

Track me but not really:

Device undercoverage and its consequences when tracking online behaviours.

ORIOL J. BOSCH | THE LONDON SCHOOL OF ECONOMICS / RECSM-UPF



THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE ■



Universitat
Pompeu Fabra
Barcelona



This project has received funding from 1) the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No849165), PI: Melanie Revilla; 2) Ministry of Economy and Competitiveness, State Programme for the Promotion of Scientific and Technical Research of Excellence (PID2019-106867RB-I00/AEI/10.13039/501100011033 (2020-2024), PI: Mariano Torcal; 3) AYUDAS FUNDACIÓN BBVA A EQUIPOS DE INVESTIGACIÓN CIENTÍFICA EN ECONOMÍA Y SOCIEDAD DIGITAL 2019, PI: Mariano Torcal.

Tracking online behaviours using a meter

Definition

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**. The meter is willingly installed on participants devices.

Tracking online behaviours using a meter

Definition

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**. The meter is willingly installed on participants devices.



Resulting data: **metered data**

Tracking online behaviours using a meter

Metered data can potentially suffer from different types of errors

Bosch, O.J., and M. Revilla (2021). **“When survey science met online tracking: presenting an error framework for metered data.”**

RECSM Working Papers Series, 62

Tracking online behaviours using a meter

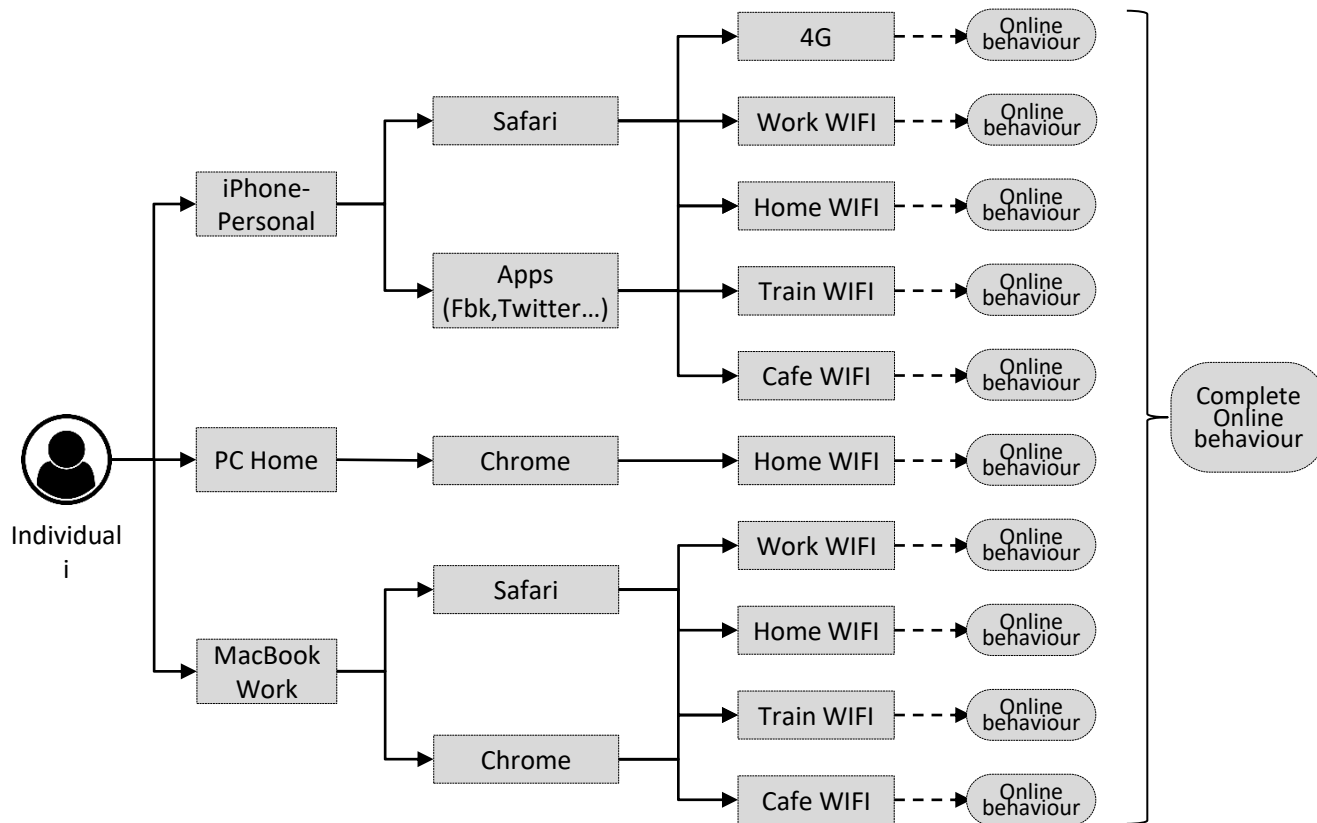
Metered data can potentially suffer from different types of errors

