

RESPONSE PROBLEMS IN THE PROTEST SURVEY DESIGN: EVIDENCE FROM FIFTY-ONE PROTEST EVENTS IN SEVEN COUNTRIES*

Stefaan Walgrave, Ruud Wouters, and Pauline Ketelaars[†]

Protest surveys are increasingly used to tackle questions related to participation in social movements. However, it is unclear whether they generate useful and valid data. This study puts the protest survey design to the test by relying on data of 51 demonstrations (2009-2011) in seven European countries. We use data on 15,000 protest participants combined with screener questionnaires and extensive debriefing records of the interviewer teams. We account for noncontact (fieldwork problems), immediate and delayed refusal, and refusal bias. Results show that fieldwork problems are frequent, that immediate refusal is low, and that delayed refusal is considerable. Systematic refusal bias is only found for age and education. Differences between countries and protest issues are small but the issue determines the composition of an event which, in turn, leads to higher or lower refusal. Researchers should be cautious when using protest survey data to compare protest events across issues. The paper pleads for standardization and constant monitoring of the data-gathering process.

Political protest is booming. The massive wave of protest against the war in Iraq in the mid-2000s was followed by the Occupy and Indignados movements at the turn of the decade, which overlapped with vast protests against austerity measures throughout Europe. At the same time, protest demonstrations became the main weapon of the disenfranchised masses in the Middle East. Rising protest participation is not a recent phenomenon, however. It is a secular trend of the last few decades (Norris, Walgrave, and Van Aelst 2005). This wave of protest raises three simple but important questions: Who protests? Why? And, how are these protesters mobilized?

Social scientists typically rely on survey designs to tackle such questions. Yet the survey method is not nearly as popular among social movement scholars as it is among sociologists and political scientists in general. According to John Crist and John McCarthy (1996) only one fifth of social movement studies published in major sociology journals relied on a form of survey design (see also Klandermans and Smith 2002). Currently, though, we are witnessing a surge in studies relying on the specific *protest* survey methodology—sampling/interviewing participants in contentious events while they are in the act of protesting—with at least a few dozen protest survey studies published during the last couple of years.¹

Remarkably, the spike in protest survey *use* has not been matched by a parallel surge in *analysis* of the validity of protest survey data. With some exceptions, extant work simply accepts that protest survey data are valid and can be used as if containing a random sample of participants. This study addresses this lacuna. We draw on a dataset across a variety of issues and on extensive debriefing data of interviewers and fieldwork supervisors for 51 demonstrations staged in seven European countries between 2009–2011. This evidence, for the first

* The authors thank their collaborators of the project Caught in the Act of Protest: Contextualizing Contestation (CCC) for being able to use the data. The CCC-project ran in the Netherlands, Belgium, Spain, Sweden, Switzerland, and the U.K. and was funded by grant number 08-ECPR-001 of the European Science Foundation.

[†] Stefaan Walgrave is Professor in the Department of Political Science at the University of Antwerp. Ruud Wouters is a Postdoctoral Fellow of the Research Foundation Flanders at the University of Antwerp. Pauline Ketelaars is a postdoctoral research in the Department of Political Science at the University of Antwerp. Please direct all correspondence to Stefaan Walgrave at Stefaan.Walgrave@uantwerpen.be.

time, yields a large-scale, comparative, and encompassing test of the data strengths and weaknesses of the protest survey design.

For students interested in who protests, why, and how participants get mobilized, the advantages of protest surveying are clear. To answer these questions it is key to take the *context* of protest participation into account. The issue at stake, the protest organizers, the political circumstances, etc., all affect which people take to the street, how they are mobilized, and why they show up. A protest of asylum seekers is very different from an event opposing the closing of a factory. Grasping this crucial context of protest participation can be done by comparing demonstrators participating in different demonstrations, on different issues, and in different nations. Norris et al., for example, compared protestors across eight issues in a single country. They found considerable cross-issue differences in the sociodemographic background, attitudes, and network embeddedness of participants. In another study, Stefaan Walgrave and Dieter Rucht compared the 2003 anti-Iraqi-war protesters in eight countries. By analyzing one issue but varying the nation, they showed that the attitudes and political beliefs of protesters strongly differ according to the political context; more diverse crowds take to the streets when media and public opinion are supportive of the movement.

In contrast, political scientists primarily rely on population surveys to assess political participation in general and protest participation in particular (e.g., the well-known study by Barnes and Kaase 1979). However, population surveys that ask citizens about their protest participation *in general* tend to lose sight of the context in which protest participation (or nonparticipation) takes place. Political protest is typically analyzed as a decontextualized phenomenon that occurs in a social vacuum. It is examined detached from the issue that spurred people to protest, isolated from the social groups that are suffering the grievance, and disconnected from the organizations or networks that stage the event (see also Verba, Schlozman, and Brady 1995). From general surveys, for example, we know that higher educated people are more likely to participate in protests than less educated people. Still, many events are not populated by the highly educated (e.g., protests by asylum seekers or the homeless). The same general surveys found that most participants protest to change the world around them—that is, for instrumental reasons. Yet, we see a lot of action directed toward a lost cause without any reasonable chance of success (e.g., the protest against the imminent war in Iraq in 2003). Population survey scholars established that recruitment for protest mainly happens in networks and organizations, but some events appear to be spontaneous and lack a formal foundation (e.g., flash mobs or demonstrations against random violence).

In sum, while yielding valuable evidence on the average protester, population surveys are not well-equipped to explain across issues and organizational and political contexts who protesters are, why they show up, and how they are mobilized. So, to understand protest participation better, we need to have an eye for the differences across events and for the interaction between individuals and the context in which they operate. Since protest events are hardly so large that their participants would constitute a sizeable share of the respondents in a population sample, these differences cannot be tackled employing population surveys. At a time when protest participation is going up and large segments of the population declare that they have participated in protest, the next step is to start thinking about the causes the participants support rather than solely focusing on participation as such.

In contrast to political scientists, sociologists have mainly employed in-depth case studies to study protest. They typically study one or a handful of movements. Their studies are rich in context but poor in generalizability. After all, what can we learn in general from studying a single movement/protest in a specific context in one nation? Political scientists studying protests have the opposite problem: they typically employ population studies that are strong in generalization but weak in context. Situated between the sociology of social movements and the political science of political participation, our comparative survey-protest design may offer the best of both worlds. Provided that a sound sample of demonstrations is covered, the approach allows researchers to make general claims and, at the same time, systematically

measure and account for the differences between demonstrators caused by characteristics of demonstrations, issues, and countries.

In a sense, comparative protest surveying also tackles the micro-macro gap, a problem in the social sciences more generally (Giddens 1987). The macro context generates a demand for protest among disaffected individuals. Organizations, the meso level, provide a supply of protest occasions by staging events (Diani and McAdam 2003; Klandermans 1997). Varying the macro and meso level by gauging different demonstrations while at the same time measuring the micro level, helps to make sense of how the micro, meso, and macro levels are interacting.

Of course, the protest survey design is not the Holy Grail for protest scholars. It only catches actual participants and does not compare participants to nonparticipants. It focuses on differences between activists, not on differences between activists and nonactivists. Also, because demonstrations are typically unplanned (or not planned long in advance), and because one simply cannot know beforehand which issues will be the object of future demonstrations, sampling protests is tricky and often unsystematic. Additionally, the scope of the design is limited. The sampling procedure of individual respondents discussed below supposes a sufficiently large and more or less orderly event; violent events are difficult to cover, which limits the generalizability of the evidence. The method seems to be best suited for peaceful street demonstrations staged in established democracies.

But, the real potential of the protest survey design, currently booming among students of protest, depends on the quality of the evidence it generates. Can we be sure that the evidence gathered among a sample of participants represents the population attending the demonstration? Knowing that the approach is essentially comparative, can evidence collected in different events be compared? Are the inevitable biases in the data systematic or random? If the protest survey method is to become an established instrument to scrutinize the who, why, and how of protest participation, we need to know more about its biases.

THE PROTEST SURVEY DESIGN

Frank Parkin's (1968) widely cited study on middle-class radicalism was the first to draw on a protest survey design. Since then, many variants of the design have been used as different sampling procedures and interview modes were employed. Often, sampling was not systematic (and thus not random) but depended on the common sense of the interviewers. Interviews were conducted on the spot or afterward. Some studies employed mail-in questionnaires; others used self-administered questionnaires completed by the participants during the event (Walgrave and Verhulst 2011).

When we discuss protest survey design in this article, we are referring to the most employed variant as spelled out by Walgrave and Verhulst (2011), which includes two phases. In the sampling phase individual protesters are selected; in the interview phase the selected participants are briefly interviewed face-to-face and/or are asked to accept a written questionnaire, fill it in at home and send it back via mail. Thus, sampling and interviewing are unconnected. Specific researchers (pointers) select the protesters to be interviewed while others (interviewers) carry out the actual interviews. This method is applied because interviewers, if left free to select their own collocutors, typically select approachable peers for the interviews, which leads to a selection bias (Dillman, Eltinge, Groves and Little 2002; Lynn and Clarke 2002; Singer 2006). The pointer leading the fieldwork is called the fieldwork supervisor.

When the demonstration is moving from point A to point B two pointers each have a team of interviewers. One team starts at the head and works its way to the back of the demonstration. The second team does the opposite and starts at the tail. It is important that the entire crowd receives full and equal coverage. If not, participants do not have the same chance of being selected and the sample is not random.

The selection of specific respondents is systematic. The fieldwork supervisor makes an estimation of the size of the crowd before the march begins and decides how many rows of people in the march are skipped before an interviewer is sent into the crowd to interview a participant. In very large crowds, a lot of rows are skipped (e.g., an interview is conducted in every twentieth row); in small crowds the number of skipped rows is small (e.g., an interview in every second row). Pointers count the rows and select a person to be interviewed in that row—alternating between the left, the right, and the middle part. Then, an interviewer is sent in to approach the selected interviewee.

