Willingness to participate in geolocation-based research.

3rd MASS workshop

June 2022

CARLOS OCHOA | Research and Expertise Centre for Survey Methodology (RECSM – UPF)

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No849165), PI: Melanie Revilla.

GEOLOCATION DATA ARE…

GREAT…

Individuals' locations collected at a frequency and level of precision inconceivable using surveys.

- Reduced burden.
- Increased accuracy.

Applications:

- Identify individuals' locations and travel patterns [1]
- Detect individuals accessing pre-specified locations [2].

[1] Geurs, Veenstra and Thomas, 2013)
[2] Clemens and Ginnis, 2017

… BUT NOT PERFECT

ERRORS

Limited precision of the technologies used to geolocate devices (e.g., GPS).

- Example: wrong coordinates, undetected visit to a location of interest.

MISSING DATA

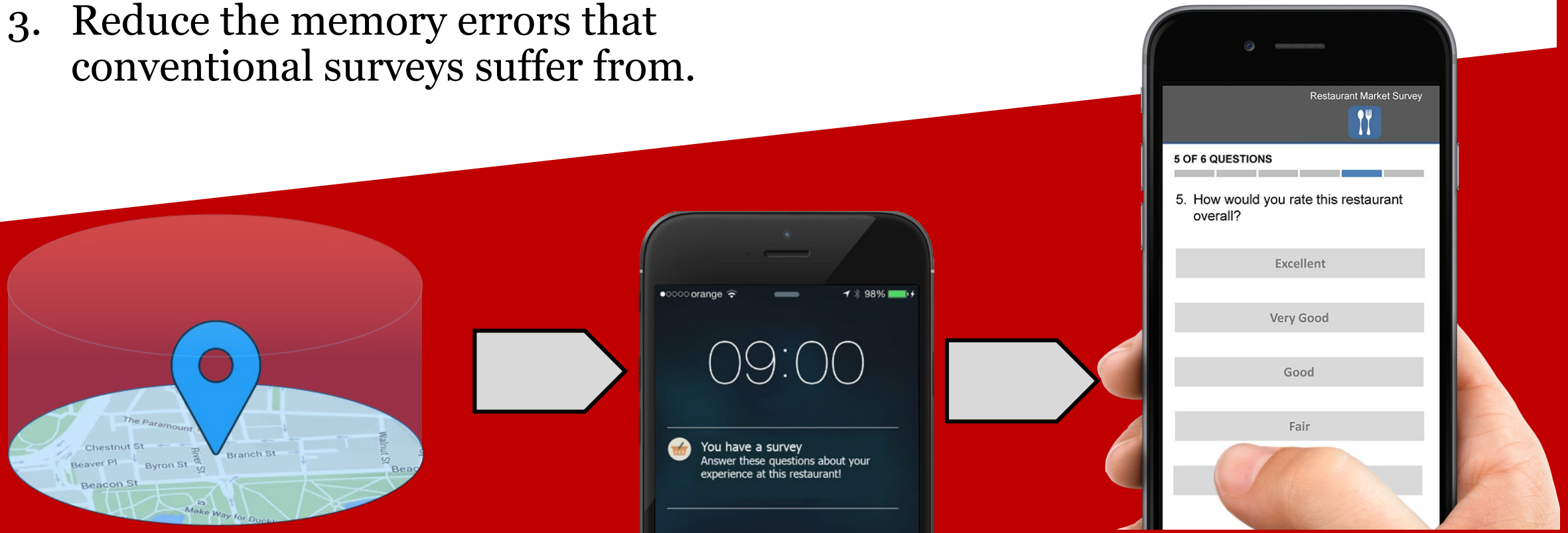
Subjective information cannot be observed using a passive tracker.

- Example: motivation of a travel, satisfaction with the mode of transport.

IN-THE-MOMENT SURVEYS

Sending a survey (to members of an online panel) right in the moment a location of interest is visited:

1. Add missing information.
2. Clarify doubtful information.
3. Reduce the memory errors that conventional surveys suffer from.



LIMITING FACTOR: WILLINGNESS TO PARTICIPATE

web data opp

Sharing geolocation data

- Already studied under different conditions. Willingness: 20% - 50%.
- Differences among participants not always consistent across studies.
- Little literature about the effect of the conditions offered to participants.

My contribution:

- Effect of project duration and incentives using a Conjoint analysis.
- More scenarios than previous literature.

In-the-moment surveys triggered by geolocation data

- A few actual experiences reported.
- No previous research on willingness to participate.
- Related research: willingness to participate in in-the-moment surveys triggered by metered data.

My contribution:

- Levels of willingness …
- … for combinations of 5 attributes.

RESEARCH QUESTIONS

RQ1 – What are the levels of **willingness to participate** in geolocation-based research among members of an online panel:

(a) share geolocation data

(b) in-the-moment surveys triggered by geolocation data.

