GESIS Leibniz Institute for the Social Sciences

Who do we not reach with online-only? Insights from a mixed-mode panel

Leibniz Association

Lena Rembser

February 26, 2026

CIPHER Conference 2026

Online data collection is fancy – but what about offliners?

- Offliners (= people who are not able or willing to participate in a survey via web mode)
- Some statistics
  - 5% of the German general population aged 18–74 do not access the internet (Destatis 2025)
    - No official statistics for those aged 75+
  - GESIS Panel.pop Population Sample (probabilistic mixed-mode panel with push-to-web design) recruitment 2023: ~18% decided that they do not want to participate online but via paper mode
- Why does that share of offliners in surveys matters?
  - Including offliners results in more accurate estimates for some socio-demographic characteristics[1] and substantive characteristics[2]

[1] Blom et al., 2017; Bosnjak et al., 2018; Cornesse & Bosnjak, 2018; Cornesse & Schaurer, 2021; Leenheer & Scherpenzeel, 2013; Revilla et al., 2016; own research
[2] Leenheer & Scherpenzeel, 2013; Pforr & Dannwolf, 2017; Schaurer & Weiß, 2020; own research

Online data collection is fancy – but what about offliners?

- Offliners (= people who are not able or willing to participate in a survey via web mode)
- Some statistics
  - 5% of the German general population aged 18–74 do not access the internet (Destatis 2025)
    - No official statistics for those aged 75+
  - GESIS Panel.pop Population Sample (probabilistic mixed-mode panel with push-to-web design) recruitment 2023: ~18% decided that they do not want to participate online but via paper mode
- Why does that share of offliners in surveys matters?
  - Including offliners results in more accurate estimates for some socio-demographic characteristics[1] and substantive characteristics[2]

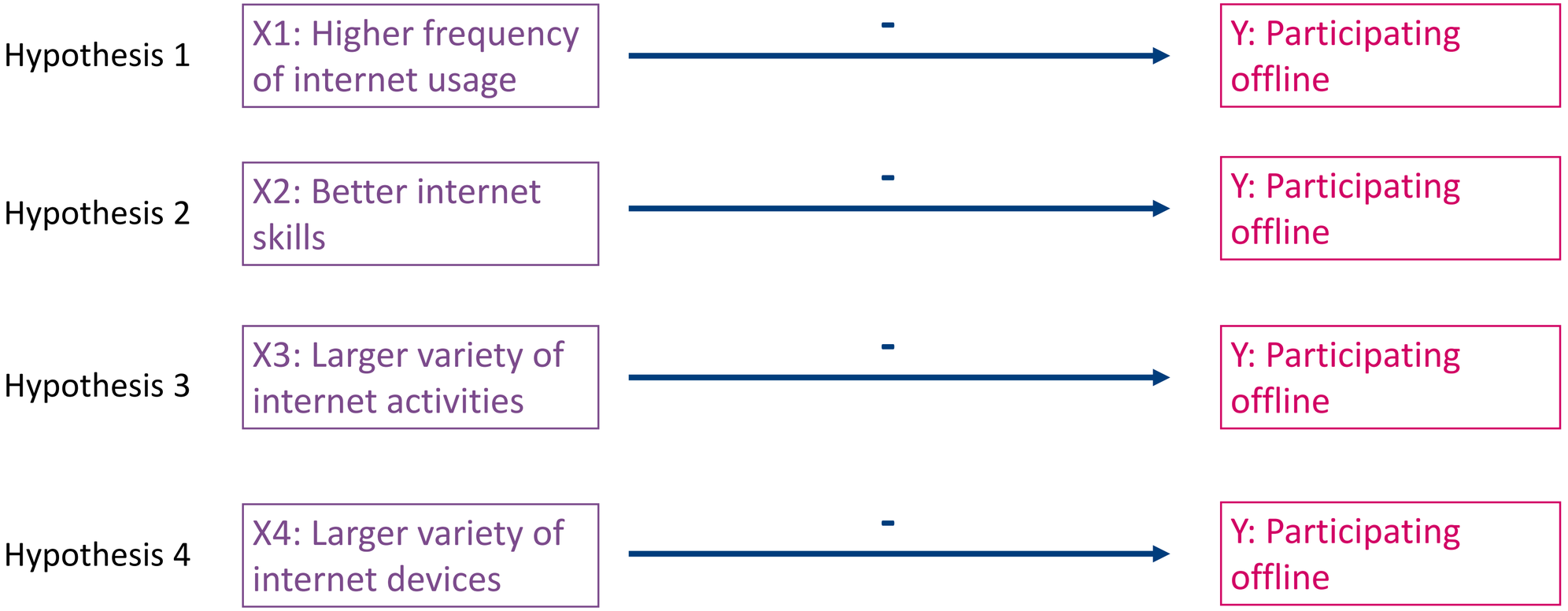
Research question

Apart from the simple effect of having or not having internet access:
How do internet-related characteristics influence mode choice for people with internet access?

