

Introducing Adaptive Design Elements in the Panel Study “Labour Market and Social Security” (PASS)

Mark Trappmann, Gerrit Müller¹

Abstract

This article gives an overview of adaptive design elements introduced to the PASS panel survey in waves four to seven. The main focus is on experimental interventions in later phases of the fieldwork. These interventions aim at balancing the sample by prioritizing low-propensity sample members. In wave 7, interviewers received a double premium for completion of interviews with low-propensity cases in the final phase of the fieldwork. This premium was restricted to a random half of the cases with low estimated response propensity and no final status after four months of prior fieldwork. This incentive was effective in increasing interviewer effort, however, led to no significant increase in response rates.

Key Words: adaptive design; panel surveys; prioritizing low-propensity cases.

1. Introduction: The panel survey “Labour Market and Social Security” (PASS)

The panel study “Labour Market and Social Security“ (PASS) (Trappmann et al. 2010, 2013) is one of the major German panel surveys. PASS is conducted by the Institute for Employment Research in co-operation with the fieldwork agency infas. It focuses on unemployment and poverty dynamics and on the life situation of poor households and benefit recipients. Since 2007 about 15.000 persons in about 10.000 households have been interviewed each year. In each household, a household interview with the head of the household is followed by a person interview with each household member aged 15 or above. Respondents are drawn from two different sampling frames. About half of the households are sampled from national registers of unemployment benefit recipients. The wave one sample target population was households in receipt in July 2006. This sample is refreshed each year by new entries into benefit receipt (subsample A). The other half of households is from a general population sample (subsample B). There are three sets of cross-sectional weights projecting subsample A to all households who received benefits on at least one of the reference dates, projecting subsample B to the German general population and projecting the combined sample to the German general population. This flexible sampling design allows for comparisons between samples as well as for integrated analyses of the combined dataset. For inference to longitudinal populations these cross-sectional weights can be multiplied with estimated probabilities to stay in the study in subsequent waves.

PASS uses a sequential mixed-mode design of computer-assisted personal interviewing (CAPI) and computer-assisted telephone interview (CATI) in order to maximize response under cost-restrictions. In the beginning, due to budgetary constraints, the default mode was CATI. Households for whom no working telephone number was available, who could not be contacted after a pre-defined number of attempts, or who requested a personal interview, were followed-up by a CAPI interviewer. From wave 2 on, households were first approached by the mode of the previous wave interview. Again, those who could not be contacted in the initial mode or who preferred a different mode were switched. Different modes occur even within a household. In addition, CAPI cases were transferred to CATI for refusal conversion. From wave 4 on the budget for the study was increased. All refreshment cases are first approached in CAPI, while the procedure for panel cases remains unchanged. As of 2013, about half of the interviews were in CAPI and the other half in CATI. PASS uses a unimode design to minimize mode-effects on measurement.

¹Mark Trappmann, Institute for Employment Research, Regensburger Str. 104, D-90478 Nuremberg, Germany, mark.trappmann@iab.de, Universität Bamberg, Feldkirchenstr. 21, 96052 Bamberg, Germany; Gerrit Müller, Institute for Employment Research, Regensburger Str. 104, D-90478 Nuremberg, Germany, gerrit.mueller@iab.de

PASS data can be linked to administrative records on mostly complete employment and unemployment histories for all respondents who provide informed consent. While the main purpose of this linkage is enhancing the dataset with detailed information that cannot be collected in the survey interview (e.g. on duration of each employment spell including exact information on wage, occupation and firm or on active labour market program participation), this linkage has been a valuable source for investigations into measurement error. Furthermore, PASS paradata can be linked to the gross sample of recipients (subsample A) enabling research on nonresponse and linkage consent. Due to these advantages PASS has extensively been used by survey methodologists in the past (e.g. Kreuter et al. 2010, 2013, West and Little 2013, West and Groves 2013, Sinibaldi et al. 2014). Detailed documentation on PASS and information on data access can be found on the website of our Research Data Center (http://fdz.iab.de/en/FDZ_Individual_Data/PASS/Working_Tools.aspx) which offers further locations in the US at Michigan University, Cornell University and Berkeley.

2. PASS paradata

Since Wave 4, detailed paradata have been available on a biweekly basis during fieldwork from CATI as well as CAPI. They include time, interviewer number and detailed disposition codes of each contact attempt. These can be linked to previous wave interview information for panel households and to information from the sampling frame for refreshment cases. From the beginning, these paradata have been used to monitor fieldwork and to intervene in case that field progress is not as expected.

