

Braucht die Wissenschaft noch Zufallsstichproben?

Möglichkeiten und Grenzen alternativer Stichprobenverfahren

Prof. Dr. Rainer Schnell

Research Methodology Group
Universität Duisburg-Essen

8. Konferenz für Sozial- und Wirtschaftsdaten
2. März 2020, Berlin

Geschichte

Theorie

Bias von Web-Surveys

Missing Data
Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Die Homogenität sozialer Kategorien als Voraussetzung für „Repräsentativität“ und Gewichtungsverfahren¹

Rainer Schnell

Fakultät für Soziologie, Universität Mannheim, Gebäude A 5, D-6800 Mannheim 1

Zusammenfassung: Die meisten Datensätze der empirischen Sozialforschung basieren auf Surveyinterviews. Das größte methodische Problem bei Surveyerhebungen sind Ausfälle (Nonresponse), insbesondere durch Verweigerungen. Zur Korrektur von Ausfällen werden in der Praxis Gewichtungsverfahren eingesetzt. Allen Gewichtungsverfahren liegen inhaltliche, soziologisch relevante Annahmen zugrunde. Diese Annahmen werden jedoch fast nie expliziert oder überprüft. Zentral für Gewichtungsverfahren ist die Annahme, daß die Gewichtungsklassen homogen sind. Auf ähnlichen Annahmen basieren auch Quota-Verfahren und sogenannte „Repräsentanzbeweise“. Anhand empirischer Daten wird gezeigt, daß diese zentrale Annahme falsch ist. Weiterhin wird mit einer Simulation systematischer Ausfälle auf der Basis der Daten des ALLBUS 1980 demonstriert, daß Standardgewichtungsverfahren nicht in der Lage sind, systematische Ausfälle zu kompensieren.

1. Einleitung

Die meisten Untersuchungen der empirischen Sozialforschung basieren auf standardisierten Interviews im Rahmen von Zufallsstichproben. Zu den methodisch bedeutsamsten Problemen bei Zufallsstichproben aus der allgemeinen Bevölkerung gehört das Nonresponseproblem: Ein Teil der ausgewählten Stichprobe fällt durch Abwesenheit, Krankheit oder Verweigerung aus der Befragung aus.

Das Ausmaß des Nonresponse scheint in den letzten Jahren in den meisten Industrieländern zugenommen zu haben. In der BRD dürften bei den meisten Zufallsstichproben der kommerziellen Institute die Ausschöpfungsquoten derzeit eher zwischen 50 und 60 Prozent liegen als bei 70 Prozent (Hansen 1988: 399). Das Ausmaß des Nonresponse scheint viele Rezipienten empirischer Sozialforschung zu beunruhigen: So scheint es Wyss (1990: 71) notwendig, darauf hinzuweisen, daß eine 70%-Ausschöpfung als Gütekriterium „wissenschaftlich“ nicht begründet werden kann. Kromrey (1990: 202) glaubt (fälschlicherweise), daß bei „Rücklaufquoten um 70 Prozent“ „auch mit den ausgefeiltesten Konzepten der Inferenzstatistik keinerlei Aussage über die Genauigkeit der Ergeb-

nisse mehr möglich“ sei.² Dies hat einige Autoren dazu gebracht, die „Repräsentativitätsproblematik“ qualitativer Methoden mit denen von Zufallsstichproben zu vergleichen (Lamnek 1989: 92) oder gar für Quota-Samples zu plädieren (Lamnek 1978: 568–569).³

Im allgemeinen Umgang mit dem Nonresponseproblem wird vor allem auf einfache Konzepte zurückgegriffen: „Repräsentativität“ zeigt sich in der Übereinstimmung von Stichprobenhäufigkeiten mit den entsprechenden bekannten Anteilen in der Grundgesamtheit.⁴ Wenn sich diese nicht ohnehin

² Nur um Mißverständnisse zu vermeiden: Gerade bei solchen Ausschöpfungsquoten gibt es nur noch eine Möglichkeit einer begründbaren Aussage: „Multiple Imputation“ (Rubin 1987), siehe dazu weiter unten.

