

Rezension: Sigrid Haunberger, 2011: Teilnahmeverweigerung in Panelstudien

Lipps, Oliver

Veröffentlichungsversion / Published Version

Rezension / review

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Lipps, O. (2012). Rezension: Sigrid Haunberger, 2011: Teilnahmeverweigerung in Panelstudien. [Rezension des Buches *Teilnahmeverweigerung in Panelstudien*, von S. Haunberger]. *Methoden, Daten, Analysen (mda)*, 6(1), 49-53. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-314523>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see: <https://creativecommons.org/licenses/by/4.0>

lungen auch Ansätze zur „Bestimmung des Standorts und des weiteren Kurses“ (S. V). In Bezug auf diese, im Geleitwort von Wilfried Seidel genannte Zielsetzung wären eventuell eine europäische oder internationale Situationsbetrachtung oder Vergleiche mit anderen statistischen Fachgesellschaften ein eigenes Kapitel wert gewesen. Aufgrund der historischen und fachlichen Synopsen ist der Band aber nicht nur „ein Buch für die Freunde der Statistik und der Deutschen Statistischen Gesellschaft“ (S. IX), sondern auch für den Einsatz in der universitären Lehre geeignet, z. B. in Einführungen in die Wirtschafts- und Sozialstatistik, und nicht zuletzt für alle Anwender und Nutzer von statistischen Verfahren und Daten interessant.

BERNHARD SCHIMPL-NEIMANNS, MANNHEIM

* * * * *



SIGRID HAUNBERGER,
2011: Teilnahme-
verweigerung
in Panelstudien.
VS-Verlag. ISBN:
978-3-531-17710-6,
254 Seiten, 39,95
EUR.

Nicht (Nonresponse) bzw. nicht mehr (Attrition) befragte Mitglieder von Zufallsstichproben unterscheiden sich in aller Regel systematisch von den Teilnehmern. Nonresponse und Attrition sind wichtige, wenn nicht sogar die wichtigsten Gründe für die Verzerrung von Stichprobenparametern und statistischen Zusammenhängen. Zudem ver-

ringern sie den Umfang der realisierten Stichprobe und bedingen daher größere Stichprobenfehler. Last but not least ist sich die mittlerweile recht umfangreiche Literatur im Bereich der Nonresponseforschung in der Einschätzung einig, dass Nonresponse tendenziell zunimmt.

Die Einhaltung gewisser Grenzen von Nonresponse stellt daher immer größere Anforderungen an Erhebungsdesign und -ressourcen. Dessen ungeachtet ist – wie die Autorin des vorliegenden Buchs treffend feststellt – „bezüglich der Erklärung systematischer Ausfälle von Teilnehmern kaum ein Fortschritt zu verzeichnen“ (S. 18). Dies liegt daran, dass nach wie vor meist ausschließlich sozio-demografische Variablen wie Alter und Geschlecht nicht nur zur Beschreibung, sondern auch zur Erklärung von Nonresponse verwendet werden. Der simple Grund ist, dass – wenn überhaupt – von Nichtteilnehmern nur solche Variablen bekannt sind; noch häufiger lediglich Randverteilungen. Im Sinne einer Erklärung der Prozesse und Mechanismen, die zur Teilnahme oder Nichtteilnahme führen, sind solche Variablen aber wenig geeignet.

Sigrid Haunbergers im VS Verlag veröffentlichte Dissertation ist ein äußerst lobenswertes Plädoyer, sich für die Erklärung von Nonresponse und Attrition von einer solchen „Variablensoziologie“ zu lösen. Die Autorin schlägt stattdessen zwei handlungstheoretische Theorien vor: die Theory of planned behavior (TOPB) und die Theory of subjective expected utility (SEU). Sie motiviert, operationalisiert und testet diese empirisch anhand einer Stichprobe von Studierenden mit Hilfe einer Onlineerhebung. Insofern dürfte sich das Buch vor allem an Erhebungsexperten aus Theorie und Praxis richten, die Interesse an der Erklärung von Nonresponse und Attrition haben. Speziell eignet sich dieses Buch sehr gut als Grundlage für weitergehende empirische Forschungsarbeiten oder Experimente, Mechanismen für Nonresponse auf der Grundlage der in

dieser Arbeit verwendeten handlungstheoretischen Theorien aufzudecken.