Bosch, O.J., and M. Revilla (2021). “**When survey science met online tracking: presenting an error framework for metered data.**”

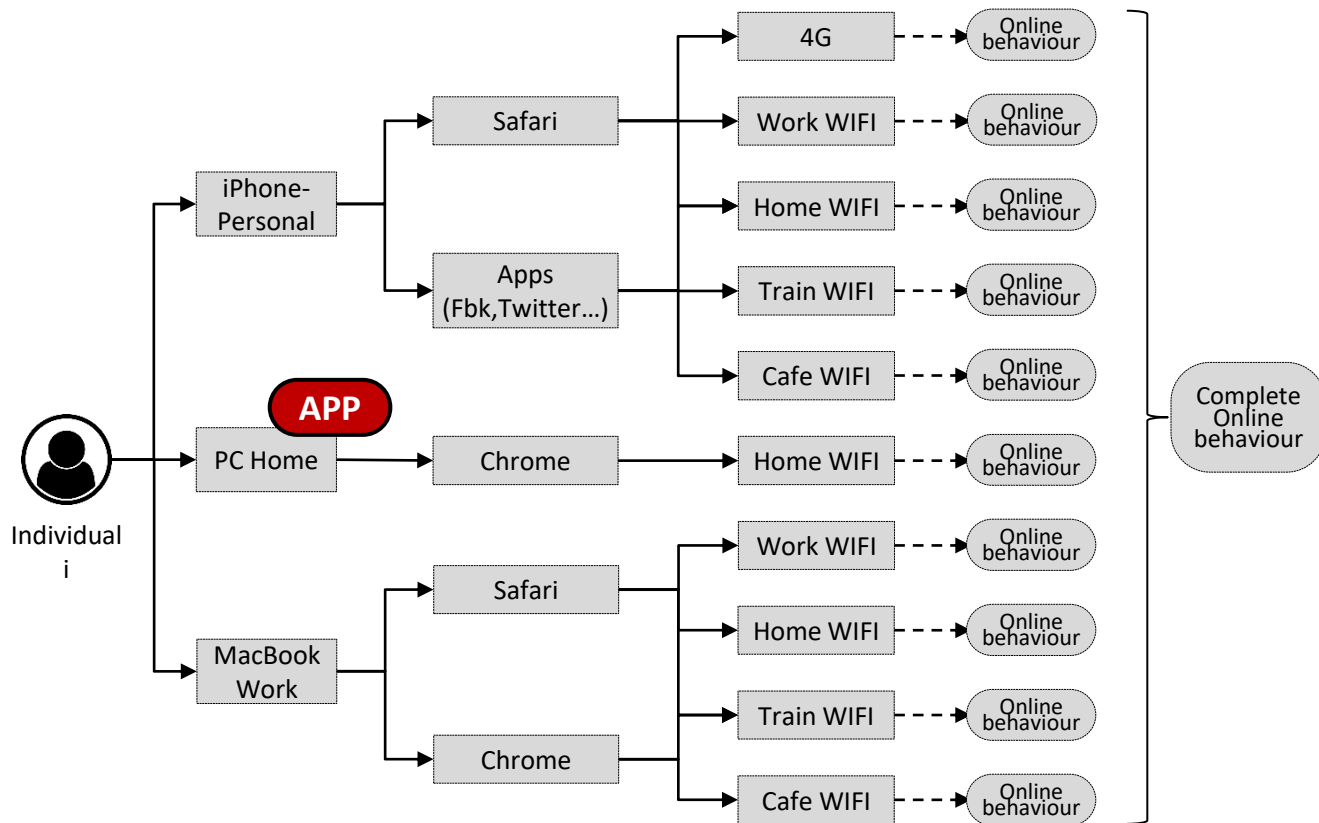
RECSM Working Papers Series, 62

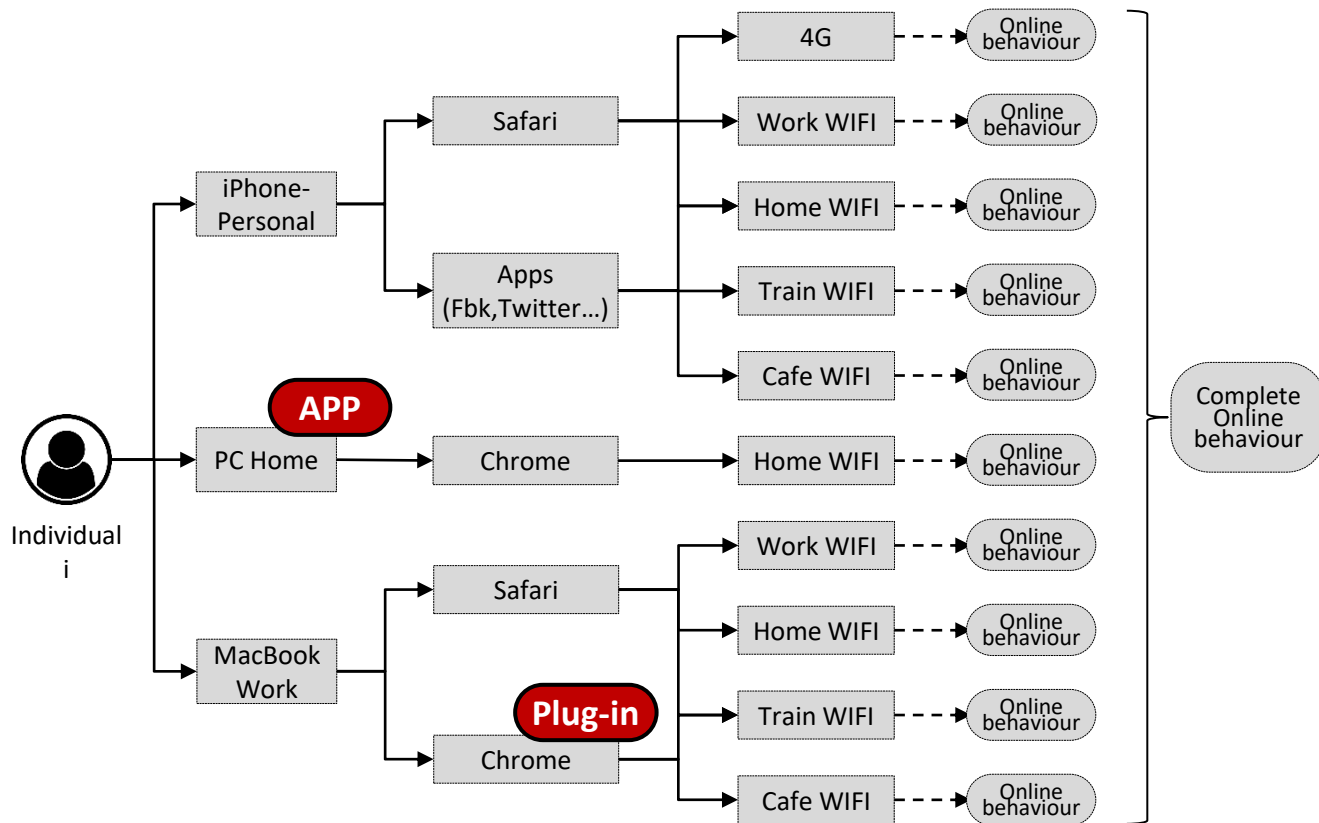