RQ2 – How the **attributes** of geolocation-based research influence the willingness to participate?

Attributes: (1) project duration, (2) survey length, (3) invitation lifetime (time to participate in the survey), (4) geolocation incentive and (5) survey incentive level (compared to a conventional survey).

RQ3 – Are there **significant differences** among panelists?

Sociodemographic variables, personality traits, attitudes/habits and panel experience.

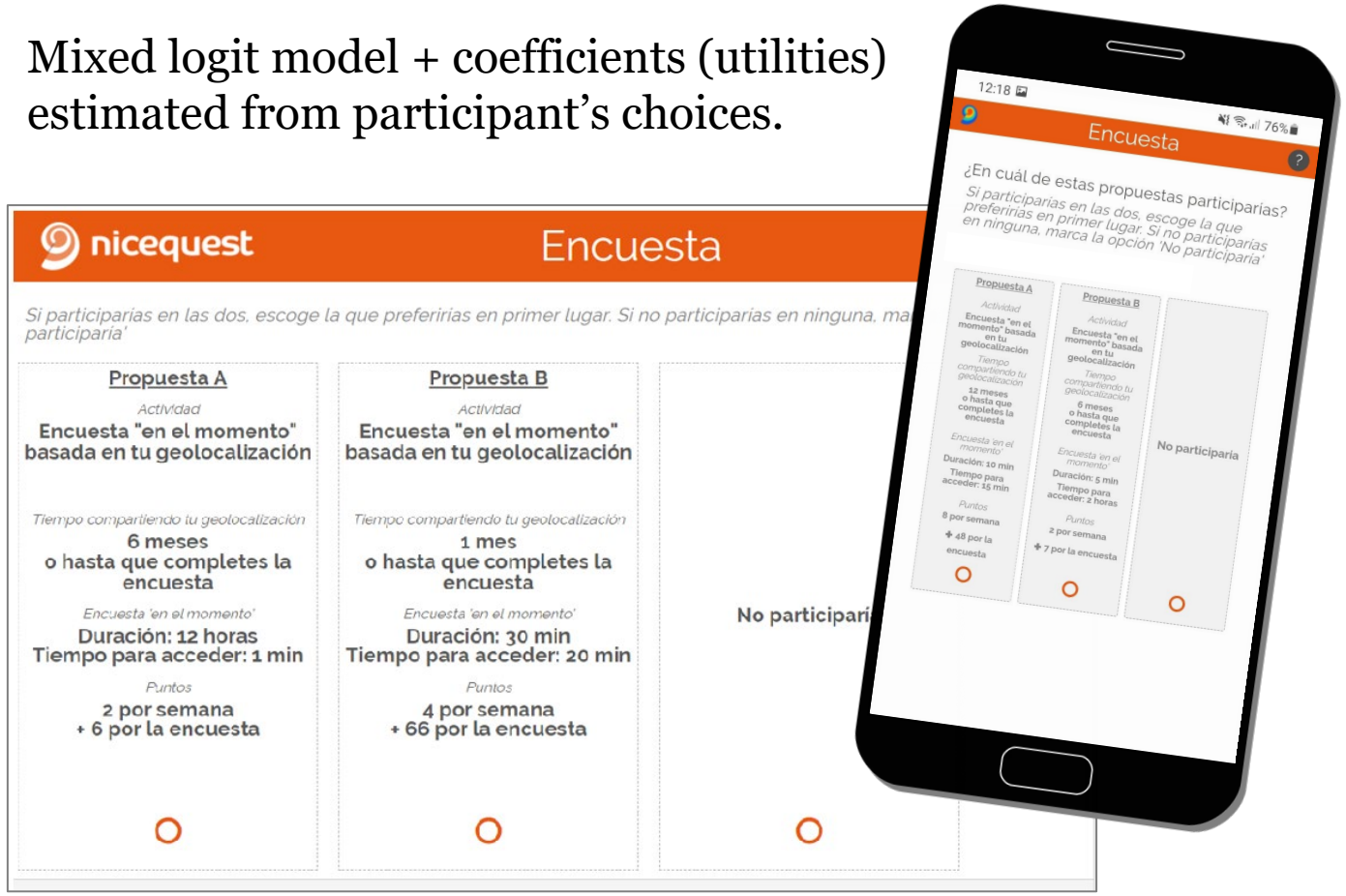
***RQ4** –Main reasons for deciding whether or not to participate stated by the panelists?*

DATA AND METHODS

- **N=1,016 valid surveys**
- Netquest opt-in online panel in Spain.
- 21st of Feb. – 7th of Mar. 2022.
- Mean survey length: 8.8 min.
- Quotas on age(3)+gender(2) and education(3).
- **<u>27% of the participants have installed a meter</u>** (already sharing online behaviors).

