Survey Research Methods (2017)
Vol. 11, No. 3, pp. 345-360
doi:10.18148/srm/2017.v11i3.6301

ISSN 1864-3361
http://www.surveymethods.org

Using Google Street View to validate interviewer observations and predict nonresponse: a Belgian case study

Anina Vercruyssen
Centre for Sociological Research
University of Leuven, Belgium

Geert Loosveldt
Centre for Sociological Research
University of Leuven, Belgium

Researchers have been looking for easily accessible and easily useable paradata and auxiliary data to improve survey data. Recently, attention has also been paid to the evaluation and validation of this external data, such as the assessment of the quality of interviewer-generated paradata. For these purposes, we investigated how useful Google Street View™ can be with regard to providing auxiliary data, and whether it allows us to assess the quality of interviewer observations on the homes and neighbourhoods of sample units. Although it is relatively simple to use Google Street View in daily life, using it to code auxiliary data for surveys proves to be more challenging than expected. Hence, this paper also offers a thorough discussion of the pitfalls of coding this auxiliary data, as well as the currently available solutions to these problems.

Keywords: auxiliary data; paradata; interviewer observations; survey nonresponse; Google Street View

1 Introduction

In the last decade, there has been a strong focus on finding and using promising types of auxiliary data to inform us about the survey data collection process in general, and about survey nonresponse in particular (for an overview see e.g. Kreuter, 2013). Interviewer observations are a type of paradata that has received a significant amount of attention. In the case of face-to-face interviews, several interviewer observations on the living conditions of the sample units are recorded on contact forms. This is standard procedure in, for example, the European Social Survey (ESS). Although the paradata from these contact forms is certainly valuable, questions have arisen about the degree of accuracy of interviewer-observed paradata, and how to evaluate this (e.g. Sinibaldi, Durrant, & Kreuter, 2013). Mostly, it is simply taken on faith that the interviewers are properly trained and honest in providing valid (para)data. Double-checking every interviewer and his/her (para)data would be a costly way to validate survey (para)data as, in essence, it would mean carrying out the fieldwork twice.

With regard to the accuracy of the interviewer observations about the home and neighbourhood of the sample units – which are required in the ESS – we might have a simpler means for double-checking and validating than physically sending someone to confirm the observations. Using Google Street View™ (GSV), we can look around an area without the need to be physically present. The main objective of this article is to test how easily GSV can be used to generate auxiliary data for surveys, and whether we can use this data to validate the interviewer-observed paradata. In addition, we compare the capacity of these two types of data for non-response analysis of the Belgian sample of the European Social Survey Round 7 (European Social Survey Round 7 Data, 2014).

Contact information: FirstName Name, Affiliation, Postal addres (email)

2 Interviewer-observed paradata and auxiliary data from GSV

In this study, we focus on two sources of external data that can be linked to the survey-data: data concerning the survey data collection process (paradata), in the form of interviewer observations, and auxiliary data generated by using GSV. Interviewer observations to assess the neighbourhood recorded in the ESS are the type and physical condition of a home, the presence of impediments to access, such as intercoms and locked gates, and the presence of litter/rubbish and vandalism/graffiti. These variables are known correlates of unit nonresponse in surveys. Sample units living in multi-unit buildings such as apartments, and in buildings with intercoms and locked gates are harder to contact (see e.g. Campanelli, Sturgis, & Purdon, 1997; Groves & Couper, 1998; Stoop, 2005). Lower response propensities are also observed for sample units living in homes in poor physical condition and those living in "bad" neighbourhoods (see e.g. Beullens, 2013; Billiet, Vehovar, Beullens, & Matsuo, 2009; Blom, de Leeuw, & Hox, 2011; Durrant, D'Arrigo, & Steele, 2013;

Durrant & Steele, 2009; Lynn, 2003). In general, one can also assume that people with different lifestyles live in different types of homes and neighbourhoods, and that this information is substantive, relevant and related to the topics in the ESS questionnaire. This type of auxiliary data, which is related to both response behaviour and substantive variables, is very useful in the analysis of nonresponse and the application of effective weighting procedures. However, the relevance of this type of data and the link between lifestyle, and home and neighbourhood characteristics, is not necessarily the same in every area and country, and can be contextual.

Recording home and neighbourhood characteristics can be considered as an additional task for interviewers, and one cannot simply assume that this data is accurate. It is possible that interviewers are not sufficiently motivated to carry out this task, and that the observations are not made on the spot before the interview. Errors can also occur because different interviewers might assess similar conditions in a different way. Some efforts have already been made to try and assess the accuracy of this type of paradata, although to date only a limited number of relevant studies have been published in peer-reviewed sources (West and Kreuter, 2013, for an overview, see; also see Pickering, Thomas, and Lynn, 2003; Sinibaldi et al., 2013; Walsh, Dahlhamer, and Bates, 2013; West, 2013). When cross-referencing interviewer observations with auxiliary census data, Sinibaldi et al. (2013) found a high degree of accuracy in interviewer observations about the type of home (97%) and the ethnic background of the household members (98%). Most of the discrepancies in the type of home were found for noncontacted sample units. Accurately judging home ownership was apparently more challenging, with 87% agreement found by Sinibaldi et al. (2013) and 46% to 89% – depending on the type of dwelling – by Pickering et al. (2003). Correctly judging the employment status of the contacted sample units and whether there were children in the household also seemed more difficult, as shown in the study by Sinibaldi et al. (2013). West and Kreuter (2013) also found that interviewers" judgements on the presence of children regularly yield false positives and false negatives.

