

Methodenpapier eBUS

Maßnahmen zur Erreichung einer verbesserten Repräsentativität im GfK eBUS

1. Ausgangssituation

Der digitale Wandel vollzieht sich seit sich einigen Jahren auf allen Gesellschaftsebenen. Die technologischen Entwicklungen sind rasant und verändern das Kauf- und Kommunikationsverhalten, die Mediennutzung als auch die Art und Weise, wie Informationen gesucht werden.

Die Marktforschungsbranche hat sich dieser Entwicklung angepasst und Interviewmethoden, die früher „der“ Standard waren, wie Face-to-Face- oder Telefoninterviews, werden zunehmend durch Online-Befragungen ersetzt. Das bietet wesentliche Vorteile für Marktforschungskunden, Institute, aber auch die Befragten. Um nur einige Vorteile zu nennen, sind Online-Befragungen schneller, flexibler und kostengünstiger durchzuführen. Bestimmte Zielgruppen, vor allem jüngere Personen, können Online besser erreicht werden als mit anderen Interviewmethoden. Zudem haben die Befragten es selber in der Hand, wann, wo und mit welchem Gerät sie an einer Befragung teilnehmen.

Jedoch sind Online-Befragungen per se nicht automatisch besser. Hier gilt es einige Qualitätskriterien zu berücksichtigen, damit ein hohes Maß an Validität und Reliabilität sowie Repräsentativität in den Befragungsergebnissen erzielt werden kann. Insbesondere der von der GfK entwickelte Maßnahmenkatalog zur Erreichung einer möglichst hohen Bevölkerungsrepräsentativität in Online-Befragungen durch den GfK eBUS soll im Folgenden näher erläutert werden.

2. Wann ist eine Stichprobe repräsentativ?

Laut Definition ist eine Stichprobe repräsentativ, wenn ihre Ergebnisse ohne systematischen Fehler auf die Grundgesamtheit hochgerechnet werden können. Dahinter steht die statistisch formulierte Anforderung nach erwartungstreuen Schätzern, konkret heißt das z.B., dass der erhobene durchschnittliche Kraftstoffverbrauch eines Haushalts in der Stichprobe nicht systematisch vom wahren Wert in der Grundgesamtheit abweicht.

Folgende Kriterien werden im Allgemeinen zur Beurteilung der Repräsentativität der Stichprobe herangezogen:

- Rekrutierung und Auswahlverfahren für die Stichprobe

Idealerweise hat jedes Individuum aus der Grundgesamtheit eine berechenbare Chance größer Null in die Stichprobe zu gelangen.

Für Online-Befragungen stellt diese Anforderung an Repräsentativität eine besondere Herausforderung dar. Zum einen ist es das Medium Internet selbst, welches eine Einschränkung der geforderten Auswahlwahrscheinlichkeit erzeugt: Befragte ohne Internetzugang sind von der Teilnahme an Online Befragungen per se ausgeschlossen. Mit zunehmender Internetpenetration in Deutschland verringert sich dieser Effekt von Jahr zu Jahr (*Pressemeldung des Stat. Bundesamts vom September 2018: Internetpenetration in Deutschland inzwischen bei 90 %), aber derzeit darf dieses Thema bei Online-Befragungen nicht außer Acht gelassen werden, insbesondere wenn es um die Verallgemeinerbarkeit der Ergebnisse auf die Gesamtbevölkerung geht.

Eine weitere Herausforderung ist der Selbstselektionseffekt, d.h. der Befragte und nicht das Institut entscheidet, ob er an einer Befragung teilnimmt oder nicht. Dies hat zur Folge, dass es zu einer Verzerrung in Form einer Überrepräsentation gewisser Personengruppen kommen und die Stichprobe damit nicht mehr als repräsentativ bezeichnet werden kann.

Als dritter Punkt ist das Non-Response Problem zu nennen, da wegen Teilnahmeverweigerung bestimmter Befragten Gruppen ausfallende Antworten mit Merkmalen korrelieren können, die Gegenstand der Untersuchung sind.

Die GfK nutzt für die eBUS-Befragung ein eigenes Online Access Panel (Nicequest by GfK, früher: askGfK), um o.g. Themen besser behandeln zu können. So tritt der Selbstselektionseffekt nur beim Eintritt in das Panel auf, nicht aber bei der eigentlichen Befragung. Die Befragten werden aus dem Nicequest by GfK Panel rekrutiert, welches über 40.000 aktive Teilnehmer (gemäß der Esomar Definition für aktive Panelisten) verfügt. Rekrutiert werden die Panelteilnehmer sowohl Online als auch Offline. Die Panelteilnehmer bekommen eine Einladung, die mit einem Link zum Nicequest by GfK-Portal versehen ist. Die Anfragen der potentiellen Panelkandidaten durchlaufen das Double-Opt-in Vorgehen sowie weitere Schritte zur Identifikationsprüfung.

