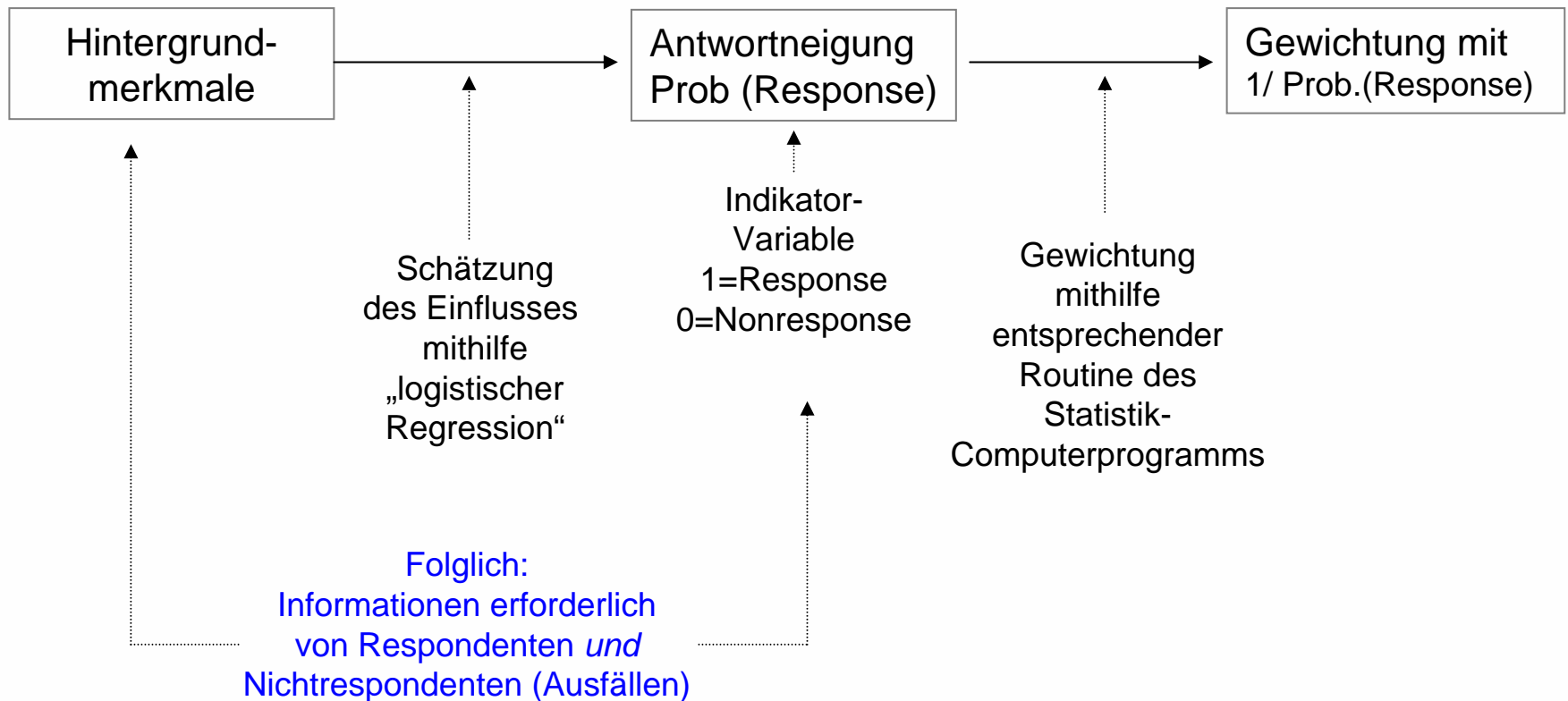


Korrekturmöglichkeiten durch geeignete Gewichtung eines Datensatzes



Gewichtung nach Antwortneigung (propensity weighting)

