Survey Research Methods (2014)
Vol. 8, No. 1, pp. 19-30
ISSN 1864-3361
http://www.surveymethods.org

Variation in Incentive Effects across Neighbourhoods An Example from The Irish Longitudinal Study of Ageing

Mark J. Hanly
University of Bristol

George M. Savva
University of East Anglia

Ian Clifford
Trinity College Dublin

Brendan J. Whelan
Trinity College Dublin

Small monetary incentives increase survey cooperation rates, however evidence suggests that the appeal of incentives may vary across sample subgroups. This observation has implications for the developing practise of adaptive survey designs, which target specific subgroups with tailored recruitment protocols. Geographic neighbourhoods are a practical level at which to vary incentives, especially in multistage address samples where the primary sampling unit typically comprises a defined area such as a postal sector, to which auxiliary information can be linked. To understand the variable effect of incentives across neighbourhoods we examine data from a randomised experiment implemented in the pilot phase of the Irish Longitudinal Study of Ageing, which assigned households to a higher (€25) or lower (€10) incentive condition. Using a random effects logistic regression model, we observe a variable effect of the higher incentive across geographic neighbourhoods. The higher incentive has the largest impact in neighbourhoods where baseline cooperation is low, as predicted by Leverage-Saliency theory. Auxiliary neighbourhood-level variables are linked to the sample frame to explore this variation further, however none of these moderate the incentive effect, suggesting that richer information is needed to identify neighbourhoods where incentive budgets should be directed.

Keywords: incentives; targeted designs; multilevel modelling; face-to-face survey

1 Introduction

The positive impact of incentives on survey response rates has frequently been reported, both for mail (Yu & Cooper, 1983; Church, 1993; Singer, 2002) and interviewer mediated surveys (Cantor, O'Hare, & O'Connor, 2008; Gelman, Stevens, & Chan, 2002; Singer, Van Hoewyk, Gebler, Raghunathan, & McGonagle, 1999). Money has been shown to be more effective than non-monetary gifts or charitable donations (Warriner, Goyder, Gjertsen, Hohner, & McSpurren, 1996) and non-contingent incentives are more effective than those conditional on cooperation (Church, 1993; J. M. James & Bolstein, 1992). In general, higher monetary amounts seem to elicit higher response rates (Singer et al., 1999; T. James, 1997), although there is some evidence to suggest a diminishing rate of returns (Gelman et al., 2002; J. M. James & Bolstein, 1992). Incentives are particularly useful in negating the effects of unappealing design features (Castiglioni, Pforr, & Krieger, 2008), sensitive topics (Singer & Bossarte, 2006) or where the respondent is initially unwilling (Trussell & Lavrakas, 2004). Incentives also reduce drop-out in longitudinal surveys (Lipps, 2010; Castiglioni et al., 2008; Jäckle & Lynn, 2007; Rodgers, 2002).

In terms of improving survey quality, incentives only reduce bias if, as well as increasing response rates, they also narrow the respondent-nonrespondent gap. As Groves (2006, p.664) points out, incentives must be "disproportionately attractive to the low propensity groups". Existing evidence suggests that monetary incentives may indeed be more attractive to specific low-propensity subgroups. Berlin et al (1992) reported that including a $20 incentive in a face-to-face literacy survey increased cooperation overall, but especially for those with the lowest literacy levels. Mack et al. (1998) found similar results from an experiment in the Survey of Income and Program Protection (SIPP). In this longitudinal study, raising the incentive from zero at the first interview to $20 for interviews 2–4, reduced nonresponse across all subgroups, but with a stronger effect observed for Black households compared to non-Black households and for households in poverty compared to those not in poverty. In both examples the incentive has a stronger impact on subgroups traditionally associated with lower response rates. More recently, Lynn (2012) reported a differential incentive effect across age groups from an Understanding Society Innovation Panel experiment. At wave 2, a €10 incentive significantly increased cooperation amongst persons over the age of 60 but not for younger age groups, an interesting result given that refusal was highest in this subgroup, driven by those over the age of 70.

The potential for incentives to have differential effects across subgroups is encapsulated in the *Leverage-Saliency*

Contact information: Mark J. Hanly, University of Bristol, UK (mark.j.hanly@bristol.ac.uk)

theory of survey response (Groves, Singer, & Corning, 2000). This model posits that the factors influencing an individual's decision to participate in a survey are weighted according to the importance bestowed on each factor by the individual (its leverage) and how prominent the factor is made in the survey request (its saliency). The final decision to participate is the net impact of the combined survey factors.

The observation that cash incentives may exert a high leverage amongst some subgroups but not others is particularly relevant given the increasing shift amongst survey methodologists away from fixed survey designs in favour of targeted or responsive designs (Lynn, 2013; Couper & Wagner, 2011; Schouten, Calinescu, & Luiten, 2011; Kreuter, Couper, & Lyberg, 2010; Wagner, 2008; Groves & Heeringa, 2006). Responsive designs vary design features for different sample subgroups with the aim of minimising costs whilst maximising survey quality, either by improving response rates and/or balancing response rates across subgroups (Lynn, 2013). The incentive amount is a design feature which can be varied in this way, and the cost-saving benefits of targeting incentives at sample subgroups where they will have the strongest impact are clear. If those for whom incentives are most attractive are also those who are otherwise least likely to take part in the survey then targeted incentives have the potential to decrease bias as well as reduce costs.

One of the challenges to targeting subgroups with specific treatments is the question of how to assign sampled units to appropriate subgroups. As Lynn (2013) suggests, this can be achieved using auxiliary information such as frame data or data from previous waves. One potential source of useful auxiliary information is aggregate data, as discussed by Smith (2011). Aggregate, or area-level data refer to information about the neighbourhood surrounding the sampled household, and are often available to methodologists, either directly from sample frames or by linking to external datasets (Smith, 2011; Smith & Kim, 2009; Cantor et al., 2008). In the context of issuing differential incentive amounts, it would seem advisable to vary incentives at a high level of aggregation. This limits the possibility that respondents within the same sampled area will learn of a neighbour receiving a higher incentive amount, a practise which may be perceived as unfair (Singer et al., 1999). There are issues with basing targeted strategies on such aggregated information. For example, data relevant to the strategy in question may not be available, or inconsistently available across countries. Data linkage may be difficult and, as such information is compiled from census information, timeliness of data may also be problematic (Smith, 2011). There is also the issue of the ecological fallacy: one cannot infer the characteristics of a household from the characteristics of its environs. Despite these limitations, the general accessibility of area-level data makes them an attractive candidate to inform targeted strategies, especially in early survey stages when information from previous waves is unavailable.