Based on the aimed-for ratio between face-to-face and postal interviews, every N^{th} respondent is briefly interviewed on the spot. The other respondents are only handed a postal questionnaire and are asked to fill it in at home and mail it back to the research team (postpaid by the addressee). At the end of their interviews, the participants who were interviewed get a postal questionnaire as well. These respondents' oral and postal questionnaires are marked to be identified and linked to each other. The postal questionnaire is rather elaborate containing eight to ten pages of questions. Longer questionnaires tend to suppress response rates (Rüdiger 2010).

RESPONSE PROBLEMS IN PROTEST SURVEYS

The general survey method literature highlights several sources of response problems which apply to protest surveys as well (Dillman et al. 2002; Lynn and Clarke 2002; Singer 2006). Broadly speaking, the first problem is that aimed-for respondents in the probability sample are not reached; this is called noncontact (e.g., targeted respondents do not pick up the phone or do not open the door). Second, when potential respondents are reached, they may refuse to collaborate (e.g., after opening the door, they close it again). This refusal can be immediate, meaning that respondents refuse to accept a questionnaire, or it can be delayed, meaning that respondents accept a questionnaire but do not complete it. Third, refusal may lead to refusal bias—i.e., a systematic difference between the population and the actual sample (e.g., the people who open the door and let interviewers in are different from the ones who do not). Noncontact and immediate or delayed refusal are only *potential* sources of bias since they do not necessarily lead to differences between the population and the sample. High levels of noncontact and refusal do tend to increase the chance of systematic differences between the population and the sample (Keeter, Miller, Kohut, Groves and Presser 2000), but recent studies find that the relationship between refusal rates and refusal bias is rather weak (Groves and Peytcheva 2008).

Because the difference in collaboration must be caused by something, it is obvious that those who collaborate and those who do not are different by definition. Bias is unavoidable and, since we do not know all of the differences between those who collaborate and those who do not, it cannot fully be measured. Rather, the question is whether bias distorts the conclusions, or, in other words, whether there is an interaction between the correlates of refusal and the variables that researchers are interested in.

Though it suffers from the same potential response problems as any other kind of survey, the protest survey design has particular features that make it different from other surveys. First, the population is unknown yet well-defined. All participants in a protest event form the population but there is no information regarding who these individuals are. Samples cannot be compared with known characteristics of the population, nor are there lists of participants from which random samples can be drawn. Second, demonstrations are temporary and short-lived events. Sampling and interviewing must be accomplished under considerable time pressure. To take maximum advantage of the small time window, protest surveys are generally carried out by large teams of interviewers. Third, demonstrations are social phenomena and participants can be cheerful, wild, angry, or aggressive. Adequately addressing participants can be a challenge. Often, circumstances are far from ideal to approach and talk to people, to ask

them questions, or to convince them to take home a questionnaire. Fourth and finally—although not unique to protest surveys—the design does not allow sending reminders to respondents as a means to convince them to collaborate, nor is there the option of pre-notification to prepare them for the interview. Because of the accumulation of all these non-conventional challenges and caveats, there are good reasons to systematically test the quality of the data generated by the protest survey design.

General survey research found that response problems are affected by factors on different levels (Groves and Couper 1998). We follow a multilevel approach here and test effects on four different levels of explanation: interviewer, demonstration, issue, and country level. The survey literature confirmed the existence of interviewer (Blom, de Leeuw, and Hox 2010) and country effects (de Leeuw and de Heer 2002). Due to the specific nature of protest we also expect the concrete issue of a demonstration, leading to a specific composition, to affect noncontact and refusal. The concrete social context in which the interview process takes place may affect response problems as well (respondents are selected and interviewed during a more or less chaotic demonstration). Therefore, we estimate multilevel models with interviewers or respondents embedded in demonstrations that are attributed issue features and country dummies. We do not offer concrete expectations about which countries or issues would be more or less plagued by some sort of response problem. We know that response rates to general population surveys in Switzerland, for example, seem to be lower than in other countries, while in Sweden they are substantially higher; the other countries in our sample fall in between (e.g. Stoop, Matsuo, Koch and Billiet 2010). All four sources of response problems—interviewers, demonstrations, issues, and countries—directly challenge the potential comparative benefit, and hence the usefulness, of protest survey evidence. Here, we therefore engage in focused quality control and set out to systematically scrutinize response problems and their consequences.

DATA AND METHODS

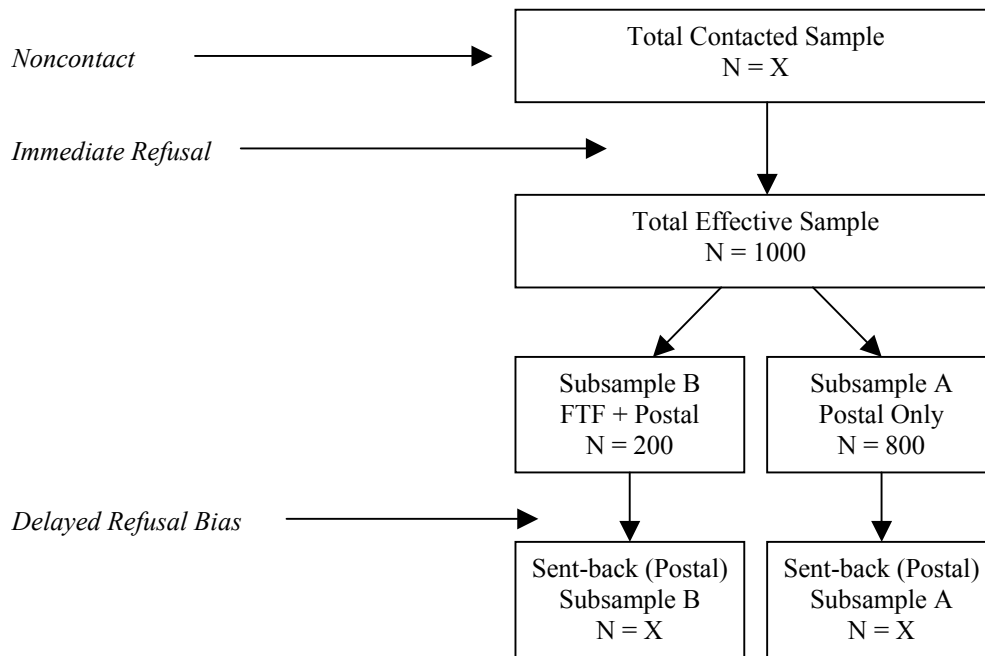
Between 2009 and 2011, 15,610 participants in 51 protest demonstrations were surveyed using the design explained above. A total of 11,414 participants returned a postal questionnaire. The dataset includes eleven demonstrations in Spain, eleven in the Netherlands, nine in the UK, seven in Belgium, six in Sweden, four in Italy, and three in Switzerland. The largest group of demonstrations ($N = 19$), dealt with budget cuts, austerity measures, and the banking crisis. This group is followed by May Day events ($N = 11$); climate change/energy demonstrations ($N = 7$); protests on reforms of the political system (nationalism, changing democratic rules, and the reform of the state) ($N = 6$); antiracist/antifascist events ($N = 4$); feminist demonstrations ($N = 2$); and single antiabortion ($N = 1$) and peace events ($N = 1$). Table A3 in the appendix presents an overview. The 51 demonstrations display a broad variety of causes, issues, movements, and countries. This variety presents a tough assessment for the method.

Figure 1 presents the basic scheme applied to all demonstrations. Interviewer teams aim to select a thousand respondents. Eight hundred of those, constituting subsample A, receive a postal questionnaire only. Two hundred other respondents, subsample B, are first interviewed briefly (3.5 minutes, 8 questions) and then asked to accept the postal questionnaire. In most cases, interviewer teams did not meet the target of distributing a thousand questionnaires. The average number of questionnaires handed-out was 716. The average number of face-to-face interviews per demonstration was 133, which was also below the target.

Figure 1 indicates which comparisons generate measurements of which response problems (printed on left in italics). The comparison between the face-to-face screener interviews (subsample B) and the postal questionnaires mailed back by that group, yields an assessment of the delayed refusal bias. As they require extra effort and are often difficult to obtain, screener data—that grasp a few respondent characteristics and precede the actual interview—are rather exceptional (Groves and Peytcheva 2008). Note that, in this study, we cannot

measure the bias caused by *immediate* refusal but only by *delayed* refusal. We did not register any features of those people who refused to collaborate altogether by declining the request to be briefly interviewed face-to-face. Immediate refusal rates are low, as we will see, but this does not mean that they are inconsequential.

Figure 1. Standard Sample Design



After the fieldwork, each interviewer completed a questionnaire resulting in 673 interviewer data points (444 interviewers) containing information such as the interviewer's age, whether he or she had lost the pointer, and so on. Additionally, per covered demonstration, the fieldwork supervisor filled in a separate questionnaire with general questions regarding fieldwork problems (for instance, the number of questionnaires distributed and the weather during the demonstration). Further, demonstrations are classified per issue, with typical old social movement issues (austerity and May Day) on the one hand, and postmaterialist or emancipatory issues typical for the new social movements (climate change, political system, antiracism, women, peace, and antiabortion) on the other. In the appendix, tables A1 and A2 present descriptive information for all variables.

