.7	54.0	46.1	41.7
East Midlands	3,452	68.8	53.1	44.6	42.3
West Midlands	3,781	62.0	50.2	42.4	39.7
East of England	4,095	66.9	54.0	45.7	43.0
Greater London	4,112	54.9	44.0	34.8	33.6
South East	5,786	65.7	54.1	45.7	43.2
South West	3,802	70.6	56.0	47.0	44.3
Wales	2,299	66.0	45.0	35.6	32.7
Scotland	3,519	59.4	47.2	39.5	36.2
Northern Ireland	2,088	62.6	50.5	41.9	38.3

Table 5. GPS Attrition: Personal Income in Quintiles

	Wave 1 (2009-11)	Wave 4 (2012-14)	Wave 7 (2015-17)	Wave 10 (2018-20)	Wave 11 (2019-21)
Bottom quintile	8,751	58.3	44.1	35.1	32.2
Second quintile	8,771	63.1	48.8	39.8	35.9
Third quintile	8,682	64.0	51.0	41.5	38.4
Fourth quintile	8,735	65.5	52.4	44.8	42.5
Top quintile	8,734	69.6	58.6	51.8	49.7

Note: Income quintiles were derived from the variable a_fimngrs_dv, gross personal monthly income as reported at wave 1.

Table 6. GPS Attrition: Employment Status

	Wave 1 (2009-11)	Wave 4 (2012-14)	Wave 7 (2015-17)	Wave 10 (2018-20)	Wave 11 (2019-21)
Self-employed	3,199	65.6	51.8	43.8	40.4
Paid employment	20,864	65.4	53.2	46.0	43.8
Unemployed	2,566	54.6	40.7	32.1	28.7
Retired	9,620	69.0	55.3	44.7	40.5
Family care or home	2,664	60.0	47.3	37.3	34.5
Full-time student	2,707	50.3	35.6	27.4	26.0
Long-term sick/disabled / others	2,043	62.8	48.0	39.0	34.8

Note: Employment status derived from a_jbstat as reported in wave 1.

Table 7. GPS Attrition: Marital Status

	Wave 1 (2009-11)	Wave 4 (2012-14)	Wave 7 (2015-17)	Wave 10 (2018-20)	Wave 11 (2019-21)
Single	13,404	55.6	42.2	34.0	31.7
Married or Civil Partnership	22,255	68.1	55.3	47.4	44.6
Separated or Divorced	5,086	67.2	54.6	45.7	42.6
Widowed	2,914	67.7	53.3	42.2	37.4

Note: Marital status derived from a_marstat as reported in wave 1.

Table 8. GPS Attrition: Highest Qualification

	Wave 1 (2009-11)	Wave 4 (2012-14)	Wave 7 (2015-17)	Wave 10 (2018-20)	Wave 11 (2019-21)
No qualifications	7,695	59.1	44.1	33.9	29.4
Other qualification	4,700	64.1	49.6	40.2	36.6
GCSE or equivalent	9,211	61.7	48.0	39.1	36.6
A-level or equivalent	8,243	62.2	48.8	40.7	38.2
Degree or other higher	13,759	69.8	58.6	51.5	49.3

Note: Highest qualification derived from a_hiqual_dv as reported in wave 1.

Table 9. GPS Attrition: Household Type

	Wave 1 (2009-11)	Wave 4 (2012-14)	Wave 7 (2015-17)	Wave 10 (2018-20)	Wave 11 (2019-21)

1 adult, no children	6,815	66.0	53.6	44.6	41.5
1 adult, children	2,320	60.4	44.2	33.6	30.3
Couple, no children	13,093	68.0	55.9	48.0	45.2
Couple, children	10,376	65.7	51.4	43.1	39.9
2+ adults, no children	7,024	57.6	45.3	37.9	36.0
2+ adults, children	4,045	58.3	44.8	36.3	33.8

Note: Household type derived from a_hhtype_dv in wave 1.

Table 10. GPS Attrition: Household Tenure

	Wave 1 (2009-11)	**Wave 4 (2012-14)**	**Wave 7 (2015-17)**	**Wave 10 (2018-20)**	**Wave 11 (2019-21)**
Owned outright	13,209	68.9	56.4	47.9	44.8
Owned with mortgage	16,895	67.8	54.8	47.1	45.1
LA/housing association rented	7,295	58.7	43.3	32.9	28.7
Rented private	5,669	49.2	37.8	30.2	27.5
Other	509	61.5	43.1	35.8	33.0

Note: Household tenure derived from a_tenure_dv in wave 1.

Annex B: Attrition in the Immigrant and Ethnic Minority Boost Sample

Notes to Annex B tables: Cells entries for wave 6 indicate the number of respondents to the adult interview in wave 6 (personal or proxy). The rest of the cells contain the response rate for the subgroup as the percentage of wave 6 respondents who completed the interview in that wave. Ineligible cases were removed from the response rates calculations and, as explained in the methods sections, further adjustments were implemented to deal with under-identified mortality. However, it is likely that remains some undetected ineligibility that might cause the underestimation of the response rates. The undetected ineligibility is likely to increase over time, especially in the oldest age groups of the sample.

