Nonresponse and Panel Attrition in a Mobile Phone Panel Survey

Marek Fuchs[1]

Darmstadt University of Technology
Marktplatz 15, 64283 Darmstadt, Germany

Abstract

Over the course of the past decades, telephone surveys have become an important data collection mode. Due to the widespread use of mobile phones, however, landline telephone surveys are prone to increasing coverage errors. Thus, survey researchers aim to compensate for coverage errors in landline telephone surveys by means of integrating mobile phone telephone numbers into their samples. Accordingly, dual frame surveys are about to become state of the art. This has the potential to evoke differential total survey error, since mobile phone surveys differ from landline surveys with respect to several properties. In order to develop a better understanding of the pros and cons of mobile phone surveys, the German Experimental Mobile Phone Panel, funded by the German Research Foundation, has been established since 2009. It offers the opportunity to study the extent and potential counter measures for various components of the total survey error in mobile phone surveys. In the present paper, we report preliminary results from an analysis of nonresponse and panel attrition in the Experimental Mobile Phone Panel. In particular, we focus on pre-call validation methods for randomly generated mobile phone numbers which help validate mobile phone numbers. Also, we focus on nonresponse bias and assess the potential benefits of various methods that aim to counteract nonresponse and nonresponse bias. In addition to text message announcements we test the effect of incentives provided to potential mobile phone respondents. In addition, we assess the potential benefits of a mobile web component for nonresponding sample members. Overall, results suggest that mobile phone surveys are prone to considerable nonresponse, however, they also offer the opportunity to fight nonresponse and panel attrition by means which were not available in the traditional landline telephone survey environment.

Introduction

Overall, telephone coverage has approached saturation in most Western countries. Landline and mobile phone penetration taken together have reached values of 99 percent and above (Busse & Fuchs, 2012). This increase is mostly due to the wide-spread use of mobile phones, which has reached about 85 percent in the general population. By contrast, landline telephone coverage has declined since reaching its maximum in the mid 1990s. Today, only about 90 percent of the general population in Germany can be reached by landline telephone. In most other Western countries the decline of landline penetration is even more severe, with Sweden being a noticeable exception: here landline coverage remained on considerable high levels above 95 percent. On the contrary, in the Baltic countries, Finland, The Czech Republic or Rumania, landline penetration has dropped to below 50 percent. Since mobile phone penetration has increased in many countries, the backdrop of landline penetration has been compensated for by the increasing mobile phone penetration. This led to a paradoxical situation where overall telephone coverage has increased while coverage for each of the communication channels (landline or mobile phones) is still limited.

While in some countries like Finland, the Czech Republic or Slovakia, mobile only rates have exceeded landline penetration rates, other countries are far behind with respect to this development. Most prominently, Sweden and the Netherlands as well as Germany house rather small mobile only populations. In Germany, the proportion of mobile onlys is currently in the range of 7 to 9 percent, depending on the method used to estimate this proportion. In the United States the substitution of landline telephones by mobile phones in households is about 30 percent at present and further increasing (Blumberg & Luke, 2011).

---

[1] Marek Fuchs is professor for social science research methods at Darmstadt University of Technology, Marktplatz 15, 64283 Darmstadt, Germany. Telephone +49.6151.16-70970, fax +49.6151.16-72070, fuchs@ifs.tu-darmstadt.de.
Support provided by Deutsche Forschungsgemeinschaft (German Research Foundation) for the "Experimental Mobile Phone Panel" is gratefully acknowledged (Grant # FU 389/16-1 and # FU 389/16-2).

The backdrop of landline penetration rates in most industrialized countries has caused coverage error of traditional landline telephone surveys resulting in severe coverage biases. Since the cell phone only population differs in certain socio-demographic characteristics from people who still have a landline telephone, fixed-line surveys are prone to coverage biases due to the underrepresentation of specific groups (Blumberg & Luke, 2009; Busse & Fuchs, 2011; Peytchev, Carley-Baxter & Black, 2010). In Europe and the United States, cell phone onlys tend to be younger and more often male, display a higher probability of being unmarried and tend to have lower incomes (Blumberg & Luke, 2011; Graeske & Kunz, 2009; Hu, Balluz, Battaglia & Frankel, 2011; Peytchev et al., 2010; Zuwallack, 2009). Consequently, pure landline telephone surveys – even when using RDD or related sampling methods – are no longer assumed sufficiently inclusive. This coverage bias has brought about the use of mobile phones in survey research. Considering the differences between sub-populations with and without landline telephone access, the need for dual frame surveys is seen as a prerequisite unavoidable to prevent potential coverage biases in telephone surveys (Gabler & Ayhan, 2007; Hu et al., 2011; Keeter, Dimock & Christian, 2010; Kennedy, 2007).