We rely on fixed-effect multilevel models with random intercepts. Respondents and interviewers are embedded in demonstrations. Countries and issues are included in the models as demonstration characteristics because there are not enough countries and issues to warrant a separate level. We make use of the STATA functions `xtmixed` (continuous dependent variables) and `xtmelogit` (binary dependent variables) with restricted maximum likelihood (REML) estimates. We provide the formulae when we present our models. Marginal effects are predicted by using the STATA `margins` command (Rabe-Hesketh and Skrondal 2008). When we account for dependent variables at the demonstration level, we run ordinary least squares linear regressions and logistic regressions. We employ different significance thresholds for variables at the respondent/interviewer level (large N ; $p < .05$) than for variables at the demonstration level (small N ; $p < .1$) and present the significant coefficients in bold.

FIELDWORK PROBLEMS AND NONCONTACT

Noncontact implies that a part of the relevant population cannot be reached (de Leeuw and de Heer 2002). In protest surveys, noncontact is the consequence of poor fieldwork performance. Systematic fieldwork sampling in moving and often boiling masses is not easy. Demonstrations are disorderly events. The shape and the size of the venue differs, even during the course of a single demonstration, and the state of mind of the crowd may make sampling and interviewing difficult. As a consequence of fieldwork problems, the systematic sampling procedure may not be carried out fully with noncontact as a consequence: a part of the crowd is not covered as it should have been. Table 1, based on the answers of the fieldwork supervisors, presents the fieldwork problems encountered when implementing the procedures to our 51 demonstrations. The table also presents one direct indicator of noncontact, the fact that a part of the crowd was not covered.

Table 1. Problems Encountered During Protest Surveying Fieldwork and Noncontact (N = 51).

	No. Demonstrations	% Demonstrations
<i>Crowd behavior</i>		
Chaotic crowd, difficult to sample	9	18
Noisy crowd (or music), difficult to interview	24	47
Dense crowd, difficult to move through	21	41
<i>Crowd size</i>		
Overestimated crowd size, sampling adaptation	16	31
Underestimated crowd size, problems covering crowd	10	20
<i>Interviewers losing pointer</i>	44	86
<i>Not entire crowd covered</i>	6	12

Protesting crowds can be chaotic (18%). They run or stop, take side streets, or walk on the sidewalk. They can be noisy (47%) with slogans or music, making it hard for interviewers to communicate with respondents. Dense crowds (41%) make approaching selected respondents difficult. In most demonstrations one of these problematic types of crowd behavior was recorded. The crowd size estimation by the fieldwork supervisor before the start of the demonstration determines the number of rows that need to be skipped to generate a random sample. In half of the cases, the size of the crowd was over- (31%) or underestimated (20%). When underestimating crowd size, the middle part of the moving column or the head/tail of the demonstration risks getting unequal coverage. When the crowd is overestimated, too few interviews are administered as the counting rule skips too many rows.

The most common problem is interviewers losing their pointer. In more than four-fifth of the demonstrations (86%), at least one interviewer became temporarily disconnected from his/her pointer. Interviewers get lost for only a short time (16 minutes on average) and most can reestablish contact (672 interviewers out of 673). On average, slightly under half of the interviewers at each demonstration lost contact with their pointer at a given point in time. In contrast to the problems caused by noisy or disorderly crowds, the lost pointer problem is something the survey team can control. Increasing pointer visibility, namely by holding something distinctive in the air (e.g., an odd-colored umbrella) proves to be effective. Results of a logistic multilevel model reported in table 2 estimating whether a specific interviewer (embedded in a specific demonstration) lost his/her pointer (0/1) indicate that experienced interviewers lose their pointer less often. First-time interviewers have a 44% chance of losing their pointer; interviewers with experience in five demonstrations only have a 20% chance. Naturally, in large crowds (+5,000 participants), interviewers lose their bearing more often (49%) than in small crowds (27%). Finally, interviewers covering Italian demonstrations lose their pointer much more often (75%) than interviewers in other countries (37%). We have no ready explanation for this last finding. It may be that the layout of the Italian cities—with narrower or more winding streets—increases the chance that interviewers lose their pointer.

Table 2. Multilevel Logistic Regression with “Interviewer Losing Pointer (No/Yes)” as the Dependent Variable (N interviewers = 609; N demonstrations = 51).

		Coef.	Std. Err.	Sig.
	Constant	22.246	42.518	.601
<i>Interviewer</i>	Gender (Female)	-.121	.195	.533
	Age (old)	.011	.021	.586
	Experience (High)	-.231	.083	.005
<i>Demonstratio</i>	Turnout > 5,000	.943	.344	.006
	Rain	.589	.390	.131
	Moving	.052	.449	.907
	Chaotic	.342	.414	.408
<i>Issue</i>	Old (ref. = New)	-.374	.295	.205
<i>Country</i>	Netherlands (ref. = Belgium)	.544	.588	.355
	Italy	1.644	.626	.009
	Spain	-.315	.451	.485
	Switzerland	-.391	.620	.527
	Sweden	.538	.690	.435
	UK	.409	.526	.437
Wald Chi ² (df)			31.30 (14)	
Prob > Chi ²			.005	
Log Likelihood (empty model = -390.929)			-375.882	
Variance demonstration level (empty model = .903)			.407	
Total explained variance			55%	

Note: Multilevel logistic regression with random intercepts:

$$\text{Logit}\{\Pr(Y_{ij} = 1 | x_{ij}, \delta_j)\} = y_{00} + y_1x_{1ij} + y_2x_{2ij} + y_3x_{3ij} + y_4x_{4j} + y_5x_{5j} + \dots + y_{15}x_{15j} + \delta_{0j}$$

y_{00} = The grand mean across all interviewers and demonstrations. δ_{0j} = demonstration-specific error component.

y_1x_{1ij} to y_3x_{3ij} = interviewer level variables (gender, year born, demonstration experience)

y_4x_{4j} to $y_{15}x_{15j}$ = demonstration-level variables (turnout to country dummies)

Difficult crowd behavior, challenging crowd sizes, and interviewers losing pointers do not lead to noncontact as such. They just slow down the sampling process. The fieldwork procedure really fails, though, if the entire crowd is not covered—that is, if the interviewer teams do not give equal coverage to the whole demonstration. This happened in twelve percent of the cases (see table 1). Table 3 reports a logistic regression with the dependent variable of whether the demonstration was fully covered or not. Due to the limited number of observations (N = 51) the model contains only a few predictors. An important predictor of full coverage is the size of the interviewer team. A team with only eight interviewers (the minimum) has a 59% chance to cover the whole crowd while this chance increases to 97% with fourteen interviewers or more. Further, large demonstrations are harder to cover than small demonstrations. In partially covered demonstrations, respondents were—according to the subjective assessment of the fieldwork supervisor—less “responsive” (effect is marginally significant; $p = .101$). Nonresponsive demonstrators take more time to address, convince, and interview, which slows the pace of the fieldwork and leads to incomplete coverage. Issues and countries do not matter for full coverage (not in table).

Table 3. Logistic Regression with “Entire Crowd Covered (No/Yes)” as Dependent Variable.

	Coef.	Std. Err.	Sig.
Number of Interviewers	.546	.277	.049
Turnout > 5,000	-1.908	1.449	.188
Demonstration Moving	2.024	1.348	.133
Interviewer - Protestor Interaction: Responsive	1.112	.677	.101
Constant	-9.183	4.199	.029
N	51		
Prob > Chi ²	.008		
Pseudo R ²	.373		

IMMEDIATE REFUSAL

Immediate refusal is the direct refusal to grant a short interview and/or accept a mail-in questionnaire. *Delayed refusal*, which will be dealt with in the next section, is the failure to send the questionnaire back by mail. In comparison with typical survey designs, immediate refusal levels in our protest surveys are particularly low while delayed refusal levels are high.

We examine two distinct types of immediate refusal: demonstrators denying an oral interview (subsample B) and demonstrators not willing to take home a questionnaire (subsamples A and B). Since both the oral interview and the asking to take home a questionnaire are oral interactions, it is best to compare our immediate refusal with benchmark noncooperation levels for face-to-face interviews. The seminal, but older, study of Joop Hox and Edith De Leeuw (1994) estimated the average nonresponse rate for face-to-face interviews to be twenty nine percent. The protest survey design does better. Immediate refusal rates per interviewer, based on the interviewer questionnaires and averaged per demonstration, reach 10%. Demonstrators are slightly more reluctant to grant a face-to-face interview (13% refusal) than to accept a questionnaire (10% refusal). These low refusal levels are probably due to high interest in the survey topic and protesters' eagerness to express themselves, which is why they took to the streets in the first place (Groves, Presser, and Dipko 2004).

Table 4 presents multilevel models that estimate, per interviewer (embedded in a demonstration), the relative amount of face-to-face (FTF) and postal refusals. Since we did not gather any information about the protesters who refused immediately, we cannot test whether

Table 4. Immediate Refusal. Multilevel Linear Regression of Refusal Rate of FTF Interviews and Postal Questionnaires (N-interviewers = 608; N-demonstrations = 51).