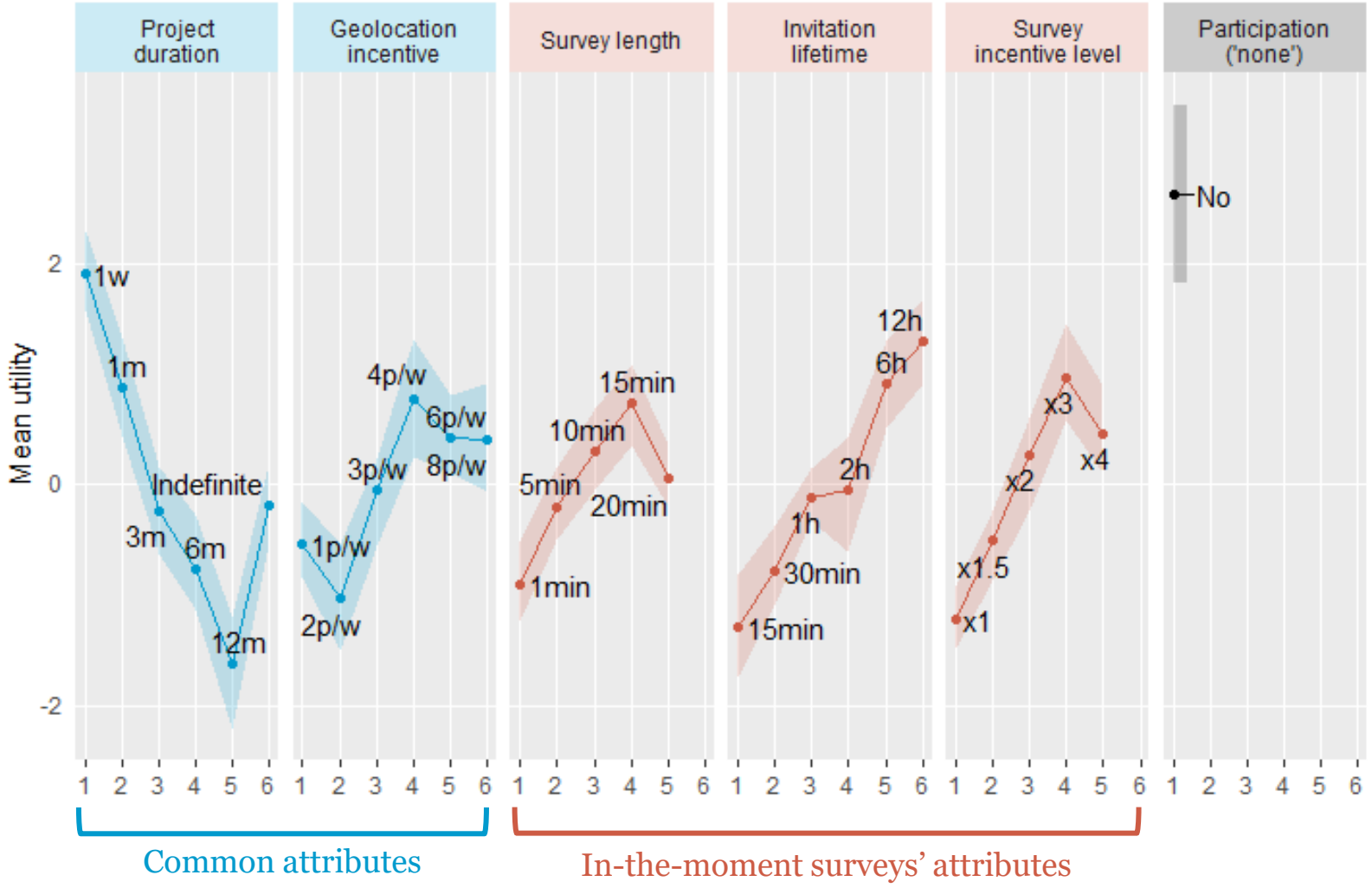
Choice based conjoint (CBC)

Mixed logit model + coefficients (utilities) estimated from participant's choices.



INFLUENCE OF EACH ATTRIBUTE-LEVEL

***Average utilities of the mixed logit model** (higher utilities = higher preference)*



Preference for:

1. Shorter project durations.
2. Larger survey lengths up to 15 min.
3. Larger invitation lifetimes.

and

4. Larger incentives (without significant differences between the two higher levels).

RESULTS

IMPORTANCE

% of variation of each attribute vs. total variation

Sharing geolocation

Attribute	Importance (%)	Percentile 2.5th	Percentile 97.5th
Project duration	64.3	54.6	74.5
Geolocation incentive	35.7	25.5	45.4

In-the-moment surveys

Attribute	Importance (%)	Percentile 2.5th	Percentile 97.5th
Project duration	29.6	24.9	37.0
Invitation lifetime	21.8	17.1	26.9
Survey incentive level	18.4	15.1	22.3
Geolocation incentive	16.4	12.3	22.6
Survey length	13.7	15.1	22.3