³ Dabei wird übersehen, daß Quota-Stichproben ebenfalls unter Nonresponse leiden: Bausch (1990: 77) schreibt explizit, daß der Vorteil des Quotenverfahrens darin besteht, daß es kein Nonresponseproblem besäße. Dies ist falsch. Das Nonresponseproblem bei Quota-Stichproben wird nur verdeckt: Da das Ausmaß des Nonresponse nicht ersichtlich ist, kann bei solchen „Stichproben“ keinerlei „Korrektur“ oder auch nur der geringste Ansatz einer Abschätzung der Verzerrung erfolgen: Man kann nur noch an die Unverzerrtheit der Ergebnisse glauben. Dies gehört nicht zu den üblicherweise akzeptierten Voraussetzungen der Anwendung einer Methode.

⁴ Entgegen weitverbreiteten Ansichten ist „Repräsentativität“ kein Begriff, der in der technischen und mathematischen Literatur zu Auswahlverfahren verwendet wird: In diesem Sinn handelt es sich um keinen „wissenschaftlichen“ Begriff, vgl. Kruskal/Moeller (1979a, 1979b; 1990c; 1980), zusammenfassend Schnell et al. (1992: 314–315).

¹ Den Anstoß für diese Arbeit gab eine Diskussion mit Bernhard von Rosenblatt anläßlich der Tagung der Methodensektion der DGS in Berlin im April 1991. Für kritische Anmerkungen danke ich Stefan Bender, Elke Esser, Frank Kalter, Thomas Klein und Johannes Kopp.

Was ist eine geeignete Stichprobe?

- Die Frage danach, was eine „repräsentative“ Stichprobe ist, ist älter als jedes Sozialforschungsinstitut, nämlich mindestens 123 Jahre.
- Es dauerte dann mehr als 30 Jahre, bis dieses Frage geklärt war.
- Die Antwort ist aus heutiger Sicht einfach: Man muss die Auswahlwahrscheinlichkeit für jedes Element der Grundgesamtheit kennen.
- Ist dies der Fall, benötigt man den Begriff nicht.
- Entsprechend spielt der Begriff in der mathematischen Literatur zu Stichproben keine Rolle.
- Vor fast 60 Jahren fragte Friedrich Wendt, wann das Quotenverfahren endlich begraben würde. Bis heute ist es nicht gelungen, dieses untote Verfahren zur endgültigen Ruhe zu legen.

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Ursachen für die Verwendung von Nonprobability-Samples

- Die Ursachen dafür sind trivial:
 - Erstens funktionieren „non-probability samples“ oft (man weiß nur nicht vorher, wann).
 - Zweitens gibt es wenige wissenschaftliche Vergleiche mit externen Daten.
- Vorgehensweisen, die selten mit der Wirklichkeit verglichen werden und bei denen dann das Versagen nicht sicher erfolgt, werden in der Regel fortgesetzt.
- Das ist bei der Anwendung willkürlicher Stichproben nicht anders als bei abergläubischen Verhalten.¹
- Entsprechend klein ist der Fortschritt in wissenschaftlichen Randbereichen, die auf empirische Tests verzichten.

¹Berlin besitzt eine psychoanalytische Universität (IPU). Und den Zentralverein homöopathischer Ärzte.

- Im Gegensatz zu 1993 können wir dank der Arbeiten von Jelke Bethlehem nun den Bias von Non-probability samples berechnen.
- Mit dem Modell von Bethlehem/Biffignandi (2012: 309–312) lässt sich beurteilen, inwieweit die Schätzung eines Mittelwerts durch einen Web-Survey von einem Populationsmittelwert \bar{Y} abweicht.²
- Da der Web-Survey eine Stichprobe aus derjenigen Population I darstellt, die das Internet nutzen bzw. nutzen können, entspricht der Bias der Differenz der beiden Populationen

$$B(\bar{y}_S) \approx \bar{Y} - \bar{Y}_I. \quad (1)$$

²Bethlehem, J./Biffignandi, S. (2012): Handbook of Web Surveys, Hoboken: Wiley.

