

Appendix: Response propensity modelling

Tables 1 and 2 below show the discriminatory power of the models considered in the preliminary analysis (i.e. prior to the adjustment for the sample design). Table 1 presents the results for a series of three stage regression analyses. Model 1 simply applies the outcome history dummies as predictors of the likelihood of being contacted. Model 2 adds selected household variables to this model and model 3 adds selected neighbourhood characteristics to complete the stages. Table 2 re-evaluates the models in paired comparisons by considering their impact on the ROC analysis. The change in ROC provides a means of discriminating between the models.

Table 1: Contact propensity models

Model No	Explanatory set	N	Pseudo R-Square	Pseudo Likelihood	ROC
1	Outcome history	16905	0.1430	-2834.8276	0.7043
2	And Household Information	16534	0.2001	-2545.6606	0.8297
3	And Neighbourhood information	16359	0.2015	-2541.1435	0.8312

Table 2: Comparison of contact propensity models

Model pairs	First ROC	Second ROC	Chi-square Statistic (df = 1)	P-Value
1 vs 2	0.7043	0.8297	60.44	0.0000
1 vs 3	0.7043	0.8312	61.58	0.0000
2 vs 3	0.8297	0.8312	2.06	0.1513

Model 2 (history of past contacts and household information) emerges as the clear favourite in terms of predictive power. This is generally to be expected given the longitudinal nature of the data. What was more surprising was the marginal (and in fact not statistically significant) role played by neighbourhood information. It is possible that as the data was collected in wave 2, the effects on contact at wave 4 were weakened sufficiently to no longer be significant. Also, the method of collection of the data (which was based on the subjective judgment of interviewers) may have carried a high degree of variability that made the predictive potential of this information less effective.

Final Contact Model

Number of strata = 9

Number of PSUs = 398

Number of obs =
16539

Population size = 16593.6

Design df = 389

F(21, 369) =
54.81

Prob > F =
0.0000

Contact Event	Odds ratio	Linearized S.E.	t value	P>t	[95% Conf. Interval]	
Contact in both wave 2 and 3	21.1	4.003	16.05	0.000	14.511	30.621
Contact in wave 2 only	2.1	0.489	3.17	0.002	1.325	3.316
Contact in wave 3 only	7.5	1.790	8.41	0.000	4.674	11.976
Voted in Last General Election	1.5	0.161	3.67	0.000	1.203	1.839
Other Housing Arrangements	0.5	0.089	-3.82	0.000	0.369	0.726
Renting	0.5	0.063	-5.38	0.000	0.409	0.660
Parents (Couple)	1.3	0.145	1.93	0.055	0.995	1.572
Parent gender: Female	1.2	0.130	1.98	0.049	1.001	1.516
Parent Ethnicity: Pakistani and Bangladeshi	0.7	0.109	-2.14	0.033	0.539	0.974
Parent Ethnicity: Black or Black British	0.7	0.125	-1.89	0.060	0.511	1.014
Parent Ethnicity: Other Ethnic Group	0.5	0.123	-2.78	0.006	0.319	0.822
Achieved qualification: NVQ5	1.6	0.164	4.44	0.000	1.292	1.940
Achieved qualification: NVQ4	1.5	0.302	1.96	0.051	0.998	2.216
Achieved qualification: NVQ3	1.8	0.405	2.60	0.010	1.153	2.798
Achieved qualification: NVQ2	5.9	1.698	6.24	0.000	3.392	10.423

Achieved qualification: NVQ1	2.2	0.780	2.12	0.034	1.059	4.390
Contact adult born after 1980	0.4	0.090	-4.06	0.000	0.275	0.639
Contact adult after 1965 and before 1981	0.6	0.113	-2.65	0.008	0.428	0.882
Accommodation: House	1.3	0.174	2.07	0.039	1.014	1.706
Wales (Status: Advantaged)	2.0	0.599	2.45	0.015	1.152	3.641
Wales (Status: Disadvantaged)	1.5	0.203	2.85	0.005	1.129	1.937