

Methoden-Newsletter vsms 2/2018

Vergleich der Datenqualität von Opt-in Panels und zufallsgesteuerten Umfragen

Opt-in Panels: Eine Alternative zu zufallsbasierten Umfragen?

Opt-in Panels sind eine kostengünstige und bequeme Alternative zu zufallsbasierten Online-Befragungen. Allerdings bilden die Stichprobenmitglieder solcher Panels, die sich oft selbst in die Stichprobe rekrutieren, substantielle Variablen der Zielpopulation nicht gut ab. In diesem Artikel untersuchen wir, ob Opt-in Panels eine Alternative sind, wenn es darum geht, die Beziehung zwischen Variablen realitätsgetreu abzubilden und ob gegebenenfalls eine Gewichtung hilft, Unterschiede auszugleichen.

Bei Umfragen in den Sozialwissenschaften konnte man in letzter Zeit einen Trend weg von persönlichen Befragungen oder Befragungen per Telefon hin zu relativ kostengünstigen Online-Befragungen feststellen. Online-Umfragen, die auf zufallsbasierten Stichproben basieren, bringen allerdings zahlreiche Probleme mit sich, die nicht nur einen Teil der finanziellen Vorteile aufwiegen, sondern auch das Risiko von Erhebungsfehlern erhöhen.

Diese Risiken beziehen sich insbesondere auf Fragen zu Stichprobenauswahl und Erfassungsbereich (coverage), Kontaktierungs- und Rekrutierungsverfahren, und relativ geringe Rücklaufquoten bei weiterhin beträchtlichen Kosten. Kostengünstig und bequem lassen sich viele dieser Probleme vermeintlich durch die Datenerfassung unter Einsatz von Access Panel Anbietern auf der Basis von Non-Probability (oder Opt-in) Panels lösen. Der grösste Nachteil dieses Verfahrens besteht darin, dass trotz des Einsatzes von Quoten und modellbasierter Gewichtungsstrategien zur Erzielung scheinbar repräsentativer Stichproben, die freiwilligen Teilnehmer an solchen Panels die Gesamtbevölkerung nicht wirklich repräsentativ abbilden. Ausserdem kann eine regelmässige Teilnahme an Panelbefragungen dazu führen, dass die Teilnehmer anders antworten als weniger erfahrene Teilnehmer. Oft sind bei Opt-in Panels Personen überrepräsentiert, die einen geringen Bildungsstand aufweisen, jung und arbeitslos sind, das Internet exzessiv nutzen und ihre Meinung gerne äussern. Ausserdem ist die fehlende Transparenz und Konsistenz der Art und Weise, wie Opt-in Panels betrieben werden, problematisch. Das gilt besonders für wissenschaftliche Befragungen, die ja reproduzierbar sein müssen. Auch wenn diese Panels meist präzise Ergebnisse liefern, weichen

diese plötzlich von den früheren Ergebnissen ohne erkennbare Gründe ab. Zufallsbasierte Befragungen liefern dagegen wesentlich konsistentere Ergebnisse; ihre Methodik ist transparent und kann später reproduziert werden.

Es ist allgemeiner Konsens, dass Opt-in Panels nicht geeignet sind, präzise Schätzer der allgemeinen Bevölkerung zu liefern (z.B. Anteil der Bevölkerung, der Meinung X hat oder Merkmal Y aufweist). Erfolgsversprechender scheinen Opt-in Panels zu sein, wenn es darum geht, Beziehungen zwischen Variablen zu erklären, was oft im Fokus der sozialwissenschaftlichen Forschung steht. Um besser zu verstehen, inwieweit sich Opt-in Panels als Alternative zu zufallsbasierten Befragungen eignen - insbesondere in der Schweiz - hat FORS die Datenqualität von zwei Referenzbefragungen (aus der Schweizer Wahlstudie - Projekt Selects 2015) und drei Opt-in Panels miteinander verglichen.

Wie vergleichen wir die Befragungen miteinander?

Eine Wahlumfrage wie Selects verfolgt zwei Hauptziele: erstens die Beschreibung der Merkmale von Wählern von verschiedenen Parteien sowie Nichtwählern und zweitens die statistische Modellierung für die wissenschaftliche Forschung. Deshalb sind Opt-in Panels dann eine sinnvolle Alternative, wenn sich die Beziehungen zwischen Variablen in Opt-in Panels und zufallsbasierten Befragungen nicht signifikant unterscheiden bzw. wenn sich dieser Unterschied mittels Gewichtung ausgleichen lässt.

Um festzustellen, ob dies der Fall ist, analysierten wir die bedingten Verteilungen von Variablenpaaren (z.B.