Studies that use GSV in a straightforward way to cross-reference and validate interviewer observations are rare in peer-reviewed studies. In the literature, using Google applications to collect auxiliary data in order to check the quality of paradata so far seems limited to evaluating the data collected by the people who list the sample frames (Eckman & Kreuter, 2013). However, GSV has already been used to cross-reference neighbourhood indicators in health and epidemiological studies (Clarke, Ailshire, Melendez, Bader, & Morenoff, 2010; Odgers, Caspi, Bates, Sampson, & Moffitt, 2012; Rundle, Bader, Richards, Neckerman, & Teitler, 2011) and geographical studies such as neighbourhood audits (e.g. Curtis, Curtis, Mapes, Szell, & Cinderich, 2013; Hara et al., 2015; Less et al., 2015). Notice that, in some of these studies, GSV was used to assess the presence of graffiti, rubbish and litter. These are the same neighbourhood characteristics as observed by the interviewers in the ESS. To date, the concordance levels between field observations and GSV images in existing studies have been relatively limited. For example, Rundle et al. (2011) found that having at least 80% correlation between the two types of information only occurred for 54.3% of the neighbourhood items. However, it must be noted that their study included neighbourhood features that are prone to change quickly over time, such as the number of pedestrians. Clarke et al. (2010) found a high level of agreement for the type of home (0.73 to 0.97 across different types) and high concordance for the absence of graffiti (0.80), but low correlations for observations of rubbish/litter/broken glass (0.35). The presence of the latter may also change somewhat quickly over time, of course. Spotting small features – such as small pieces of litter – and tracking temporarily variable features, which can also be the case for litter, is still less effective using GSV than with a real audit of the neighbourhood (for overviews, see Curtis et al., 2013; Vandeviver, 2014).

The abovementioned studies do show that spotting graffiti, garbage, observing the physical condition of the property, etc. using GSV is clearly not new in (social) research. However, to date there is little information on how GSV images can be coded into useable auxiliary data for surveys. The current study pays extra attention to this process, before setting out to explore how useful GSV data can be as auxiliary data to validate interviewer observations and to predict nonresponse in the Belgian data for the ESS Round 7 (2014). If the coded GSV data contains the same information as the interviewer observations, consideration can be given to reducing the interviewers" tasks and relying on the GSV software instead.

3 Data and Methods

3.1 Data

For this study, we use the interviewer-observed paradata variables from the contact forms of the ESS concerning the type and condition of homes and the condition of the neighbourhood. The same variables were constructed using auxiliary data from GSV.

The ESS is a biannual effort by about 20 European countries to "check the pulse" of Europeans" social attitudes and behaviour[1]. The fieldwork for the seventh round of the ESS for Belgium took place in the autumn of 2014 (see Barbier, Wuyts, Italiano, & Loosveldt, 2016). A stratified random sample of 3,204 sample units was drawn from the national register. To explore the usefulness of GSV data, we took a 20% random sample (N = 640) from the total ESS7BE

[1] www.europeansocialsurvey.org

sample. This subsample is a stratified sample, with the same distribution as the main sample with regard to the final disposition codes (see Table 1). We used anonymized address lists to code the GSV data. No other variables were shown in this file to enable blind coding, thereby avoiding any potential influence by knowledge of the final disposition codes or the actual interviewer observations. This can be considered as a general instruction for GSV coding.

The GSV auxiliary data was obtained through maps.google.be. In 2007, Google launched Street View, providing free access to high-resolution 3D-imaging of streets, as if you were physically present. Many European countries, including Belgium, have a high coverage of streets and the facades of homes. Using GSV to provide auxiliary data to evaluate the dwelling and neighbourhood of the sample units has some clear advantages: 1) the coding can be done in a controlled setting and can be closely supervised; 2) GSV allows us to view a home and its neighbourhood without requiring physical presence, which saves time and money compared with sending a second person into the field to double-check the observations; 3) GSV is freely available to anyone who has an internet connection in a censorship-free country. Google Maps provides several application programming interfaces (APIs) to retrieve data from Google Maps, and allow for either simple use or extensive customization[2].

The cost of employing someone to do the actual coding needs to be considered, however. This is especially relevant, as we discovered that using GSV to create supplementary data equivalent to interviewer observations is not as easy and straightforward as expected (see Section 4.1. regarding the pitfalls).

Table 1
Response and nonresponse rates for the ESS7BE and the GSV 20% subsample

	ESS 7, BE		GSV sample	
	N	%	*N*	%
Interview	1769	55.2	354	55.3
Noncontact	172	5.4	34	5.3
Refusal	837	26.1	168	26.3
Other nonresponse	324	10.1	64	10.0
Ineligible	102	3.2	20	3.1
Total	3204	100.0	640	100.0

3.2 Variables

ESS7 interviewer observations and contact forms. For each sampling unit, one contact form should be filled out. As is recurrently the case in the ESS, whenever an address is visited for the first time, interviewers are required to record in the contact forms (see Appendix B) the type of home, the presence of any impediments to access (intercom, locked door/gate or both), the physical condition of the property (5-point scale from "very good" to "very bad"), the presence of vandalism/graffiti and the presence of litter/rubbish (4-point scale from "very large amount" to "none or almost none"). The formulation of the questions and answer categories may have been changed slightly over the years, but the use of contact forms has always been, and will remain, a fixed part of the ESS methodology. During the interviewer briefing, there is a special part with instructions about how to identify different types of homes and impediments to access, and how to assess the neighbourhood characteristics. One example of the instructions is that interviewers should consider only the space in front of the building (e.g. the house or entire apartment block) plus about fifteen metres to each side.

The contact form should also indicate who the respondent is, when and by whom contact attempts were made, and what the result of each contact attempt was. This information is used to monitor fieldwork activities, calculate accurate response rates and assess different types of respondents (noncontacts, refusals and ineligibles). In our analysis, we focus on whether the interviewer observations can help us predict each case's final outcome: interview obtained, contact, refusal (Table 1), as defined by the AAPOR guidelines (American Association for Public Opinion Research, 2009). Unfortunately, there were too few "ineligible" cases in our subsample to predict ineligibility in the regression models.

The interviewer observations are available for all sample units. Our stratified subsample contains interviewer observations from 126 different interviewers, each with one to eight sample units.

Google Street View observations. GSV is accessed by entering the address into Google Maps. Whereas the interviewer-observed paradata comes from 126 different interviewers, we had one single coder who first recorded how each address was found in GSV. Several pitfalls were encountered, as described in detail in Section 4.1 below. The month and year of the GSV images, available in the left upper corner of the GSV-screen, were also coded. We define images as being outdated if they are from before the year of the fieldwork for the ESS7BE, which took place in 2014. Once the home had been identified, the same instructions and coding scheme were followed as given to the interviewers to describe the property and neighbourhood (see Appendix B).

