Evaluating the immediate and longer term impact of a refusal conversion strategy in a large scale longitudinal study

Lisa Calderwood
Centre for Longitudinal Studies
UCL Institute of Education
London, UK

Ian Plewis
School of Social Sciences
University of Manchester
Manchester, UK

Sosthenes Ketende
Johns Hopkins Bloomberg School
of Public Health
Baltimore, USA

Tarek Mostafa
Centre for Longitudinal Studies
UCL Institute of Education
London, UK

Refusal conversion is one of the fieldwork strategies commonly used to minimise non-response in surveys. There is, however, relatively little evidence about the effectiveness of this strategy, particularly for face-to-face longitudinal surveys. Moreover, much of the existing evidence is based on observational studies. This paper evaluates the effectiveness of a fieldwork strategy – intensive re-issuing – to convert refusals using evidence from an intervention on a random sub-sample of refusals implemented in wave four of a large scale longitudinal study in the UK: the Millennium Cohort Study. We show that intensive re-issuing is an effective way of reducing the refusal rate. We also show that refusal conversion led to a modest reduction in non-response bias in the survey estimates for several key variables. The longer term impact of refusal conversion is also a key concern in longitudinal surveys. We demonstrate that, although the majority of converted refusals go on to participate in the subsequent wave of the study, there is no overall effect of intensive re-issuing on sample size at this wave.

Keywords: non-response; fieldwork intervention; longitudinal study; representativity; Millennium Cohort Study

1 Introduction

Most surveys typically devote considerable resources to maximising response rates because non-response is the main component of non-observational error in sample surveys (Groves, 1989). In the context of declining response rates over recent years (de Leeuw & de Heer, 2002), survey organisations have had to make increased efforts to maintain response rates (Stoop, 2005)) and there has been a growth in methodological research into the effectiveness of fieldwork strategies to minimise non-response.

Higher response rates do not, however, necessarily imply lower levels of non-response bias. In their meta-analysis, Groves and Peytcheva (2008) point out that the extent to which higher response rates generate less bias depends in part on the correlation between the predictors of survey participation and the substantive survey variables and conclude that the level of bias can differ between surveys with similar response rates and between variables of interest within the same survey.

Refusal conversion is one of the fieldwork strategies commonly used to minimise non-response. This paper evaluates the effectiveness of refusal conversion in an experimental framework within an ongoing longitudinal study by assessing its impact on response rate and non-response bias at the implementation wave, and participation at the subsequent wave.

The next section reviews the literature in relation to refusal conversion and non-response bias. Section 3 presents the design of the intervention to convert refusals and provides details of its implementation and the study, the UK Millennium Cohort Study, on which it was carried out. Section 4 provides results from the intervention. Section 5 concludes and reflects on the implications of the findings for fieldwork strategies in longitudinal studies.

2 Refusals, refusal conversion and non-response bias

There is a large body of theoretical and empirical literature about refusals (e. g. Groves, Cialdini, & Couper, 1992;

Groves & Couper, 1998) and it is well established that survey design factors such as respondent incentives (e. g. Laurie & Lynn, 2009; Singer, 2002) and advance letters (de Leeuw, Callegaro, Hox, Korendijk, & Lensvelt-Mulders, 2007) can successfully reduce refusal rates. Refusal conversion is another well established fieldwork strategy for reducing refusal rates (e. g. Stoop, 2004). This practice involves re-approaching a sampled person who has initially refused to take part and attempting to get them to reconsider their decision, i.e. to 'convert' them from a refusal to a successful interview. Conversion is usually, but not necessarily, attempted by a different and often a more experienced interviewer. The theoretical rationale for attempting refusal conversion is that a sample member's decision to co-operate is influenced by their interaction with the interviewer. Hence using a different interviewer will lead to a different interaction and, hopefully, a positive decision about participating. It is also based on evidence that refusals often result from the particular circumstances of the sample member when contacted by the interviewer and that the same or a different interviewer recontacting on a different occasion may mean that the circumstances which led to the refusal are no longer pertinent. It is well-established (and evidenced by non-monotonic or arbitrary response patterns) that refusal to participate at a particular wave of a longitudinal survey, conditional on participation at the baseline wave, does not necessarily mean that the sample member will not take part in future waves. The theoretical rationale for using more experienced interviewers for refusal conversion is that it is well known that interviewer characteristics can have a considerable impact on refusal rates (Hox & de Leeuw, 2002). Converted refusals constitute a significant minority of completed interviews in many surveys. Lynn, Clarke, Martin, and Sturgis (2002) report that converted refusals accounted for between 1.2% and 8% of all completed interviews on six UK face-to-face surveys conducted between 1995 and 1998. Curtin, Presser, and Singer (2000) report that on the Survey of Consumer Attitudes, a long-running repeated cross-sectional telephone survey in the US, the proportion of interviews from refusal conversions doubled from 7.4% in 1979 to 14.6% in 1996.

Higher response rates increase precision but, as noted earlier, they do not necessarily result in less non-response bias. There are several examples in the literature from cross-sectional telephone surveys in the US which demonstrate that, although refusal conversion (and other extended field efforts to maximise response) can have a positive impact on response rates, there is little or no evidence that this is beneficial in terms of reducing non-response bias (Curtin et al., 2000; Keeter, Miller, Kohut, Groves, & Presser, 2000). One implication of these papers is that the additional resources devoted to response maximisation on these surveys may not be justified as they appear to bring little or no benefit in terms of bias reduction. The focus of this paper is, however, on

large-scale longitudinal studies where analyses are usually based on multiple observations of the same unit and so bias reduction at a single point in time is less relevant.