Responsepropensities ρ

- Da nicht immer alle Personen, die Zugang zum Internet haben, an allen Web-Surveys teilnehmen, kann man vermuten, dass jede Person eine Wahrscheinlichkeit besitzt, an einer Befragung teilzunehmen.
- Diese Wahrscheinlichkeit wird Responsepropensity genannt und mit ρ bezeichnet.
- Der Mittelwert der Responsepropensities in der Population, die das Internet nutzt, entspricht $\bar{\rho}$.
- Die Standardabweichung der Responsepropensities ist S_{ρ} .

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

- Schließlich kann man sich einen Zusammenhang zwischen der Responsepropensity und der zu schätzenden Variablen Y vorstellen.
- Die Korrelation der Responsepropensity mit Y wird mit $R_{\rho Y}$ bezeichnet.
- Es lässt sich zeigen, dass der Bias geschätzt werden kann mit

$$B(\bar{y}_S) = \frac{R_{\rho Y} S_{\rho} S_Y}{\bar{\rho}}, \quad (2)$$

wobei S_Y die Standardabweichung von Y ist.

Also hängt der Bias von drei Größen ab:

- der Korrelation zwischen Responsepropensity und der zu schätzenden Variablen,
- der Streuung der Responsepropensities und
- der Streuung des interessierenden Merkmals.

Daher wird der Bias kleiner werden,

- wenn die Beteiligung am Web-Survey steigt oder
- wenn es keinen Zusammenhang mit der Responsepropensity und der interessierenden Variablen gibt oder
- die Streuung des interessierenden Merkmals klein ist.
- Schließlich wird der Bias umso größer werden, je größer die Streuung der Responsepropensities ist.

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Unter welchen Bedingungen ist das Ergebnis einer Nichtzufallsstichprobe dem Ergebnis einer Zufallsstichprobe ähnlich?

- Der Unterschied zwischen den Schätzungen hängt von vier Größen ab.
- Der Unterschied zwischen den Stichproben wird größer je:
 - (1) stärker es von Merkmal, das man schätzen will, abhängt ob jemand antwortet oder nicht,
 - (2) mehr sich die Wahrscheinlichkeit zu antworten zwischen den Personen unterscheidet,
 - (3) größer die Unterschiede zwischen den Personen bei dem zu schätzenden Merkmal sind,
 - (4) geringer die Wahrscheinlichkeit zu antworten ist.

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data
Mechanisms

Empirie

Demographie
Internetnutzer
Altersaufbau
JSSM 2020,1
Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Beispiel: Altersheim als Population 1/3

- Das lässt sich anhand eines Beispiels erläutern: Wir möchten das Alter der Bevölkerung schätzen.
 - Als Stichprobe nehmen wir die Bewohner eines Altersheims.
 - Wir nehmen an, in einem Altersheim antworten alle Personen, die wir fragen.
 - Außerhalb des Altersheims fragen wir niemanden, daher antwortet auch niemand außerhalb des Altersheims.

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data
Mechanisms

Empirie

Demographie
Internetnutzer
Altersaufbau
JSSM 2020,1
Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Beispiel: Altersheim als Population 2/3

- Junge Personen leben nicht im Altersheim, damit hängt die Wahrscheinlichkeit einer Teilnahme von dem Merkmal ab, das uns interessiert (1).
- Personen außerhalb des Altersheims haben keine Chance der Teilnahme, nur die Personen in einem Altersheim. Also gibt es deutliche Unterschiede in der Wahrscheinlichkeit der Teilnahme (2).
- Weiterhin unterscheiden sich Personen in der Bevölkerung deutlich im Alter (3).
- Schließlich ist nur ein kleiner Teil der Bevölkerung im Altersheim (4).

