YOUTH IN FOCUS PROJECT DISCUSSION PAPER SERIES
No. 3, April 2008

Does the effect of incentive payments on survey response rates differ by income support history?

Juan Barón
Robert Breunig
Deborah Cobb-Clark
Tue Gørgens
Anastasia Sartbayeva

youthinfocus.anu.edu.au

Jointly funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs and the Australian Research Council and carried out by The Australian National University

Youth in Focus Project Discussion Paper Series
ISSN 1835-4025

Youth in Focus
Research School of Social Sciences
ANU College of Arts and Social Sciences
Coombs Building 9
The Australian National University
Canberra ACT 0200

Does the effect of incentive payments on survey response rates differ by income support history?

Juan D. Barón, Robert V. Breunig*
Deborah Cobb-Clark, Tue Gørgens and Anastasia Sartbayeva

Research School of Social Sciences
Australian National University

April 24, 2008

Abstract

This paper asks which sub-groups of the population are affected by the payment of a small cash incentive to respond to a telephone survey. We find that an incentive improves response rates primarily amongst those individuals with the longest history of income support receipt. Importantly, these individuals are least likely to respond to the survey in the absence of an incentive. The incentive thus improves both average response rates and acts to equalize response rates across different socio-economic groups, potentially reducing non-response bias. Interestingly, the main channel through which the incentive appears to increase response rates is in improving the probability of making contact with individuals in the group with heavy exposure to the income support system.

KEYWORDS: Survey response; incentive payments; income support

Acknowledgements: We appreciate financial support from the Australian Research Council (Linkage Grant Number LP0347164) and the Australian Commonwealth Department of Families, Housing, Community Services and Indigenous Affairs. Bruce Packard at Roy Morgan Research, Australia has been helpful throughout the data gathering process. All errors and opinions are those of the authors.

*2036 H.C. Coombs Building (#9), Research School of Social Sciences, Australian National University, Canberra ACT 0200, Australia. Robert.Breunig@anu.edu.au.

1 Introduction

Incentive payments are often used in conjunction with surveys to increase response rates and/or to improve data quality. In this paper, we examine whether the effect of such payments is related to the socio-economic status of respondents. Specifically, we examine whether a past history of income support receipt is correlated with refusal rates and response rates in a telephone survey.

There is a large literature using randomized experiments to assess the impact of incentives on response rates. Most of it is based on mail-out surveys, though a number of studies have looked at incentives in telephone and face-to-face surveys. Both monetary and non-monetary incentives have been assessed. Church (1993) and Singer et al. (1999) discuss the literature and conduct meta-analyses of a large number of papers. They conclude, generally, that incentives raise response rates, that prepaid incentives are better than incentives which are paid only upon survey completion, and that monetary incentives are more effective at increasing response rates and data quality than gifts or lotteries.

In this paper, we approach the question from a slightly different angle. Incentives may increase response rates, but do they do so in a uniform way across all socio-economic groups? Using detailed administrative data about the income support receipt of individuals (and their families) from 1993 to 2006, we examine whether the intensity and recentness of income support receipt are related to responsiveness to incentives. Very few studies have gone beyond examining the response in average survey response rates to incentives to actually address the question of who it is who responds to incentives. Shettle and Mooney (1999) point out that if incentives disproportionately motivate people already predisposed to respond, then non-response bias could increase rather than decrease with the use of incentives. Alternatively, if incentives disproportionately lead those generally disinclined to respond to in fact respond, non-response bias would fall.

We will examine the relationship between socio-economic status and the effect of incentives on four specific questions. The first is whether incentives increase the probability of making contact with a target population. The second is whether the incentive makes it more likely that those who are contacted will agree to participate in the survey. Thirdly, we are interested in whether the payment of an incentive makes individuals more likely to consent to data linking.[1] Lastly, we examine whether the

[1] Specifically, this paper forms part of a larger research project in which survey and administrative data are matched to better understand the inter-generational transmission of economic disadvantage. As part of that project, we ask respondents for their permission to match their survey responses to detailed, government administrative data from the income support system.

incentive has an effect on the likelihood that participants will return a self-completion questionnaire in follow-up to the telephone interview.

To foreshadow our detailed results, we find differences in our ability to contact people and in refusal rates across individuals with different relationships to the income support system. Those with long histories of income support receipt are more difficult to contact than those with no history of income support receipt. Moreover, those in families with distant and only moderate histories of income support are more likely to refuse to participate in the survey once contacted relative to those with no income support history and those with large exposure to the income support system. Incentives work to counter-act both of these effects. Even though the incentive payment is quite small, $15 AUD, it has the effect of making the probability of contacting targeted individuals equal across all categories of past income support receipt. Likewise for response rates, where the incentive produces the largest increases in response rates precisely amongst those groups which are least likely to respond in the absence of an incentive.

The concern of Shettle and Mooney (1999), therefore, does not manifest itself in our results. To the contrary, inasmuch as non-response bias arises from differences in observable income support histories, our results suggest that the payment of an incentive *reduces* non-response bias in addition to increasing overall response rates. This is quite encouraging.

In what follows, we provide a brief background to the research project of which this paper forms a part and describe the administrative and survey data in detail. We then describe our four questions and the results of each in detail. We discuss our results in the context of the literature and provide some concluding comments in section 4.

2 The Youth in Focus Project

The data come from the pilot of the Youth in Focus (YIF) Project.[2] The YIF Project relies upon an administrative data set extracted from the Australian government social security system. The administrative data were constructed by choosing all individuals appearing in the data with a birth date between 1 October 1987 and 31 March 1988, forming a birth cohort of young people. Individuals may appear in the administrative data because they received an income support payment themselves or because a family member or other relative received a payment the amount of which was determined by the individual's relationship to the payee or to the presence of the individual in the payee's household. Using this information, we constructed administrative 'families'

[2]More information on the project may be found at `http://youthinfocus.anu.edu.au/home.htm`.

of young people by linking to all adults ('parents') who had ever claimed or received a payment on behalf of the young person, to partners and spouses of the 'parents' identified in the administrative records, and to other young people ('siblings') for whom the 'parent' also claimed or received a payment.

The Australian income support system is almost universal, with some payments such as Child Care Benefit having no income test, and other payments, such as Family Tax Benefit, being denied only to families in the top 20 per cent of the income distribution. (See Centrelink (2007) for more information on the Australian income support system.) As the administrative data are of high quality going back to at least 1993 (when the young adults who were aged 18 on 31 March 2006 were five or six years old) we have 12-year period during which a young adult might appear in the data. Comparing the number of young adults in the administrative data to census information, we believe that we have over 98 per cent of all Australians born between January 1986 and March 1986 in our administrative data. (See Breunig et al. (2007) for more information on the data.)