Mit der Offline Rekrutierung soll u.a. gewährleistet werden, dass auch weniger online-affine Bevölkerungsgruppen im Panel vertreten sind.

Die Stichprobe für den GfK eBUS wird anhand des Quota-Verfahrens aus dem Pool der Panelisten gezogen. Neben weiteren qualitätssichernden Maßnahmen des Panelmanagements wird automatisiert sichergestellt, dass Panelisten nicht zu häufig für Befragungen in einer bestimmten Periode gezogen werden.

Für den GfK eBUS gilt, dass Panelisten nach frühestens vier Monaten wieder dafür eingeladen werden.

- Stichprobenstruktur

Die Stichprobe muss in ihrer Zusammensetzung und in der Struktur der relevanten Merkmale der Grundgesamtheit möglichst ähnlich sein. Strukturmerkmale sind üblicherweise soziodemographische Variablen wie Geschlecht, Alter, Region, Haushalts- und Ortsgröße etc. Gerade in Bezug auf Online Panels werden häufig Bedenken geäußert, dass die Online Population anders strukturiert sei als die Gesamtbevölkerung. Untersuchungen haben ergeben, dass sich unter allen Internetnutzern im Vergleich zur Allgemeinbevölkerung proportional mehr jüngere, einkommensstärkere, besser gebildete, sowie mehr Männer als Frauen befinden.

Daher werden beim GfK eBUS folgende Maßnahmen ergriffen, um die systematischen Verzerrungen durch das Fehlen der Offliner zu verringern:

- Quotenplan mit soziodemographischen Strukturmerkmalen (Geschlecht, Alter, Region, Haushalts- und Ortsgröße sowie Schulbildung des Haupteinkommensbeziehers) für deutschsprachige Personen im Alter von 18 -74 Jahren.
Für die Sollvorgaben der soziodemographischen Merkmale stehen amtliche Statistiken von z.B. des Statistischen Bundesamts (Mikrozensus) oder der AGMA (Media-Analyse) zur Verfügung, um die Merkmalsverteilungen der Grundgesamtheit mit den Verteilungen in der Stichprobe abzugleichen. Diese Informationen werden für die Quotenaussteuerung in den eBUS-Befragungen genutzt und regelmäßig aktualisiert.

- Um die angenommene Schiefelage beim Bildungsniveau auszugleichen, wird beim GfK eBUS zusätzlich zu den ansonsten üblichen Quotierungsvariablen die Bildung des Haupteinkommensbeziehers quotiert und gewichtet. Gerade diese Variable korreliert sehr stark mit der gesellschaftlichen Stellung eines Haushalts bzw. einer Person und ist dadurch für die Verbesserung der Repräsentativität der Stichprobe entscheidend.

- Gewichtung

Da in der Regel nur wenige Merkmale in einer Befragung quotiert werden können, ist die Gewichtung ein weiterer wichtiger Schritt, um Schiefen in nicht quotierten Merkmalen auszugleichen und die Repräsentativität der Ergebnisse zu erhöhen. Für eine Gewichtung müssen, wie auch für die Quotenvorgaben zur Stichprobenziehung, entsprechende Informationen über die Merkmalsverteilungen in der Grundgesamtheit vorliegen.

Die GfK hat ein spezielles Gewichtungsverfahren auf der Basis eines mathematisch-analytischen Ansatzes entwickelt, welches die Berücksichtigung vieler Gewichtungsvariablen erlaubt. Neben diskreten Merkmalen wie Alterskategorien und Geschlecht können zudem auch metrische Variablen verarbeitet werden. Da die Anzahl an Gewichtungsvorgaben schnell zunehmen kann, hat sich diese Gewichtungsverfahren für die Aussteuerung systematischer Schiefen im Sample als besonders vorteilhaft erwiesen. Die GfK berücksichtigt bei der Gewichtung der Stichprobe des eBUS insbesondere die Merkmalskombination Geschlecht x Alter neben den Randverteilungen für die Merkmale Geschlecht, Haushaltsgröße, Schulbildung des Haushaltsvorstands, Ortsgröße und Region. Die entsprechenden Gewichtungsvorgaben werden gezielt im GfK eBUS eingesetzt, um die Repräsentativität der Studienergebnisse zu erhöhen.