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Beispiel: Altersheim als Population 3/3

- Damit sind alle genannten Voraussetzungen für eine verzerrte Schätzung gegeben.
- Niemand käme auf die Idee, ein Altersheim zur Schätzung des Alters einer Bevölkerung zu verwenden.
- Obwohl die Situation bei Websurveys logisch identisch ist, werden diese von manchen Anwendern zur Schätzung der allgemeinen Bevölkerung verwendet.

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data
Mechanisms

Empirie

Demographie
Internetnutzer
Altersaufbau
JSSM 2020,1
Pre-election polling

Zusammenfassung

Vorschläge

Literatur

- Die Wahrscheinlichkeit, bei einem Websurvey zu antworten, falls man gefragt wird, ist klein (4).
- Die Wahrscheinlichkeit einer Teilnahme variiert zwischen Personen erheblich (z.B. nach Bildung und Alter) (2).
- Ob die Unterschiede zwischen Antwortenden und Nichtantwortenden klein sind (1), hängt vom Merkmal ab: Die Vorliebe für eine Farbe ist vermutlich von der Teilnahme an einer Befragung unabhängig. Ob dies für die Vorliebe für eine Partei gilt, hängt von vielen Umständen ab.
- Entsprechendes gilt für die Unterschiede zwischen den Personen in Hinsicht auf das zu schätzende Merkmal (3).

Der Bias kann vom Thema abhängen

- Die Vorliebe für Kuhspeichel als Getränk dürfte in jeder Bevölkerung konstant Null sein.
- Die Vorliebe für Milch hingegen schwankt stärker.
- Den Bedarf nach Kuhspeichel kann man mit einem Websurvey vorhersagen. Die Vorhersage für Milch ist schwieriger.
- Damit sind 2 Voraussetzungen für Verzerrungen bei Websurveys nahezu immer erfüllt: (4) und (2).
- Die beiden anderen Voraussetzungen (1, 3) müssen in jedem Einzelfall geprüft werden. Genau dies geschieht in der Praxis nicht.

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Missing Data Mechanisms³

- For example, suppose $K = 2$, $Y_1 = \text{age}$, and $Y_2 = \text{income}$.
- If the probability that income is missing is the same for all individuals, regardless of their age or income, then the data are **MCAR**.
- If the probability that income is missing varies according to the age of the respondent but does not vary according to the income of respondents with the same age, then the data are **MAR**.
- If the probability that income is recorded varies according to income for those with the same age, then the data are **MNAR**.

³Little,R./Rubin,D.B. (2002): Statistical Analysis with Missing Data, 2nd. ed. Hoboken:Wiley.

MD Summary by Donald Rumsfeld⁴

...as we know,
there are known knowns;
there are things we know we know
We also know there are known unknowns;
that is to say,
we know there are some things
we do not know.
But there are also unknown unknowns –
the ones we don't know we don't know.
12.2.2002, DoD newsbriefing)

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data
Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

⁴van Belle (2008): Statistical Rules of Thumb, 2nd ed., Hoboken

- Alle bisher verwendeten Verfahren setzen voraus, das wir nicht NMAR haben, auch:
 - Propensity weighting with calibration
 - Multilevel regression with poststratification
 - Bayesian Additive Regression Trees (BART)⁵
- Genau das ist aber das Problem.
- Daher versagen die Gewichtungungsverfahren häufig unvorhergesehen.

⁵Rafei, A./Flannagan, C.A.C./Elliott, M.R. (2020): Big Data for Finite Population Inference: Applying Quasi-Random Approaches to Naturalistic Driving Data Using Bayesian Additive Regression Trees, Journal of Survey Statistics and Methodology, 8, 1, 148-180.

Allgemeine Bevölkerung und Availability Bias

- Menschen leiden unter einem Availability Bias.
- In der Folge halten wir Personen mit anderen Merkmalen, die wir nicht sehen, für selten.
- Die objektive Größe ausgeschlossener Populationen⁶ wird oft unterschätzt.

⁶Schnell, R./Noack, M. (2015): Stichproben, Nonresponse und Gewichtung für Viktimisierungsstudien“. In: Viktimisierungsbefragungen in Deutschland, Band 2: Methodik und Methodologie, Bundeskriminalamt, Wiesbaden, S.8-75.