3.3 Methods

The use of GSV to create auxiliary data starts with the coding of the visual information. We first report the perceived problems during the search and coding process including the success in finding each of the 640 selected ad-

[2]https://developers.google.com/maps/faq#whatis

dresses. The optimal case is when the address can be confirmed visually by spotting the house number. However, alternative strategies needed to be used when this was not possible, either by using circumstantial information (e.g. counting back from a readable neighboring house number) or by using other functions of Google Maps (e.g. Earth View showing the house from above instead of Street View showing the facade) (section 4.1). In the next step, we calculate the concordance of the GSV data with interviewer observations of the same characteristics (section 4.2). The level of agreement between both types of observations can be considered as an assessment of the concurrent validity of this information. Lastly, we evaluate the ability of the GSV data to predict nonresponse error analysis (section 4.3). For these analyses, we specify logistic regression models with response, contacts and refusals as independent variables and the information from GSV or interviewer observation as independent variables.

4 Results

Before we could start the analysis, we needed to overcome a few challenges during the coding of the GSV observations. We describe these problems below, together with possible solutions.

4.1 Locating addresses using GSV and pitfalls of coding GSV data and possible solutions

Using Google Maps and GSV may be simple in daily life, but it is more challenging to use them to generate auxiliary data. Entering an address in Google Maps does not guarantee finding the house (easily) in GSV. Visual confirmation of the exact house number or that of a direct neighbour was only possible in 399 of the cases (62.3%; see Table 2, row "total of exact numbers"). Some 23.6% of the properties (151 homes) were found in more indirect ways, and for 14.1% (90 homes), we could not find the property or did not have useable images in GSV. For 38 of these, the type of home was assessed using the aerial view in Google Earth, bringing the total to 588 (= 399 + 151 + 38) out of 640 cases (92%) for which we could determine the type of property (Table 3). For the vast majority of these, we could also observe the state of the home and the presence of litter or vandalism. Spotting impediments to access, however, was only possible for 76% of the 588 identified homes (Table 3, column: "total N") due to the problem of pixellation.

Finding an address in GSV does not necessarily mean all the observations can be made. We discuss the pitfalls we encountered and our (current) solutions:

Coding auxiliary data from GSV is more time consuming than expected. Each address needs to be entered and searched for individually. On average, twelve homes were coded per hour, but this also includes time for measuring geographical distances to public transportation and motorways. No relevant relationships were found between these distances and the survey outcomes, therefore the focus of this study remains on the visual GSV data. In the case of the ESS7BE, which includes 3204 sample units, coding the whole sample would take one coder about 33 working days of eight hours, which equates to more than six regular working weeks. Without measuring the geographical distances, this would be likely to take a little more than half of this amount of time, still requiring a minimum of 16 full working days or more. Hence, the coding process is more time consuming than expected, although using GSV is still significantly faster and cheaper than having to physically send a second person into the field.

Solution. Strictly focussing on cross-checking the interviewer-generated paradata for the home and neighbourhoods without measuring distances can reduce the time needed by about two to three minutes per case. We were not able to use algorithms for entering addresses in GSV automatically. Such a technological feature would obviously also decrease the time needed for finding the homes to some degree. Computerized street number recognition might also be helpful. However, for GSV the current assessment of street numbers with such software indicates a coverage of 83% at 99% accuracy and 89% coverage at 98% accuracy (Goodfellow, Bulatov, Ibarz, Arnoud, & Shet, 2014), indicating that this technology is still developing. Manual assessment of the images remains necessary.

GSV can be completely unavailable for some streets. Although GSV has extremely high coverage in Belgium, images for a few streets in more remote villages are currently not available. GSV data was not available for 44 cases: in 31 cases the whole street was unavailable, in the other 13 only a part of the street was missing.

Solution. In some cases, it was still possible to see homes from a distance with enough detail to continue coding. In other cases, we were unable to see anything of a home with GSV, but switching to Google Earth view provided some information, for example, enabling us to determine the type of property at the suggested location. For 16 of the 44 cases it was not possible to determine the house type using Google Earth. Currently, the only alternative would be to send somebody to the location, but the problem of unavailable (parts of) streets is likely to gradually be resolved with updates to GSV.

GSV can be off by a few houses. When GSV is available for the address entered in Google Maps, it is possible to end up a few houses away or facing the wrong side of the street. In 38 cases, we were a few houses off and in 44 cases we simply did not have enough information to determine the exact or approximate location of the house using GSV (Table 2, rows "exact but further" and "not found").

Solution:. the 3D rendering in GSV offers a 360ř panoramic view and a limited vertical tilt to the imagery. If

Table 2
Availability and year of images in the GSV 20% subsample of the ESS7BE

	All years		N in Single years						
	N	%	2009	2010	2011	2012	2013	2014	Missing
Exact house number visually confirmed									
Exact number	247	61.9	85	22	1	0	74	65	0
Exact but further	38	9.5	12	2	0	0	17	7	0
Direct neighbour	114	28.6	38	11	1	2	38	24	0
Total of exact numbers	399	100.0	135	35	2	2	129	96	0
% of 640	-	62.3	21.1	5.5	0.3	0.3	20.2	15.0	0
Exact house number not visible									
Further neighbour	40	26.5	18	7	0	0	8	7	0
Commercial/ public building	18	11.9	4	2	0	0	8	4	0
Strong similarities	75	49.7	34	6	0	0	25	10	0
Physical barriers	18	11.9	6	0	0	0	6	6	0
Total 4 less exact categories	151	100.0	62	15	0	0	47	27	0
% of 640	-	23.6	9.7	2.3	0	0	7.3	4.2	0
House not visible									
Censored	2	2.2	2	0	0	0	0	0	0
Not found	44	48.9	33	6	0	0	3	0	2
No GSV	44	48.9	11	2	0	0	0	0	31
Total 3 categories no GSV	90	100.0	46	8	0	0	3	0	33
% of 640	-	14.1	7.2	1.3	0	0	0.5	0	5.2
Total									
N	640	-	243	58	2	2	179	123	33
% of 640	-	100	38.0	9.1	0.3	0.3	28.0	19.2	5.2

Table 3
Concordance between interviewer observations and GSV observations in the ESS7BE

Observations	Exact match	Lenient match	Extra lenient	Only 1 wrong	Total N	Missing
Type of house	70.9	78.7	84.2	-	588	8.1
Condition of house	44.6	78.3	-	-	561	12.3
Litter	75.2	76.7	-	-	576	10.0
Vandalism	87.1	88.0	-	-	575	10.1
Impediments to access	43.4	83.9	-	-	447	30.2
First 4 observations	46.6	-	-	82.9	543	-
All 5 observations	38.9	-	-	74.6	445	-

Exact match: exactly the same coding; Lenient match: similar coding at a more general level; Extra lenient: correcting for potential visual difficulties in identifying the type of home in GSV; Missing values are calculated on $N = 640$

the house numbers are readable, it is relatively easy to turn around or move further down the street to find the exact address. If the house numbers are not readable, more elaborate efforts need to be made to find the property (see pitfall 5).