WILLINGNESS TO PARTICIPATE

Willingness to participate in three difference scenarios

Research activity	Scenario	Mean willingness (%)	Percentile 5th	Percentile 95th
Sharing geolocation	Best	50.1	46.8	53.7
	Average	43.2	41.1	45.1
	Worst	37.6	35.6	39.6
In-the-moment surveys	Best	57.1	55.2	59.3
	Average	47.2	46.6	47.8
	Worst	34.4	32.4	36.2

<u>Best scenario:</u> Survey / duration: 1 week / invitation lifetime: 12h / survey length: 15 min / 4 points per week / x3 survey incentive

<u>Worst scenario:</u> Geoloc / duration: 1 year / invitation lifetime: 15min / survey length: 1 min / 2 points per week / x1 survey incentive

DIFFERENCES AMONG PARTICIPANTS

Sociodemographic variables
- Moderate effects (5.5% <-> 12.1%)

Personality traits (attitudes)
- Moderate effects (4.5% <-> 15.9%)

Panel experience
- Past participations: only in the last 3 months (+10.8%)
- Metered panelist: +18.1%

Attitudes/habits
- Large effects:
  - Survey privacy concerns: -26.1%
  - Survey safety concerns: -22.9%
  - Sharing contents in SM: +38.7%
  - Installing apps: +26.2%
  - Google maps: +28.1%

PRELIMINARY CONCLUSIONS

1. In-the-moment surveys triggered by geolocation data:
   - Feasible in terms of willingness to participate.
   - But actual participation may differ substantially due to practical issues (not seeing the invitation in time).

2. To ensure high levels of willingness:
   - Short project durations with reasonable invitation lifetimes.
   - Up to 15 min survey length
   - Incentives are still key

3. When using quota sampling, variables other than sociodemographic variables should be considered

4. Developing geolocation-based research on "panelists already sharing online behaviors" may be effective and would allow us to research offline and online events.

Thanks!

