The Effect of Events Between Waves on Panel Attrition

Tobias Gramlich (Universität Duisburg-Essen)
Frauke Kreuter (University of Maryland, JPSM)
Alexander Mosthaf (IAB – Institute for Employment Research)
Mark Trappmann (IAB – Institute for Employment Research)

4th User Conference, FDZ der BA im IAB, Nürnberg, April, 8th, 2011

Outline

1. Introduction

2. The PASS Panel

3. Research Questions

4. Research Design

5. Results

6. Future work and discussion

1. Introduction (1)

- Panel attrition can never be completely avoided

- If dropout process is not MCAR estimates for longitudinal populations and cross-sectional populations at later waves will be biased

- Many panel studies use some kind of propensity weighting to correct for attrition bias (e.g. BHPS, ECHP, GSOEP, HILDA, PASS)

  - Usually probability of location/contact and probability of participation given contact are modelled separately (logit / probit / weighting classes)

  - Predictors are usually taken from previous waves (referring to person attributes and interview situation) and from fieldwork of the current wave (Watson and Wooden 2009, Tortora 2009)

- Propensity weighting works if the dropout process is MAR: i.e. conditional on observed covariates it is random

- Adequacy of the MAR assumption can be doubted
- In particular: Changes in important attributes **between waves** could affect the participation decision
  - Heller & Schnell (2000) for health
  - Neukirch (2002) for main activity, working hours, marital status
  - van den Bergh et al. (2006) for employment status
- If this holds: Amount of change might be strongly underestimated (unless variables used in the propensity models are good predictors of change)

2 The PASS Panel (1)

- Designed by the Institute for Employment Research (IAB); fieldwork agency: infas

- Household panel survey for research on unemployment, poverty and the welfare state in Germany

- Two subsamples:

  - (i) recipients of unemployment assistance (UB II)

  - (ii) general population (stratified by status)

- Mixed mode (CATI/CAPI)

- Waves 1 to 3 available via RDC
  http://fdz.iab.de/de/FDZ_Individual_Data.aspx#PASS

