The Impact of Mixed Modes on Measurement Error

Joseph W. Sakshaug
Institute for Employment Research (IAB)
University of Mannheim

Alexandru Cernat
University of Manchester

21st June 2019

Background

- Mixing modes of data collection becoming standard practice in surveys

- Examples
  - American community survey
    - Web, mail, telephone, FtF
  - The UK Household Longitudinal Study (Understanding Society)
    - Web, FtF

Advantages of Mixed-Mode Designs

- Potential cost savings
- Improve coverage of target population
- Can improve response rates, bring in different kinds of respondents

Drawbacks of Mixed-Mode Designs

- Different modes can evoke different responses to the same questions by the same group of respondents

- Potential for differential measurement errors
  - Self-administered modes tend to elicit more honest answers than interviewer modes
  - Other systematic measurement errors (e.g. acquiescence, extreme response reporting) shown to vary by mode

Mixing Modes and Measurement Error

- Mixing modes with inherently different ME properties produces at least two *unwanted* effects:

1) Compromises the accuracy of comparisons of respondents interviewed in different modes
2) Compromises the accuracy of comparisons to other surveys employing single or different modes

- Especially problematic in longitudinal studies
  - measures of change may reflect measurement error effects rather than actual changes

Methods of Studying Measurement Error

- Record check
- Repeated measurement
- Multitrait-Multimethod
- Qualitative settings (focus groups, lab studies)
- …among others

Common Limitations of ME Studies

- Investigate only one type of ME at a time
  - Different ME types assumed to be *independent* of each other

- Confounding of selection and measurement
  - ME effects could actually be due to differences in respondent composition by mode

Research Questions

- To what extent does a mixed-mode (Web-face-to-face) data collection approach produce different measurement errors, relative to a single mode (face-to-face) approach?

- To what degree does face-to-face versus Web produce different measurement errors?

Survey Implementation

- Understanding Society Innovation Panel
  - Waves 7, 8, and 9

- Quasi-experimental design
  - 1/3 sample -> unimode (FtF)
  - 2/3 sample: sequential mixed-mode (Web-FtF)

Multitrait-Multierror

- MTME experiment implemented using six questions regarding attitudes towards immigrants
  - Split ballot

- Types of errors estimated
  - Social desirability
  - Acquiescence
  - Method effect

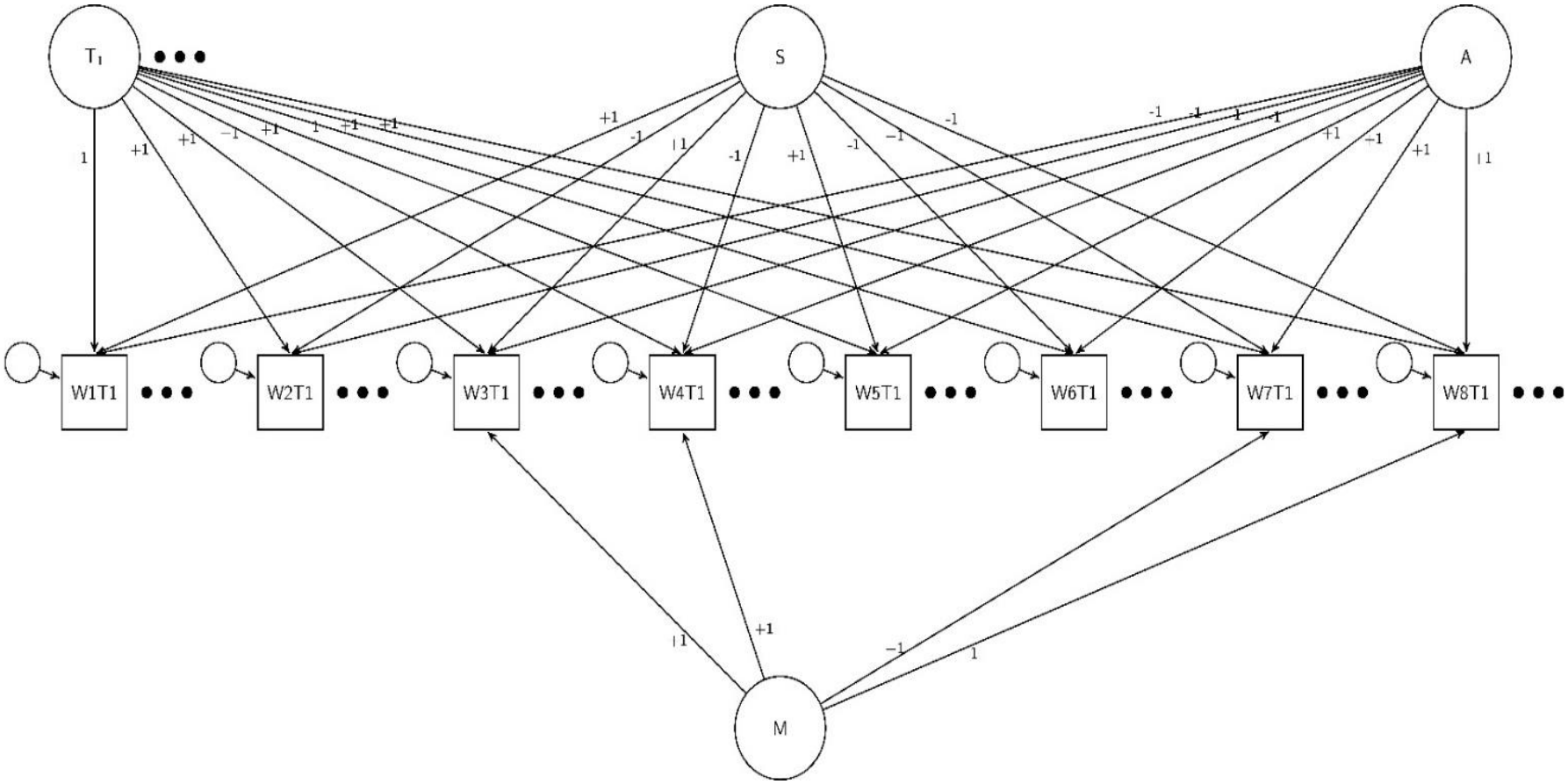
Items measuring attitudes towards immigrants

