

Statistische Analysen und Studien Nordrhein-Westfalen

Band 55

Impressum

Herausgeber :
Information und Technik
Nordrhein-Westfalen (IT.NRW),
Geschäftsbereich Statistik

Redaktion:
Bianca Oswald, Hans Lohmann

Preis der gedruckten Ausgabe: 4,00 EUR

Die Statistischen Analysen und Studien
NRW finden Sie als PDF-Datei
zum kostenlosen Download in unserer
Internet-Rubrik „Publikationsservice“.

Erscheinungsfolge: unregelmäßig

Bestellungen nehmen entgegen:

Information und Technik NRW,
Postfach 10 11 05,
40002 Düsseldorf,
Mauerstraße 51,
40476 Düsseldorf
Telefon: 0211 9449-2172/3516
Telefax: 0211 442006
Internet: <http://www.it.nrw.de>
E-Mail: poststelle@it.nrw.de

sowie der Buchhandel.

Pressestelle:
0211 9449-2521/2518

Zentraler Informationsdienst:
0211 9449-2495/2525

© Information und Technik
Nordrhein-Westfalen,
Düsseldorf, 2009

Vervielfältigung und Verbreitung,
auch auszugsweise, mit Quellenangabe
gestattet.

Bestell-Nr. Z 08 1 2009 51

ISSN 1619-506X

Inhalt

**Ergebnisse einer nordrhein-westfälischen Zusatzbefragung
der Interviewerinnen und Interviewer im Mikrozensus
zu ihren Erfahrungen in der Feldarbeit**

Diplom-Mathematiker Paul Berke

15

Ergebnisse einer nordrhein-westfälischen Zusatzbefragung der Interviewerinnen und Interviewer im Mikrozensus zu ihren Erfahrungen in der Feldarbeit

Dipl.-Math. Paul Berke

Im Jahr 2005 wurden der Mikrozensus und die integrierte Arbeitskräftestichprobe der EU auf eine kontinuierliche Erhebung umgestellt. Seitdem richten sich an diese Stichprobenerhebung viele neue inhaltliche und qualitative Anforderungen. Abgesehen von verschiedenen Problemen, die sich aus den Umstellungen der Konzepte selbst ergaben, zeigte sich insbesondere eine geringere Ausschöpfung geringfügiger Beschäftigungsverhältnisse im Vergleich zu anderen Quellen. Die fehlende Übereinstimmung wurde zum Anlass genommen, die Qualität der Feldarbeit im Rahmen eines Aktionsplans zu überprüfen. Die in diesem Beitrag vorgestellten Ergebnisse einer ergänzenden Interviewerbefragung sind daher nicht nur als Bestätigung der erfolgreichen Arbeit für die Interviewerinnen und Interviewer selbst von Interesse, sondern auch als öffentliche Werbung für die Überbrückung der Gegensätze zwischen Theorie und Praxis zu verstehen.

Vor der ausschließlich beschreibenden Darstellung der Einschätzungen der befragten Interviewerinnen und Interviewer erfolgt zunächst eine ausführliche Beschreibung des Befragungskonzeptes und des Aufbaus der Basisdaten der Interviewerbefragung. Die detaillierte Darstellung ist notwendig, weil weder das Befragungskonzept noch die Gewichtung der Antworten den sonst bei Testerhebungen üblichen Vorgehensweisen entspricht. Nebenbei erhält der Leser dieses Beitrags einige Informationen über den nordrhein-westfälischen Interviewerstamm.

Auch ohne statistische Tests ergibt sich aus den Ergebnisbeschreibungen als Fazit, dass wesentliche Qualitätssteigerungen der Statistik nur über eine Vertiefung des Dialogs zwischen Fragern und Befragten zu erreichen sind.

Einleitung

Im Jahr 2005 wurden der Mikrozensus und die integrierte Arbeitskräftestichprobe der EU auf eine kontinuierliche Erhebung umgestellt. Die vorher auf eine einzige Berichtswoche im Frühjahr bezogenen Interviews wurden nach einem Zufallsverfahren auf das ganze Jahr verteilt, womit sich die Basis für die Quartalshochrechnungen verringert hat. Der wichtigste Grund für diese Umstellung war die Forderung der EU, für die Fragen zur Erwerbstätigkeit nicht nur Jahresergebnisse, sondern auch Ergebnisse für alle Quartale des Jahres zu erhalten. Da aus der Umstellung vom Berichtswochenkonzept auf Quartalergebnisse und erstmalig mögliche Jahresdurchschnitte ohnehin Brüche in den langen Zeitreihen zu erwarten waren, wurde das Erhebungsprogramm modifiziert und teilweise noch erweitert. Zu den Modifizierungen zählte die Umstellung des Familienkonzeptes auf das Konzept der neuen Lebensfor-

men. Eine wesentliche Erweiterung betraf die Befragung zu Staatsangehörigkeiten und dem Migrationshintergrund.

Zusammen mit den inhaltlichen und terminlichen Erweiterungen ergaben sich höhere Anforderungen an die Qualität der erhobenen Daten. Die höchste Anforderung an die Qualität der Daten ergab sich aus der monatlichen Messung des ILO-Erwerbsstatus¹⁾ im Mikrozensus für Zwecke der Erwerbstätigenrechnung. Bereits in den ersten Quartalergebnissen zeigte sich, dass auch nach der Redukti-

1) Der ILO-Erwerbsstatus – ILO ist die Internationale Arbeitsorganisation – bezieht sich nicht auf die klassische Einteilung der Bevölkerung in Schüler, Erwerbstätige oder Rentner, sondern bezieht die Messung der Beteiligung am Erwerbsleben auf eine Berichtswoche. Als erwerbstätig werden Personen eingestuft, die in der Berichtswoche mindestens eine Stunde gegen Entgelt gearbeitet haben. Hinzu kommen Urlauber oder mithelfende Familienangehörige. Personen, die in der Berichtswoche nicht gearbeitet haben, gelten definitionsgemäß nur dann als erwerbslos, wenn sie aktiv nach Arbeit suchen und zeitnah für die Aufnahme einer Arbeit zur Verfügung stehen.

on auf die pro Quartal kleinere Hochrechnungsbasis die Ergebnisse des Mikrozensus zwar mit der Zahl der sozialversicherungspflichtig Beschäftigten oder der Zahl der bei den Arbeitsagenturen gemeldeten Personen immer noch gut übereinstimmten, dass aber im Vergleich zu anderen Datenquellen und Expertenschätzungen eine deutliche Unterschätzung der geringfügig Beschäftigten sichtbar wurde. Daraufhin wurde ein Aktionsplan erstellt, der fast ausschließlich auf die Prüfung und die Verbesserung der Datenerhebung konzentriert war. Nicht geprüft wurde dabei die Genauigkeit der Vergleichswerte oder die Eignung der statistischen Definition für die beabsichtigten Messungen.

Ein wichtiger Punkt des Aktionsplans war eine Nachbefragung von Haushalten, die im Mikrozensus im Jahr 2008 letztmalig befragt worden waren. Auf die Ergebnisse dieser Nachbefragung kann in diesem Beitrag nicht eingegangen werden. Dieses Projekt ist aber deshalb erwähnenswert, weil die eigentlich für die Befragung von Haushalten entwickelten Thesen Anlass für die durchgeführten Interviewerbefragungen waren. Da die Gefahr bestand, dass auf der Basis der Befragung weniger Haushalte und theoretisch abgeleiteter Thesen weitere einschneidende Eingriffe in die Erhebungs- und Aufbereitungsstrukturen des Mikrozensus gefordert werden könnten, war es nötig geworden, die Erfahrungen der Interviewerinnen und Interviewer nicht nur in der Feldarbeit, sondern auch für methodische Entscheidungen zu nutzen. Die drei Thesen mit möglichen Schlussfolgerungen für die zukünftige Organisation des Mikrozensus werden in den folgenden Abschnitten kurz dargestellt.

Das Konzept der Nachbefragung der Haushalte war wesentlich durch die

sogenannte **Hauptstatus-These** bestimmt. Diese These geht davon aus, dass befragte Personen in der Berichtswoche geleistete Arbeitsstunden nicht angeben, weil sie – gemessen an ihrem Status als überwiegend nicht erwerbstätige Rentner/-innen oder Student(inn)en – sich entweder nicht erinnern oder die geleistete Arbeit nicht für wichtig halten. Die Hauptstatus-These hatte bereits vor der empirischen Überprüfung dazu geführt, dass die Fragen zur Erwerbstätigkeit in den Reihenfolgen verändert und im Zusammenhang mit geringfügigen (marginalen) Beschäftigungen deutlich verstärkt worden sind. Zum Beispiel wird nach der Antwort, dass jemand nicht erwerbstätig ist, noch einmal nachgefragt, ob er oder sie nicht doch in der Berichtswoche gearbeitet hat. Der Befragte kann dabei den Eindruck gewinnen, dass auch seine möglicherweise nach bestem Wissen gegebene erste Antwort angezweifelt wird. Das Ausmaß und die Folgen dieser Störung des Interviewablaufs sind noch nicht hinreichend untersucht worden, da bei allen vorgenommenen Änderungen des Frageprogramms und der Formulierungen nur der Wunsch, jede geleistete Arbeitsstunde in die Ergebnisse einzubringen, im Vordergrund stand.

