Adam Rybak

Possibilities of using declarative and behavioral data in predicting respondents survey mode preference in Poland

CSDI workshop 19.03.2019

Institute of Sociology

Overview

- **Previous research:** who can be pushed to mail mode? as based on the mixed-mode experiment results

- Survey mode preference – introduction

- Description of 2 datasets

- Method – multinomial logistic regression

- **Main question:** Whether declarations of mode preference are a reliable source of information in case of creating mixed-mode design based on the model?
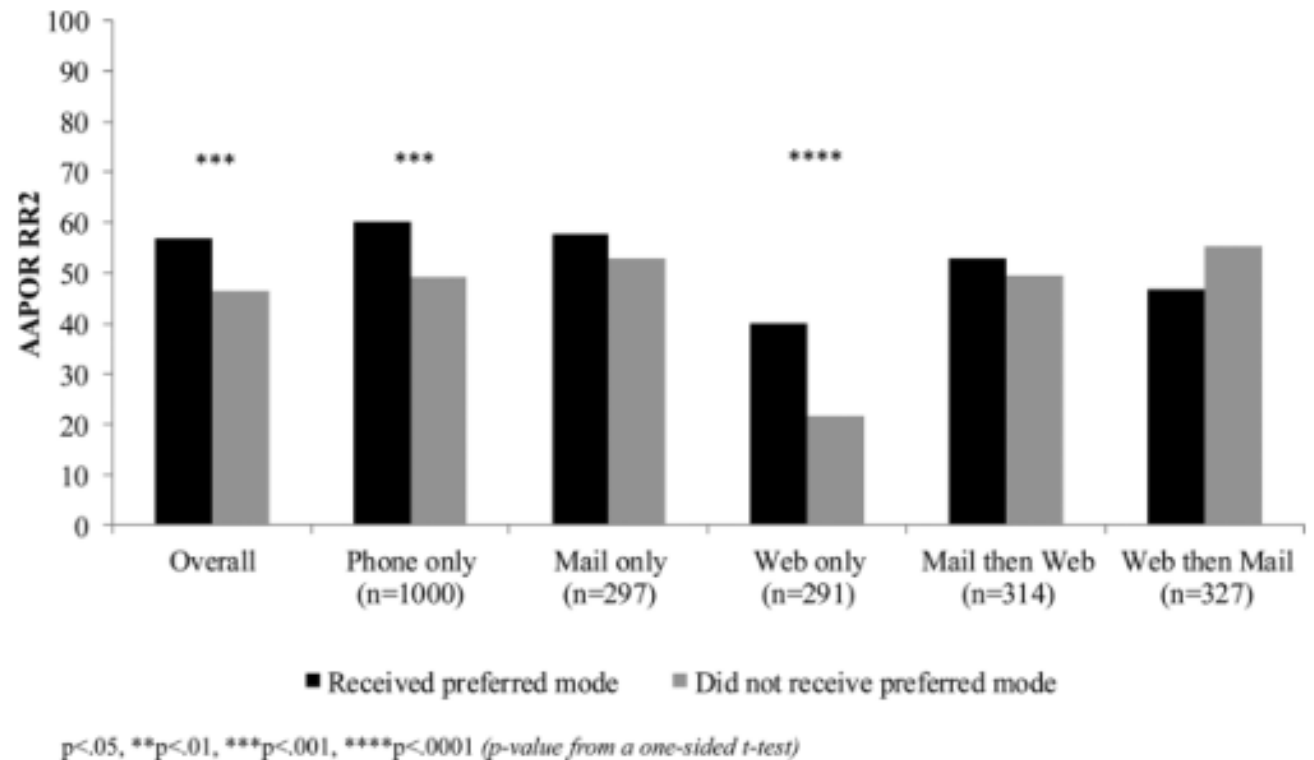
Survey Mode Preference

- Discovered by Groves and Kahn in 1979
- Relatively constant preference toward one mode at the expense of other
- Conscious?
- Does it even exist? – Relationship between mode of question and answer
- Effect on nonresponse (error), measurement error (satisficing?)