Dependent variable		Refusal FTF Interview			Refusal Postal		
		Coef.	Std. Err.	Sig.	Coef.	Std. Err.	Sig.
Constant		-406.384	269.210	.131	-393.772	161.190	.015
Interviewer	Gender (Female)	-3.221	1.188	.007	-.776	.704	.270
	Age (old)	-.213	.135	.116	-.204	.081	.012
	Experience (High)	.217	.423	.608	-.250	.257	.332
Demonstr.	Turnout > 5,000	-3.657	2.395	.127	-4.612	1.301	.000
	Rain	2.236	2.876	.437	1.499	1.160	.337
	Moving	-4.276	3.197	.181	.279	1.760	.874
	Chaotic	6.355	3.021	.035	.905	1.627	.578
Issue	Old (ref. = New)	4.922	2.107	.019	4.300	1.142	.000
Country	NL. (ref. = Belg.)	-3.013	4.203	.473	-2.285	2.300	.321
	Italy	-5.978	4.311	.166	-2.118	2.340	.365
	Spain	-1.693	3.229	.600	-1.887	1.745	.279
	Switzerland	13.259	4.567	.004	4.088	2.454	.096
	Sweden	-4.015	4.682	.391	-.396	2.600	.879
	UK	-1.379	3.823	.718	-.284	2.071	.891
Wald Chi ² (df)		44.35 (14)			45.48 (14)		
Prob > Chi ²		.000			.000		
Demo. level	empty model	58.499			17.995		
	full model	27.696			7.299		
Interv. level	empty model	183.495			62.270		
	full model	180.735			61.851		
Log	empty model	-2486.033			-2109.784		
Explained variance interviewer		1.5%			0.6%		
Explained variance demo's		52.7%			59.4%		

Note: Multilevel linear regression with random intercepts:

$$Y_{ij} = y_{00} + y_1x_{1ij} + y_2x_{2ij} + y_3x_{3ij} + y_4x_{4j} + y_5x_{5j} + \dots + y_{15}x_{15j} + \varepsilon_{ij} + \delta_{0j}$$

Y_{ij} = refusal rate of interviews/booklets for interviewer i in demonstration j

y_{00} = the grand mean across all interviewers and demonstrations. δ_{0j} = demonstration-specific error component.

y_1x_{1ij} to y_3x_{3ij} = interviewer level variables (gender, year born, demonstration experience)

y_4x_{4j} to $y_{15}x_{15j}$ = demonstration-level variables (turnout to country dummies). ε_{ij} = interviewer-specific error component.

immediate refusal is affected by particular respondent characteristics. Yet, we can test whether refusal to grant an interview or accept a questionnaire is due to interviewer and demonstration characteristics. The survey literature has found noncooperation levels to vary between countries and to diverging “survey climates” (de Heer 1999). Table 4 shows the same to be true for protest surveys.

Demonstrators clearly prefer to talk to female interlocutors instead of male interlocutors (refusal FTF interview: men 15%; women 11%), but they accept take-home questionnaires from both. Older interviewers experience fewer refusals for the postal questionnaires: a ten-year-older interviewer has on average 2% less refusal. With an older interviewer, respondents may perceive the task to be of greater importance. Interestingly, getting someone to accept an interview or questionnaire does not come with experience. Although, at first sight, the age and sex effects do not seem to be particularly strong, one has to keep in mind that immediate refusal rates are particularly low (10%–13%) and that small shifts actually matter. Nevertheless, a large proportion of the variance at the interviewer level cannot be explained by our variables.

Large demonstrations (+5,000) witness substantially lower (5% less) postal refusal rates than smaller demonstrations. For FTF refusal, turnout fails to reach significance but marginal effects estimation indicates that large demonstrations result in 3% lower FTF refusals as well. We suspect that interviewers are more conspicuously present in smaller demonstrations, which may result in a more noncollaborative attitude among participants. Also, participants in larger demonstrations may feel emboldened by the turnout and therefore more eager to speak out. Chaotic conditions (according to the fieldwork supervisors) substantially affect FTF refusal: FTF refusal goes up from 12% (nonchaotic) to 18% (chaotic). In irregular and rough demonstrations, potential respondents, perhaps due to increased anger or excitement, are less willing to talk to interviewers.

A strong variable in both models is the issue dichotomy distinguishing old social movement demonstrations from new social movement events. FTF refusal in old events is, on average, 15%; it is 10% for new events. For postal refusal, the figures are 12% and 8%, respectively. Bread-and-butter issues generate significantly higher refusal, but during typical middle-class demonstrations, interviews are granted and questionnaires are accepted more easily. This finding is the first indication of the possibly distinct nature of response problems across issues—a possible cause for concern when using the evidence for comparing across issues.

There is one substantial country effect: Swiss demonstrators are much more unlikely to grant a FTF interview or to accept a questionnaire. Compared to the other countries, the Swiss immediate refusal rate is 13% higher for FTF interviews and 4% higher for postal questionnaires. As mentioned earlier, in other surveys as well, Swiss refusal rates have been found to be exceptionally high (Stoop et al. 2010).

DELAYED REFUSAL BIAS

Apart from immediate refusal, the protest survey design has a delayed refusal component: the failure to send an accepted questionnaire back. In general, refusal rates for surveys are rising, but not every interview mode or target group is equally hurt (Groves and Peytcheva 2008; de Leeuw and de Heer 2002). Surveys among specific populations typically have higher response rates than general surveys. Nevertheless, the specific nature of the protest survey design with, for example, no possibility of sending follow-up reminders (Dillman 1991) leads us to expect relatively low response rates.

In general, across all 51 demonstrations the delayed refusal rate is about 64%. One third of the respondents accepting a questionnaire returns it. Previously, we found that 90% of the targeted respondents accept the questionnaire. Here, we find that only 36% of those initially cooperative people effectively mail back the questionnaire. Together, this results in a total cooperation rate of 32 percent. Compared to average mail survey response rates, this rate is low. Whereas the reasons for these low rates are easily understandable—especially the fact

that respondents are completely unaccounted for once they get a questionnaire in their hands—it cannot be denied that noncooperation is a serious problem in protest surveying.

Apart from increasing the odds of getting problematic refusal bias, the substantial loss of respondents is worrisome to the extent that it increases the costs of the fieldwork and diminishes the return. Researchers need to distribute multiple questionnaires to end up with a relatively small number of units. Aggregating the unsuccessful distribution of the target number of questionnaires due to fieldwork problems (noncontact) and the two kinds of refusal, the following average picture emerges: of the 1,000 aimed-for questionnaires, 209 were not distributed due to noncontact, 75 were not handed out because protestors refused to accept one, 716 were taken home, and 223 were effectively sent back.

Features of demonstrations, especially the demonstrators populating them, are associated with refusal rates. Table 5 presents a small model with delayed refusal rate per demonstration as the dependent variable. Note that we use average features of respondents who *did* respond to predict refusal rates (and thus the noncooperative behavior of refusers).

Table 5. Delayed Refusal. OLS Regression of Delayed Postal Refusal Rate per Demonstration

	Coef.	Std. Err.	Sig.
Age demonstrator (average)	-.499	.183	.009
Education demonstrator (average)	-7.200	2.133	.002
Political efficacy demonstrator (average)	-2.463	1.313	.067
Italy	10.401	4.224	.018
Switzerland	14.632	4.839	.004
United Kingdom	12.830	2.924	.000
Constant	82.875	29.283	.007
N	51		
Prob > Chi ²	.000		
Adj. R ²	.448		

Demonstrations with older demonstrators register lower refusal rates. The refusal rates of events with a mean age of 45 are 10% lower than the refusal rates of demonstrations wherein activists are on average 25 years old. When demonstrators are more educated and consider themselves to be politically efficacious, refusal rates are lower as well. Demonstrations with highly educated participants (i.e., with, on average, a college degree) have a response rate of 40%, while the response rate is 20% for demonstrations with less educated participants (i.e., with, on average, an upper secondary education). These results suggest that skilled and politically aware demonstrators are more willing to go through the intellectually demanding task of filling in a political questionnaire. Although we did not find any significant issue effect on top of the composition of an event, the fact that demonstrations with competent and politically aware participants lead to higher response suggests that issues *do* matter (since issues affect the composition). Again, this challenges the use of the protest survey design for comparing demonstrations across issues.

Departing from the idea that the concrete circumstances in which the data are collected may be correlated with response rates (James and Couper 2002) we tested a series of demonstration features (size, chaotic character, rain, moving), but none of the variables was significant (not reported in table). Matching a recurring finding in the broad survey literature (de Heer 1999), we find systematic differences in delayed refusal across nations with higher refusal in Italy, Switzerland, and the United Kingdom. Comparing protest survey evidence across nations must be done with caution.

Considering an average low response rate of 32%, and systematic refusal rate differences between demonstrations, issues, and countries, can we use protest survey evidence to compare demonstrators turning out for different issues, in different countries, and in different demonstrations? For a long time, the survey literature held that high refusal rates are a sure sign of refusal bias (Armstrong and Overton 1977), but recent work found that the correlation

between refusal rate and refusal bias is weak (Groves and Couper 1998; Groves and Peytcheva 2008; Singer 2006). As such, the low response rates we record for protest surveys are not necessarily problematic for the quality of the data.

To examine delayed refusal bias, we compare the short, screener interviews in subsample B—for which we record very high cooperation levels—with the respondents who sent back their mail-in questionnaire. We test three sociodemographics—age, gender, and education—expecting that the older, higher educated, and female participants would be more willing to collaborate. We incorporate three attitudes—political interest, eagerness to participate in the demonstration, and satisfaction with democracy—expecting that people with more political interest (Voogt and Saris 2003), higher levels of motivation, and more satisfaction with democracy would be more responsive to a political survey. And, we have two measures of political behavior, membership of staging organization and past participation in demonstrations, expecting that members of protest organizations and experienced protesters would be more collaborative (for a list of very similar screener variables see Matsuo, Billiet, Loosveldt, Berglund and Kleven 2010).

We estimate a multilevel logistic regression with postal response (no/yes) of the FTF-interviewed respondents as the dependent variable (table 6). The model clusters participants in demonstrations and considers issues and countries to be situated at the demonstration level. Only respondents who were interviewed *and* who accepted to take home a questionnaire are included.

Three individual-level demonstrator characteristics are associated with delayed refusal, which is in line with previous survey research (Matsuo et al. 2010), especially education and age-effect response. Ph.D. holders (45% response rate) are three times more likely to respond than are demonstrators without formal education (15%). Predicted probabilities for age decrease

Table 6. Delayed Refusal. Multilevel Logistic Tgression with Response (No/Yes) as the Dependent Variable (N-respondents = 3,666; N-demonstrations = 32).