2 The PASS Panel (2)
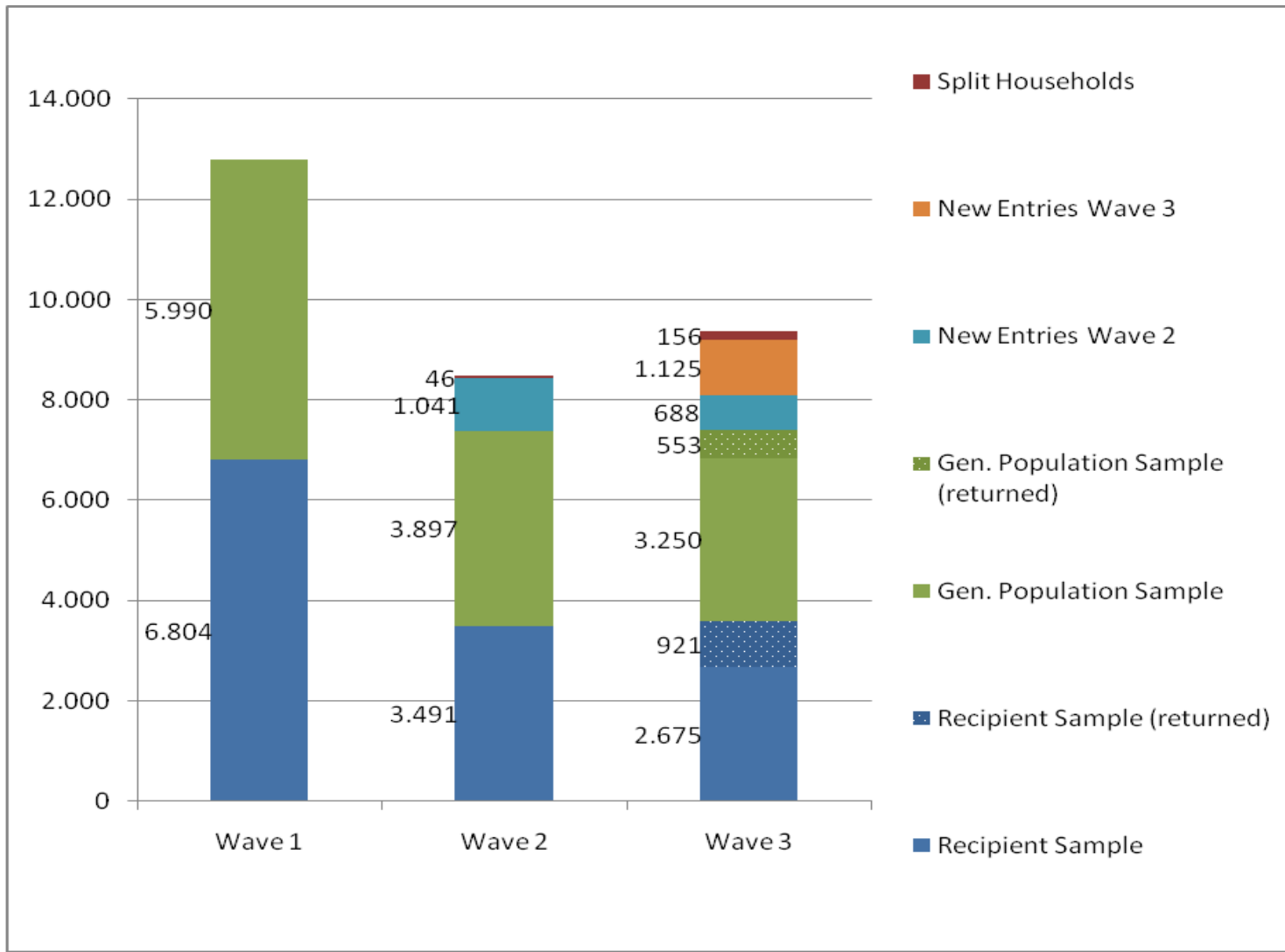
Number of particpating households (waves 1-3)

2 The PASS Panel (3)

- **Administrative data from Federal Employment Agency (BA) can be linked to**
    - the PASS survey data for all respondents who gave consent (about 80% of participants)
    - survey paradata (including participation indicator for each wave) for the whole UB II gross sample (on the household level)
- **Administrative data contain information on:**
    - Employment (BeH), UBII-recipiency, sanctions in UB II,, household composition (all LHG/X-LHG), UB I recipiency (LeH), ALMP participation (MTH), address (various sources), …
    - Focus on attributes where data quality can be considered high: Social security payments are based on this information
    - Usually time-lag of 1-2 years until ALL data are available for research
- **No linked SUF yet**
    - FB E3, KEM, FDZ and Uni Du-E are currently preparing a wave 1 release

2 The PASS Panel (4): Correction of Attrition

- **Propensity weighting**
  - Separate logit models for contact and cooperation (TNS Infratest, Büngeler et al. 2008)
  - Person and household attributes: age, gender, nationality, language, education, working hours, income, UB II recipiency, self rated health, life satisfaction, children/age groups, house ownership (ct)
  - Regional context: state, municipal size
  - Interview situation: mode, length, missing values, participation of other household members, subsample
  - Fieldwork of current wave: number of contact attempts in CATI / CAPI (cp)

3 Research Questions (1)

- To what extent do events between waves (according to admin data) influence contact and co-operation rates in wave 2?
  - Events: person/household moved, separation of couples, change of employment status, change of benefit recipiency status, sanctions
- Does this still lead to biased estimates of the amount of change after the propensity weights of the survey are applied?
- If yes: Can additional variables (usually not used in propensity weighting models) help reduce this bias?
  - variables from wave 1 that might help predict those changes (e.g. indicators for quality of partnership; satisfaction with dwelling)
  - information from contact protocols of wave 2 (changes in patterns)

4 Research Design (1)

- Cases for analysis
    - 9.386 persons from UBII sample participated in wave 1
    - For 7.273 (77.5%) admin. data were successfully linked
    - 3.503 (48.2%) of these persons participated in wave 2
1. Calculate attrition rates for persons with / without certain events between waves according to admin data
    - differentiate between contact and cooperation given contact
2. Calculate attrition bias as difference between proportion of wave 1 participants with event and wave 2 participants with event
3. Weigh wave 2 by PASS propensity weights and calculate bias again
4. Add predictors of change from previous wave and from fieldwork of current wave to improve models, weigh again and calculate bias