1. Regression des $[0,1]$ Indikators für unit response r_i auf ein geeignetes, **zugleich für Respondenten und Nichtrespondenten verfügbares** Set von Hintergrundmerkmalen (logistische Regression);
2. Berechnung *erwarteter* Antwortwahrscheinlichkeiten auf der Basis dieser Regression, und
3. Gruppierung dieser Wahrscheinlichkeiten zu fünf oder sechs Werten, um
4. über den so gruppierten *propensity score* jeweils eine Adjustierungs- bzw. Gewichtungsklasse C_j zu bilden. (response propensity stratification)

Alternative

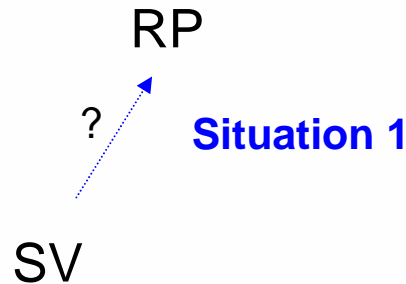
Verzicht auf Schritt 4 (Gruppierung) und gleich

Gewichtung der Respondenten i mit der Inversen
des geschätzten propensity scores,

$$[\hat{p}_i]^{-1}$$

Informationen über Nichtrespondenten

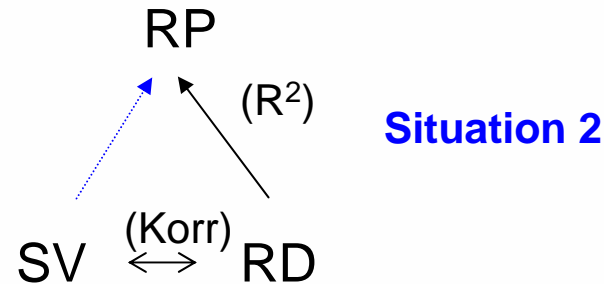
- über separate follow-up Surveys
- über Registerinformationen
 - Einwohnermeldeamt
 - Access Panels (wenn Sampling Frame selbst unverzerrt ist)
- über Informationen aus anderen Wellen einer Panelstudie
- über „Proxy – Nichtrespondenten“



Es ist unbekannt, inwieweit RP von von SV abhängt, Da keine Informationen über Nicht-Respondenten vorliegen.

Folglich ist etwaiger Nonresponse bias in SV basierten Schätzungen unbekannt

RP response propensity
SV study variable(s)
RD register/regional data



Einbezug externer Quelle(n), hier von RD, und Gewichtung auf dieser Basis

Effektivität des Ansatzes hängt von der Höhe von (R^2) und $(Korr)$ ab:

- Je stärker RP durch RD erklärt wird, desto kleiner wird im Prinzip der Spielraum für eigenständige Effekte von SV auf RP;
- Je stärker SV und RD korreliert sind, desto stärker wird ein Gewichtungsausgleich nach RD-Effekten auf RP auch einen indirekten Gewichtungsausgleich nach etwaigen SV-Effekten bewirken (da beide qua Korrelation miteinander konfundiert wären). Falls $(Korr)$ gleich null ist, kann es keinen solchen indirekten Gewichtungs-Effekt geben. Eine RD basierte Gewichtung würde etwaigen Bias in SV basierten Schätzungen unberührt lassen.
- Komplikation: $(Korr)$ kann nur für Respondenten berechnet werden; für die Gruppe der Nichtrespondenten fehlt SV ja gerade (nur RD ist für beide Teilgruppen verfügbar)

Befragungen im Rahmen von Bremen's „Stadt-der-Wissenschaft“ - Projekt

Wahrscheinlichkeit der Umfrageteilnahme

Erklärung durch	Erklärungsbeitrag**
Geschlecht, Alter	1,1%
Geschlecht, Alter, Arbeitslosenquote des Ortsteils	3,2%
Geschlecht, Alter, Ortsteilzugehörigkeit	9,7%

Zum Vergleich (Bürgerbefragungen der Polizei Bremens 2004 und 2005):

Stadtteilzugehörigkeit	6,1%
Charakteristika der Stadtteile*	2,5%



fr_6 an veranstaltungen teilgenommen

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig 1 ja	34	6,6	10,3	10,3
2 nein	298	57,6	89,3	99,6
9 weiß nicht	1	,3	,4	100,0
Gesamt	334	64,5	100,0	
Fehlend -77	6	1,1		
System	178	34,4		
Gesamt	184	35,5		
Gesamt	518	100,0		