		Coef.	Std. Err.	Sig.
	Constant	51.335	5.633	.000
<i>Respondent</i>	Gender (Female)	.005	.078	.948
	Age (old)	.027	.003	.000
	Education (High)	.228	.028	.000
	Motivation (High)	.132	.044	.003
	Organization Member (Yes)	-.086	.090	.340
	Political Interest (High)	.064	.056	.253
	Democratic Satisfaction (High)	.011	.016	.513
	Past Participation (Frequent)	-.021	.034	.538
<i>Demonstration</i>	Turnout (> 5,000)	.316	.216	.143
	Rain	-.104	.264	.693
	Moving	-.610	.277	.027
	Chaotic	-.284	.242	.241
<i>Issue</i>	Old (ref. = New)	-.195	.187	.297
<i>Country</i>	Netherlands (ref. = Belgium)	-.313	.345	.364
	Italy	-.504	.476	.289
	Spain	.004	.257	.988
	Switzerland	-.574	.395	.146
	Sweden	.236	.432	.585
	UK	-.466	.301	.123
	Wald Chi ² (df)		188.11 (19)	
	Prob > Chi ²		.000	
	Log Likelihood (empty model = -2261.218)		-2,161.037	
	Variance demonstration level (empty model = .240)		.129	
	Total explained variance		46.3%	

from 44% response for respondents born in 1949 to 22% for respondents born in 1989. Motivation to participate (decided early) affects delayed refusal as well, but to a lesser extent. Respondents who decided to participate more than a month before the demonstration are 8% more likely to mail in the postal questionnaire compared to those who decided on the demonstration the day itself. Notwithstanding the high number of observations ($N = 3,666$) none of the other individual-level variables reaches significance.

The features of demonstrations, countries, and issues correlate less with delayed refusal. Questionnaires distributed during moving demonstrations are returned less than during static events (28% vs. 40%). The practicalities of a moving march probably make protestors lose their booklets more frequently. Most importantly, holding the composition of the event constant, delayed refusal differs neither between old vs. new social movement events nor between countries. It is not the issue or country as such, but rather the differential *composition* of events regarding different issues and countries that leads to dissimilarities in refusal.

Table 7. Delayed Refusal Bias. Number of Significant Differences, per Demonstration, between Respondents Who Sent Back the Questionnaire and Those Who Did Not.

	Total No. Demonstrations	No. Significant Differences	% Significant Differences
Gender (female)	51	4	7.8
Age (old)	48	15	31.3
Education (high)	45	10	22.2
Political Interest (high)	51	7	13.7
Motivation (high)	51	8	15.7
Democratic Satisfaction (high)	42	1	2.4
Organizational Membership (yes)	47	4	8.5
Past Participation (frequent)	42	5	11.9

Note: Due to missing values we cannot calculate the bias for all these eight variables for all demonstrations. As a consequence the N for each row is different.

For each of the eight individual-level test variables and for each demonstration, we calculate Independent Samples Mann-Whitney U Tests that compare willing to reluctant respondents. Table 7 presents the aggregate results. Per demonstration, we find a significant difference between cooperative and noncooperative respondents for on average 1.38 of the eight test variables. This relatively small refusal bias at the demonstration level is, of course, affected by the fact that subsample B is, on average, small ($N = 132$).

Table 7 confirms the multilevel estimates. Refusal bias is most frequent with respect to age (31% of the demonstrations) and education (22%). We record intermediary levels of refusal bias for two attitudinal variables—political interest (14%) and motivation (16%). According to the general survey literature, these particular variables are prone to refusal bias (Groves and Peytcheva 2008). There is a low incidence of refusal bias regarding organizational membership (9%), gender (8%), and democratic satisfaction (2%).

An alternative take on the incidence of delayed refusal bias at the demonstration level is, first, calculating a simple additive scale of the number of significant delayed refusal biases per demonstration (0-8). Then, we take this scale as the dependent variable in an OLS model (table 8). With only 32 demonstrations with complete data, we limit the number of reported independent variables (all other variables were tested as well).

Again, the composition of a demonstration—measured through the average features of its participants—correlates with the amount of significant delayed refusal biases registered. The results, once more, confirm that demonstrations populated by politically savvy and confident protestors are less affected by refusal bias. When the average demonstrator has a high political efficacy and when he/she has been mobilized via the mass media (which is an indirect measure

Table 8. Delayed Refusal Bias. OLS Regression Predicting the Number of Significant Biases per Demonstration (0-8)

	Coef.	Std. Err.	Sig.
Media mobilization demonstrator (average)	-1.985	.762	.015
Political efficacy (average)	-.642	.192	.002
Issue: Old (ref. = New)	-.000	.323	.999
Italy	2.484	.917	.012
Switzerland	1.793	.672	.013
Constant	15.786	4.218	.001
N	32		
Prob > F	.001		
Adj. R ²	.448		

of political sophistication), the protest survey design produces fewer biases. The finding underscores that the design is better suited to tackle marches with politically aware citizens. On top of those composition effects, and in line with the results presented in table 5, we record significantly more delayed refusal bias in Italy and Switzerland.

Wrapping up, we record systematic delayed refusal bias in the 51 protest surveys under scrutiny. Relying on screener interviews, which is a method that is prone to find more refusal bias than alternative refusal-bias methods, our findings are in line with the general survey literature: age, education, and motivation are correlated with a bias in response. The few additional effects of attitudinal and political measures suggest that political awareness is associated with higher collaboration. Overall, the individual-level composition of demonstrations is substantially correlated with refusal bias. In contrast, in the models the direct statistical effect of demonstration features, issues, or countries is limited. While at first sight, this may seem to be good news for comparing protest evidence across demonstrations, issues, and countries, one must remain cautious. One should be particularly watchful of issues that are correlated with refusal bias *via* the differential composition of their protests.

DISCUSSION

Noncontact, immediate refusal, and delayed refusal bias in protest surveying correlate with features of interviewers, demonstrators, demonstrations, issues, and countries. Are certain features connected to several response problems at the same time leading to an accumulation of problems for specific demonstrations? Table 9 summarizes our significant findings.

Evidence is scattered. Different blocks of independent variables are associated with different response problems. *Interviewer* features affect noncontact and immediate refusal but are unrelated to delayed refusal bias. Once the questionnaire is distributed, respondents are no longer affected by who delivered it. *Demonstrator* properties, both on the individual level and aggregated on the demonstration level, lead to delayed refusal bias. *Demonstration* features have a dispersed effect on the entire process. For *issues*, we see that old social movement events lead to higher immediate refusal. *Countries* exert no concentrated effects either, with Italy particularly bad for noncontact and Switzerland for immediate refusal. The same countries, though, together with the UK, perform worse for delayed refusal bias as well. These findings are encouraging for the protest survey design. Response problems are not cumulatively reinforcing each other. Correlations between the different response problems with demonstrations as the units of analysis are low (results not shown), confirming the non-cumulativity of the varying response problems.

Table 9. Predictors of Response Problems. Overview of Multivariate Effects.

	Noncontact Losing Pointer	Noncontact Incomplete Coverage	Immediate Refusal FTF Interview	Immediate Refusal Postal Questionnaire	Delayed Refusal Bias
<i>Interviewers</i>					
Age				–	
Gender (female)			–		
Experience (high)	–				
Size interviewer team		–			
<i>Demonstrators</i>					
Age					–
Education					–
Motivation					–
<i>Demonstrators (average)</i>					
Age					–
Education					–
Political efficacy					–
Media mobilization					–
<i>Demonstrations</i>					
Turnout	+			–	
Chaotic			+		
Responsive		–			
Rain					
Moving					–
<i>Issues</i>					
			+	+	
<i>Countries</i>					
Italy	+				+
Switzerland			+	+	+
UK					+

Our evidence suggests that the quality of protest survey data is not very much directly affected by higher-level variables such as demonstrations, issues, and countries. These are precisely the key context variables the comparative protest survey design aims to compare, which reinforces confidence in the protest survey procedure. There are notable demonstration, issue, and country effects on refusal bias (table 7), but they seem to be minor compared to the more pervasive effects of individual-level demonstrator features (aggregated on the demonstration level).

However, there are valid reasons to remain cautious when comparatively employing protest survey data. The direct effect of demonstrations, issues, and countries on refusal bias may be weak and dispersed in comparison to the effect of demonstrator features, but the latter are a *function* of demonstration, issue, and country features. The individual-level composition of demonstrations is affected by the type of demonstration, by the type of issue, and by the country in which the protest takes place. Older, highly educated, motivated, and politically aware people collaborate better; some demonstrations include systematically more people with those characteristics. Yet, assessing differences in the composition of protest events is exactly what the comparative study of protest participation is all about. Hence, there *is* a systematic interaction between the correlates of refusal and the variables of interest to students of political protest.

CONCLUSION

Protest surveys are increasingly used by students of protest. Theoretically, they offer several advantages for investigating who shows up, why, and how. The design's most important asset is that protest is studied in context. The truism that protest does not happen in a vacuum but is affected by the issue at stake, by the type of organization staging the event, by the national context, and by the political situation can be effectively studied by comparing across protest events. Protest surveying allows scholars to get a systematic grasp on the contextual variables affecting protest composition.

This study intended to test the validity of the individual protester evidence generated by the protest survey design. While it is potentially promising, it remains unclear whether the produced evidence can stand a critical quality test. Drawing on evidence covering 51 demonstrations in seven countries, we scrutinized different sources of bias drawing on features of the interviewers, demonstrators, demonstrations, issues, and countries.