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

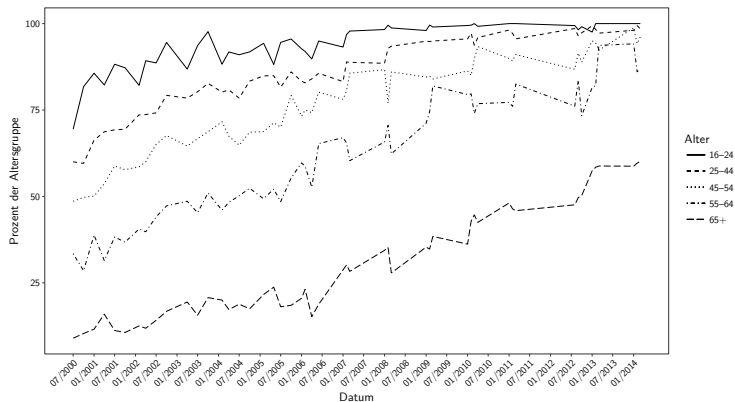
Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Prozent der Internetnutzer („ever used“)



UK National Omnibus-Survey 2000–2014

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data
Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020.1

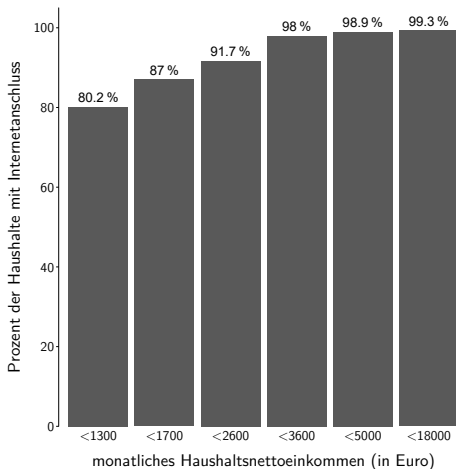
Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Haushalte mit Internet nach Einkommen



Datenbasis: Destatis, laufende Wirtschaftsrechnung 2017

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

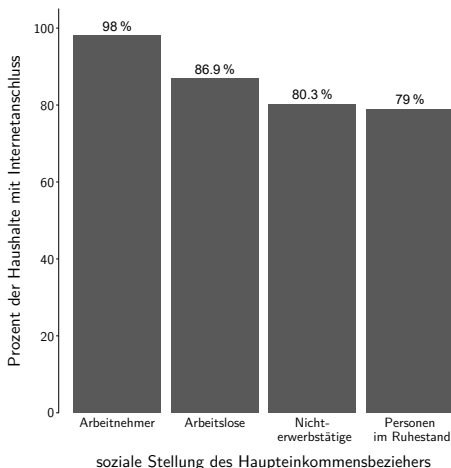
Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Haushalte mit Internet nach sozialer Stellung



Datenbasis: Destatis, laufende Wirtschaftsrechnung 2017

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

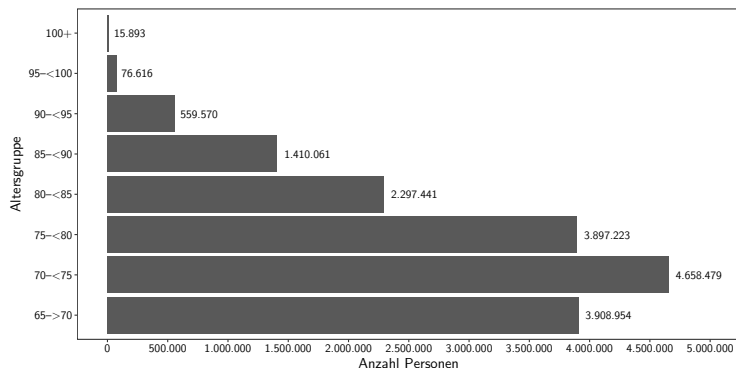
Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Altersgruppen in Deutschland (65+)



Datenbasis: Destatis, Sonderauswertung

- Fast 17 Millionen Personen (mehr als ein Fünftel der Bevölkerung) sind älter als 65.
- Die Wahlbeteiligung liegt bei den über 70-Jährigen bei 75.8% (Destatis: Wista 3/2018).