Hypotheses

Research question

How do internet-related characteristics influence mode choice for people with internet access?



Data

- GESIS Panel.pop Population Sample
- Population register-based samples
  - Two survey modes (web-based and complementary paper mode)
    - Internet users: online as default, paper mode only as alternative (push-to-web)
  - Analysis Sample: The most recent waves containing the for the analyses necessary variables after recruitment for the cohorts recruited in 2021 and 2023 (Waves jc & jd (2022) and lc & ld (2024))
  - N≈1,400

Method

- Binary logistic regressions with interpretation of Average Marginal Effects (AMEs)
- Standard-demographic control variables (education, age, gender)
- Variable thresholding for the variables **X1: frequency of internet usage** and **X2: internet skills** to identify potential thresholds

Results Hypothesis 1

X1: Higher frequency of internet usage → - → Y: Participating offline

Binary logistic regression of **X1: frequency of internet usage** on mode choice (0 = web mode, 1 = paper mode) for respondents with internet access
Coefficients: Average Marginal Effects (AMEs)

	Model 1.1 AME	Model 1.2a AME	Model 1.2b AME	Model 1.2c AME
frequency of internet usage *(metric; 6 answer categories)*	-3.3***			
Dummy-Variables frequency of internet usage				
[1] more than once a month *Ref.: Up to once a month*		-17.9*		
[1] more than once a week *Ref.: up to once a week*			-23.8***	
[1] daily or more *Ref.: up to several times a week*				-11.0***
Control Variables (education, age, gender)	Included	Included	Included	Included

* significant with p < 0.1, ** significant with p < 0.05, *** significant with p < 0.01

Interpretations:

Model 1.1: With every increasing frequency level of internet usage, respondents have on average a 3.3 percentage points lower probability of participating offline.

Models 1.2: The group of respondents using the internet more often has on average a 17.9 / 23.8 / 11.0 percentage points lower probability of participating offline.

Results Hypothesis 2

X2: Better internet skills → - → Y: Participating offline

Binary logistic regression of **X2: internet skills** on mode choice
(0 = web mode, 1 = paper mode) for respondents with internet access
Coefficients: Average Marginal Effects (AMEs)

	Model 2.1 AME	Model 2.2a AME	Model 2.2b AME	Model 2.2c AME
Internet skills *(metric; 5 answer categories)*	-7.3***			
Dummy-Variables internet skills				
[1] moderately or better *Ref.: bad + very bad*		-17.5***		
[1] good or very good *Ref.: moderately or worse*			-13.5***	
[1] very good *Ref.: good or worse*				-8.1***
Control Variables (education, age, gender)	Included	Included	Included	Included

* significant with p < 0.1, ** significant with p < 0.05, *** significant with p < 0.01

Interpretations:

Model 2.1: With every increasing internet skill level, respondents have on average a 7.3 percentage points lower probability of participating offline.

Models 2.2: The group of respondents with better internet skills has on average a 17.5 / 13.5 / 8.1 percentage points lower probability of participating offline.

Results Hypothesis 3

X3: Larger variety of internet activities → - → Y: Participating offline

Binary logistic regression of **X3: variety of internet activities** on mode choice (0 = web mode, 1 = paper mode) for respondents with internet access

Coefficients: Average Marginal Effects (AMEs)

	Model 3 AME
Variety of internet activities (metric) *Index of 11 internet activities*	-3.3***
Control Variables (education, age, gender)	Included

* significant with p < 0.1, ** significant with p < 0.05, *** significant with p < 0.01

Interpretation:

Model 3: With every additional internet activity, respondents have on average a 3.3 percentage points lower probability of participating offline.