A similar approach to assessing the impact of refusal conversion (and other extended interviewer efforts) on non-response bias was taken by Lynn et al. (2002). They found that refusal conversion did appear to bring some benefit for the six cross-sectional surveys they examined in terms of bias reduction for survey estimates relating to financial variables as statistically significant differences were found between those initially interviewed and converted refusals. This conclusion was not, however, replicated for health or attitude variables.

Longitudinal surveys are typically better placed to assess the impact of refusal conversion on non-response bias as, unlike cross-sectional surveys, information about most if not all sample members is available from prior waves. However, the context of refusal conversion is different for a longitudinal study compared with a cross-sectional survey as refusal conversion can take place both within and across waves of data collection i.e. refusals can be re-approached at subsequent waves as well as (or instead of) during the current wave of data collection. The report by AAPOR (2014) on survey refusal includes a section on refusal conversion although it does not cover longitudinal studies. But it points to the importance of being able to take informed decisions about when it is most cost-effective to devote resources to refusal conversion. The balance of risks and rewards in relation to refusal conversion in the current wave of data collection is different in longitudinal surveys when the objective of securing participation needs to be repeated at each wave of data collection and there will be another opportunity to convert refusals at a future wave. So, while maximising response at a particular wave is important, this short-term aim needs to be balanced against the desire not to jeopardise participation in future waves. Similarly, the long-term impact of refusal conversion on non-response bias and response rates in longitudinal surveys cannot be fully assessed until future waves of data collection have taken place.

Burton, Laurie, and Lynn (2006) evaluate the long-term effectiveness of within-wave refusal conversion procedures on a household panel survey. They use pooled data from waves 4–13 of the British Household Panel Survey (BHPS) and report that within-wave refusal conversion was attempted for around 36% of refusals (ranging from 22% to 48% across the 10 waves) and 37% (25% to 47% across waves) of these were converted to a face-to-face, proxy or telephone interview giving an overall unconditional conversion rate of 13%. They also show that the majority of converted refusals went on to participate at the subsequent wave, and that between 24% and 47% were still participating at wave 13. The wave 13 sample included an extra 584 cases (6.7% of the total) as a result of refusal conversion across waves four to 13. The

paper shows that refusal conversion on the BHPS improved the representation in the sample of certain groups (such as the geographically mobile, self-employed and local authority renters) and, to the extent that these variables are correlated with other variables of interest, notably measures of change, may be expected to reduce non-response bias. Other than Burton et al. (2006), there appears to be no published evidence about the effectiveness of refusal conversion attempts in longitudinal studies.

A limitation of almost all of the literature assessing the impact of refusal conversion on bias reduction is that it is based on observation of fieldwork procedures rather than randomised interventions. There are only a few studies which experimentally evaluate the effectiveness of different refusal conversion techniques and these have been for telephone surveys (Basson & Chronister, 2006; Keeter et al., 2000). This is problematic because estimates of the effectiveness of refusal conversion attempts, including their impact on non-response bias, will be influenced by selection effects. In particular, there is a concern that allowing field staff in the office discretion over which refusals to re-issue may mean that a minority of cases are re-issued and the field staff are likely to re-issue cases which they judge most likely to respond, which in practice are likely to be those with similar characteristics to those who have already been interviewed. Although this may result in an increase in response rate, it may not result in a reduction in bias. The intervention described in the following section was specifically designed to address this limitation of the existing evidence.

3   Design and implementation of the intervention

3.1   Survey context

The intervention described in this paper was developed for and carried out on the fourth wave of the Millennium Cohort Study (MCS). The MCS, which is following over 19,000 children born in 2000/1, is one of four national birth cohort studies in the UK. The sample, which was recruited through records of those in receipt of a universal benefit paid to parents called Child Benefit, is disproportionately stratified and clustered at the level of electoral ward. The stratification is based on UK country and the characteristics of the sampled wards. In Scotland, Wales and Northern Ireland, there are two strata in each country: disadvantaged and advantaged. In England, there are three strata: minority ethnic, disadvantaged and advantaged. The probabilities of selection differ by strata. More details of the MCS sample design can be found in Plewis (2007). The data collection for the study takes place in the home and involves face-to-face interviews with multiple informants in each family. Interviews have been sought with up to two co-resident parents at every wave and, from wave two onwards, the child has also been asked to participate directly though the nature and extent of their

participation has changed at each wave as the children get older. There have been five waves of the study so far: at nine months (2001–2), age three years (2003–4), age five (2006), age seven (2008) and age 11 (2012).

The MCS employs a range of response maximisation techniques, including between-wave mailings, advance letters and leaflets, and multiple call-backs. All interviewers working on the project have received the fieldwork agency's standard training on probability surveys which covers refusal avoidance and additional project-specific training. Although interviewers are encouraged to make first contact by telephone for families who took part in the most recent wave, they are trained to withdraw if the family appears reluctant to take part on the phone and they then make a personal visit to the family at a later date. More generally, interviewers are trained not to return cases as refusals until reasonable efforts have been made to persuade the sample member to take part. These would typically include visiting the household on more than one occasion (with the exception of extremely firm refusals), making multiple attempts to re-arrange broken appointments and attempting to speak to all household members eligible for interview.