Die einzelnen Kapitel bauen sinnvoll aufeinander auf; das Buch liest sich leicht, was auch einigen Wiederholungen von komplexeren Sachverhalten zu verdanken ist. Die Arbeit startet in Kapitel A mit einer sehr guten Dokumentation der zunehmenden Relevanz, Komplexität, und Popularität des Themas Nonresponse. Die Autorin beanstandet, dass sich in der umfangreichen Literatur nur wenige Autoren finden, die Nonresponse nicht mittels soziodemografischen (und anderen gerade zur Verfügung stehenden Korrelaten) erklären, sondern handlungstheoretisch fundierte Modelle verwenden. Aber auch falls dies der Fall ist, ist nach Ansicht der Autorin die Operationalisierung der Theorie unzureichend. Die geeignete Operationalisierung der beiden Theorien stellt – neben dem empirischen Test – in der vorliegenden Arbeit die Hauptzielsetzung dar. Kapitel B bespricht Panelstudien mit ihren analytischen Vorteilen, erhebungsspezifischen Problemen und der Vielschichtigkeit der Ausprägungen von Nonresponse. Insbesondere können bei Panelerhebungen Stichprobenmitglieder nicht nur in der ersten Befragungswelle ausfallen, sondern auch erst nach einer gewissen Teilnahmedauer endgültig oder temporär nicht (mehr) teilnehmen (Attrition). Kapitel C diskutiert den aktuellen Forschungsstand zu Nonresponse, wobei zwischen Nichterreichbarkeit und Verweigerung differenziert wird. Diese Differenzierung ist wichtig, da Motivation und Hintergrund für diese beiden Ursachen verschieden sind. Die Autorin kritisiert, dass die verwendeten empirischen Modelle zur Erklärung von Nichterreichbarkeit und Verweigerung nur selten theoriegeleitet sind und dementsprechend besonders bei befragtenbezogenen Faktoren heterogene Ergebnisse produzieren. Die „Theorien“ zur Erklärung der Wirkungsmechanismen von Nonresponse werden erläutert. Bemängelt wird, dass sie meist bei Einzelaspek-

ten ansetzen (etwa Incentives) und oft die Handlungsdispositionen der Stichprobenmitglieder ignorieren. Die Autorin stellt fest, dass Gründe für Verweigerungen oft situativ sind. Dies vor allem aufgrund der Tatsache, dass „eine Teilnahmeentscheidung ... aus einem Zustand der Indifferenz heraus erfolgt“ (S. 96). Entsprechend müsste eine handlungstheoretisch bedeutsame Erklärung die subjektive Situation als relevante Variable operationalisieren. In Kapitel D werden die beiden handlungstheoretischen Modelle des Teilnahmeverhaltens vorgestellt: Die SEU enthält die Bewertungen von und Erwartungen an möglichst alle zur Verfügung stehenden Handlungsalternativen als handlungserklärende Variablen; die TOPB ‚verhaltensbezogene Einstellung‘, ‚subjektive Norm‘, und ‚wahrgenommene Verhaltenskontrolle‘. Wichtig sind die Prämissen, dass Subjekte nicht voll rational/informiert handeln, und dass das Handeln aus einer bestimmten Situation heraus erklärt werden muss. Zusätzlich spielen Routinen und in ähnlichen Situationen bewährte Handlungsweisen eine Rolle. Die Autorin erläutert die Operationalisierung der beiden Theorien zu quantitativen Modellen. Kapitel E beschreibt den empirischen Test der beiden handlungstheoretischen Modelle. Die Stichprobe besteht aus Studierenden der Universität Bern aus verschiedenen, meist sozialwissenschaftlichen Vorlesungen. Diese werden zunächst in einer Paper & Pencil „Null“messung über ihre Basischarakteristika sowie – in zwei Gruppen randomisiert – ihre handlungstheoretisch relevanten Teilnahmedispositionen bezüglich der Komponenten der SEU und der TOPB befragt. Zudem werden die E-Mail Adressen erfragt, wobei hier (Angabe oder eben Nichtangabe der E-Mail Adresse) eine erste Selektion erfolgt. Einige Zeit danach werden die Befragten per E-Mail gebeten, bei einer zwei Wellen umfassenden Online-Erhebung über Drogen teilzunehmen. Hierbei erfolgt die zweite und dritte Möglichkeit der Nonresponse. In Kapitel F werden schließlich die Ergebnisse interpretiert und Schluss-