Standardized plots have been developed and are automatically generated for each new paradata set. They show how indicators of volume and quality of fieldwork evolve over time by different groups of target households. This can be visualized relative to the previous wave. Figure 2-1. shows plots for percentage point differences in four different indicators when compared to the same week in the previous wave. The Figure displays contact rates ('CON'), cooperation rates ('COOP_MAX'), response rates ('RR_MIN') and rate of appointments and other codes ('WTA') by five different groups which systematically differ in contactability and cooperativeness: Panel households with previous wave interview ('Panel'), panel households without previous wave interview ('Temp. Dropout'), new households formed by persons who moved out of respondent households in the previous wave ('Split'), households from a refreshment sample ('Refreshment'), and new households formed by persons who moved out of respondent households in the current wave ('Split (current wave)').

Other plots focus on nonresponse bias and show how bias with respect to previous wave attributes (for panel households) or frame variables (for refreshment sample households) changes during fieldwork. Figure 2-2 shows an example. It displays the cumulative proportion of heads of households with the highest educational degree across weeks of fieldwork (red curved line). The blue horizontal line shows the target value, which is the proportion of heads of households with the highest educational degree in the sample. In addition to these single variable plots, plots for different variables are summarized in plots of R-indicators (Schouten et al. 2009) across weeks of fieldwork.

Figure 2-1
Fieldwork indicators compared to previous wave

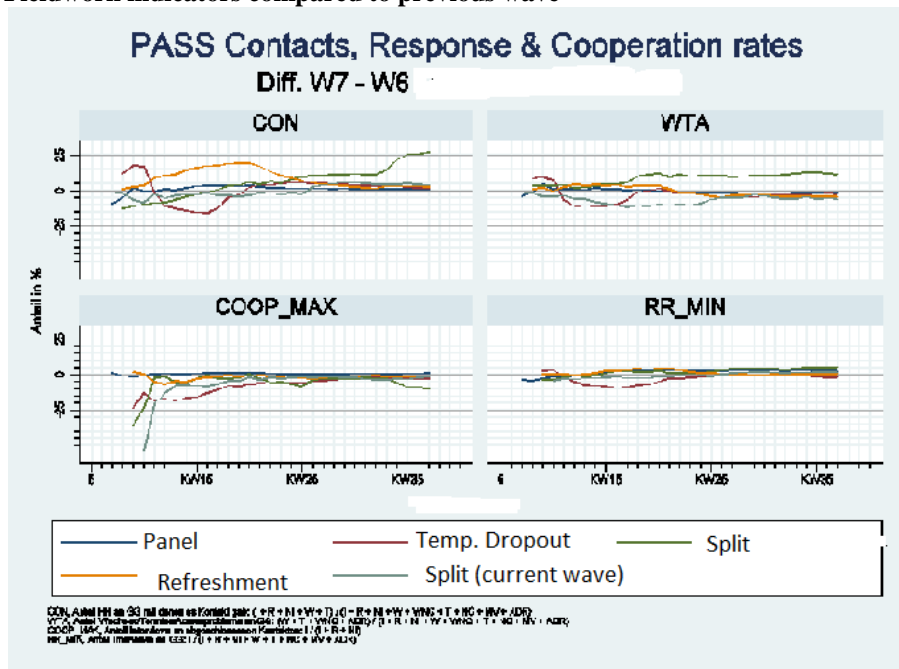
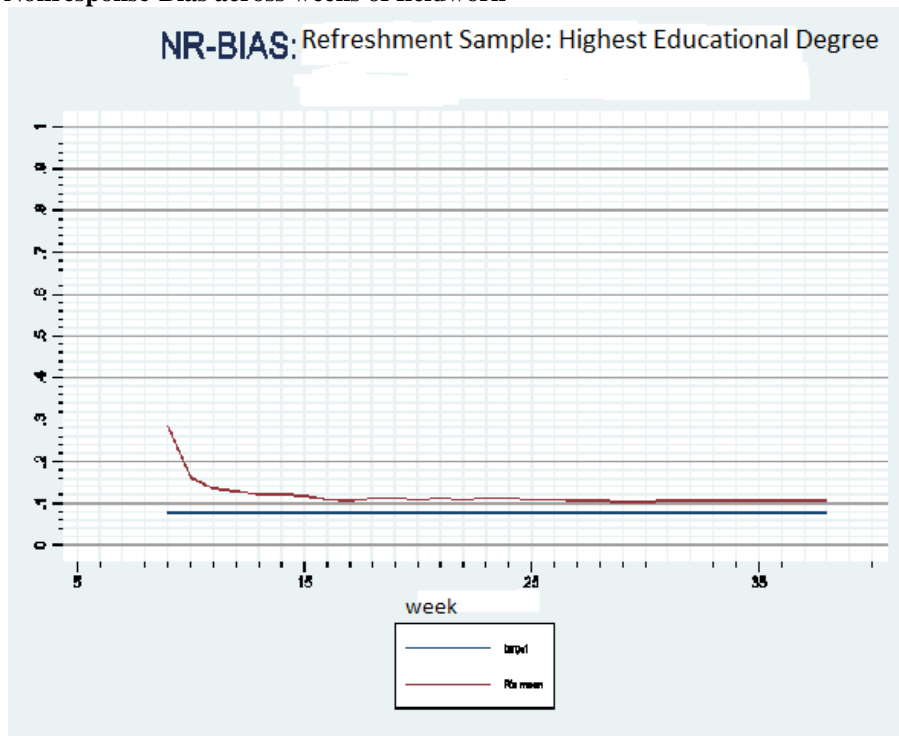