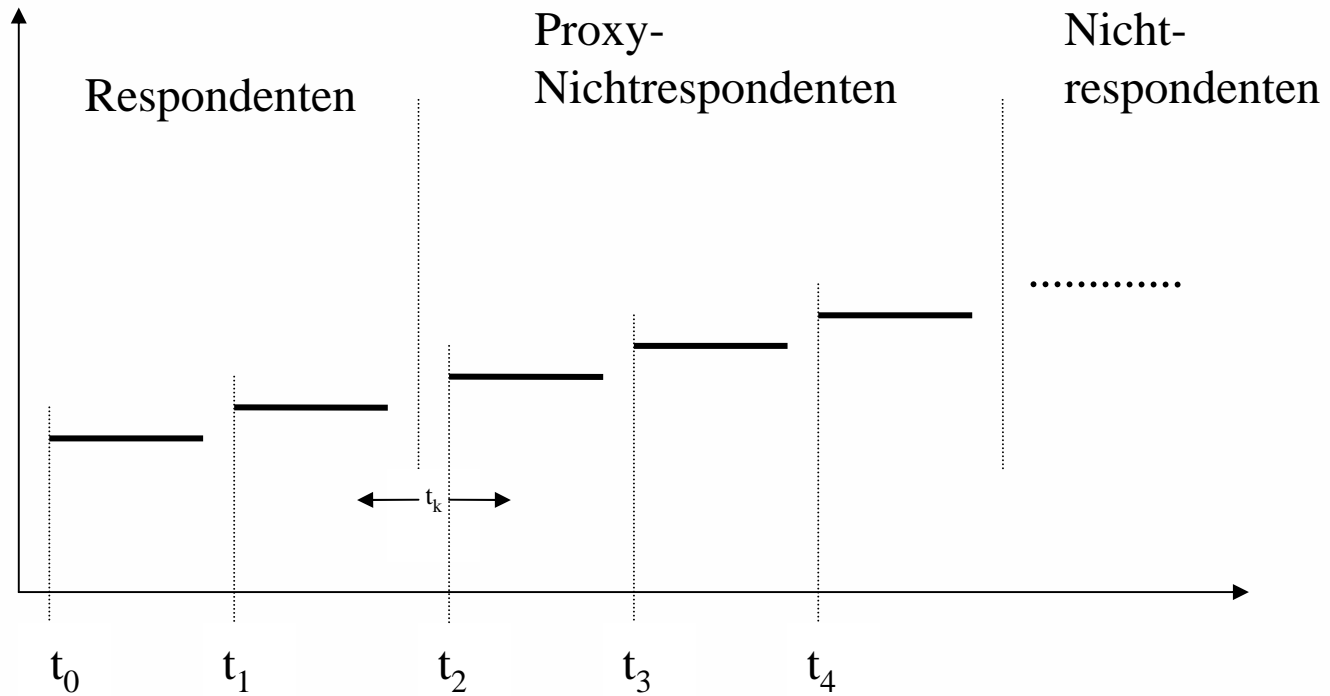
gewichtet

fr_6 an veranstaltungen teilgenommen

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig 1 ja	40	7,7	12,0	12,0
2 nein	291	56,2	87,4	99,4
9 weiß nicht	2	,4	,6	100,0
Gesamt	333	64,3	100,0	
Fehlend -77	7	1,4		
System	178	34,4		
Gesamt	185	35,7		
Gesamt	518	100,0		

ungewichtet

*Gewichtet mit Inverse der mit logistischer Regression geschätzten Antwortwahrscheinlichkeit (auf n=518 normiert)



Flexible Survey Designs bieten ..

- zugeschnittenen („tailored“) Interviewer-vermittelten Kontakt (im Sinne der „leverage-saliency“ Theorie)

in Abhängigkeit von der Stärke der Survey Kooperation:

- Einsatz einer langen, intermediären, kurzen Fragebogenversion
- Flexibles Zeitmanagement von Interviews