3.2   Design of the intervention

The intervention presented here was part of an experiment, described in detail in Calderwood, Plewis, Ketende, and Taylor (2010), which was designed to test the efficacy of strategies to reduce non-response in an ongoing longitudinal study. There were two aspects of that experiment: (i) whether a leaflet designed to encourage sample members to take part by addressing reasons for refusal that are commonly reported on the study would lead to higher conversion rates for refusal cases who were re-issued; (ii) whether re-issuing all refusals increased the proportion of refusals who were converted to productive interviews more than the standard re-issuing strategy of the fieldwork agency i.e. re-issuing a non-random sub-set of refusals. These two aspects were crossed to create a 2x2 factorial design and cases were randomly assigned to one of these four groups at the point when cases were initially issued to the field. Calderwood et al. (2010) provide further details about the randomisation process and its outcomes (in Appendix B).

Calderwood et al. (2010) show that the leaflet did not have any effect in terms of refusal conversion. In addition, the fieldwork agency did not, in fact, implement their standard strategy in the usual way so that only seven cases were re-issued in what was meant to be a control condition but turned out, in essence, to be a no re-issue group. Thus, the focus of this paper is on the effectiveness of intensive re-issuing. In that sense, the paper is an extension of Burton et al. (2006) for a birth cohort study but with one important difference: the cases subject to intensive re-issuing were not purposively selected by field staff, rather they were a randomly selected

Table 1
Refusals, exclusions and actual treatment by group

	Intensive re-issue		No re-issue	
	n	*%*	*n*	*%*
Number of Refusals[a]	803	48.3	857	51.6
Excluded as 'hard' refusal'[b]	180	22.4	128	14.9
Outcome code[b]	107	13.3	123	14.3
Office field staff decision[b]	73	9.1	5	0.6
Excluded for other reason[b]	13	1.6	9	1.1
Number of treated refusals[b]	610	76.0	7	0.8

[a] Percentage of refusals (overall)
[b] Percentage of refusals (within intervention group)

50% of all refusals. Randomisation ensured internal validity but the external validity or generalisability of the results is conditional on the fieldwork agency's policy of refusal conversion being carried out by different and more experienced interviewers.

Although the intention of the intensive re-issuing treatment was to re-issue all refusals, some were classified as 'hard' refusals and therefore not considered for re-issuing. The majority of 'hard' refusals were excluded automatically on the basis of their outcome code alone. Specifically, office refusals, refusals during tracking and refusals during the interview were excluded automatically. It was agreed with the fieldwork agency that is was not appropriate to consider re-issuing these cases, and hence they were not reviewed individually by office field staff. All other refusal outcomes – the most common of which were refusals given to the interviewer in person – were considered for re-issue. Field staff (not interviewers) reviewed these cases individually and decided whether or not they should be re-issued, making their judgement based on available information, in particular interviewer notes on reasons for refusal given by the cohort family. It was felt to be important that field staff retained the ability to do this to avoid interviewers re-approaching families where it would not have been appropriate to do so. In general cases were re-issued unless there was a clear reason not to do so e.g. threatening or aggressive behaviour towards the interviewer, notes from the interviewer to indicate an extremely firm and unequivocal refusal or serious illness in the family.

3.3 Implementation of the intervention

In total there were 1660 refusals (11% of the issued sample of 15350 cases in Great Britain; Northern Ireland is omitted for operational reasons). These are household rather than individual level refusals i.e. no interviews were conducted. Typically the refusal would be given by one or both of the parents on behalf of the family. Families in which some but not all individual interviews were conducted are classified as partially productive and were not considered for re-issuing.

Table 1 shows the number of refusals, exclusions and actual treatment by experimental group.

Overall, the exclusions reduced the number of intensively re-issued cases by 24%, from 803 to 610. The non-random exclusion of cases from the intervention group introduced an element of non-random selection into the actual treatment given. Although this has the potential to jeopardise the validity of the conclusions based on the initial randomisation, the standard solution to this problem is to estimate treatment effects for all cases for which there was an intention to treat as well as to estimate the effects of treatment on the treated (Shadish & Cook, 2009) i.e. compare conversion rates for all refusals as well as treated refusals.

In the intensive re-issuing group, a very high proportion of refusals were re-issued (76%), a much higher proportion than were re-issued on previous waves of the MCS and in other comparable surveys. For example, in wave two of MCS, around 13% of refusals were re-issued and on the BHPS waves 4–13, 36% of refusals were re-issued (Burton et al., 2006).

4 Results

4.1 What proportion of refusals was converted to productive interviews in the intervention group?

Table 2 shows the final survey outcome for all treated refusals. Fully productive cases are those in which all individual interviews were conducted, partially productive where some but not all interviews were carried out, and unproductive cases are those in which no interviews were done. Treated refusals are defined, as explained in Table 1, as those refusals which were not excluded. Table 2 shows that when intensive re-issuing was administered almost a quarter of treated (i.e. re-issued) refusals (and 17.3% of all refusals) were successfully converted to a productive interview. Compared with Burton et al. (2006), the conversion rate for treated refusals (22.8%) in the intensive re-issue group is lower than the equivalent proportion reported for the BHPS (37%). However, the unconditional conversion rate i.e. the

Table 2
Refusals, exclusions and actual treatment by group

	Intensive re-issue	
	n	*%*
Fully productive	98	16.1
Partially productive	41	6.7
Unproductive	471	77.2
Total (*N*)	610	
Productive (%)[a]	22.8	
Productive (%)[b]	17.3	

[a] Percentage based on all treated refusals. Productive includes partially productive.
[b] Percentage based on all refusals.

overall proportion of refusals converted (17.3%) is higher than the equivalent unconditional conversion rate reported by Burton et al. (2006) of 13.5%.

