

Proxy Respondents in Longitudinal Ageing Studies

Can they help to reduce bias in ageing research?

Daniele Zaccaria & Michael Bergmann

ESRA 2025, *18 July 2025*

Motivation

- Health surveys on older adults are at risk of underestimating poor health conditions as...
- Healthier individuals are more likely to participate and to remain in panel studies (Rothman et al., 2008; Chatfield et al., 2005)
- Excluding frail or institutionalised older adults can lead to biased results

The Role of Proxies

- Proxy respondents are used when older adults are unable to respond themselves
- Common reasons: cognitive impairment, physical frailty (Elliott et al., 2008)
- Proxies help reduce attrition and maintain the representativeness of longitudinal data in an ethically acceptable way (Wolf et al., 2019; Bourke, 2020)

Key Concerns

- Risk of measurement error and lower data quality in general (e.g. Magaziner et al., 1997)
- Objective indicators (e.g., physical functions) are better captured by proxies than subjective (e.g., emotions, pain) states (Bowling & Windsor, 2008)
- The relationship between the proxy and respondent may influence data accuracy (Neumann et al., 2000)
- Changing proxies over waves may affect the reliability of longitudinal analyses (Lee et al., 2020)

Research Questions

1. Could using proxy respondents help reduce the underestimation of health problems?
2. How does the use of proxy respondents affect data quality (e.g. item non-response)?
3. What is the impact of proxy respondents on substantive analyses?

SHARE in a nutshell

Survey of Health, Ageing and Retirement in Europe

- **Biennial** representative cross-national household **panel survey** of the **50+ population** in **27 European countries & Israel**
- **Launched** in **2004**, Wave 10 fieldwork started in 2024
- Around 616.000 **face-to-face interviews (CAPI mode)** with over 160.000 respondents
- ~8000 interviews with respondents in nursing/care homes
- **Ex-ante harmonisation** – centralised procurement, questionnaire development, sampling, data collection instrument, and database management
- **Broad range of topics** (e.g., health, economic and living conditions, social relationships) and measures
- **Free data access for researchers:** <http://www.share-eric.eu>



Data

- Wave 9 (2021/2022) with 28 countries
- Interviewers document assistance by a proxy after most modules (exceptions: questions on cognition, physical test)
- Focus on physical health questionnaire module

	Interview in a private household	Interview in a nursing/care home	Total
No assistance in answering PH questions	66,429	428	66,857
Assistance in answering PH questions	2,278	207	2,485
Total	68,707	635	69,342

Data: SHARE Wave 9, Release 9.0.0. Significance level: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

Preliminary Results I

Characteristics of Respondents with and without Assistance in Answering Questions

	All Respondents (n=69,342)	Proxy Respondents Private Households (n=2278)
Age	69.3	7.9***
Sex: female	57.4	-8.8***
Education: low	30.5	25.7***

Data: SHARE Wave 9, Release 9.0.0. Significance level: ***: $p < .001$, **: $p < .01$, *: $p < .05$.
Entries for proxy and regular respondents are differences in %-points/years compared to all respondents.

Preliminary Results I

Characteristics of Respondents with and without Assistance in Answering Questions

	All Respondents (n=69,342)	Proxy Respondents Private Households (n=2278)	Proxy Respondents Nursing/Care Homes (n=207)
Age	69.3	7.9***	14.8***
Sex: female	57.4	-8.8***	-14.1***
Education: low	30.5	25.7***	30.0***

Data: SHARE Wave 9, Release 9.0.0. Significance level: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

Entries for proxy and regular respondents are differences in %-points/years compared to all respondents.

Preliminary Results I

Characteristics of Respondents with and without Assistance in Answering Questions

	All Respondents (n=69,342)	Proxy Respondents Private Households (n=2278)	Proxy Respondents Nursing/Care Homes (n=207)	Regular Respondents (n=66,857)
Age	69.3	7.9***	14.8***	-0.3***
Sex: female	57.4	-8.8***	-14.1***	0.3
Education: low	30.5	25.7***	30.0***	-1.0***

Data: SHARE Wave 9, Release 9.0.0. Significance level: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

Entries for proxy and regular respondents are differences in %-points/years compared to all respondents.

Preliminary Results I

Characteristics of Respondents with and without Assistance in Answering Questions

	All Respondents (n=69,342)	Proxy Respondents Private Households (n=2278)
Poor health	10.5	36.4***
GALI	48.8	33.7***
IADL	18.5	50.1***
Pain	46.1	14.8***
Chronic disease	53.6	18.3***
Mobility limitations	50.3	27.5***
Taking drugs	79.0	11.1***
Long-term illness	54.8	25.7***
Frailty	41.5	25.2***

Data: SHARE Wave 9, Release 9.0.0. Significance level: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

Entries for proxy and regular respondents are differences in %-points compared to all respondents.