Results Hypothesis 4

X4: Larger variety of devices → (-) Y: Participating offline

Binary logistic regression of **X4: variety of devices** on mode choice (0 = web mode, 1 = paper mode) for respondents with internet access

Coefficients: Average Marginal Effects (AMEs)

	Model 4 AME
Variety of internet devices *Index of 4 devices (PC, laptop, smartphone, tablet)*	-5.9***
Control Variables (education, age, gender)	Included

* significant with p < 0.1, ** significant with p < 0.05, *** significant with p < 0.01

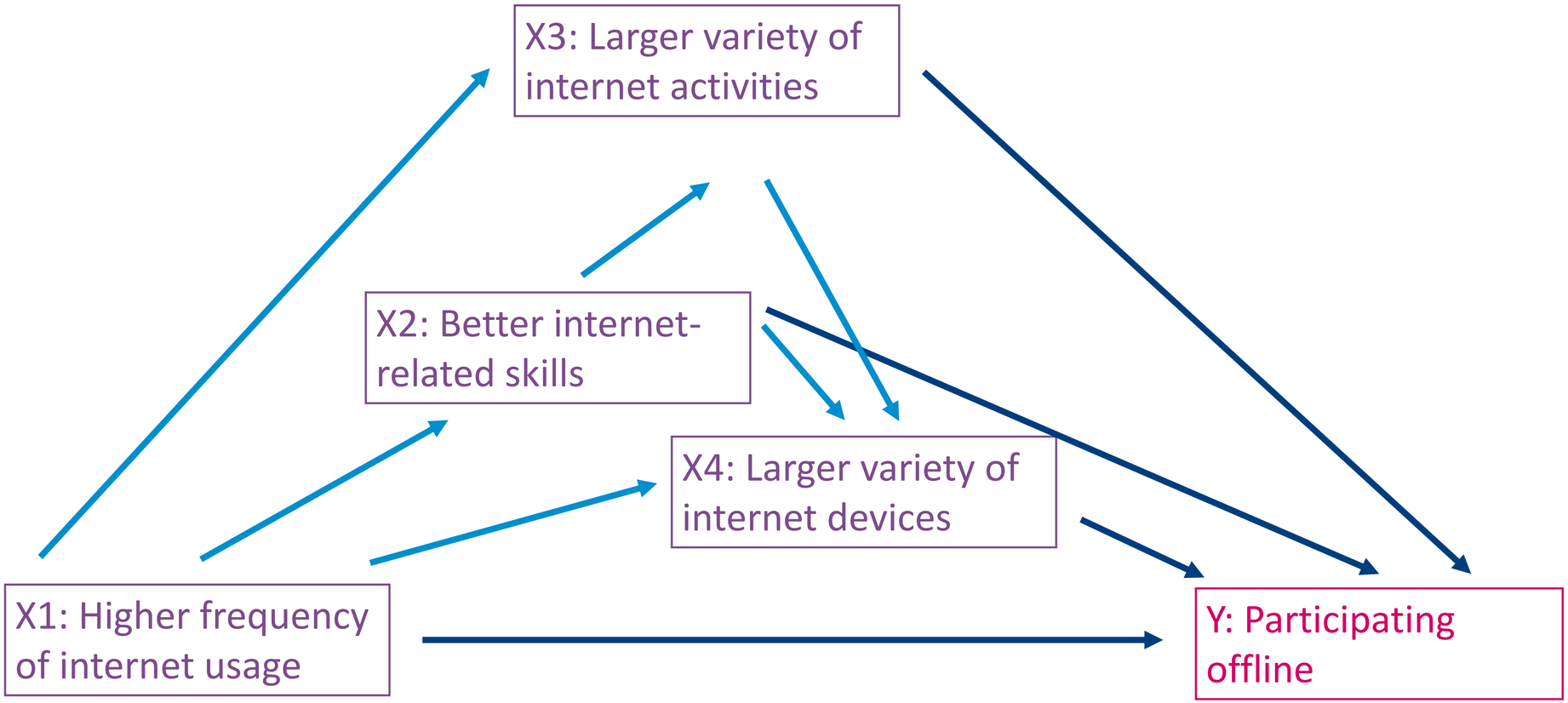
Interpretation:

Model 4: With every additional internet device, respondents have on average a 5.9 percentage points lower probability of participating offline.

Can the effects of the previous slides be interpreted as causal?

Or are they overestimated or underestimated by undercontrol / overcontrol biases stemming from the internet-related characteristics among each other?

Presumably most plausible dominant causal directions I argue for, visualized as DAG



Results

Binary logistic regression of internet-related characteristics on mode choice (0 = web mode, 1 = paper mode) for respondents with internet access

Coefficients: Average Marginal Effects (AMEs)

	Model 5.X1 AME	Model 5.X2 AME	Model 5.X3 AME	Model 5.X4 AME
frequency of internet usage (metric)	-3.3***	-1.4**	0.3	-0.4
Internet skills (metric)		-6.6***	-4.6***	-4.4***
Variety of internet activities (metric) *Index of 11 internet activities*			-2.6***	-2.3***
Variety of internet devices *Index of 4 devices (PC, laptop, smartphone, tablet)*				-2.0*
Control Variables (education, age, gender)	Included	Included	Included	Included

Presumably interpretable as causal effect

* significant with p < 0.1, ** significant with p < 0.05, *** significant with p < 0.01

Hypotheses

Research question

How do internet-related characteristics influence mode choice for people with internet access?