4.2   What impact did re-issuing have on the overall achieved sample size and refusal rate?

Intensive re-issuing reduced the refusal rate by around two percentage points (from 11% to 9%). This indicates that, if the intensive re-issuing treatment had been carried out on the whole sample, the increase in achieved sample size would be expected to have been double what was actually observed i.e. 278 cases rather than 139.

4.3   Were converted refusals less likely than those initially interviewed to complete all of the survey elements?

The survey consisted of several different data collection elements. Families who were interviewed initially were much more likely to have completed all of the survey elements they were eligible for (i.e. to be fully productive) than families who were converted refusals (87% compared with 72%). The ratio of fully productive to partially productive families was 2.4 for converted refusals compared with 7.7 for those who did not refuse initially.

The most striking difference between the two groups was in the proportion of partner respondents who were interviewed. In productive families who were initially interviewed, almost 85% of eligible partners were interviewed compared with 63% among productive families who were converted refusals.

4.4   How were the refusal, re-issuing and conversion rates related to prior response history?

Prior response history is strongly associated with both initial refusal and refusal conversion rates at wave four. Families who did not take part at wave three were more likely to refuse (33% compared with 7%) and less likely to be converted to a productive interview (10% of all refusals compared with 18%) at wave four than those who did take part at wave three. Similarly, families who had taken part in all prior waves they were eligible for were less likely to refuse (6% compared with 27%) and more likely to be converted to a productive interview (20% of all refusals compared with 11%) at wave four than those who had missed at least one wave.

4.5   How were the refusal, re-issuing and conversion rates related to sample characteristics?

Table 3 shows refusal rates, conversion attempt rates and conversion success rates for all cases in the intensive re-issue group (n = 610) who were interviewed at the previous wave (wave three; n = 531; 87%) by a range of characteristics observed at wave three. The wave three characteristics chosen have been shown to be related to attrition by Plewis, Ketende, Joshi, and Hughes (2008), and to key outcome variables on MCS. This approach of examining the impact of refusal conversion on characteristics which are *predictors* of key outcome variables, rather than examining outcome variables directly, is appropriate given the multi-disciplinary and multipurpose nature of the MCS. This approach was taken by Burton et al. (2006) and our chosen characteristics are intended to maximise comparability with their study.

The aim of the intensive re-issuing treatment was to increase the proportion of refusals which were reissued – the conversion attempt rate – with the intention that the observable demographic characteristics which are related to initial refusal would not be taken into account in the decisions about whether or not to re-issue the case. In terms of bias reduction, it is preferable for sample members with a higher propensity to refuse also to have higher conversion rates as this is likely to lead to a reduction in bias on this variable. Conversely, if the conversion rate is highest for sample members who have a low propensity to refuse this can lead to an increase in bias in the achieved sample.

Table 3 shows, as Plewis et al. (2008) did for wave two, that the main respondent's education level, employment status, health, voting behaviour as well as family type and housing tenure are all statistically significantly related to the refusal rate. Refusals were more common among those with no or lower (levels 1 and 2) educational qualifications, those who were not in work, in fair or poor health, those who did not vote in the last election, lone parent and other family types and those in rented or other tenure accommodation. The refusal rate did not vary significantly with the ethnic group of the cohort member suggesting a relative lack of bias in this variable.

The table also shows that for all of the characteristics associated with refusal rates there were no statistically significant differences in both the conditional and unconditional conver-

Table 3
Refusal and conversion rates at wave four by wave three characteristics

	Number of refusals (wave four) N	Refusal rate[a] %	Conversion attempt rate[b] %	Conditional conversion rate[c] %	Unconditional conversion rate[d] %
	531	6.9	72.9	25.3	18.4
Main respondent's educational qualifications					
No qualifications	108	13.5	75.3	16.4	12.4
Level 1	49	8.4	83.1	28.1	23.3
Level 2	153	6.9	78.9	28.5	22.5
Level 3	66	5.8	71.6	20.0	14.3
Level 4	111	5.3	62.5	32.2	20.1
Level 5	26	5.6	61.8	27.3	16.9
Overseas qualifications only	17	6.6	82.8	8.4	6.9
F-statistic[e]		8.64	1.89	1.16	0.97
p-value		<.001	<.1	>.3	>.4
Whether main respondent is in work (including on leave)					
Yes	255	5.9	71.2	29.8	21.2
No	273	8.2	74.5	20.9	15.5
F-statistic		14.15	0.49	3.19	2.18
p-value		<.001	>.4	<.1	>.1
Cohort member's ethnic group					
White	418	6.7	73.0	26.1	19.1
Non-white	113	8.1	72.2	20.4	14.7
F-statistic		2.29	0.02	0.85	0.91
p-value		>.1	>.8	>.3	>.3
Whether main respondent voted in last general election					
Yes	255	5.3	64.5	25.0	16.1
No	262	9.1	80.2	26.6	21.2
F-statistic		28.42	9.44	0.07	1.59
p-value		<.001	<.003	>.7	>0.2

Continues on next page

sion rate. This clearly shows that, in relation to these characteristics, comparable conversion rates can be achieved among cases with high initial refusal rates as among those with low initial refusal rates, and indicates that re-issuing may have led to a reduction in bias in these variables (although it should be remembered that we do not take into account non-response before wave three here and our focus is on bias arising from refusal rather than from not being located or contacted).