Als weitere Ursache für die Untererfassung geringfügiger Beschäftigungen wurde aus dem Blickwinkel der Theorie die **Effizienz-These** formuliert. Die Effizienz-These besagt, dass Interviewer/-innen und Interviewte ein Interesse an einer schnelleren Durchführung der Interviews haben. Aufgrund der Länge der Frageblöcke zur Erwerbstätigkeit bestünden folglich Anreize, kleinere Tätigkeiten sowie die Suche danach nicht korrekt anzugeben. Während bei der Hauptstatus-These nur das Wissen der Befragten in Frage gestellt war, wird hier den befragten Haushalten und den Interviewerinnen und Interviewern ein wesentliches Abweichen von den Vorgaben unterstellt. Selbst wenn es gelingen würde, diese These in Einzelfällen zu bestätigen, wäre sie immer noch für eine methodische Weiterentwicklung der Mikrozensus-Befragungen völlig ungeeignet: Wenn die Aktivität zur bewussten Verkürzung der Interviews messbar von den Befragten aus-

gehen sollte, ist dies nur als Bestätigung des Verdachtes zu interpretieren, dass die Befragten durch die aktuelle zeitliche Länge des Interviews bereits überfordert sind. In diesem Fall würde nur eine zurzeit nicht durchsetzbare Verkürzung des Gesamtinterviews weiterhelfen. Die Feststellung gezielter Verkürzungen der Interviews durch Interviewer/-innen ist nur durch direkte Kontrollen möglich. Schon deshalb war sie zumindest als Ausgangspunkt für eine Nachbefragung der Haushalte ungeeignet.

Die **Proxy-These** ging davon aus, dass vor allem bei zeitlich stark variablen Merkmalen von einem erhöhten Messfehler auszugehen ist, wenn Personen über andere, nicht anwesende Haushaltsmitglieder Auskunft geben. Auch aus Unwissenheit kann es hier zu falschen Auskünften kommen. Unter dem Eindruck dieser These wurde in der Methodendiskussion bereits der Vorschlag gemacht, die Haushaltsbefragungen durch die Befragung jeder einzelnen Person zu ersetzen. Der Versuch, auch alle generell selten anzutreffenden Personen eines Haushalts persönlich zu befragen, ist aber schon wegen der erheblichen Mehrkosten und des hohen Zeitaufwandes nur dann zu rechtfertigen, wenn eine deutliche Steigerung der Qualität der Selbstauskünfte im Vergleich zu den – in diesem Beitrag auch als Fremdauskünfte bezeichneten – Proxy-Antworten nachgewiesen werden kann. Theoretisch wird davon ausgegangen, dass Auskünfte über sich selbst genauer sind als Auskünfte über eine andere Person. Auch eine Nachbefragung der Haushalte konnte zu dieser Frage nicht mehr leisten als einen einfachen Vergleich der Fremdauskünfte mit den als zutreffend angenommenen Selbstauskünften. Dass auch dieser Ansatz sehr viele Vereinfachungen beinhaltet, ist bereits an den ohnehin erforderlichen Antworten der Eltern für ihre minderjährigen Kinder zu sehen.

In der Erwartung eines deutlich differenzierteren Gesamtbildes der Feldarbeit schlugen die statistischen Landesämter vor, die Haushaltsbefragung um eine Befragung der Interviewer zu ergänzen. Nur die Interviewerinnen und Interviewer haben die Gelegen-

heit, die Haushalte schon zum Zeitpunkt der Entstehung der Antworten auf die Fragen des Mikrozensus unmittelbar zu beobachten. Deshalb wurde davon ausgegangen, dass eine Abfrage dieser Beobachtungen nicht nur verschiedene Einzelheiten ergänzt, sondern unter bestimmten Voraussetzungen auch Bewertungen für die Ergebnisse der Haushaltsbefragungen liefert. Voraussetzung dafür war aber, die mit der Erhebung der Daten Beauftragten nicht selbst als Befragte und erst recht nicht als Testpersonen anzusprechen, sondern als Personen, die über noch nicht ausgeschöpftes Expertenwissen verfügen. Entgegen der eigentlichen Intention des Vorschlags umfasste aber die bundesweit durchgeführte Befragung der Interviewer/-innen nur wenige Fragen zur Interviewsituation und zur Wirkungsweise der Fragen. Stattdessen gab es überwiegend Fragen zum Interviewereinsatz selbst, den Schulungen und Materialien, zur Organisation der Feldarbeit und zur Zusammenarbeit mit den statistischen Landesämtern. Beeinflusst von den Thesen für die Haushaltsbefragung war ein Frageprogramm entstanden, welches eher auf Schulungs- und Organisationsfragen ausgerichtet war als auf das eigentliche Befragungsziel. Auch die Formulierung der Inhalte war nicht mehr durchgängig auf Experten für die Einschätzung der Reaktionen der befragten Haushalte zugeschnitten.

Deshalb wurde in Nordrhein-Westfalen die Befragung der Interviewer durch ein Zusatzprogramm ergänzt, welches ausschließlich auf die Gewinnung von Informationen über Erfahrungen und Einschätzungen der Interviewerinnen und Interviewer zur Entstehung und Genauigkeit der Mikrozensus-Ergebnisse abzielte. Es war offensichtlich, dass die Befragung der Interviewerinnen und Interviewer nicht durch die sonst üblichen Testmethoden vor Gefälligkeitsantworten oder Beschönigungen bewahrt werden konnte. Stattdessen wurde versucht, die Interviewerinnen und Interviewer für eine kooperative Preisgabe ihres Expertenwissens zu gewinnen. Vor der im nächsten Abschnitt folgenden Beschreibung der Inhalte und Beweggründe des Ansatzes der Zusatz-

befragung ist deshalb noch einmal zu betonen, dass der Befragungsansatz nicht die Qualität der Feldarbeit messen sollte, sondern ausschließlich auf die Sammlung von Expertenwissen ausgerichtet war.

Beschreibung des Konzeptes der nordrhein-westfälischen Zusatzbefragung

Die Zusatzbefragung war so aufgebaut, dass für 11 Beispielfragen jeweils 20 Einzelereignisse hinsichtlich der Häufigkeit des Auftretens abgefragt wurden. Zur Gruppierung der Einzelereignisse war es erforderlich, in jedem der vier in der Übersicht 1 dargestellten Aspekte sowohl eine logische Reihenfolge zu konstruieren, als auch die Ereignisse so zu beschreiben, dass dem Interviewer eine Identifikation möglich ist. Da die Reaktionen der befragten Haushalte auch davon abhängig sind, wie die konkrete Frage transportiert wird, konnte schon mit Blick auf die Effizienz-These auf eine Abfrage des Interviewerverhaltens nicht völlig verzichtet wer-

den. Um aber den Akzent der Befragung nicht direkt auf das Verhalten der Interviewer/-innen zu legen, wurden im ersten Block unter dem Motto „Befragungsaufwand“ alle Abfragen zusammengefasst, die die Interviewerin bzw. den Interviewer unmittelbar betreffen. Im idealisierten Ablauf des Interviews wird die Frage vorgelesen, die befragte Person versteht und akzeptiert die Frage und kann sie für sich selbst und bei Bedarf auch für weitere Haushaltsmitglieder zutreffend beantworten. Unter den Bedingungen persönlicher Interviews besteht zusätzlich die Möglichkeit, die gegebene Antwort noch vor der Übernahme in das Befragungsergebnis zu bewerten. Wenn die Frage im Wortlaut gestellt und vom Befragten ohne Zögern kurz beantwortet wurde, hat der Interviewer normalerweise keine Veranlassung die Richtigkeit der Antwort in Zweifel zu ziehen. Deshalb konnte der normale Ablauf der Beantwortung einer Einzelfrage sowohl hinsichtlich der Genauigkeit als auch der Akzeptanz nur durch fehlendes Zögern oder kurze Antworten charakterisiert werden.

Störungen oder Verzögerungen des Interviews sind dagegen leichter zu beschreiben und auch für den Interviewer einfacher vom normalen Ablauf zu unterscheiden, kommen aber seltener vor. Massive Störungen, wie z. B. die Antwortverweigerung gegenüber der/dem bereits anwesenden Interviewer/-in, sind sogar so selten, dass einige Interviewerinnen und Interviewer diese Situation trotz mehrjähriger Feldarbeit höchstens einmal erlebt haben. Deshalb schien es erforderlich, die bundesweit verwendeten Antwortmöglichkeiten „oft“, „manchmal“, „selten“ und „nie“ in der Zusatzbefragung um die Antwortkategorie „nur einmal“ zu erweitern. Dadurch war es den befragten Interviewerinnen und Interviewern möglich, seltene Ereignisse anders zu bezeichnen als absolute Ausnahmefälle. Die Ereignisse wurden – wie in der ersten Übersicht zu sehen ist – in vier Themenbereiche gruppiert, um den befragten Interviewerinnen und Interviewern die Orientierung zu erleichtern.

