Validating paradata and predicting unit nonresponse with police statistics: a test of the 'broken windows' hypothesis for face-to-face surveys.

Anina Vercruyssen – Celine Wuyts – Geert Loosveldt

KULeuven

Survey nonresponse is affected by individual motivations, the interaction between interviewer and sample unit, societal factors and local survey climate (see e.g. Groves & Couper, 1998; Groves et al., 2009; Stoop, 2012 for overviews). With regard to the societal factors, European Social Survey (ESS) systematically gathers interviewer observations of the immediate vicinity of the houses of sample units about visible signs of neighbourhood disorder. These observations of neighbourhood disorder or decay can be linked to the 'broken windows' hypothesis about the increased likelihood of more serious crimes and avoidance behaviour of its inhabitants (Wilson & Kelling, 1982). The current paper tests this 'broken windows' hypothesis in the realm of survey methodology by using the interviewer observations from the 6th Round of European Social Survey in Belgium (ESS6BE) and comparing them with police zone crime statistics to assess if we can validate the interviewer-observed paradata. Moreover, we investigating the relationship between these auxiliary data and nonresponse in ESS6BE.

The 'broken windows' hypothesis and survey participation

Social exchange theory is one of the dominant theories for explaining survey (non-) participation. It refers to the norm of reciprocating favours and to the notion of obligations and expectations between individuals and societal institutions (Dillman, 2000; Goyder, 1987; Groves et al., 1992; 2002). The underlying assumption is that interactions are based on a subjective comparison of costs and benefits. As follows, people will accept a request to participate in surveys if the social benefits outweigh the costs. Social connectedness, involvement and responsibility are, of course, influenced by the communities in which people live. With regard to the latter, the 'broken windows' hypothesis was formulated by Wilson & Kelling (1982) as a theory on social control in communities that has gotten wide attention in criminology. A broken (house) window that remains unfixed can be seen as a sign that no one cares for the neighbourhood. It will attract more neighbourhood decay, incivilities and disorderly behaviour, which in their turn will increase crime rates, fear and avoidance behaviour in communities. As follows, we can expect more fear, lower social connectedness and more avoidance behaviour towards unknown individuals – such as interviewers – in "bad" neighbourhoods. Hence, 'broken windows' might tell us something about the likelihood of cooperation in face-to-face interviews.

Without straightforwardly referring to the 'broken windows' hypothesis, quite some studies found evidence for the effects of neighbourhood incivilities and decay on survey non-participation. Studies based on ESS show that sample units living in neighbourhoods with poor physical state of buildings were harder to contact and less likely to participate in most countries in Round 1 of ESS (ESS1, Blom et al., 2011) and Round 2 (ESS2, Billiet et al., 2009; Cincinatto et al., 2008). Kreuter et al. (2010), however, only found very small correlations between interviewer observations of litter and the response propensity in ESS1 when investigating only Greece, Poland and Portugal. In ESS2, the presence of litter and/or vandalism did lead to more non-contact, initial and final refusals in most of the investigated countries (Billiet et al., 2009; Cincinatto et al., 2008). Similar results can be found for ESS5

(Beullens, 2013). Billiet et al. (2009) also found that the presence of vandalism/graffiti and litter/ rubbish in neighbourhoods correlated with the presence of buildings in a bad state, which proves that neighbourhood decay and incivilities indeed attract each other as the 'broken windows' hypothesis predicts. For ESS5, the results with regard to the effects of interviewer observations are more mixed throughout the countries, but significant effects of poor neighbourhoods and housing quality on participation remain in e.g. Belgium (Beullens, 2013). In studies on other survey data, the literature also shows that sample units living in neighbourhoods with buildings in poor physical state were more likely to refuse than sample units from well-maintained neighbourhoods (Durrant et al., 2009; 2013; Lipp et al., 2005; Lynn, 2003; Stoop, 2005).

The availability of police zone data in Belgium allows us to test the 'broken windows' hypothesis more into detail by cross-referencing and potentially validating the interviewer observed paradata and for assessing if 'broken windows' can tell us something about nonresponse in "bad" neighbourhoods.