GSV images are restricted by physical barriers. GSV camera cars can only go as far as the roads allow them to. This means that some car-free areas need to be covered using portable cameras. Although Google sends people with portable cameras to no-car zones, much of this work remains undone. This means that sometimes GSV can only provide a (limited) view from a distance, as was the case for 18 addresses in our subsample (Table 2, row "physical barriers"). These were apartment blocks with a park or large area of grass around them, apartments located in a private inner court, and houses with tall hedges. It may be possible to obtain an overall impression of the type of building and its location, but obtaining detailed information, such as being able to see if there is a locked front door with an intercom, is likely to be almost impossible.

Solution. If the address is located in a private inner court or behind tall hedges, GSV unfortunately hits a barrier again with regard to its usefulness for double-checking interviewer-observed auxiliary data. In most cases, Google Earth still allowed us to assess what type of building was at the location from an aerial view. This may also give an indication about the type of neighbourhood. However, for gathering data about the condition of the home and the neighbourhood, a double-check can only be carried out by sending someone there.

GSV images can be too low resolution, leaving smaller features too blurred to identify. Issues concerning spotting small details are frequently mentioned in studies using GSV (for overviews, see Curtis et al., 2013; Vandeviver, 2014). On many occasions GSV offers close-up imagery of homes, but even then it may still not be possible to decipher the house number because it is too small or even absent. In the cases where images are relatively low resolution, it is often impossible to identify intercoms or spot multiple doorbells that indicate a multi-unit dwelling. GSV getting too close to homes can also lead to problems with the resolution, because parts of the images then end up in the blurred rendering zone immediately around the GSV camera car.

Solution. This problem may be resolved in the future with technological improvements and updates of the GSV images. However, at present we need to put some extra effort into gathering GSV data. We were only able to find 285 (Table 2, rows "exact number" + "exact but further" = $247 + 38$) out of the 640 homes by viewing the actual house number and an additional114 by a readable house number for a directly neighbouring property (Table 2, row: "direct neighbour"). In 40 cases, we needed to "go the extra mile" in GSV, both literally and figuratively, before finding a readable house number and then counting back from there to find the right property (Table 2, row "further neighbour"). In some other cases, we looked for public or commercial buildings in the street that can be "Googled" to find their exact address, and then counted back from there (18 cases; Table 2, row "commercial/public building"). This counting back approach also requires additional information about the street numbering: identification of the odd and even side, finding out if all the house numbers exist (for example, numbers can be become redundant after two small houses are merged into one) and if all the houses are individual (for example, indexes for the same house in the sample list may hint at multi-unit buildings with multiple numbers at one location). The last two, however, are hard to determine.

When the abovementioned strategies proved unsuccessful, we were often able to narrow it down to a few potential buildings with a high degree of similarity. In this way, we coded 75 cases, as identical codes would have been given to all the potential homes (Table 2, row "strong similarities"). Unfortunately, for 44 other cases this approach did still not provide enough information to locate the exact home (Table 2, row "no GSV"). In all the cases where there was any doubt, the variable was coded as missing. The missing code was most frequently used due to inability to determine impediments to access (30.2% missing values). Only 447 out of 640 cases had sufficient visibility for coding (see Table 3). Spotting intercoms in particular was often hampered because the doorbell was too pixellated in the images or because intercoms tend to be located behind the first entry door, out of reach of GSV.

GSV can be censored. Images of homes can be deliberately pixellated. People have the right to demand that their houses are blurred on GSV to respect their privacy (Google, 2015a). We only encountered this twice in our subsample (Table 2, row "censored"). Although we do not have exact numbers, the actual subsample gives us the impression that Belgian people rarely request this privacy protection.

Solution. Coding houses with censored GSV images will normally be impossible. Sometimes a censored home might be part of a street full of almost identical properties. This exceptional case would allow coding based on strong similarities with the neighbouring houses. However, to obtain an accurate second assessment of the type and condition of the home, we would need to send someone to the actual address. Google Earth view can help in determining the type of property and it would still be possible to code the neighbourhood observations using GSV.

GSV can be out of date. Homes can be demolished, rebuilt, renovated or altered. Neighbourhoods can be transformed through gentrification, the change of designated land use (for example, turning farmland into housing estates), etc. Although GSV is regularly updated, we only had up-to-date images from the year of the fieldwork (2014) in 19.2% of the cases (Table 2, row "total"). Often, 2009 was the most recent year (38.0%, Table 2). These older images can be problematic, as we may end up coding an incorrect (outdated) condition for the property and neighbourhood, as this might be from a time when it had other occupants who subsequently moved out. Moreover, 15 properties did not exist or were under construction at the time the GSV pictures were taken. Hence, extra attention will need to be paid to the age of the images in analyses.

Solution. GSV does not provide the option to look at previous versions and to "go back in time" . We can select images that are closest to the actual fieldwork period, but in many cases, we will be left with imagery a few years older. Old does not necessarily mean outdated, but that is not an assumption we should readily make. With more regular updates to GSV, this problem will be likely to reduce in the future. Until then, up-to-date second assessments of homes may simply have to be done in person at the location.

4.2 Validating interviewer observations with auxiliary data from GSV

The abovementioned pitfalls illustrate that using GSV to provide auxiliary data is not as easy and straightforward as using it in regular daily life. In this section, we will assess how useful the obtained data from GSV is to validate interviewer observed data.

To validate interviewer observations, we evaluate how much concordance we find between the interviewer-observed paradata and the GSV auxiliary data. Some 70.9% of the property types show an exact match between our GSV observations and those of the interviewers (Table 3 and Appendix A1). If we are more lenient and ignore misclassifications of the subtypes of detached houses (farms versus regular detached homes) and multi-unit buildings (apartments, flats, student houses, or retirement homes) that can be hard to visually distinguish, even when physically present at the location, we reach 78.7% concordance. With the sometimes pixelated quality of the images when zooming in on doorbells, we may have classified certain properties as normal terraced and detached houses, whereas they are in fact multi-unit dwellings. If the possibility of them being multi-unit holds true and we correct for this, then the resulting extra-lenient match would leave us with 84.2% concordance.