Effect of preference on response rate
Kristen Olson, Jolene Smyth, Heather Wood, 2012

	Model 1 ($n = 1,329$)	
	B	SE
Intercept	0.041	(0.080)
Mode preference		
Preferred mode	0.323***	(0.090)
Not preferred mode	–	
Mode treatment		
Mail only	0.038	(0.129)
Web only	−1.187****	(0.148)
Mail web	−0.093	(0.127)
Web mail	0.013	(0.128)
Phone	–	

OR= 1,38

p<.05, **p<.01, ***p<.001, ****p<.0001 *(p-value from a one-sided t-test)*

Satisficing
Jolene Smyth, Kristen Olson, Alian Kasabian, 2014

- **Positive effect** of preferred mode on avoiding satisficing 'traps'
- **Only** Mail and Web Modes

29. Thinking about your personal finances, please indicate whether or not you have done each of the following in the past 12 months.

	Yes	No
Delayed/canceled purchasing a home	☐	☐
Delayed/canceled plans to buy a car	☐	☐
Delayed/canceled plans to make a major household purchase	☐	☐
Cut back on vacation spending	☐	☐
Cut back on eating out	☐	☐
Cut back on home internet access	☐	☐
Cut back on home landline telephone services	☐	☐

29. Thinking about your personal finances, which of the following have you done in the past 12 months? Check all that apply.

- ☐ Delayed/canceled purchasing a home
- ☐ Delayed/canceled plans to buy a car
- ☐ Delayed/canceled plans to make a major household purchase
- ☐ Cut back on vacation spending
- ☐ Cut back on eating out
- ☐ Cut back on home internet access
- ☐ Cut back on home landline telephone services

Frequency of 'Yes'

32. All things considered, during the next year, what do you think will be your biggest challenges?

32. All things considered, during the next year, what do you think will be your biggest challenges?

Length of answer No Answer

14. Please indicate how satisfied or dissatisfied you are with the availability of each of the following in your community.

	Very Dissatisfied	Dissatisfied	Neither Satisfied nor Dissatisfied	Satisfied	Very Satisfied
Parks	☐ Very Diss.	☐ Diss.	☐ Neither	☐ Sat.	☐ Very Sat.
Bicycle paths	☐ Very Diss.	☐ Diss.	☐ Neither	☐ Sat.	☐ Very Sat.
Outdoor areas to hunt, fish, or hike	☐ Very Diss.	☐ Diss.	☐ Neither	☐ Sat.	☐ Very Sat.
Sporting events	☐ Very Diss.	☐ Diss.	☐ Neither	☐ Sat.	☐ Very Sat.
Restaurants	☐ Very Diss.	☐ Diss.	☐ Neither	☐ Sat.	☐ Very Sat.
Fine arts (museums and theatres)	☐ Very Diss.	☐ Diss.	☐ Neither	☐ Sat.	☐ Very Sat.
Cell phone service	☐ Very Diss.	☐ Diss.	☐ Neither	☐ Sat.	☐ Very Sat.
Internet access	☐ Very Diss.	☐ Diss.	☐ Neither	☐ Sat.	☐ Very Sat.
Libraries	☐ Very Diss.	☐ Diss.	☐ Neither	☐ Sat.	☐ Very Sat.

14. Please indicate how satisfied or dissatisfied you are with the availability of each of the following in your community.

	Very Dissatisfied	Dissatisfied	Neither Satisfied nor Dissatisfied	Satisfied	Very Satisfied
Parks	☐	☐	☐	☐	☐
Bicycle paths	☐	☐	☐	☐	☐
Outdoor areas to hunt, fish, or hike	☐	☐	☐	☐	☐
Sporting events	☐	☐	☐	☐	☐
Restaurants	☐	☐	☐	☐	☐
Fine arts (museums and theatres)	☐	☐	☐	☐	☐
Cell phone service	☐	☐	☐	☐	☐
Internet access	☐	☐	☐	☐	☐
Libraries	☐	☐	☐	☐	☐

Straightlining

Mode preference predictors

(Smyth, Olson and Millar, 2014)