CARLOS OCHOA | RECSM - UPF

Carlos.ochoa@upf.edu

https://www.upf.edu/web/webdataopp

Universitat Pompeu Fabra Barcelona

RECSM Research and Expertise Centre for Survey Methodology

webdataopp

INDEX

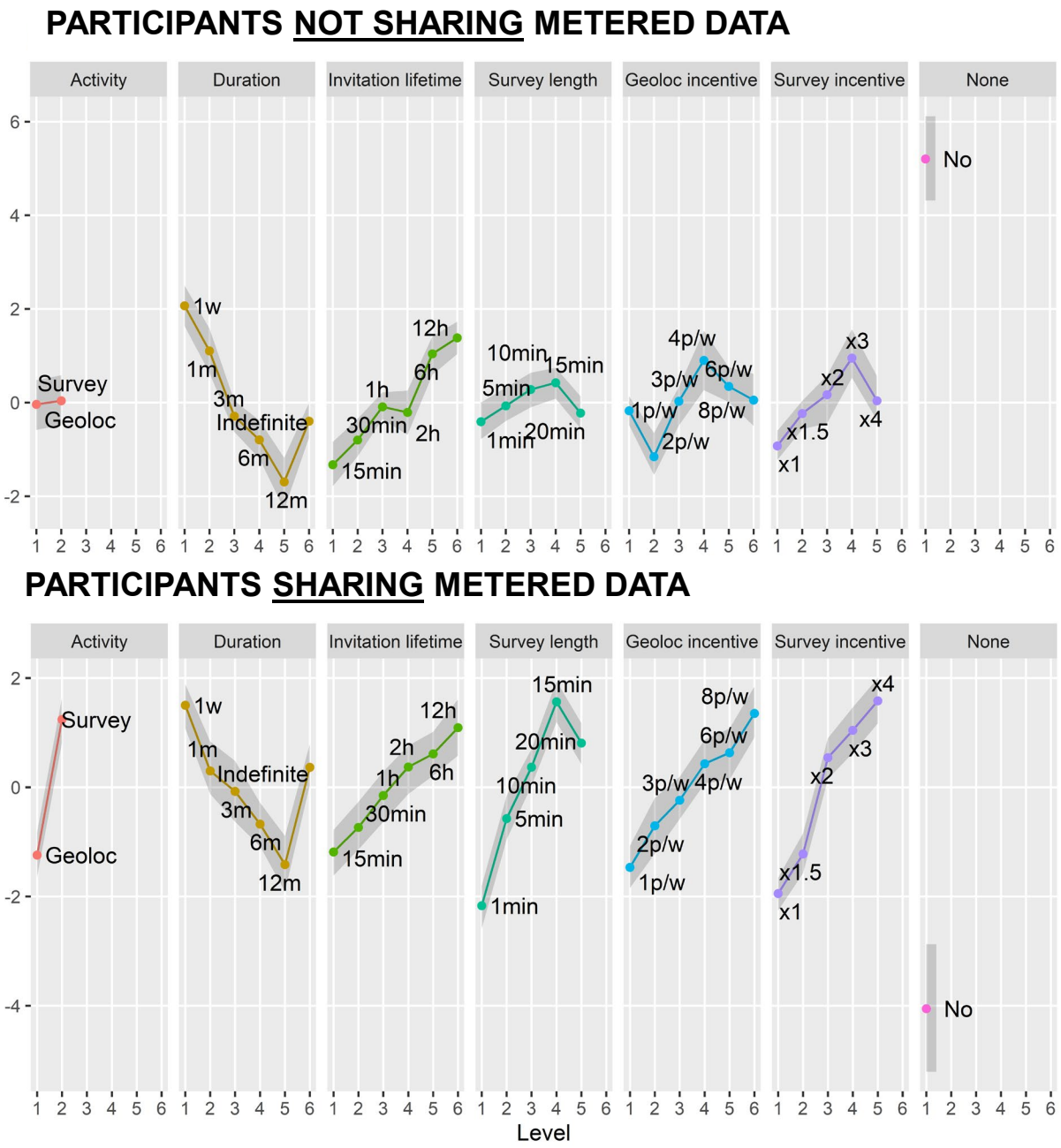
RESULTS

DIFFERENCES

Few differences among participants…

BUT being sharing metered data makes a big difference (as expected).

	Avg. WTP
Metered panelists	40.2%
Non-metered panelists	58.9%



WILLINGNESS TO PARTICIPATE IN GEOLOCATION-BASED RESEARCH

IN THE MOMENT SURVEYS

moment
d.

ion.

s that
er from

Introduction: the problem

About how we remember

Major classes of memory problems[2]

1. **Non-encoding**

   We may never form a representation of an event in our memory

2. **Post-encoding errors**

   Errors introduced after the original encoding.

3. **Retrieval failures**

   We cannot remember the information that is there.

4. **Reconstruction errors**

   We fill in missing details based on our general knowledge.

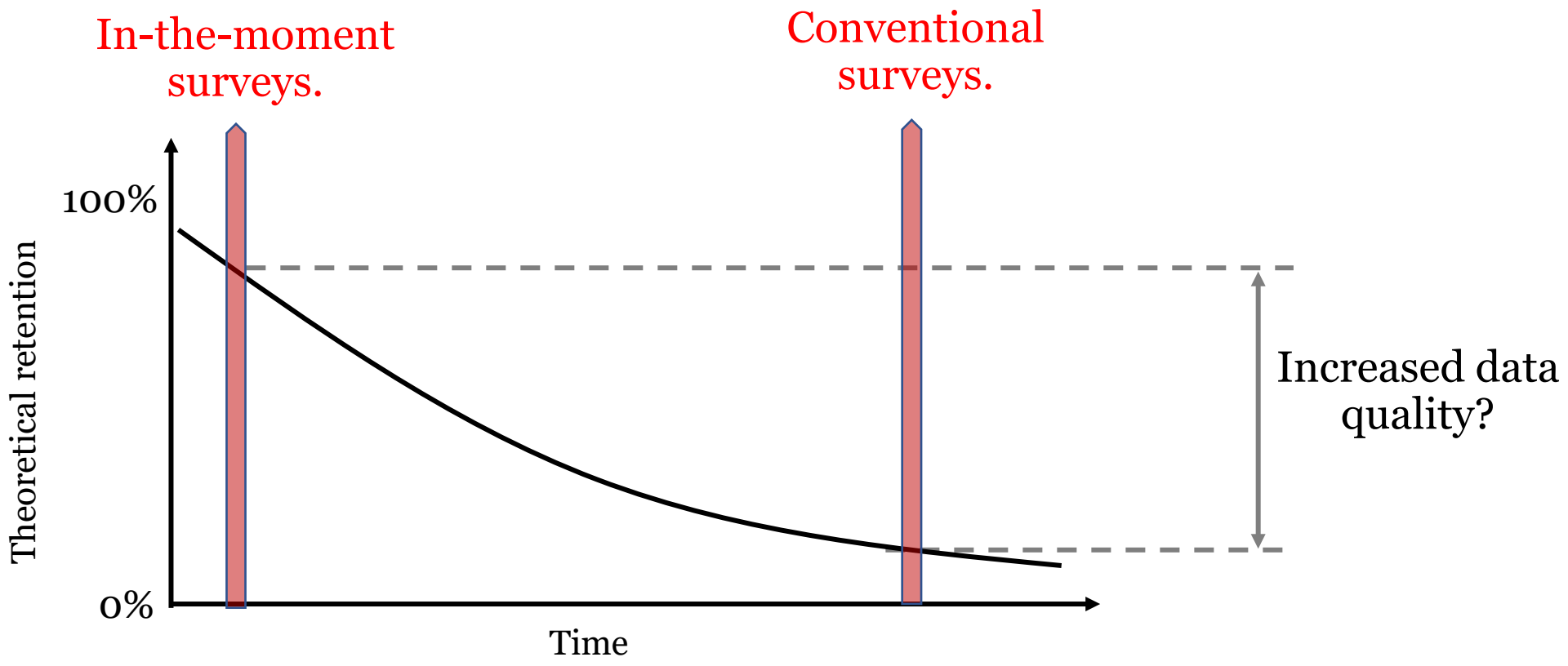
Factors increasing the chances of suffering memory errors:

+ Many events of the same category (e.g., supermarket visits)
+ Low distinctiveness
+ Low emotional impact
+ Short duration
+ Non-rehearsal (time spent thinking or talking about the event).

+TIME!

In-the-moment surveys

Surveying a sample of individuals **right in the moment** – or short time after – an event of interest happens may reduce memory errors.



SURVEY INCENTIVE LEVEL - CLARIFICATION

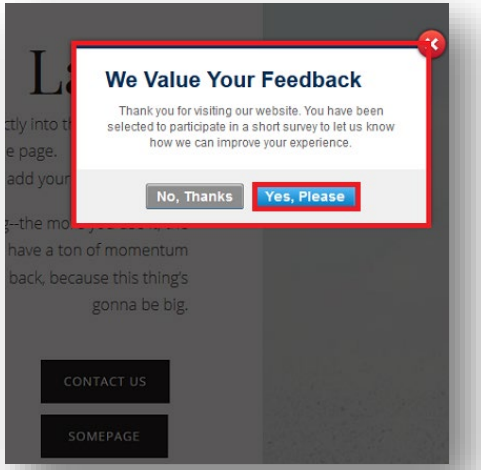
- Survey incentive level: 1, 1.5, 2, 3 and 4 times a conventional survey.

Survey incentive levels were not shown as such to participants. Instead, they saw the number of points that they would get for answering the survey. For instance, a conventional 10-minute survey would be rewarded with 12 points according to the existing panel policy; if the incentive level was 1, the participants saw that they would get 12 points, whereas if the incentive level was 2, they saw that they would get (2x12=)24 points. This design does not allow measuring the effect of the total number of points on the willingness to participate as it was not the purpose of this research. Only the effect of the incentive level (compared to a conventional survey) is measurable (see also Ochoa and Revilla, 2022b).

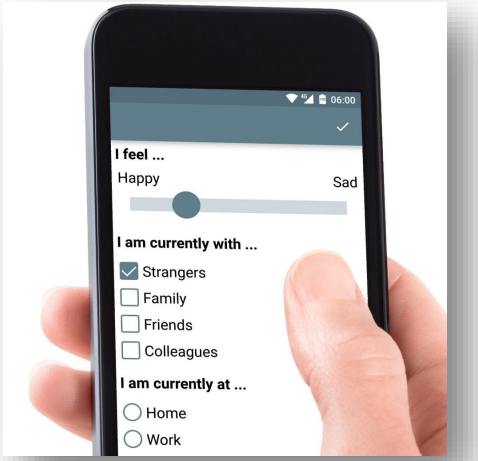
Existing in-the-moment surveys

In-the-moment surveys are used nowadays (and were used in the past), but only in very specific environments (proprietary databases, no control on the sample., one-shot…)

Satisfaction surveys in public transportation.



Online satisfaction surveys.



Experience Sampling Method

Coincidental surveys: “are you listening to the radio?” instead of “did you listen to radio last week?”

New type of surveys: ppt-in online panel + passive + in-the-moment

To overcome existing limitations of conventional surveys, I propose **a new type of in-the-moment surveys**.

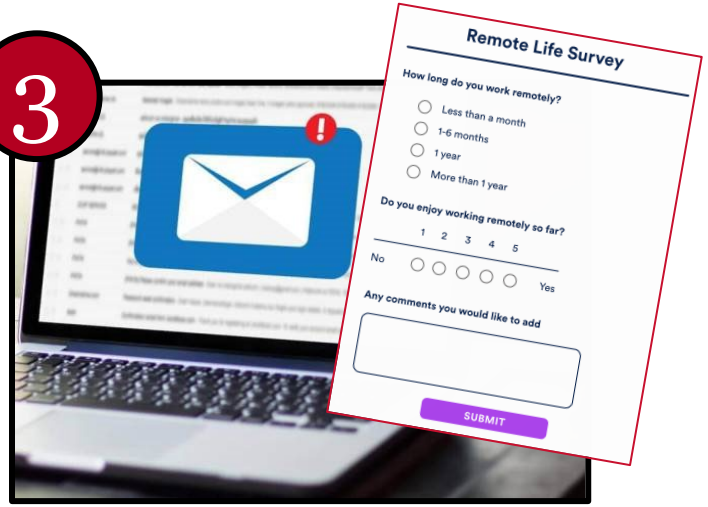
1

Opt-in online panels
Communities of people that voluntarily participate in research activities in exchange of reward.

