Passive Meter: some methodological discussions

Oriol J. Bosh and Melanie Revilla

TRI-POL research project.

A general explanation

Metered data, also called “web log data”, “digital trace data”, “passive data”, “online behavioural data” or “web-tracking data” is obtained from a meter willingly installed or configured by a sample of participants on their devices (PCs, tablets and/or smartphones). A meter refers to a heterogeneous group of tracking technologies that allow sharing with the researchers, at least, information about the URLs of the web pages visited by the participants. Depending on the technology used, HTML, time or device information can also be collected.

Metered data has been considered as the best option to track individual’s online behaviour, since it can track the real behaviour in an inobtrusive way. However, different errors can affect metered data, most of them related to behaviours not being tracked. For instance, technology errors (e.g. the meter stopping working in low battery mode), technology limits (e.g. in-app behaviours not being trackable), behaviours being actively hidden (e.g. participants stopping the meter when visiting sensitive websites) or not tracking all the devices.

Let us consider the error of not tracking all devices, which could be considered the most problematic of all. When using a metered panel like Netquest, researchers can obtain samples of individuals with the meter already installed in at least one device. However, researchers do not have control over which devices panellists have the meter installed, nor which technology is being used. In practice, individuals can use multiple devices to access the internet, from their personal smartphone or laptop to their work computer to a shared computer in their home or elsewhere. To gain a comprehensive picture of their online activity, trackers should be installed on all devices an individual use. However, we know that in practice this does not happen. Barthel et al. (2020) found that, for those Ipsos’ Knowledge Panel panellists who accepted to install a meter in at least one of their devices, only 28% stated having all the devices that they use to access the internet metered. Revilla, Ochoa and Loewe (2017),

surveying the metered panellists of Netquest in Spain, found that almost 57% of the panellists have the meter installed in only one device whereas only 4% of them use only one device to go online. Hence, it seems that a big proportion of metered panellists might be partially tracked, with an unknown amount of information being lost. Ideally, we would only aim to sample individuals with the meter installed in all their devices, although this could increase the selection bias. However, this information is not normally available by metered panels, and it fluctuates over time (e.g. a panellist buying a new device).

Observing only part of the online behaviour of individuals is an important problem on itself. However, because of the very nature of metered data, there is an extra problem that we should take care of: **understanding if a lack of behaviour tracked is a real lack of behaviour or if there is no behaviour tracked because the behaviour happened on an untracked device**. This is important because, if we know that a lack of behaviour recorded for an individual is provoked by a tracking error and not by its real behaviour, the individual should be excluded from the analyses (we cannot consider the lack of behaviour as real, so we do not know the real value for the variable of interest). If the lack of behaviour is a real behaviour, then it should be considered as so (e.g. 0 minutes visiting elpais.com during the last 15 days). An easy way to illustrate this is comparing it with what happens with surveys. Let us imagine that we want to know how many minutes someone has visited elpais.com during the last 15 days. In a survey, we would directly ask a question to the participant. The participant could answer and provide a number or could not answer/not provide a number (for whatever reason). Those not providing a number/answer would be considered as missing, and since no information would be available from them, they would be excluded from the specific analyses using information about the time spent visiting elpais.com. When using metered data, instead of asking a question we are observing their behaviour using a meter. If the meter does not observe an individual visiting the URL (elpais.com), this can be a real behaviour (the person never visited elpais.com) or be provoked by tracking errors (e.g. the participant visited elpais.com but with an untracked device). However, we do not have enough information to determine if the lack of behaviour is real or provoked by errors. How can we decide if we should consider the lack of behaviour as 0 minutes or as a missing? We need information in order to consider the absence of behaviour as a missing. Let us say that for this example we suspect that the only reason for data to be missing is that an individual does not have the meter installed in all his or her devices. In this case, we would need to know if the individual has all his devices tracked. If the individual has all his or her devices tracked, it is

safe to assume that the absence of behaviour should be treated as true. Hence, we would input a value of 0 minutes visiting elpais.com. If it does not have the meter installed in all his or her devices, all the behaviours done through the untracked devices will not be observed. In this case, we can either consider that absence as a real absence of behaviour or as a missing. Both options have perils: we can falsely exclude an individual with a real absence of behaviour, or we can underreport the real behaviour of an individual with a real but unknown behaviour. If we decide not to consider it as missing, the produced variable might be affected by measurement errors i.e. the value obtained from a sampled unit deviates from the true value that the measurement should have if no errors happened when collecting the data.