Übersicht 1: Aufbau und Reihenfolge der Einzelereignisse

<p>Abfrage des Bearbeitungsaufwands:</p> <p>A) Ich lasse mir die im Zusammenhang erkennbar richtige Antwort von der befragten Person bestätigen.</p> <p>B) Die befragte Person liest die Frage selbst vom Bildschirm ab und antwortet.</p> <p>C) Ich lese die Frage vor.</p> <p>D) Ich erläutere die Frage mit eigenen Worten oder gebe Beispiele.</p> <p>E) Auch nach zusätzlichen Hinweisen und Beispielen ergeben sich weitere Rückfragen.</p>	<p>I) Der Haushalt reagiert in einer Weise, dass die weitere Durchführung des Interviews kaum noch möglich ist.</p> <p>J) Der Haushalt lehnt die Beantwortung weiterer Fragen ab.</p>	<p>Abfrage der Genauigkeit bei Fremdauskünften (abwesende Personen):</p>
<p>Abfrage der Akzeptanz:</p> <p>F) Die Frage wird direkt und ohne Zögern beantwortet.</p> <p>G) Das Interview verzögert sich durch zusätzlichen Erläuterungsbedarf zum Zweck der Frage.</p> <p>H) Der Haushalt kritisiert die Frage, kann aber zur weiteren Mitarbeit bewegt werden.</p>	<p>Abfrage der Genauigkeit bei Selbstauskünften (anwesende Personen)</p> <p>K) Die Frage wird sofort verstanden und kurz beantwortet.</p> <p>L) Die befragte Person verweist bei dieser oder bei ähnlichen Fragen auf bereits gegebene Antworten.</p> <p>M) Die befragte Person ändert die ursprüngliche Antwort aufgrund einer Nachfrage oder Erläuterung.</p> <p>N) Die befragte Person kann die Antwort nur nach Aufforderung zu einer Schätzung geben.</p> <p>O) Die befragte Person besteht auf der Eintragung einer vermutlich falschen Antwort.</p>	<p>P) Die Frage wird ohne Zögern auch für abwesende Haushaltsmitglieder beantwortet.</p> <p>Q) Die Antwort für abwesende Haushaltsmitglieder wird gegeben, aber mit vorsichtigen Worten wie z. B. „soweit mir bekannt“ eingeschränkt.</p> <p>R) Anwesende Haushaltsmitglieder haben nur lückenhafte Vorstellungen über die richtige Antwort für eine abwesende Person, entschließen sich aber letztendlich selbst, eine Antwort zu geben.</p> <p>S) Die lückenhaften Vorstellungen der anwesenden Haushaltsmitglieder über die richtige Antwort führen dazu, dass eine Vermutung oder Schätzung als Antwort hingenommen werden muss.</p> <p>T) Die Beantwortung der Frage für eine abwesende Person wird abgelehnt</p>

Die Abfrage des Bearbeitungsaufwandes weicht von den folgenden Abfragen bereits dadurch ab, dass die neutrale bzw. problemlose Rubrik nicht am Anfang, sondern in der Mitte der Einzelereignisse steht. Dies ergab sich aus der angestrebten Reihenfolge. Eine weitere, mit Absicht vorgenommene befragungstechnische Abweichung von den drei anderen Abfragen und auch der sonst üblichen Vorgehensweise ist darin zu sehen, dass in der ersten Rubrik A) für das Einzelereignis bereits eine Begründung in der Formulierung vorhanden ist. Dass in der Praxis persönlicher Interviews (face-to-face) sich das wörtliche Vorlesen jeder Frage sogar verbietet, ist am Beispiel des Merkmals „Geschlecht“ bereits zu sehen. Auch wenn die Antworten auf andere Fragestellungen für die Interviewerin bzw. den Interviewer vielleicht nicht ganz so offensichtlich sind, befindet sich die Vorgabe, alle Fragen explizit zu stellen, stets in Konkurrenz mit der Empfehlung, die Befragten nicht übermäßig mit dem Wortlaut der statistischen Definitionen zu belasten. Dass es durch diese Form der Abfrage eher gelingt, die Interviewerinnen und Interviewer als Kooperationspartner mit Expertenwissen anzusprechen, zeigte auch der Vergleich der Häufigkeitsangaben mit denen der bundesweiten Interviewerbefragung.

In der bundesweit durchgeführten Interviewerbefragung wurde als Filterfrage gefragt, ob er oder sie alle Fragen stellt, auch wenn die Antwort aus dem Interviewverlauf schon bekannt ist. Anschließend wurde – ähnlich wie in der Zusatzerhebung – gefragt, wie häufig es bei bestimmten Themenblöcken vorkommt, dass er oder sie eine Antwort einträgt, ohne explizit danach zu fragen. Zur Herstellung der Vergleichbarkeit wurden aus der bundesweiten Erhebung nur die Interviewerinnen und Interviewer berücksichtigt, die auch an der Zusatzerhebung teilgenommen hatten. Außerdem wurden bei allen Interviewern, die angegeben hatten, dass sie immer alle Fragen stellen, die Antworten zu den Themenblöcken auf „nie“ gesetzt. Im Gegenzug musste in der Zusatzbe-

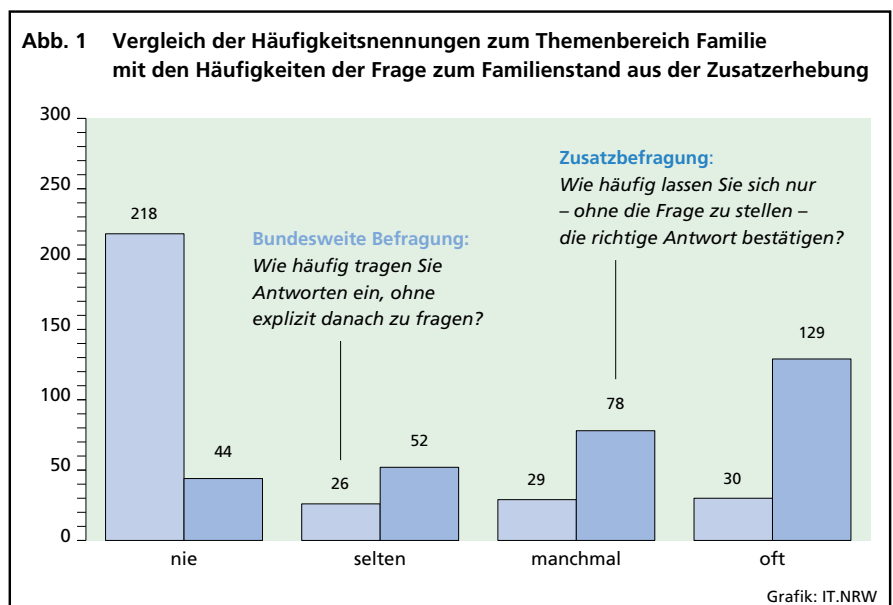
fragung die Häufigkeitsbezeichnung „einmal“ mit „selten“ für den Vergleich zusammengefasst werden. In der Zusatzbefragung zeigt sich die Wirkung der in der Frage mitgelieferten Begründung für das Abweichen von der allgemeinen Vorgabe, dass alle Fragen vorgelesen werden sollen.

Im Abbildung 1 ist zu sehen, dass die gleiche Interviewergruppe, die eine nennenswerte Zahl an Situationen erlebt, die ein explizites Stellen der Frage nicht angebracht erscheinen lassen, das eigene – schon auf dem ersten Blick sachgerechte – Verhalten nicht wahrnimmt, wenn es unvermittelt und mit der Fixierung auf die Einhaltung der methodischen Vorgaben abgefragt wird. Aus nachträglichen Gesprächen mit einzelnen Interviewerinnen und Interviewern ist bekannt, dass erst die Kenntnis der Intention der Zusatzbefragung einige Interviewerinnen und Interviewer dazu bewegt hatte, die Filterfrage in der bundesweiten Befragung zu verneinen und die Ausnahmen von der Grundvorgabe, dass alle Fragen vorzulesen sind, zu benennen. In den Ergebnissen der Zusatzbefragung wird sich außerdem noch zeigen, dass die vorweg genommene Begründung für das Übergehen von Fragen im Verlauf des Interviews tatsächlich ein Motiv der Interviewführung war.

Auf der anderen Seite ist ein Abweichen vom schlichten Vorlesen der

Frage nötig, wenn die befragte Person die Frage nicht oder nicht richtig versteht. In diesen Fällen kann zwar eine gewisse Standardisierung dadurch erreicht werden, dass die Interviewerin bzw. der Interviewer – wie auch im Mikrozensus – mit zusätzlichen Erläuterungen und vorgefertigten Beispielen ausgestattet wird. Aber schon bei einer Erläuterung mit eigenen Worten ergeben sich nicht kontrollierbare Effekte in den Ergebnissen. Unverständnis bei befragten Personen tritt immer dann auf, wenn die Frage sich auf Definitionen bezieht, deren Abgrenzungen in den Antworten besonders präzise eingehalten werden sollen. In diesem Moment kann die Interviewerin oder der Interviewer nur noch helfen, wenn es ihr oder ihm gelingt, dem Befragten den Sinn der Frage zu erläutern, ohne sich dabei zu weit von den methodischen Vorgaben zu entfernen. An dieser Stelle ist darauf hinzuweisen, dass der Einsatz eigener Worte oder zusätzliche Hinweise und Erläuterungen das Interview im Vergleich zum einfachen Vorlesen der Frage zeitlich verlängern. Dieser Aspekt führte letztlich zu der Entscheidung, die Abfrage zum Interviewereinsatz in der aus Übersicht 1 ersichtlichen Form unter das Motto „Bearbeitungsaufwand“ zu stellen.

Weniger schwierig waren die Abstufungen der Akzeptanz der Fragen und der Genauigkeit von Selbstausskünften und Fremdausskünften zu



definieren. Für die/den Interviewer/-in wahrnehmbare Einschränkungen der Akzeptanz ergeben sich aus verzögerten Antworten über kritische Bemerkungen der Auskunftspflichtigen bis zur offenen Ablehnung der Beantwortung. Ungenaue Antworten sind bereits an einschränkenden Bemerkungen der Befragten erkennbar. Hinzu kommen Unsicherheiten bei der Beantwortung bis hin zum logischen Widerspruch zu den Antworten auf andere Fragen.

Bei genauen Antworten auf akzeptierte Fragen ist davon auszugehen, dass nur das Fehlen der vorher genannten Einschränkungen bemerkbar ist. Deshalb konnte für die Situationsbeschreibungen F), K) und P) in der Übersicht 1 eigentlich nur der geringe Zeitbedarf herangezogen werden. Bei der Abfrage der Akzeptanz und der Genauigkeit von Fremdauskünften wurde dabei auf den Zeitraum bis zum Beginn der Beantwortung und bei den Selbstauskünften auf die Länge der Antwort abgestellt. Dass diese feine Unterscheidung bei der Beantwortung bemerkt wird, war aber nicht von großer Bedeutung, weil im Erhebungsvordruck auch die Überschriften der Abfragen zu jeder Beispielfrage wiederholt wurden.