The methodological implications of conducting cell phone surveys have received considerable attention. Over the course of past five to ten years, several studies have assessed the pros and cons of mobile phone surveys using the terminology and conceptual approach of the total survey error design (Groves et al, 2009). Amongst others, coverage (Blumberg & Luke, 2009, 2011; Busse & Fuchs, 2012) as well as sampling issues (Wolter, Smith & Blumberg, 2011; Busse & Fuchs, 2010) and measurement error (Vehovar, Berzelak & Manfreda, 2010; Peytchev & Hill, 2010; Lynn & Kaminska, 2010) have been assessed. Since nonresponse is a crucial error component for telephone surveys that has the potential to inflict data quality considerably it is also subject to studies concerning mobile phone surveys (Kennedy, 2010). In the remainder of this paper, we will focus on this component of the total survey error.

Since the early days of survey research, response rates are declining (Brehm, 1994; Steeh, 1981; Hox & de Leeuw, 1994; Harris-Kojetin & Tucker, 1999). In a comprehensive assessment of response rates and its components in surveys conducted in the 1990s by statistical agencies in Europe, North America and Australia, de Leeuw and de Heer (2002) found a considerable average increase in nonresponse rates of roughly 0.5 percent per year. About 0.2 percent was due to increasing non-contacts while an increase of about 0.3 percent per year was attributed to accelerating refusals. Shrinking response rates have raised concerns regarding the accuracy of survey estimates due to nonresponse bias. Even though there is little empirical evidence for a strong correlation of nonresponse and nonresponse bias (Groves & Peytcheva, 2008), large response rates are still assumed to limit maximum nonresponse bias in surveys (Vehovar, 2007).

In the early days of mobile phone surveys, response rates seem to be slightly higher compared to similar landline surveys. This finding has been attributed to a novelty effect of surveys conducted by mobile phones. In recent years, however, the response rates of cross-sectional mobile phone samples have dropped considerably and are nowadays even below the response rates of landline telephone surveys (Steeh, Buskirk & Callegaro, 2007; Brick et al, 2002). In this paper, we will assess whether and to what extent nonresponse in a mobile phone survey causes nonresponse bias. Based on this analysis, we will focus on traditional means of increasing cooperation that help decrease nonresponse and nonresponse bias (incentives and advanced messages). Prior to the analysis of nonresponse and nonresponse bias, we describe a methodology that helps improve the quality of randomly generated mobile phone samples. Since the early days of telephone surveys, random digit dialing has been considered the only feasible method to yield random samples of households and individuals in the general population. Previous research could identify crucial drawbacks of random digit dial samples and has particularly focused on hit rates (Buskirk, Callegaro & Rao, 2010). The proportion of working numbers among all randomly generated telephone numbers is typically low. This proportion is even lower in the mobile phone frame since the proportion of working numbers among all generated numbers is typically smaller—at least in Europe. In addition to the necessary field efforts due to the low hit rate, many of the randomly generated mobile phone numbers remain unknown eligible numbers even after multiple contact attempts. This phenomenon is mostly caused by non-informative provider messages when calling numbers that are not immediately answered by users or voice boxes. Many of those non-informative messages do not provide clear indication whether this number is in fact no longer in use or has not been assigned a user or whether those numbers are actually in use but currently out or reach of the network. Since numbers of unknown eligibility induce uncertainty in the computation of response rates, the reduction of the proportion of numbers of unknown eligibility is in the best interest of survey researchers and helps increase the informative value of response rates.

Methods

In order to promote methodological research and gain inside into the methodological implications of mobile phone surveys as a supplement to traditional landline telephone surveys, the Experimental Mobile Phone Panel (Busse & Fuchs, 2010) has been set up and conducted since 2009. As part of the panel, a sample of about 3,000 cell phone respondents has been recruited at two points in time. Sample members were re-interviewed into subsequent panel waves (see Table 1). The panel was predominantly set up to answer methodological questions.

Table 1. Data collection in the Experimental Mobile Phone Panel

			complete interviews			
			recruitment 2009	refreshment 2011	Total	
1	recruitment	spring 2009	1,451	-	1,451	(8%)[a]
2	wave 1	fall 2009	812	-	812	67%
3	wave 2	spring 2010	304	-	304	67%
4	wave 3	fall 2010	208	-	208	75%
5	refreshment	winter 2010/11	-	1,577	1,577	12%
6	wave 4	spring 2011	192	924	1,116	68%
7	Wave 5	fall 2011	842		842	63%

Note: Response rates according to AAPOR RR1. The rate for wave 1 is based on the field work in the call center of GESIS Mannheim, only. Comparable information for the field work in the call center in Duisburg-Essen is not available.