We uncovered some structural weaknesses. The disorderly character of demonstrations causes frequent fieldwork problems leading to noncontact. Generally low response rates on the postal survey phase can only be countered by employing large interviewer teams, but that makes for a costly undertaking. Refusal rates largely vary across countries and demonstrations making it nearly impossible to attain fixed amounts of useful interviews. And, most importantly, there are systematic biases in refusal that are related to the context variables in which most protest survey researchers are interested. This finding challenges the comparative use of protest survey evidence.

Indeed, although net differences between countries and issues remained limited, we discovered recurring response rate and response bias problems related to age, education, protest motivation, and political awareness. Since the individual-level composition of a protest event is strongly affected by the protest issue, researchers must be especially cautious when comparing demonstrations on different issues. The demonstrator features leading to high response rate and low refusal bias are typical for events staged by the new social movements in contrast to events staged by the old social movements. Data quality of evidence gathered on new social movement events is on average better than that of evidence collected on old social movement events. Our findings quite strongly suggest that the protest survey design is best suited to cover events of the former instead of events by the latter. When comparing participants across events on diverging issues, researchers should be aware of the fact that the differences found between events may be partly due to biases in the evidence rather than to real differences between the populations at the event. Refusal-bias problems seem to be more limited in regard to comparisons across demonstrations and across nations.

Apart from the thorny problem of comparison across issues, we can conclude that the refusal bias engrained in the protest survey design is, in general, not exceptionally large. The encountered refusal-bias problems are neither different nor worse than the problems affecting any type of survey.

A notable weakness of the design we relied upon for this study is the absence of any data about the immediate refusers—those who declined the oral, screener interview or refused to accept a written questionnaire without being interviewed first. Although this group is relatively small—around ten percent—collecting observational data on these immediate refusers would yield more complete evidence on the entire refusal-bias problem.

Our results suggest at least two practical guidelines for researchers willing to employ the protest survey design. First, it is advisable to rely on female, older, and experienced interviewers. With such interviewers, noncooperation rates are lower and the fieldwork is less afflicted by interviewers losing their pointers. The acquired data are quantitatively more adequate when pointers make better crowd size assessments. Experienced teams make for better estimations and less noncontact.

Second, and more importantly, for cross-issue comparison researchers should consider weighing their protest survey data. If they decide to do so, a good start would be to system-

atically compensate for age and education since these variables generate the most recurrent biases. As far as we know, *none* of the earlier cited, recent studies employing the protest survey design has bothered to use weights. Note that the decision to weigh the data has to be taken cautiously, as weighing may actually cause more problems than that it solves. But since the found response biases in protest surveys are overall pretty small, the weights would be rather modest and the chance that they skew the data are quite small.

Either way, we strongly advocate the use of screener interviews as we have done here, following the procedure proposed by Walgrave and Verhulst (2011). Screener interviews allow researchers to measure when and where refusal bias is problematic and to correct for distortions. As there are a number of response problems that do materialize frequently, constant scrutiny and monitoring are necessary. Even when strict rules are applied by experienced teams of researchers, some eradicable problems remain. If used correctly and cautiously, and as data collected via standardized procedures accumulate, protest surveys may become a valuable instrument to increase our knowledge of who protests, why they protest, and how they were mobilized. Especially, they can heighten our understanding of variations in protest participation across events, issues, and countries.

NOTES

¹ Some of this work dating from the last decade is Adler and Mittelman (2004); Bédoyan, Van Aelst, and Walgrave (2004); Bennett, Breunig, and Givens (2008); Blocq, Klandermans, and Van Stekelenburg (2012); Boekkooi, Klandermans, and Van Stekelenburg (2011); Botetzagias and Boudourides (2004); Chan and Lee (2009); della Porta, Andretta, Mosca, and Reiter (2006); della Porta and Reiter (2012); Diani (2009); Eggert and Giugni (2012); Eggert and Giugni (2015); Fillieule and Blanchard (2010); Fisher, Stanley, Berman, and Neff (2005); Gómez-Román and Sabucedo (2014); Goss (2003); Heaney and Rojas (2007); Ketelaars, Walgrave, and Wouters (2014); Klandermans, Stekelenburg, Damen, Van Troost, and Van Leeuwen (2014); Nomiya (2009); Norris, Walgrave, and Van Aelst (2005); Olcese, Saunders, and Tzavidis (2014); Peterson, Wahlström, Wennerhag, Christancho, and Sabucedo (2012); Rootes and Saunders (2007); Saunders (2014); Saunders, Grasso, Olcese, Rainsford, and Rootes (2012); Stekelenburg (2006); Stekelenburg and Klandermans (2014); Stekelenburg, Klandermans, and Van Dijck (2009); Stekelenburg, Walgrave, Klandermans, and Verhulst (2012); Van Aelst and Walgrave (2001); Van Laer (2010); Verhulst and Walgrave (2009); Wahlström and Wennerhag (2014); Walgrave, Bennett, Van Laer, and Breunig (2011); Walgrave and Rucht (2010); Walgrave and Verhulst (2009); Walgrave and Wouters (2014); Walgrave, Wouters, Van Laer, Verhulst, and Ketelaars (2012); Wilson and Dunn (2011).

APPENDIX

Table A1. Dependent Variables

Variable Name	Question Wording	LEV*	N	Mean	SD	Min	Max
<i>Fieldwork Problems (Noncontact)</i>							
Losing Pointer (0/1)	Did the interviewer lose his/her pointer? (no,yes)	I	673	.40	.49	0	1
Entire crowd covered (0/1)	Did the protest event get entirely covered or not? (no, yes)	D	51	.12	.33	0	1
<i>Refusal rates (Immediate & Delayed)</i>							
Immediate Refusal FTF (%)	How many FTF interviews were refused?	I	673	12.86	15.50	0	100
Immediate Refusal Postal (%)	How many protesters refused to accept an envelope?	I	673	9.97	8.83	0	100
Delayed Refusal (%)	How many of the distributed booklets were not send back?	D	51	64.36	10.20	0	100
Delayed Refusal (0/1)	Did the respondent send his/her postal questionnaire back? (no, yes)	R	5542	.36	.01	0	1
<i>Refusal Bias</i>							
Delayed Refusal Bias	Do the people who send back the postal differ significantly from the people who did not send it back? (8 possible biases)	D	32	1.38	1.16	0	8

Notes: * LEV = Level of Analysis; D = Demonstration level; I = Interviewer level; R = Respondent level

Table A2. Independent variables

Variable Name	Question Wording	LEV*	N	Mean	SD	Min	Max
<i>Countries and Issues</i>							
Belgium	Belgium? (no, yes)	D	51	.14	.348	0	1
Italy	Italy? (no, yes)	D	51	.08	.272	0	1
Netherlands	The Netherlands? (no, yes)	D	51	.22	.415	0	1
Spain	Spain? (no, yes)	D	51	.22	.415	0	1
Sweden	Sweden? (no, yes)	D	51	.12	.325	0	1
Switzerland	Switzerland? (no, yes)	D	51	.06	.238	0	1
UK	United Kingdom? (no, yes)	D	51	.18	.385	0	1
Old	Old issue (1) or new issue? (0)	D	51	.61	.493	0	1
<i>Demonstrator Characteristics (FTF)</i>							
Gender	Male, Female	R	6492	1.47	.50	1	2
Year born	In which year were you born?	R	6312	1969	15.33	1910	1992
Education	Highest educational degree?	R	5781	5.31	1.54	1	8
Motivation	When did you make a firm decision to participate?	R	6690	2.92	1.01	1	4
Organization Member	Are you a member of any of the organizations that organized this demonstration?	R	6303	0.47	0.49	0	1
Political Interest	How interested are you in politics?	R	6688	3.30	0.79	1	4
Democratic Satisfaction	In general, how satisfied or dissatisfied are you with the functioning of democracy in your country?	R	5549	4.21	2.67	0	10
Past Participation	How many times have you in the past taken part in a demonstration?	R	5588	3.25	1.47	1	5
<i>Demonstrator Characteristics (Aggregated)</i>							
Media Mobilization	Which of the above information channels was the most important? (Answers: Radio or Television, Newspapers, Alternative online media, Online Social networks)	D	51	.26	.18	0	1
Age	In which year were you born?	D	51	46.16	6.84	24.53	56.9
Education	What is your highest educational qualification? (scale 0-8)	D	51	5.72	.61	0	8
Political Efficacy	To what extent do you agree or disagree with the following statements? Scale 0-30: Politicians promises; no use of voting; I impact; organization impact; different countries impact; sides of argument.	D	51	21.93	.94	0	30
<i>Demonstration Characteristics</i>							
Interviewer-Protector Interaction (Interviewer)	How would you evaluate the protestors in their interaction with you as interviewer? (scale 1-5)						
	Responsive	I	673	4.22	.698	1	5
	Hostile	I	673	1.79	1.02	1	5
	Friendly	I	673	4.15	.728	1	5
	Open	I	673	4.05	.762	1	5
Demonstration Rainy	Did it rain during the demonstration? (no, yes)	D	51	.157	.367	0	1
Demonstration Moving	Was it a moving demonstration? (no, yes)	D	51	.686	.469	0	1
Turnout	How many people showed up according to you?	D	51	38,989	143,289	350	1,000,000
<i>Fieldwork Problems</i>							
Chaotic	Was the demonstration chaotic or disorganized? (no, yes)	D	51	.18	.385	0	1
Noisy	Was there a lot of noise or music at the demonstration? (no, yes)	D	51	.47	.504	0	1
Dense	Was the demonstration dense, a high concentration of protestors? (no, yes)	D	51	.41	.497	0	1
Overestimated size	Did you overestimate the size of the crowd? (no, yes)	D	51	.22	.415	0	1
Underestimated size	Did you underestimate the size of the crowd? (no, yes)	D	51	.08	.272	0	1
<i>Interviewer Characteristics</i>							
Gender	Are you male or female? (0: male, 1: female)	I	673	.593	.492	0	1
Year born	In which year were you born?	I	673	1986	4.84	1958	1993
Experience	At how many demonstrations have you carried out interviews before?	I	673	.93	1.43	0	11
No. of interviewers	How large was the interview team?	D	51	13.20	3.06	8	20

Note: * LEV = Level of Analysis; D = Demonstration level; I = Interviewer level; R = Respondent level.