Considerations for this specific project

Information about the devices

As it will be shown in further sections, some errors are only found on specific devices, operating systems (OSs), OS versions or browsers. However, the proportion of individuals using those devices in the given countries is sometimes significantly different. Besides, within the Netquest metered panel, the proportion metered with different devices is also significantly different. Therefore, before detailing the potential errors, we first provide some contextual information.

In terms of the devices in which participants install the meter, Table 1 presents the proportion of panellists in each country with the metered installed in at least one type of device. Data comes from Revilla et al. (2021).

Table 1. Presence of different device types

	Spain	Portugal	Italy	Argentina	Chile
At least one PC	61.0	81.4	NA	81.2	65.5
At least one smartphone	71.0	48.6	NA	69.0	77.5
At least one tablet	21.7	17.4	NA	16.1	11.8

Within mobile devices (smartphone and tables), there can be variation between the OSs used. Table 2 presents the information about the OSs for both the national market and the metered

panel. Information from the market is from January 2021 (https://gs.statcounter.com/os-market-share/mobile). Metered panel data comes from Revilla et al. (2021).

Table 2. Proportion of OSs out of all mobile devices

	Spain	Portugal	Italy	Argentina	Chile
Android market share	79.9	72.3	73.5	92.0	85.1
Android metered panel	78.9	85.4	NA	92.3	81.1
iOS market share	19.7	27.3	25.6	7.7	14.7
Other market share	.4	.4	.9	.3	.2

For Android devices, it is also important to know the proportion of devices running a version of the OS higher or lower than 10. Table 3 presents the market share of each OS versions, per country. Information is from January 2021 (https://gs.statcounter.com/browser-market-share).

Table 3. Proportion of OS versions out of all Android devices

	Spain	Portugal	Italy	Argentina	Chile
10 or higher	56.7	57.3	51.8	29.9	51.6
Lower than 10	43.3	42.7	48.2	70.1	48.4

Limitations also vary across browsers. Table 4 presents the market share of each browser, for mobiles and desktops separately, for the different countries. Information is from January 2021 (https://gs.statcounter.com/browser-market-share).

Table 4. Proportion of browser type out of all browsers

	Spain	Portugal	Italy	Argentina	Chile
Desktop					
Chrome	71.6	73.6	67.0	81.9	78.4
Safari	8.8	6.9	9.5	4.0	4.0
Firefox	8.7	6.2	12.9	5.5	5.2
Edge	6.5	8.1	6.5	3.9	4.7
Internet Explorer	1.3	1.3	1.5	.8	.8

	Opera	2.4	3.4	1.9	3.4	5.8
	Other	.7	.5	.7	,5	1.1
Mobile						
	Chrome	74.3	68.3	66.1	87.0	81.3
	Safari	18.2	24.7	24.0	6.5	12.9
	Firefox	.6	.5	.5	.2	.3
	Edge	.2	.0	.0	.1	.0
	Samsung	5.8	4.9	8.1	5.5	4.7
	Opera	.6	.5	.4	.4	.3
	Other	.3	1.1	.9	.3	.5

Potential problems and how to identify them.

There are different sources of missing data. Not all sources provide enough information to safely identify missing data. Hence, to decide in which ways missing data can be accurately identified, we first need to understand which problems can generate missing data for this specific project. In this project the main interest is in domain-level (e.g. twitter.com) and subdomain-level information in the use of social media (e.g. twitter.com/pabloiglesias), and main outlets (e.g. elpais.com) and subdomain-level (elpais.com/politica). Data can be missed completely or partially. Completely would mean that no data was recorded. Partially would mean that some data was recorded, but not all. Partial missingness will always produce measurement errors. Conversely, complete missingness can be identified as missing with enough information. There are three general sources of error affecting both domain and subdomain level information. We present the three sources of error as well as specific proposals to identify missing data:

1) The behaviour is not tracked because the person decided to stop the meter.