Figure 2-2
Nonresponse-Bias across weeks of fieldwork



3. Adaptive design elements in PASS

Since Wave 6 (2012), fieldwork in PASS is divided into three phases of approximately 2 months each and experiments have been introduced into the fieldwork (primarily in the first phase). They enable the PASS team to implement informed and more systematic interventions in order to optimize fieldwork later in the same wave ('adaptive survey design', Wagner 2008) or in later waves. Past experiments have included respondent incentives for notifying the field work agency of address changes (which showed a significant effect on the number of notifications but not on the effort and final outcome of the fieldwork) and a different Season's Greetings Card for temporary dropouts (emphasizing the unconditional incentive they received compared to the importance of their continued participation for the goals of the study in the regular cards). Both experiments led to no improvements and fieldwork was kept unchanged.

The focus here will thus be on two additional experiments, one concerning optimal contact times for panel cases and the other interviewer incentives for low propensity cases in the refreshment sample.

3.1 Optimal contact times

In the first experiment, implemented in phase one of wave 6, contact times in the CATI part of the study were tailored to the day of the week and time of the day of the interview in the previous wave. The results of this experiment have been published in Kreuter and Müller (2014).

Previous research e.g. by Lipps (2012) has indicated that tailoring contact attempts in panel surveys according to previous wave contact times might be a successful strategy. Non-experimental evidence from the previous two PASS waves seemed to confirm this: Respondents who were by chance first contacted in the same time window² in wave 5 as they were interviewed in wave 4 had a 4.6 percentage points higher response propensity (average marginal effect as estimated from a logistic regression model) in wave 5 on first contact as compared to panel cases who were by chance first contacted at a different time (Kreuter and Müller 2014, 4).

Thus, in phase one of the fieldwork, 80% of the panel households assigned to CATI mode received this treatment, while 20% were scheduled randomly by an autodialler. The treatment was limited to the first three contact attempts which were performed on the exact day of the week and in the time window (before 12:00, 12:01-17:00, after 17:00) of the previous wave interview (Kreuter and Müller 2014, 4f.).

The experiment was evaluated estimating intention to treat as well as local average treatment effects (cf. Angrist and Pischke 2009). Results showed that response at first contact could only marginally and non-significantly be improved. While there was weak evidence that efficiency could be increased – the number of call attempts until first contact was reduced on average by about .36 in the treatment group – detailed analysis showed that this was mainly due to a small number of outliers with a very large number of contact attempts (Kreuter and Müller 2014, 7-8). The authors estimated the potential for savings due to this reduction in number of contact attempts at about 1.800 US\$ for the 6,000 households in the CATI panel sample (Kreuter and Müller 2014, 9).

Nevertheless, the fieldwork agency decided to stick with this scheduling of calls in wave 6 and all subsequent waves. As a side effect, this scheduling proved to be superior to the autodialler in securing that every household in the panel sample receives sufficient call attempts already during the first week of the fieldwork³.