Using this administrative information on young people and their families as our frame, we stratified the administrative data into six strata based upon the intensity and recentness of income support. We adopt the Australian government definition that Family Tax Benefit, which is an income tax credit to families with children, is not an income support payment. (Currently, a family with two children would receive income support even if the family earns $105,000 AUD.) Forty per cent of families in the administrative data have only ever received Family Tax Benefit or Child Care Benefit and have had no history of income support receipt.

The most commonly received income support payments in this population are unemployment benefits (Newstart Allowance) or payments to low-income parents with children (Parenting Payment Single or Parenting Payment Partnered). Table 1 provides information on the strata definitions, population percentages in each strata, and the code letters A-F by which we refer to the six strata in what follows.

Of particular interest in this paper will be the comparison between the group of respondents who have not received any income support (stratum A) and those who have received income support for more than six years out of the last twelve (stratum B). We will refer to this later group as those who have had heavy exposure to the income support system.

Table 1: Income Support Stratification Categories

Strata identifier	Stratification category	Proportion in population	Target Proportion in sample
A	No parental income support history	40.9%	25.0%
B	Heavy exposure to income support programs– family spent more than six (out of 12) total years on income support	27.5%	34.9%
C	First exposure prior to 1994 and less than six total years on income support	9.5%	12.1%
D	First exposure to income support system after 1998	8.5%	10.7%
E	First exposure to income support system between 1994 and 1998 and less than three total years on income support	8.5%	10.8%
F	First exposure to income support system between 1994 and 1998 and more than two but less than six total years on income support	5.1%	6.5%

2.1 Survey data, the incentive payment, and matching survey responses with administrative records

From this administrative data we drew a stratified random sample following the sample proportions given in the last column of Table 1. We selected a total of 1,400 youths with matched parents from this administrative data for the pilot survey prior to wave 1.[3] A small number of youth and parents called to opt out of the survey, an option they were given in the initial approach letter. We exclude these individuals from the sample. We also exclude any observations for whom the initial approach letter was returned to sender.

Table A1 in the appendix describes our sample in detail. For the purposes of this paper, we are interested in the 1,123 parents and 1,080 youth who we believe were obtainable through the telephone interview process. We exclude those who were unobtainable. The main reason that an individual was unobtainable was that the person answering the phone told us that this was not a valid phone number for the named sampled respondent (i.e. they did not know the named 18 year old or parent.)[4] This happened in 154 cases. There were also 100 cases in which the phone call was terminated before we could determine whether or not we had the right phone number/respondent. There were 54 that were terminated because the person answering the phone could not speak English sufficiently well to determine whether or not we

[3]Less than two per cent of the young adults had no parent identifiable in the administrative data and for this group there is only a young adult in the sample.

[4]Recall that our sampling frame was a list of named individuals, not households. Thus making contact with a household was not sufficient for that household or its members to be included in the sample. We required that the household contain a particular individual.

had the right phone number/respondent. The total of other exclusions is less than 10 and is detailed in panel 2 of Table A1. All of these unobtainable categories are marked with "?" in Table A1 and are excluded from our analysis.

The pilot had several purposes including testing the survey instrument and testing the ability of the survey design to produce interviews with matched pairs of youth and parents. For the purposes of this paper, we will focus on the incentive payment which was tested during the pilot. Fifty per cent of respondent pairs (parents and youths) were selected into an incentive sample by strata. The other half of the sample was not offered nor paid an incentive. On the basis of the pilot study and the results presented here, the incentive was incorporated into the entire sample for the main project.

The offered incentive payment was $15 AUD for completing the survey. In the case of the parents, this payment was paid upon completion of a 30 minute phone survey. For the youth, the payment was made upon completion of a 25-minute phone survey and receipt of a self-completion questionnaire which took approximately 10 minutes to complete. The self-completion questionnaire could be mailed back or completed online over a secure web site. For those in the sub-sample who were paid an incentive, participants were told in the initial approach letter that there was an incentive payment which would be paid upon survey completion. They were also reminded of this at the beginning of the phone interview.

In the survey, respondents were also asked to give permission to university researchers to link their administrative income support data with their survey responses. It was made clear to respondents that their survey responses would not be given back to Centrelink, the government agency which manages income support payments. The exact question was

> "Do you agree to having your survey answers linked by researchers at the Australian National University to information from your Centrelink records. This linking would be done at the Australian National University and your survey responses would not be given to Centrelink."

In addition to looking at the effect of the incentive on contactability and on response rates, we will also look at its effects on these last two elements of our survey design–the return of the self-completion questionnaire and the agreement to linking of administrative and survey records. We now turn to the detailed results.

3 Results

In this section, we look at four questions regarding response rates and data quality which might be related to the payment of an incentive. These are

1. Does an approach letter which includes information about an incentive payment increase the probability of being able to contact selected individuals? Does this effect vary based upon an individual's income support history?

2. Does payment of an incentive decrease the probability that a person who is contacted will refuse an interview? Does this effect vary by income support history?

3. Does payment of an incentive increase the probability that respondents will agree to having their survey responses matched to their administrative records? Does this vary by income support history?

4. Does payment of an incentive increase the probability that respondents will complete a self-completion questionnaire after a phone interview? Does this vary by income support history?

Our general approach will be to estimate probit models of the probability of each outcome. The exact definitions of the outcome variables in terms of the actual survey/contact outcome are provided in Table A1 of the appendix. We generally control for gender, age, marital status, number of kids, current income support status, and whether the individual is an immigrant to Australia.[5] For the youth, we also add a dummy variable equal to one if the he or she receives Youth Allowance which is a government payment with two variants. The first variant is an unemployment benefit which is paid to young people. Receipt of this benefit obliges the young person to engage in monitored job search or training activity. The second variant is paid to young people who are independent of their parents but who are studying full-time. We cannot distinguish, in our data, between these two types of youth allowance receipt.

We do not analyze partial response or incomplete response as 100 per cent of those participating completed the survey. This was despite survey lengths which went beyond what is considered acceptable for phone interviews. We attribute this, for the parents, to a great willingness to spend time on the phone talking about their kids. For the youth, the questionnaires included a range of questions which solicited their opinions on personal and societal values and respondents reported finding the process of answering the questionnaire to be an interesting one.

We discuss our results in detail in the next four sub-sections.

[5]Table A2 in the appendix provides descriptive statistics for each variable for the three main subsamples used in our analysis. Table A3 provides a cross-tabulation of the control variables by 0/1 outcome for each of the three main models we estimate. Table A4 provides precise definitions of the independent variables.