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

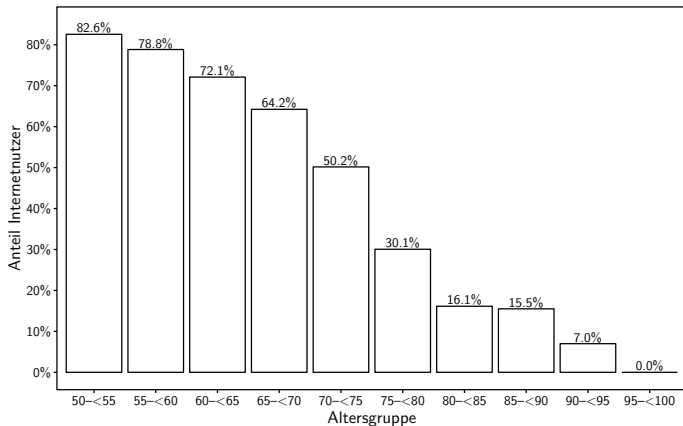
Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Internetnutzung 2015 in Deutschland (50+)



Datenbasis: SHARE-6

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzter

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

- Die empirische Leistungsfähigkeit von Non-probability samples ist wenig überzeugend, auch nach Gewichtung.⁷
- Das gilt auch für Gesundheitssurveys.⁸
- Noch nicht einmal die Art der Zusammenhänge bleibt immer erhalten.⁹

⁷Cornesse et al. (2020): A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research, *Journal of Survey Statistics and Methodology*, 8, 1, 4–36.

⁸Unangst, J.H. et al. (2020): A Process for Decomposing Total Survey Error in Probability and Nonprobability Surveys, *Journal of Survey Statistics and Methodology*, 8, 1, S. 62–88.

⁹Pasek, J./Krosnick, J.A. (2020): Relations Between Variables and Trends Over Time in RDD Telephone and Nonprobability Sample Internet Surveys, *Journal of Survey Statistics and Methodology*, 8, 1, S. 37–61

- Laien (und auch einige Lehrbücher der empirischen Sozialforschung) betrachten „pre-election polls“ als Test der Güte von Befragungen und/oder Gewichtungen.
- Die tatsächlichen Fehlerquoten dieser Art von Veröffentlichungen liegen in allen empirischen Studien immer weit über den falsch berechneten Konfidenzintervallen.¹⁰
- Das gilt auch für Civey.

¹⁰Schnell, R./Noack, M. (2014): The Accuracy of Pre-Election Polling of German General Elections, methods, data, analyses, |8(1), S. 5-24

Fehler bei der Landtagswahl Hessen 2018

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

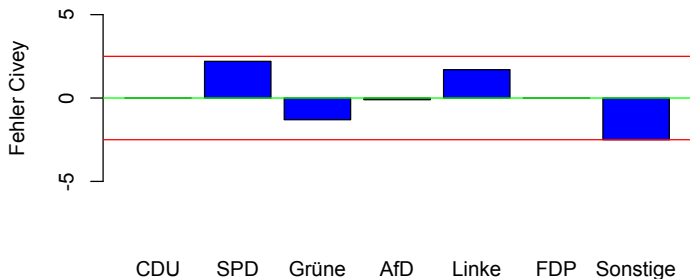
JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

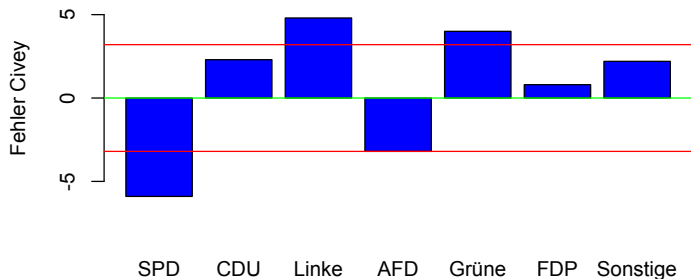
Literatur



— Angebener Fehler: 2.5%

Umfragedatum: 25.10.2018; Datum der Landtagswahl: 28.10.2018

Fehler bei der Landtagswahl Brandenburg 2019



– Statistischer Fehler: 3.2%

Umfragedatum: 28.08.2019; Datum der Landtagswahl: 01.09.2019

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Fehler bei der Landtagswahl Sachsen 2019