Data & Methods

For this study, the Belgian data of ESS6 (see Tirry & Loosveldt, 2013) are used as well as the Belgian police crime statistics from the same year (2012). The police statistics are freely accessible online (Belgian Federal Police, 2012) but the platform is not always active. Each police zone covers a handful of adjacent municipalities. Out of the 194 police zones in Belgium, 148 are represented in both our gross sample and our response sample. The police statistics provide statistics about different types of crimes ranging from petty theft to assault and from graffiti to vandalism. Especially the latter two are of interest when it comes to cross-referencing the interviewer observations for vandalism/graffiti on neighbourhood level. We also use the overall crime number per 1,000 inhabitants per zone. As such, we take into account that more densely populated areas may have a higher absolute number of crimes. To avoid potential problems with the scaling of these skewed count variables, we standardized them as well. The interviewer observations were recoded to deal with small categories. For the *physical condition of buildings* we had to group bad and very bad house state together, the other categories (satisfactory, good, very good) were big enough. For *vandalism/graffiti* and *litter/rubbish*, we had to create binary variables as neither occurred very often.

Given that sample units are hierarchically clustered in police zones, this random effect is included in each model. For each model, only one indicator of crime is included at a time to avoid multicollinearity due to the extremely high correlations of the police zone statistics among themselves (see Table 1). As women and older people tend to be more afraid of crime (e.g. Gainey et al., 2011), we also control for interviewer gender and age.

Results

Table 1 shows the Spearman correlations based on the data on individual sample unit level. Interestingly, we see rather modest but significant Spearman correlations in the expected direction on the individual level between all police zone crime numbers and interviewer observations. In the multilevel models (Table 2), we continue to see the significant relationships between police zone crime statistics and interviewer observations: more police zone total crime rates as well as more police zone reports of vandalism and graffiti (per thousand inhabitants) significantly predict a higher odds of interviewers reporting that the state of the sample unit's house is (very) bad and a higher odds that the interviewer is going to observe vandalism/graffiti and litter/rubbish in the neighbourhood. As such, the substantial content of the 'broken windows' hypothesis gets confirmed: small signs of neighbourhood decay and incivilities do related to more serious crimes in Belgium. And it shows some validation of the interviewer-observed auxiliary data given the significant relationships with official police zone statistics.

The binary logistic multilevel models (Table 3) show that neither police zone statistics nor interviewer observations of vandalism/graffiti and litter/rubbish have significant effects on the survey outcomes for ESS6BE. We also did not observe any significant random intercepts, meaning police zones do not significantly differ in average contact, refusal, nonresponse, and ineligibility rates either. A (very) bad state of the houses as observed by the interviewers, however, does significantly predicts noncontact, nonresponse and ineligibility, but not refusal. As such, we do see how interviewer observations on neighbourhood decay can say something about survey participation but the lack of significant effects of refusal suggests that it does not seem to be a deliberate act of avoidance behaviour, unless noncontacts are actually sample units who were at home but refused to open the door. We also find it remarkable that sample units are more likely to be identified as ineligible when they are male and when they are living in "bad" neighbourhoods.

Discussion and conclusion

Our study confirms that police zone crime statistics and interviewer observations on neighbourhood decay/incivilities relate with each other in the expected directions. Although the correlations between observations on neighbourhood decay/incivilities and disaggregated police statistics are significant, they are rather modest. Additionally, there are considerable scale differences between police zones and neighbourhood levels. Hence, stating that police zone data is a viable alternative for the neighbourhood observations would be one step too far. As for using police statistics to validate the interviewer observed paradata on the neighbourhood of sample units, the same objections on the modest correlations and scale differences can be uttered: the correlations with the neighbourhood observations – though significant on the individual sample unit level – are too modest and the Belgian police zones seem to be too big to be able to say we can completely validate this interviewer-generated paradata with the external police data. Still, the significant relations between police zone statistics and interviewer observations of neighbourhood decay/incivilities provide us with proof that the 'broken windows' hypothesis as originally intended as a theory on social control and connectedness in neighbourhoods also applies in Belgium: small incivilities relate to more serious crime.

Both types of auxiliary data also correlate significantly with nonresponse, cooperation, contactability and ineligibility. In the multilevel models, however, we only see significant effects of the interviewer-observed state of the houses on these components of survey (non-) participation. The absence of significant correlations and effects of any of the interviewer observations or police zone statistics on refusal suggests that 'broken windows' do not lead to increased avoidance behaviour – unless some of the noncontacts are actually caused by people refusing to even open the door for an unknown interviewer. But given that 'broken windows' do significantly affect the other dimension of survey non-participation, we should definitely proceed with collecting these interviewer observations when conducting face-to-face surveys. The significant effect on eligibility is also intriguing as it leads us to wonder whether people in more dilapidated neighbourhood more often move away from such areas or whether interviewers more easily classify a sample unit from a dilapidated neighbourhood as ineligible so they do not have to spent too much time in such a “bad” area. Future research should look into that by, for example, including a question in the contact forms on whether interviewer suspects safety issues in the neighbourhood of the sample unit.