Trait number	Item formulation
T1	The UK should allow more people of the same race or ethnic group as most British people to come and live here
T2	UK should allow more people of a different race or ethnic group from most British people to come and live here
T3	UK should allow more people from the poorer countries outside Europe to come and live here
T4	It is generally good for UK's economy that people come to live here from other countries
T5	UK's cultural life is generally enriched by people coming to live here from other countries
T6	UK is made a better place to live by people coming to live here from other countries

Experimental Variations (1 example trait)

Wording number	Social desirability	Number of scale points	Agree or Disagree	Item formulation (using trait 1 as an example)
W1	Higher	2	AD	The UK should allow fewer people of the same race or ethnic group as most British people to come and live here
W2	Lower	2	AD	The UK should allow more people of the same race or ethnic group as most British people to come and live here
W3	Higher	11	AD	The UK should allow fewer people of the same race or ethnic group as most British people to come and live here
W4	Lower	11	AD	The UK should allow more people of the same race or ethnic group as most British people to come and live here
W5	Higher	2	DA	The UK should allow more people of the same race or ethnic group as most British people to come and live here
W6	Lower	2	DA	The UK should allow fewer people of the same race or ethnic group as most British people to come and live here
W7	Higher	11	DA	The UK should allow more people of the same race or ethnic group as most British people to come and live here
W8	Lower	11	DA	The UK should allow fewer people of the same race or ethnic group as most British people to come and live here

Bayesian Structural Equation Model



Modeling Approach

	Mode design comparison (RQ1)	Single mode comparison (RQ2)
Exploration	Compare means and variances of ME using posterior distributions	Compare means and variances of ME using posterior distributions
Inference	a. Regress measurement error on mode design b. Investigate mode design regression coefficient and R^2	a. Regress measurement error on control variables b. Regress measurement error on control variables and mode c. Investigate mode regression coefficient and change in R^2

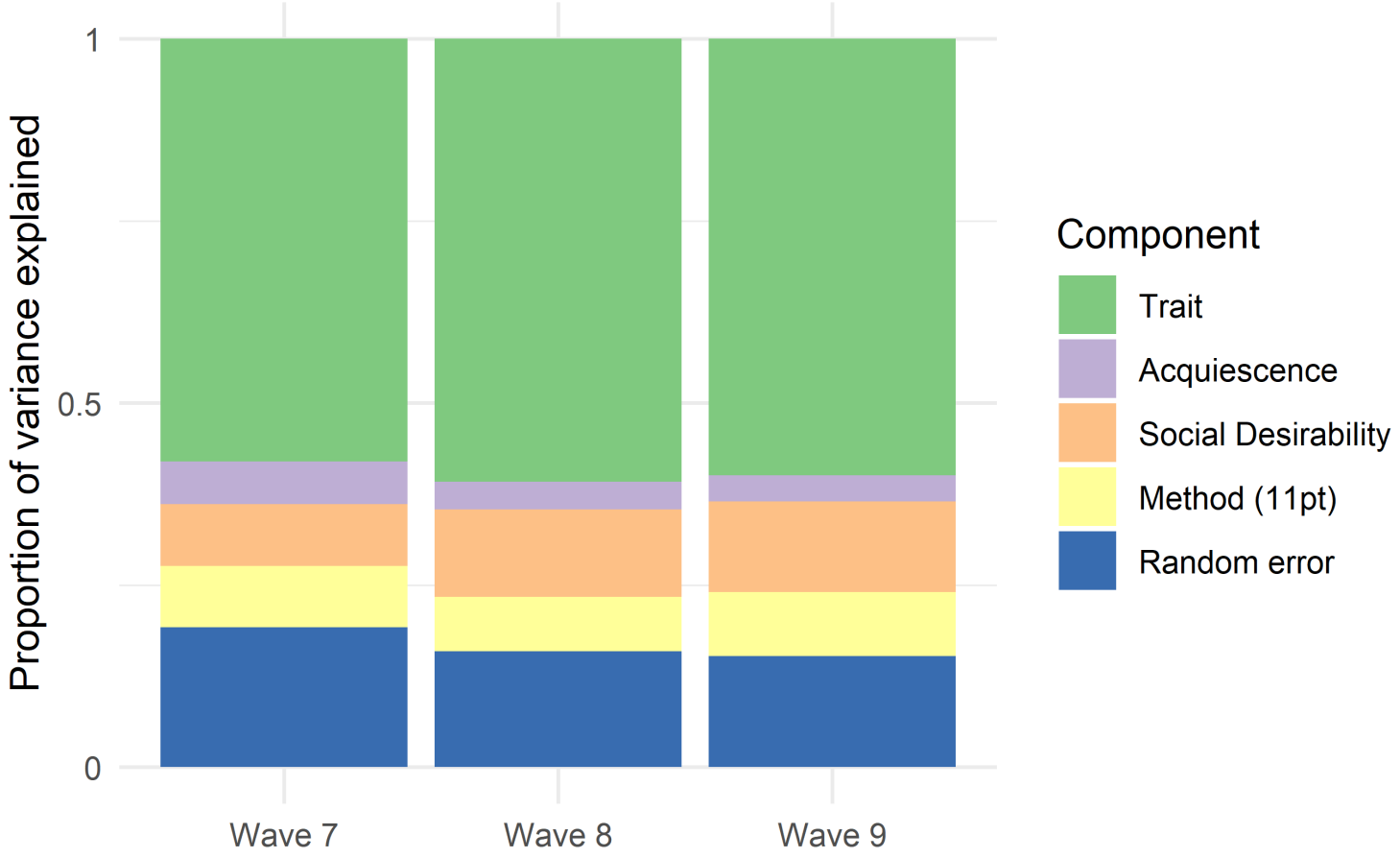
$$Y = \beta_0 + \beta_1 * mode\ (design) + \boldsymbol{\beta_2} * controls + \epsilon$$