Source 1 might be redundant for your project since most of the domains that you are interested in might not be considered as sensitive. We can assume that not many people are going to stop their meter when visiting, for instance, Twitter or news outlets, since it might not be considered as a sensitive webpage for most people. Besides, the meter can only be stopped for 15 minutes. It is unlikely that any participant is going to systematically stop the

meter every time that visits twitter or a news outlet. Hence, most likely this might produce partial missingness i.e. measurement error.

2) The behaviour is not tracked due a malfunction/limitation of the tracker

Source 2 can be provoked by different types of errors. Most of them affect iOS devices.

- Proxy errors. Some behaviours collected using a proxy might not be collected. These affect mostly iOS since proxies are used to track them. It is unclear how common errors due to using a proxy are. **And no information is available to assess this**. It seems unlikely, anyway, that all the visits to general domains as twitter or news outlets will be lost due to proxy errors, since errors should not systematically affect some domains and not others. Besides, as presented in Table 1, a substantially lower proportion of iOS are tracked. Hence, proxy errors might be more likely to cause measurement errors than missing data errors.
- Incognito session. Information from incognito sessions cannot be collected from iOS or Edge/Explorer. It is unlikely that the behaviours that this project is interested in are going to happen in their entirety only on incognito sessions, but it should be considered. It could be asked which proportion of the time that they spend visiting Twitter is done through the incognito session. However, **we consider that asking a question is not necessary since the incidence of this problem is expected to be low and most likely it will not provoke complete missingness.**

3) The behaviour happens on an untracked device/web browser/network

Source 3 is the most problematic one. It is safe to assume that Twitter and news are consumed through different devices. For Twitter, it is also safe to assume that a non-redundant part of the users only consumes it through their smartphones, specifically though the app. If participants do not have the meter installed in all their devices/web browsers/networks, it is then likely that data is missing. An approach to obtain relevant information is to ask which devices they use to access the internet (including using apps like Twitter) in an average day or during the last X days, and then cross-reference that with the information of the devices tracked.

met1: During the last 15 days, from how many of these different types of devices have you accessed the Internet (including using apps like Facebook, Twitter or YouTube)? Please, type the number of devices in the respective boxes.

met1a: Computer with Windows operating system: **[PROGRAMMER: NUMERIC OPEN BOX]**

met1b: Apple computer(s) (MAC): **[PROGRAMMER: NUMERIC OPEN BOX]**

met1c: Smartphone or tablet with Android operating system: **[PROGRAMMER: NUMERIC OPEN BOX]**

met1d: Apple smartphone or tablet (iPhone or iPad): **[PROGRAMMER: NUMERIC OPEN BOX]**

met1e: Others: **[PROGRAMMER: NUMERIC OPEN BOX] (IF >0: "Please, specify: [OPEN TEXT BOX]")**

However, even if devices are tracked, the behaviour happening in some browsers might not be tracked. For PCs, behaviours can be either tracked with plug-ins or with a desktop app. Those using the desktop are completely tracked, regardless of the browsers used. Those using plug-ins are only tracked on those browsers with the plug-in installed. It is not clear which proportion of participants might use non-metered browsers. For Android devices with an OS version equal or higher to version 10, behaviours are only tracked on Chrome, Firefox and Samsung browsers. As seen in Table 3, except in Argentina, around 50% of Android devices run a version 10 or higher. However, the market share of non-trackable browsers is really low (around 1%). If this is also the case for the metered panel, we should expect a low proportion of participants affected by this. Regardless of this, to gather information, for the devices affected, we should ask about the browser that they use.