Table A3. Covered Demonstrations, Issue, Country and Number of Face-to-face and Postal Interviews

Demonstration	Country	Date	Issue	Demonstrators	Number Postals Distributed	Number Send Back	Number FTF Interviews
Climate Change	Belgium	5/12/09	Climate	15,000	777	334	144
March for Work	Belgium	29/01/10	Austerity	30,000	466	302	138
1 st of May March	Belgium	1/05/10	May Day	2,000	837	216	141
No Government, Great Country	Belgium	23/01/11	Political System	45,000	717	365	122
Nonprofit Demonstration	Belgium	29/03/11	Austerity	15,000	634	200	134
Not in Our Name	Belgium	7/05/11	Pol. System	700	619	202	57
We have alternatives	Belgium	2/12/11	Austerity	70,000	767	169	150
Retirement Demonstration	Netherlands	21/11/09	Austerity	7,000	676	242	113
Climate demonstration	Netherlands	12/12/09	Climate	3,500	662	270	101
Student demonstration (1)	Netherlands	21/05/10	Austerity	2,000	514	161	100
Student demonstration (2)	Netherlands	21/01/11	Austerity	15,000	943	280	183
Culture demo (Amsterdam)	Netherlands	20/11/10	Austerity	15,000	353	174	73
Culture demo (Utrecht)	Netherlands	20/11/10	Austerity	2,500	385	171	57
Together Strong for Public Work	Netherlands	17/02/11	Austerity	8,000	1000	339	188
Stop racism and exclusion	Netherlands	19/03/11	Racism	350	387	124	80
Antinuclear Demonstration	Netherlands	16/04/11	Climate	2,500	855	450	154
Military Demonstration	Netherlands	26/05/11	Austerity	4,000	387	142	164
Stop budget cuts (care and welfare)	Netherlands	19/09/11	Austerity	4,500	940	290	189
Euromayday	Italy	1/05/2011	May Day	5,000	993	127	167
May Day	Italy	1/05/2011	May Day	500	408	110	84
General Strike	Italy	6/05/2011	Austerity	15,000	987	234	188
Marcia Perugia-Assisi	Italy	25/09/11	Peace	150,000	1000	264	195
Against language decree	Spain	21/01/10	Pol. System	40,000	1000	324	191
Against Europe of Capital, Crisis, War	Spain	28/01/10	Austerity	1500	300	77	94
Against Abortion	Spain	7/03/10	Abortion	15,000	871	302	172
1st May, Labor Day	Spain	1/05/10	May Day	8,000	700	180	136
Self-determination is democracy	Spain	12/06/10	Pol. System	10,000	655	301	147
Against the new labor law	Spain	30/06/10	Austerity	10,000	780	168	124
We are a nation, we decide	Spain	12/07/10	Pol. System	1,000,000	980	309	192
Against Labor Law	Spain	29/09/10	Austerity	55,000	900	308	172
For employment, not capital reforms	Spain	1/05/11	Austerity	2,000	340	168	119
Celebration May Day	Spain	1/05/11	May Day	15,000	263	66	47
Real Democracy Now!	Spain	15/05/11	Austerity	25,000	776	350	179
May 1 March, Left Party	Sweden	1/05/10	May Day	4,200	430	167	106
May Day, Social Democratic Party	Sweden	1/05/10	May Day	3,000	429	176	106
Against racist politics	Sweden	4/10/10	Racism	2,800	668	193	131
Antinuclear	Sweden	26/04/11	Climate	1,000	718	283	133
May Day (Left Party)	Sweden	1/05/11	May Day	2,000	388	142	70
May Day (SAP/LO)	Sweden	1/05/11	May Day	900	281	97	47
World March of Women	Switzerland	13/03/10	Women	6,000	420	320	60
May 1st Demonstration	Switzerland	1/05/10	May Day	8,000	861	135	163
Antinuclear Manifestation	Switzerland	22/05/11	Climate	20,000	980	148	154
National Climate March	UK	5/12/09	Climate	50,000	606	243	90
May Day Labour March	UK	1/05/10	May Day	5,000	977	178	116
No to Hate Crime Vigil	UK	23/10/10	Racism	2,000	993	169	115
Unite Against Fascism	UK	6/11/10	Racism	3,000	992	194	182
Fund Our Future: Stop Education Cuts	UK	10/11/10	Austerity	30,000	995	147	157
Take Back Parliament	UK	15/05/10	Pol. System	2,000	991	351	160
Climate March 2010	UK	04/12/10	Climate	1,500	966	360	150
Million Women Rise	UK	5/03/11	Women	3,000	973	178	180
TUCs March for the Alternative	UK	26/03/11	Austerity	250,000	993	214	175
TOTAL				1,988,450	36,533	11,414	6,790

REFERENCES

- Adler, Glenn, and James H. Mittelman. 2004. "Reconstituting 'Common-Sense' Knowledge: Representations of Globalization Protests." *International Relations* 18(2): 189–211.
- Van Aelst, Peter, and Stefaan Walgrave. 2001. "Who Is That (wo)man in the Street? From the Normalisation of Protest to the Normalisation of the Protester." *European Journal of Political Research* 39(4): 461–86.
- Armstrong, Scott, and Terry Overton. 1977. "Estimating Nonresponse Bias in Mail Surveys." *Journal of Marketing Research* 14(3): 396–402.
- Barnes, Samuel, and Max Kaase. 1979. *Political Action, Mass Participation in Five Western Democracies*. London: SAGE-publications.
- Bédoyan, I., Peter Van Aelst, and Stefaan Walgrave. 2004. "Limitations and Possibilities of Transnational Mobilization: The Case of EU Summit Protesters in Brussels, 2001." *Mobilization* 9(1): 39–54.
- Bennett, Lance W., Christian Breunig, and Terri Givens. 2008. "Communication and Political Mobilization: Digital Media and the Organization of Anti-Iraq War Demonstrations in the U.S." *Political Communication* 25(3): 269–89.
- Blocq, D., Bert Klandermans, and Jacquélien Van Stekelenburg. 2012. "Political Embeddedness and the Management of Emotions." *Mobilization* 17(3): 319–34.
- Blom, Annelies, Edith de Leeuw, and Joop Hox. 2010. *Interviewer Effects on Nonresponse in the European Social Survey*. Colchester, Essex: ESRC.
- Boekkooi, Marije, Bert Klandermans, and Jacquélien Van Stekelenburg. 2011. "Quarrelling and Protesting: How Organizers Shape a Demonstration." *Mobilization* 16(2): 498–508.
- Botetzagias, Iosif, and Moses Boudourides. 2004. "The Iraq War and Greek Anti-War Organisations." *Bad Subjects* 65: 15–17.
- Chan, Joseph, and Francis Lee. 2009. "Who Can Mobilize Hong Kong People to Protest? A Survey-Based Study of Three Large-Scale Rallies." Pp. 14–37 in *Politics and Government in Hong Kong*, edited by Ming Sing. Abingdon, Oxon: Routledge.
- Crist, John, and John McCarthy. 1996. "If I Had a Hammer: The Changing Methodological Repertoire of Collective Behavior and Social Movements Research." *Mobilization* 1(1): 87–102.
- Della Porta, Donatella, Massimiliano Andretta, Lorenzo Mosca, and Herbert Reiter. 2006. *Globalization from Below: Transnational Activists and Protest Networks*. Minneapolis: University of Minnesota Press.
- Diani, Mario. 2009. "The Structural Bases of Protest Events Multiple Memberships and Civil Society Networks in the 15 February 2003 Anti-War Demonstrations." *Acta Sociologica* 52(1): 63–83.
- Diani, Mario, and Doug McAdam. 2003. *Social Movements and Network. Relational Approaches to Collective Action*. Oxford: Oxford University Press.
- Dillman, Don. 1991. "The Design and Administration of Mail Surveys." *Annual Review of Sociology* 17: 225–49.
- Dillman, Don, John Eltinge, Robert Groves, and Roderick Little. 2002. "Survey Nonresponse in Design, Data Collection, and Analysis." Pp. 3–26 in *Survey Nonresponse*, edited by Don Dillman, John Eltinge, Robert Groves, and Roderick Little. New York: Wiley-Interscience.
- Eggert, Nina, and Marco Giugni. 2012. "Homogenizing 'Old' and 'New' Social Movements: A Comparison of Participants in May Day and Climate Change Demonstrations." *Mobilization* 17(3): 335–48.
- Eggert, Nina, and Marco Giugni. 2015. "Does the Class Cleavage Still Matter? The Social Composition of Participants in Demonstrations Addressing Redistributive and Cultural Issues in Three Countries." *International Sociology* 30(1): 21–38.
- Fillieule, Olivier, and Philippe Blanchard. 2010. "Individual Surveys in Rallies (INSURA). A New Tool for Exploring Transnational Activism?" Pp. 186–210 in *The Transnational Condition: Protest Dynamics in an Entangled Europe*, edited by Simon Teune. New York: Berghahn Books.
- Fisher, Dana, Kevin Stanley, David Berman, and Gina Neff. 2005. "How Do Organizations Matter? Mobilization and Support for Participants at Five Globalization Protests." *Social Problems* 52(1): 102–121.
- Giddens, Anthony. 1987. *Social Theory and Modern Sociology*. Cambridge: Polity Press.
- Gómez-Román, Cristina, and José-Manuel Sabucedo. 2014. "The Importance of Political Context: Motives to Participate in a Protest before and after the Labor Reform in Spain." *International Sociology* 29(6): 546–64.
- Goss, Kristin A. 2003. "Rethinking the Political Participation Paradigm." *Women & Politics* 25(4): 83–118.