folgerungen gezogen. Während mit beiden Theorien die Teilnahmeintention relativ gut erklärt wird, fällt die Erklärungskraft für das tatsächliche Verhalten deutlich geringer aus. Bei der TOPB sind es vor allem modellexogene Faktoren, wie die subjektive Entscheidungssicherheit und vergangene Teilnahmehäufigkeit, bei der SEU Kostenbefürchtungen und Situationsmerkmale, die das tatsächliche Teilnahmeverhalten erklären. Die Autorin folgert, dass die Theorien trotz ihrer geringen Erklärungskraft zur Erklärung des Teilnahmeverhaltens geeignet sind. Sie schließt mit einer Kritik der Resultate und der Diskussion weiterer Forschungsmöglichkeiten.

Die besonderen Stärken der Arbeit liegen in der gut begründeten Kritik der konventionellen Modelle zur Erklärung des Teilnahmeverhaltens und der Motivation und Operationalisierung der beiden Theorien SEU und TOPB. Was ebenfalls bei einer Dissertation nicht unbedingt erwartet werden kann, ist der große Umfang der zitierten Literatur. Als kleiner Wermutstropfen kann einzig festgehalten werden, dass einzelne neuere Forschungsergebnisse unerwähnt bleiben. So hätte man etwa in Kapitel B, §1.2 einen Hinweis darauf begrüßt, dass bei Stichprobenmitgliedern gerade Veränderungen – die mit Panelerhebungen ja insbesondere gemessen werden sollen – zu erhöhter Attrition führen (Ansätze z. B. in Voorpostel/Lipps 2011). Dies zieht tendenziell eine Unterschätzung von Veränderungen nach sich. Auch wäre in Kapitel B, §2.6 (und auf S. 59) die Zitierung der zwischenzeitlich etablierten Standardcodes von finalen Ausfallsgründen der American Association for Public Opinion Research (AAPOR 2011) sinnvoll. Zudem wäre eine stärkere Differenzierung von Nonresponse bei Querschnittstudien einerseits, die im Vergleich damit höhere Nonresponse bei der ersten Welle einer Panelstudie andererseits, und schließlich Panelattrition wünschenswert. Auch wäre die Erwähnung von Ansätzen zur Abschätzung von Verzerrungen durch Nonresponse und Attrition auf

Individualebene mit Hilfe von Registerdaten möglich (z. B. im CHINTEX Projekt mit ECHP Daten, vgl. Ehling et al. 2003; Røed 2006). Im Zusammenhang mit dem Vergleich von Nonresponse Raten und Nonresponse Verzerrung (Kapitel B, §3.2) hätte die Studie von Groves und Peytcheva (2008) erwähnt werden können. Zudem wäre ein Hinweis auf R-Indikatoren (Schouten et al. 2009) als Beispiel eines neuen Qualitätsindikators für die Repräsentativität von Befragungen willkommen. Entgegen der Kritik der mangelhaften Dokumentation der Stichprobenrealisierung bei Erhebungen gibt es Gegenbeispiele, z. B. beim European Social Survey (vergleiche die Methodenberichte zur Feldarbeit im ESS auf der ESS Website <http://ess.nsd.uib.no/>). Im Kapitel C hätte eine kurze Diskussion des Ausmaßes der Verzerrung durch Nichtkontakt einerseits und Verweigerung andererseits erfolgen können (etwa Lynn/Clarke 2002). Am Ende hätte man sich einen Anhang mit den verwendeten Fragebögen (inklusive der Studienbeschreibung bei der Nullmessung, vgl. S. 173) gewünscht; möglicherweise auch ein Glossar.

Allerdings muss ausdrücklich betont werden, dass die vorliegende Arbeit meines Erachtens die üblichen Anforderungen an eine Dissertation sowohl in Bezug auf den empirischen Teil als auch insbesondere auf den theoretischen Teil weit mehr als erfüllt. Etwaige Kritik muss im Kontext der insgesamt sehr hohen Qualität der Arbeit und des vermutlich geringen zur Verfügung stehenden finanziellen Budgets für die Datenerhebung beurteilt werden. Kleinere Mängel im Text (Schreibfehler, Bezüge auf Tabelleninhalte, vergessene Aufnahme zitierten Literatur im Verzeichnis) und beim Fragebogen (etwa Verlabelung der Endpunkte als „sicher“, nicht nur als „sehr wahrscheinlich“ wie auf S. 178) dürfen nicht überbewertet werden. Ich erlaube mir lediglich eine Überlegung dazu, warum die Modellvariablen insbesondere auf das tatsächliche Teilnahmeverhalten einen so geringen Einfluss haben: Erstens werden in der Nullmessung