[PROGRAMMER: GO TO met2 IF THE PARTICIPANT IS TRACKED AT LEAST IN ONE WINDOWS PC AND IT IS NOT TRACKED WITH THE DESKTOP APP. THEREFORE, IF BROWSER PLUGIN= YES & OS= WINDOWS. IF NOT TRACKED ON THAT DEVICE AND WITH THAT TECHNOLOGY, MOVE TO met3]

met2: During the last 15 days, have you used any of the following web browsers to access the Internet through a computer with Windows operating system?

		Yes	No
met2a	Internet Explorer	o	o
met2b	Chrome	o	o

met2c	Firefox	o	o
met2d	Edge, Opera or others	o	o

[PROGRAMMER: GO TO met3 IF THE PARTICIPANT IS TRACKED AT LEAST ON ONE MAC AND IT IS NOT TRACKED WITH THE DESKTOP APP. THEREFORE, IF BROWSER PLUGIN= YES & OS= OSX. IF NOT TRACKED ON THAT DEVICE AND WITH THAT TECHNOLOGY, MOVE TO met4]

met3: During the last 15 days, have you used any of the following web browsers to access the Internet through an Apple computer (MAC)?

		Yes	No
met3a	Internet Explorer	o	o
met3b	Safari	o	o
met3c	Chrome	o	o
met3d	Firefox	o	o
met3e	Edge, Opera or others	o	o

[PROGRAMMER: GO TO QUESTION met4 IF THE PARTICIPANT IS TRACKED AT LEAST ON ONE ANDROID DEVICE. THEREFORE, IF ANDROID=YES. IF NOT TRACKED ON THAT DEVICE, MOVE TO met5]

met4: During the last 15 days, have you used any of the following web browsers to access the Internet through smartphone or tablet with Android operating system?

		Yes	No
met4a	Chrome	o	o
met4b	Samsung browser	o	o
met4c	Firefox	o	o
met4d	Edge, Opera or others	o	o

For iOS devices, participants need to configure the proxy for all the networks that they use. For instance, if a participant uses a 4G mobile data as well as a home and work WIFI, three networks should be configured. If the work WIFI is not configured, all the online behaviour

happening in that given network would not be tracked. However, Netquest only knows if an individual is being tracked by a proxy but does not know in how many or which types of networks the participant is being tracked. This lack of information prevents us from knowing which participants are completely or partially tracked when it comes to networks for iOS devices.

Moving to **subdomains**, when measuring exposure to specific subdomains, data can be affected by the previously mentioned problems and some other specific problems. These specific problems can provoke that only information at the domain level is recorded (twitter.com) so no subdomains will be recorded (twitter.com/pabloiglesias). The specific sources are:

1) The behaviour happens in an Android device with the OS version lower than 10.

For those using browsers others than Chrome, Firefox and Samsung, only domain level (e.g. twitter.com) information is captured. Said differently, no information at the subdomain level (e.g. twitter.com/pabloiglesias) is captured. Combining information from question 2c and the device information provided by Netquest, we can identify participants in this situation. As seen in tables 3 and 4, around 50% of Android devices run a version lower than 10. Nevertheless, the market share of other browsers is really small. Hence, if this is reproduced on the metered panel, we should expect this problem to affect a low proportion of individuals.

2) The behaviour happens in an iOS device.

HTTPs information at the subdomain-level information cannot be collected from iOS devices. Hence, for Twitter or most of the main national news outlets, subdomain-level information cannot be collected from iOS devices. Netquest provides information on the devices tracked. Hence, we will be able to know which participants are tracked on an iOS device without having to ask any question.

3) The behaviour happens in a smartphone app

Information of what individuals do inside their apps is not available. Therefore, all the information about visiting specific profiles or news inside an app is not available either. There are, anyway, ways to address this. For instance, for Twitter:

a) For those using smartphone devices, we can check using metered data the time that they spent on Twitter.

a. **For those who spent 0 or close to 0 minutes on Twitter through the app we can be confident that no data is missing because of behaviours happening on an app**.
b. **For those who only spent time on the Twitter app, we can be confident that all information will be missing**. Hence, it can be considered as missing.
c. For those who spent time on Twitter through the app but also on the browser, we know that behaviours of interest might not be tracked.