3. Qualitätschecks

Alle Anstrengungen, die Repräsentativität der Befragten sicherzustellen, laufen natürlich ins Leere, wenn nicht zusätzlich intensiv die Qualität der gegebenen Antworten geprüft würde. Dies soll sicherstellen, dass zum einen eine reale Person den Fragebogen ausgefüllt hat und zum anderen auch das Antwortverhalten dieser Person valide ist. Für die Kontrollen werden sowohl GfK-Eigenentwicklungen verwendet als auch am Markt gängige Tools der Anbieter ATLAS oder Orion.

- Für ersteres Kriterium wird durch ein GfK eigenes Captcha-Tool verhindert, dass Bots automatisiert Fragebögen befüllen. Außerdem wird sichergestellt, dass eine IP-Adresse nicht zweimal die gleiche Befragung beantworten darf. Somit kann eine Person, ggf. unter Nutzung unterschiedlicher soziodemografischer Merkmale, nicht fälschlicherweise mehrfach den Fragebogen beantworten. Darüber hinaus wird auch noch sichergestellt, dass die IP-Adresse aus dem Land stammen muss, in dem die Studie tatsächlich durchgeführt wird („Geofencing“).
- Für zweites Kriterium gilt die Annahme, dass das Antwortverhalten einer Person als valide anzusehen ist, wenn die zur Beantwortung benötigte Zeit sinnvoll lang ist und wenn es eine Varianz im Antwortverhalten gibt. Dabei wird die Dauer zur Beantwortung sowohl auf der Ebene der Einzelfrage als auch auf der Ebene des Gesamtfragebogens mit der durchschnittlichen Beantwortungszeit verglichen, um sogenannte „Speeder“ zu identifizieren, die dann entfernt werden. Darüber hinaus werden Befragte, die bei den Antworten zu einzelnen Frageblöcken keine oder zu wenig Varianz aufweisen, als „Straightliner“ markiert und vom Datensatz gelöscht. Zur Festlegung der Antwortzeit greifen wir auf GfK Erfahrung aus zahllosen Studien zurück.

Auch wenn es natürlich selbstverständlich ist, ist noch hinzuzufügen, dass alle von der GfK angewandten Checks DSGVO-konform sind und regelmäßig erweitert und ergänzt werden, um mit der technologischen Entwicklung Schritt zu halten bzw. ihr zuvorzukommen.

4. Untersuchung auf Unterschiede in Befragungsergebnissen auf Basis von Trackingstudien

Da ein Umstieg von Befragungsmethoden wie z.B. CAPI auf CAWI grundsätzlich unterschiedliche Befragungsergebnisse erwarten lässt, haben wir untersucht, wie stark der Effekt der Befragungsmethode bei langjährigen Trackingstudien ausfällt.

Die Untersuchung basiert auf der Erhebung von Trackingdaten mit dem CAPI-Bus (von 2013 bis Ende 2018), ergänzt mit jeweils neuen Datenpunkten, die im eBUS in 2019 erhoben wurden. Im Fokus der Untersuchung stand, ob die Ergebnisse aus dem CAPI-Bus konsistent weitergeführt werden können oder sich deutliche Brüche in den Zeitreihen ergeben.

Insgesamt standen 97 Zeitreihen zur Verfügung. Darin enthalten waren verschiedene Aspekte, wie Image-Items, gestützte und ungestützte Bekanntheit, Käuferreichweiten, Wiederkaufverhalten etc.

Mittels linearer Trendmodelle pro Ausprägung konnte gezeigt werden, dass lediglich 10% der gemessenen Werte im eBUS nicht im Vorhersageintervall der CAPI-Bus Zeitreihen lagen und damit die Zeitreihen in 90% der Fälle konsistent weitergeführt werden konnten - trotz des Umstiegs auf den eBUS.

5. Fazit

Anhand der vorliegenden Ergebnisse konnte gezeigt werden, dass die Ergebnisse von GfK-Online-Befragungen repräsentativ für Markenwahrnehmung, Einstellungen, Konsumverhalten und ähnlicher Ausprägungen der Konsumenten sind.

Bestimmte Effekte, die durch die Interviewmethode bedingt sind, wie z.B. der Einfluss des Interviewers auf den Befragten können weder durch die Aussteuerung oder Auswahl der Stichprobe noch durch die Gewichtung ausgeglichen werden. Daher empfiehlt sich bei wiederkehrenden Studien, die zuvor mit Offline Interviewmethoden erhoben wurden und jetzt auf CAWI umgestellt werden, die Durchführung eines Methodentests, um potenzielle Effekte von systematischen Verzerrungen besser quantifizieren und ggfs. minimieren zu können.

Auf Basis der durchgeführten Analysen kann festgehalten werden, dass eine Kombination sinnvoller Maßnahmen, wie sie für den GfK eBUS beschrieben und angewandt werden, die beste Strategie ist, um repräsentativere und qualitativ hochwertig Ergebnisse aus Online-Befragungen zu erhalten.