3.1 Do incentives help in contacting people?

Table 2 presents the results from a model in which we assess the probability that a person who is selected into the sample is contactable. For the parents, we have 1,080 individuals who are potentially obtainable. Of that group, we made contact with 691. For the youth, we had 1,123 in the sample of potentially obtainable people. We made contact with 755 of those. Telephone contact was attempted with all individuals in the sample at least eight times. For each individual, contact attempts included at least some attempts in the evenings and some on weekends. We model the probability of making contact as being a function of gender, age, marital status, number of children, immigrant status, and income support status. We estimate a model for parents, a model for youths, and a combined model with an indicator variable equal to one if the respondent is a youth and zero if the respondent is a parent. For each model we present weighted and unweighted estimates. We will primarily discuss the weighted estimates.[6]

Youth in families who have never been exposed to income support are 22 per cent more likely than individuals in families with heavy exposure to income support to be contactable in the survey in the non-incentive sample. (This is the difference between the coefficient on the dummy variable for stratum A and the coefficient on the dummy variable for stratum B.) This effect is highly significant.[7] There is also a large difference in the contactability of those in the intermediate income support exposure categories compared with those in the heavy exposure category. Those with less than three years exposure to the income support system and only since 1998 are 17 per cent more likely than individuals in families with heavy exposure to be contactable in the non-incentive sample. At the lower end of the spectrum, those whose first exposure was pre-1998 but who have less than three years are about 12 per cent more likely than individuals in families with heavy exposure to income support to be contactable in the non-incentive sample. Those with no exposure to the income support system are between four and ten per cent more likely to be contactable. These differences are fairly small and are only occasionally significant.

How does the promise, in an approach letter, of payment of an incentive change the picture? It dramatically and significantly reduces the gap in the probability of making contact with youth in the heavy exposure vs. no exposure to income support categories. With incentives, the difference in the contact rates of youths with no

[6] Appendix two discusses the procedure we used for weighting. Table A7 in appendix two provides information about the population sizes which were used in the calculation of the regression weights.

[7] Table A5 in the appendix provides the stratum-by-stratum comparison and the standard errors of the differences between stratum based upon the estimated coefficients from the weighted models of Table 2.

exposure and youths with a heavy exposure to the income support system is only eight per cent instead of 22 per cent. This eight per cent difference is not statistically significant.

This is a very important result. Without incentives, we are much more likely to make contact with those people from the wealthier end of the socio-economic spectrum. With incentives, we eliminate most of that difference. Sending an approach letter which mentions the incentive may make those to whom the incentive represents a larger fraction of their income proportionately more interested in responding to the survey. One can speculate as to how this effect might work. Individuals are looking out for the phone call instead of trying to avoid the interviewer and perhaps take the call rather than claiming that the sampled person is not at home.

We see a similar result when we look at the results for the parents. Those in the heavy exposure to income support category are 18 per cent more likely to be contactable when the incentive is proposed in the initial approach letter. The initial difference of 24 per cent in contactability between the no income support and heavy exposure categories is eliminated–it is less than 6 per cent and not significant.

We find a resounding yes to our first question–payment of an incentive improves the probability of getting the respondent on the telephone. The increased probability of response happens amongst those least well-off who are the most difficult to contact. Differences in the probability of making contact with sampled individuals in different socio-economic groups are eliminated by the incentive payment.

3.2 Do incentives help in reducing refusal?

Table 3 presents results for a probability model of refusal. The sample here includes only individuals who were contacted, 691 parents and 755 youth. Of these, 231 youth agreeded to being interview whereas 524 youth refused. For the parents, 266 agreed to be interviewed and 425 refused. On average, the refusal rate was much higher for young adults than for parents, which matches our a prior expectation that eighteen year-old young adults are a difficult group to interview.

We find significant differences in refusal rates, in the non-incentive sample between categories E and F on one hand and A, B, and C on the other.[8] These differences are difficult to explain on the basis of income support histories since the response rates of heavy exposure and no exposure look similar to each other but different to those with small amounts of exposure to the income support system. The heavy exposure group is less likely to refuse, which may be explained by the fact that they are frequently

[8] Table A6 in the appendix provides all of the differences between stratum and their standard errors based upon the estimated coefficients from the weighted models of Table 3.

surveyed and are perhaps used to the intrusion into their lives.

We find overall that the incentive does reduce refusal rates. The effect is concentrated in strata E and F–these groups had their first exposure to the income support system between 1994 and 1998. The first group has spent less than three years since 1994 on income support whereas the second group has spent between three and six years on income support between 1994 and 2006.

Once incentives are offered, the initial differences across strata in the non-incentive group are eliminated. In the incentive group, there is no difference across strata in refusal rates amongst those with whom we made contact. The patterns are similar for youth and parents and we can see this in the pooled model of columns six and seven of table 3.

3.3 Do incentives affect an individual's willingness to consent to linking survey and administrative data?

Broadly, we find that they do not. Table 4 presents the results from a model of the probability to agree to matching survey to administrative data. Here we use the sample of 497 youth and parents who agreed to being interviewed and completed a full interview.

Interestingly, youth were about 23 per cent less likely to refuse matching their survey responses to their administrative data than were parents. However, this difference was only significant at the 20 per cent level. Those currently on income support, again perhaps due to being more accustomed to government intrusion in their lives, were more likely to accept matching.

Due to the small sample sizes, we did not separately estimate models for young adults and parents. Very few individuals refused the match–only 25 out of 497. The failure to find much significant difference across strata or across incentives is perhaps due to the small number of refusals.

3.4 Do incentives encourage the return/completion of a self-completion questionnaire?

In a simple model, we find that incentives have no effect on the probability of returning the self-completion questionnaire. The self-completion questionnaire was only completed by the youth. Consequently, we estimate this model on the 231 youth who completed the phone questionnaire. Of these 152 returned the self-completion questionnaire, while 79 failed to return it. Those in the incentive sample were about 7 per cent less likely to return the self-completion questionnaire, but the difference was not significant. The sample size is quite small, so this is perhaps not surprising.

One theory which could justify a negative effect is that the promise of an incentive encouraged youth who otherwise might not have responded to complete the telephone questionnaire but that the additional effort of completing the self-completion questionnaire outweighed the benefit of the small cash incentive.

Table 5 also presents the results by strata, but, not surprisingly given the quite small stratum-specific sample sizes, it is difficult to discern any particular patterns in the data.

4 Conclusion

We tested the payment of a small cash incentive, $15 AUD, for completing a telephone survey on a sample of individuals drawn from administrative records related to income support and family tax credit data in Australia. The sample included matched 18-year-old young adults and their parents (usually their natural mother.) Overall, despite the small size of the cash incentive, we found a large and significant effect on overall response rates to the survey from payment of the incentive. Of the original sample to whom we sent approach letters, 33 per cent of parents responded to the survey in the absence of an incentive. For those who were offered an incentive, we find that 40 per cent responded. This represents a significant increase in response rates.