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

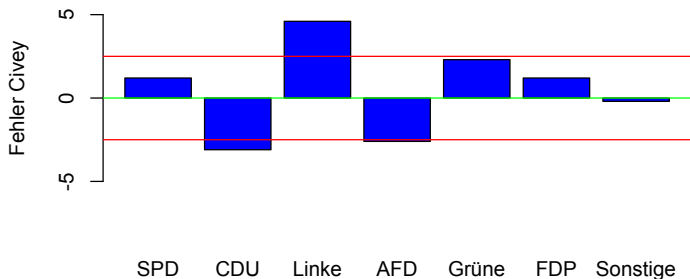
JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

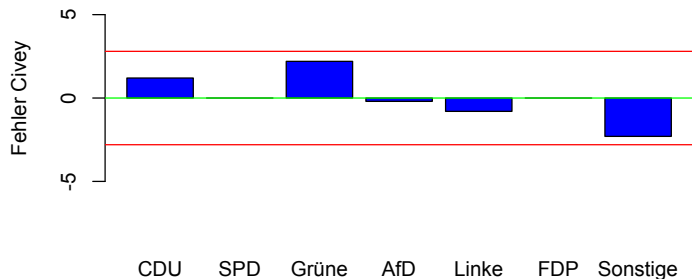
Literatur



— Angebener Fehler: 2.5%

Umfragedatum: 28.08.2019; Datum der Landtagswahl: 01.09.2019

Fehler bei der Landtagswahl Thüringen 2019



— Angebener Fehler: 2.8%

Umfragedatum: 23.10.2019; Datum der Landtagswahl: 27.10.2019

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Fehler bei der Bürgerschaftswahl Hamburg 2020

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

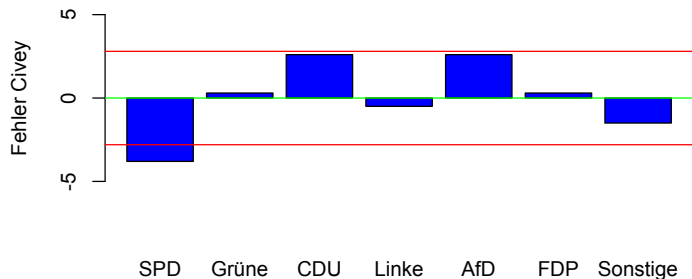
JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur



— Angebener Fehler: 2.8%

Umfragedatum: 19.02.2020; Datum der Bürgerschaftswahl: 23.02.2019

Die empirische Fehlerquote von Civey

No more SRS?

Rainer Schnell

A 95% confidence interval is often explained heuristically: If we take samples from our population over and over again, and construct a confidence interval using our procedure for each possible sample, we expect 95% of the resulting intervals to include the true value of the population parameter.

S.Lohr (2019): Sampling, 2nd ed., Boca Raton, S.41

- Betrachten wir die letzten 5 Landtagswahlen, (und ignorieren alle Probleme, z.B. Verwendung unabhängiger Binomialverteilungen), dann haben wir 35 Vorhersagen.
- Davon liegen 7 außerhalb der angegebenen Grenzen (wie auch immer diese ermittelt wurden).
- Das ist eine empirische Quote von **20%** falschen Vorhersagen.
- Die naive Interpretation ist noch naiver als man glaubt.