Matching the coding for the physical condition of homes was more difficult, which may be due to the relatively subjective nature of these judgements. If we are more lenient and distinguish only between "(very) good" versus "reasonable" to "(very) bad" condition, we end up with 78.3% concordance (Table 3 and Appendix A3). The level of agreement for the exact type(s) of impediments to access is similarly low, but lenient matching, by distinguishing whether or not any impediment was observed, results in 84% concordance for the cases without missing codes for the GSV data.

Higher levels of agreement are found for the presence of litter and vandalism. The differences between the exact match (categorical: amounts of litter and vandalism) and the lenient match (binary: presence of litter and vandalism or not) are very small. It should be noted, however, that neither the interviewers nor the GSV coder frequently observed litter or vandalism. Obtaining concordance for all five observations combined only occurred in 38.9% of the 445 usable cases. When disregarding the difficult-to-code variable 'impediments to access," the four combined observations matched in 46.6% out of 543 cases. Three out of four matches occurred in 82.9% of the cases, and four out of five in about 74.6% of them.

This level of agreement is not a poor result, but for the types of dwellings, we question why 125 (21.3% no exact or lenient match) out of 588 houses have different codes, because the type of home should be more stable over time than the presence of litter or vandalism, and less subjective than the judgement of the condition. The most common mismatches occurred with multi-unit buildings (Appendix A1). We might have failed to recognize multi-unit buildings because we missed the details of doorbells ($N = 25$), but using GSV we also coded some multi-unit buildings ($N = 34$) where the interviewers did not. The latter may also be caused by us misjudging doorbells, but it is also possible that there are real discrepancies between our GSV observations and the interviewer observations. Another recurrent difference is the interviewers recording a terraced house, whereas based on GSV it was classified as a detached or semi-detached house ($N = 34$) and the other way around ($N = 19$). It is possible that houses with the same building style appear as one large house in GSV. Other mismatches occurred for buildings that comprise a commercial property with one living unit (5 in the interviewer data, 12 in the GSV data), and with "other" (respectively 1 and 5) and "don't know" ($N = 3$).

Table 4 shows the extent of matching by how we located the home using GSV. Having a match for the house type (with the lenient variable that disregards differences between the subtypes of types of dwelling) did not significantly differ with regard to how the property was located using GSV (Table 4). We achieved 78.9% matching for the type of property when spotting the exact address, and 68.4% when we spotted the exact address after GSV placed us far from the actual home. Whether there was a match for the type of home also did not significantly differ with regard to whether the GSV imagery was out of date (Table 5).

We would expect that having up-to-date GSV imagery from 2014 and accurate visual confirmation of the exact house number – or that of a direct neighbour – combined with being certain of the direction of the numbering in the street, would provide us with the most trustworthy data. Some 70 out of the 96 cases that fulfil these criteria have a matching house type (73%, not tabulated). This means that we are left with 26 mismatches with up-to-date images and visual confirmation of the – or a direct neighbour's – house number.

Table 4
Concordance between interviewer observations and GSV observations in the ESS7BE for type of house by how homes were found in GSV

	No match		Match		Total	
	N	%	N	%	N	%
Exact address	52	21.1	195	78.9	247	100
Exact address but further	12	31.6	26	68.4	38	100
Direct neighbour's house number	21	18.4	93	81.6	114	100
Further neighbour's house number	4	10.0	36	90.0	40	100
Neighbour commercial or public building	7	38.9	11	61.1	18	100
Strong similarities	14	18.7	61	81.3	75	100
Censored	0	0.0	2	100	2	100
Physical barriers	5	31.3	11	68.8	16	100
Exact address not found	2	20.0	8	80.0	10	100
No GSV	8	28.6	20	71.4	28	100
Total	125	21.3	463	78.7	588	100

$\text{Chi}^2 = 12.043$, $p = 0.211$

Table 5
Concordance between interviewer observations and GSV observations in the ESS7BE for type of home by date of GSV imagery

	No match		Match		Total
	N	%	N	%	N
2009	39	18.3	174	81.7	213
2010	9	17.3	43	82.7	52
2011	1	50.0	1	50.0	2
2012	1	50.0	1	50.0	2
2013	38	21.6	138	78.4	176
2014	31	25.4	91	74.6	122
< 2014	88	19.8	357	80.2	445
2014	31	25.4	91	74.6	122
Total	119	21.0	448	79.0	567

$\text{Chi}^2 = 4.853$, $p = 0.434$ (upper panel); $\text{Chi}^2 = 1.568$, $p = 0.211$ (lower panel)

It seems unlikely that all these 26 houses had been rebuilt or transformed during 2014. Also somewhat suspicious to observe are six cases with matching house types, and up-to-date and accurate GSV imagery, in which we clearly spotted impediments to access whereas the interviewers had not. Similarly, there were eight cases the other way round, where we did not spot any impediments (not tabulated). In the case of any doubt when scrutinizing the GSV images, the GSV data would have been coded with a missing value. The cause of this lack of concordance despite up-to-date and spot-on GSV images remains unclear: either the GSV coding was wrong or the interviewers made mistakes.

4.3 Predicting nonresponse with auxiliary data from GSV

In the previous sections, we discussed some quality issues of GSV data and interviewer observed data. Paradata in the form of interviewer observations are often used to evaluate the randomness of non-response (Kreuter & K., 2013). Therefore, we compare the utility of the interviewer and GSV observations in response propensity models to establish which works better. In Table 6 the results of logistic regressions for contact, refusals and response are presented separately for GSV data and interviewer observations. Both the GSV-data and the interviewer observations predict the same: the presence of impediments lead to significantly less contact in the ESS7BE. That impediments hamper contact is also in line with literature about contactability (see e.g. Blom et al., 2011; Groves & Couper, 1998). Additionally, when it was not possible to view if there were any impediments in GSV, the GSV-model also significantly predicts less contact. No significant effects of impediments on response (given contact) are found nor significant effects on refusal.