Das Ziel, Angaben zur Genauigkeit der Selbst- und der Fremdauskünfte ausschließlich aus – für die/den Interviewer/-in objektiv wahrnehmbaren – Sachverhalten abzuleiten, konnte mit Rücksicht auf die logische Abstufung der Abfragen nicht völlig durchgehalten werden. Bei den Selbstauskünften betrifft dies die Position O) am oberen Rand der Ungenauigkeit. Hier war es aber erforderlich, den Interviewerinnen und Interviewern die Unterscheidung von geschätzten und definitiv falschen Antworten zu ermöglichen. Bei den Fremdauskünften ist die mittlere Position R) besonders subjektiv gefärbt. Aus Gründen der Symmetrie dient diese Position hauptsächlich dazu, die Positionen Q) und S) deutlicher voneinander zu trennen und der Interviewerin bzw. dem Interviewer im Einzelfall eine klare Unterscheidung zu ersparen.

Aus nahe liegenden Gründen wurde das eben beschriebene Abfrageraster nicht auf ganze Themenbereiche, sondern auf die in Übersicht 2 aufgeführten Beispielfragen bezogen. Da es für ein geschlossenes Gesamtbild erforderlich war, alle 20 Einzelsituationen für jede Beispielfrage vollständig abzuarbeiten, war es nicht möglich, sehr viele Beispielfragen auszuwählen. Insgesamt 11 Einzelfragen waren aber als Minimum erforderlich, weil nicht nur unterschiedliche Themenbereiche, sondern auch für spezifische Situationen typische Fragestellungen abzudecken waren. Im Einzelnen wurden die Beispielfragen nach folgende Kriterien ausgesucht:

Die erste Frage nach dem Familienstand steht als Frage 10 des Laptop-Interviews fast am Anfang der Befragung, sodass weder vom Inhalt noch von der Länge des Interviews her gesehen besondere Komplikationen erwartet werden können. Deshalb bestand während der Konzeption der Interviewerbefragung die Absicht, diese Frage als neutralen Maßstab für die spezifischen Verteilungen der übrigen Fragen zu nutzen.

Die zweite Frage aus dem Thema Erwerbstätigkeit zur Existenz marginaler Beschäftigungen hat durch das Bemühen, eine möglichst hohe Ausschöpfung zu erreichen, einen fast suggestiven Charakter angenommen. Sie schien deshalb nicht nur wegen der zu erwartenden emotionalen Barrieren, sondern auch als Kristallisationspunkt der Effizienz-These unverzichtbar zu sein. Die im Interview unmittelbar folgende direkte Frage nach einer geringfügigen Beschäftigung in der Berichtswoche dient in diesem konkreten Zusammenhang der Ermittlung von Vergleichsstrukturen. Außerdem wurde auch bei dieser Frage unterstellt, dass die Antworten aus Fremdauskünften deutlich ungenauer ausfallen als Selbstauskünfte.

Die am Anfang bereits erwähnte Proxy-These stand auch bei der Auswahl der drei folgenden Beispielfragen im Vordergrund. Die Frage 4 nach dem Wirtschaftszweig des Be-

triebs erlaubt dabei zusätzliche Beobachtungen, weil wegen der sofortigen Codierung der Antworten die befragende Person nicht nur auf eine ausführlichere Antwort angewiesen ist, sondern auch noch darauf achten muss, dass keine zu lange Suche der zutreffenden Codierung im Thesaurus der Wirtschaftszweige entsteht. Die Frage 5 zum Wunsch nach Mehrarbeit bietet die Möglichkeit, eine oberflächliche und nicht überprüfbare Antwort für abwesende Haushaltsmitglieder zu geben, erfordert aber für eine exakte Beantwortung sehr genaue Kenntnisse über den aktuellen Standpunkt der betroffenen Person zu dieser Frage. Dies gilt in geringerem Maße auch für die Frage 6 nach der Zahl der Arbeitsstunden. Das Schätzrisiko (z. B. aus der irrtümlichen Einbeziehung von Arbeitswegen) in den Fremdauskünften ist bei dieser Frage sogar etwas geringer, weil die Zahl der tatsächlich gearbeiteten Stunden in vielen Fällen mit den normalen Arbeitszeiten übereinstimmt.

Zum Thema Arbeitssuche wurden aus dem Frageprogramm des Mikrozensus die Fragen 7 und 8 ausgesucht, weil auch die Frage nach der Suche einer anderen Tätigkeit weitere Fragen nach sich zieht und daher ebenfalls die Effizienz-These stützende Antwortstrukturen aufweisen könnte. Die auf eine Erfassung der sogenannten Stillen Reserve abzielende zweite Frage löst zwar keine Folgefragen aus, birgt aber ein größeres Potenzial bei der Auskunft über Dritte entweder nicht oder nur sehr oberflächlich beantwortet zu werden.

Der hohe Anteil von Jugendlichen und jungen Erwachsenen an den Fremdauskünften zeigt, dass für die Analysen eine nicht arbeitsmarktbezogene Frage zur Ermittlung des Wissens oder Nichtwissens von Eltern über die Lebensverhältnisse ihrer fast erwachsenen Kinder erforderlich ist. Dazu wurde die Frage 9 ausgewählt. Die Frage 10 nach dem Geburtsland steht fast am Ende des Interviews und ergänzt daher die auch in anderer Hinsicht problematische Frage 11 nach dem Nettoeinkommen

Übersicht 2: Liste der Beispielfragen

1. Zur Frage 10: Welchen Familienstand haben Sie? (Familienstand)	hört der Betrieb an, in dem Sie tätig sind? (Branche)	oder Berufstätigkeit suchen, würden Sie denn trotzdem gern arbeiten? (Stille Reserve)
2. Zur Frage 23: Auch wenn man eigentlich nicht erwerbstätig ist, kann man ja z. B. als Schüler/-in, Hausfrau/Hausmann oder Rentner/-in Tätigkeiten ausüben, mit denen man Geld hinzuverdient. Wie ist das bei Ihnen: Üben Sie regelmäßig oder gelegentlich eine solche bezahlte Tätigkeit aus? (Hinzuverdienst)	5. Zur Frage 46: Haben Sie den Wunsch, normalerweise mehr Stunden pro Woche als derzeit zu arbeiten? (Wunsch nach Mehrarbeit)	9. Nachdem die Frage zum Schul- oder Hochschulbesuch mit „Ja“ beantwortet wurde: Zur Frage 88a: Um welche Schule oder Hochschule handelte es sich dabei? (Art der Schule)
3. Zur Frage 24: Haben Sie in der letzten Woche (zusätzlich) eine geringfügige Beschäftigung, Minijob oder einen Ein-Euro-Job ausgeübt? (Minijob)	6. Zur Frage 47: Und wie viele Stunden (einschließlich Überstunden) haben Sie in der letzten Woche tatsächlich gearbeitet? (Wochenstunden)	10. Frage 98: Sind Sie auf dem heutigen Gebiet der Bundesrepublik Deutschland geboren? (Land der Geburt)
4. Zur Frage 37: Welchem Wirtschaftszweig ge-	7. Zur Frage 65: Haben Sie in den letzten 4 Wochen eine andere oder weitere Tätigkeit gesucht? (Andere Tätigkeit gesucht)	11. Zur Frage 111: Wie hoch waren Ihr persönliches Nettoeinkommen und Ihr Haushaltseinkommen im letzten Monat? (Nettoeinkommen)
	8. Zur Frage 73: Auch wenn Sie keine Erwerbs-	

um die Chance Effekte zu entdecken, die ausschließlich auf die Länge des gesamten Interviews zurückzuführen sind.

Aus Kontakten mit Interviewerinnen und Interviewern war außerdem bereits bekannt, dass die gewählten Beispielfragen charakteristische Reaktionen in den befragten Haushalten auslösen, sodass eine Verknüpfung der Reaktionen der Haushalte mit der verursachenden Frage in der Erinnerung erwartet werden konnte. Auch die Beteiligung der Interviewerinnen und Interviewer an der Auswahl der Beispielfragen ist unter dem Blickwinkel zu bewerten, dass die Interviewerbefragung nicht als Erhebung von Daten für statistische Tests²⁾, sondern als Expertenbefragung betrachtet wurde. Außerdem wäre die Abfrage der Häufigkeiten von zwanzig verschiedenen Situationsbeschreibungen mit zum Teil nur auf den zweiten Blick erkennbaren Unterschieden für willkürlich oder

2) Statistische Signifikanztests und Fehlerrechnungen basieren auf dem Vergleich der Ergebnisse mit zufälligen Verteilungen. Wenn die Befragten an der Auswahl der Inhalte und der Form der Auswertung beteiligt werden, können die Ergebnisse systematische Abweichungen vom Zufallsprinzip aufweisen.

zufällig ausgewählte Beispielfragen nicht möglich gewesen. Zur Vorbereitung der verwendeten Gewichtung der Ergebnisse der Interviewerbefragung erfolgt im nächsten Abschnitt eine etwas ausführlichere Beschreibung der befragten Interviewerinnen und Interviewer. Einblicke in die Struktur der Rohdaten werden zeigen, weshalb die anschließend beschriebene Gewichtung der Antworten zweckmäßig war.

