

Can Weighting Compensate for Sampling Issues in Online Surveys?

Jerry J. Vaske

Colorado State University, USA

Maarten H. Jacobs

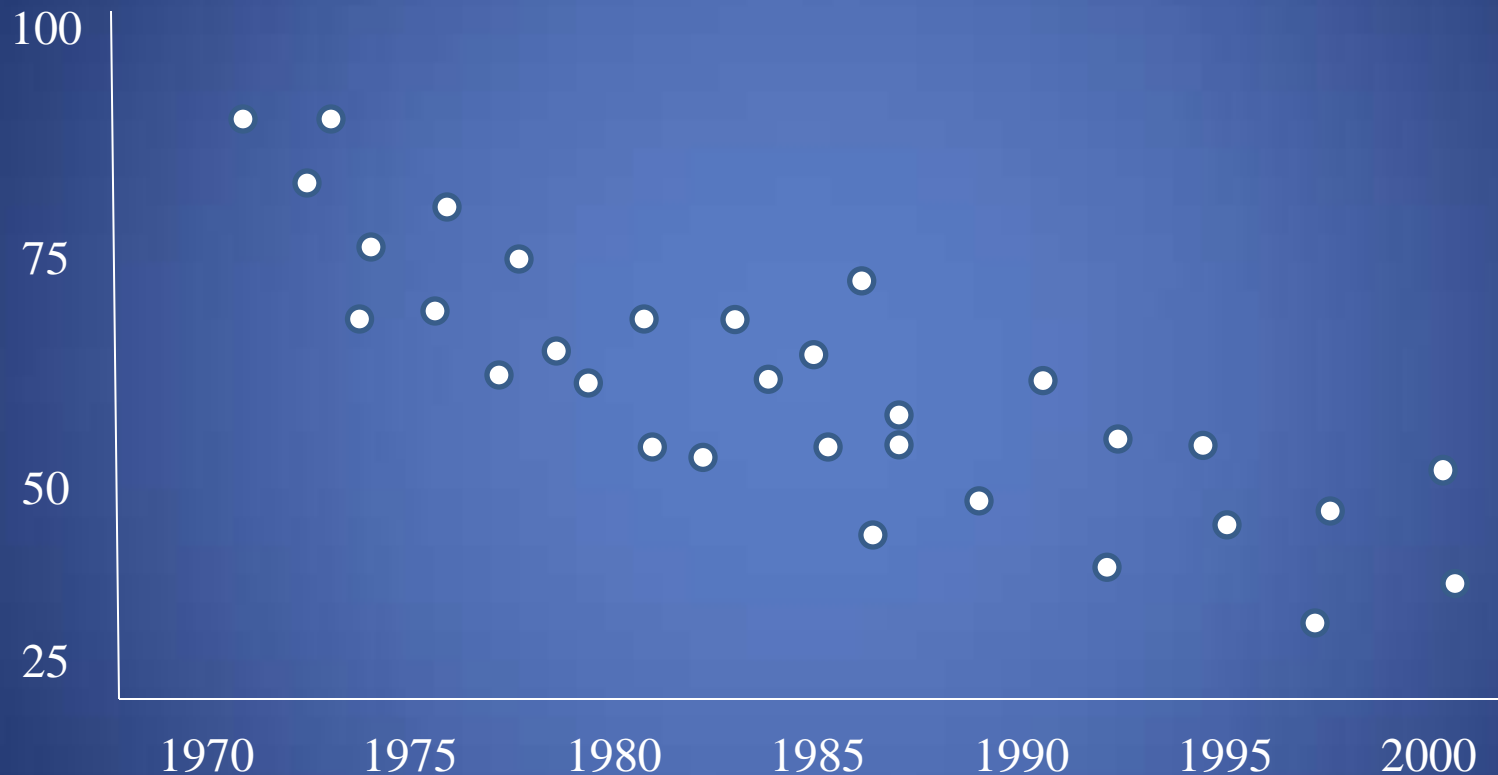
Wageningen University, NL

Mette T. J. Sijtsma

Wageningen University, NL

Trend 1

- Survey response rates have declined over time



- If sample not representative of target population data may need to be weighted

Trend 2

Online surveys are becoming more popular

- Advantages
 - Eliminate mailing costs
 - Minimize data entry errors
 - Reduce project completion times
- Disadvantages
 - May not be representative of population (e.g., online respondents are sometimes younger, urban, males)

Basic Research Question

Can weighting the data compensate for the lack of representativeness in:

- Mail surveys?
- Online surveys?