Assuming area-level information is available, or can be linked, it still remains unclear which variable(s) will predict where incentives will have the maximum impact. This research focuses on identifying suitable indicators which can be used to target areas where incentives will be most effective. Drawing on data from a randomised incentive experiment incorporated into the pilot phase of The Irish Longitudinal Study of Ageing (TILDA), we use multilevel logistic regression to reveal the variable effect of incentives across neighbourhoods. Linking to census statistics aggregated at the neighbourhood level, we test the power of five separate area-level variables to explain the observed variation in the incentive effect.

Before presenting our results, we introduce our data and the design of the incentive experiment, as well as motivating the selection of the area-level variables we examine.

2 TILDA and the Pilot Incentive Experiment

The Irish Longitudinal Study on Ageing (TILDA)

The Irish Longitudinal Study on Ageing (TILDA) is a nationally representative, prospective study of persons aged fifty years and over and resident in the Republic of Ireland (RoI). A national pilot study was carried out between April and June 2009. The pilot sample followed a multi-staged, clustered design selected from the GeoDirectory, a comprehensive database of all addresses in RoI compiled by An Post (The Irish postal service) and Ordnance Survey Ireland. A multistage sample was selected using the *RANSAM* program (Whelan, 1979). The Primary Sampling Unit (PSU) consisted of a geographic cluster, or neighbourhood, 20 of which were chosen from a total of 3101, with a probability of selection proportionate to the number of addresses with persons aged over fifty in the area. Prior to selection the neighbourhoods were sorted by education and occupation profiles and location, thus the sample was stratified by socioeconomic status (SES) and geography. Subsequently, 60 addresses were selected within each neighbourhood. This resulted in a total of 1200 addresses to be issued to interviewers.

As there was no information on occupants' age contained within the sample frame, the first job of the fieldwork team was to approach each address and identify whether an eligible person (i.e. anybody over the age of fifty) was resident. An oversample was drawn to compensate for the fact that many households did not have an eligible occupant. The final stage of selection took place on the doorstep, where the 'next-birthday' rule was employed by interviewers to select the target respondent from eligible occupants. Once identified, the target respondent was invited to take part in a face-to-face CAPI interview which took place in the respondents' home. This interview lasted one hour and eighteen minutes on average. Respondents were also invited to complete a pencil and paper questionnaire and to attend a health assessment at a dedicated health centre, although completion of the CAPI component alone was sufficient to constitute cooperation.

The Incentive Experiment

It is unclear what constitutes the optimal incentive in the context of an older population in the Republic of Ireland. In order to assess the relative effectiveness of differing incentive amounts, half the pilot sample was randomly assigned to receive a €10 incentive while the other half was assigned to €25. As the aim was to identify an effective incentive level, rather than decide whether or not to issue an incentive, a zero incentive condition was not included. The randomisation of the incentives took place within neighbourhoods, prior to issuing addresses to the fieldwork team. Thus, within each neighbourhood, 30 addresses were assigned to the €10 incentive condition and 30 addresses were assigned to the €25 condition. The incentives were offered by the interviewer, contingent on participation, during recruitment efforts on the doorstep. Interviewers were aware of the assigned incentive group prior to approaching an address. With a few exceptions, there was a one-to-one relationship between interviewers and neighbourhoods so each interviewer administered both incentive conditions. In order to maximise recruitment to the pilot, interviewers were authorised to offer those who refused the €10 incentive an additional €15 for their cooperation. While the original incentives were given in cash, the additional €15 was given in the form of a promissory note that the additional amount would be posted out by cheque. Interviewers reported that respondents had no issue accepting the promissory note instead of cash. As the additional offer was not offered systematically, our dependent variable here is the response to the initial survey request with the original incentive offer, rather than to the response to subsequent calls offering the additional amount.

Fieldwork and Interviewers

A professional market research company was employed to carry out the fieldwork for the pilot study. All interviewers attended a two day briefing session which explained the sample design and interview structure and provided an opportunity for interviewers to familiarise themselves with the CAPI script. Twenty-six interviewers were briefed for the pilot study and in total 23 of these took part in the fieldwork. An advance letter, explaining the study and inviting sample members to participate, was issued to 60% of the original sample one week prior to the interviewer approaching the address. As occupants' names were unavailable, the letter was only issued to "unique" addresses which contained a street or apartment number. The high proportion of non-unique addresses is not unusual, particularly in rural Ireland. This is not an issue in the field, as addresses are linked to geo-coordinates which are included in the sampling frame. These coordinates were entered into satellite navigation systems provided to interviewers, allowing the correct address to be found in the absence of unique identifiers such as names, street or house numbers. In the case where a unique address was unavailable, interviewers delivered a copy of the advance letter on the first call to the household. It was also used as a reminder in the case of unique addresses. Interviewers also carried an information sheet and an FAQ booklet which provided more details about the survey components, confidentiality, the funders and those running the survey.

3 Experiment Results

Sample Breakdown and Cooperation Rate

Of the initial 1200 sampled addresses 156 (13.0%) were unoccupied or non-residential and 535 (44.6%) did not have a resident over the age of 50. Interviewers failed to establish contact at a further 100 addresses (8.3%) leaving 409 (34.1%) households where contact was made with an eligible occupant and the potential for an incentive offer arose. Of these, 33 households (2.8%) were inadmissible for the pilot and not offered an incentive, primarily due to the respondent being ill, away during the period of the fieldwork or not a fluent English speaker.

From the initial sample, there remained 376 households where an incentive could be offered to an eligible respondent. The average number of eligible households per neighbourhood was 18.8 with a range of 2 to 29 households. Of these, 179 (48%) had been pre-assigned to the €10 incentive level while the remaining 197 (52%) had been pre-assigned to the €25 level. The sample characteristics of those assigned to each group were compared for house type, age and sex of respondent and aggregated neighbourhood statistics. While there were minor profile differences between the two incentive groups none of these differences were statistically significant. Based on the response to the initial incentive offer, 121 of 197 (61.4%) offered €25 agreed to participate compared to 62 of 179 (34.6%) offered €10 (Odds Ratio=3.0; 95% CI: 2.0–4.6; $p < 0.001$). After refusers in the €10 group were offered the full €25 an additional 33 respondents were recruited. These figures correspond to a final cooperation rate of 52.8% to the national pilot (based on *Coop1*, AAPOR (2011)), where the cooperation rate corresponds to the number of interviewed sample members over all eligible sample members ever contacted.