Using two different call-centers in Germany, an initial sample of 1,451 panelists was recruited using random digit dial mobile phone samples provided by GESIS Mannheim (Gabler & Häder, 2008). Part of the field work was conducted in the call-center of the University of Duisburg-Essen as part of a recruitment effort for a German online access panel. The other half of the sample was recruited in the call-center of GESIS Mannheim. Both field work enterprises yielded response rates below 10 percent according to AAPOR standards. In fall 2009, the first follow-up panel wave was administered in the recruited sample. Wave 1 yielded a response rate of 67 percent and resulted in 812 completes. After panel wave 1, half of the sample (cases that were recruited in the call-center of the University of Duisburg-Essen) had to be dropped from the panel due to the termination of a cooperation contract of the University of Bremen and Darmstadt University of Technology. Consequently, the gross sample for wave 2 was considerably reduced. Panel wave 2 was conducted in spring 2010 and yielded again a response rate of 67 percent providing 304 completes. In fall 2010, panel wave 3 was administered, this time in the call-center of Darmstadt University of Technology. 208 completes could be achieved reflecting a response rate of 75 percent. Since the net sample of the panel was considerably reduced due to the termination of the cooperation contract as well as due to panel attrition, we conducted a refreshment study in winter 2010/2011. During that refreshment survey, 1,577 cases were recruited again using an RDD-like sample following the Gabler-Häder procedure (Gabler & Häder, 2008). This refreshment survey yielded a response rate of 12 percent using AAPOR RR1. In spring 2011, we conducted wave 4 of the panel which yielded 192 cases from the “old” panel and 924 completes from the refreshment sample. Overall, a response rate of 68 percent was achieved. The final panel wave was conducted in fall 2011 yielding 842 completes yielding a response rate of 63 percent.

As part of the Experimental Mobile Phone Panel, two methods were assessed in order to identify non-working numbers prior to field work (pre-call validation methods). We used a so-called home location register lookup (HLR) in order to determine the working status of randomly generated mobile phone numbers and at the same time experimented with return codes of text messages sent to randomly generated mobile phone numbers prior to field work. Both methods provide return codes which can be used by survey researchers in order to determine the working status of a phone number (for details see Kunz & Fuchs, 2012). In the result section of this paper, we report results from a simulation study using the final disposition codes of the field work as a gold standard for an assessment of the methods to determine the working status of a randomly generated number prior to field work. Since the proportion of mobile phone numbers of unknown eligibility has a negative impact on response rates, we consider the two methods a contribution to increase response rates. Even though the two methods do not increase the number of completed interviews, they provide more reliable information concerning all numbers that could not be turned into a completed or partial interview (in particular the proportion of numbers of unknown eligibility is reduced). Accordingly, the resulting response rates have higher informative value compared to similar studies that do not use pre-call validation methods.

Using data from the longitudinal panel, we were able to assess the magnitude of nonresponse on various stages of the data collection process. Given the available socio-demographic information in the Experimental Mobile Phone Panel and using the European Social Survey as a reference study for the mobile phone population in Germany (European Social Survey, 2006), we focused on four socio-demographic variables for the assessment of nonresponse bias and panel attrition bias: gender, age, employment status, and education. In the result section of this paper, we will compare response distributions for the four socio-demographic variables in the refreshment survey as well as in panel waves 4 and 5 to the response distribution in the European Social Survey which is assumed to represent the German mobile phone population.[2]

Even though noncontact is the prime component of overall nonresponse in mobile phone surveys, refusals also contribute to nonresponse and potentially cause nonresponse bias. Accordingly, survey researchers are interested in mechanisms that may help increase cooperation among mobile phone respondents. As part of the Experimental Mobile Phone Panel, we assessed the feasibility and efficiency of two traditional mechanisms to increase cooperation. On the one hand, we sent text messages as advanced letters to mobile phone numbers in order to make potential respondents aware of the upcoming survey call. On the other hand, we experimented with pre-paid cash incentives in order to boost cooperation. As part of panel wave 5 about 1,100 respondents from wave 4 and 300 respondents from prior waves that could not be reached in wave 4 were contacted for a ten minute mobile phone interview. About 200 randomly selected respondents received a 5 Euro give coupon as a code sent to them by text message prior to the first survey call. Respondents could redeem the gift coupon at a website where most of the large-scale online businesses in Germany accept gift coupons. Also, a random sub-sample of about 600 respondents received a text message announcing the upcoming survey call prior to the first contact attempt.

Results

Number validation

In order to identify non-working numbers prior to field work activities, we assess the efficiency of two methods: home location register lookup (HLR) and return codes from bulk text messages sent to the sample mobile phone numbers. As part of the refreshment survey in winter 2010/2011, about 25,000 randomly generated mobile phone numbers were used during field work. About 3,300 of these numbers were randomly assigned to various experimental splits prior to field work. One of the groups (n = 1,100) underwent HLR testing prior to field work. Consequently, we had return codes from the home location register lookup as well as final disposition codes from field work (up to 15 contact attempts). Thus, we were able to assess the efficiency of the home location register lookup as compared to the final disposition codes. The HLR test is a rather cost-efficient method (four Euro cents per number) and can be administered online in a short period of time. In order to assess this method, we computed

[2] The European Social Survey is a face-to-face interview study using area probability sampling. Accordingly, this survey is not prone to the same nonresponse biases as assessed in this paper. Even though other nonresponse biases might have an effect of the response distributions concerning the four socio-demographic variables assessed, we still consider these distributions as a standard for comparison.

working number rates for the experimental sample (including all working numbers leading to a business or to other non-household entities). We first assessed the proportion of working numbers among all numbers in this sub-sample. Table 2 indicates that 57 percent of all numbers were in fact working numbers. We then simulated the proportion of working numbers in the net sample if we had excluded all numbers that could be identified as non-working prior to field work using home location register lookup return codes. We simulated the effect of two different screening conditions: a rather strict screening condition excluding all numbers where the return codes did not indicate that a particular cell phone number was in fact a working number, and a rather loose screening condition where only those numbers were excluded that were definitely not working based on the HLR return codes. Both screening conditions increased the proportion of working numbers significant and yielded a working number rate of about 91 percent (see Table 2, first section).