These strategies can be applied to other subdomains.

What to do with this information

With the information gathered we can have some information to assess our data and, differentiate (to some extent) between missingness and measurement error. Using questions 1, 2 and 3, it is possible to identify those who are completely or partially tracked in terms of devices and browsers. Therefore, the following proportions can be calculated from the sample:

- % With all devices tracked
- % With all browsers tracked, for those devices tracked

Using this information, three groups can be created:

1) % Completely covered (device & browser)
2) % Completely undercovered
3) % Partially covered

Using this information, we already can make some decisions regarding when no information should be treated as missing or "0".

Completely covered.

– For information at the domain level, absence of behaviour should be treated as 0, always. We are sure that we are tracking all their behaviour (except some minor errors). If information is not there, we can consider that as a real absence of behaviour.
– For information at the subdomain level, it depends on other factors.

- If the webpage is HTTPs and a participant is not tracked on an iOS device, absence of behaviour should be treated as 0. There is no reason to expect data to be missing.
- If the webpage is HTTPs, the participant is tracked on an iOS device (only or among other devices) and no behaviour has been tracked at the domain level, the absence of behaviour should be treated as 0. There is no reason to expect data to be missing at the subdomain level since no domain level behaviour has been reported.
- If the webpage is HTTPs, a participant is tracked on an iOS device (only or among other devices) and behaviour at the domain level has been tracked, absence of behaviour at the subdomain level should be treated as missing. We know that the subdomain level behaviour of interest could have only happened in the iOS device. But subdomain-level information for HTTPs webpages cannot be collected. Therefore, the potential absence of behaviour cannot be accurately considered as 0. The absence of behaviour, in this case, could be treated as a missing.
- If the participant is not tracked on an Android with an OS version lower than 10, absence of behaviour should be treated as 0. There is no reason to expect data to be missing.
- If the participant is tracked on an Android with an OS version lower than 10 (only or among other devices), we know that he or she uses a browser other than chrome/FF/Samsung and no behaviour has been tracked at the domain level, the absence of behaviour should be treated as 0. There is no reason to expect data to be missing at the subdomain level since no domain level behaviour has been reported.
- If the participant is tracked on an Android with an OS version lower than 10 (only or among other devices), we know that he or she uses a browser other than chrome/FF/Samsung, no subdomain level information has been tracked in other devices/browsers, and behaviour at the domain level has been tracked in the device*browser in question, we know that the subdomain level behaviour of interest could have only happened in the given device. But subdomain-level cannot be collected. Therefore, the potential absence of behaviour cannot be accurately considered as 0. The absence of behaviour, in this case, could be treated as a missing.

- If the participant is tracked on a mobile device, entirely or among other devices, and a 0% of the behaviour happens on an app, the absence of behaviour should be considered as 0. We know that no information could be missed.
- If the participant is tracked on a mobile (only or among other devices), no subdomain level information has been tracked in other devices/browsers, and behaviour at the domain level has been tracked on a mobile device's app, the subdomain level behaviour of interest could have only happened in the app of a mobile device. But within app information cannot be collected. Therefore, the potential absence of behaviour cannot be accurately considered as 0. The absence of behaviour, in this case, could be treated as a missing.

Completely undercovered.

– For information at the domain and subdomain level, an absence of behaviour should be treated as missing, always. We are sure that we are not tracking any of the devices. All information should be missing.

Partially covered.

- With the available information it is not possible to determine accurately if an absence of a general behaviour (which can happen in all devices, browsers or apps) should be treated as a missing or as a 0. For specific behaviour which only happen in a specific device, browser, OS or app, we can sometimes be confident to interpret lack of information as 0 or missing. This, however, must be analysed case by case. For instance, if we are interested in a specific mobile device internet behaviour (e.g. total time using Tinder during the last 15 days), those with no behaviour tracked that self-report using no mobile devices, we can safely assume that they should be considered as 0. For those which say that use a mobile device but are not tracked with a mobile device, we can safely treat them as missing.

What to do with the unsolved problems?