die Effekte von fast allen befragungsbezogenen Variablen (Ausnahme Incentives) auf das selbst eingeschätzte hypothetische Teilnahmeverhalten erhoben und *nicht* bei der eigentlichen Befragung experimentell variiert. Damit ist es wohl wahrscheinlich, dass die Modellvariablen einen Einfluss auf das intendierte Teilnahmeverhalten haben. Der Einfluss auf das tatsächliche Verhalten aber steht und fällt dann mit dem Zusammenhang zwischen dem intendierten und dem tatsächlichen Verhalten. Dieser Zusammenhang ist aber aufgrund der vermuteten starken Wirkung situativer¹ Umstände nicht zu erwarten. Außerdem überlegt die Autorin, ob „das Teilnahmeverhalten ... für Studierende habituell abläuft ... als Gewohnheit anzusehen [ist, so dass] das Intentionsmaß irrelevant [wird]“ (S. 223).² Zweitens zeigt ein Vergleich der bestehenden Anwendungen des SEU Modells (Kapitel D, §4.4) und des TOPB Modells (Kapitel D, §5.3, Kap. F, §2.1) mit der Entscheidung, an einer (kurzen) Onlineerhebung teilzunehmen, dass letztere Entscheidung relativ unbedeutend ist. Aus diesem Grund würde man sich eine möglichst große Varianz der modellierten befragten- und befragungsbezogenen Einflussvariablen wünschen. Nun ist die gewählte Stichprobe äußerst homogen.³ Mehr Varianz bei den Befragten hätte etwa durch den Einbezug von mehr Universitäten, mehr Nicht-Sozialwissenschaftlern, oder auch Aufnahme von Nichtstudierenden realisiert werden können. Beim Befra-

gungsdesign hätte eine *tatsächliche* Variation von befragungsbezogenen Elementen wie eine Randomisierung des Sponsors oder des Befragungsthemas erfolgen können. Natürlich setzt die geringe Stichprobengröße hierbei enge Grenzen.⁴ Zudem wäre eine – von der Autorin wegen unvollständigem Rücklauf kritisierten – Nonresponse follow-up bei nicht angegebenen E-Mail Adressen möglich gewesen. Diese Kritiken werden aber zum Teil bereits von der Autorin problematisiert.

Eine vorsichtige Schlussfolgerung der Ergebnisse könnte sein, dass insgesamt nicht so sehr auf die (totale) Erklärungskraft des tatsächlichen Teilnahmeverhaltens durch die Variablen der beiden Modelle fokussiert wird, sondern auf Einzelbefunde. So könnte die Berücksichtigung der Tatsache, dass Kostenkomponenten für verschiedene Stichprobenmitglieder eine unterschiedliche Wirkung haben können, interessant sein. Zur besseren Erklärung des Teilnahmeverhaltens könnte in Zukunft mehr Gewicht auf das habitualisierte Verhalten und situative Umstände gelegt werden. Die Autorin zeigt, dass in Bezug auf das habitualisierte Verhalten bei Teilnehmern der ersten Welle bereits eine gewisse Verpflichtung zur weiteren Teilnahme an der zweiten Welle zu bestehen scheint, falls die Teilnahme in der ersten Welle angenehm verlief. Diese Erkenntnis kann zur Verbesserung von Panelstudien ausgenutzt werden: Möglicherweise sollte die erste Welle weniger und angenehmere Fragen enthalten und der Fragebogen in Folgewellen schrittweise verlängert werden. Generell scheint die Situation, in der sich Stichprobenmitglieder zum konkreten Befragungszeitpunkt befinden, eine weitaus bedeutendere Rolle für die Teilnahmeentscheidung zu spielen als vorab bekundete Teilnahmeintentionen. Für ein

1 Befragungsthema, -dauer und Auftraggeber würde ich nicht als situativ sondern als befragungsbezogen betrachten.

2 Diese Hypothese ließe sich anhand der Zeitspanne zwischen Erhalt der E-Mail Aufforderung zur Teilnahme und dem Zeitpunkt des Beginns der Befragung testen. Bei den Teilnehmern, die die Befragung spontan ausfüllen, sollte die Intention geringer sein als bei denen, die sich länger Zeit lassen.