For young adults, we find almost identical results. In the absence of an incentive, 32.6 per cent respond in the absence of an incentive, whereas almost 39 per cent respond once offered an incentive. Again the difference is statistically and methodologically significant. There is a large statistical literature on the positive effect of incentives on response rates; e.g. Berk et al. (1987), Brick et al. (2005), Dawson and Dickinson (1988), Godwin (1979), James and Bolstein (1992), McDaniel and Rao (1980), Singer et al. (1999), Teisl et al. (2005), Singer and Kulka (2002). Our results are consistent with the standard results in the literature.

We have two findings which we believe are unique and which add to this literature. The first is that the effect of incentives appears to work in two distinct ways. The first is that the promise of incentives in an approach letter increases the probability that contact will be made with a selected individual in the sample quite apart from whether the individual chooses to respond to the survey or not. The second is the traditional result that respondents who are contacted are more likely to respond if they are paid an incentive. We find statistically significant effects for both of these channels.

Secondly, we find that incentives work to reduce response bias related to socio-economic characteristics. Our data are drawn from income support and tax credit records. We stratify the data by the intensity and recency of the family's receipt of

income support since 1993. We find that in the non-incentive sample there are large differences (20 per cent and greater) in the probability of contacting those in the group who have had heavy exposure to income support relative to those who have received no income support in the previous 12 years. The wealthier group is much easier to contact. Importantly, the payment of an incentive almost completely removes this effect. The those with relatively high socio-economic status are not much affected by the incentive, but the contactability of the group with heavy exposure to income support increases so much that there is no longer any significant differences between these two groups. This is good news for the use of incentives–not only does it increase response rates, it also reduces selection bias.

Once people were contacted, we find higher refusal rates amongst those with moderate levels of contact with the income support system in the distant past (over six years ago). These higher refusal rates are relative both to those with heavy exposure to the income support system and those with no exposure to the income support system. Interviewers began the interview by explaining the source of the data and these individuals with moderate past exposure may have found it odd to be contacted based upon something that was over eight years old and this may have raised suspicions about the purpose or scientific validity of the survey. It was precisely amongst this group that the incentive payments had the largest positive effect on response rates. Again, incentive payments not only increased the average response rate, but the promise of the incentive increased the response rates amongst those groups that had the lowest response rates. Again this is good news both in terms of average response rates and in terms of bias reduction.

The literature is quite convincing regarding the positive effects of incentives. This literature has mostly focused on average effects. Here, we confirm those results and extend them. Our extension is important in that we show that it is precisely amongst the groups that are most difficult to contact and most likely to refuse that incentives work the most. Fears have been expressed that incentives could exacerbate response bias if it increases response rates more amongst those who are already responding more (see Shettle and Mooney (1999).) Our results argue that in fact exactly the opposite is happening. Incentives reduce refusals and improve contactability in a way that also reduces response bias from differential response rates across socio-economic categories.

References

Berk, M. L., Mathiowetz, N. A., Ward, E. P., and White, A. A. (1987). The effect of prepaid and promised incentives: results of a controlled experiment. *Journal of Official Statistics*, 3(4):449–457.

Breunig, R., Cobb-Clark, D., Gorgens, T., and Sartbayeva, A. (2007). User's guide to the youth in focus data, version 1.0. Youth in focus project discussion paper series, no. 1, Australian National University. Available at: `http://youthinfocus.anu.edu.au/pdf/YIF%20Technical%20Paper%201%2019%20October%2007.pdf`.

Brick, J. M., Montaquila, J., Hagedorn, M. C., Roth, S. B., and Chapman, C. (2005). Implications for RDD design from an incentive experiment. *Journal of Official Statistics*, 21(4):571–589.

Centrelink (2007). A guide to Australian Government payments. Centrelink Publication, Commonwealth of Australia 2007. Available at: `http://www.centrelink.gov.au/internet/internet.nsf/filestores/co029_0702/$file/co029_0702en.pdf`.

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

Dawson, S. and Dickinson, D. (1988). Conducting international mail surveys: the effect of incentives on response rates with an industry population. *Journal of International Business Studies*, 19(3):491–496.

Godwin, R. K. (1979). The consequences of large monetary incentives in mail surveys of elites. *The Public Opinion Quarterly*, 43(3):378–387.

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

McDaniel, S. W. and Rao, C. P. (1980). The effect of monetary inducement on mailed questionnaire response quality. *Journal of Marketing Research*, 17(2):265–268.

Shettle, C. and Mooney, G. (1999). Monetary incentives in U.S. Government surveys. *Journal of Oficial Statistics*, 15(2):231–250.

Singer, E. and Kulka, R. A. (2002). *Studies of Welfare Populations: Data Collection and Research Issues*, chapter Paying Respondents for Survey Participation, pages 105–128. Committee on National Statistics, Division of Behavioral and Social Sciences and Education, National Research Council.

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

Teisl, M. F., Roe, B., and Vayda, M. (2005). Incentive effects on response rates, data quality, and survey administration costs. *International Journal of Public Opinion Research*, 18(3):364–373.

Table 2: **Dependent variable is whether the person was CONTACTED. Probit Marginal Effects.**

	(1) Youth				*(2) Parents*				*(3) Youth and Parents*			
	No weights		*Weights*		*No weights*		*Weights*		*No weights*		*Weights*	
Variable	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*
Male	.040	(.029)	.069	(.038)	−.107	(.085)	−.167	(.104)	.021	(.027)	.038	(.035)
Currently on Income Support	−.298	(.132)	−.419	(.156)	.002	(.041)	.063	(.049)	−.012	(.037)	.019	(.047)
Married or partnered	−.013	(.210)	−.024	(.237)	.070	(.037)	.101	(.048)	.056	(.033)	.081	(.043)
Receiving Youth Allowance	.282	(.097)	.379	(.096)					.035	(.045)	.038	(.059)
Number of kids	.242	(.087)	.264	(.103)	−.002	(.010)	−.006	(.012)	−.001	(.009)	−.004	(.012)
Immigrant	−.131	(.063)	−.133	(.081)	−.076	(.036)	−.039	(.046)	−.088	(.031)	−.059	(.040)
Age					.002	(.003)	.006	(.004)	.001	(.003)	.004	(.004)
Strata A	.197	(.039)	.205	(.049)	.053	(.144)	−.148	(.191)	.118	(.125)	−.037	(.183)
Strata B	.015	(.051)	−.012	(.057)	−.152	(.162)	−.380	(.186)	−.081	(.150)	−.265	(.188)
Strata C	.144	(.043)	.119	(.049)	.054	(.146)	−.174	(.204)	.089	(.132)	−.087	(.193)
Strata D	.177	(.042)	.162	(.044)	.032	(.146)	−.181	(.200)	.096	(.129)	−.068	(.190)
Strata E	.124	(.044)	.107	(.048)	.069	(.138)	−.133	(.196)	.084	(.129)	−.080	(.188)
Strata F	.162	(.042)	.144	(.046)	.072	(.140)	−.140	(.199)	.106	(.127)	−.061	(.189)
Strata A × incentive	−.120	(.076)	−.118	(.075)	−.008	(.075)	−.012	(.075)	−.066	(.054)	−.069	(.053)
Strata B × incentive	.017	(.066)	.020	(.068)	.166	(.058)	.183	(.060)	.087	(.045)	.093	(.046)
Strata C × incentive	−.034	(.069)	−.034	(.071)	−.025	(.072)	−.010	(.073)	−.034	(.050)	−.029	(.051)
Strata D × incentive	.045	(.068)	.047	(.070)	.019	(.071)	.020	(.072)	.034	(.049)	.036	(.049)
Strata E × incentive	.027	(.068)	.027	(.070)	.020	(.071)	.018	(.072)	.026	(.049)	.025	(.050)
Strata F × incentive	−.021	(.071)	−.029	(.074)	−.056	(.073)	−.058	(.074)	−.035	(.051)	−.035	(.051)
Youth Indicator									.060	(.092)	.151	(.119)
Joint Test for significance of interactions: χ^2_6 and (p−value)	3.61	(.73)	3.58	(.73)	7.12	(.31)	8.05	(.23)	6.66	(.35)	7.06	(.31)
Log−likelihood	−693		−697		−690		−688		−1,391		−1,401	
Observations	1,123		1,123		1,080		1,080		2,203		2,203	