Differential Incentive Effect

The results of the incentive experiment presented above suggest that the higher incentive amount was more effective in recruiting respondents to the pilot survey, as expected from existing literature. In order to model the effect of the higher incentive across the sampled clusters a series of multilevel models were fitted to the experimental data. The use of multilevel models allowed us to test whether the observed impact of the higher incentive amount was constant or whether it varied between areas. Model improvements were assessed by comparing the Deviance Information Criterion (DIC), which incorporates information about the fit and complexity of a model (Spiegelhalter, Best, Carlin, & Linde, 2002). Data preparation and descriptive analysis was carried out using Stata 12. All models were fitted using Markov Chain Monte Carlo estimation in MLwiN (Rasbash, Steele, Browne, & Goldstein, 2009) using the Stata command *runmlwin* (Leckie & Charlton, 2011).

Table 1 Multilevel Logistic Regression Coefficients, Standard Errors, Variance Parameters and Fit Statistics for Models of Survey Cooperation (1=Cooperation/0=Refusal).

	Model 0	Model 1	Model 2	Model 3
Fixed Part				
Intercept	−0.64***	−0.03	−0.63**	−0.71*
	(0.16)	(0.19)	(0.23)	(0.34)
Incentive (Ref:€10)	1.10***		1.13***	1.23***
	(0.21)		(0.22)	(0.33)
Intercept				
Variance of Intercept		0.45	0.44	1.53
		(0.28)	(0.28)	(0.93)
Intercept – Incentive Covariance				−0.89
				(0.75)
Variance of Incentive Effect				0.89
				(0.68)
Model Fit				
DIC	497.6	503.8	481.4	477.4

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, standard errors in parentheses

The results are presented in Table 1. To begin, a single-level logistic model is fitted to the response outcome controlling only for the incentive effect (Model 0). The significant, positive effect of the incentive value can be seen. The log odds of the intercept (−0.64) corresponds to an estimated response probability of 34.6% among the low incentive group. Similarly, the log odds of the intercept plus the coefficient for the higher incentive (−0.64 + 1.10) corresponds to 61.5%, the estimated probability of response among high incentive households.

Model 1 presents the baseline multilevel model, in which a random effect is included to allow for the clustering of cooperation behaviour within neighbourhoods. Here, the intercept variance of 0.45 indicates that a 1 standard deviation increase in the underlying factors represented by the random effect would correspond to a two-fold increase in the odds of cooperation. The higher incentive dummy is again included in Model 2. The large decrease in the DIC statistic between Model 2 and Model 0 supports the use of a multilevel model and indicates that there was significant variation in response rates across neighbourhoods.

While Model 2 allows response rates to vary, it constrains the effect of the higher incentive to be fixed for all neighbourhoods. In Model 3 this assumption is relaxed and the effect of the higher incentive is allowed to vary across neighbourhoods under a random coefficients multilevel model. Again, the drop in the DIC statistic between Models 2 and 3 indicates that the more complex model is a better fit to the data. In other words, not only is there a significant effect of the higher incentive, the magnitude of this effect varies significantly across neighbourhoods. The assumptions underlying Models 2 and 3 can be seen clearly in Figure 1. The graph shows the predicted probabilities of response from Model 2 (left panel) and Model 3 (right panel). The average predicted probability of response within each neighbourhood for € 10 and € 25 are represented by the dots and X's respectively. The average predicted response rates across the whole sample for € 10 and € 25 are represented by the horizontal dashed lines. For ease of viewing, the neighbourhoods are ranked according to predicted response rates in both panels.

Focusing first on Model 2 (left panel), there is considerable variation across neighbourhoods with probabilities of response ranging from 20.1% to 59.2% in the € 10 group and from 44.7% to 81.6% in the € 25 group. The average predicted cooperation rates within the € 10 and € 25 groups are 34.8% and 62.2% respectively. Notice here that the additional effect of the higher incentive is similar in each neighbourhood, and equal on the log-odds scale. The effect of the higher incentive corresponds to an increase in probability of response of approximately 26 percentage points.

The second panel depicts the predicted probabilities from Model 3, where the effect of the incentive is no longer constrained to be equal across all neighbourhoods. Here the true variation in the magnitude of the incentive effect across neighbourhoods is clear. While the sample average predicted probabilities of cooperation in the € 10 and € 25 groups remain similar at 32.9% and 62.6%, the predicted effect of the higher incentive varies significantly across the twenty neighbourhoods. In the most receptive areas, the predicted probability of response increases by 30 to 40 percentage points when the higher incentive is offered. Conversely, in the least receptive areas the increase in response due to the higher incentive is less than 10 percentage points.

The impact of the € 25 incentive is most pronounced in neighbourhoods where the predicted probability of response was generally low. This relationship is reflected in the high negative covariance between the random effect of the incentive and the random intercept observed in Model 3. This observation fits with the previously discussed Leverage-Saliency theory of survey participation (Groves et al., 2000) which suggests that incentives will be most effective in the

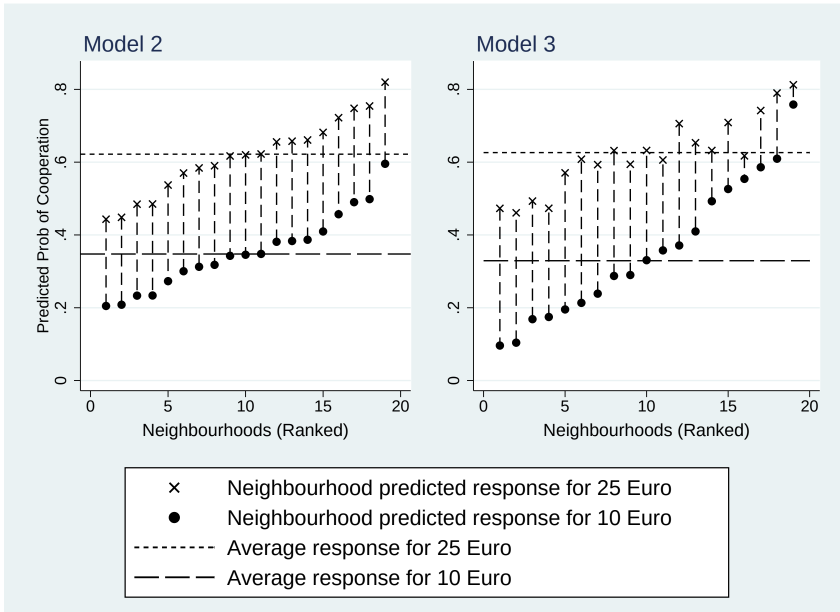

Figure 1. Predicted probability of cooperation, by neighbourhood and incentive amount

absence of other motives to participate. While Leverage-Saliency theory is presented at the level of the individual, we extend it here to a neighbourhood level and surmise that the higher incentive amount was particularly effective in areas where other impetuses to respond were lacking. The question remains: is there any quantifiable ecological variable which explains this variation and characterises areas where the incentive was effective or ineffective?