How to correct for an absence of behaviour for those partially covered?

For those cases in which the available information is not enough to accurately determine if the absence of behaviour should be treated as a missing or as a 0, more information can be obtained. However, this information needs to be concept specific, which can increase substantially the number of questions. We propose to use the following question, which can help to obtain information for different domains or group of domains in the same question.

For each domain asked, if a partially tracked participant answers "yes" we will know that some behaviour has not been tracked. Hence, those partially tracked with an absence of behaviour in a given domain answering "yes" should be considered as missing. For those answering "no", an absence of behaviour can safely be considered as 0. This could be applied to subdomains, but it would require a more complex question and adding many more questions, which I would not recommend.

[PROGRAMMER: FOR THE FOLLOWING QUESTION WE ARE GOING TO COMBINE PARADATA FROM THE METER AND THE PREVIOUS ANSWERS FROM QUESTIONS 1-4.]

[PROGRAMMER: GO TO met5 IF ANY OF THESE CONDITIONS ARE TRUE:

IF met1a>= 1 & OS =! WINDOWS

IF met1b>= 1 & OS =! MAC

IF met1c>= 1 & OS =! ANDROID

IF met1d>=1 & OS =! iOS

IF met1e>= 1

IF met2a== YES

IF met2b== YES & KIND =! CHROME_PLUGIN

IF met2c== YES & KIND =! FIREFOX_PLUGIN

IF met2d== YES

IF met3a== YES

IF met3b== YES & KIND =! SAFARI_PLUGIN

IF met3c== YES & KIND =! CHROME_PLUGIN

IF met3d== YES & KIND =! FIREFOX_PLUGIN

IF met3e== YES

IF met4d== YES & OS VERSION >= 10

IF NONE OF THESE CONDITIONS ARE TRUE, GO TO QUESTION met6]

met5: During the last 15 days, have you used another device or browser apart from **[INSTER DEVICE(S)]** to visit the following web pages or apps:

		Yes	No
met5a	Twitter	o	o
met5b	Facebook	o	o
met5c	La Vanguardia	o	o
met5d	El Mundo	o	o
met5e	ABC	o	o
met5f	El Pais	o	o
met5g	20minutos	o	o
met5h	El Espanol	o	o
met5i	El Confidencial	o	o
met5j	OkDiario	o	o
met5k	El HuffingtonPost	o	o
met5l	Eldiario.es	o	o

[PROGRAMMER: IN met5 WE NEED TO INSERT THE DEVICES THAT PARTICIPANTS ARE TRACKED WITH. TO DO SO, WE NEED AGAIN TO USE PARADATA FROM THE METER. WE HAVE PREPARED THE PHRASES TO USE DEPENDING ON THE DEVICE THAT THEY ARE TRACKED WITH.

WHAT [DEVICE(S)] TO CONSIDER AND THE WORDING FOR EACH [DEVICE] IS PROVIDED IN THE TABLE BELOW

DEVICE	WORDING
[If Desktop= yes & OS = windows]	a computer with Windows
[If Desktop= yes & OS = MAC]	an Apple computer (MAC)
[If Browser plugin= yes & OS= windows & kind = chrome_plugin]	a Chrome browser on a computer with Windows
[If Browser plugin= yes & OS= windows & kind = firefox_plugin]	a Firefox browser on a computer with Windows
[If Browser plugin= yes & OS= MAC & kind = chrome_plugin]	a Chrome browser on an Apple computer (MAC)
[If Browser plugin= yes & OS= MAC & kind = firefox_plugin]	a Firefox browser on an Apple computer (MAC)
[If Browser plugin= yes & OS= MAC & kind = safari_plugin]	a Safari browser on an Apple computer (MAC)
[If OS = android & manufacturer= X]	a [manufacturer] smartphone or table with Android
[If OS = osx]	an Apple smartphone or tablet (iPhone or iPad)

IF THERE ARE TWO DEVICES, SEPRATARE THEM USING THE FOLLOWING STRUCTURE: [DEVICE] and [DEVICE]

IF THERE IS MORE THAN ONE DEVICE, SEPARATE THEM USING THE FOLLOWING STRUCTURE: [DEVICE], [DEVICE],… and [DEVICE]

FOR EXAMPLE, FOR SOMEONE WITH: [If Browser plugin= yes & OS= windows & kind = chrome_plugin], [If OS = android & manufacturer= samsung] and [If OS = osx]: During the last 15 days, have you used another device or browser apart from a Chrome browser on a computer with Windows, a Samsung smartphone or tablet with Android and an Apple smartphone or tablet (iPhone or iPad) to visit…

What to do with Twitter subdomain information?