Notes: Robust standard errors in parentheses. See Appendix Table A1 for a definition of *CONTACTABLE*. I use Stata's `mfx` command. For dummy variables, the marginal effects are calculated as the difference in probability when the dummy variable is set to one and when it is set to zero. Appendix two discusses the weights used in estimation. See Appendix Tables A2 and A3 for descriptive statistics and cross tabulations.

Table 3: **Dependent variable is whether the person REFUSED being interviewed. Probit Marginal Effects.**

	(1) Youth				*(2) Parents*				*(3) Youth and Parents*			
	No weights		*Weights*		*No weights*		*Weights*		*No weights*		*Weights*	
Variable	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*
Male	.081	(.037)	.086	(.037)	−.102	(.105)	−.107	(.106)	.066	(.035)	.070	(.035)
Currently on Income Support	.314	(.234)	.343	(.232)	.005	(.056)	.004	(.056)	.032	(.052)	.034	(.052)
Married or partnered	.036	(.315)	.018	(.308)	.020	(.048)	.026	(.049)	.037	(.047)	.041	(.047)
Receiving Youth Allowance	−.335	(.196)	−.366	(.191)					−.066	(.061)	−.077	(.061)
Number of kids	−.231	(.206)	−.227	(.212)	−.030	(.014)	−.033	(.014)	−.035	(.014)	−.037	(.014)
Immigrant	.196	(.084)	.216	(.083)	.026	(.047)	.024	(.047)	.074	(.041)	.078	(.042)
Age					−.005	(.004)	−.004	(.004)	−.004	(.004)	−.004	(.004)
Strata A	−.108	(.059)	−.110	(.059)	.159	(.210)	.122	(.212)	.157	(.201)	.127	(.204)
Strata B	−.153	(.069)	−.154	(.069)	.210	(.216)	.179	(.220)	.149	(.205)	.123	(.207)
Strata C	−.137	(.063)	−.138	(.063)	.128	(.212)	.094	(.213)	.126	(.203)	.098	(.205)
Strata D	−.075	(.064)	−.076	(.064)	.333	(.190)	.299	(.196)	.255	(.193)	.228	(.198)
Strata E	.001	(.066)	−.001	(.066)	.277	(.192)	.246	(.197)	.264	(.189)	.238	(.194)
Strata F	.101	(.068)	.098	(.068)	.198	(.203)	.164	(.206)	.285	(.189)	.259	(.194)
Strata A × incentive	−.105	(.083)	−.104	(.083)	.016	(.092)	.015	(.092)	−.048	(.062)	−.048	(.062)
Strata B × incentive	−.013	(.098)	−.011	(.098)	−.012	(.108)	−.010	(.108)	−.010	(.072)	−.009	(.072)
Strata C × incentive	.010	(.090)	.010	(.091)	−.007	(.093)	−.004	(.093)	−.001	(.065)	−.0	(.065)
Strata D × incentive	−.080	(.081)	−.081	(.080)	−.171	(.075)	−.169	(.075)	−.121	(.056)	−.121	(.055)
Strata E × incentive	−.229	(.069)	−.229	(.069)	−.171	(.073)	−.173	(.072)	−.195	(.051)	−.197	(.050)
Strata F × incentive	−.228	(.068)	−.226	(.067)	.003	(.089)	.003	(.089)	−.123	(.055)	−.123	(.055)
Youth Indicator									−.185	(.127)	−.171	(.128)
Joint Test for significance of interactions: χ^2_6 and (p−value)	18.75	(.005)	18.66	(.005)	8.69	(.19)	8.67	(.19)	20.74	(.002)	21.03	(.002)
Log−likelihood	−495		−493		−455		−454		−958		−956	
Observations	755		755		691		691		1,446		1,446	

Notes: Robust standard errors in parentheses. See Appendix Table A1 for a definition of *REFUSED*. I use Stata's `mfx` command. For dummy variables, the marginal effects are calculated as the difference in probability when the dummy variable is set to one and when it is set to zero. Appendix two discusses the weights used in estimation. See Appendix Tables A2 and A3 for descriptive statistics and cross tabulations.

Table 4: **Dependent variables is whether the person REFUSED the MATCH of administrative with survey data. Probit Marginal Effects.**

	(1) *No weights*		*(2)* *Weights*	
Variable	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*
Incentive	.018	(.016)	.013	(.017)
Male	−.011	(.022)	−.011	(.023)
Currently on Income Support	−.043	(.023)	−.042	(.024)
Married or partnered	−.029	(.024)	−.027	(.023)
Receiving Youth Allowance	.114	(.090)	.104	(.086)
Number of kids	.004	(.008)	.004	(.008)
Immigrant	−.006	(.024)	−.007	(.023)
Age	−.006	(.002)	−.006	(.002)
Strata A	.335	(.427)	.333	(.420)
Strata B	.261	(.423)	.239	(.378)
Strata C	.295	(.430)	.290	(.416)
Strata D	.333	(.441)	.349	(.460)
Strata E	.199	(.376)	.198	(.377)
Strata F	.241	(.418)	.247	(.430)
Youth Indicator	−.241	(.141)	−.232	(.137)
Log−likelihood	−93.29		−93.29	
Observations	497		497	

Notes: Robust standard errors in parentheses. See Appendix Table A1 for a definition of *REFUSED MATCH*. The model does not include the interaction terms between strata and incentive group because some of them perfectly predict the outcome. For dummy variables, the marginal effects are calculated as the difference in probability when the dummy variable is set to one and when it is set to zero. See Appendix Tables A2 and A3 for descriptive statistics and cross tabulations.