Table 2. Results from the simulation study concerning pre-call validation methods

			applied screening rules	
	validation method	no screening	(a) loose	(b) strict
working number rate (%)	(a) number validation	56	$91^{*, (2b), (3a), 3b}$	$91^{*, (2b), (3a), 3b}$
	(b) text messaging	54	$91^{*, (3a), 3b}$	$97^{*, \dagger, (1a), (1b)}$
	(c) combination	53	$95^{*, (1a), (1b), (2a)}$	$98^{*, \dagger, 1a, 1b, 2a}$
contact rate (%)	(a) number validation	26	$44^{*, (2a), 2b, 3a, 3b}$	$68^{*, \dagger, 2a, 2b, 3a, 3b}$
	(b) text messaging	28	$54^{*, (1a), 1b, 3b}$	$80^{*, \dagger, 1a, 1b, 3a}$
	(c) combination	28	$55^{*, 1a, 1b, 2b}$	$84^{*, \dagger, 1a, 1b, 2a}$
interview rate (%)	(a) number validation	6	$11^{*, 2b, 3b}$	$17^{*, \dagger}$
	(b) text messaging	7	$14^{*, (3b)}$	$22^{*, \dagger, 1a, (3a)}$
	(c) combination	7	$14^{*, (2b)}$	$22^{*, \dagger, 1a, (2a)}$
call attempts[a]	(a) number validation	106	87	43
	(b) text messaging	90	62	33
	(c) combination	100	68	32
overall call duration (minutes)[a]	(a) number validation	71	65	50
	(b) text messaging	63	53	43
	(c) combination	70	58	44

Note. [a] Data on call attempts and overall call duration referred to mean per completed interview. No significance tests were calculated for these two indicators because calculations involved derived values based on the number of completed interviews, and not on the original elements of the sample.
Calculations were based on chi-squared goodness-of-fit tests for dependent samples:
* $p < .001$ compared to the respective 'no screening' condition,
† $p < .001$ compared to the respective loose screening condition.
Calculations were based on chi-squared tests for independent samples:
^{1a} $p < .001$ compared to condition 1a, $^{(1a)}$ $p < .05$ compared to condition 1a,
^{1b} $p < .001$ compared to condition 1b, $^{(1b)}$ $p < .05$ compared to condition 1b,
^{2a} $p < .001$ compared to condition 2a, $^{(2a)}$ $p < .05$ compared to condition 2a,
^{2b} $p < .001$ compared to condition 2b, $^{(2b)}$ $p < .05$ compared to condition 2b,
^{3a} $p < .001$ compared to condition 3a, $^{(3a)}$ $p < .05$ compared to condition 3a,
^{3b} $p < .001$ compared to condition 3b, $^{(3b)}$ $p < .05$ compared to condition 3b.

In addition to the home location register lookup, we also employed bulk text messaging services in order to obtain return codes for a random sub-sample of telephone numbers used during field work (n = 1,100). Again, we had two screening conditions: a rather strict screening condition where only those numbers were included that were identified working numbers based on the available return codes, while in the loose screening condition all numbers were included that were not definitely non-working. Again, we used a simulation approach where all numbers were put into the field. About 54 percent of this sub-sample was identified as working numbers based on the final disposition codes from field work. When simulating the effect of pre-call validation screening (by excluding all numbers that would have been excluded from field work if pre-call validation methods had been applied), the proportion of working numbers increased to 91 percent in the loose screening condition and yielded even higher values of 97 percent in the strict screening condition (both values are significantly different from the control group).

In a third experimental group (n = 1,100), we applied a combination of home location register lookup and text messaging return codes. Interestingly, the combined application of both methods yielded only a slight increase of the working number rate to 95 percent in the loose screening condition and 98 percent in the strict screening condition.

When looking at the contact rates among the experimental sub-samples, we again observed considerable increases in the loose and strict screening conditions for both pre-call validation methods as well as for the combination of the two methods. In all cases the strict screening condition yielded significantly higher contact rates compared to the loose screening condition. Nevertheless, even the loose screening condition yielded noticeable higher contact rates compared to the control condition. When using the home location register lookup, contact rates could be increased to up to 68 percent; when using text messaging return codes, the contact rate yielded 80 percent. The combination of the two methods brought about a contact rate of about 84 percent.

These considerable increases in working number rates and contact rates transformed into an increase of the interview rates as well. When applying the home location register lookup in the loose screening condition, the interview rate increased to 11 percent (compared to 6 percent in the control condition). When applying the strict screening condition in the home location register lookup condition, the interview rate increased to 17 percent (all differences statistically significant). In the text messaging return code condition, the increases in interview rates were even more pronounced: here, in the loose condition, the interview rate increased to 14 percent, and in the strict screening condition the interview rate was 22 percent. Interestingly, the combination of the home location register lookup method and the text messaging return code method did not improve the interview rate above the sole application of the text messaging return methodology.