4 Further Analysis: Explaining the Variable Incentive Effect

Neighbourhood-Level Correlates of Nonresponse

Answering the question posed above would offer a significant contribution to the field of survey methodology. Identifying neighbourhoods where incentives are likely to be most effective would allow surveyors to use targeted designs to distribute their incentive budgets to maximise impact. We propose to model the incentive effect using census data aggregated at the level of the sampled neighbourhoods, guided by Cantor et al. (2008) who posit that such neighbourhood characteristics are a cost effective way of targeting low propensity areas in the absence of other information about households. From a practical point of view, such aggregated data are relatively easy to obtain in many countries (Smith, 2011; Kim, Smith, Kang, & Sokolowski, 2006). Moreover, the neighbourhood level is the most sensible one at which to vary incentive amounts to avoid offering different sums to closely neighbouring households and the potential discord this might generate (Lynn, 2013; Singer et al., 1999).

Our approach to analysing individuals' response to an incentive using area-level data could be criticised as suffering from the ecological fallacy. However, as shown from the initial analysis, differential effects were observed at the level of the neighbourhood. We are simply asking what, if any, ecological variables can explain this observed difference. Thus, we are explaining neighbourhood-level variation with neighbourhood-level variables. This research question follows naturally from existing nonresponse literature. Previous analyses of contextual effects have shown that area-level characteristics can be related to response outcomes (House & Wolf, 1978; Couper & Groves, 1996; Groves & Couper, 1998; T. P. Johnson, Cho, Campbell, & Holbrook, 2006; Kim et al., 2006) and several incentive experiments have shown that different incentive amounts are relatively effective at recruiting from populations with different characteristics (Groves, Presser, & Dipko, 2004; Martin, Abreau, & Winters, 2001; Groves et al., 2000; T. James, 1997; Mack et al., 1998).

From reviewing this literature we have selected five candidate area-level variables which may moderate the effect of an incentive in a neighbourhood: population density; presence of children; volunteering behaviour; deprivation and age profile. All of these variables are derived from the Irish Census 2006, and we believe are a good example of the sorts of area-level data available in other countries. We link these variables to the following contextual factors hypothesised to relate to response propensity: social disorganisation; social cohesion; civic engagement; socioeconomic status and topic interest. Below we review the selection of these constructs. Where applicable, we provide evidence from existing research of an interaction effect between monetary incentives and these characteristics. We also define the aggregated statistics used as proxies for these constructs. We then individually test the ability of each variable to explain the variation in the effectiveness of the higher incentive.

Urbanicity and Social Disorganisation

One of the most commonly observed neighbourhood-level correlates of participation is urbanicity: response rates are lower in large cities compared to small towns and rural areas (Groves & Couper, 1998; Goyder, Lock, & McNair, 1992). Support for this assertion is found internationally across differing survey modes and varying populations of interest. House and Wolf (1978) report a strong monotonic relationship between refusal rates and eight ordered classifications of community size for five nationally representative presidential election surveys from 1956–1972. Similarly, using data from six national household surveys, Couper and Groves (1996) report a statistically significant association between cooperation and a nominal measure of urbanicity.

While the relationship between urbanicity and response is well established, the reason for this relationship is unclear: as Groves and Couper (1998, p.176) remark, the observation that response is usually higher in rural areas than in cities, "does little to explain *why* people in different size communities differ in their likelihood of cooperation". Attempts to explain the association between urbanicity and low participation rates are guided by classical theories of social disorganisation, whereby correlates of urban areas such as high population density, heterogeneity and increasing crime rates are thought to weaken neighbourhood and social ties (Fischer, 1982). Lack of social organisation is in turn thought to reduce willingness of individuals to become involved in community activities such as surveys (Groves & Couper, 1998). Typical indicators of social disorganisation include ethnic heterogeneity and presence of multiunit accommodation. Residential stability and presence of children have been used as indicators of the positive pole of this construct, referred to as social cohesion by Couper and Groves (1996).

Existing evidence suggests that crime rates (Couper & Groves, 1996; House & Wolf, 1978) and population density (T. P. Johnson et al., 2006; Kim et al., 2006; Couper & Groves, 1996; House & Wolf, 1978) do indeed explain variation in response rates across communities. However evidence to link other indicators of social disorganisation to variation in cooperation rates is weaker, perhaps because unlike population density and crime statistics this is a construct which is less easy to define. House and Wolf (1978) found that ethnic heterogeneity failed to explain any variation in refusal rates. Drawing on the work of House and Wolf (1978), Goyder et al. (1992) used factor analysis to combine indicators such as migration, crime rate, population density, single parent families and apartment dwellers into a factor labelled social disorganisation. However this factor failed to explain variation in cooperation rates across neighbourhoods in three Canadian cities.

Couper and Groves (1996) included the effect of five environmental indicators of social cohesion: proportion of persons in group quarters, proportion of homes owner occupied, proportion of persons of minority race, proportion of single detached units and proportion of persons under 20 years old. Of these variables only the proportion of persons under 20 years old was significantly associated with cooperation. The significant association between presence of children in an area and higher cooperation rates was replicated by Kim et al. (2006) for the General Social Survey (GSS) in the United States, however, inconsistent effects were found for other indicators of social disorganisation, including residential stability and ethnic heterogeneity.