² Time windows were categorized into five categories: Weekdays before 12:00, weekdays 12:01 to 17:00, weekdays after 17:00, weekends before 12:00, weekends after 12:00.

³ The autodialler tended to get jammed⁴ due to the callback rules which include a callback on the next day when no one answers, a callback within a quarter of an hour when the line is busy and generally gives priority to appointments over cases that have not been worked yet. In this way some panel households did not receive any contact attempts for weeks.

3.2 Prioritizing low-propensity sample members

The second experiment was first implemented in the third phase of wave 7. It aims at balancing the refreshment sample of the study (new entries to unemployment benefit recipients) with respect to variables available on the sampling frame (national register of recipients): age, sex, education, highest school degree, nationality, household type, household size, state and regional agency type. These cases are approached by CAPI interviewers first. So the experiment was limited to CAPI.

First experiences with prioritizing low-propensity sample members in PASS had been made in wave 6. There, response propensities were estimated for all CAPI cases during fieldwork based on these frame data. In the last phase of data collection (phase 3), the interviewers' premiums for successful interviews were doubled for completing cases with a low predicted response propensity in the bottom half⁴. In wave 6 this led to equal response rates in the final phase for low and high propensity cases (Müller and Trappmann 2013). These results are inconclusive as they might be the results of a positive effect of the increased interviewer incentive as well as of an unpredictable propensity model.

This led to the experiment in wave 7 where only half of the low-propensity cases were assigned to the interviewer incentive while the other half served as control group. Interviewers were not aware that these control group cases were any different from high propensity cases. They were told that the incentive was paid because given the information on the sampling frame we were more interested in these cases than in others. Fieldwork was even more progressed at the beginning of wave 3 than in the previous wave. Only 731 cases were left, 365 of which were low propensity cases. These were equally split between treatment (182) and control (183) group.

Table 3.2-1
Results from wave 7 prioritizing experiment

Final Code	Low Propensity Treatment		Low Propensity Control		High Propensity	
	n	%	n	%	n	%
Interview	22	12.1	20	10.9	70	19.1
Refusal	70	38.5	50	27.3	104	28.4
Appointment	35	19.2	40	21.9	82	22.4
Noncontact	29	15.9	40	21.9	59	16.1
Address problem	18	9.9	22	12.0	32	8.7
Other	8	4.4	11	6.0	19	5.2
Total	182	100.0	183	100.0	366	100.0
Mean (attempts)		1.85		.88		1.11

Results are presented in Table 3.2-1. Note that all results refer to those difficult cases without a final disposition code after four months of fieldwork. Within this group, high propensity cases reach a significantly higher response rate than low propensity cases (one-sided, $p=.026$)⁵ while among the latter there is no difference between treatment and control group. While prioritizing low propensity cases does not seem to have an effect on the likelihood of an interview, there is evidence that interviewers try harder on these cases to reach a final result. The sum of interviews

⁴ Only cases with no final status at the beginning of phase 3 and without known address problems were included. This amounted to $n=914$ out of which 455 were low propensity.

⁵ All significance tests are one-sided and adjusted for clustering within interviewers.

and refusals is significantly larger for the prioritized low-propensity cases than for the control group (one-sided, $p=.027$), implying that more cases in the control group have codes indicating an open state (appointment, noncontact, address problem) by the end of fieldwork. Furthermore the number of contact attempts in the final phase is significantly larger in the treatment group than in the control group (one-sided, $p=.004$).

4. Plans for the future: Developing the PASS adaptive design

The prioritizing experiment indicates that interviewers try harder on low propensity cases, but in a very difficult group (the bottom half of those still open after four months) are not more successful at gaining cooperation. As a consequence, in wave 8, we repeated the prioritization experiment at an earlier stage (phases 2 and 3, phase 2 starting on April 28th, 2014) giving interviewers more time and supposedly a better sample composition with respect to contactability and cooperativeness. Unfortunately with respect to this experiment, fieldwork progressed even faster than in the two previous waves so that only about 1,000 households in the refreshment sample were without a final state towards the end of phase 1. In wave 8, we complement interviewer incentives with additional prepaid respondent incentives of 10 Euro per household. The two treatments were implemented orthogonally with about 125 cases in each of the following groups: 1) control, 2) only interviewer incentive, 3) only respondent incentive, 4) interviewer and respondent incentive. Results are not available yet.