Preliminary Results I

Characteristics of Respondents with and without Assistance in Answering Questions

	All Respondents (n=69,342)	Proxy Respondents Private Households (n=2278)	Proxy Respondents Nursing/Care Homes (n=207)
Poor health	10.5	36.4***	39.0***
GALI	48.8	33.7***	47.2***
IADL	18.5	50.1***	78.4***
Pain	46.1	14.8***	7.5*
Chronic disease	53.6	18.3***	24.2***
Mobility limitations	50.3	27.5***	42.4***
Taking drugs	79.0	11.1***	16.6***
Long-term illness	54.8	25.7***	36.5***
Frailty	41.5	25.2***	31.6***

Data: SHARE Wave 9, Release 9.0.0. Significance level: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

Entries for proxy and regular respondents are differences in %-points compared to all respondents.

Preliminary Results I

Characteristics of Respondents with and without Assistance in Answering Questions

	All Respondents (n=69,342)	Proxy Respondents Private Households (n=2278)	Proxy Respondents Nursing/Care Homes (n=207)	Regular Respondents (n=66,857)
Poor health	10.5	36.4***	39.0***	-1.3***
GALI	48.8	33.7***	47.2***	-1.3***
IADL	18.5	50.1***	78.4***	-1.9***
Pain	46.1	14.8***	7.5*	-0.5
Chronic disease	53.6	18.3***	24.2***	-0.7*
Mobility limitations	50.3	27.5***	42.4***	-1.0***
Taking drugs	79.0	11.1***	16.6***	-0.4
Long-term illness	54.8	25.7***	36.5***	-1.0**
Frailty	41.5	25.2***	31.6***	-0.9**

Data: SHARE Wave 9, Release 9.0.0. Significance level: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

Entries for proxy and regular respondents are differences in %-points compared to all respondents.

Preliminary Results II

Item Non-response of Respondents with and without Assistance in Answering Questions

	All Respondents (n=69,342)	Proxy Respondents (n=2485)
Poor health	92	60***
GALI	96	52***
IADL	118	72***
Pain	130	92***
Chronic disease	103	40***
Mobility limitations	114	72***
Taking drugs	180	78***
Long-term illness	100	56***
Frailty	138	105***

Data: SHARE Wave 9, Release 9.0.0. Significance level: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