References

Belgian Federal Police [Federale Politie België] (2012). *CGOP/B-Politiële criminaliteitstatistieken 2012*. http://www.politie.be/fed/nl/statistieken. Consulted April 16th, 2015.

Beullens, K. (2013). *The use of paradata to assess survey representativity. Cracks in the nonresponse paradigm.* Leuven: KULeuven.

Billiet, J., Vehovar, V., Beullens, K., & Matsuo, H. (2009). Nonresponse Bias in Cross-national Surveys: Designs for Detection and Adjustment in the ESS. *ASK. Research&Methods*, (18), 3-43.

Blom, A. G., de Leeuw, E. D. and Hox, J. J. (2011) Interviewer effects on nonresponse in the European Social Survey. *J. Off. Statist.*, 27, 359–377.

Cincinatto, S., Beullens, K., & Billiet, J. (2008). *Analysis of observable data in call records ESS* – R2. Deliverable no 6 of Joint Research Actions 2 of ESSi.

Couper, M. P., & Groves, R. M. (1996). Social environmental impacts on survey cooperation. *Quality and Quantity*, 30(2), 173-188.

Couper, M. P., Singer, E., & Kulka, R. A. (1998). Participation in the 1990 Decennial Census Politics, Privacy, Pressures. *American Politics Research*, 26(1), 59-80.

Dillman, D. (2000). *Mail and Internet Surveys. The Tailored Design Method* (Second edition). New York: John Wiley & Sons, Inc.

Durrant, G. B., D'Arrigo, J., & Steele, F. (2013). Analysing interviewer call record data by using a multilevel discrete time event history modelling approach. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(1), 251-269.

Durrant, G. B., & Steele, F. (2009). Multilevel modelling of refusal and non-contact in household surveys: evidence from six UK Government surveys. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(2), 361-381.

Gainey, R., Alper, M., & Chappell, A. T. (2011). Fear of crime revisited: Examining the direct and indirect effects of disorder, risk perception, and social capital. *American Journal of Criminal Justice*, 36(2), 120-137.

Goyder, J., Boyer, L., & Martinelli, G. (2006). Integrating exchange and heuristic theories of survey nonresponse. *Bulletin de méthodologie sociologique*, 92(1), 28-44.

Groves, R., Cialdini, R., & Couper, M. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56(4), 475-495.

Groves, R. & Couper, M. (1998). *Nonresponse in household interview surveys*. New York: Wiley.

Groves, R., Dillman, D., Eltinge, J., & Little, R. (2002). *Survey Nonresponse*. New York: Wiley series in survey methodology.

Kreuter, F., Olson, K., Wagner, J., Yan, T., Ezzati-Rice, T.M., Casas-Cordero, C., Lemay, M., Peytchev, A., Groves, R.M., & Raghunathan, T.E. (2010). Using proxy measures and other correlates of survey outcomes to adjust for nonresponse: examples from multiple surveys. *Journal of the Royal Statistical Society. Series A,* 173 (part 2), 389-407.

Lipps, O., Benson, G., & Panel, S. H. (2005). Cross-national contact strategies. *Proceedings of the Survey Research Section of the American Statistical Association, American Statistical Association, Alexandria, VA*.

Lynn, P. (2003). PEDAKSI: Methodology for collecting data about survey non-respondents. *Quality and Quantity*, 37(3), 239-261.

Stoop, I. (2005). *The hunt for the last respondent: Nonresponse in sample surveys*. The Hague: Social and Cultural Planning Office.

Tirry, D., & Loosveldt, G. (2013). *European Social Survey Round 6 Belgium Process evaluation for the data collection*. Onderzoeksverslag Centrum voor Sociologisch Onderzoek (CeSO) Survey Methodology CeSO/ SM /2013-4.

Wilson, J. Q., & Kelling, G. L. (1982). Broken windows. *Atlantic monthly*, 249(3), 29-38.