Aufbau der Basisdaten aus der Interviewerbefragung

Von den 374 im Jahr 2008 für den Mikrozensus in Nordrhein-Westfalen einsetzbaren Interviewerinnen und Interviewern nahmen 323 an der bundesweiten Interviewerbefragung und 324 an der nordrhein-westfälischen Zusatzbefragung teil. Die hohe Ausschöpfung erklärt sich zum Teil aus dem Sachverhalt, dass für die Beteiligung an beiden Befragungen eine kleine Aufwandsentschädigung gezahlt wurde. Das wichtigere Motiv für die hohe Beteiligung schien aber die Verbundenheit der Beteiligten mit ihrer Tätigkeit zu sein. Die hohe Beteiligungsquote

von mehr als 85 % kam vermutlich auch dadurch zustande, dass aus dem Interviewerstamm nur die im Jahr 2008 Aktiven im Rahmen der jährlich durchgeführten Schulungsveranstaltungen angesprochen wurden. Die für einige Auswertungen notwendige Zusammenführung der Informationen aus beiden Befragungen war für weniger als 10 Datensätze nicht zweifelsfrei möglich. Außerdem musste auf weitere Stichprobenfälle verzichtet werden, weil die Häufigkeitsangaben zu geringe Differenzierungen aufwiesen, um in der geplanten Gewichtung brauchbare Proportionen zu erzeugen, sodass sich die Auswertungsbasis schon im Vorfeld auf 303 Stichprobenfälle reduzierte.

Bei den verbliebenen Antworten musste die Originalangabe zur Zahl der im Jahr 2007 bearbeiteten Auswahlbezirke wegen fehlender bzw. unplausibler Werte in 20 Einzelfällen durch den errechneten Mittelwert von 26 Auswahlbezirken ersetzt werden. Die mit dieser Prüfung verbundene Summenkontrolle ergab, dass von den im Jahr 2007 insgesamt bearbeiteten Auswahlbezirken (8 387) einschließlich der erwähnten Schät-

zung 98 % in der Befragung berücksichtigt sind. Der Einsatz des Mittelwertes für die wenigen Imputationen war erforderlich, weil in der Basiserhebung die hier ermittelte Durchschnittszahl der monatlich erhobenen Haushalte nicht auf das Kalenderjahr, sondern auf eine nicht bestimmbare Zahl „aktiver“ Monate bezogen war. Die zweite Gewichtungskomponente (Anzahl der Jahre der Tätigkeit) wies nach Bearbeitung in nur 7 Einzelfällen einen Schätzwert auf: In diesen Fällen wurde das erste Jahr der Tätigkeit ermittelt und unterstellt, dass in jedem Jahr seit der ersten Beteiligung Interviews geführt wurden. Der direkte Vergleich mit dem Jahr des Beginns der Interviewertätigkeit zeigt aber, dass dies insbesondere bei den längerfristig aktiven Interviewerinnen und Interviewern nicht immer der Fall war.

Während die Gewichtungskomponenten direkt aus der Zusatzerhebung stammen, wurden weitere Charakteristiken der Interviewerinnen und Interviewer der bundesweit durchgeführten Befragung entnommen. Für die weiteren Auswertungen sind allerdings nur die Gliederung nach Geschlecht, Alter (25 – 44 Jahre/45 Jahre und älter) und Erwerbstätigkeit (ja/nein) nutzbar. Da in Nordrhein-Westfalen die jüngste Interviewerin bereits 26 Jahre alt war, blieb die auf Bundesebene erfragte untere Altersklasse völlig leer. Auf der anderen Seite gab es nach der Umstellung auf die kontinuierliche Erhebung mit Laptops nur noch 15 aktive Interviewerinnen und Interviewer im Alter von 65 und mehr

Jahren. Hinsichtlich der sozialen Stellung wiesen die Rohdaten der Interviewerbefragung 3 Student(inn)en, 24 Rentner/-innen, 21 Hausfrauen/-männer, 3 Arbeitslose und 11 „sonstige Nichterwerbstätige“ auf, die aufgrund der geringen Besetzungen nur insgesamt eine Vergleichsgruppe zu den Berufstätigen darstellen können.

Auch die für die Bewertung der Interviews möglicherweise viel wichtigere Charakterisierung der Auswahlbezirke konnte nur in sehr allgemeiner Form erfolgen, weil in Verbindung mit der Zusatzbefragung die detaillierte Abfrage der Struktur der Auswahlbezirke ohne völlige Überforderung der Befragten nicht möglich gewesen wäre. So konnte in der Befragung nur ermittelt werden, dass die Einschätzungen der Interviewerinnen und Interviewer sich gemäß der Gewichtung der Zusatzbefragung zu 14 % auf Haushalte in ländlichen Regionen des Landes Nordrhein-Westfalen beziehen. Doppelt so viele Haushalte wurden über die Auswahlbezirke einer kleinstädtischen Umgebung zugeordnet. Die Anteile für Großstädte und der zumindest teilweise ähnlich strukturierten Mischgebiete überwogen demnach mit 58 % in der Summe.

Die Tabelle 1 zeigt im Überblick die Gewichtungskriterien, die für die Schätzung der Gesamtzahl der von dieser Auswertung betroffenen Interviews genutzt wurden. Bereits in der Tabelle ist zu sehen, dass keine statistische Abhängigkeit der beiden Kriterien einer multiplikativen Verknüpfung zu einem Messwert entgegen-

genstand.³⁾ Eine Gewichtung ist aber nicht nur möglich, sondern auch erforderlich, weil sonst der Interviewer, der im Jahr 2007 nur 2 Auswahlbezirke bearbeitet hatte, das gleiche Gewicht bekommen hätte, wie die Kollegin mit der Maximalzahl von 67 Auswahlbezirken. Im Idealfall hätte sich die Gewichtung der Antworten auf die Zahl der Haushalte beziehen müssen, mit denen die Interviewer/-innen Kontakt hatten. Eine entsprechende Abfrage bei den Interviewerinnen und Interviewern war aber für Gewichtungszwecke zu ungenau, weil nur wenige Interviewer/-innen für ihre persönlichen Abrechnungen entsprechende Aufzeichnungen führen. Auf die technisch machbare Rekonstruktion der Zahl der von jeder Interviewerin bzw. jedem Interviewer in 2007 befragten Haushalte aus den Begehungs- und Verteilungsinformationen wurde wegen des hohen Aufwandes verzichtet. Der Aufwand hätte sich auch nicht gelohnt, weil die Ereignisabfrage für einen festgelegten Zeitraum eine weitere Belastung des Erinnerungsvermögens der Befragten erzwungen hätte. Die Abfrage ohne direkten Zeitbezug erfolgte außerdem mit Rücksicht auf die Abfrage seltener Ereignisse, die möglicherweise in der Jahresarbeit eines einzelnen Interviewers überhaupt nicht vorgekommen sind. Auch aus diesem Grund ist neben dem aktuellen Befragungsvolumen

3) Der Chi-Quadrat-Test ergab auf der Zusammenfassung benachbarter Größenklassen zu einer 3x3-Tabelle einen Wert von 1,41. Bei einer Irrtumswahrscheinlichkeit von 5 % hätte für die Annahme eines statistischen Zusammenhangs dieser Wert bei mindestens 9,49 liegen müssen.

1. Anzahl der Interviewer/-innen nach Zahl der Jahre ihrer Interviewertätigkeit und der betreuten Auswahlbezirke im Jahr 2007								
Im Jahr 2007 betreute Auswahlbezirke	Interviewer/-innen							
	insgesamt	davon mit einer Interviewertätigkeit von ... bis unter ... Jahren						
		unter 5	5 – 10	10 – 15	15 – 20	20 – 25	25 – 30	30 und mehr
unter 15	17	3	5	2	3	1	1	2
15 bis unter 20	35	5	7	6	10	6	–	1
20 bis unter 25	91	6	28	15	20	11	8	3
25 bis unter 30	92	8	26	15	15	15	6	7
30 bis unter 35	30	5	4	7	8	3	1	2
35 bis unter 40	17	4	4	2	2	3	1	1
40 und mehr	21	5	6	3	3	–	2	2
Insgesamt	303	36	80	50	61	39	19	18

nur der Zeitraum der Zugehörigkeit zum Kreis der aktiven Mikrozensusinterviewer in die Gewichtung eingeflossen.

Wegen der geringeren Bedeutung weit zurückliegender Jahre verbietet es sich, die Zahl der aktiven Jahre in vollem Umfang in die Schätzbasis einfließen zu lassen. Deshalb wurde die Gesamtzahl der Basishaushalte für jede(n) Interviewer/-in nach der in der Übersicht 3 enthaltenen For-

mel bestimmt. Durch die Formel wird der Einfluss der Zahl der aktiven Jahre dahingehend gedämpft, dass erst eine achtjährige Interviewertätigkeit zu einer Verdoppelung des Gewichtes gegenüber einem Anfänger mit der gleichen Zahl an Auswahlbezirken führt. Der konstante Faktor 3,5 bewirkt, dass die Summe der insgesamt und für den einzelnen Interviewer nur fiktiven Haushaltszahlen mit rund 60 000 Haushalten das Niveau eines Jahresbestandes erreicht.

Hauptmotiv für die Einführung eines konstanten Faktors in die Gewichtung war aber, den Einfluss der individuellen Interviewerdaten nicht zu groß werden zu lassen. Nach dieser Berechnung ergab sich für die Antworten jedes Befragten ein mittleres Gewicht von 202 fiktiven Haushalten. Die Spanne reichte von 51 bis 478 Haushalten, letzteres aufgrund der fehlenden Beschränkung auf ein Beobachtungsjahr.