As a result of the increases in working number rates, contact rates and interview rates, survey efforts during field work could be reduced. We computed the average number of contact attempts per complete as well as the average call duration (including all call attempts) in minutes per complete. Both values decrease considerably for the two pre-call validation methods in both screening conditions. More pronounced cost-savings could be achieved in the strict screening conditions. Here, the average number of contact attempts dropped to about 40 percent compared to the control condition. Concerning the average duration per complete, cost savings were also in the range of 30 to 40 percent. Further details of the two pre-call validation methods can be found in a paper by Kunz and Fuchs (2011).

Even though the positive effects of the two pre-call validation methods in terms of survey cost are quite convincing, survey researchers have to be careful when adopting these techniques to their surveys. At this point in time it is not clear whether all excluded numbers are in fact non-working numbers. In order to estimate potential biases introduced by the application of the two pre-call validation methods, we assessed screening biases using the basic question technique proposed by Kersten and Bethlehem (1984). Interviewers were advised to estimate socio-demographic variables for all members of the gross sample at the beginning of the interview. Also, all respondents who initially refused to take part in the survey interview were asked to answer at least two basic questions: concerning employment status and education. We then were able to compare the socio-demographic variables in the gross sample to the distribution of these variables in the net sample. Given our simulation approach, we had either answers provided by respondents or interviewer estimates of the socio-demographic variables for all working mobile phone numbers that would have been excluded if the pre-call validation methods had been applied prior to field work (false negatives). Accordingly, we are able to simulate the screening bias for age group and gender (employment status and education were prone to nonresponse and could not be estimated by interviewers).

Table 3. False negative cases due to the pre-call validation methods and resulting screening bias

screening condition	false negatives % (N)	bias due to false negatives: age		gender	
HLR test					
loose screening condition	4 (22)	0.6	n.s.	0.0	n.s.
strict screening condition	41 (232)	0.4	n.s.	0.9	n.s.
text message return code					
loose screening condition	17 (106)	0.6	n.s.	0.3	n.s.
strict screening condition	46 (281)	0.4	n.s.	0.6	n.s.
combination of both methods					
loose screening condition	17 (103)	0.5	n.s.	1.3	n.s.
strict screening condition	52 (311)	1.0	n.s.	2.9	n.s.

Note. n.s. = not significant

Overall, the methods tested to identify nonworking mobile phone numbers prior to field work differ considerably with respect to the proportion of phone numbers that would have been excluded even though final disposition codes from field work suggest that these numbers are in fact working numbers. In the loose screening condition, the home location register lookup yielded about 4 percent false negatives; in the text message return code method, the loose screening condition yielded about 17 percent false negatives. In both methods, the strict screening condition produced tremendously higher false negative rates. The home location register lookup yielded 41 percent false negatives and in the text message return code method, the strict screening condition had a false negative rate of 46 percent. The combination of the two methods yielded slightly higher false negative rates. However, only in the strict screening condition, the combination of the two methods had a visibly higher false negative rate than each of the single methods (52 percent). Accordingly, all methods involved the risk of excluding working numbers from the sample. This risk was lowest in the home location register lookup using the loose screening condition which excluded only those numbers that were definitely not working according to the return codes from the home location register lookup. However, even using the loose screening condition in the home location register lookup, 4 percent of all excluded numbers were false negatives. In the strict screening conditions of the home location register lookup and the text messaging return code method more than four in ten numbers were excluded even though final disposition codes from field work after 15 call attempts indicate that these numbers might in fact be working numbers.

With respect to the magnitude of the resulting screening bias (see Table 3) it is safe to say that the potential impact of the screening bias on the estimates for age and gender are rather small. Analyses reveal only small and non-significant screening biases for both variables and both pre-call validation methods. Only in the combined application of home location register lookups and text messaging return codes the screening bias for gender seems to be slightly larger compared to the sole application of one of the pre-call validation methods. Nevertheless, even those larger biases remain on a non-significant level. Even though the small-scale scope of the present simulation study limits the generalizability of results, we found no indication of significant screening biases due to the false negatives in the two pre-call validation methods.

Nonresponse bias

In order to assess the potential nonresponse bias due to nonresponse and panel attrition, we compared response distributions for four socio-demographic variables: gender, age group, employment status and highest level of education. We used data from the refreshment study conducted in winter 2010/2011 and compared response distributions to population data taken from the European Social Survey (restricted to the respondents who own a mobile phone). Results indicate considerable nonresponse biases introduced into the refreshment survey. The proportion of male respondents is considerably higher compared to the mobile phone population. Panel attrition in waves 4 and 5 of the panel further increased the proportion of male respondents, suggesting that panel attrition bias and nonresponse bias shifted the composition of the net sample into the same direction. With respect to age group, it is interesting to note that the initial nonresponse bias in the refreshment survey overrepresented younger respondents compared to middle-age and older respondents. However, panel attrition seems to reduce this overrepresentation of younger respondents to the benefit of middle-age respondents. The initial nonresponse bias with respect to the elderly respondents could not be reduced by means of attrition bias.