Overall, the evidence from Table 3 lends support to the fieldwork strategy of re-issuing a high proportion of refusals, including groups with high refusal rates, as it is clear that high conversion rates can be achieved among some of these groups.

4.6 What impact did re-issuing have on bias in the achieved sample?

The discussion in the previous section provides some indicative evidence on the impact that re-issuing may have had on bias by examining the correlates of refusal rates and conversion rates. This section aims to assess this more directly in two ways. The first is by comparing the distributions of those initially interviewed, converted refusals and unconverted refusals on the same wave three characteristics as reported in Table 3. Clearly, this is a relative assessment of bias i.e. how much bias there is at wave four compared with wave three and does not account for any bias existing at wave three. As for Table 3, this analysis is confined to cases who were interviewed at wave three and so excludes the small proportion

Table 3, continued from previous page

	Number of refusals (wave four) N	Refusal rate[a] %	Conversion attempt rate[b] %	Conditional conversion rate[c] %	Unconditional conversion rate[d] %
	531	6.9	72.9	25.3	18.4
Main respondent's general health					
Excellent/Very good	286	6.3	70.9	26.3	18.7
Good	145	6.6	70.4	24.1	17.0
Fair/Poor	88	8.9	82.0	26.4	21.7
F-statistic		3.26	1.77	0.07	0.29
p-value		<.1	>.1	>.9	>.7
Family Type					
Married or cohabiting natural parents	378	6.4	69.7	27.0	18.8
Lone natural mother	111	8.1	84.5	21.1	17.9
Other family type	42	9.2	72.3	23.1	16.7
F-statistic		3.28	3.95	0.49	0.06
p-value		<.05	<.03	>.6	>.9
Housing Tenure					
Own	278	5.4	67.0	26.7	17.9
Rent/Other	240	9.7	80.4	24.7	19.9
F-statistic		44.94	10.6	0.13	0.24
p-value		<.001	<.002	>.7	>.6

[a] This is based on the total sample responding at wave three (n = 6806).

[b] The conversion attempt rate is the proportion of refusals which are re-issued.

[c] The conditional conversion rate is the proportion of re-issued refusals which are interviewed.

[d] The unconditional conversion rate is the proportion of all refusals which are interviewed.

[e] Design-based F tests (degrees of freedom omitted) were used to test the null hypothesis of no relationship between each of the variables and the refusal rate, conversion attempt rate, conditional conversion rate and unconditional conversion rate. The analysis was carried out using the svy commands in Stata to adjust for the sample design.

of cases (13%) that were not interviewed at wave three but were interviewed at wave four. The second way is to use R (for representativity)-indicators (Schouten, Cobben, & Bethlehem, 2009) to compare the actual achieved sample at wave four with the reduced sample that would have been achieved had there been no intensive re-issuing. This is essentially the approach taken by Plewis, Calderwood, and Mostafa (2016), to assess the value of a sample maintenance strategy in terms of its potential to reduce bias in estimates of interest.

The first approach taken to assessing bias reduction is similar to the approach taken by Burton et al. (2006) and involves comparing the characteristics of both converted refusals and unconverted refusals with initially interviewed cases. If (i) the distributions for converted refusals and the initially interviewed are different and the distributions of unconverted refusals and the initially interviewed are the same or (ii) the distributions of both converted and unconverted refusals are different from the distributions for those initially inter-

viewed, this indicates that there is potential bias in the survey estimates for this variable and that the refusal conversion attempts are likely to have led to a reduction in this bias. Conversely, if the distributions of converted refusals and the initially interviewed are similar, and the distributions of unconverted refusals and the initially interviewed are different, this indicates that there is bias in the survey estimates for this variable and that refusal conversion attempts are unlikely to have led to a reduction in this bias. If the distributions of both converted and unconverted refusals are similar to the distribution of the initially interviewed, this indicates that there is little or no bias in the survey estimates for this variable.

Bias reduction was assessed by a series of binary, unordered and ordered logistic regression models with the characteristics of interest as the dependent variables and the three interview outcomes included as two dummy explanatory variables in a single model with initially interviewed as the reference category. The results are given in Table 4 as

estimated parameters from the models along with their standard errors.

Table 4 shows that the results are mixed and differ between variables. For employment status, education level, ethnic group and family type, the distributions of unconverted refusals are significantly different from the initially interviewed cases but there was no significant difference between converted refusals and initially interviewed cases for these variables. This indicates that there could be bias in this variable at wave four which refusal conversion has done little or nothing to remove. For health, the distributions for converted refusals and unconverted refusals were not significantly different from the distribution for the initially interviewed cases. This indicates that there is little or no bias in this variable at wave four. However, for voting and housing tenure, there is evidence that refusal conversion has led to a removal of bias. Both converted refusals and unconverted refusals have significantly different distributions from the initially interviewed cases and similar distributions to each other: tests of the equality of the coefficients in the two columns of estimates in Table 4 gave p-values >.2 and >.6 for voting and tenure. This indicates that the refusal conversion process has removed some bias in these variables (although our assessment of bias must, perforce, rely on internal comparisons rather than any external benchmark).