Table 5: **Dependent variables is whether YOUTH returned Self-Completion Questionnaire given that she completed the phone interview. Probit Marginal Effects.**

	(1)		*(2)*		*(3)*	
	Only characteristics		*Strata and incentive interactions*		*All variables*	
Variable	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*	*Mg. E.*	*S.E.*
Incentive	−.072	(.064)				
Male	−.169	(.062)			−.179	(.064)
Currently on Income Support	.081	(.068)			.082	(.077)
Immigrant	−.154	(.191)			−.083	(.199)
Strata A			.201	(.082)	.279	(.078)
Strata B			.256	(.084)	.282	(.081)
Strata C			.000	(.108)	.084	(.114)
Strata D			.111	(.104)	.195	(.104)
Strata E			.231	(.096)	.271	(.076)
Strata F			.251	(.098)	.284	(.089)
Strata A × incentive			−.134	(.144)	−.178	(.149)
Strata B × incentive			−.176	(.185)	−.229	(.184)
Strata C × incentive			.127	(.128)	.114	(.134)
Strata D × incentive			.082	(.133)	.079	(.139)
Strata E × incentive			−.306	(.177)	−.287	(.174)
Strata F × incentive			−.063	(.217)	−.054	(.212)
Joint Test for significance of interactions: χ^2_6 and (p−value)			5.94	(.430)	6.89	(.331)
Log−likelihood	−143.0		−143.3		−138.6	
Observations	231		231		231	

Notes: Robust standard errors in parentheses. Self-completion questionnaires were not administered to parents. Results do not include weights. Twenty five (25) people who REFUSED the MATCH of survey with administrative data are excluded.

Appendix One: Variable definitions and descriptive statistics

Table A1: **Definition of dependent variables and sample sizes**

	Contactable	*Refused*	*Refused Match*	*Obs.*
Agreed to interview	1	0	**?**	350
Answering machine	0	·	·	61
Complete	1	0	0	231
Complete, match refused	1	0	1	25
Completed	1	0	0	241
Engaged	0	·	·	1
Fax / modem	0	·	·	2
Fax modem	0	·	·	2
General appointment	1	1	·	127
No reply	0	·	·	14
Not willing to participate at SCR2	1	1	·	215
Number tried 3+ times engaged / no reply / answer or 10+ times called with no reply last	0	·	·	236
Refusal	1	1	·	231
Respondent can not provide information (Code 2 in Q7C)	0	·	·	4
Terminate - other not specified	**?**	·	·	100
Termination - Business number	0	·	·	1
Termination - Hearing difficulty / very elderly / drunk	**?**	·	·	2
Termination - No-one in household fits introduction criteria	0	·	·	1
Termination - hearing difficulty / very elderly / drunk	**?**	·	·	1
Termination - language problem	**?**	·	·	54
Termination - named sample respondent not at this number	0	·	·	154
Termination - respondent did not wish to continue interview	1	1	·	23
Termination - respondent wants to be sent new letter	1	0	**?**	3
Unobtainable	0	·	·	281
Missing values (several causes)	·	·	·	482
Total observations	2203	1446	497	2842

Notes: Each column represents the definition of a dependent variable (except for the first one). Zeros and ones mean that the variable takes those values (usable observations); "·" means that those observations are excluded; and "**?**" means that there is some ambiguity as to how observations in these categories are to be classified (we exclude all these observations from the analysis).

Table A2: **Descriptive statistics by sample.**

	Contactable				*Refused*				*Refused Match*			
	Mean	*Std. Dev.*	*Min*	*Max*	*Mean*	*Std. Dev.*	*Min*	*Max*	*Mean*	*Std. Dev.*	*Min*	*Max*
Contactable	.656	.475	0	1								
Refused					.412	.492	0	1				
Refused Match									.050	.219	0	1
Incentive	.502	.500	0	1	.503	.500	0	1	.559	.497	0	1
Male	.291	.454	0	1	.301	.459	0	1	.264	.441	0	1
Currently in Income Support	.281	.450	0	1	.265	.441	0	1	.296	.457	0	1
Married or partnered	.323	.468	0	1	.331	.471	0	1	.35	.477	0	1
Receiving Youth Allowance	.148	.355	0	1	.149	.357	0	1	.157	.364	0	1
Number of kids	1.54	1.91	0	17	1.49	1.88	0	17	1.67	1.99	0	12
Immigrant	.145	.352	0	1	.128	.334	0	1	.113	.317	0	1
Age	31.52	14.26	18	74	31.20	14.23	18	68	32.26	14.33	18	60
Strata A	.161	.368	0	1	.167	.373	0	1	.227	.420	0	1
Strata B	.159	.366	0	1	.129	.336	0	1	.155	.362	0	1
Strata C	.172	.377	0	1	.17	.376	0	1	.173	.379	0	1
Strata D	.168	.374	0	1	.181	.385	0	1	.181	.385	0	1
Strata E	.168	.374	0	1	.176	.381	0	1	.145	.352	0	1
Strata F	.172	.377	0	1	.176	.381	0	1	.119	.324	0	1
Strata A $\times$ incentive	.082	.275	0	1	.082	.274	0	1	.125	.331	0	1
Strata B $\times$ incentive	.078	.268	0	1	.069	.254	0	1	.085	.278	0	1
Strata C $\times$ incentive	.085	.279	0	1	.082	.274	0	1	.078	.269	0	1
Strata D $\times$ incentive	.088	.283	0	1	.096	.295	0	1	.121	.326	0	1
Strata E $\times$ incentive	.084	.278	0	1	.09	.286	0	1	.093	.290	0	1
Strata F $\times$ incentive	.085	.279	0	1	.085	.279	0	1	.058	.235	0	1
Obs	2203				1446				497			

Notes: See Table A1 for definitions of *CONTACTABLE*, *REFUSED*, and *REFUSED MATCH*.