Insgesamt waren – auf die 11 Beispielfragen verteilt – 220 Häufigkeitsnennungen von jeder befragten Interviewerin bzw. jedem befragten Interviewer zu verarbeiten, von denen jeweils fünf in einer logischen Verbindung zueinander stehen. Unter der Voraussetzung, dass alle Interviewerinnen und Interviewer die 20 Einzelereignisse der Übersicht 1 für jede der 11 Beispielfragen nach einheitlichen Kriterien⁴⁾ gezielt zugeordnet hätten, wäre es möglich gewesen, die entstandenen Antwortmuster selbst für die Analyse der Ergebnisse zu nutzen. Bei gleicher Wortwahl und vergleichbaren Bezügen wäre theoretisch zu erwarten gewesen, dass in jeder Kombination von jeweils 5 Alternativen als Häufigkeit des Auftretens mindestens einmal „oft“ oder mindestens zweimal „manchmal“ genannt wird. Die restlichen Alternativen hätten dann auf eine der Angaben von „selten“ bis „nie“ entfallen müssen. Diese idealtypischen Kombinationen zeichnen sich dadurch aus, dass durch die Nutzung von mindestens 3 der 5 Antwortmöglichkeiten ein hoher Differenzierungsgrad erreicht ist und gleichzeitig ein geringerer Anpassungsbedarf der in der Übersicht 3 beschriebenen Gewichtung an die Summe der Basisgewichte besteht. Mit 8 700 Antwortkombinationen erfüllten aber nur 65 % der Antworten beide Kriterien. Auf der anderen Sei-

4) Eine provisorische Analyse der Antwortmuster hat unter anderem ergeben, dass die Bezüge für die Häufigkeitsnennungen variabel sind. So erklärt sich z. B. dass ein(e) Befragte(r) das Wort „selten“ angibt, weil die Antwort auf alle Interviews bezogen wird, während ein(e) Andere(r) mit der Antwort „manchmal“ oder „oft“ alle Fälle gedanklich ausklammert, bei denen die Fragestellung unzutreffend ist. Auf entsprechende Vorgaben wurde aber auch in diesem Zusammenhang bewusst verzichtet, um den Befragungsdruck zu senken.

Übersicht 3: Die Gewichtung der Interviewerantworten

Schätzung der Gesamtzahl der betroffenen Interviews:

Gesamtzahl der fiktiven Haushalte (Basishaushalte) des Interviewers i:

$$H_i = \sqrt[3]{J_i \cdot A_i \cdot F}$$

mit

A_i = Zahl der Auswahlbezirke des Interviewers i im letzten Jahr

J_i = Zahl der aktiven Jahre des Interviewers i

F = 3,5 als durchschnittliche Zahl der befragten Haushalte je Auswahlbezirk

Gesamtzahl der Basishaushalte:

$$H = \sum_{i=1}^n H_i$$

mit n = Anzahl der befragten Interviewer/-innen

Basisgewichte der angegebenen Häufigkeiten für die Alternativen $k = 1$ bis 5 :

$b_k = 0$, wenn Antwort „noch nie“

$b_k = 1$, wenn Antwort „nur einmal“

$b_k = 2$, wenn Antwort „selten“

$b_k = 20$, wenn Antwort „manchmal“

$b_k = 75$, wenn Antwort „oft“

Schätzung der von der Alternative k betroffenen Haushalte des Interviewers i

Nach Abzug der Haushalte mit direkt abzählbarer Häufigkeit:

$$H_i^* = H_i - \sum_{b_k \leq 1} b_k$$

ergibt sich aus der Berechnung:

$$H_{ik} = b_k \text{ wenn } b_k \leq 1$$

$$H_{ik} = \frac{H_i^* \cdot b_k}{\sum_{b_k > 1} H_i^* \cdot b_k} \cdot H_i^* \text{ wenn } b_k > 1$$

der Zähler für die relative Häufigkeit des Ereignisses k beim Interviewer i : $w_{ik} = \frac{H_{ik}}{H_i}$

te mussten aber nur 326 Antwortkombinationen (2,4 %) hingenommen werden, die weder hinsichtlich des Anpassungsbedarfs noch bezüglich des Differenzierungsgrades akzeptabel waren. Davon entfielen 86 Fälle auf die Kombination von „manchmal“ mit viermal „nie“. Diese Kombination ist problematisch, weil in dieser Kombination die Angabe „manchmal“ ein Gewicht von 100 % erhält. Obwohl es wenig wahrscheinlich ist, dass die von den Befragungspersonen nicht wahrgenommen Alternativen tatsächlich nicht vorgekommen sind, wurden auch diese Angaben in der Gewichtung unverändert berücksichtigt. Eine Veränderung dieser Kombination wäre auch nur durch eine Anpassung an die Antwortmuster anderer Interviewerinnen und Interviewer möglich gewesen. Dieser Weg war aber aus zwei Gründen nicht gangbar: Aus methodischer Sicht ist zweifelhaft, ob das Ergebnis einer Gewichtung durch die Modifizierung von Basisdaten vorweg genommen werden darf. Außerdem bestand die Gefahr, die mit nur 55 bereits geringe Zahl an unterschiedlichen Antwortmustern noch weiter zu reduzieren. Bei dieser Zählung der Antwortmuster wurden die Angaben „nie“ und „einmal“ nicht unterschieden, weil beide Angaben keinen direkten Einfluss auf die Quantifizierung der Klartextangaben haben.

Das in der Übersicht im Detail dargestellte Schätzverfahren zur Gewichtung der Klartextangaben nutzt im Wesentlichen die Antwortkombinationen und nicht die einzelne Antwort. Insofern ist der kritische Punkt dieser Interviewerbefragung und des zugehörigen Auswertungskonzeptes in der Häufigkeit des als Ein-Punkt-Verteilung in Erscheinung tretenden Antwortmusters „oft-nie-nie-nie-nie“ zu sehen. Dieses Antwortmuster erreichte einen Anteil von 17,5 % an allen Antwortkombinationen. Die Gewichtung der Antworten führt in dieser Kombination dazu, dass ein bestimmtes Einzelereignis allen befragten Haushalten einer Interviewerin oder eines Interviewers zugeordnet wird. Da dies aber nur in wenigen Fällen als richtig anzunehmen

ist, zeigt sich eher die Grenze der Belastbarkeit des Erinnerungsvermögens der Befragten. Hier werden die in der Fragestellung angebotenen Differenzierungsmöglichkeiten nicht mehr genutzt. Die gegenteilige Möglichkeit, den Befragungsdruck durch eine pauschale Antwort zu reduzieren oder auch um unsichere Festlegungen zu vermeiden, war, allen Alternativen eines Themas die gleiche Häufigkeit beizumessen. Von dieser Möglichkeit wurde dagegen nur sehr selten (18 Einzelfälle) und fast ausschließlich im Zusammenhang mit der Beurteilung von Fremdauskünften Gebrauch gemacht. Für die gewichteten Ergebnisse bedeutet das deutlich häufigere Auftreten von Ein-Punkt-Verteilungen im Vergleich zu den wenigen Gleichverteilungen, dass seltener vorkommende Ereignisse in der Gewichtung immer noch unterschätzt werden. Die mit diesem Antwortmuster verbundene Betonung des Fehlens von Komplikationen bezog sich allerdings fast ausschließlich auf die Abfragen zur Akzeptanz und Genauigkeit. Gegen eine bewusste Umgehung von differenzierten Antworten spricht die Tatsache, dass nur 7 % der Ein-Punkt-Verteilungen auf das Thema Bearbeitungsaufwand entfielen. Dies ist umso bemerkenswerter, als in diesem Ereignisraster der direkte Bezug zum Verhalten der Interviewer nicht ganz verborgen werden konnte. In den übrigen Abfragen befanden sich die Interviewerinnen und Interviewer dagegen in der Rolle neutraler Beobachter, so dass anzunehmen ist, dass in diesem Umfang tatsächlich keine Komplikationen erinnert wurden. Hinzu kommt, dass sich mehr als 50 % der 100 %-Aussagen⁵⁾ auf nur 4 eher unscheinbare Beispielfragen bezogen. Dies waren die Fragen zum Familienstand, zur Art der Schule, dem Land der Geburt und zur Suche nach einer anderen Tätigkeit.

Trotz der großen Abhängigkeit der Schätzungen von den Antwortmus-

5) Wenn nur eine der Alternativen mit „oft“ und alle anderen mit „nie“ bezeichnet wurden, ergibt sich aus der Gewichtung der Antworten, dass diese Alternative bei der einzelnen Interviewerin bzw. dem Interviewer immer aufgetreten ist.

tern mussten einige Korrekturen und Ergänzungen vor der Gewichtung der Antworten vorgenommen werden. Der Einfluss auf die Ergebnisse dürfte aber nicht sehr groß sein. In 285 Einzelfällen fehlende Häufigkeitsangaben der Interviewerinnen und Interviewer wurden im Rahmen der Plausibilitätsprüfungen auf den Wert gesetzt, der von allen übrigen Interviewer(inne)n am häufigsten angegeben war. Ein Verzicht auf diese Setzungen hätte bedeutet, dass alle übrigen Angaben der zugehörigen Abfrage nicht mehr nutzbar gewesen wären. Zum Beispiel fehlte bei zwei Interviewern nur die Antwort auf die Frage, wie häufig sie die Frage nach dem Familienstand im Wortlaut vorlesen. Da die Rohdaten bereits in 269 Fällen die Angabe „oft“ und in nur 25 Fällen die Angabe „manchmal“ aufwiesen, wurde hier der Wert „oft“ eingesetzt. Offensichtliche Irrtümer oder möglicherweise auch Übertragungsfehler waren in 670 Einzelfällen zu korrigieren. Um die Antwortmuster mit Rücksicht auf die Schätzergebnisse so wenig wie möglich zu verändern, wurde bei diesen Korrekturen nicht die häufigste Nennung, sondern die in der Sortierung am nächsten liegende plausible Häufigkeit eingesetzt. Zum Beispiel wurde eine unplausible Nennung „nie“ nicht auf den häufigsten Wert „oft“ gesetzt, wenn der näher benachbarte Wert „manchmal“ ebenfalls als plausibel angenommen werden musste. Dabei wurden aber nur Werte verändert, wenn weniger als 10 Interviewer/-innen die gleiche Angabe gemacht hatten. Einschließlich der Schätzung unbekannter Werte wurden demnach 955 Korrekturen durchgeführt. Das entspricht einer Korrekturquote von 1,4 % der Angaben aus dem Blickwinkel der isolierten Nennung. Durch die Plausibilisierung wurde nur die relativ neutrale Nennung „manchmal“ etwas erhöht.