3 Natürlich hat die gewählte Stichprobe den Vorteil der kosten- und aufwandgünstigen Erhebung der Probanden vor Ort sowie eine nahezu 100% Teilnahme bei der Nullmessung.

4 Eine größere Analysetichprobe hätte man aber eventuell erzielen können, wenn alle Probanden beide Theorie-Komponenten beantwortet hätten.

besseres Verständnis sollten bei zukünftigen Befragungen diese situativen Umstände (Zeitpunkt, Haushaltskontexte, Stimmung, Wetter, etc.) tatsächlich erhoben werden.

Diese Arbeit zeigt, dass für die Erforschung von Nonresponse und Attrition noch ein erhebliches Nachholpotential besteht. Einige äußerst interessante Anregungen werden von der Autorin am Beispiel einer Onlineerhebung von Studierenden gegeben. Es wäre zu begrüßen, wenn diese Anregungen in weiterführenden Arbeiten aufgenommen würden.

Literatur

- AAPOR (The American Association for Public Opinion Research) 2011: Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th edition (6.3.2012).
- Ehling, M., U. Rendtel et al., 2003: CHINTEX (The Change from Input Harmonisation to Ex-post Harmonisation in National Samples of the European Community Household Panel – Implications on Data Quality) – Synopsis. <http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Wissenschaftsforum/Chintex/ResearchResults/Downloads/Synopsis.psm1> (5.3.2012).
- Groves, R. und E. Peytcheva, 2008: The Impact of Non-Response Rates on Non-Response Bias: A Meta-Analysis. *Public Opinion Quarterly* 72 (2): 167-189.
- Lynn, P. und P. Clarke, 2002: Separating Refusal Bias and Non-Contact Bias: Evidence from UK National Surveys. *The Statistician* 51 (3): 319-333.
- Røed K., 2006: Longitudinal Administrative Registers in Economic Research – A Norwegian Experience. Unpublished Presentation for the Conference on Longitudinal Surveys in International Perspective, Montreal, 25-27 January 2006.
- Schouten, B., F. Cobben und J. Bethlehem, 2009: Indicators for the Representativeness of Survey Response. *Survey Methodology* 35 (1): 101-113.
- Voorpostel, M. und O. Lipps, 2011: Attrition in the Swiss Household Panel: Is Change Associated With Later Drop-Out? *Journal of Official Statistics* 27 (2): 301-318.

OLIVER LIPPS, LAUSANNE



STEPHAN BÜTTNER,
HANS-CHRISTOPH
HOB OHM, LARS
MÜLLER (Hg.),
2011: Handbuch
Forschungsdaten-
management.
Bock + Herchen
Verlag, Bad Honnef,
ISBN 978-3-88347-
283-6, 24,90 EUR.
Online Version:
<http://www.forschungsdatenmanagement.de/>

Man trifft immer wieder auf die teilweise paradoxe Situation, dass die Vorteile und Stärken wissenschaftlicher Methoden mit nicht erbringbaren Leistungen legitimiert werden oder dass zentrale Aspekte von Forschungsprozessen kaum oder nur ungenügend thematisiert werden, obwohl auf sie beträchtliche Teile des wissenschaftlichen Arbeitsprozesses entfallen. Beispielsweise gehört es zum Kern des wissenschaftlichen Selbstverständnisses, intersubjektiv zugängliche und reproduzierbare Ergebnisse zu produzieren. Aber ähnlich den Vergessenskurven von Hermann Ebbinghaus gehen quer über die verschiedenen Wissenschaftsfelder sehr schnell die meisten relevanten Datenbestände verloren. Wissenschaftliche Resultate erweisen sich daher gerade *nicht* als intersubjektiv zugänglich oder als reproduzierbar. Ein anderer Punkt betrifft den Nutzen wie den Vorteil von Sekundäranalysen sowie das mit der Zeit wachsende Ergebnispotential solcher Sekundäranalysen. Wegen der vielfach unzugänglichen Daten wird auf dieses Potential weitgehend verzichtet und stattdessen auf jeweils neue Datenerhebungen gesetzt, denen jeweils weitere neue Datenerfassungen folgen. Die akkumulierten Datenbestände aus der Vergangenheit, obwohl sie aus kontextuellen oder aus langfristigen Gesichtspunkten eine starke analytische Bereicherung darstellen würden, bleiben in ihrer überwiegenden Mehrzahl