Our focus: problems of **undercoverage**

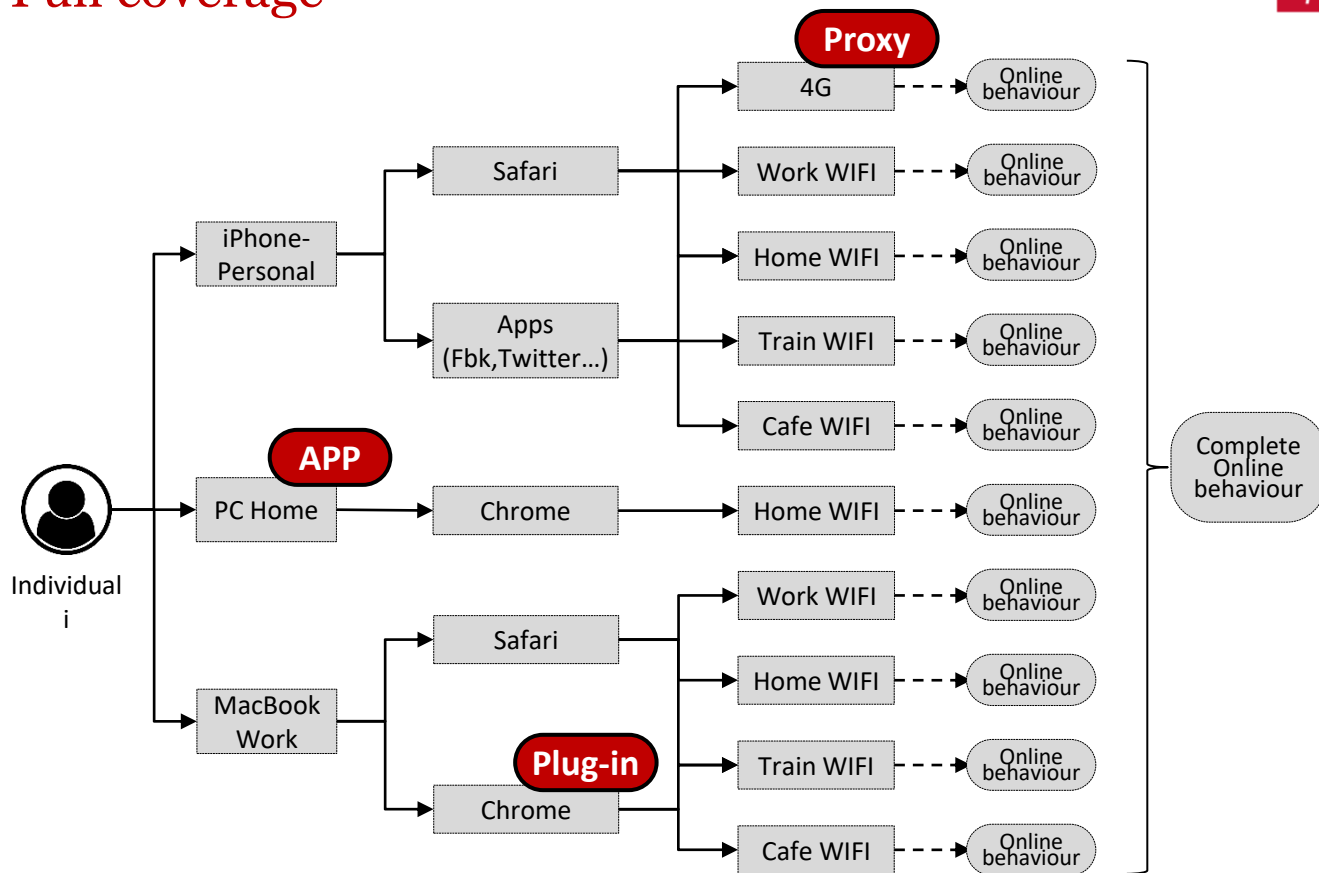


Full coverage