5 Discussion

Using Google Street View seems like a straightforward, cost-saving approach to collecting auxiliary data for surveys and cross-checking interviewer-observed paradata. However, in practice, we encountered quite a few pitfalls, as detailed in Section 4.1. Some of these problems were relatively easy to overcome, others were not. In addition to GSV sometimes simply not being available, there were difficulties in finding homes, and if found, problems in identifying all the details in the images. The images were frequently a few years old, risking the information being out of date. However,

Table 6
Predicting contact, refusals and response in the 20% subsample of the ESS7BE with GSV observations versus interviewer observations

	GSV				Interviewer observations			
	Model 1		Model 2		Model 3		Model 4	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Response (1=response)								
Constant	0.264**	0.100	0.331**	0.139	0.204*	0.102	0.463***	0.141
Litter	0.070	0.246	0.107	0.259	0.078	0.323	0.134	0.355
Vandalism	−0.646	0.358	−0.261	0.401	0.205	0.431	0.433	0.474
Condition of home[a]	0.056	0.271	0.197	0.295	0.021	0.226	0.043	0.248
Impediments to access	-	-	−0.339	0.201	-	-	−0.775***	0.205
Nagelkerke	0.008		0.013		0.002		0.048	
Contact (1=contact)								
Constant	3.063***	0.235	3.931***	0.468	2.865***	0.223	4.476***	0.597
Litter	−0.831*	0.426	−0.709	0.495	0.918	0.925	0.794	1.064
Vandalism	−1.132*	0.518	−0.703	0.647	−1.143	1.015	−0.691	1.169
Condition of home[a]	0.298	0.577	1.478	1.053	−0.448	0.440	−0.485	0.503
Impediments to access	-	-	−1.535**	0.525	-	-	−2.331***	0.633
Nagelkerke	0.049		0.128		0.014		0.156	
Refusal (1=refusal)								
Constant	−0.940***	0.111	−0.881***	0.151	−0.898***	0.112	−0.923***	0.153
Litter	−0.436	0.293	−0.409	0.307	0.089	0.356	−0.064	0.391
Vandalism	0.302	0.387	0.177	0.449	−0.530	0.513	−0.561	0.564
Condition of home	−0.036	0.306	0.088	0.323	−0.331	0.266	−0.250	0.284
Impediments to access	-	-	−0.137	0.226	-	-	0.123	0.224
Nagelkerke	0.007		0.008		0.011		0.013	
N	531		430		531		430	

[a] reasonable to very bad, ref. cat.: very good
* $p \leq 0.05$ ** $p \leq 0.01$ *** $p \leq 0.001$

the differences between GSV data and interviewer-observed paradata cannot all be explained by outdated images or by how the homes were found in GSV. Even when we had clear visual confirmation of the exact address and up-to-date images, we sometimes still ended up with different codes to those from the interviewers. Given the pitfalls of working with GSV, it is hard to determine the actual cause of these discrepancies. This leaves us to wonder whether the GSV images are still not accurate enough, or whether the interviewers made real mistakes.

The advantage of using GSV is that we can inspect homes and neighbourhoods without having to physically send a second person to carry out independent observations in order to double-check the interviewer observations. Having one person do all the coding from GSV images should also help overcome the potentially subjective interpretations by many different interviewers regarding the condition of a home and the amount of litter and vandalism. Whereas judging and expressing the condition of a home and the amount of litter and vandalism may be influenced by subjective interpretations, the type of home and observations about impediments to access should be less susceptible to interpretation. However, the current typology for the type of property may nevertheless be susceptible to this issue. In 12 cases, GSV "took us" to farm buildings in very rural areas, where there was mostly farmland around. The interviewers had coded these houses as detached. This is of course also a correct code for the (main) farm building, but clearly shows that the house type categories used in the ESS are not necessarily mutually exclusive and can be debated. This example illustrates that the coding of home and neighbourhood characteristics is more complex than it appears at first glance. Reality is not easy to capture with a limited set of codes.

We also need to note that the GSV coder focused only on the observations, whereas the interviewers also needed to establishing contact and, even more importantly, on obtaining an interview. Accordingly, the GSV coder may have more time and motivation to create the supplementary data. Moreover, the GSV coder registers the observations immediately when locating a home, whereas interviewers might not always fill in the contact forms at the location straight away. Although we cannot provide evidence for all these suggestions, we do believe that coding of the type and condition of dwellings, the presence of impediments to access and the condition of the neighbourhood by one, single coder results in more reliable data compared with interviewer observations made by a large group of different individuals. Having two or three coders who each process all the sample units" addresses would be even better, as it would allow the testing of between-coder reliability. On the one hand, it is easier to control and assess the quality of the coding done by a small group of GSV coders than by a large group of interviewers. Interviewers, on the other hand, do their coding work close to reality.

Our conclusion is the current pitfalls of using GSV data limit the ease of use and the quality of this auxiliary data. The coding takes more time than expected, and the issues of outdated and pixellated images do not currently allow us to accurately validate the interviewer-observed paradata. Moreover, the GSV coverage in other countries participating in the ESS is not always as good as in Belgium, so implementing the GSV coding procedure in some other countries may not be possible until the coverage (and the quality of the imagery) improves further. Also, different countries have different privacy regulations with regard to the level of detail that GSV can publish (Google, 2017b). Belgium does not limit the use of GSV (Privacy Commission Belgium, 2017), but as in all countries with GSV, there is an established right to have your house blurred by Google removed (Google, 2015a).

Non-commercial use of GSV, such as for research reports and other related professional documents, is allowed as long as the fair use principle of Google is honoured (Google, 2015b). Under that stipulation, images from Google Maps, Google Earth and Google Street View are even allowed to be imbedded in printed documents, given correct attribution of the source of the images. For the use of GSV as auxiliary data for ESS, images of respondents" houses will obviously never released to the public. All ESS-data is required to be completely anonymized and the ESS-protocols guarantee this protection of privacy. The coder of the GSV data in this study also had legal access to the Belgian ESS-sample information and more specifically to the address list at the time of the data collection. We highly recommend informing about the privacy rules regarding use of the sample and the local terms of service regarding the use of GSV when intending the use of GSV for the same or similar purposes as presented in this study. If GSV was to be used on a larger scale than presented in this study, the fair use principle of the Google products allows a maximum of 25,000 free map loads per day – additional loads require a small fee (Google, 2017a). Google also offers application programming interfaces (APIs) that can allow bulk entry of addresses and customizing of the maps within the terms of service of Google (Google, 2017a).