Table A3: **Cross-tabulations of dependent variables with explanatory variables.**

	Contactable[a]			*Refused*[b]			*Refused Match*[c]		
	Mean	*No*	*Yes*	*Mean*	*No*	*Yes*	*Mean*	*No*	*Yes*
Incentive	.497	.496	.498	.511	.559	.457	.559	.680	.553
Male	.288	.276	.295	.291	.264	.321	.264	.200	.267
Currently in Income Support	.290	.319	.273	.275	.296	.251	.296	.240	.299
Married or partnered	.329	.307	.343	.343	.350	.334	.350	.400	.347
Receiving Youth Allowance	.147	.149	.145	.145	.157	.132	.157	.160	.157
Number of kids	1.55	1.6	1.52	1.56	1.67	1.42	1.67	1.96	1.66
Immigrant	.161	.172	.154	.126	.113	.141	.113	.120	.112
Age	31.74	31.89	31.64	31.75	32.26	31.18	32.26	32.32	32.26
Strata A	.159	.151	.165	.194	.227	.157	.227	.320	.222
Strata B	.166	.222	.133	.139	.155	.121	.155	.120	.157
Strata C	.173	.169	.175	.166	.173	.159	.173	.160	.174
Strata D	.167	.145	.180	.191	.181	.202	.181	.240	.178
Strata E	.164	.151	.171	.160	.145	.177	.145	.080	.148
Strata F	.171	.163	.176	.150	.119	.184	.119	.080	.121
Strata A $\times$ incentive	.081	.083	.079	.098	.125	.067	.125	.200	.121
Strata B $\times$ incentive	.080	.097	.069	.074	.085	.063	.085	.040	.087
Strata C $\times$ incentive	.086	.087	.085	.081	.078	.083	.078	.080	.078
Strata D $\times$ incentive	.084	.068	.094	.107	.121	.092	.121	.200	.117
Strata E $\times$ incentive	.083	.075	.088	.082	.093	.070	.093	.080	.093
Strata F $\times$ incentive	.084	.086	.083	.070	.058	.083	.058	.080	.057
Obs	2360	883	1477	943	497	446	497	25	472

Notes: [a] The figures in column *No* give the average of the variables on the left (i.e. incentive, male, etc.) for the subsample that are *Not Contactable*. The figures in column *Yes* give the average of the left variables (i.e. incentive, male, etc.) for the subsample that are *Contactable*.
[b] The figures in column *No* give the average of the variables on the left(i.e. incentive, male, etc.) for the subsample that *Did Not Refused* the interview. The figures in column *Yes* give the average of the variables for the subsample that *Refused* the interview.
[c] The figures in column *No* give the average of the variables on the left(i.e. incentive, male, etc.) for the subsample that *Did Not Refused Match* of information. The figures in column *Yes* give the average variables for the subsample that *Refused Match* of information. For all samples, *Mean* is the average of the variables on the left for the whole subsample (*Yes* and *No*).

Table A4: **Definition of covariates.**

Variable	Description	Notes
Incentive	= 1 if person was offered a monetary incentive, 0 otherwise	
Male	= 1 if person is male	
Currently in Income Support	= 1 if person is currently receiving income support of any type, 0 otherwise.	As of January, 2006.
Married or partnered	= 1 if currently married or partnered, 0 otherwise	Those with missing marital status (624) or unknown (123) are assumed to be single (all of them are youth). Also take into account that for those people not receiving income support we don't know their actual (as of January 2006) marital status.
Receiving Youth Allowance	= 1 if person is currently receiving income support of the Youth allowance type, 0 otherwise.	As of January, 2006.
Number of kids	Number of individuals FTB/FTA children associated with partner or spouse, and 0 otherwise.	Missing values set to zero.
Immigrant	= 1 if NOT born in Australia, 0 otherwise	Missing values are set to Australians (688). From administrative data.
Age	age in years (integer numbers) at April 1, 2006	
Strata A	= 1 if person is in Strata A, 0 otherwise	
Strata B	= 1 if person is in Strata B, 0 otherwise	
Strata C	= 1 if person is in Strata C, 0 otherwise	
Strata D	= 1 if person is in Strata D, 0 otherwise	
Strata E	= 1 if person is in Strata E, 0 otherwise	
Strata F	= 1 if person is in Strata F, 0 otherwise	
Strata A × incentive	= 1 if person is in Strata A and was offered incentive, 0 otherwise	
Strata B × incentive	= 1 if person is in Strata B and was offered incentive, 0 otherwise	
Strata C × incentive	= 1 if person is in Strata C and was offered incentive, 0 otherwise	
Strata D × incentive	= 1 if person is in Strata D and was offered incentive, 0 otherwise	
Strata E × incentive	= 1 if person is in Strata E and was offered incentive, 0 otherwise	
Strata F × incentive	= 1 if person is in Strata F and was offered incentive, 0 otherwise	

Table A5: Differences in strata dummy variables from weighted models of Table 2 and p-value for test of equatliy across strata

Youth **Parent**

Non-incentive model

Strata i = Strata j

	A	B	C	D	E
B	.608 (.205) {.003}				
C	.225 (.205) {.271}	-.383 (.185) {.039}			
D	.082 (.205) {.69}	-.526 (.20) {.008}	-.143 (.20) {.473}		
E	.264 (.201) {.189}	-.344 (.194) {.076}	.038 (.195) {.844}	.182 (.20) {.363}	
F	.14 (.207) {.50}	-.468 (.19) {.014}	-.086 (.192) {.656}	.058 (.204) {.776}	-.124 (.199) {.533}

(std. error) , {p-value}

Strata i = Strata j

	A	B	C	D	E
B	.606 (.227) {.008}				
C	.053 (.206) {.797}	-.553 (.203) {.006}			
D	.072 (.198) {.715}	-.534 (.222) {.016}	.019 (.202) {.924}		
E	-.05 (.20) {.804}	-.656 (.217) {.003}	-.103 (.20) {.609}	-.122 (.198) {.538}	
F	-.032 (.20) {.873}	-.638 (.209) {.002}	-.085 (.195) {.663}	-.104 (.197) {.597}	.018 (.196) {.928}

(std. error) , {p-value}

Incentive model

Strata i + Strata I * Incent. = Strata j + Strata j * Incent

	A	B	C	D	E
B	.241 (.197) {.221}				
C	.006 (.196) {.974}	-.235 (.192) {.222}			
D	-.361 (.193) {.062}	-.602 (.202) {.003}	-.367 (.20) {.066}		
E	-.122 (.193) {.527}	-.363 (.198) {.067}	-.128 (.197) {.515}	.239 (.198) {.227}	
F	-.095 (.195) {.627}	-.336 (.195) {.086}	-.101 (.193) {.601}	.266 (.199) {.181}	.027 (.197) {.889}

(std. error) , {p-value}

Strata i + Strata I * Incent. = Strata j + Strata j * Incent

	A	B	C	D	E
B	.036 (.224) {.872}				
C	.046 (.207) {.823}	.01 (.195) {.958}			
D	-.015 (.195) {.937}	-.051 (.218) {.813}	-.062 (.202) {.76}		
E	-.131 (.196) {.504}	-.167 (.212) {.431}	-.177 (.198) {.37}	-.115 (.193) {.549}	
F	.087 (.199) {.661}	.051 (.199) {.797}	.041 (.189) {.829}	.103 (.195) {.598}	.218 (.192) {.256}