Weighting Data

- Weighting by nonresponse checks
- Weighting by population proportions
 - With *wildlife surveys* population ratios might be based on a single variable (e.g., license sales)
 - With *general public surveys* population ratios might use multiple variables (e.g., census data)

Understanding Weighting

- Assume

mail survey:

Population

– Response rate	40%	
– % Males	30%	60%
– % Females	70%	40%

- Suggests results

- May not be representative of population
- Need to weight the data

Understanding Weighting – Single Variable

$$\text{Weight} = \frac{\text{Population}\%}{\text{Sample}\%}$$

	Sample	Population
Males	30%	60%
Females	70%	40%

$$\text{Weight (males)} = \frac{0.60}{0.30} = 2.0$$

$$\text{Weight (females)} = \frac{0.40}{0.70} = 0.57$$

Understanding Weighting – Multiple Variables

Sex	Residence	Age	Pop %	Sample %	Weight
Male	Rural	20-45	9	9	1.00
		46-65	8	4	2.00
		> 65	8	5	1.60
	Urban	20-45	10	12	0.83
		46-65	11	14	0.79
		> 65	8	11	0.73
Female	Rural	20-45	5	2	2.50
		46-65	7	10	0.70
		> 65	4	2	2.00
	Urban	20-45	14	17	0.82
		46-65	10	9	1.11
		> 65	6	4	1.50

Study Context

- General population survey in the Netherlands
- Survey focus:
 - Wildlife value orientations
 - Demographics
- Two versions of survey
 - Random sample mail survey
($n = 353$; response rate = 18%)
 - Online internet survey
($n = 216$; response rate = unknown)

Weighting Strategy

- Given the response rate,
weighted data based on Dutch census data
- 3 Weighting Variables
 - Sex: Male, Female
 - Current
residence: Rural, Urban
 - Age: 20 to 45, 46 to 65, 66 +

Dutch Population & Sample Data

Sex	Residence	Age	Pop %	Mail Sample %	Online Sample %
Male	Rural	20-45	3.4	3.7	8.3
		46-65	3.1	8.8	4.4
		> 65	1.4	5.1	0
	Urban	20-45	14.4	7.9	27.1
		46-65	10.1	20.4	6.6
		> 65	4.6	12.8	0
Female	Rural	20-45	3.3	4.5	7.7
		46-65	2.9	5.9	4.4
		> 65	1.8	1.4	0
	Urban	20-45	14.3	15.0	36.5
		46-65	10.1	9.9	5.0
		> 65	6.5	5.1	0

Dutch Population & Sample Data

Sex	Residence	Age	Pop %	Mail Sample %	Online Sample %
Male	Rural	20-45	3.4	3.7	8.3
		46-65	3.1	8.8	4.4
		> 65	1.4	5.1	0
	Urban	20-45	14.4	7.9	27.1
		46-65	10.1	20.4	6.6
		> 65	4.6	12.8	0
Female	Rural	20-45	3.3	4.5	7.7
		46-65	2.9	5.9	4.4
		> 65	1.8	1.4	0
	Urban	20-45	14.3	15.0	36.5
		46-65	10.1	9.9	5.0
		> 65	6.5	5.1	0

Dutch Population & Sample Data

Sex	Residence	Age	Pop %	Mail Sample %	Online Sample %
Male	Rural	20-45	3.4	3.7	8.3
		46-65	3.1	8.8	4.4
		> 65	1.4	5.1	0
	Urban	20-45	14.4	7.9	27.1
		46-65	10.1	20.4	6.6
		> 65	4.6	12.8	0
Female	Rural	20-45	3.3	4.5	7.7
		46-65	2.9	5.9	4.4
		> 65	1.8	1.4	0
	Urban	20-45	14.3	15.0	36.5
		46-65	10.1	9.9	5.0
		> 65	6.5	5.1	0

Population vs. Samples

	%
	Pop.
Gender	
Male	49
Female	51
Age	
20 – 45	24
46 – 65	61
> 65	15

Population vs. Samples

	%	Mail %	
	Pop.	Unweighted	Weighted
Gender			
Male	49	58	49
Female	51	42	51
Age			
20 – 45	24	31	24
46 – 65	61	45	58
> 65	15	24	18

Population vs. Samples

	%	Mail %		Online %	
	Pop.	Unweighted	Weighted	Unweighted	Weighted
Gender					
Male	49	58	49	47	50
Female	51	42	51	53	50
Age					
20 – 45	24	31	24	80	57
46 – 65	61	45	58	20	43
> 65	15	24	18	0	0

Population vs. Samples

	%	Mail %		Online	
	Pop.	Unweighted	Weighted	Unweighted	Weighted
Residence					
Rural	21	30	21	25	20
Urban	79	70	79	75	80

Population vs. Samples

	%	Mail %		Online	
	Pop.	Unweighted	Weighted	Unweighted	Weighted
Residence					
Rural	21	30	21	25	20
Urban	79	70	79	75	80
Education					
≤ H.S.		48	45	12	19
Higher Degree		52	55	88	81

Impact on Wildlife Value Orientations

	Mail %	Online %
Domination	21	27
Neither	35	52
Mutualism	44	21

$\chi^2 = 25.90, p < .001, \text{Cramer's } V = .215$

Conclusions

- Neither mail nor online surveys reflected Dutch pop.
- Weighting strategy resolved the problem for:
 - Random sample mail survey
 - But *not* for online survey
- Explanation
 - Can't weight for empty cells (e.g., older respondents)
 - Large percentages in some cells (e.g., urban & 20–45)
 - 81% in Higher Degree response category
- *Summary*: Weighting did *not* compensate for online

Questions