Despite the current limits in using GSV to validate of replace interviewer observations, it is a promising, freely available source of auxiliary data that has provided us with some insights into potential improvements with regard to filling in the contact forms correctly. These potential improvements should be addressed in the upcoming interviewer training sessions for the ESS in Belgium (e.g. more detailed discussion of specific examples which are illustrated with GSV screens or photographs). It is clear that some of the observed differences between GSV information and the information collected by the interviewers are very informative with regard to improving the interviewer instructions for collecting observable variables.

Although the focus of this paper is on the coding of GSV data and matching this to interviewer observations, we also evaluated whether this data can predict nonresponse, contacts and refusals. The GSV data is found to make similar predictions than the interviewer observations. In that way, the GSV-data are not more advantageous than using the interviewer observations. However, the GSV data seems to enable the detection of potential contactability issues before the actual fieldwork starts. Although it does not seem to help us to tackle the bigger nonresponse problem of refusal (26.1% of the ESS7BE sample) in comparison with contactability issues (5.4%), GSV can still be useful. It could help in predicting contact issues in advance of the fieldwork and foreseeing extra contact attempts when litter, vandalism or impediments to access are identified. This kind of applications may prove particularly useful after further updates and technological improvements of GSV. As such, GSV auxiliary data may soon become easier to process and more reliable to use, in combination with a well-elaborated coding scheme that is sufficiently adapted to the local environment.

Acknowledgements

The author wants to thank the colleagues at the Center for Survey Methodology at KU Leuven for helpful comments at the start of this project. The author also wishes to express gratitude to the editor Stephanie Eckman for her valuable feedback and for the constructive cooperation. And finally, a special thanks goes out to Dr. Maximilian Grasl at Universität Regensburg.

References

American Association for Public Opinion Research. (2009). Standard definitions: final dispositions of case codes and outcome rates for surveys. Retrieved from http://www.aapor.org/

Barbier, S., Wuyts, C., Italiano, P., & Loosveldt, G. (2016). *European Social Survey Round 7 Belgium – process evaluation for the data collection.* Leuven: Centre for Sociological Research, KULeuven.

Beullens, K. (2013). *The use of paradata to assess survey representativity. Cracks in the nonresponse paradigma*. 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*(1), 3–43.

Blom, A. G., de Leeuw, E. D., & Hox, J. J. (2011). Interviewer effects on nonresponse in the European Social Survey. *Journal of Official Statistics*, *27*, 359–377.

Campanelli, P., Sturgis, P., & Purdon, S. (1997). *Can you hear me knocking? An investigation into the impact of interviewers on survey response rates.* London: National Centre for Social Research (UK).

Clarke, P., Ailshire, J., Melendez, R., Bader, M., & Morenoff, J. (2010). Using Google Earth to conduct a neighborhood audit: reliability of a virtual audit instrument. *Health & Place*, *16*(6), 1224–1229.

Curtis, J. W., Curtis, A., Mapes, J., Szell, A. B., & Cinderich, A. (2013). Using Google Street View for systematic observation of the built environment: analysis of spatio-temporal instability of imagery dates. *International Journal for Health Geographics*, *12*(1), 53.

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.

Eckman, S. & Kreuter, F. (2013). Undercoverage rates and undercoverage bias in traditional house listing. *Sociological Methods & Research*, *42*(3), 264–293.

European Social Survey. (2014). ESS 7 source contact forms. Retrieved from http://www.europeansocialsurvey.org/docs/round7/fieldwork/source/ESS7_source_contact_forms.pdf

European Social Survey Round 7 Data. (2014). *Data file edition 1.0.* Norway: NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.

Goodfellow, I., Bulatov, Y., Ibarz, J., Arnoud, S., & Shet, V. (2014). Multi-digit number recognition from street view imagery using deep convolutional neural networks. In *Iclr2014*.

Google. (2015a). Image acceptance & privacy policies. Retrieved from https://www.google.com/maps/streetview/policies/

Google. (2015b). Using Google Maps, Google Earth and Street View. Retrieved from https://www.google.com/permissions/geoguidelines.html

Google. (2017a). Google maps APIs. Retrieved from https://developers.google.com/maps/pricing-and-plans/%5C#details

Google. (2017b). Google maps help. Blur or remove 360 photos. Retrieved from https://support.google.com/maps/answer/7011973?hl=en%5C&ref%5C_topic=6275604

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

Hara, K., Azenkot, S., Campbell, M., Bennett, C. L., Le, V., Pannella, S., … Froehlich, J. E. (2015). Improving public transit accessibility for blind riders by crowdsourcing bus stop landmark locations with Google Street View: an extended analysis. *ACM Transactions on Accessible Computing (TACCESS)*, *6*(2), 5.

Kreuter, F. (Ed.). (2013). *Improving surveys with paradata: analytic uses of process information.* Hoboken, NJ: John Wiley & Sons.

Kreuter, F. & K., O. (2013). Paradata for nonresponse error investigation. In F. Kreuter (Ed.), *Improving surveys with paradata: analytic uses of process information* (pp. 13–42). Hoboken, NJ: John Wiley & Sons.

Less, E. L., McKee, P., Toomey, T., Nelson, T., Erickson, D., Xiong, S., & Jones-Webb, R. (2015). Matching study areas using Google Street View: a new application for an emerging technology. *Evaluation and program planning*, *53*, 72–79.

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

Odgers, C. L., Caspi, A., Bates, C. J., Sampson, R. J., & Moffitt, T. E. (2012). Systematic social observation of children's neighborhoods using Google Street View: a reliable and cost-effective method. *Journal of Child Psychology and Psychiatry*, *53*(10), 1009–1017.

Pickering, K., Thomas, R., & Lynn, P. (2003). *Testing the shadow sample approach for the English house condition survey.* London: National Centre for Social Research (UK).

Privacy Commission Belgium. (2017). FAQ Google Street View. Retrieved from https://www.privacycommission.be/nl/faq-page/387%5C#t387n3759

Rundle, A. G., Bader, M. D., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using Google Street View to audit neighborhood environments. *American journal of preventive medicine*, *40*(1), 94–100.

Sinibaldi, J., Durrant, G. B., & Kreuter, F. (2013). Evaluating the measurement error of interviewer observed paradata. *Public Opinion Quarterly*, *77*(S1), 173–193.

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

Vandeviver, C. (2014). Applying Google Maps and Google Street View in criminological research. *Crime Science*, *3*(1), 1–16.