(std. error) , {p-value}

Table A6: Differences in strata dummy variables from weighted models of Table 3 and p-value for test of equatliy across strata

Youth **Parent**

Non-incentive model

Strata i = Strata j

	A	B	C	D	E
B	.122 (.244) {.616}				
C	.076 (.227) {.737}	-.046 (.238) {.847}			
D	-.092 (.221) {.677}	-.214 (.242) {.376}	-.168 (.227) {.459}		
E	-.289 (.222) {.194}	-.411 (.244) {.092}	-.365 (.228) {.11}	-.197 (.225) {.383}	
F	-.539 (.223) {.016}	-.661 (.238) {.006}	-.615 (.225) {.006}	-.446 (.225) {.047}	-.25 (.226) {.269}

(std. error) , {p-value}

Strata i = Strata j

	A	B	C	D	E
B	-.143 (.298) {.632}				
C	.073 (.25) {.77}	.216 (.28) {.441}			
D	-.455 (.238) {.056}	-.312 (.288) {.278}	-.528 (.242) {.029}		
E	-.315 (.238) {.187}	-.172 (.283) {.544}	-.388 (.239) {.104}	.141 (.231) {.543}	
F	-.106 (.237) {.655}	.037 (.282) {.896}	-.179 (.238) {.452}	.349 (.231) {.131}	.209 (.228) {.36}

(std. error) , {p-value}

Incentive model

Strata i + Strata I * Incent. = Strata j + Strata j * Incent

	A	B	C	D	E
B	-.125 (.254) {.622}				
C	-.225 (.241) {.35}	-.10 (.248) {.686}			
D	-.156 (.225) {.489}	-.031 (.241) {.899}	.07 (.226) {.759}		
E	.087 (.237) {.714}	.212 (.25) {.396}	.312 (.237) {.187}	.243 (.222) {.274}	
F	-.17 (.234) {.468}	-.045 (.243) {.854}	.056 (.23) {.809}	-.014 (.219) {.949}	-.256 (.23) {.264}

(std. error) , {p-value}

Strata i + Strata I * Incent. = Strata j + Strata j * Incent

	A	B	C	D	E
B	-.078 (.268) {.771}				
C	.123 (.249) {.621}	.201 (.251) {.423}			
D	.056 (.233) {.809}	.134 (.262) {.608}	-.067 (.246) {.786}		
E	.209 (.23) {.363}	.288 (.256) {.262}	.086 (.242) {.721}	.153 (.229) {.504}	
F	-.074 (.238) {.756}	.004 (.252) {.987}	-.197 (.241) {.413}	-.13 (.237) {.581}	-.284 (.232) {.222}

(std. error) , {p-value}

Appendix Two: Weighting

This section describes the way in which we calculate weights to be included in estimation. We calculate weights for Youth and Parents separately. For each of these groups we calculate weights for each strata ($A, B, C, D, E,$ and F).

Weights for the CONTACTABLE model

For this model we calculate weights as:

$$P_y\left(Strata_i\right) = \frac{\textit{Youths Selected for Pilot in Strata}_i}{\textit{Total Youth in Strata}_i} \tag{1}$$

where $Strata_i$ represents the different strata (e.g. $i = A, B, C, D, E,$ and F). To calculate weights for the parent sample using equation (1) replace $P_y()$ by $P_p()$ and Youth by Parent. See Table for the information used to calculate the weights.

Weights for the REFUSED model

For the REFUSED model we take into account the fact that in order to refuse being interviewed people must be contacted first. That is, Refusing or Not Refusing the interview is conditional on being contacted. We calculate weights as:

$$P_y\left(Strata_i | Contacted\right) = \frac{P_y\left(Strata_i, Contacted\right)}{P_y\left(Contacted\right)}, \tag{2}$$

where $P_y()$ denotes probabilities calculated for the youth sample and $Strata_i$ represents the different strata (e.g. $i = A, B, C, D, E,$ and F). $P_y\left(Contacted\right)$ is calculated as the proportion of the sample (of youth) that was Contacted, and $P_y\left(Strata_i, Contacted\right)$ as the proportion of the sample (of youth) in $Strata_i$ that was Contacted. To calculate weights for the parent sample replace $P_y()$ by $P_p()$.

Weights for the REFUSED MATCH model

For this model we calculate weights as:

$$P_y\left(Strata_i|Contacted, Refused\right) = \frac{P_y\left(Strata_i, Refused|Contacted\right)}{P_y\left(Refused|Contacted\right)}. \tag{3}$$

In this case we exclude from the calculations 193 youth and 160 parents. The majority of these people, 350 in total, agreed to an interview but where not actually interviewed. The other 3 people were contacted (they did not explicitly refuse), but requested a a new approach letter to be sent. In these two cases, it is impossible to know whether these people would have allowed the match of survey and administrative data; so, we exclude them.

Weights for the YOUTH returned Self-Completion Questionnaire

Self-Completion Questionnaires (SCQ) were only administered to youth to collect extra information about them. Youth were asked to complete SCQ once they were Contacted, did not Refused the interview or Match with administrative data, and actually completed the phone interview. Let $\mathscr{I}$ represent the events (i) Contacted and (ii) Not Refusing the Interview or Match; and let $\mathscr{C}$ denote the event Completed Phone Interview. Then, we can write the weights as

$$P_y\left(Strata_i|\mathscr{C}, \mathscr{I}\right) = \frac{P_y\left(Strata_i, \mathscr{C}|\mathscr{I}\right)}{P_y\left(\mathscr{C}|\mathscr{I}\right)}. \tag{4}$$

The information used to calculate weights for this and all other models is reported in Table (). Once we calculate these probabilities we take their inverse and use Stata's `pweight` option in the estimation of probit models.

(Date: April 24, 2008)

Table A7: **Information used to calculate weights for different models**

	Contacted Model		*Refused Interview Model*		*Refused Match Model*		*Returned SCQ model*	
Strata	*Selected for pilot*	*Selected and Not Selected*	*Contacted*	*Contacted and Not Contacted*	*Refused Interview*	*Refused and Not Refused Int.*	*Completed Ph. Interview*	*Completed and Not Completed Ph. Inter.*
Youth								
A	203	17,869	127	184	49	105	34	52
B	228	12,032	101	185	37	75	25	35
C	226	3,549	129	193	52	88	20	35
D	218	3,692	140	187	60	109	32	47
E	217	4,262	125	183	52	88	21	36
F	228	2,171	133	191	70	97	20	26
Total	1,320	43,575	755	1,123	320	562	152	231
Parent								
A	199	16,889	115	171	45	102		
B	232	11,878	86	165	34	73		
C	224	3,500	117	185	39	89		
D	218	3,636	122	184	56	97		
E	219	4,217	129	187	52	88		
F	225	2,139	122	188	50	82		
Total	1,317	42,259	691	1,080	276	531		