Preliminary Results II

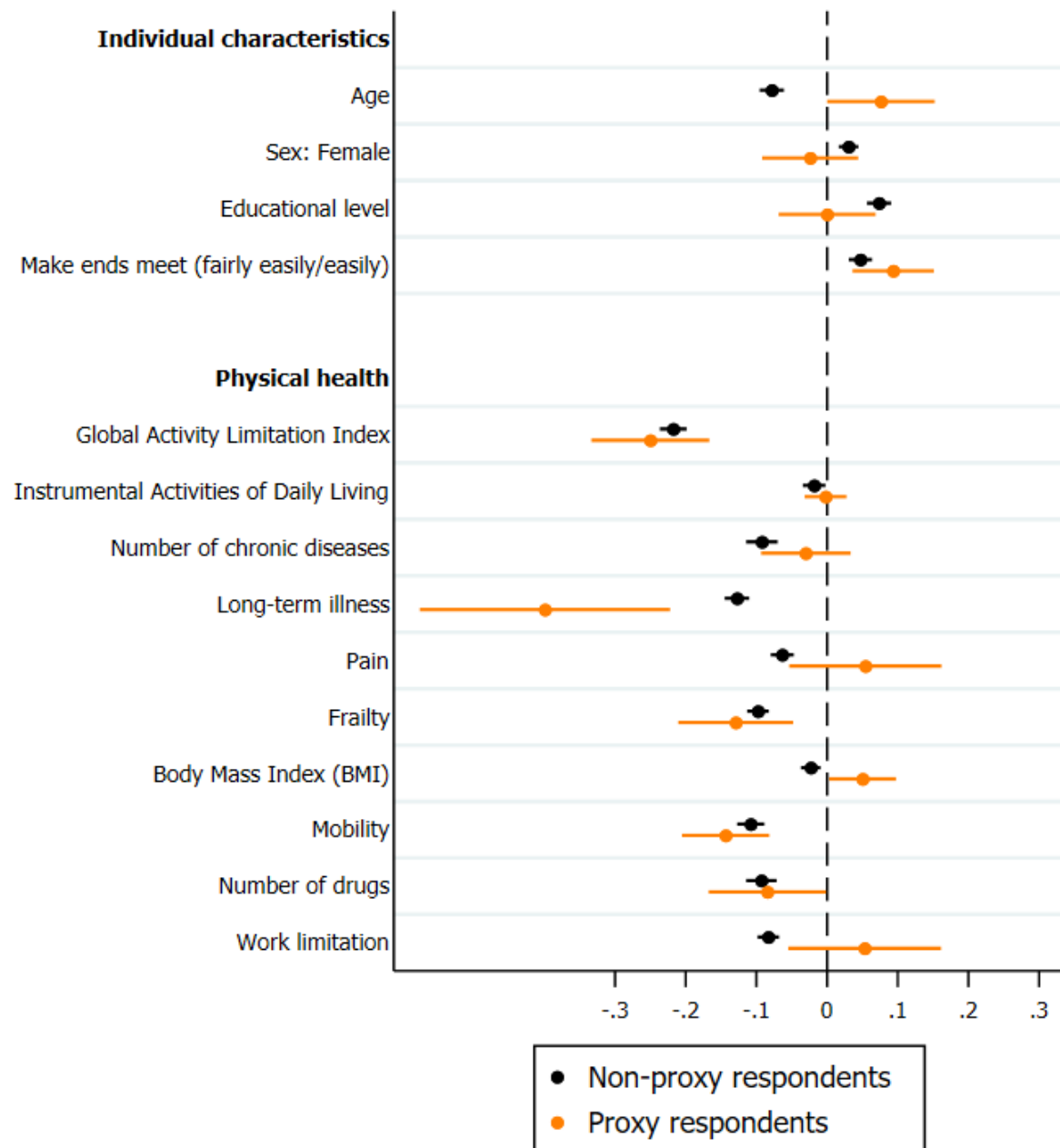
Item Non-response of Respondents with and without Assistance in Answering Questions

	All Respondents (n=69,342)	Proxy Respondents (n=2485)	Regular Respondents (n=66,857)
Poor health	92	60***	32***
GALI	96	52***	44***
IADL	118	72***	46***
Pain	130	92***	38***
Chronic disease	103	40***	63**
Mobility limitations	114	72***	42***
Taking drugs	180	78***	102***
Long-term illness	100	56***	44***
Frailty	138	105***	33***

Data: SHARE Wave 9, Release 9.0.0. Significance level: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

Preliminary Results III

Determinants of Self-rated Health by Proxy Respondent Status



Data: SHARE Wave 9, release version: 9.0.0 (n=64,811, weighted) with 95%-confidence intervals. R^2 : .49.
Note: Shown are average marginal effects (AMEs). Country controls are included but not shown.

Conclusion

1. Could using proxy respondents help reduce the underestimation of health problems?
2. How does the use of proxy respondents affect data quality (e.g. item non-response)?
3. What is the impact of proxy respondents on substantive analyses?

Conclusion

1. Could using proxy respondents help reduce the underestimation of health problems? → Yes (at least to some degree)
2. How does the use of proxy respondents affect data quality (e.g. item non-response)?
3. What is the impact of proxy respondents on substantive analyses?

Conclusion

1. Could using proxy respondents help reduce the underestimation of health problems? → Yes (at least to some degree)
2. How does the use of proxy respondents affect data quality (e.g. item non-response)? → Proxy respondents are responsible for a large part of item non-response
3. What is the impact of proxy respondents on substantive analyses?

Conclusion

1. Could using proxy respondents help reduce the underestimation of health problems? → Yes (at least to some degree)
2. How does the use of proxy respondents affect data quality (e.g. item non-response)? → Proxy respondents are responsible for a large part of item non-response
3. What is the impact of proxy respondents on substantive analyses? → Differences in associations are mostly insignificant; however, including proxies can contribute to more meaningful analyses

Limitations & Next Steps

Limitations:

- No information on actual responses of respondents who need assistance
- Not entirely clear which questions have been answered by a proxy and which not

Next Steps:

- Analyze additional indicators of data quality (e.g. extreme values)
- Look closer into the difference between subjective and objective questions
- Proxy continuity effect: Assess consistency and reliability of proxy data over time

Thanks!

Have you got any questions?

mbergmann@share-berlin.eu

Bluesky



LinkedIn



www.share-eric.eu

www.share-berlin.eu