As previously seen, for subdomain level information, there is a higher chance of not being able to determine if an absence of behaviour should be considered as 0 or missing. This can be especially problematic for Twitter, which is an HTTPs web as well as a mobile app. For instance, for those accessing Twitter through the app and the browser version, it might be difficult to know if no recorded visit to twitter.com/pabloiglesias means a real lack of behaviour or a tracking problem. In the case of Twitter, for those in the experimental condition, we can use as a proxy the total time spent during the days of the experiments in twitter.com and compare it with preceding time using the same domain. The increase in the time spent on the domain/app Twitter might be inferred to be produced by the treatment condition i.e. they spend more time on twitter getting information about the issues/profiles presented. This is, however, an imperfect proxy: participants might increase their time for other reasons. Besides, they might as well spend more time checking other things, not the ones asked.

What to do with iOS devices?

Compared with other devices, iOS devices are the ones presenting more problems:

- The accuracy of proxies is most likely lower than other tracking technologies,
- Incognito browsing cannot be tracked,
- Subdomain level information cannot be collected for HTTPs domains,
- Network undercoverage cannot be assessed.

We should expect, therefore, to experience more problems for those tracked on iOS devices. An approach to avoid this would be to exclude individuals tracked on iOS devices from the

pool of sampling participants. By doing so, problems related to iOS devices would be avoided. Nevertheless, this could introduce other problems which might be more problematic than the ones introduced by including iOS devices:

- By default, not all the online behaviour is being tracked. All the constructs of interest exclude iOS devices, assuming from the beginning that no inference can be made about the general online behaviour but only about the behaviour for specific devices.
- It can exclude participants which are only tracked on iOS devices. It has been found that ownership of smartphones with specific OSs is correlated with several sociodemographic and substantive variables. For instance, in Germany, iPhone owners constitute a much smaller and more unique subpopulation, both in terms of sociodemographic characteristics and in terms of attitudes and behaviors (e.g., Keusch et al., 2020; Götz et al. 2017; Pryss et al. 2018; Shaw et al. 2016; Ubhi et al. 2017). Excluding iOS owners might make the sample of participants even less representative. Besides, weighing strategies have been found not to completely eliminate the bias introduced by OS ownership (Keusch et al., 2020).
- Excluding participants tracked on iOS devices can increase the chance to sample undercovered individuals. Those which are not tracked on an iOS device can be 1) non-iOS owners or 2) undercovered iOS owners. Individuals undercovered in one device might have a higher likelihood to be undercovered in other devices/OS. If this is the case, excluding individuals tracked on an iOS might increase the probability of tracking individuals which have a higher likelihood of being undercovered.

All in all, it is unclear which option can have a more negative impact on data quality. Therefore, we consider that it is best to sample all participants and assess the impact of including or excluding them. Sampled participants tracked on iOS devices can always be excluded from final analyses (with proper weighting adjustments) if needed. However, if they are excluded by design, no information whatsoever will be available.

Consideration about robustness

We expect the different problems exposed in this document to affect to some extent the quality of the data and, consequently, the substantive results. Nevertheless, with the information available to date, it is not possible to stablish the exact effect that the different

errors have on the results. Similarly, the strategies proposed here to correct for some of these potential errors have not been tested yet. How these can affect results is also difficult to predict. Therefore, once the data is available, we propose exploring the robustness of the results when using uncorrected data or different correction strategies. Only this way we will be able to understand how much of a difference it makes to correct results, using different approaches.