5 Results (1): Attrition Rates by Events

Event	n	Contact rate	Cooperation rate	Response rate
Events defined for whole UB II sample				
total	*7273*	*68,1%*	*73,1%*	*48,2%*
no employment -> employment	645	65,3%ns	70,3%ns	44,8%ns
change to full time job	462	62,5%**	67,7%*	41,3%**
UB II exit	1128	59,6%**	66,8%**	38,7%**
UB II entry	220	71,7%ns	75,0%ns	51,8%ns
UB II continuous	4696	70,0%**	75,8%**	51,4%**
no UB II continuous	1229	67,7%ns	67,0%**	43,8%**
UB II sanction	54	70,4%ns	65,8%ns	46,3%ns
Moved to other district	290	42,6%**	72,7%ns	30,3%**
Events defined only for UB II continuous				
total	*4696*	*67,7%*	*67,0%*	*51,4%*
Family status to single/divorced/widow	47	36,2%**	76,5%ns	27,7%**
household size smaller	320	60,5%**	74,7%ns	44,4%**
household size bigger	255	64,7%*	68,9%*	43,5%**

** $p < 0,01$

* $p < 0,05$

5 Results (2): Bias and Bias Reduction by Weighting

Attrition Bias	(1) w1 resp.	(2) w2 resp.	(3) (2)+ Infratest weights	
	proportion	proportion	proportion	bias red.
change to full time job	6,35%	5,45%	5,77%	**35,56%**
UB II exit	15,51%	12,45%	12,47%	0,65%
UB II continuous	64,57%	68,94%	67,46%	33,87%
no UB II continuous	16,90%	15,36%	16,16%	51,95%
Moved to other district	6,18%	**3,65%**	3,45%	-7,91%
Family status to single/divorced/widow	1,00%	**0,54%**	0,51%	-6,52%
household size smaller	7,09%	6,01%	6,79%	72,22%
household size bigger	5,65%	4,70%	4,81%	11,58%

Institute for Employment Research
The Research Institute of the Federal Employment Agency
IAB

Propensity Models	(3) p(contact)	(3) p(success)	(4) p(contact)	(4) p(success)	(5) p(contact)	(5) p(success)
Infratest model	included	included	included	included	included	included
Predictors: hhsize / Family status / moved						
conflicts in household			0.918	1.103	0.936	1.107
			(0.281)	(0.347)	(0.423)	(0.332)
partner outside household			0.916	1.139	0.913	1.140
			(0.224)	(0.185)	(0.232)	(0.183)
person aged 17-20 in household			1.016	1.177	1.037	1.175
			(0.882)	(0.184)	(0.747)	(0.187)
not satisfied with dwelling			0.787	1.078	0.793	1.077
			(0.071)	(0.702)	(0.102)	(0.707)
pays rent regularly			1.054	1.007	1.208	1.001
			(0.699)	(0.966)	(0.183)	(0.993)
state of dwelling			0.924	0.936	0.924	0.937
			(0.374)	(0.574)	(0.396)	(0.579)
applied for job >100km			0.888	1.237	0.912	1.227
			(0.320)	(0.228)	(0.478)	(0.248)
Concession for job: move			0.806	1.296	0.820	1.299
			(0.065)	(0.121)	(0.107)	(0.118)
Predictors: Found a job / UB II exit						
replied to job ads (≥10)			0.811	0.854	0.789	0.851
			(0.103)	(0.388)	(0.078)	(0.379)
placed 'employment wanted' ad (≥1)			0.789	1.720	0.781	1.721
			(0.131)	(0.030)	(0.127)	(0.030)
asked for jobs at companies (≥10)			1.004	0.960	0.967	0.961
			(0.973)	(0.821)	(0.803)	(0.828)
unsolicited applications (≥5)			0.952	1.088	1.031	1.089
			(0.628)	(0.563)	(0.780)	(0.560)
at least 1 job interview last month			0.806	1.296	0.820	1.299
			(0.065)	(0.121)	(0.107)	(0.118)
initial adress/phone number wrong					0.082	1.541
					(0.000)	(0.104)
additional variables (not displayed)	not included	not included	included	included	included	included
Pseudo R^2:	0.069	0.073	0.074	0.085	0.142	0.086