Based on the literature, crime rates, population density and to a lesser extent presence of children appear to be the most important components of what has been labelled social disorganisation. We use population density, defined as the number of persons per square kilometre, as an indicator of social disorganisation. Presence of children, defined as the proportion of the cluster aged 19 years and younger was included as an indicator of social cohesion. Unfortunately crime rates were not available at the level of our PSU, but from the literature we would expect the effects of crime rate to operate in much the same way as population density.

Civic Engagement

Another potential reason for responding positively to a survey request is a sense of altruism or civic duty (Dillman, 2007), where taking part in a survey is felt to contribute towards the greater good of society. Declining rates of survey participation have been attributed to (amongst other things) lower levels of civic engagement (Groves & Couper, 1998; T. Johnson & Owens, 2003). In their discussion of Leverage-Saliency theory, Groves et al. (2000) report a significant association between community involvement and the propensity to respond to a follow-up mail survey in Detroit. They also report a significant negative interaction between the presence of a $5 incentive and community involvement. That is, the positive effect of the incentive on response propensity was diminished amongst those who were active within the community, who tended to respond at higher rates regardless. This example illustrates that an incentive has a stronger impact in the absence of other motivations to participate.

Here we include a measurement of volunteering behaviour as an indicator of civic engagement. Volunteering behaviour is defined as the proportion of the neighbourhood regularly engaged in one or more community activities. Activities included being a member of a charitable organisation, a religious group, a sporting club a political organisation or any other voluntary activity.

Socioeconomic Status

Goyder et al. (1992, p.39) refer to socioeconomic status as, "the key socio-demographic bias" resulting from survey nonresponse. Previous evidence suggests a middle class bias: cooperation is often lower amongst the disadvantaged. This pattern has been observed both at a cluster (Kim et al., 2006; Goor, Jansma, & Veenstra, 2005) and household level (Goor & Rispens, 2004; Goyder, Warriner, & Miller, 2002). Some analysis also suggests that cooperation may also be low in very advantaged areas. For example, Johnson et al. (2006) report that households in areas of concentrated affluence as well as those in concentrated poverty were significantly less likely to participate in a telephone survey on substance use in Illinois.

Previous research indicates that incentives may operate differently across socioeconomic groups, although again the results are mixed. In a discussion of surveying low-income populations, Singer and Kulka (2002) suggest that more money may be required to recruit and maintain participation amongst high-income groups. Elsewhere results suggest that higher amounts are more effective amongst the disadvantaged, for example James (1997) and Mack et al. (1998) who report a positive impact of $20 but not of $10 in high poverty areas.

Groves and Couper (1998) lay out competing frameworks for the association between survey participation and SES. The first argues that individuals of low SES may have higher levels of indebtedness to the government than those of high SES, and consequently may be more inclined to participate (in government sponsored surveys at least). The second model suggests that those of lower SES may have feelings of inequity towards the survey interviewer, who is a representative of a more advantaged group, while at the same time those of higher SES may experience similar feelings of unfairness if they are repeatedly approached for 'contributions of time and money' (Groves & Couper, 1998, p.127). Perhaps the safest conclusion is that the relationship between SES and survey response will vary depending on survey design features such as interview mode or topic (Groves, 2006).

We use a measure of neighbourhood deprivation as a proxy for socioeconomic status. Deprivation was measured using the national deprivation index for health and health services research (Kelly & Teljeur, 2007), a neighbourhood-level measure of relative material deprivation across the Republic of Ireland, developed by the Small Area Health Research Unit (SAHRU). The index is based on a weighted combination of five characteristics believed to represent material disadvantage: unemployment; low social class; car ownership; presence of rented accommodation and overcrowding.

Topic Interest

Groves et al. (2004) hypothesised that individuals presented with a survey request on a topic of interest to them would be more likely to respond. Furthermore, in line with Leverage-Saliency theory, they posited that offering monetary incentives would attenuate the relationship between topic interest and cooperation. To test these theories, the authors devised an experiment which cross-classified survey topic with specific sample frames. The topics of education, childcare, politics and Medicare were selected to reflect the likely interests of the independent sample frames of teachers, new parents, political contributors and the over 65s respectively. Within each topic-frame combination there was a random allocation of a pre-issued $5 incentive. The analysis showed that when the survey topic matched the likely interests of the sample frame, cooperation was on average 38 percent higher than when it did not. The effect of the incentive on the interaction between topic interest and cooperation was present, and in the hypothesised direction, although it did not reach statistical significance.

Roose et al. (2007) reported similar effects with respect to topic interest and follow-up efforts in a mail survey in Belgium. Here, the authors suggest that the positive effects of increased follow-up efforts are diminished as interest in the survey increases.

As we are dealing with a survey of ageing we hypothesise that interest will be greatest in communities with older age profiles. Thus, the proportion of the cluster above the national retirement age of sixty five is used as a marker for topic interest.

Data Summary

The five contextual attributes which have been hypothesised to moderate the effect of the higher incentives are summarised in Table 2 below. All of the proxy variables are derived from the Small Area Population Statistics (SAPS) based on the Irish Census of the Population 2006. With the exception of the deprivation index, all data were downloaded from the archive of the Central Statistics Office (www.cso.ie). The deprivation index was provided directly from the Small Area Health Research Unit (www.sahru.tcd.ie). All variables were available at the level of our sampling PSU which is the electoral district, the smallest geographical unit for which data is released.

For the purpose of analysis, each scale variable was transformed into a ternary categorical variable. Multilevel logistic regression was used to model the impact of the incentives on household nonresponse across the twenty neighbourhoods. Each variable was inspected individually by adding it to the random effects logistic model of cooperation (Model 3 above). Adding each as a main and interactive effect with the higher incentive allowed both the direct effect on cooperation and the impact of the incentive in different contexts to be assessed. In each case the lowest tertile was the reference category. Main effects may explain variation in the intercept, while the interactions can explain variation in the random slope (which can be thought of as interactions between the incentive and unobserved area-level characteristics, commonly called cross-level interactions). Positive main effects combined with negative interactive effects (or negative main effects combined with positive interactions, depending on the variable) would indicate that the effect of the incentive was lower in the type of area where cooperation was higher, in line with Leverage-Saliency theory.

Results

Results from the five multilevel logistic regressions are presented in Table 3. Model 3, the random effects model of cooperation without any area-level predictors, is included as a baseline for comparison.

The first point to emerge from Table 3 is that, with one exception, none of the added area-level predictors are statistically significant at conventional levels, either as main effects or as interactive effects with the higher incentive amount.