Insgesamt sollte mit den vorstehenden Detailbeschreibungen verdeutlicht werden, dass die entstandenen Antwortmuster als Basis für die Schätzung von relativen Häufigkeiten ausreichen, weil durch die Ge-

wichtung im Zusammenhang einige im Detail vorhandene Unschärfen ausgeglichen werden können. Diese Unschärfen hätten sich bei einer direkten Auswertung der Antwortmuster⁶⁾ unangenehm bemerkbar machen können. Außerdem sollte gezeigt werden, dass die notwendigen Plausibilisierungen nicht zur Vorwegnahme von Ergebnissen geführt haben. Um diesbezügliche Einwände noch etwas weiter zu entkräften, ist darauf hinzuweisen, dass das gesamte Schätzverfahren einschließlich der Gewichtung der Antworten bereits vor der Befragung festgelegt und unverändert beibehalten wurde. Auch wenn im Folgenden die Schätzergebnisse zur Vereinfachung der Darstellungen nur noch als Ergebnisse bezeichnet werden, bedeutet dies aber nicht, dass der Unterschied zwischen einer direkten empirischen Messung und der aus einer Expertenbefragung abgeleiteten Schätzung aus dem Auge verloren wurde.

Ergebnisse zum Befragungsaufwand und zur Akzeptanz

Über alle Beispielfragen gerechnet zeigte sich nach der im vorigen Abschnitt dargestellten Gewichtung, dass im Durchschnitt in nur 62,5 % der Fälle die Frage vorgelesen wurde, ohne dass zusätzliche Beispiele gegeben oder Rückfragen der befragten Haushalte beantwortet werden mussten. Zusätzliche Erläuterungen oder Beispiele waren in 14,3 % der Fälle erforderlich, die bei weiteren 3,3 % der Fälle erst nach der Beantwortung zusätzlicher Rückfragen der befragten Haushalte ausreichten. Die über das Vorlesen der Frage im Wortlaut hinausgehenden Befragungsformen werden im Folgenden unter der Bezeichnung „individualisierte“ Befragung

6) Denkbar war auch eine direkte Analyse der Antwortmuster z. B. über Ähnlichkeitsmessungen. Dabei wären aber alle Unterschiede zwischen Bezugsgrößen und dem Antwortverhalten von Interviewer/-in zu Interviewer/-in zu berücksichtigen gewesen. Erst die nivellierenden Effekte der Gewichtung stellen sicher, dass zwei gleiche Häufigkeitsbezeichnungen von verschiedenen Interviewern nicht unbedingt auch das Gleiche beinhalten müssen, um zu brauchbaren Ergebnissen zu kommen.

gung⁷⁾ zusammengefasst. Die Zusammenfassung der Wiedergabe der Frage mit eigenen Worten mit dem Geben von Beispielen wurde aus zwei Gründen vorgenommen. Zum einen erhöht sich im Vergleich zum einfachen Vorlesen der überwiegend kurzen Fragen der Bearbeitungsaufwand. Zweitens muss man davon ausgehen, dass ein Beispiel eine ähnlich steuernde Wirkung entfalten kann, wie eine umformulierte Frage. Die Erläuterung von Fragen mit eigenen Worten, das Geben von Beispielen und erst recht die Beantwortung zusätzlicher Rückfragen deutet darauf hin, dass schon wegen des erhöhten Zeitaufwandes diese Steuerung des Ablaufs nicht ohne entsprechenden Bedarf bei den Befragten erfolgt. Stattdessen ist davon auszugehen, dass auch bei den einzelnen Beispielfragen ein hoher Anteil der individualisierten Befragung als besonderes Bemühen der Interviewerinnen und Interviewer um eine richtige Antwort zu interpretieren ist.

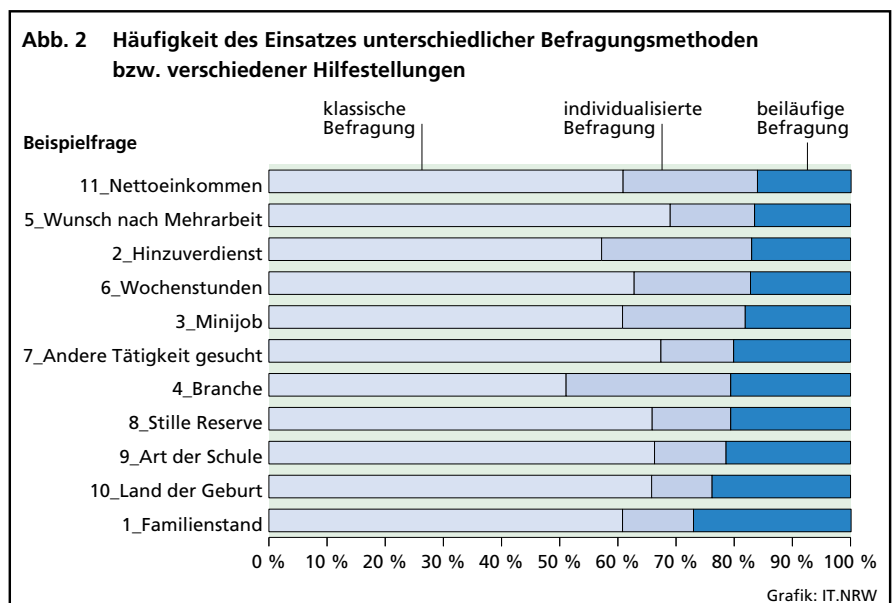
Auf der anderen Seite ließen sich in 17,4 % der Fälle die Interviewer/-innen die erkennbare Antwort nur von den Befragten bestätigen. In weiteren 2,5 % der Fälle hatten die Haushaltsmitglieder die Frage vor der Beantwortung selbst vom Bildschirm abgelesen. Weil bei beiden

7) Durch die Wortwahl soll auch zum Ausdruck gebracht werden, dass selbst notwendige Abweichungen vom Wortlaut der Frage die Gefahr der Abweichung von den methodischen Vorgaben in sich bergen.

Formen der Befragung die Frage nicht gestellt wird, werden sie unter der Bezeichnung „beiläufige“ Befragung zusammengefasst. Die Interpretation dieser Formen der Befragung kann aber besser anhand der einzelnen Beispielfragen entwickelt werden.

Von allen Beispielfragen wurde die Frage nach dem Familienstand mit 25,5 % der (fiktiven) Haushalte am häufigsten beantwortet, ohne dass die Frage gestellt wurde. Hinzu kam in 1,6 % der Fälle, dass die befragte Person die Frage selbst vom Bildschirm abgelesen hatte. Bei den übrigen Beispielfragen lag diese Form der Beantwortung nur bei der Frage zum Nettoeinkommen (5,8 %) deutlich über dem Durchschnitt von 2,5 %. Dieser verhältnismäßig hohe Anteil erklärt sich vermutlich aus der sonst für diese Frage üblichen Vorlage einer Antwortliste in Papierform. Im Vergleich zum expliziten Vorlesen der Frage ist die Interviewerin bzw. der Interviewer zwar nicht unbedingt weniger aktiv, aber in beiden Situationen wird die Frage den Befragten doch eher beiläufig vermittelt.

Mit Blick auf die Effizienz-These ist davon auszugehen, dass nur ein willkürliches Abweichen von der klassischen Befragung zur beiläufigen Befragung geeignet wäre, Interviews zu verkürzen. Sobald eine Schlüssel- oder Filterfrage explizit gestellt wurde, kann nur noch die befragte Per-



son den weiteren Verlauf des Interviews bestimmen. Aus zusätzlichen Erläuterungen, der Beantwortung von Verständnisfragen und möglicherweise auch aus (theoretisch suspekten) Umformulieren ist eher zu schließen, dass der Fragensteller sich bemüht hat, eine objektiv richtige Antwort auch dann zu erhalten, wenn sie Folgefragen nach sich zieht. Von den Beispielfragen sind davon nur die Minijob-Frage⁸⁾ und die Hinzuverdienst-Frage⁹⁾ betroffen. Da bei der zuletzt genannten Frage fast schon etwas suggestiv anmutende Elemente der individualisierten Befragung in den Wortlaut der Frage integriert worden sind, ist es bemerkenswert, dass trotzdem ein vergleichsweise kleiner Anteil (57,2 %) der Befragungen auf das einfache Vorlesen der Frage beschränkt war. Gemessen an der Befragungsintensität rangiert diese Frage mit einem Anteil von 22,8 % an zusätzlichen Erläuterungen und weiteren 3 % an Beantwortung von Verständnisfragen auf dem 3. Rang hinter dem Wunsch nach Mehrarbeit und dem Nettoeinkommen.