Geschichte

Theorie

Bias von Web-Surveys

Missing Data

Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Schlussfolgerungen

- Für Zufallsstichproben gibt es keinen Ersatz.
- Wissenschaftliche Anwendungen können auf Zufallsstichproben nicht verzichten.
- Laien vergessen oft, dass dies auch für extrem große Datenbestände gilt.
 - Big Data ist nicht die Lösung, sondern ein Problem.
 - Eine Instrumententheorie hierzu ist unverzichtbar.¹¹
- Es mangelt an der Ehrlichkeit, festzustellen, dass viele Anwendungen von Befragungen keine Anwendungen einer wissenschaftlichen Methode darstellen.
- Dazu gehört das Tagesgeschäft vieler ADM und Nicht-ADM-Institute.
- Wir brauchen eine klare Kennzeichnung dieser nichtwissenschaftlichen Unternehmungen.
- Will man freundlich sein, schlage ich Infotainment vor.

¹¹Amaya,A./Biemer,P.B./Kinyon,D. (2020): Total Error in a Big Data World, Journal of Survey Statistics and Methodology, 8,1,S. 89-119

- Wünschenswert ist eine Dreiteilung und Kennzeichnung von Befragungsformen:
 - ① wissenschaftliche Studien mit echten Zufallsstichproben (SOEP, PASS).
 - ② Markt- und Sozialforschung für die mittleren 80% der Bevölkerung (Konsum-, Werbe- und Einstellungsforschung).
 - ③ Infotainment auf der Basis willkürlicher Stichproben (ZDF-Politbarometer, Civey).
- Ohne zentrale Instanz wird der Markt dies nicht regeln.
- Wissenschaftler haben keine Regelungsmacht.
- „We have the power to blame“ (Sir Roger Jowell).

Weiterführende Literatur

Schnell, R. (1993). "Die Homogenität sozialer Kategorien als Voraussetzung für „Repräsentativität“ und Gewichtungsverfahren". In: *Zeitschrift für Soziologie* 22.1, S. 16–32.

Schnell, R. (1997). *Nonresponse in Bevölkerungsumfragen*. Opladen: Leske + Budrich.

Schnell, R. und M. Noack (2016). "Stichproben, Nonresponse und Gewichtung für Viktimisierungsstudien". In: *Viktimisierungsbefragung in Deutschland, Band 2: Methodik und Methodologie*. Hrsg. von N. Guzy, C. Birkel und R. Mischkowitz. Wiesbaden: Bundeskriminalamt, S. 8–75.

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data
Mechanisms

Empirie

Demographie
Internetnutzer
Altersaufbau
JSSM 2020,1
Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Weiterführende Literatur

Schnell, R., M. Noack und S. Torregroza (2017). “Differences in General Health of Internet Users and Non-users and Implications for the Use of Web Surveys”. In: *Survey Research Methods* 11.2, S. 105–123.

Schnell, R. (2018a). “„Big Data“ aus wissenschaftssoziologischer Sicht: Warum es kaum sozialwissenschaftliche Studien ohne Befragungen gibt”. In: *Erklärende Soziologie und soziale Praxis*. Hrsg. von D. Baron, O. Arránz und D. Lois. Wiesbaden: Springer, S. 101–125.

Schnell, R. (2018b). “Repräsentativität: Wenn der Scharfschütze sein Ziel selber malt”. In: *planung & analyse*.

No more SRS?

Rainer Schnell

Geschichte

Theorie

Bias von Web-Surveys

Missing Data
Mechanisms

Empirie

Demographie

Internetnutzer

Altersaufbau

JSSM 2020,1

Pre-election polling

Zusammenfassung

Vorschläge

Literatur

Schnell, R. (2018c). *Warum ausschließlich 'online' durchgeführte Bevölkerungsumfragen nicht 'repräsentativ' sind*. Technischer Bericht 11/2018. Research Methodology Group, University of Duisburg-Essen.

Schnell, R. (2019a). "Repräsentativität". In: *60 Notizen zur Zukunft*. Hrsg. von M. Smid. Bonn: Infas, S. 109.

Schnell, R. (2019b). *Survey-Interviews: Methoden der standardisierten Befragungen*. 2. Aufl. Wiesbaden: Springer.