Table 4. Distribution of socio-demographic variables in the refreshment survey 2010/11 and in subsequent panel waves 4 and 5 compared to the target population

	mobile phone users in Germany	respondents in the refreshment survey 2010/11	wave 4	wave 5
	Gender			
Male	52%	59%	60%	62%
Female	48%	41%	40%	38%
	age group			
< 40	37%	45%	40%	37%
40 – 60	43%	43%	48%	52%
> 60	20%	12%	12%	11%
	employment status			
Working	60%	76%	78%	79%
not working	40%	24%	22%	21%
	highest level of education			
Student	4%	1%	1%	1%
no degree	1%	1%	0%	0%
primary school	26%	17%	17%	17%
secondary school (middle track)	36%	34%	34%	33%
higher education entrance certificate	33%	47%	49%	49%

Note. Population data taken from the Germany sample of the European Social Survey 2008 (respondents who own a mobile phone).

With respect to employment status, the initial nonresponse bias which overrepresented working respondents considerably was further increased by means of panel attrition. In panel wave 5, the proportion of non-working respondents was only 21 percent compared to about 40 percent in the population. Finally, we assessed the highest level of education for respondents in the initial refreshment study as well as in panel wave 4 and 5. Results indicate that respondents holding a higher education entrance qualification were overrepresented among respondents in the refreshment survey indicating nonresponse bias. The proportion of respondents with a higher education entrance certificate was even further increased in waves 4 and 5. Consequently, the initial nonresponse bias with respect to education was not reduced by means of panel attrition.

Incentives and text message announcements

Overall, the response rate in wave 5 reached 72 percent among respondents who have been contacted in wave 4 (see Table 5, lower section, column denoted as control group). As expected, the response rate among panel members that had not been reached in wave 4 was rather low. Only 15 percent of respondents who last participated prior to wave 4 completed the survey in wave 5. However, it is noteworthy that panel attrition is also substantial in the group of respondents who at the end of the wave 4 interview agreed to be re-contacted for panel wave 5.

Table 5. Contact rates, cooperation rates and response rates of the control group compared to experimental text message announcement group and the incentive group

	control group	text message announcement	incentive
	contact rate		
last interview prior to wave 4 (n=301)	51%	45%	40%
last interview in wave 4 (n=1,087)	89%	88%	91%
total wave 5 (n=1,388)	81%	79%	80%
	cooperation rate		
last interview prior to wave 4 (n=301)	30%	24%	32%
last interview in wave 4 (n=1,087)	81%	84%	87% **
total wave 5 (n=1,388)	74%	76%	80% *
	response rate		
last interview prior to wave 4 (n=301)	15%	11%	13%
last interview in wave 4 (n=1,087)	72%	74%	80% **
total wave 5 (n=1,388)	60%	60%	64% *

Note. * < .05; ** < .01. Chi² test of independence compared to the control group. Contact rates, cooperation rates and response rates according to AAAPOR (2011).

Results indicate that incentives have the potential to increase response rates compared to the non-treatment control group. The response rate in the incentive group was 4 percentage points higher than in the control group ($p < .05$). This effect occurred for respondents who participated in wave 4 (8 percentage points); for respondents who last participated prior wave 4, no effect of the incentive occurred. As expected, the positive effect of the prepaid incentive on the response rate was predominantly due to higher cooperation rates among respondents who received the incentive. The increase of about 6 percentage points for the cooperation rate was not reflected in the contact rate. Even though we have to admit that all respondents had participated in prior panel waves and thus response rates cannot directly be compared to other cross-sectional mobile phone surveys, we are confident that the positive effect of incentives can be generalized.

By contrast, a simple text message announcing the upcoming survey call has no positive effect on response rates. Also, no significant increase of cooperation rates was to be noticed. Interestingly, neither incentives nor advanced text messages had the potential to increase contact rates. This came as a surprise since it was expected that either incentives or advanced messages would increase the willingness of respondents to answer the incoming survey call.

Mobile web alternative mode for nonrespondents

For mobile phone surveys as for traditional landline telephone surveys, a key problem during field work arises from the fact that the respondents' willingness to cooperate is fundamentally bound to the situation that they are in when reached by the interviewer. Due to the synchronized character of the data collection process in telephone interviews, nonresponse is to a large extent bound to the timing of survey calls. By contrast, self-administered modes offer respondents the opportunity to answer a survey at a time that is convenient for them. Accordingly, the experimental mobile phone panel was also concerned with self-administered components in a mobile phone survey environment. Given the increasing penetration of mobile web devices, we used text messages sent to non-contacted cases and refusals after field work to ask respondents to complete a short version of our core questionnaire using mobile web technology. In total, 444 respondents that could not be reached during field work or who refused to cooperate in panel wave 5 (soft refusals only) took part in this experiment. Each nonresponding panel member received a text message with a URL to a short mobile web questionnaire containing core questions from the wave 5 questionnaire. In order to administer the mobile web survey we used Globalpark panel survey software including the Mobile Extension.