Hypothesis 1	X1: Higher frequency of internet usage	- →	Y: Participating offline	Evidence
Hypothesis 2	X2: Better internet skills	- →	Y: Participating offline	Evidence
Hypothesis 3	X3: Larger variety of internet activities	- →	Y: Participating offline	Evidence
Hypothesis 4	X4: Larger variety of internet devices	- →	Y: Participating offline	Not enough evidence based on p-value

Conclusion

Research question

Apart from the simple effect of having or not having internet access: How do internet-related characteristics influence mode choice for people with internet access?

- The more frequent respondents use the internet, the higher their variety of internet activities and the higher their skills, the less often they participate offline
  - It might also be that variety of internet devices influences the probability of participating offline negatively
  - The effect of frequency of internet usage is the most substantial (~25 percentage points)

- Internet-related characteristics do not determine mode choice -> further research needed
  - Targeted design with only considering internet-related characteristics would not capture the complexity of mode choice
  - Mode choices are more complex than being only influenced by internet related-characteristics

Outlook

- Who are “typical” offliners? What are the profiles of offliners?
- How do attitudes (privacy concerns, affinity for technology, risk perception) interplay in the influence of the examined internet-related characteristics on mode choice?

Thank you!

Leibniz
Association

References

- Blom, A. G., Herzing, J. M. E., Cornesse, C., Sakshaug, J. W., Krieger, U., & Bossert, D. (2017). Does the Recruitment of Offline Households Increase the Sample Representativeness of Probability-Based Online Panels? Evidence From the German Internet Panel. *Social Science Computer Review, 35(4)*, 498–520. https://doi.org/10.1177/0894439316651584
- Bosnjak, M., Dannwolf, T., Enderle, T., Schaurer, I., Struminskaya, B., Tanner, A., & Weyandt, K. W. (2018). Establishing an Open Probability-Based Mixed-Mode Panel of the General Population in Germany: The GESIS Panel. *Social Science Computer Review, 36*(1), 103–115. https://doi.org/10.1177/0894439317697949
- Cornesse, C., & Bosnjak, M. (2018). Is there an association between survey characteristics and representativeness? A meta-analysis. *Survey Research Methods, 12*(1), Article 1. https://doi.org/10.18148/srm/2018.v12i1.7205
- Cornesse, C., & Schaurer, I. (2021). The Long-Term Impact of Different Offline Population Inclusion Strategies in Probability-Based Online Panels: Evidence From the German Internet Panel and the GESIS Panel. *Social Science Computer Review, 39*(4), 687–704. https://doi.org/10.1177/0894439320984131
- Destatis. (2025). Informations- und Kommunikationstechnologien privater Haushalte (Mikrozensus-Unterstichprobe—IKT). Endergebnis 2025 (EVAS-Nummer 12331). Destatis. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/IT-Nutzung/Publikationen/Downloads-IT-Nutzung/statistischer-bericht-ikt-privater-haushalte-2150400257005.xlsx?__blob=publicationFile&v=3
- GESIS. (2025). GESIS Panel.pop Population Sample – Standard Edition (Version (ZA5665; Version 60.0.0)). GESIS (Cologne). https://doi.org/https://doi.org/10.4232/1.14587
- Leenheer, J., & Scherpenzeel, A. (2013). Does It Pay Off to Include Non-Internet Households in an Internet Panel? *International Journal of Internet Science, 8*, 17–29.
- Pforr, K., & Dannwolf, T. (2017). What do we Lose with Online-Only Surveys? Estimating the Bias in Selected Political Variables Due to Online Mode Restriction. *Statistics, Politics and Policy, 8.* https://doi.org/10.1515/spp-2016-0004
- Revilla, M., Cornilleau, A., Cousteaux, A.-S., Legleye, S., & de Pedraza, P. (2016). What Is the Gain in a Probability-Based Online Panel of Providing Internet Access to Sampling Units Who Previously Had No Access? *Social Science Computer Review, 34*(4), 479–496. https://doi.org/10.1177/0894439315590206
- Schaurer, I., & Weiß, B. (2020). Investigating selection bias of online surveys on coronavirus-related behavioral outcomes. *Survey Research Methods, 14*(2), Article 2. https://doi.org/10.18148/srm/2020.v14i2.7751