Table 11. IEMB Attrition: General Health Status

	Wave 6 (2014-16)	Wave 8 (2016-18)	Wave 10 (2018-20)	Wave 11 (2019-21)
Excellent	1,035	49.4	36.6	31.0
Very good	1,211	56.4	40.5	36.2
Good	1,067	58.2	42.1	35.4
Fair	383	63.4	47.2	35.1
Poor	188	54.7	44.5	31.5

Note: General health status was in the self-completion questionnaire and in the main questionnaire for proxy interviews, so for this analysis we combined both variables.

Table 12. IEMB Attrition: Government Office Region (GOR)

	Wave 6 (2014-16)	Wave 8 (2016-18)	Wave 10 (2018-20)	Wave 11 (2019-21)
North of England & Scotland	620	57.9	47.8	31.9
Yorks & Humber	464	57.4	46.6	40.2
East Midlands	77	53.3	40.3	29.6
West Midlands	492	60.9	44.5	36.4
East of England	157	59.2	41.0	37.5
Greater London	2,095	50.2	33.2	28.7
South East	285	61.3	41.0	41.5
South West	111	57.7	46.8	37.2

Note: GOR had small counts in some cells, such as Scotland or North East, due to the sampling design of the IEMB, so these groups were combined with North West.

Table 13. IEMB Attrition: Personal Income in Quintiles

	Wave 6 (2014-16)	**Wave 8 (2016-18)**	**Wave 10 (2018-20)**	**Wave 11 (2019-21)**
Bottom quintile	861	47.5	32.3	25.9
Second quintile	860	53.2	40.1	33.0
Third quintile	860	55.0	39.8	31.6
Fourth quintile	860	59.6	43.1	36.8
Top quintile	860	57.6	41.5	36.2

Note: Income quintiles were derived from the variable f_fimngrs_dv, gross personal monthly income as reported at wave 6.

Table 14. IEMB Attrition: Employment Status

	Wave 6 (2014-16)	**Wave 8 (2016-18)**	**Wave 10 (2018-20)**	**Wave 11 (2019-21)**
Self-employed	413	54.3	35.9	32.0
Paid employment	1,915	54.9	40.0	33.7
Unemployed	412	50.9	33.2	26.1
Retired	342	66.5	52.3	42.4
Family care or home	474	61.0	44.2	34.1
Full-time student	547	46.5	32.6	28.5
Long-term sick or disabled & others	172	51.5	41.6	34.4

Note: Employment status derived from f_jbstat as reported in wave 6.

Table 15. IEMB Attrition: Marital Status

	Wave 6 (2014-16)	**Wave 8 (2016-18)**	**Wave 10 (2018-20)**	**Wave 11 (2019-21)**
Single	1,683	50.1	34.0	29.2
Married or Civil Partnership	2,148	58.4	43.9	35.6
Separated or Divorced	293	54.6	38.8	34.1
Widowed	120	61.9	46.0	36.0

Note: Marital status derived from f_marstat as reported in wave 6.

Table 16. IEMB attrition: Highest Qualification (ISCED11)

	Wave 6 (2014-16)	**Wave 8 (2016-18)**	**Wave 10 (2018-20)**	**Wave 11 (2019-21)**
Primary or less	332	52.2	38.4	27.3
Secondary and post-secondary	1,688	54.7	39.1	31.9
Other higher	854	54.0	41.3	35.5
Degree	1,032	57.1	40.5	36.2
Other	292	58.0	39.3	30.3

Note: A substantive part of the IEMB obtained their qualifications out of the UK and they were asked using ISCED 11, an international classification developed by UNESCO. The education variable is a combination of the ISCED 11, for those getting their qualifications abroad, and the highest qualification obtained in the UK.

Table 17. IEMB Attrition: Household Type

	Wave 6 (2014-16)	Wave 8 (2016-18)	Wave 10 (2018-20)	Wave 11 (2019-21)
1 adult, no children	577	51.5	35.4	29.5
1 adult, children	296	55.3	39.9	32.7
Couple, no children	573	52.9	43.3	34.9
Couple, children	1,075	62.5	44.1	37.1
2+ adults, no children	1,018	47.3	34.3	28.7
2+ adults, children	752	57.0	39.4	32.7

Note: Household type derived from f_hhtype_dv in wave 6.

Table 18. IEMB Attrition: Household Tenure Status

	Wave 6 (2014-16)	Wave 8 (2016-18)	Wave 10 (2018-20)	Wave 11 (2019-21)
Owned outright	643	68.3	57.4	49.3
Owned with mortgage	816	62.4	46.3	42.7
LA/housing association rented	1,110	57.0	37.9	29.4
Rented private	1,131	46.2	31.6	24.6
Other	113	36.7	30.5	21.9

Note: Household tenure derived from f_tenure_dv in wave 6.