

Summary of biases and steps to redress them in the design and implementation of survey

<i>Bias</i>	<i>Defined</i>	<i>Steps to redress bias</i>
I. ENDOGENEITY:		
1. Omitted variable bias	Omitted variables – either unseen or seen but unmeasured – are not included. <u>Consequence:</u> biases coefficients by the sign and size of the indirect effect.	Measure variables only if they likely affect X & Y; otherwise try to assess direction & magnitude of bias
		Panel; first difference
1.a. Attrition	Subjects leave sample for non-random reason; see also self-selection. Only problem with panel; in cross sectional study, akin to refusal. <u>Consequence:</u> besides omitted variable bias problems, creates loss of statistical power and incorrect inferences about univariate distribution of variables	Aggressive tracking of attriters; at baseline ask subject about name/address of neighbors/close kin who could be asked in the future the whereabouts of the subject.
		Over sample subjects likely to leave, but weigh on right var
1.b. Self selection	Only subjects with certain characteristics included in sample (e.g., in study of schooling, wealthy parents send kids to school). <u>Consequence:</u> parameter estimated biased by omitted variables driving self-selection.	Compare attributes of people in sample with sample of people of similar characteristics who decide not to join sample and test for differences
2. Reverse causality	Y and X reciprocally related. <u>Consequence:</u> unsure of which way causality goes.	Instrumental variables or quasi experimental design
		Lag x's
3. Random measurement error	X or Y variables measured with random error (not systematic error). Random measurement of x called "classical measurement error". <u>Consequence:</u> in X, attenuation bias; in Y, larger standard errors	Instrumental variables
II. SAMPLING		
1. Sampling	Sample ≠ representative of larger population; lack of external validity. Pop of interest=small share of total pop, so numerically uninteresting	Significance of pop in larger pop of relevance. Share of sample in larger pop
	Not enough statistical power (observations) to say something meaningful about the population	Do power calculation before study
2. Additions	Subjects enter the sample after start (e.g., birth, in-marrying); only affects panel/experimental study.	Inclusion rules; if in, note they are additions and do analysis +/- them to see if results change
III. SYSTEMATIC MEASUREMENT ERROR (Only in panel; 2-5 could be random as above)		
1. Contamination	Subjects in treatment (T) and control (C) interact; T tells C about intervention. Only a problem in experimental study	Physical separation; include module asking about frequency of interaction between T & C
2. Response burden	Subjects get tired by length of interview and give short answers to finish soon	Be sensitive to length of survey
3. Conditioning	Subjects learn how to respond and provide you with more accurate (e.g., if question involves tests) or less accurate answers (e.g., if subjects know that if they answer in a certain way they will have to answer more questions); only in panel	Use two parallel fresh, look alike cross sectional surveys and compare change in your panel with changes in the cross sectional surveys
4. Hawthorne	Only in experimental studies; subjects given treatment respond favorably because they like the attention they receive from researcher	Include 2 ND intervention; include modules asking subjects about attitudes to researcher
5. Halo	Only in experimental studies; subjects react positively to intervention because it comes from an outsider who knows best	Ask about attitudes toward outsiders