5. Discussion

One open question is whether prioritization should be extended to panel households. Prioritization of low-propensity cases is a rational strategy in cross-sectional surveys or in initial panel waves, where the goal is not to maximize response rates, but to minimize bias with respect to target variables. In later panel waves, however, maximizing response rates (i.e. panel retention) becomes a primary goal because the statistical power of longitudinal as well as later wave cross-sectional analyses depends on the number of cases who can be re-interviewed. By prioritizing panel cases with a low response propensity resources might be taken off other panel cases.

Another form of extension might be to use the lesson of a prior wave, in order to prioritize cases with low expected response propensity from the beginning. Currently prioritizing sets in late in the fieldwork (about two thirds of the cases already have a final status assigned) and therefore any effect can only be limited. Before implementing this, it would be important to check how stable response propensity model coefficients are across waves.

Then there is the general issue applying to all attempts of balancing the sample: We can only balance on variables that are known for respondents and nonrespondents alike. These very same variables could be used for postsurvey adjustments. Bias after adjustment should be the same whether responsive design is applied or not. Therefore it all boils down to the question whether the balancing reduces the variance of weights and thus effective sample size enough to justify the costs.

References

- Angrist, J. Pischke, J. (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.
- Kreuter, F., Müller, G. (2014), "A Note on Improving Process Efficiency in Panel Surveys with Paradata", *Field Methods*, DOI: 10.1177/1525822X14538205.
- Kreuter, F., Müller, G., and Trappmann, M. (2010), "Nonresponse and Measurement Error in Employment Research: Making use of Administrative Data", *Public Opinion Quarterly*, 74, pp. 880–906.
- Kreuter, F., Müller, G., and Trappmann, M. (2014), "A Note on Mechanisms Leading to Lower Data Quality of Late or Reluctant Respondents", *Sociological Methods and Research*, 43, pp. 452-464.

- Lipps, O. (2012), "A Note on Improving Contact Times in Panel Surveys", *Field Methods*, 24, pp. 95-111.
- Müller, G., and Trappmann, M. (2013), "Einführung von Responsive-Design-Elementen im PASS und erste Ergebnisse", paper presented at Gemeinsame Tagung der ASI und der Methodensektion in der DGS: Nonresponse Bias: Qualitätssicherung sozialwissenschaftlicher Umfragen, Berlin, Germany.
- Rosen, J., Murphy, J., Peytchev, A., Riley, S., Lindblad, M. (2011), "The Effects of Differential Interviewer Incentives on a Field Data Collection Effort", *Field Methods*, 23, pp. 24-36.
- Rosen, J., Murphy, J., Peytchev, A., Holder, T., Dever, J.A., Herget, D.R., and Pratt, D.J. (2014), "Prioritizing Low-Propensity Sample Members in a Survey: Implications for Nonresponse Bias", *Survey Practice*, 7.
- Sinibaldi, J., Trappmann, M., and Kreuter, F. (2014), "Which is the Better Investment for Nonresponse Adjustment: Purchasing Commercial Auxiliary Data or Collecting Interviewer Observations?", *Public Opinion Quarterly*, 78, pp. 440-473.
- Schouten, B., Cobben, F., and Bethlehem, J. (2009), "Indicators for the Representativeness of Survey Response", *Survey Methodology*, 35, pp. 101-113.
- Trappmann, M., Beste, J., Bethmann, A., and Müller, G. (2013), "The PASS Panel Survey after Six Waves", *Journal for Labour Market Research*, 46, pp. 275-281.
- Trappmann, M., Gundert, S., Wenzig, C., and Gebhardt, D. (2010), "PASS – A Household Panel Survey for Research on Unemployment and Poverty", *Schmollers Jahrbuch: Journal of Applied Social Science Studies*, 130, pp. 609-622.
- Wagner, J.R. (2008), *Adaptive Survey Design to Reduce Nonresponse Bias*, Ann Arbor: ProQuest.
- West, B.T., and Groves, R.M. (2013), "A Propensity-Adjusted Interviewer Performance Indicator", *Public Opinion Quarterly*, 77, pp. 352-374.
- West, B.T., and Little, R.J.A. (2013), "Non-Response Adjustment of Survey Estimates Based on Auxiliary Variables Subject to Error", *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 62, pp. 213-231.