- CATI

- 2008 Nebraska Annual Social Indicators Survey

Preference	Model 4[b] (base = phone)	
	Internet	Mail
Media familiarity/access		
Female	1.14	1.48[*]
Age	0.97[***]	0.98[*]
Income (25–49 K)	1.02	1.00
Income (50–74 K)	0.94	0.77
Income (75+K)	1.38	0.71
Nonwhite	1.08	1.33
Uses internet	10.60[**]	0.65
Has cell phone	1.02	1.00
Poor computer ability	0.13[***]	1.10
Lives in country	1.89	0.86
Lives in city	1.50	0.93
External distractions		
Married	0.94	1.24
Has children	1.09	0.80
Employed	1.70	0.87
Cog./Phys. abilities		
Education (some college)	2.03[*]	0.98
Education (college degree)	2.27[**]	0.93
Poor hearing	0.42	1.13
Poor health	1.06	0.78
Normative concerns		
CESD > 8	0.98	1.20
Safety concerns		
Often fear identity theft	1.12	1.40[*]
Often fear crime	0.52	0.80
Constant	0.04[**]	1.57
Observations	1,730	1,730

Downsides of NASIS Experiment for implementation in polish context

- No face-to-face mode included
- Age – only linear relationship
- No real ‚size of settlement' variable

- Different context:
  - norms of hospitality
  - attitude to strangers
  - sincerity
  - perceived legitimacy of mailed documents
  - access to post offices
  - Internet penetration

Main ways of mode preference research

- Asking about declarations – preference must be conscious, mode of question have impact on answer

- Registering choices – in mixed-mode parallel design survey – giving choice significantly decrease response rate

- Comparing declaration and then response rate between experimental groups – panel survey

- Comparing declaration and choices – in panel survey or in two similar surveys

Main ways of mode preference research

- Asking about declarations – preference must be conscious, mode of question have impact on answer

- Registering choices – in mixed-mode parallel design survey – giving choice significantly decrease response rate

- Comparing declaration and then response rate between experimental groups – panel survey

- Comparing declaration and choices – in panel survey or in two similar surveys

Data

- European Social Survey 7 Mixed-Mode Experiment in Poland (Sztabiński, Sztabiński, 2015) – Behavioral data

- European Social Survey 8 in Poland (additional set of ‚national questions') – Declaration data

Similarities

- Frame (PESEL – personal register)
- Sampling Design
- Research Institution (ORBS IFiS PAN)
- Main questionnaire

ESS 7 Mix-Mode experiment

- P. Sztabiński 2015 In Poland
- Parallel to main ESS Survey

Sample size	800	ESS7	2715
Response Rate	55,9%	ESS7	65,8%

- Same questionnaire as in ESS7
- Very similar fieldwork period as in ESS7

ESS 7 Mix-Mode experiment

- Design:

1. Mailing of both postal questionnaire and link to web questionnaire – choice

2. Postal Reminder

3. Face to face Interview

4. Second postal questionnaire with persuasion letter – no web choice

- Mixed-mode design was not parallel (not strictly sequential either)

ESS 8 in Poland

- Only face-to-face mode

Sample size	2675
Response rate	69.63%

- Additional question in Poland (**P1**)

If in the future you would be offered participation in a study similar to this one, what form of participation would suit you the most

o In direct personal conversation, like this one

o In direct conversation, but by phone

o By filling out the questionnaire sent by post

o By completing the questionnaire on the website

o Or whether it would be indifferent to you?

Completed interviews

	Behavioral data	
Choice	Frequency	Percentage
Face2face	231	55,2
Mail	164	39,1
Web	24	5,7
Total	419	100

Completed interviews

Behavioral data		
Choice	Frequency	Percentage
Face2face	231	55,2
Mail	164	39,1
Web	24	5,7
Total	419	100

Declaration Data		
Declaration	Frequency	Percentage
Face2face	1213	71,6
Mail	96	5,7
Web	190	11,2
Phone	27	1,6
Not interested	74	4,4
No preference	81	4,8
Refusal	2	0,1
No answer	11	0,7
Total	1694	100

Completed interviews

Behavioral data		
Choice	**Frequency**	**Percentage**
Face2face	231	55,2
Mail	164	39,1
Web	24	5,7
Total	419	100

Declaration Data		
Declaration	**Frequency**	**Percentage**
Face2face	1213	71,6
Mail	96	5,7
Web	190	11,2
Phone	27	1,6
Not interested	74	4,4
No preference	81	4,8
Refusal	2	0,1
No answer	11	0,7
Total	1694	100