Overall, these results show that it is plausible that refusal conversion has led to a statistically significant reduction in bias in the survey estimates for two of the seven characteristics considered i.e. voting and housing tenure. This is supported by the evidence in the previous section (4.5) showing that for these variables the groups with the highest refusal rates, i.e. non-voters and those living in rented accommodation, also had high conversion rates. Although the refusal rate did not vary significantly with ethnic group, the analysis in this section shows that the converted refusals were similar to the initially interviewed and unconverted refusals were different from the initially interviewed for this variable. This indicates that the refusal conversion process may have introduced or exacerbated existing bias for this variable. For health, employment status and education level, there was indicative evidence in the previous section that refusal conversion may have led to a reduction in bias as the groups with higher refusal rates also had high conversion rates. However, the analysis in this section reveals that this was not the case for health, as there was no difference in the distributions for both converted refusals and unconverted refusals and the initially interviewed indicating a lack of bias in the variable, or for employment status or education level, as the converted refusals were similar to those initially interviewed and unconverted refusals were different from those initially interviewed.

Burton et al. (2006) also consider the impact of refusal conversion in relation to some similar variables: employment status, housing tenure, health and political preference. In relation to housing tenure, they also found that refusal conversion led to a reduction in bias. They found that refusal conversion led to a reduction in bias in relation to employment status, which we do not find. However, we use a binary indicator of whether the sample member is in work or not and they use a more detailed employment status variable which also distinguishes self-employed and retired. They find no clear pattern in relation to health which is not inconsistent with our finding of a lack of bias for this variable. We find evidence of bias reduction in relation to voting behaviour and they do not find any evidence of this in relation to political preference, although again the variables used are not directly comparable. It should also be borne in mind that BHPS is a study of all households whereas MCS is a study of families with young children.

Our second approach to assessing bias reduction is to generate R-indicators from the response propensity model described in Plewis et al. (2008) which is based on a logistic regression model using wave one variables. The R-indicator is estimated by:

$$\widehat{R}_\rho = 1 - 2\widehat{S}_\rho$$

where ρ is the probability of responding, estimated from the response propensity model, and $\widehat{S}_\rho$ is the standard deviation of these estimated probabilities.

We find that, for the issued cases at wave four, the estimate of R increases from .769 when the converted refusals from the intensive re-issuing are treated as non-responders (i.e. had there been no re-issuing) to .775 for the achieved sample (with intensive re-issuing for half the sample). This is a modest increase in representativity but it is a lower bound in that, if intensive re-issuing had been applied to all the refusing cases rather than just to the random 50%, the increase in representativity (reduction in sample bias) would have been more marked. This is because an expected twice as many cases would have been added to the sample, and these additional cases would have been expected to have the same characteristics as those that were converted (and, as we have seen, different from those that were initially interviewed). These estimates of R are model-dependent but they do reinforce the conclusion that intensive re-issuing can lead to some reduction in bias in variables of interest.

4.7 What impact did re-issuing have on participation at the subsequent wave?

In order to assess the longer-term effectiveness of refusal conversion we examined the fieldwork outcomes at wave five (in 2012) for both converted refusals and unconverted refusals in the intensive re-issue group at wave four (in 2008). For comparison, we also examined wave five outcomes for the refusals in the no re-issue group at wave four.

The results in Table 5 show that around 62% of the converted refusals took part again at wave five compared with

Table 4
Wave three characteristics by outcome at wave four

	Initially interviewed vs. Converted refusals		Initially interviewed vs. Unconverted refusals	
	Estimate	SE	Estimate	SE
Main respondent's educational qualifications (unordered)				
No qualifications	0.21	0.37	0.89	0.16
Level 1	0.26	0.45	0.24	0.20
Level 2 (ref.)	-	-	-	-
Level 3	−0.66	0.39	−0.15	0.18
Level 4	−0.42	0.32	−0.31	0.15
Level 5	−0.54	0.58	−0.12	0.25
Overseas qualifications only	−1.18	0.68	0.32	0.32
F-statistic	$F(6, 328) = 1.56$		$F(6, 328) = 9.73$	
p-value	>.1		<.001	
Whether main respondent is in work (including on leave)				
Yes (ref.)	-	-	-	-
No	0.07	0.23	0.45	0.10
t-statistic	0.31		4.40	
p-value	>.7		<.001	
Cohort member's ethnic group				
White (ref.)	-	-	-	-
Non-white	0.007	0.31	0.38	0.14
t-statistic	0.02		2.73	
p-value	>.9		<.01	
Whether main respondent voted in last general election				
Yes (ref.)	-	-	-	-
No	0.89	0.24	0.58	0.12
t-statistic	3.70		4.92	
p-value	<.001		<.001	
Main resp. general health	0.26	0.24	0.16	0.11
t-statistic	1.07		1.43	
p-value	>.2		>.1	
Family Type				
Married or cohabiting natural parents (ref.)	-	-	-	-
Lone natural mother	0.27	0.28	0.40	0.14
Other family type	0.33	0.46	0.56	0.21
F-statistic	$F(2, 332) = 0.65$		$F(2, 332) = 6.34$	
p-value	>.5		<.01	

Continues on next page

Table 4, continued from previous page

	Initially interviewed vs. Converted refusals		Initially interviewed vs. Unconverted refusals	
	Estimate	SE	Estimate	SE
Housing Tenure				
Own (ref.)	-	-	-	-
Rent/Other	0.79	0.24	0.65	0.10
t-statistic	3.28		6.26	
p-value	<.01		<.001	

Note: Separate models fitted for each dependent variable using svy mlogit, svy logit and svy ologit in Stata for the unordered, binary and ordered dependent variables.