Table 6. Results of the mobile web follow-up study of nonrespondents in wave 5

disposition code after wave 5	text message received	online questionnaire accessed	questionnaire completed	consent to be contacted in wave 6
soft refusal	37% (166)	59% (13)	67% (12)	67% (6)
non-contact	27% (119)	27% (6)	17% (3)	11% (1)
other nonresponse	1% (5)	-	-	-
unknown eligibility	35% (154)	14% (3)	17% (3)	22% (2)
total	100% (444)	100% (22)	100% (18)	100% (9)

Only about 5% (22 individuals) of invited respondents actually accessed the mobile web questionnaire. Of these, only 18 persons have completed the core mobile web questionnaire. Interestingly, we gained considerable cooperation among those respondent who had refused during regular field work of panel wave 5. In this group, we could convince about 8 percent of respondents to complete the mobile web questionnaire. Half of this group also agreed to be recontacted in panel wave 6. Overall using the mobile web alternative mode, the response rate of panel wave 5 could be increased by two percentage points. Even though results are not fully promising, we gained first inside in the methodological implications of conducting a mobile web survey among nonrespondents in an ongoing panel.

Summary and Discussion

Rising mobile only rates in Germany as well as the increasing usage of mobile communication devices in the general population cause serious challenges for landline phone surveys. Coverage error, sampling methods, differential measurement error as well as potentially protective usage patterns and negative attitudes to mobile phone surveys leading to nonresponse are concerns of increasing importance for overall survey quality. In order to explore the potentials of conducting mobile phone surveys, the Experimental Mobile Phone Panel was set up in 2009 predominantly to assess methodological questions. In this paper, we report selected results concerning nonresponse in mobile phone surveys.

Since hit rates in samples of randomly generated mobile phone numbers are traditionally rather low, we experimented with two methods for pre-call validation of randomly generated mobile phone numbers. We applied (1) home location register lookups and assessed (2) return codes of bulk text messages sent to mobile phone numbers. Both methods seem productive and efficient in order to reduce the proportion of non-working numbers put into the field. Working number rates as well as contact rates and interview rates could be increased considerably and at the same time the number of call attempts per complete and the overall call duration per complete could be reduced. However, both methods imply the risk of false negatives (i.e. numbers that were excluded from field work even though they are working numbers), and consequently a screening bias arises. Further research is necessary in order to determine the magnitude of potential screening bias when applying to pre-call validation methods.

As expected, mobile phone surveys were prone to nonresponse bias similar to traditional landline telephone surveys. The proportion of male respondents who cooperate was considerably higher, and also the proportion of young respondents in the net sample was increased compared to population data. The same was true for employment status: Working respondents were overrepresented, and also people holding a higher education entrance certificate were overrepresented in the net sample. Interestingly, most of these biases were further increased by panel attrition with the exemption of age. Here, the proportion of younger respondents in the net sample decreased from the recruitment survey to panel wave 5. For age, employment status and education, the initial nonresponse bias remained about the same or was further increased by means of panel attrition.

In order to reduce nonresponse and panel attrition, we employed two methods that have proven effective in traditional landline telephone surveys. We provided a five Euro gift coupon as an incentive to a random sub-sample of panel wave 5 respondents and also sent advanced text messages announcing the upcoming survey call to another random sub-sample of respondents. Interestingly, advanced text messages seemed to have no productive effect on contact rates, cooperation rates and response rates. By contrast, incentives have the potential to increase cooperation and consequently increase response rates. No effect of incentives on contact rates was noticed. Results indicate that cooperation rates and response rates were significantly increased in the incentive group compared to the control group.

Finally, we assessed a mobile web alternative mode for nonrespondents. Results indicated that only a small proportion of nonrespondents were willing to answer a mobile web questionnaire with core questions from the telephone survey questionnaire. Nevertheless, the mobile web alternative was able to boost the response rate of panel wave 5 by 2 percentage points.

References

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

Blumberg, S. J., & Luke, J. V. (2009). Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2009. Hyattsville, MD: National Center for Health Statistics.

Blumberg, S. J., & Luke, J. V. (2011). Wireless substitution: early release of estimates from the National Health Interview Survey, January–June 2011. Washington: U.S. Department of Education, Office of Educational Research and Improvement.

Brehm, J. (1994). Stubbing our toes for a foot in the door? Prior contact, incentives and survey response. International Journal of Public Opinion Research, 6(1), 45-63.

Brick, J. M., Brick, P. D., Dipko, S., Presser, S., Tucker, C., & Yuan, Y. (2007). Cell phone survey feasibility in the U.S.: Sampling and calling cell numbers versus landline numbers. Public Opinion Quarterly, 71(1), 23-39.

Buskirk, T. D., Callegaro, M., & Rao, K. (2010). "N the network?" - Using internet resources for predicting cell phone number status. Social Science Computer Review, 28(3), 271-286.