13

5 Results (4): Bias and Bias Reduction by Weighting

Attrition Bias	(1) w1 resp.	(2) w2 resp.	(3) (2)+ Infratest weights		(4) (3)+predictors prev. wave		(5) (4)+ predictors contact protocoles	
	proportion	proportion	proportion	bias red.	proportion	bias red.	proportion	bias red.
change to full time job	6,35%	5,45%	5,77%	**35,56%**	5,55%	11,11%	5,48%	3,33%
UB II exit	15,51%	12,45%	12,47%	0,65%	12,36%	-2,94%	13,20%	**24,51%**
UB II continuous	64,57%	68,94%	67,46%	33,87%	67,24%	38,90%	66,63%	**52,86%**
no UB II continuous	16,90%	15,36%	16,16%	51,95%	16,45%	**70,78%**	16,08%	46,75%
Moved to other district	6,18%	**3,65%**	3,45%	-7,91%	3,45%	-7,91%	3,48%	-6,72%
Family status to single/divorced/widow	1,00%	**0,54%**	0,51%	-6,52%	0,51%	-6,52%	0,52%	-4,35%
household size smaller	7,09%	6,01%	6,79%	72,22%	6,87%	**79,63%**	6,66%	60,19%
household size bigger	5,65%	4,70%	4,81%	11,58%	5,02%	33,68%	5,27%	**60,00%**

6 Future Work and Discussion (1)

- Preliminary Conclusions
  - Events between waves influence panel attrition
    - In benefit recipient population: Finding a full-time job and overcoming recipiency lead to lower contact and cooperation rates
    - Moving and changes in household composition and family status lead to lower contact rates
  - This leads to biased estimates of key variables of the survey
    - PASS propensity weighting reduces bias for most variables
    - Additional variables from the previous wave that predict change in propensity models are only a slight improvement
      - prediction is weak
    - Additional indicators from the fieldwork of the survey seem more promising

5 Future Work and Discussion (2)

- Limitations
  - Wave 2 attrition rates unusually high: In particular high noncontact rate
  - Only (former) benefit recipients
- Future Directions
  - Re-analysis with wave 3
  - Add linked cases from general population sample
  - Sequence analysis of contact data to infer change (Kreuter & Jäckle 2008)

References

Büngeler, Kathrin; Gensicke, Miriam; Hartmann, Josef; Jäckle, Robert & Tschersich, Nikolai (2009): IAB-Haushaltspanel im Niedrigeinkommensbereich Welle 2 (2007/08). Methoden- und Feldbericht. FDZ Methodenreport, 08/2009, Nürnberg.

Heller, G. & Schnell, R. (2000). The Choir Invisible. Zur Analyse der gesundheitsbezogenen Panelmortalität im SOEP; in: Helmert,U., Bamman,K., Voges,W., Müller,R. (Eds), Müssen Arme früher sterben? Soziale Ungleichheit und Gesundheit in Deutschland, München (Juventa), pp.115-134.

Kreuter, F. & Jäckle, A. (2008). Are contact protocol data informative for potential nonresponse and nonresponse bias in panel studies? A case study from the Northern Ireland subset of the British Household Panel Survey. Paper presented at the first Panel Survey Methods Workshop, Colchester, 14-15 July 2008.

Tortora, R.D. (2009). Attrition in consumer panels (pp. 234-250). In Lynn, P. (ed.). Methodology of Longitudinal Surveys. Chichester: Wiley.

Van den Berg, G.J. Lindeboom, M. & Dolton, P.J. (2006). Survey non-response and the duration of unemployment. Journal of the Royal Statistical Society: Series A,

Watson, N. & Wooden, M. (2009). Identifying factors affecting longitudinal survey response (pp. 157-182). In Lynn, P. (ed.). Methodology of Longitudinal Surveys. Chichester: Wiley.