Table 2 Summary of Area-Level Variables Included in the Analysis

Construct	Variable	Definition	Mean	Min	Max
Social Disorganisation	Population Density	Number of persons per square Kilometre	1842	5	11111
Social Cohesion	Presence of Children	Proportion of cluster aged under twenty	27%	10%	34%
Civic Engagement	Volunteering	Proportion of cluster regularly participating in one or more voluntary activities, including being a member of a charitable organisation, a religious group, a sporting club, a political organisation or any other voluntary activity.	14%	7%	21%
Socioeconomic Status	Deprivation Index	The national deprivation index for health and health services research based on a weighted combination of five characteristics believed to represent material disadvantage: unemployment, low social class, car ownership, presence of rented accommodation and overcrowding.	1.1	-2	6.5
Topic Interest	Elderly Population	Proportion of the cluster aged sixty-five and over	11%	2%	18%

Moreover, none of the area-level predictors reduce the variance of the random effects or the DIC statistic to a large degree. Examining the models individually reveals some interesting patterns. The neighbourhood-level indicator of population density (Model 4) behaves as would be expected from previous results, with predicted cooperation lower in moderate and high density neighbourhoods relative to low density areas. Moreover, the interaction effects are both positive indicating that the higher incentive had the largest impact in more densely populated areas, as would be expected from Leverage-Saliency theory. However, while in the hypothesised direction, none of these parameters reach significance.

A small decrease in the neighbourhood variance, accompanied by a small drop in the DIC statistic is observed when volunteering behaviour is controlled for (Model 6). This decrease is driven by a significantly higher probability of response in clusters with moderate compared to low levels of volunteering. The associated interactive effect between moderate volunteering clusters and the higher incentive amount indicates that the incentive was less effective in these areas. While this agrees with Leverage-Saliency theory and the results of Groves et al. (2000), it is difficult to give too much credence to this observation given that the interaction parameter does not reach traditional levels of statistical significance and the fact that a comparable effect is not observed for clusters with the highest proportion of volunteers.

The effects of the other neighbourhood-level indicators; presence of children (Model 5), deprivation (Model 7) and large elderly population (Model 8) are inconsistent and difficult to interpret. The main and interactive parameters are insignificant and none of these variables explain any of the variation in incentive effects across communities.

5 Discussion

Our secondary analysis of an incentive experiment incorporated into the TILDA pilot study showed that cooperation was significantly higher amongst households offered €25 compared to those offered €10. Subsequent analyses showed that this positive effect of the higher incentive varied across areas and, in line with Leverage-Saliency, the higher incentive amount had the largest impact where baseline cooperation rates were low. Driven by survey nonresponse theory, five area-level indicators were selected and individually included into a multilevel logistic regression model of cooperation to test whether they could explain the variation in the incentive effect across communities.

The results of the analyses reveal that none of the examined variables could adequately explain the observed variation across neighbourhoods. The most likely candidates appear to be population density and volunteer activity which both produce effects in the hypothesised direction. However, neither could be considered strong indicators: the parameters fail to reach statistical significance in the case of population

Table 3 Multilevel Logistic Regression Coefficients, Standard Errors, Variance Parameters and Fit Statistics for Models of Survey Cooperation (1=Cooperation/0=Refusal) Controlling for Neighbourhood Characteristics.

	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
		Variable Controlled For (Reference Category Low):				
		Density	% Under 20	% Volunteers	Deprivation	% Over 65
Fixed Part						
Intercept	-0.71^{*}	-0.33	-1.01^{+}	-1.39^{*}	-0.98^{+}	-0.54
	(0.34)	(0.68)	(0.56)	(0.54)	(0.57)	(0.58)
Incentive (Ref: €10)	1.23^{***}	0.97	1.54^{**}	1.64^{**}	1.66^{**}	1.42^{*}
	(0.33)	(0.60)	(0.53)	(0.55)	(0.55)	(0.56)
Area Characteristics (Ref: Low)						
Moderate		-0.48	0.87	1.99^{*}	0.88	-0.97
		(0.95)	(0.83)	(0.79)	(0.84)	(0.88)
High		-0.60	-0.04	0.21	0.02	0.28
		(0.91)	(0.87)	(0.92)	(0.90)	(0.82)
Area-Incentive Interactions						
Incentive X Moderate		0.30	-1.00	-1.49^{+}	-1.06	0.33
		(0.86)	(0.79)	(0.78)	(0.79)	(0.84)
Incentive X High		0.46	0.11	0.17	-0.29	-0.79
		(0.85)	(0.83)	(0.89)	(0.83)	(0.78)
Fixed Part						
Variance of Intercept	1.53	1.79	1.55	1.13	1.72	1.50
	(0.93)	(1.21)	(1.06)	(1.00)	(1.12)	(1.02)
Intercept - Incentive Effect Covariance	-0.89	-1.09	-0.86	-0.56	-1.02	-0.87
	(0.75)	(0.98)	(0.87)	(0.83)	(0.90)	(0.84)
Variance of Incentive Effect	0.89	1.08	0.87	0.65	0.99	0.90
	(0.68)	(0.89)	(0.80)	(0.75)	(0.81)	(0.77)
Model Fit						
DIC	477.4	481.2	480.1	476.1	479.7	479.4

Notes: $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, standard errors in parentheses density and are somewhat inconsistent in the case of volunteer activity.

Identifying aggregate variables which predict areas where incentives will have a strong impact would be a useful advance for survey methodologists. Such an understanding would allow incentive budgets to be effectively distributed by targeting low-propensity areas. It is intriguing that none of the variables we examined adequately explained the variation in the incentive effect. One possible cause is the sparsity of our data: only two incentive levels were employed across twenty clusters, some of which had only a small number of sampled households. Another possibility is that the underlying ecological factors which explain variation across neighbourhoods are not easily captured with the types of aggregated census data we examine here. Campanelli and O'Mhuircheartaigh (1999) come to a similar conclusion regarding the effects of both geographic areas and interviewers in recruitment to the British Household Panel Study (BHPS).

The role of interviewers is in itself worth considering. The interviewer represents a critical link between the survey organisation and the sampled households (Campanelli & O'Mhuircheartaigh, 1999) and there is evidence that interviewer characteristics, albeit largely intangible ones, may influence response rates (Campanelli, Sturgis, & Purdon, 1997). It is possible that what we interpret here as the effect of area is in fact a direct effect of the interviewer. As with many other surveys, interviewers and areas are confounded in the TILDA pilot study meaning that it is impossible to empirically disaggregate the independent interviewer and area effects.