Busse, B., & Fuchs, M. (2011). One mobile phone = one person? – Findings from two studies addressing consequences of cell phone sharing for sampling in mobile phone surveys. Paper presented at the NTTS Conferences on New Techniques and Technologies for Statistics.

Busse, B., & Fuchs, M. (2012). The components of landline telephone survey coverage bias. The relative importance of no-phone and mobile-only populations. Quality & Quantity, 46(4), 1209-1225.

European Social Survey (2006) Documentation Report. The ESS Data Archive Edition 3.2.

Fuchs, M. (2008). Total survey error. In P. J. Lavrakas (Ed.), Encyclopedia of survey research methods (Vol. 2, pp. 896-902). Thousand Oaks, CA: Sage.

Gabler, S., & Ayhan, Ö. (2007). Gewichtung bei der Erhebung im Festnetz und über Mobilfunk. Ein Dual Frame Ansatz. ZUMA-Nachrichten Spezial(13), 39-46.

Gabler, S., & Häder, S. (2009). Die Kombination von Mobilfunk- und Festnetzstichproben in Deutschland. In M. Weichbold, J. Bacher & C. Wolf (Eds.), Umfrageforschung. Herausforderungen und Grenzen (pp. 239-252). Meppel: Krips b.v.

Graeske, J., & Kunz, T. (2009). Stichprobenqualität der CELLA-Studie unter besonderer Berücksichtigung der Mobile-Onlys. In M. Häder & S. Häder (Eds.), Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung eienr Strategie zur Datenerhebung (pp. 57-70). Wiesbaden: VS-Verlag für Sozialwissenschaften.

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

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). Survey Methodology (2 ed.). Hoboken, New Jersey: Wiley.

Harris-Kojetin, B. A., & Tucker, C. (1998). Longitudinal nonresponse in the current population survey (CPS). ZUMA-Nachrichten Spezial, 4, 263-272.

Hox, J. J., & De Leeuw, E. (1994). A comparison of nonresponse in mail, telephone, and face-to-face surveys. Quality & Quantity, 28(4), 329-344.

Hu, S. S., Balluz, L., Battaglia, M. P., & Frankel, M. R. (2011). Improving public health surveillance using a dual-frame survey of landline and cell phone numbers. American Journal of Epidemiology, 173(6), 703-711.

Keeter, S., Dimock, M., & Christian, L. (2010). Republican vote share bigger in landline-only surveys. The growing gap between landline and dual fram election polls. Washington, D.C.: The Pew Research Center for the People and the Press.

Kennedy, C. (2007). Evaluating the effects of screening for telephone service in dual frame RDD surveys. Public Opinion Quarterly, 71(5), 750-771.

Kennedy, C. K. (2010). Nonresponse and measurement error in mobile phone surveys. University of Michigan, Michigan.

Kersten, H. M. P., & Bethlehem, J. G. (1984). Exploring and reducing the nonresponse bias by asking the basic question. Statistical Journal of the United Nations ECE 2, 2, 369-380.

Kunz, T., & Fuchs, M. (2011). Pre-call validation of RDD cell phone numbers. A field experiment. Paper presented at the Annual Conference of the American Association for Public Opinion Research.

de Leeuw, E. D., & De Heer, W. (2002). Trends in household survey nonresponse: A longitudinal and international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge & R. J. A. Little (Eds.), Survey nonresponse (pp. 41-54). New York: Wiley.

Lin, I.-F., & Schaeffer, N. C. (1995). Using survey participants to estimate the impact of nonparticipation. Public Opinion Quarterly, 59(2), 236-258.

Lynn, P., & Kaminska, O. (2010). The impact of mobile phones on survey measurement error. Paper presented at the Mobile Research Conference 2010.

Peytchev, A., Carley-Baxter, L. R., & Black, M. C. (2010). Coverage bias in variances, associations, and total error from exclusion of the cell phone-only population in the United States. Social Science Computer Review, 28(3), 287-302.

Peytchev, A., & Hill, C. A. (2010). Experiments in mobile web survey design. Similarities to other modes and unique considerations. Social Science Computer Review, 28(3), 319-335.

Steeh, C. G. (1981). Trends in nonresponse rates, 1952-1979. Public Opinion Quarterly, 45(1), 40-57.

Steeh, C., Buskirk, T. D., & Callegaro, M. (2007). Using text messages in U.S. mobile phone surveys. Field Methods, 19(1), 59-75.

Vehovar, V. (2007). Non-response bias in the European Social Survey. In G. Loosveldt, M. Swyngedouw & B. Cambré (Eds.), Measuring meaningful data in social research (pp. 335-356). Leuven: Acco.

Vehovar, V., Berzelak, N., & Manfreda, K. L. (2010). Mobile phones in an environment of competing survey modes: Applying metric for evaluation of costs and errors. Social Science Computer Review, 28(3), 303-318.

Wolter, K. M., Smith, P., & Blumberg, S. J. (2010). Statistical foundations of cell-phone surveys. Survey Methodology, 36(2), 203-215.

Zuwallack, R. (2009). Piloting data collection via cell phones: results, experiences, and lessons learned. Field Methods, 21(4), 388-406.