Table 1: Spearman correlations of disaggregated police zone data, interviewer observations, and ESS6BE survey (non-)participation on sample unit level

		Mean, mode or proportion	**reported crime police zone**				**interviewer observations**			**ESS6BE survey outcome**			
			1	2	3	4	5	6	7	8	9	10	11
1	police zone total crime	774.751	1										
2	police zone vandalism + graffiti	46.412	**.984**	1									
3	police zone graffiti	1.448	**.892**	**.903**	1								
4	police zone vandalism	44.965	**.984**	**1**	**.898**	1							
5	physical condition house	good state	**.133**	**.130**	**.100**	**.130**	1						
6	vandalism/graffiti	0.081	**.203**	**.195**	**.161**	**.195**	**.293**	1					
7	litter/rubbish	0.169	**.184**	**.177**	**.137**	**.178**	**.387**	**.586**	1				
8	Response	0.587	**-.055**	**-.054**	**-.053**	**-.054**	**-.087**	-.027	**-.039**	1			
9	Cooperation	0.629	**-.044**	**-.044**	**-.044**	**-.044**	**-.050**	-.008	-.021	**1**	1		
10	Contact	0.934	**-.048**	**-.042**	**-.041**	**-.042**	**-.135**	**-.066**	**-.067**	**.316**	-	1	
11	Eligibility	0.974	**-.037**	**-.037**	**-.044**	**-.036**	**-.118**	**-.086**	**-.074**	-	-	-	1

Note: correlations in **bold** are significant at least at 0.05-level.

Table 2: Multilevel models predicting interviewer observations with police statistics

	State house (ref =very bad)				Vandalism/graffiti (ref = no)				Litter/rubbish (ref = no)			
model term	**coefficient**	**sig.**	**coefficient**	**sig.**	**coefficient**	**sig.**	**coefficient**	**sig.**	**coefficient**	**sig.**	**coefficient**	**sig.**
Threshold house very good	-0.604	0.000	-0.611	0.000								
Threshold house good	1.319	0.000	1.313	0.000								
Threshold house satisfactory	3.168	0.000	3.161	0.000								
Intercept					-3.226	0.000	-3.208	0.000	-1.964	0.000	-1.945	0.000
Police zone total crimes	0.311	0.000			0.669	0.000			0.386	0.001		
Police zone vandalism+graffiti			0.309	0.000			0.668	0.000			0.377	0.000
Gender interviewer = female	-0.101	0.410	-0.104	0.398	-0.107	0.789	-0.117	0.771	-0.106	0.676	-0.111	0.661
Age interviewer	0.037	0.586	0.035	0.608	-0.007	0.972	-0.013	0.945	-0.052	0.710	-0.055	0.692
Variance intercept (police zones)	0.454	0.000	0.456	0.000	1.617	0.000	1.613	0.000	0.951	0.000	0.947	0.000
BIC	20613.563		20613.87		10656.24		10656.18		9229.743		9228.850	

Table 3: multilevel binary logistic regressions for survey (non-)participation in ESS6BE

	Noncontact (ref = contact)		Refusal (ref = no refusal)		Nonresponse (ref = response)		Ineligible (ref = eligible)	
model term	**coefficient**	**sig.**	**coefficient**	**sig.**	**coefficient**	**sig.**	**coefficient**	**sig.**
Age	-0.39467	0.000	0.05995	0.163	0.11053	0.00325	0.15339	0.196740
Female	-0.26899	0.08544	0.09181	0.292	0.10905	0.14787	-0.58589	0.015966
Physical condition house (ref = (very bad)								
house satisfactory	-0.10889	0.70273	0.06016	0.798	-0.06731	0.71208	-1.16564	0.000532
house good	-0.83830	0.00523	0.31513	0.166	-0.26436	0.14195	-1.60066	0.000000
house very good	-1.03734	0.00108	0.11537	0.620	-0.44569	0.01576	-2.06135	0.000000
Litter/rubbish (ref. = none)	-0.07711	0.75600	0.03270	0.833	0.04420	0.73635	0.05292	0.882066
Vandalism/graffiti (ref. = none)	0.31891	0.28693	-0.36139	0.115	0.05141	0.77305	0.62540	0.108393
Police zone level								
Total reported crimes per 1000p	0.12165	0.27739	-0.04659	0.554	0.04037	0.53582	-0.03226	0.767823
Variance intercept (police zones)	0.3797	0.6162	0.2034	0.4511	0.1461	0.3823	0	
BIC	1417.339		3390.797		4233.448		752.6327	