Completed interviews

Behavioral data		
Choice	**Frequency**	**Percentage**
Face2face	231	55,2
Mail	164	39,1
Web	24	5,7
Total	419	100

Declaration Data		
Declaration	**Frequency**	**Percentage**
Face2face	1213	71,6
Mail	96	5,7
Web	190	11,2
Phone	27	1,6
Not interested	74	4,4
No preference	81	4,8
Refusal	2	0,1
No answer	11	0,7
Total	1694	100

Method

- Multinomial logistic regression
- Socio-demographic variables from frame (PESEL) included only
- Age and Urbanicity treated as categorical variables – nonlinear relationships?
- Two bases (declaration and behavioral) merged – variable 'type of data' included
- Number of included observations – only completed cases stating preference toward F2F, mail, web (1918)
- **Main question:** Whether declarations of mode preference are a reliable source of information in case of creating mixed-mode design based on the model?

Multinomial Logistic Regression

Pseudo R2=0.1824

Pseudo R2 without data type variable=0.0938

Model (base=F2F)	MAIL RRR	WEB RRR
Female	2.149**	1.196
Age (base 15-19)		
20-24	0.806	1.298
25-34	0.886	0.692
35-44	0.965	0.562*
45-54	1.248	0.322**
55-64	0.964	0.190**
65-74	0.730	0.0637**
75+	0.332*	0.0464**
Urb (base village)		
city <10k	1.330	0.969
city 10-19k	1.859*	2.069*
city 20-49k	1.950**	1.479
city 50-99k	2.456**	3.517**
city 100-199k	1.806*	2.744**
city 200-499k	3.186**	3.363**
city 500k+	3.547**	7.164**
Type of data (base declarative)		
Behavioral data	9.302**	0.620
Constant	0.0332***	0.174***

** p<0.01, * p<0.05

Multinomial Logistic Regression

Pseudo R2=0.1824

Pseudo R2 without data type variable=0.0938

Model (base=F2F)	MAIL RRR	WEB RRR
Female	2.149**	1.196
Age (base 15-19)		
20-24	0.806	1.298
25-34	0.886	0.692
35-44	0.965	0.562*
45-54	1.248	0.322**
55-64	0.964	0.190**
65-74	0.730	0.0637**
75+	0.332*	0.0464**
Urb (base village)		
city <10k	1.330	0.969
city 10-19k	1.859*	2.069*
city 20-49k	1.950**	1.479
city 50-99k	2.456**	3.517**
city 100-199k	1.806*	2.744**
city 200-499k	3.186**	3.363**
city 500k+	3.547**	7.164**
Type of data (base declarative)		
Behavioral data	9.302**	0.620
Constant	0.0332***	0.174***

** p<0.01, * p<0.05

Multinomial Logistic Regression

Pseudo R2=0.1824

Pseudo R2 without data type variable=0.0938

Model (base=F2F)	MAIL RRR	WEB RRR
Female	2.149**	1.196
Age (base 15-19)		
20-24	0.806	1.298
25-34	0.886	0.692
35-44	0.965	0.562*
45-54	1.248	0.322**
55-64	0.964	0.190**
65-74	0.730	0.0637**
75+	0.332*	0.0464**
Urb (base village)		
city <10k	1.330	0.969
city 10-19k	1.859*	2.069*
city 20-49k	1.950**	1.479
city 50-99k	2.456**	3.517**
city 100-199k	1.806*	2.744**
city 200-499k	3.186**	3.363**
city 500k+	3.547**	7.164**
Type of data (base declarative)		
Behavioral data	9.302**	0.620
Constant	0.0332***	0.174***

** $p<0.01$, * $p<0.05$

Multinomial Logistic Regression

Pseudo R2=0.1824

Pseudo R2 without data type variable=0.0938

| Model (base=F2F)	MAIL RRR	WEB RRR
Female	2.149**	1.196
Age (base 15-19)		
20-24	0.806	1.298
25-34	0.886	0.692
35-44	0.965	0.562*
45-54	1.248	0.322**
55-64	0.964	0.190**
65-74	0.730	0.0637**
75+	0.332*	0.0464**
Urb (base village)		
city <10k	1.330	0.969
city 10-19k	1.859*	2.069*
city 20-49k	1.950**	1.479
city 50-99k	2.456**	3.517**
city 100-199k	1.806*	2.744**
city 200-499k	3.186**	3.363**
city 500k+	3.547**	7.164**
Type of data (base declarative)		
Behavioral data	9.302**	0.620 p=0.0501…
Constant	0.0332**	0.174**