Full coverage

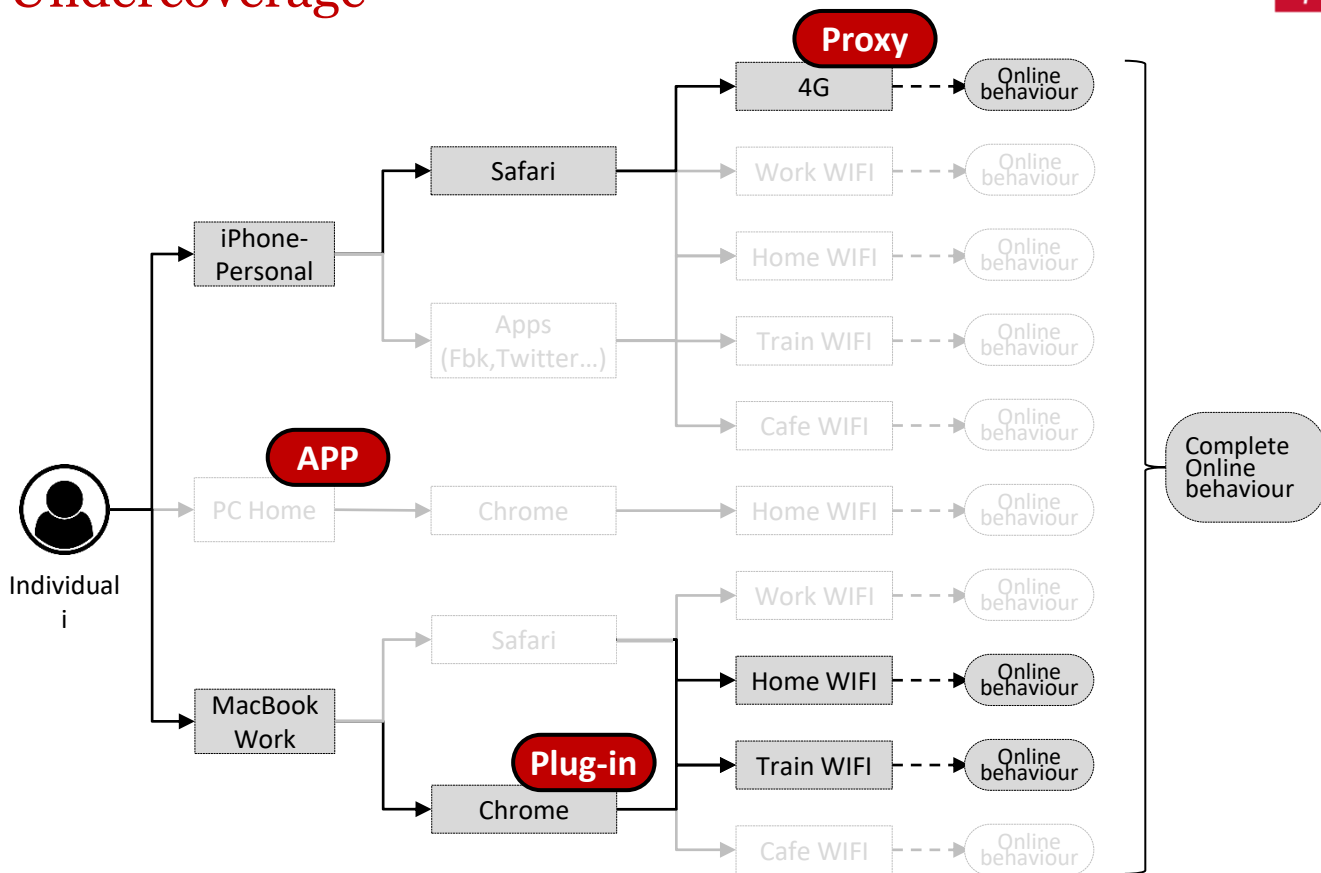


Undercoverage can prevent tracking the complete online behavior.