Walsh, R., Dahlhamer, J., & Bates, N. (2013). Assessing interviewer observations in the NHIS. Retrieved from http://www.Amstat.org

West, B. T. (2013). An examination of the quality and utility of interviewer observations in the National Survey of Family Growth. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(1), 211–225.

West, B. T. & Kreuter, F. (2013). Factors affecting the accuracy of interviewer observations evidence from the National Survey of Family Growth. *Public Opinion Quarterly*, *77*(2), 522–548.

Appendix A
Tables

(Appendix Tables follow on next page)

Table A1
Interviewer observations versus GSV observations for the type of home in the 20% subsample of the ESS7BE

	Interviewer observations																					
	Farm		Detached		Semi-detached		Terraced house		Comm. property		Multi-unit house/flat		Student apartm.		Retirem. home		Other		Dont't know		Total GSV	
GSV	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Farm	0	0	12	6	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	13	2
Detached	2	67	157	82	15	17	8	5	0	0	1	1	0	0	0	0	3	60	1	33	187	32
Semi-Detached	0	0	14	7	53	58	26	17	0	0	2	2	0	0	1	25	0	0	1	33	97	17
Terraced house	0	0	4	2	15	17	99	66	0	0	17	13	2	40	0	0	1	20	1	33	139	24
Comm. Property	0	0	4	2	2	2	2	1	1	20	3	2	0	0	0	0	0	0	0	0	12	2
Multi-unit flat/house	1	33	1	1	6	7	16	11	4	80	104	81	3	60	2	50	1	20	0	0	138	24
Students apartments	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	25	0	0	0	0	1	0
Other	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0
Total	3 (1%)		192 (33%)		91 (16%)		151 (26%)		5 (1%)		129 (22%)		5 (1%)		4 (1%)		5 (1%)		3 (1%)		588 (100)	

Table A2
Interviewer observations versus GSV observations for the condition of the home, presence of litter, vandalism and impediments to access in the 20% subsample of the ESS7BE

	Interviewer observations											
	1		2		3		4		5		Total	
GSV	N	%	N	%	N	%	N	%	N	%	N	%
Physical condition of the home												
(1) very good	108	60	106	43	16	16	1	6	1	50	232	43
(2) good	65	36	112	46	56	58	9	53	0	0	242	45
(3) satis-factory	8	4	23	9	22	23	7	41	1	50	61	11
(4) bad	0	0	4	2	3	3	0	0	0	0	7	1
(5) very bad	0	0	0	0	0	0	0	0	0	0	0	0
total	181	100	245	100	97	100	17	100	2	100	542	100
Amount of litter												
(1) very large	0	0	0	0	0	0	1	0	-	-	1	0
(2) large	0	0	2	11	3	4	1	0	-	-	6	1
(3) small	1	33	5	26	17	23	64	13	-	-	87	15
(4) (almost) none	2	67	12	63	54	73	414	86	-	-	482	84
total	3	100	19	100	74	100	480	100	-	-	576	100
Amount of vandalism												
(1) very large	0	0	0	0	0	0	1	0	-	-	1	0
(2) large	1	50	0	0	1	3	3	1	-	-	5	1
(3) small	1	50	2	20	5	14	26	5	-	-	34	6
(4) (almost) none	0	0	8	80	31	84	496	94	-	-	535	93
total	2	100	10	100	37	100	526	100	-	-	575	100
Presence of impediments to access												
(1) entry phone	22	24	0	0	6	9	8	3	-	-	36	8
(2) locked gate/ door	26	28	12	44	35	52	24	9	-	-	97	22
(3) both	30	33	0	0	21	31	6	2	-	-	57	13
(4) neither	14	15	15	56	5	7	223	85	-	-	257	57
total	92	100	27	100	67	100	261	100	-	-	447	100

Appendix B
Interviewer observations in the contact form of the ESS7

ESS DOCUMENT DATE 03/04/2014

NEIGHBOURHOOD CHARACTERISTICS FORM

- **ONE FORM TO BE COMPLETED FOR EACH ADDRESS**
- **COMPLETE DURING DAYLIGHT WHEREVER POSSIBLE**
- **MUST BE COMPLETED FOR ALL SAMPLE UNITS INCLUDING ALL NON CONTACTS, ALL REFUSALS, ALL OTHER TYPES OF NONRESPONSE UNITS AS WELL AS ALL INTERVIEWS**

N1. What type of house does the (target) respondent live in?

1 Farm

Single-unit:

2 Detached house

3 Semi-detached house

4 Terraced house

5 The only housing unit in a building with another purpose (commercial property)

Multi-unit:

6 Multi-unit house, flat

7 Student apartments, rooms

8 Retirement house

Other:

9 House-trailer or boat

10 Other (SPECIFY)……………………

88 Don't know

N2. Before reaching the (target) respondent's individual door, is there an entry phone system or locked gate / door?

INTERVIEWER: Record whether there is a gate / door that is locked at the time that the neighbourhood characteristics form is completed.

1. Yes – entry phone system
2. Yes – locked gate / door
3. Yes – entry phone system AND locked gate / door
4. No – neither of these

5

ESS DOCUMENT DATE 03/04/2014

N3. What is your assessment of the overall physical condition of this building/house?

NOTE TO INTERVIEWER:

Consider the following issues when assessing the overall physical condition of this building/house.

1. Roof problems (e.g. sagging roof, missing roofing material)
2. Problems with windows (e.g. boarded up or broken windows)
3. Other problems (e.g. sloping outside walls, broken plaster or peeling paint, guttering problems)

1. Very good
2. Good
3. Satisfactory
4. Bad
5. Very bad

NOTE TO INTERVIEWER:

For the remaining two questions (N4 & N5) please give your overall opinion about the 'immediate vicinity' of the building/house of the target respondent. Look to the left and the right of the building/house taking into account a distance of about 2 normal sized houses on either side (approximately 15 metres on either side). Only include this area and the property of the target respondent when answering these questions.

There may not be other properties on either side of the building so just estimate the space that about 2 'normal' size houses on either side would take up.

Note that in the case of blocks of flats refer to the space on either side of the whole building and NOT just the individual flat where the target respondent lives.

N4. In the immediate vicinity, how much litter and rubbish is there?

1 Very large amount
2 Large amount
3 Small amount
4 None or almost none

N5. In the immediate vicinity, how much vandalism and graffiti is there?

1 Very large amount
2 Large amount
3 Small amount
4 None or almost none

6

(European Social Survey, 2014)