- Groves, Robert, and Mick Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: John Wiley.
- Groves, Robert, and Emilia Peytcheva. 2008. "The Impact of Nonresponse Rates on Nonresponse Bias. A Meta-Analysis." *Public Opinion Quarterly* 72(2): 167–189.
- Groves, Robert, Stanley Presser, and Sarah Dipko. 2004. "The Role of Topic Interest and in Survey Participation Decisions." *Public Opinion Quarterly* 68(1): 2–31.
- Heaney, Michael T., and Fabio Rojas. 2007. "Partisans, Nonpartisans, and the Antiwar Movement in the United States." *American Politics Research* 35(4): 431–64.
- De Heer, Wim. 1999. "International Response Trends: Results of an International Survey." *Journal of Official Statistics* 15(2): 129–142.
- Hox, Joop, and Edith De Leeuw. 1994. "A Comparison of Nonresponse in Mail, Telephone, and Face-to-Face Surveys. Applying Multilevel Modeling to Meta-Analysis." *Quality & Quantity* 28(4): 329–44.
- James, Lepkowski, and Mick Couper. 2002. "Nonresponse in the Second Wave of Longitudinal Household Surveys." Pp. 259–279 in *Survey Nonresponse*, edited by Robert Groves, Don Dillman, John Eltinge, and Roderick Little. New York: Wiley.
- Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert Groves, and Stanley Presser. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." *Public Opinion Quarterly* 64(2): 125–148.
- Ketelaars, Pauline, Stefaan Walgrave, and Ruud Wouters. 2014. "Degrees of Frame Alignment: Comparing Organisers' and Participants' Frames in 29 Demonstrations in Three Countries." *International Sociology* 29(6): 504–24.
- Klandermans, Bert. 1997. *The Social Psychology of Protest*. Oxford: Blackwell Publishers.
- Klandermans, Bert, Jacquelin van Stekelenburg, Marie-Louise Damen, Dunya van Troost, and Anouk van Leeuwen. 2014. "Mobilization Without Organization: The Case of Unaffiliated Demonstrators." *European Sociological Review* 30(6): 702–16.
- Klandermans, Bert, and Jacky Smith. 2002. "Survey Research: A Case for Comparative Designs". Pp. 3–31 in *Methods of Social Movement Research*, edited by Bert Klandermans and Suzanne Staggenborg. Minnesota: University of Minnesota Press.
- Van Laer, Jeroen. 2010. "Activists Online and Offline: The Internet as an Information Channel for Protest Demonstrations." *Mobilization* 15(3): 347–66.
- De Leeuw, Edith, and Wim de Heer. 2002. "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison." Pp. 41–54 in *Survey Nonresponse*, edited by Robert Groves, Don Dillman, John Eltinge, and Roderick Little. New York: Wiley.
- Lynn, Peter, and Paul Clarke. 2002. "Separating Refusal Bias and Non-Contact Bias: Evidence from UK National Surveys." *The Statistician* 51(3): 319–33.
- Matsuo, Hideko, Jaak Billiet, Geert Loosveldt, Frode Berglund, and Øyvén Kleven. 2010. "Measurement and Adjustment of Non-Response Bias Based on Non-Response Surveys: The Case of Belgium and Norway in the European Social Survey Round 3." *Survey Research Methods* 4(3): 165–178.
- Nomiya, Daishiro. 2009. "Under a Global Mask: Family Narratives and Local Memory in a Global Social Movement in Japan." *Societies Without Borders* 4(2): 117–140.
- Norris, Pippa, Stefaan Walgrave, and Peter Van Aelst. 2005. "Who Demonstrates? Anti-State Rebels or Conventional Participants? Or Everyone?" *Comparative Politics* 2(37): 251–75.
- Olcese, Cristiana, Clare Saunders, and Nikos Tzavidis. 2014. "In the Streets with a Degree: How Political Generations, Educational Attainment and Student Status Affect Engagement in Protest Politics." *International Sociology* 29(6): 525–45.
- Parkin, Frank. 1968. *Middle Class Radicalism: The Social Bases of the British Campaign for Nuclear Disarmament*. New York: Praeger.
- Peterson, Abby, Mattias Wahlström, Magnus Wennerhag, Camilo Christancho, and José-Manuel Sabucedo. 2012. "May Day Demonstrations in Five European Countries." *Mobilization* 17(3): 281–300.
- Della Porta, Donatella, and Herbert Reiter. 2012. "Desperately Seeking Politics: Political Attitudes of Participants in Three Demonstrations for Worker's Rights in Italy." *Mobilization* 17(3): 349–361.
- Rabe-Hesketh, Sophia, and Anders Skrondal. 2008. *Multilevel and Longitudinal Modeling Using Stata*. College Station, TX: Stata Press Publication.
- Rootes, Chris, and Clare Saunders. 2007. "Demonstrations of Democracy. The Make Poverty History and Stop Climate Chaos Marches." Paper presented at ECPR conference, May 7–12, Helsinki.
- Rüdiger, Wolfgang. 2010. "Assessing Nonresponse Bias in Activist Surveys." *Quality & Quantity* 44: 173–180.

- Saunders, Clare. 2014. "Anti-Politics in Action? Measurement Dilemmas in the Study of Unconventional Political Participation." *Political Research Quarterly* 67(3): 574-588.
- Saunders, Clare, Maria Grasso, Cristiana Olcese, Emily Rainsford, and Chris Rootes. 2012. "Explaining Differential Protest Participation: Novices, Returners, Repeaters, and Stalwarts." *Mobilization* 17(3): 263-80.
- Singer, Eleanor. 2006. "Introduction: Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70(5): 637-45.
- Stekelenburg, Jacquélien van. 2006. "Promoting or Preventing Social Change?: Instrumentality, Identity, Ideology and Group-Based Anger as Motives of Protest Participation." Doctoral dissertation, Department of Social Psychology, VU-University Amsterdam.
- Stekelenburg, Jacquélien van, and Bert Klandermans. 2014. "Fitting Demand and Supply: How Identification Brings Appeals and Motives Together." *Social Movement Studies* 13(2): 179-203.
- Stekelenburg, Jacquélien van, Bert Klandermans, and Wilco W. Van Dijk. 2009. "Context Matters: Explaining How and Why Mobilizing Context Influences Motivational Dynamics." *Journal of Social Issues* 65(4): 815-38.
- Stekelenburg, Jacquélien van, Stefaan Walgrave, Bert Klandermans, and Joris Verhulst. 2012. "Contextualizing Contestation: Framework, Design, and Data." *Mobilization* 17(3): 249-62.
- Stoop, Ineke, Hideko Matsuo, Achim Koch, and Jaak Billiet. 2010. "Paradata in the European Social Survey: Studying Nonresponse and Adjusting for Bias." Paper presented at conference of the American Statistical Association, July 31-August 5, Vancouver.
- Verba, Sydney, Kay Schlozman, and Henri Brady. 1995. *Voice and Equality: Civic Voluntarism in American Politics*. Cambridge: Harvard University Press.
- Verhulst, Joris, and Stefaan Walgrave. 2009. "The First Time Is the Hardest? A Cross-National and Cross-Issue Comparison of First-Time Protest Participants." *Political Behavior* 31(3): 455-84.
- Voogt, Robert, and Willem Saris. 2003. "To Participate or Not to Participate. The Link Between Survey Participation, Electoral Participation, and Political Interest." *Political Analysis* 11(2): 164-179.
- Wahlström, Mattias, and Magnus Wennerhag. 2014. "Alone in the Crowd: Lone Protesters in Western European Demonstrations." *International Sociology* 29(6): 565-83.
- Walgrave, Stefaan, Lance Bennett, Jeroen Van Laer, and Christian Breunig. 2011. "Multiple Engagements and Network Bridging in Contentious Politics: Digital Media Use of Protest Participants." *Mobilization* 16(3): 325-50.
- Walgrave, Stefaan, and Dieter Rucht. 2010. *The World Says No to War: Demonstrations against the War on Iraq*. Minneapolis: Minnesota University Press.
- Walgrave, Stefaan, and Joris Verhulst. 2009. "Government Stance and Internal Diversity of Protest: A Comparative Study of Protest against the War in Iraq in Eight Countries." *Social Forces* 87(3): 1355-87.
- Walgrave, Stefaan, and Joris Verhulst. 2011. "Selection and Response Bias in Protest Surveys." *Mobilization* 16(2): 203-22.
- Walgrave, Stefaan, and Ruud Wouters. 2014. "The Missing Link in the Diffusion of Protest: Asking Others." *American Journal of Sociology* 119(6): 1670-1709.
- Walgrave, Stefaan, Ruud Wouters, Jeroen Van Laer, Joris Verhulst, and Pauline Ketelaars. 2012. "Transnational Collective Identification: May Day and Climate Change Protesters' Identification with Similar Protest Events in Other Countries." *Mobilization* 17(3): 301-17.
- Wilson, Christopher, and Alexandra Dunn. 2011. "Digital Media in the Egyptian Revolution: Descriptive Analysis from the Tahrir Data Sets." *International Journal of Communication* 5:1248-1272.