** p<0.01, * p<0.05

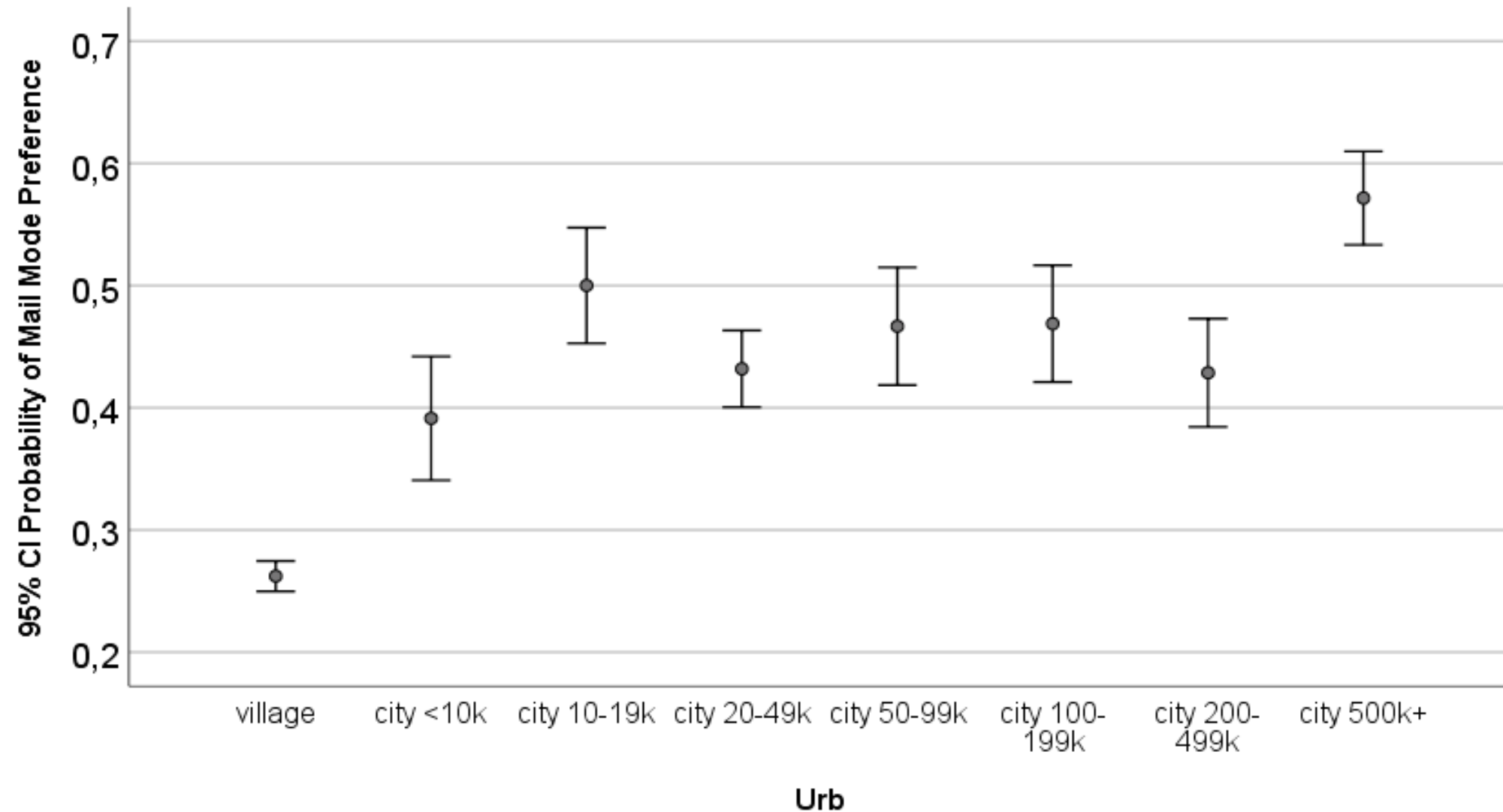# Probabilities of preference toward mail mode Behavioral data
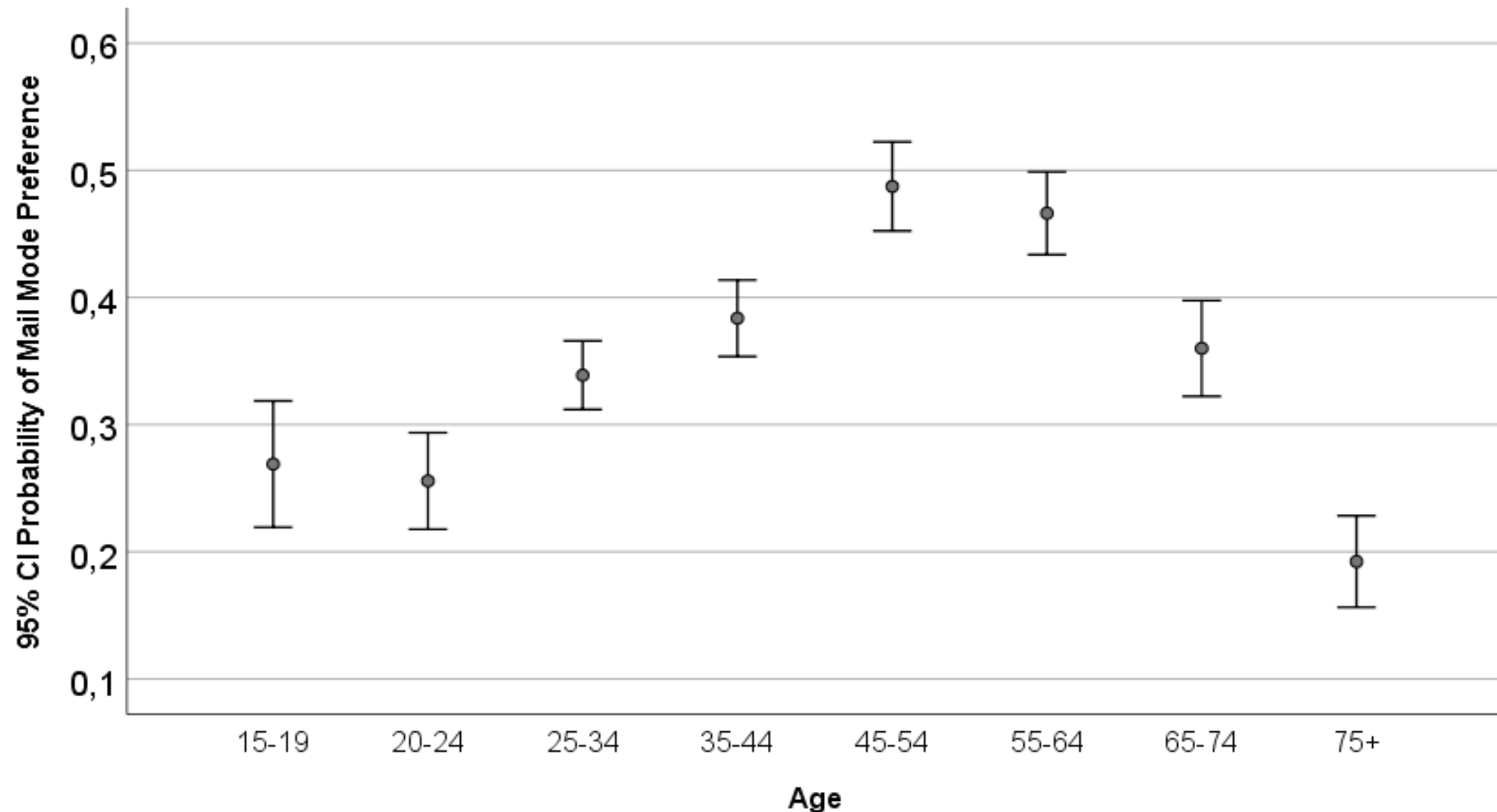
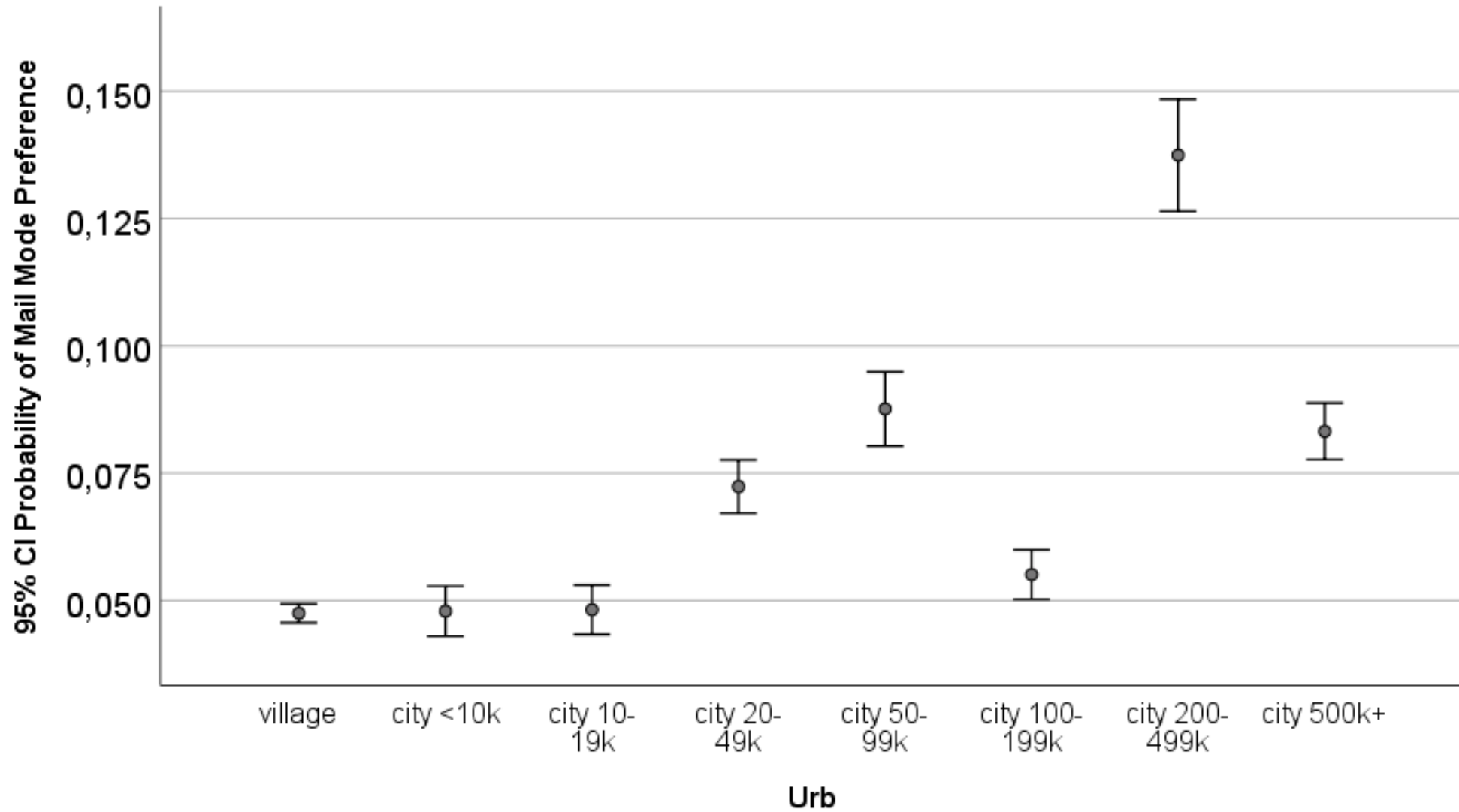# Probabilities