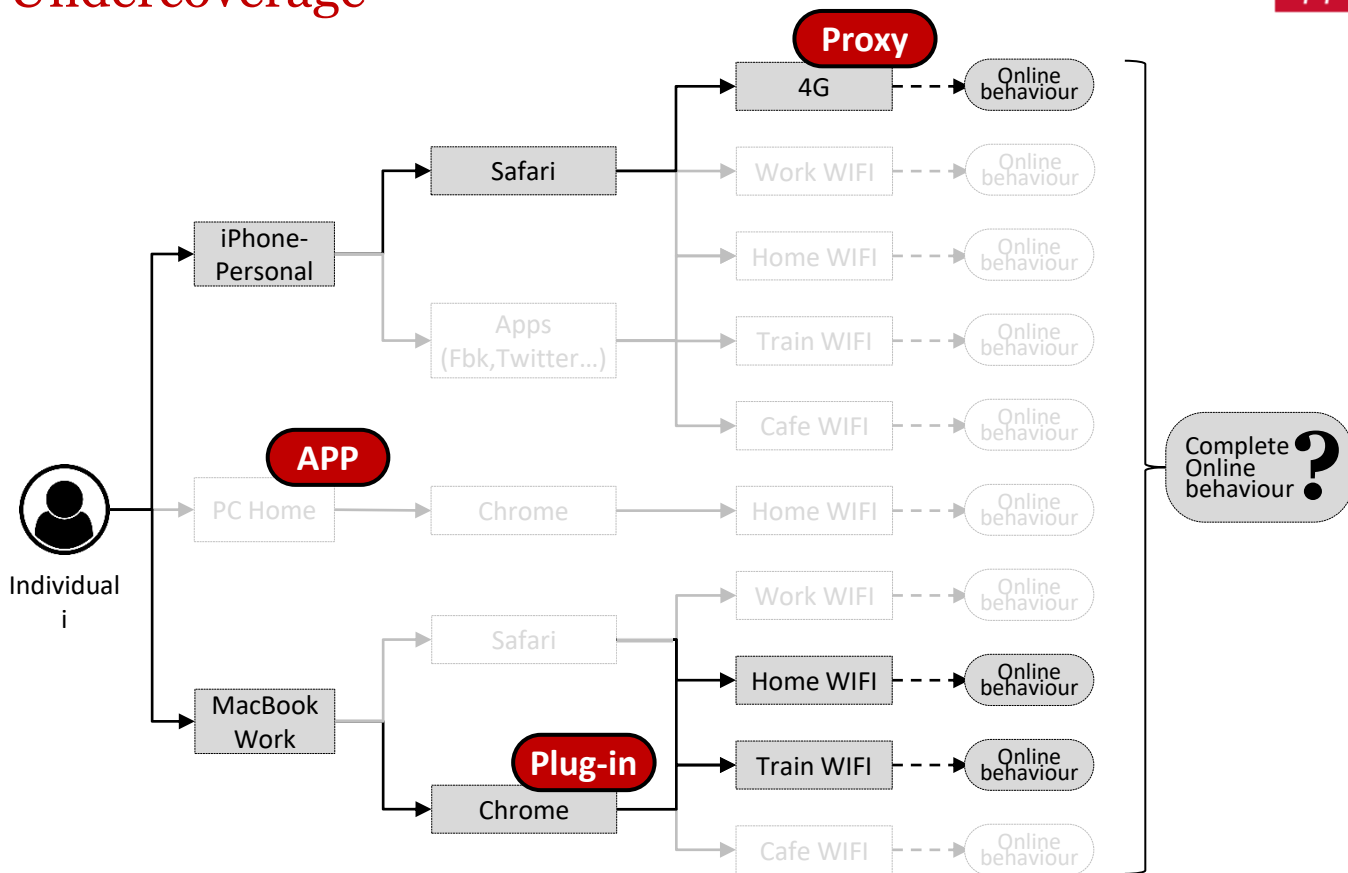
Different levels of undercoverage.

- **Device:** at least one device used by a participant is not tracked
- **Browser:** at least one web-browser used by a participant is not tracked
- **In-app:** the behaviours happening inside apps are not tracked.
- **Network:** at least one network from which a participant connect to the Internet is not tracked

Undercoverage



Undercoverage



Undercoverage

57%

meter installed
in only one
device.

4%

used only one
device to access
the Internet.

