

Alternative Selectionmechanisms in Online Samples

Franz Prücklmair*

Ulrich Rendtel†

June 13, 2022

Manuscript under construction

Do not cite

Contents

1	Online surveys und Ihre Probleme	3
1.1	Arten von Online Surveys	3
1.2	Probleme von Online Surveys	3
1.3	Selektionmechanisms	3
2	Korrektur verfahren	3
2.1	Quasi Randomisation	3
2.2	Superpopulation	3
2.3	Bayesian Approach	3
2.4	Machine Learning	3
3	Method	3
4	Results	3
5	Discussion	3

Abstract

*Universität Bamberg

†FB Wirtschaftswissenschaft, Freie Universität Berlin

Online-Surveys gehören aufgrund der relativ günstigen und schnellen Umsetzung zu den weitverbreitetsten Datenerhebungsmethoden. Allerdings weisen die daraus gewonnenen Daten, sogenannte Non-probability Samples einige Probleme auf. Dazu gehören mögliche Verzerrungen durch Selektivität, also des systematischen Ausschlusses von Personengruppen aus der Grundgesamtheit. Auch Erhebungsmethoden, wie z.B. River-Sampling, bei dem Befragte oftmals durch Implementation von Widgets auf Webseiten unkontrolliert rekrutiert werden, führen dazu, dass die für die Gewichtung benötigte inclusion probability des Befragten unbekannt ist. Somit ist die klassische design-basierende Inferenz wie im Probability Sample für Non-probability Samples unmöglich. Der Ansatz der Quasi-randomisation versucht durch Kombination eines Online-Samples mit einem Probability (Reference) Sample den Selektionsmechanismus zu modellieren und mithilfe von Pseudogewichtung die Verzerrung zu korrigieren. Ein häufig modellierter Selektionsmechanismus ist bislang der Internetzugang. Mit fortschreitendem Netzausbau ist jedoch davon auszugehen, dass der Ausschluss von Personen ohne Internetzugang immer mehr vernachlässigbar wird. Es könnte sich daher lohnen, den QR-Ansatz unter der Modellierung alternativer Selektionsmechanismen zu untersuchen. Für Analysen dient die 9. Runde des European Social-Survey (ESS) aus dem Jahr 2018 als Datengrundlage. Dieser übernimmt im QR-Ansatz die Rolle als klassisches Reference Sample und dient gleichzeitig, auf Basis alternativer Selektionsmechanismen wie der „Mindestdauer im Internet“ oder der „Bereitschaft, digitale Inhalte zu verbreiten“, als Quelle für die Simulation der Online-Samples. Die daraus resultierenden pseudo-gewichteten Ergebnisse zur Bundestagswahl 2017 sollen anhand der vom ESS geschätzten und dem tatsächlichen Wahlergebnis verglichen werden. Zudem sollen die Korrekturen mit klassischen Kalibrationsmethoden wie dem Raking gegenübergestellt werden. Erste Ergebnisse zeigen, dass der klassische Internetzugang nicht mehr selektiv ist. Dies spricht für die Verwendung alternativer Selektionsmechanismen, um den QR-Ansatz zu evaluieren. Die alternativen Modelle zeigen zudem, dass sich starke Verzerrungen in den Wahlergebnissen teilweise sehr gut korrigieren lassen. Gleichzeitig können durch den Ansatz auch vergleichbare Grenzen bei der Korrektur von Einstellungen und Meinungen deutlich gemacht werden.

Abstract: Keywords: Sampling frame, Access Panel, River Sampling, Non-probability samples

1 Online surveys und Ihre Probleme

Allgemein Haben Online Surveys als sogenannte Nonprobability samples Probleme Welche Wieso? usw.
Es gibt aber unterschiedliche Arten von Online Surveys

1.1 Arten von Online Surveys

Classical sampling and change to online mode Sampling via Mail Sampling from Accesspanel River
sampling via Widgets

Entscheidend hierbei ist vor allem der Selektionsmechanismus

2 Selektionmechanismen und Online Samples und Korrekturverfahren

2.1 Selektionsmechanismen

Internet Zugang nicht mehr aussagekräftig neuer Selektionsmechanismen entscheidend

2.2 Quasirandomisation

2.3 Wahlumfragen und ihre Aussagekraft

3 Method

3.1 Datenstz der ESS

3.2 Simulationaufbau

4 Results

ESS Vergleich SAR Abs Bias Parteipräferenz Rel Bias Partei präf Boxplots Gewichts betrachtung

5 Discussion

Netzzugang nicht mehr aussagekräftig Korrektur funktioniert bei etablierten Parteien sehr gut bei kleineren
weniger. Messungen von Einstellungen schwierig Gilt für alle Selektionsmechanismen Alternative Selektionsmechanismen Identifizieren