2

Passive measurement
- Metered data -> online events
- Geolocation data -> offline events

3



In-the-moment survey
When an event of interest is detected (e.g., visiting a political party Facebook page) a survey is sent.

Potential use cases

Examples of potential uses of these new in-the-moment surveys triggered by metered data:

FAKE NEWS

Nyhan and Reifler (2018)[6] used **meter data to research consumption of fake news**: do Trump's supporters read more fake news? Surveys used only to profile participants.

In-the-moment surveys answer:

- "Do you give credibility to this news?"
- "Read this fact-checking information, do you still give credibility to…?

TRAVEL RESEARCH

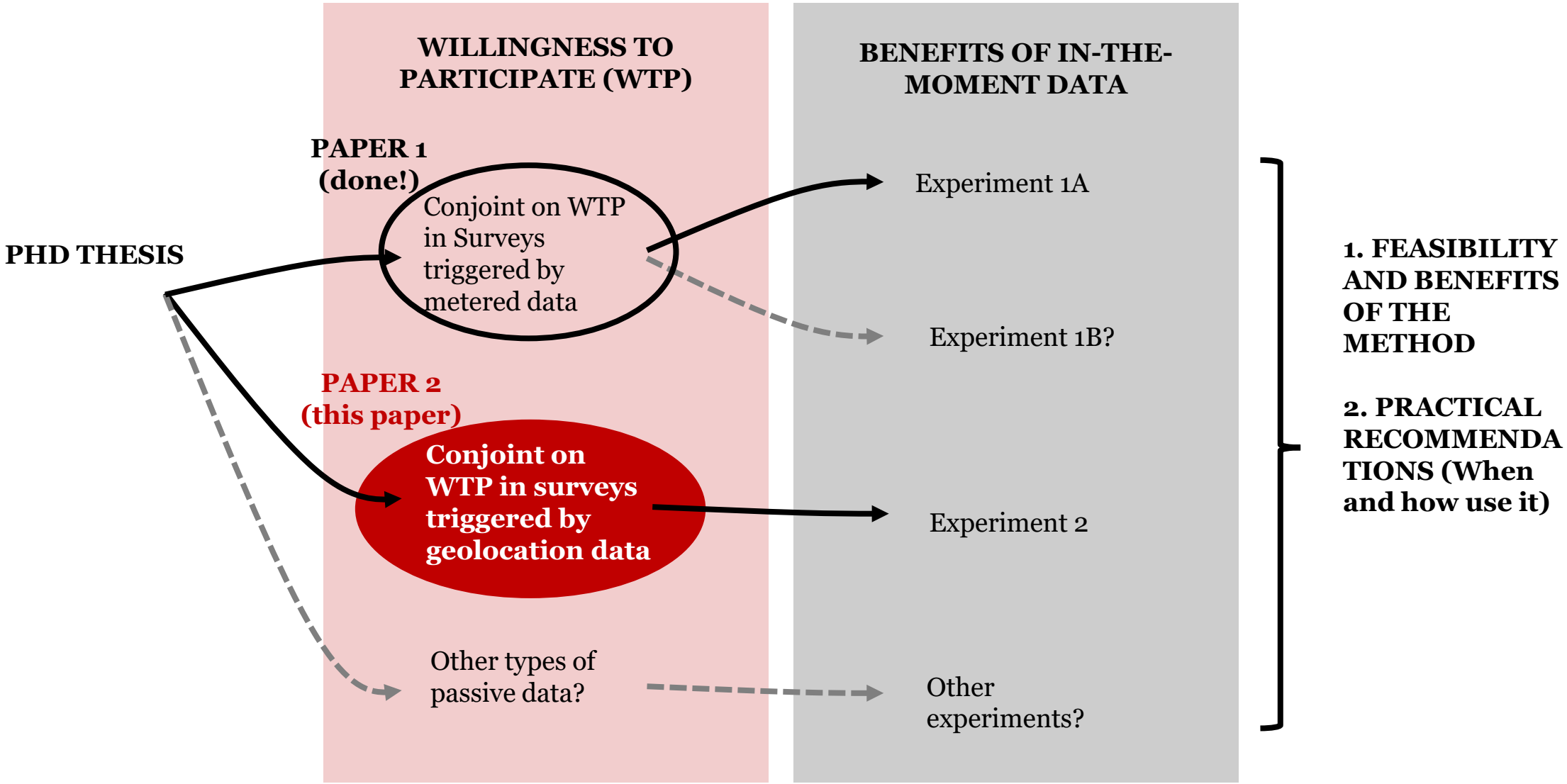
Detecting when someone has travelled using geolocation data + in-the-moment survey to ask:

- The purpose of the travel.
- Satisfaction with the mode of transport.
- Confirm whether a particular location was visited.