Netquest (Spain)

(Revilla, Ochoa and Loewe, 2017)

28%

fully covered.

Ipsos Knowledge Panel (USA)

(Pew Research Center, 2020)

Main goals

#1 **Assess** prevalence of device, browser and in-app undercoverage

- Who is fully covered?
- How does undercoverage look like?
- What drives undercoverage?

#2 **Investigate** how to deal with the consequences of undercoverage.

Data from TRI-POL project

- Data collection planned for the end of this spring
 - Surveys, metered data and paradata
 - Netquest metered panels
 - Italy, Portugal and Spain
- Questionnaire about polarization, political trust and political communication
- Focus on 6 questions asked to assess undercoverage and its consequences
- Survey information is combined with device paradata and metered data

THIS STUDY

Assess prevalence (of device, browser and in-app undercoverage)



Combining survey and paradata

Combining survey and paradata

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.

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

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

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

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

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

Combining survey and paradata

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.

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

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

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

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

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

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
Internet Explorer	<input type="radio"/>	<input type="radio"/>
Chrome	<input type="radio"/>	<input type="radio"/>
Firefox	<input type="radio"/>	<input type="radio"/>
Edge, Opera or others	<input type="radio"/>	<input type="radio"/>

Combining survey and paradata

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.

Computer with Windows operating system: [NUMERIC OPEN BOX]

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

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

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

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

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?

Internet Explorer	
Chrome	
Firefox	
Edge, Opera or others	

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
Internet Explorer	<input type="radio"/>	<input type="radio"/>
Safari	<input type="radio"/>	<input type="radio"/>
Chrome	<input type="radio"/>	<input type="radio"/>
Firefox	<input type="radio"/>	<input type="radio"/>
Edge, Opera or others	<input type="radio"/>	<input type="radio"/>

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

	Yes	No
Chrome	<input type="radio"/>	<input type="radio"/>
Samsung browser	<input type="radio"/>	<input type="radio"/>
Firefox	<input type="radio"/>	<input type="radio"/>
Edge, Opera or others	<input type="radio"/>	<input type="radio"/>

Combining survey and paradata

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.

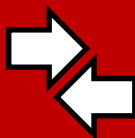
Computer with Windows operating system: [NUMERIC OPEN BOX]

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

Smartphone

Apple

Other



*Compare this information with device **paradata**: Information about **all** the devices and browsers in which they are tracked .*

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

Internet Explorer	
Chrome	
Firefox	
Edge, Opera or others	

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
Internet Explorer	<input type="radio"/>	<input type="radio"/>
Safari	<input type="radio"/>	<input type="radio"/>
Chrome	<input type="radio"/>	<input type="radio"/>
Firefox	<input type="radio"/>	<input type="radio"/>
Edge, Opera or others	<input type="radio"/>	<input type="radio"/>

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

	Yes	No
Chrome	<input type="radio"/>	<input type="radio"/>
Samsung browser	<input type="radio"/>	<input type="radio"/>
Firefox	<input type="radio"/>	<input type="radio"/>
Edge, Opera or others	<input type="radio"/>	<input type="radio"/>

What can we get from this?

- We can estimate the undercoverage at different levels:
 - % fully covered;
 - % fully covered for device/browser/in-app;
 - % with 1, 2, 3, ... n devices/browsers not covered;
 - % with type of device (e.g. iPhone) or browser (e.g. Safari) not covered.
- We can investigate the characteristics of those partially covered.

How to deal with the consequences of undercoverage

How should we consider a lack of behaviour?

	CodPanelista	News_Views	Avg_News_V-s	News_Time	Avg_News_T-e
1	000140569144c377	79	11	8520	1217
2	0007a79c5a045b66	23	6	855	214
3	0017a3da81690a8d	62	6	213	19
4	001b5fdac4fa8d7f
5	002714f25b438c0d	46	4	2701	225
6	003520bac736b3fa
7	003c160d88367c30	23	3	161	20
8	00704967cf90b380
9	00a3b5fa16f2b567	227	11	6025	287
10	00c227fc0078a550	48	4	926	77
11	00d2c3acfa54d30a	71	4	2398	133
12	00e14c6e57df3fa6	85	7	5108	393
13	011d33ab6124791b	40	4	1532	139
14	013b734186ad95ab

How to deal with the consequences of undercoverage

How should we consider a lack of behaviour?