Table 5
Wave five outcome by wave four re-issue outcome

	Converted Refusals (Wave four)		Unconverted Refusals (Wave four)			
	Intensive re-issue		Intensive re-issue		No re-issue	
	n	*%*	*n*	*%*	*n*	*%*
Productive	85	61.6	122	26.0	239	33.6
Unproductive	47	34.1	239	51.0	310	43.5
Not issued	6	4.3	108	23.0	163	22.9
Total (*N*)	138		469		712	

around 26% of unconverted refusals in the intensive re-issue group (χ^2 = 65.0, p <.001 for this comparison). This demonstrates that this was not just a short-lived increase in sample size, for the converted refusals. However, Table 5 also shows that a significant minority of unconverted refusals went on to take part again at the next wave. Burton et al. (2006) show that a similar proportion of converted refusals took part at the subsequent wave; around 60 per cent for refusals converted at waves 4–7, dropping to around 40 per cent for waves 8–9 and around 30 per cent for waves 10–13. However, when we combine the two intensive re-issue groups from Table 5 and compare them with the no re-issue group (i.e. the experimental contrast) in terms of the percentage of productive interviews at wave five, we find no difference for the intensive re-issue group (34.0%) versus the no re-issue group (33.6%) (and this similarity holds if we include 'hard' refusals from wave four, some of whom participated at wave five). Hence, there is no effect of intensive re-issuing on sample size at the subsequent wave although the composition of the two groups might be different.

5   Conclusions

We have clearly shown that devoting additional field resources to converting refusals on the fourth wave of the MCS brought some benefits to the study in terms of an increased achieved sample size and a lower refusal rate. The randomi-sation also implies that if intensive re-issuing had been applied to the study as a whole, rather than being carried out on half of the sample only, the magnitude of these benefits would be roughly doubled. Despite the small number of converted refusals, there is also some evidence that re-issuing refusals may have led to a reduction in bias in the cross-sectional survey estimates on a few key variables. There was also a potential longer-term benefit, as the majority of the converted refusals went on to take part at the subsequent wave.

We have made two important contributions to the methodological evidence on refusal conversion. Firstly, by providing evidence from a randomised intervention, we have addressed a limitation in the existing research in this area which was largely based on observational studies. Secondly, we have also made a significant contribution to improving knowledge in an area of methodological research in which there is relatively little existing evidence i.e. the immediate and longer-term effectiveness of refusal conversion strategies in longitudinal surveys.

As well as showing the important role refusal conversion can play in response maximisation and bias reduction, this paper has also brought into focus the fact that these two aims may sometimes conflict and that the approach taken to re-issuing may differ depending on which of these two aims survey managers wish to prioritise.

If the focus is solely on response maximisation, the aim of refusal re-issuing should be to maximise the unconditional conversion rate in the most cost-effective way. As demonstrated in section 4, the unconditional conversion rate is a product of the proportion of refusals that are re-issued and the conversion rate among re-issued refusals. It is clear that there is an inverse relationship between the proportion of refusals that are re-issued and the proportion that will be converted. In general, field staff tend to re-issue a relatively low proportion of refusals and to re-issue cases which they believe are most likely to be converted. This tends to be because there is a relatively high fixed cost associated with re-issuing cases, regardless of the outcome. However, this paper has clearly shown that it is possible to achieve relatively high conversion rates for groups with high refusal rates. The clear implication of this finding is that, where response maximisation is the objective, decisions about which cases to re-issue for refusal conversion should be based on empirical data on conversion rates and information about the marginal cost of re-issuing in order to find the optimal balance between maximising the unconditional conversion rate and minimising the field costs associated with re-issuing.

However, if the main objective of refusal re-issuing is to minimise response bias, the optimal strategy will be to re-issue higher proportions of refusals from groups who are known to be under-represented in the achieved sample (which will usually be those with higher refusal rates) and lower proportions of refusals from groups who are known to be over-represented in the achieved sample (which will usually be those with lower refusal rates). In practice, this is likely to mean that refusals from groups with lower than average refusal rates should not be re-issued (even though there may be a relatively high chance that they will be converted to a successful interview). This paper has shown that some groups with higher than average refusal rates also have higher than average conversion rates i.e. it can be worthwhile to re-issue these cases. It is also clear that it is the conversion of these refusals which has the strongest impact on bias reduction (and conversely that, for some variables, the re-issuing of cases from groups with low refusal rates may actually serve to increase bias). As well as reducing the proportion of refusals re-issued among groups that are over-represented in the achieved sample, it is also necessary to increase the proportion of re-issued refusals that are converted to an interview among groups with high refusal rates and low conversion rates. As noted by Groves and Peytcheva (2008), the level of bias can differ between variables of interest in the same survey so, although it is generally acknowledged that focusing on bias reduction may be methodologically preferable to focusing on response maximisation, it can be difficult to do this in practice, particularly for multi-purpose studies which collect data on many different domains such as the MCS.