of preference toward mail mode Declaration data
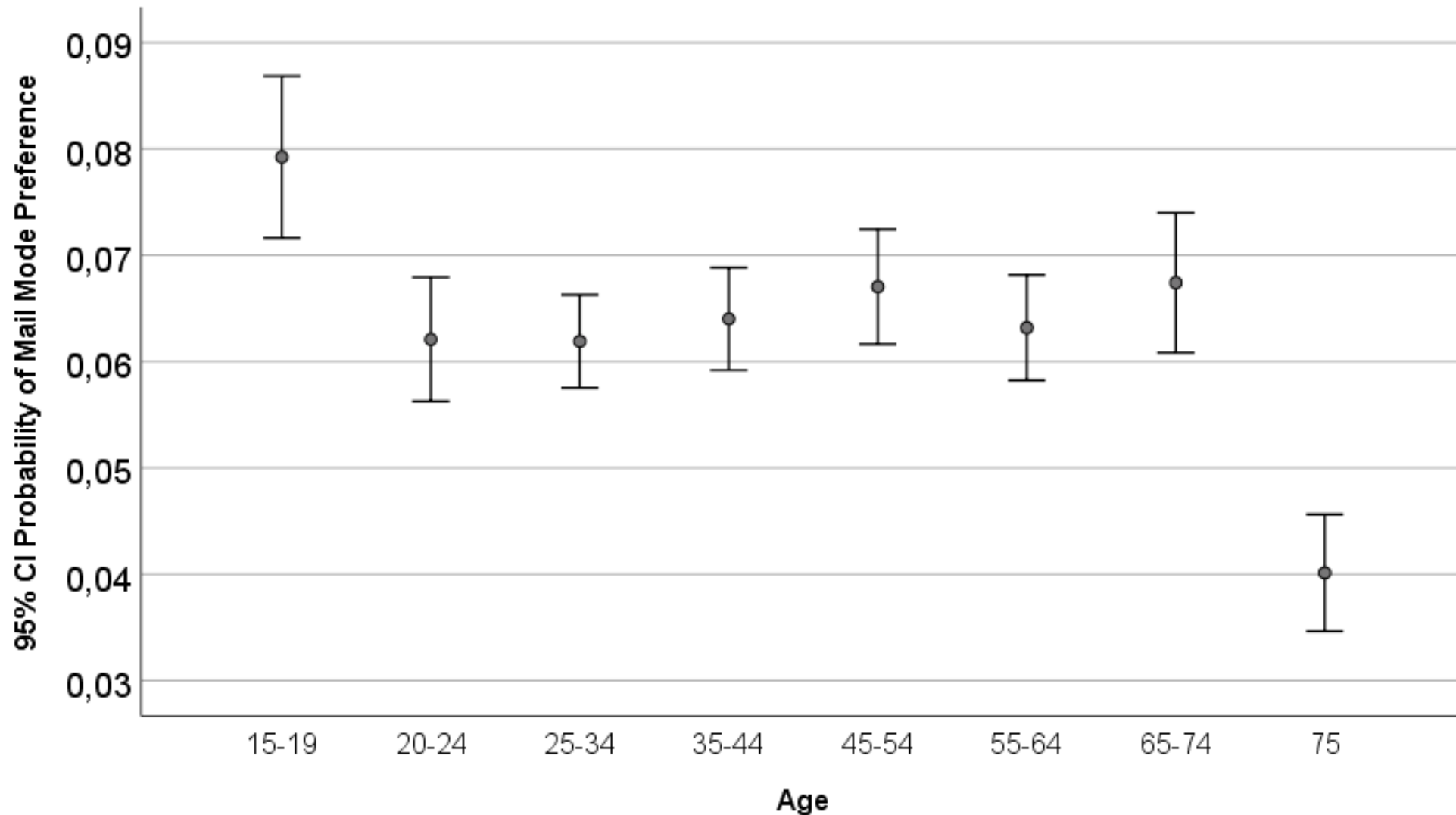
Probabilities of preference toward mail mode
Behavioral data

Probabilities of preference toward mail mode
Declaration data

Conclusion

- Socio-demographic predictors are significant
- But there is huge difference in declaration of mail mode preference rate, and mail choices rate

- It is better to gather data from mixed-mode parallel design with choice of mode
- To construct effective mixed-mode design with mode pushing

Thank you for your attention

Adam Rybak
adam.rybak@amu.edu.pl

Institute of Sociology