Means/Variances of 3 Types of ME

	Means			Variances		
Wave	Soc. des.	Method	Acq.	Soc. des.	Method	Acq.
7	-0.18	5.13	0.25	0.29	0.86	0.42
8	-0.13	4.94	0.16	0.40	0.75	0.60
9	-0.33	5.03	0.26	0.98	0.88	0.44

- Mean of ME indicates how much the average response on a trait is changed by an experimental condition. Can bias point estimates.
- Variance of ME indicates how much the average response on that trait varies within an individual. Can bias multivariate analyses.

Variance Decomposition by ME Type



- Trait represents about 60% of total variance.
- Random error is the largest source of non-trait variance followed by social desirability, method, and acquiescence

Regression Slope and R^2 (Difference b/w Mixed-Mode and Unimode)

Wave	ME	Est.	Lower C.I.	Upper C.I.	R^2
7	**Social desirability**	**-0.22**	**-0.43**	**-0.08**	**2.5**
	Method	-0.05	-0.22	0.11	0.1
	Acquiescence	0.01	-0.09	0.12	0.1
8	**Social desirability**	**-0.18**	**-0.36**	**-0.06**	**1.4**
	Method	-0.11	-0.26	0.04	0.4
	Acquiescence	-0.03	-0.17	0.11	0.1
9	**Social desirability**	**-0.37**	**-0.66**	**-0.17**	**2.7**
	Method	-0.02	-0.19	0.15	0.1
	Acquiescence	**-0.12**	**-0.25**	**0.00**	**0.8**

The mixed-mode respondents have more extreme levels of social desirability compared to respondents allocated to the single-mode design

Regression Slope and R^2 (Difference b/w Web and FtF; w/ controls)

Wave	ME	Est	Lower C.I.	Upper C.I.	R^2 extra
7	**Social desirability**	**0.48**	**0.29**	**0.69**	**3.8**
	Method	0.00	-0.12	0.26	0.3
	Acquiescence	-0.01	-0.12	0.10	0.1
8	**Social desirability**	**0.51**	**0.23**	**0.81**	**1.1**
	Method	0.04	-0.15	0.22	0.3
	Acquiescence	0.01	-0.19	0.20	0
9	**Social desirability**	**0.77**	**0.45**	**1.11**	**1.7**
	Method	-0.05	-0.23	0.13	0.4
	Acquiescence	0.08	-0.08	0.24	0.8

The expected mean for CAPI respondents on the social desirability variable is *lower* than for the Web respondents

Conclusions (I)

- We investigated the impact of mode and mode design on ME using combination of experimental designs and statistical modeling
- Experimental design allows us to control for confounding of ME and selection
- MTME approach allows simultaneous estimation of multiple types of ME
  - Social desirability, acquiescence, method effect

Conclusions (II)

- Small differences in ME across mode (designs)
  - Reassuring for survey practice
- Social desirability was systematically different by mode (design)
  - But explains only small amount of variance
  - Surprising: Mean effect of SD larger in Web mode
- No mode (design) differences wrt acqueisence and method effects

Thank you for your attention

Questions? / Comments? / Collaborations?

Contact:

joe.sakshaug@iab.de

alexandru.cernat@manchester.ac.uk