One notable exception is the aforementioned study by Campanelli and O'Mhuircheartaigh (1999). Here, an interpenetrated sample experiment was designed, which randomised interviewers within geographic pools of PSUs. The results of this analysis indicated that there were modest independent effects of both interviewers and areas and that, in the case of cooperation, the effect of interviewers was greater. However, multivariate analysis indicated that the variation in response across interviewers could not be explained in terms of characteristics such as age, sex or experience.

Another possibility is that the variation we observe is not a direct effect of the interviewer but rather an effect of the higher incentive operating through the interviewer. Existing evidence suggests that more confident interviewers have lower refusals rates (Durrant, Groves, Staetsky, & Steele, 2010), and we might hypothesise that higher incentive amounts increase interviewer confidence. Previous research, however, points to the contrary. Singer et al. (2000) designed an experiment to test whether incentive payments created expectation effects on interviewers in a random digit dial telephone survey. A \$5 prepaid incentive was issued to a subset of sampled households, but interviewers were only made aware of this payment in half of the cases. The results revealed that the positive effects of the incentive on cooperation were the same regardless of whether or not interviewers were aware of the payment, leading the authors to conclude that "the incentive exerts its effect directly on respondents, rather than being mediated through interviewer expectations" (Singer et al., 2000, p.177).

Similar results were found for a face-to-face study. Drawing on qualitative data gathered from survey fieldworkers in the UK, Lynn (2001) reported that interviewers' attitudes towards the presence of an incentive were either neutral or negative. Despite this perception, incentives did increase cooperation rates significantly, suggesting that the positive effect of the incentive was not moderated by interviewers' attitudes.

Thus, we cannot entirely dismiss an interviewer effect, although the literature suggests this will act directly through the interviewer rather than via the incentive. Doubts over a potential interviewer effect on this experiment are somewhat assuaged by the use of a small, professional and experienced fieldwork team, all of whom had been employed on previous household studies and received specific training from core TILDA team members.

In conclusion, we observed a significant variation in the effect of incentives across neighbourhoods, but we were unable to explain this variation using neighbourhood-level indicators which have been previously related to nonresponse. This result does not support the use of aggregated census data as a means to assign sampled units to different incentive treatments. However this was a small study with only two incentive amounts randomised across a small number of areas. To understand how different incentives can be employed to minimise survey costs, future research should examine multiple incentive amounts randomised within many clusters. Ideally interviewers and clusters will be cross-classified to distinguish separate interviewer and neighbourhood effects. Assigning sampled units to incentive treatments using paradata collected during early fieldwork stages may also be a useful area for future investigation. For example, response propensity models based on initial interviewer observations may be used to identify low-propensity households where an incentive might most effectively be issued in subsequent phases (Kreuter et al., 2010; Groves & Heeringa, 2006). Low-propensity groups may also be identified using non-model-based approaches, such as the typology of households identified by Pollien and Joye (2011) using sequence analysis of call record data.

Acknowledgments

We would like to thank two annonymous reviewers for helpful comments on earlier drafts of this paper. The first author would particularly like to thank Paul Clarke and Fiona Steele, for their generous advice and support.

References

AAPOR. (2011). *Standard definitions: Final dispositions of case codes and outcome rates for surveys.*

Campanelli, P., & O'Mhuircheartaigh, C. (1999). Interviewers, interviewer continuity and panel survey nonresponse. *Quality and Quantity*, *33*, 59-76.

Campanelli, P., Sturgis, P., & Purdon, S. (1997). *Can you hear me knocking: An investigation into the impact of interviewers on survey response rates*. London: Social and Community Planning Research.

Cantor, D., O'Hare, B., & O'Connor, K. (2008). The use of monetary incentives to reduce non-response in random digit dial telephone surveys. In J. M. Lepkowski et al. (Eds.), *Advances in telephone survey methodology.* New York: Wiley.

Castiglioni, L., Pforr, K., & Krieger, U. (2008). The effect of incentives on response rates and panel attrition: Results of a controlled experiment. *Survey Research Methods*, *2*(3), 151-158.

Church, A. H. (1993). Estimating the effect of incentives on mail survey response rates: A meta-analysis. *The Public Opinion Quarterly*, *57*(1), 62-79.

Couper, M. P., & Groves, R. M. (1996). Social environment impacts on survey cooperation. *Quality and Quantity*, *30*, 173 - 188.

Couper, M. P., & Wagner, J. (2011). *Using paradata and responsive design to manage survey nonresponse.* (Invited paper presented to the World Statistics Congress of the International Statistical Institute, Dublin, Ireland)

Dillman, D. A. (2007). *Mail and internet surveys: The tailored design method*. New Jersey: John Wile and Sons.

Durrant, G., Groves, R., Staetsky, L., & Steele, F. (2010). Effects of interviewer attitudes and behaviours on refusal in household surveys. *Public Opinion Quarterly*, *74*, 1 - 36.

Fischer, C. S. (1982). *To dwell among friends: Personal networks in town and city*. Chicago: The University of Chicago Press.

Gelman, A., Stevens, M., & Chan, V. (2002). Regression modeling and meta-analysis for decision making: A cost-benefit analysis of incentives in telephone surveys. *Journal of Business and Economic Statistics*, *21*(2), 213 - 215.

Goor, H. V., Jansma, F., & Veenstra, R. (2005). Differences in undercoverage and nonresponse between city neighbourhoods in a telephone survey. *Psychological Reports*, *96*, 867-878.

Goor, H. V., & Rispens, S. (2004). A middle class image of society: A study of undercoverage and nonresponse bias in a telephone survey. *Qualtiy and Quantity*, *38*, 35-49.

Goyder, J., Lock, J., & McNair, T. (1992). Urbanization effects on survey nonresponse: A test within and across cities. *Quality and Quantity*, *26*, 39-48.

Goyder, J., Warriner, K., & Miller, S. (2002). Evaluating socioeconomic status (SES) bias in survey nonresponse. *Journal of Official Statistics*, *18*(1), 1-11.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, *70*(5), 646–675.

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

Groves, R. M., & Heeringa, S. G. (2006). Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *169*(3), 439–457.

Groves, R. M., Presser, S., & Dipko, S. (2004). The role of topic interest in survey participation decisions. *Public Opinion Quarterly*, *68*(1), 2-31.

Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: description and an illustration. *The Public Opinion Quarterly*, *64*(3), 299–308.

House, J. S., & Wolf, S. (1978). Effects of urban residence on interpersonal trust and helping behaviour. *Journal of Personality and Social Psychology*, *36*(9), 1029-1043.

Jäckle, A., & Lynn, P. (2007). Respondent incentives in a multimode panel survey: Cumulative effects on nonresponse and bias (no. 32). *Colchester: University of Essex*.

James, J. M., & Bolstein, R. (1992). Large monetary incentives and their effect on mail survey response rates. *Public Opinion Quarterly*, *56*(4), 442–453.

James, T. (1997). Results of the wave 1 incentive experiment in the 1996 *s*urvey of *i*ncome and *p*rogram *p*articipation. In *Proceedings of the survey research methods section* (p. 834 - 839). Alexandria, VA: American Statistical Association.

Johnson, T., & Owens, L. (2003). Survey response rate reporting in the professional literature. In *Proceedings of the survey research methods section.* Alexandria, VA: American Statistical Association.

Johnson, T. P., Cho, Y. I., Campbell, R. T., & Holbrook, A. L. (2006). Using community-level correlates to evaluate nonresponse effects in a telephone survey. *Public Opinion Quarterly*, *70*(5), 704 - 719.

Kelly, A., & Teljeur, C. (2007). *The national deprivation index for health & health services research.* (SAHRU Technical Report, Trinity College Dublin)

Kim, J., Smith, T. W., Kang, J., & Sokolowski, J. (2006). *Community context and survey response*. Chicago: National Opinion Research Center.

Kreuter, F., Couper, M., & Lyberg, L. (2010). The use of paradata to monitor and manage survey data collection. In *Proceedings of the survey research methods section* (pp. 282–296). Alexandria, VA: American Statistical Association.

Leckie, G., & Charlton, C. (2011). runmlwin: Stata module for fitting multilevel models in the MLwiN software package [Computer software manual]. Bristol, UK.

Lipps, O. (2010). Effects of different incentives on attrition and fieldwork effort in telephone household panel surveys. *Survey Research Methods*, *4*(2), 81–90.

Lynn, P. (2001). The impact of incentives on response rates to personal interview surveys: Role and perceptions of interviewers. *International Journal of Public Opinion Research*, *13*(3), 326 - 336.

Lynn, P. (2012). *The propensity of older respondents to participate in a general purpose survey* (Working Paper No. 2012 03). Colchester: ISER, University of Essex.

Lynn, P. (2013). *Targeted response inducement strategies on longitudinal surveys* (Working Paper No. 2013 02). Colchester: ISER, University of Essex.

Mack, S., Huggins, V., Keathley, D., & Sundukchi, M. (1998). Do monetary incentives improve response rates in the *s*urvey of *i*ncome and *p*rogram *p*articipation. In *Proceedings of the section on survey research methods.* Mack98: American Statistical Association.

Martin, E., Abreau, D., & Winters, F. (2001). Money and motive: Effects of incentive on panel attrition in the Survey of Income and Program Participation. *Journal of Official Statistics*, *17*(2), 267-284.

Pollien, A., & Joye, D. (2011). A la poursuite du répondant? Essai de typologie des séquences de contact dans les enquêtes. In *Parcours de vie et insertions sociales* (p. 189-212). Zürich: Seismo.

Rasbash, J., Steele, F., Browne, W. J., & Goldstein, H. (2009). A user's guide to MLwiN, v2.10 [Computer software manual]. Bristol, UK.

Rodgers, W. L. (2002). Size of incentive effects in a longitudinal study. In *Proceedings of the survey research methods section.* Alexandria, VA: American Statistical Association.

Roose, H., Lievens, J., & Waege, H. (2007). The joint effect of topic interest and follow-up procedeures on the response in a mail questionnaire: An empirical test of the leverage-saliencey theory in audience research. *Socialogical Methods and Research*, *35*(3), 410 - 428.

Schouten, B., Calinescu, M., & Luiten, A. (2011). *Optimizing quality of response through adaptive survey designs* (Discussion Paper). The Hague/Heerlen: Statistics Netherlands.

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

Singer, E., & Bossarte, R. M. (2006). Incentives for survey participation: When are they coercive? *American Journal of Preventive Medicine*, *31*(5), 411–418.

Singer, E., Hoewyk, J. V., & Maher, M. P. (2000). Experiments with incentives in telephone surveys. *Public Opinion Quarterly*, *64*(2), 171 - 188.

Singer, E., & Kulka, R. A. (2002). Paying respondents for survey participation. In M. V. Ploeg, R. A. Moffitt, & C. F. Citro (Eds.), *Studies of welfare populations: Data collections and research issues.* Washington, D.C.: National academy Press.

Singer, E., Van Hoewyk, J., Gebler, N., Raghunathan, T., & McGonagle, K. (1999). The effect of incentives on response rates in interviewer-mediated surveys. *Journal of Official Statistics*, *15*(2), 217–230.

Smith, T. W. (2011). The report of the international workshop on using multi-level data from sample frames, auxiliary databases, paradata and related sources to detect and adjust for nonresponse bias in surveys. *International Journal of Public Opinion Research*, *23*(3), 389 - 402.

Smith, T. W., & Kim, J. (2009). *An assessment of the multi-level integrated database approach* (GSS Methodological Report No. No. 116). Chicago: NORC.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Linde, A. van der. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B*, *64*(4), 583 - 639.

Trussell, N., & Lavrakas, P. J. (2004). The influence of incremental increases in token cash incentives on mail survey response: Is there an optimal amount? *Public Opinion Quarterly*, *68*(3), 349-367.

Wagner, J. R. (2008). *Adaptive survey design to reduce nonresponse bias*. Unpublished doctoral dissertation, University of Michigan.

Warriner, K., Goyder, J., Gjertsen, H., Hohner, P., & McSpurren, K. (1996). Charities, no; lotteries, no; cash, yes: Main effects and interactions in a Canadian incentives experiment. *Public Opinion Quarterly*, *60*(4), 542-562.

Whelan, B. J. (1979). RANSAM: A national random sampling design for Ireland. *The Economic and Social Review*, *10*(2).

Yu, J., & Cooper, H. (1983). A quantitative review of research design effects on response rates to questionnaires. *Journal of Marketing Research*, *20*(1), 36-44.