Altersgruppen innerhalb der beiden Geschlechterkategorien oder Einkommensgruppen innerhalb der Kategorien Wähler und Nichtwähler). Insgesamt verglichen wir drei Gruppen bedingter Verteilungen von Variablenpaaren 1.) die soziodemographischen Variablen Geschlecht, Alter, Haushaltsgrösse und Einkommen 2.) diese soziodemographischen Variablen kombiniert mit politischer Partizipation und 3.) diese soziodemographischen Variablen kombiniert mit der Stimme für die grösste Partei (SVP). Wir verwendeten Daten basierend auf dem Personenstandsregister als Benchmark, um die soziodemographische Repräsentativität der erhobenen Stichproben in den verschiedenen Befragungen zu beurteilen. Wir verwenden zur Beurteilung die nach der Wahl erfolgte Selects-Befragung, wenn politische Variablen berücksichtigt werden.

Ergebnisse

Unsere eindimensionalen Verteilungen bestätigen, dass die Ableitung von Bevölkerungsschätzern aus Opt-in Panels hoch problematisch ist. Das Verhältnis zwischen denjenigen, die zur Wahl gehen und Wählern verschiedener Parteien unterscheidet sich von einem Opt-in Panel zum anderen gewaltig und kann nicht durch die unterschiedliche soziodemographische Charakteristika der Stichproben einschliesslich Einkommen und Bildung erklärt werden, die stark mit den relevanten Variablen korrelieren.

Hinsichtlich der Abweichung der bedingten soziodemographischen Variablen (z.B. Repräsentation der Altersgruppe nach Geschlecht) stellen wir bei den nicht-zufallsbasierten Befragungen einen grösseren Gesamtbias fest als bei den zufallsbasierten Befragungen. Besonders auffallend ist der extrem unterschiedliche Bias bei Opt-in Panels in Abhängigkeit von den untersuchten Variablenpaaren, während sich bei den beiden zufallsbasierten Befragungen viel konsistentere Muster ergeben. Die Ergebnisse folgen dem gleichen Muster, wenn wir Wahlbeteiligung und gewählte Partei anhand soziodemographischer Variablen vergleichen.

Zuletzt wendeten wir die inverse Propensity-Gewichtung an, um die bedingten Verteilungen in den Opt-in Panels zu verbessern. Grundlage dieser Gewichtungen war die registerbasierte Stichprobe; die Berechnung erfolgte anhand von Alter, Haushaltsgrösse, Geschlecht, Sprachregion und Einkommen. Doch auch mit diesem Verfahren verbesserte sich die bedingte Repräsentativität nur marginal.

Schlussfolgerung

Unsere Untersuchung bestätigt das bereits weithin bekannte Ergebnis, dass Opt-in Panels für präzise Bevölkerungsschätzer nicht geeignet sind. Die Ergebnisse variieren extrem - sowohl innerhalb eines Panelbetreibers, als auch unter den verschiedenen Panelbetreibern. Ausserdem lassen sich diese Probleme durch Gewichtung der Daten nicht zufriedenstellend lösen. Nach unseren Analysen sind Opt-in Panels zudem unzuverlässige Quellen, wenn es darum geht, wie verschiedene Gruppen über etwas denken oder ein bestimmtes Verhalten ausdrücken, z.B. für welche Partei sie stimmen würden. Unsere Analysen zur bedingten Verteilung zeigen, wie Beziehungen zwischen Variablen durch verschiedene Erhebungsmethoden beeinflusst werden und wie sich die Gewichtung auf diese Beziehungen auswirkt. Die Unterschiede in den Beziehungen zwischen Variablen wirken sich wiederum auf statistische Modelle und möglicherweise auch auf die daraus gezogenen Schlussfolgerungen aus.

Da sich dieser Bereich rasch entwickelt, sehen sich Betreiber von Opt-in Panels zunehmender Konkurrenz ausgesetzt, und es wird möglicherweise immer schwerer, Teilnehmer zu gewinnen. Ausserdem sind die von den Betreibern angebotenen Anreize oft nicht so attraktiv, dass sich eine Teilnahme lohnt. Deshalb dürften die Probleme, vor denen traditionelle Befragungen stehen, genauso auch Opt-in Panels betreffen. Insbesondere online werden die Menschen zunehmend mit Informationen und Aufforderungen zur Beteiligung bei Umfragen bombardiert. Es dürfte immer schwieriger werden, potenzielle Teilnehmer davon zu überzeugen, sich freiwillig dieser Zusatzbelastung auszusetzen. Unserer Ansicht nach liefern zufallsbasierte Stichproben nach wie vor eine bedeutend höhere Datenqualität für Bevölkerungserhebungen, und wir gehen nicht davon aus, dass sie in naher Zukunft in nennenswertem Umfang von Opt-in Panels abgelöst werden.

Text:

Nicolas Pekari, FORS

Oliver Lipps, FORS und Universität Bern

Caroline Roberts, Universität Lausanne

Georg Lutz, FORS

Tel. +41 21 692 37 65, nicolas.pekari@fors.unil.ch