Refusal reissuing is an expensive way to obtain an interview. Data from the fieldwork agency shows that the cost of achieving an interview was over three times as high for converted refusals as it was for families who did not refuse initially. However, it should also be borne in mind that this cost was incurred for a very small proportion of the achieved sample and a small number of cases overall. So, in absolute terms, the additional total cost of the extra interviews achieved via refusal conversion is small compared with the fieldwork costs for the study as a whole. More generally, targeted interventions such as refusal conversion are likely to be more cost-effective than universal ones, such as incentives given to all respondents, as resources are not then wasted on sample members who would participate without this intervention. It should also be noted that the approach taken for the experiment i.e. re-issuing all refusals, is expensive as there is a high fixed cost associated with re-issuing a case regardless of outcome and, as discussed earlier, this approach may not be the most cost-effective way either to maximise response or to reduce bias. For this reason, it is expected that a more refined approach to refusal re-issuing, which takes into account all available information, may lead to a reduction in the cost of an interview achieved through refusal re-issuing.

In addition, for a longitudinal study such as MCS, the evaluation of the cost-effectiveness of refusal conversion should also take into account whether or not converted refusals continue to participate more than unconverted refusals at later waves of the study. Longitudinal studies are concerned about response maximisation over the long-term rather than at only one point in time, and about bias reduction in estimates of change rather than in cross-sectional estimates. Decisions about resource allocation should also be taken with a long-term perspective. This implies that decisions about refusal conversion on a particular wave should, if possible, be informed by evidence about the likely impact on survey costs, response rates and bias at future waves. Our evidence indicates that converted refusals do participate again at the subsequent wave – but so do unconverted refusals. Hence, the longer-term benefits, in terms of sample size, are questionable.

References

AAPOR. (2014). Current knowledge and considerations concerning survey refusals. Retrieved from http://www.aapor.org/AAPOR%5C_Main/media/MainSiteFiles/RefusalTF%5C_FINAL090814.pdf

Basson, D. & Chronister, M. (2006). Recordings of prior refusals: do they improve later conversion attempts? Paper presented at Methodology of Longitudinal Surveys conference, Essex, England, July 2006.

Burton, J., Laurie, H., & Lynn, P. (2006). The long-term effectiveness of refusal conversion procedures on longitudinal surveys. *Journal of the Royal Statistical Society Series A*, *169*(3), 459–478.

Calderwood, L., Plewis, I., Ketende, S. C., & Taylor, R. (2010). Experimental testing of refusal conversion strategies in a large-scale longitudinal study. CLS Working Paper 2010/9. London: Centre for Longitudinal Studies. Retrieved from http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=939&sitesectiontitle=CLS+working+paper+series

Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly*, *64*(4), 413–428.

de Leeuw, E., Callegaro, M., Hox, J., Korendijk, E., & Lensvelt-Mulders, G. (2007). The influence of advance letters on response in telephone surveys: a meta-analysis. *Public Opinion Quarterly*, *71*(3), 413–443.

de Leeuw, E. & de Heer, W. (2002). Trends in household survey nonresponse: a longitudinal and international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & A. Little (Eds.), *Survey nonresponse*. New York: Wiley.

Groves, R. M. (1989). *Survey errors and survey costs*. New York: Wiley.

Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, *56*(4), 475–495.

Groves, R. M. & Couper, M. P. (1998). *Nonresponse in household interview surveys*. New York: Wiley.

Groves, R. M. & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public Opinion Quarterly*, *72*(2), 167–189.

Hox, J. & de Leeuw, E. (2002). The influence of interviewers' attitude and behavior on household survey nonresponse: an international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & A. Little (Eds.), *Survey nonresponse*. New York: Wiley.

Keeter, S., Miller, C., Kohut, A., Groves, R. M., & Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly*, *64*(2), 125–148.

Laurie, H. & Lynn, P. (2009). The use of respondent incentives on longitudinal surveys. In P. Lynn (Ed.), *Methodology of longitudinal surveys*. Chichester: Wiley.

Lynn, P., Clarke, P., Martin, J., & Sturgis, P. (2002). The effects of extended interviewer efforts on nonresponse bias. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & A. Little (Eds.), *Survey nonresponse*. New York: Wiley.

Plewis, I. (2007). *Millennium Cohort Study: technical report on sampling* (4th ed.). London: Centre for Longitudinal Studies. Retrieved from http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=878&sitesectiontitle=Technical+Reports

Plewis, I., Calderwood, L., & Mostafa, T. (2016). Can interviewer observations of the interview predict future response? *Methods, Data, Analyses*. (in press).

Plewis, I., Ketende, S., Joshi, H., & Hughes, G. (2008). The contribution of residential mobility to sample loss in a british birth cohort study: evidence from the first two waves of the Millennium Cohort Study. *Journal of Official Statistics*, *24*(3), 364–385.

Schouten, B., Cobben, F., & Bethlehem, J. (2009). Indicators of the representativeness of survey response. *Survey Methodology*, *35*, 101–113.

Shadish, W. R. & Cook, T. D. (2009). The renaissance of field experimentation in evaluating interventions. *Annual Review of Psychology*, *60*, 607–629.

Singer, E. (2002). The use of incentives to reduce nonresponse in household surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & A. Little (Eds.), *Survey nonresponse*. New York: Wiley.

Stoop, I. A. L. (2004). Surveying nonrespondents. *Field Methods*, *16*, 23–54.

Stoop, I. A. L. (2005). *The hunt for the last respondent: nonresponse in sample surveys*. The Hague: Social and Cultural Planning Office.