	CodPanelista	News_Views	Avg_News_V-s	News_Time	Avg_News_T-e
1	000140569144c377	79	11	8520	1217
2	0007a79c5a045b66	23	6	855	214
3	0017a3da81690a8d	62	6	213	19
4	001b5fdac4fa8d7f
5	002714f25b438c0d	46	4	2701	225
6	003520bac736b3fa
7	003c160d88367c30	23	3	161	20
8	00704967cf90b380
9	00a3b5fa16f2b567	227	11	6025	287
10	00c227fc0078a550	48	4	926	77
11	00d2c3acfa54d30a	71	4	2398	133
12	00e14c6e57df3fa6	85	7	5108	393
13	011d33ab6124791b	40	4	1532	139
14	013b734186ad95ab

How to deal with the consequences of undercoverage

How should we consider a lack of behaviour?

	CodPanelista	News_Views	Avg_News_V-s	News_Time	Avg_News_T-e
1	000140569144c377	79	11	8520	1217
2	0007a79c5a045b66	23	6	855	214
3	0017a3da81690a8d	62	6	213	18
4	001b5fdac4fa8d7f
5	002714f25b438c0d	46	4	2701	225
6	003520bac736b3fa
7	003c160d88367c30	23	3	161	20
8	00704967cf90b380
9	00a3b5fa16f2b567	227	11	6025	287
10	00c227fc0078a550	48	4	926	77
11	00d2c3acfa54d30a	71	4	2398	133
12	00e14c6e57df3fa6	85	7	5108	393
13	011d33ab6124791b	40	4	1532	139
14	013b734186ad95ab

0 minute visiting
news outlets?

Missing
information?

How to deal with the consequences of undercoverage

Better identifying real lack of behaviour and missings

- For those completely covered, we can be quite certain that a lack of behaviour should be considered as 0.
- For those partially undercovered, it is still not clear what to do.

How to deal with the consequences of undercoverage

Better identifying real lack of behaviour and missings

- For those completely covered, we can be quite certain that a lack of behaviour should be considered as 0.
- For those partially undercovered, it is still not clear what to do

We need more data

We (still) need more data

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
Twitter	<input type="radio"/>	<input type="radio"/>
Facebook	<input type="radio"/>	<input type="radio"/>
The Guardian	<input type="radio"/>	<input type="radio"/>
BBC	<input type="radio"/>	<input type="radio"/>
CNN	<input type="radio"/>	<input type="radio"/>

We (still) need more data

During the last 15 days, have you used another device or browser apart from [INSTER DEVICE(S)] to visit the following website(s)?



If the person did not use another device or browser to visit the pages/apps of interest -> undercoverage does not affect this measure

If the person did use another device or browser to visit the pages/apps of interest -> undercoverage affects this measure

We (still) need more data

During the last 15 days, have you used another device or browser apart from [INSTER DEVICE(S)] to visit the following...



If the person did not use another device or browser to visit the pages/apps of interest -> undercoverage does not affect this measure

If the person did use another device or browser to visit the pages/apps of interest -> undercoverage affects this measure

Most likely cannot be done for every web page/app of interest

Undercoverage might be common and problematic

- Evidence suggests that a high proportion suffers from device undercoverage.
- More research is needed to understand the extent of this.
- Although undercoverage can potentially bias metered data measures, no research has been conducted.

Identifying when a lack of behaviour is real or a product of undercoverage is key

- Confounding both phenomena can inflate measurement and missing data errors.
- Gathering information is necessary to both correct for this issue and be transparent about the quality of the sample.

Thank you *Questions?*

ORIO L J. BOSCH | THE LONDON SCHOOL OF ECONOMICS / RECSM-UPF



o.bosch-jover@lse.ac.uk



[orioljbosch](https://twitter.com/orioljbosch)



<https://orioljbosch.com/>



THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE ■



Universitat
Pompeu Fabra
Barcelona