This paper

Willingness to participate in geolocation-based research

Research plan



LITERATURE ABOUT GEOLOCATION

- Willingness to share geolocation data: **20%-40%.**
- Studied factors:
  - Offline samples vs. online panels
  - One-time capture vs. continuous sharing.
  - Willingness to participate vs. actual participation
  - Others: country, specific online panel…

HOWEVER

- Little research about the **effect of the conditions offered to individuals to participate** (incentive, duration of the project).

LITERATURE ABOUT IN-THE-MOMENT SURVEYS

- 1st paper of this PhD: "willingness to participate in in-the-moment surveys triggered by online behaviors of metered panelists".
- Four survey attributes studied:
  1. Survey length
  2. Invitation lifetime (maximum time allowed to participate)
  3. Incentivization level (compared to a conventional survey)
  4. Triggering activity (that causes to be invited to take a survey)

MAIN FINDINGS

- High willingness to participate (**68.5%** to **94.7%).**
- Preference for longer surveys and longer times to participate.
- The tracked activity that triggers the survey plays a minor role.
- Survey length + incentive level = **75.9% of the importance**.
- Few differences among panelists.

RESEARCH QUESTIONS

RQ1 – What are the levels of **willingness to participate** in geolocation-based research:

(a) share geolocation data

(b) in-the-moment surveys triggered by geolocation data.

RQ2 – How the **attributes** of geolocation-based research influence the willingness to participate in such surveys?

RQ3 – Are there **significant differences** among panelists?

RQ4 –Main reasons for deciding whether or not to participate stated by the panelists?

ABOUT THE ATTRIBUTES

We study the effect of 6 attributes, 2-6 levels per attribute.

Research activity:
Sharing geolocation
vs.
In-the-moment surveys triggered by geolocation

Project duration:
1 week
1 month
3 month
6 month
1 year
Indefinite

Invitation lifetime*:
15 min
30 min
1 h
2 h
3 h
6 h
12 h

Geolocation incentive:
1 point/week
2 points/week
3 points/week
4 points/week
6 points/week
8 points/week

Survey incentive level:
X 1 (normal)
X 1.5
X 2
X 3
x 4

Length of the interview:
1 min
5 min
10 min
15 min
20 min

* =*maximum time to participate*

METHODS

Choice Based Conjoint analysis:

- A method to assess the influence of each attribute by the analysis of choices.
- 10 questions per participant: 2 proposals + “I would not participate”.
- Orthogonal design (minimum correlation between attribute-levels)
- Multinomial model + Bayesian analysis using simulation (MCMC*).
- “Utilities” (coefficients) used to estimate **importance** of attributes and **willingness to participate** in each scenario, for each participant.

** =Markov Chain Monte Carlo*

DATA

- Data collection: 21st of February – 7th of March 2022.
- Netquest opt-in online panel in Spain.
- **1,016 valid surveys** (2,306 invited, 1,847 started the survey, 461 discarded due to quotas and filters)
- Survey length: mean = 8.8 min.
- Quotas on age(3)+gender(2) and education(3), representative of the Spanish online population.
- **27% of the participants have installed a meter** (already sharing online behaviors).

Preliminary results
(*work in progress)*

References

1. Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., Tourangeau, R., (2009). Survey Methodology, 2nd Edition, Wiley series in survey methodology. Wiley
2. Tourangeau, R. (1999). Remembering What Happened: Memory Errors and Survey Reports. Chapter from the book "The Science of Self-report". Psychology Press.
3. Gray, P. G. (1955) The Memory Factor in Social Surveys, Journal of the American Statistical Association, 50:270, 344-363, DOI: 10.1080/01621459.1955.10501269
4. Walker, W. Richard; Skowronski, John J. (November 2009). The Fading affect bias: But what the hell is it for? (PDF). Applied Cognitive Psychology. 23 (8): 1122–1136. doi:10.1002/acp.1614.
5. Kanheman, D. (2011). Thinking, Fast and Slow. Farrar, Straus and Giroux.
6. Guess, A. M., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. Nature human behaviour, 4(5), 472–480. https://doi.org/10.1038/s41562-020-0833-x
9. Sun, B., Morwitz, V.G. (2010). Stated intentions and purchase behavior: A unified model, International Journal of Research in Marketing, Volume 27, Issue 4, Pages 356-366, ISSN 0167-8116, https://doi.org/10.1016/j.ijresmar.2010.06.001
10. Klein, S. B., Babey, S. H., & Sherman, J. W. (1997). The functional independence of trait and behavioral self-knowledge: Methodological considerations and new empirical findings. Social Cognition, 15, 183–203.