Da – abgesehen von der Frage nach dem Nettoeinkommen – das Ablesen der Frage vom Bildschirm bei allen Fragen kaum vorkommt, ist etwas vereinfachend davon auszugehen, dass im Durchschnitt in rund 20 % der Fälle die Antwort benannt wird, noch bevor die Frage ausgesprochen war. Zulässig ist dies, wenn offenkundige Konstellationen aus dem Befragungszusammenhang erkennbar sind. Zum Nettoeinkommen und insbesondere zu den Filterfragen zur Erwerbstätigkeit ist anzumerken, dass diese Konstellationen zwar nicht so häufig sind wie beim Familienstand, aber auch für fehlendes Einkommen¹⁰⁾ oder den fehlenden

Wunsch nach Mehrarbeit kann es Anhaltspunkte geben, die ein mechanisches Vorlesen der Frage nicht angebracht erscheinen lassen. Vor diesem Hintergrund überrascht es nicht, dass Interviewerinnen die Frage nach dem Familienstand in 31,6 % der Fälle nicht stellen, während Männer nur zu 24,1 % auf das Vorlesen der Frage verzichten. Auch bei den Fragen zum Land der Geburt (Differenz: 5,1 Prozentpunkte) und der Art der Schule (Differenz: 5,0 Prozentpunkte) sind deutliche Unterschiede zwischen Interviewerinnen und Interviewern zu sehen, zumal in dieser groben Zusammenfassung nivellierende Effekte nicht auszuschließen sind. Überraschend ist hingegen, dass dieser Unterschied zwischen Interviewerinnen und Interviewern beim Nettoeinkommen (0,3 Prozentpunkte) und bei der Frage nach Hinzuverdiensten (0,7 Prozentpunkte) fast nicht vorhanden ist. Ähnliche Abweichungen und Übereinstimmungen zeigen sich auch im Vergleich der älteren Interviewerinnen und Interviewer mit den Jüngeren. Das bedeutet, dass der Spielraum für beiläufige Befragungen, der sich sowohl aus konkreten Einzelbeobachtungen als auch aus den Wiederholungsbefragungen¹¹⁾ ergibt, bei der Frage zum Hinzuverdienst fast genau so wenig genutzt wird wie beim Nettoeinkommen. Dies ist zwar noch keine Widerlegung der Effizienz-These, aber bereits ein deutliches Zeichen dafür, dass diese Frage im Interview stark betont wird. Die Beispielfrage zum Hinzuverdienst erreichte außerdem mit 22,8 % den höchsten Anteil an Erläuterungen und Beispielen. Die Erläuterung von Fragen mit eigenen Worten oder das Geben von Beispielen verlängert das Interview. Deshalb ist in dieser Beobachtung das genaue Gegenteil der in der Effizienz-These unterstellten Tendenz zur Verkürzung der Interviews zu sehen. Noch bevor untersucht wird, inwieweit das Verhalten der Interviewerinnen und Interview-

11) Die Haushalte werden im Mikrozensus im Jahresabstand viermal hintereinander befragt. Bei den Wiederholungsbefragungen in den Folgejahren müssen mit Rücksicht auf die räumliche Nähe der Auswahlbezirke häufig die gleichen Interviewer/-innen eingesetzt werden, wie im Vorjahr.

wer auch von den Reaktionen der Befragten abhängt, ist bereits festzuhalten, dass bei den Schlüssel- und den Filterfragen, die Gegenstand der Effizienz-These waren, Erfahrung und Intuition nicht genutzt wurden, um von einer formellen und teilweise sogar eindringlichen Befragung abzuweichen.

Aus der Gewichtung der Antworten zum Thema Akzeptanz der Fragen bei den Haushalten ergab sich, dass im Durchschnitt nur 2 von 1 000 Haushalten eine der Fragen zum Anlass nehmen, das Interview abubrechen. Relative Häufungspunkte ergaben sich bei der Frage zum Einkommen (1,6 %) und zum Familienstand (0,2 %).

Die gegenüber allen übrigen Beispielfragen höhere Abbruchquote beim Familienstand ist vermutlich dadurch bedingt, dass diese Frage als Repräsentant für den Beginn des Interviews genommen wurde. Wenn man berücksichtigt, dass für alle Beispielfragen eine Auskunftspflicht besteht, muss man davon ausgehen, dass Kritik an der Fragestellung bereits mit einer Ablehnung gleichzusetzen ist. Deshalb wurden in Abbildung 3 alle abgefragten Formen der Kritik bis zur völligen Ablehnung zusammengefasst. Im Mittel ergab sich über alle Beispielfragen ein Anteil von 6,5 %. Bezüglich weiterer 11,8 % der Antworten wurde von den Interviewerinnen und Interviewern berichtet, dass sich das Interview durch zusätzlichen Erläuterungsbedarf zum Zweck der Frage verzögert hatte. Dabei wurde bewusst nicht unterschieden, ob dieses Hinterfragen der betroffenen Haushalte als Ausdruck allgemeinen Interesses oder bereits als Zweifel am Zweck der Frage zu werten ist. Insofern ist dieser Anteil der Reaktionen der Haushalte auf die Fragen als der Anteil der Haushalte zu sehen, die weder ohne zu zögern geantwortet haben, noch die gestellten Fragen direkt kritisiert haben. Unter dem Blickwinkel der Akzeptanz wiesen das Haushaltseinkommen (24,9 %) und die Branche (23,0 %) die höchsten Anteile in diesem Bereich (Situation G) auf. Beim Einkommen deutet

8) Haben Sie in der letzten Woche (zusätzlich) eine geringfügige Beschäftigung, Minijob oder einen Ein-Euro-Job ausgeübt? – 9) Auch wenn man eigentlich nicht erwerbstätig ist, kann man ja z. B. als Schüler/-in, Hausfrau/Hausmann oder Rentner/-in Tätigkeiten ausüben, mit denen man Geld hinzuverdient. Wie ist das bei Ihnen: Üben Sie regelmäßig oder gelegentlich eine solche bezahlte Tätigkeit aus? – 10) Im Mikrozensus ergibt sich das Haushaltsnettoeinkommen aus der Summe der persönlichen Einkommen der Haushaltsmitglieder. Deshalb müssen Geldtransfers (z. B. Taschengeld) innerhalb des Haushalts nicht berücksichtigt werden.

bereits der hohe Anteil der direkten Kritik (31,3 %) darauf hin, dass auch die Nachfragen zum Zweck der Frage bereits in hohem Maße skeptisch motiviert sind. Bei der Frage zum Wirtschaftszweig der Arbeitsstätte bezieht sich die Verzögerung der Antwort vermutlich eher auf die vermeintliche Redundanz, die sich aus der unmittelbar vorausgegangenen Frage nach der ausgeübten Tätigkeit bzw. nach dem Beruf ergeben könnte. Auch der Sachverhalt, dass im Zusammenhang mit der Ermittlung des Wirtschaftszweigs der Arbeitsstätte der Name des Betriebs als Hilfsmerkmal für die Zuordnung abgefragt wird, könnte den Erläuterungsbedarf zum Zweck der Frage erhöht haben. Aus dem Themenkreisen Erwerbstätigkeit und Arbeitssuche war die an Erwerbstätige gerichtete Frage zur Suche einer anderen oder weiteren Tätigkeit die einzige Beispielfrage, die in mehr als 90 % der Fälle ohne Zögern beantwortet wurde. Die hohe Akzeptanz dieser Frage erklärt sich aus der geringen Belastung der Befragten, weil zumindest die überwiegende Mehrheit aller nachhaltig Erwerbstätigen diese kurze Frage einfach verneinen kann. Entsprechend gering war der Anteil (nur 12,5 % im Vergleich zu 25,8 % beim Hinzuverdienst) an Erläuterungen und anderen Formen der individualisierten Befragung. Alle übrigen Fragen zur Erwerbstätigkeit weisen eine spürbar geringere Akzeptanz auf als die Beispielfragen zum Familienstand, dem Land der Geburt oder der Art der Schule.

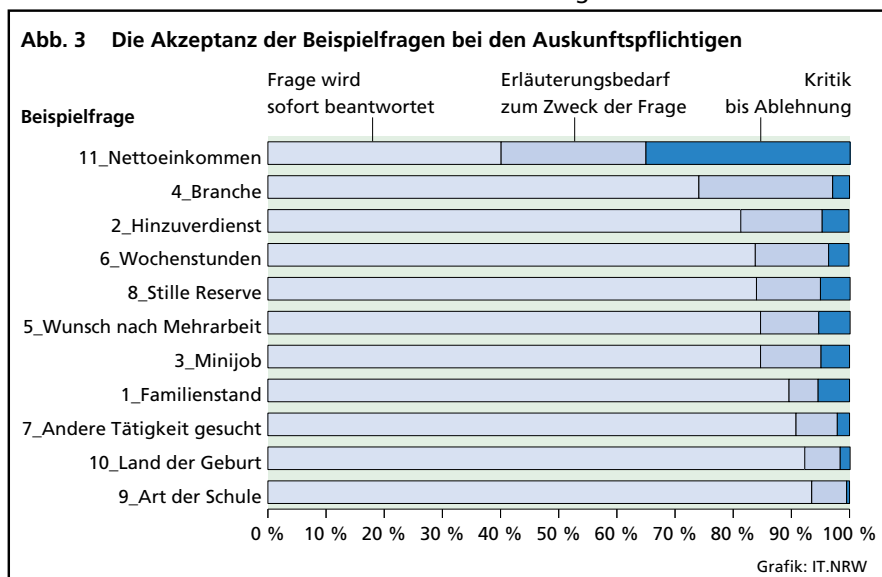
Aus vielen Einzelberichten von Interviewerinnen und Interviewern war bekannt, dass Akzeptanzprobleme nicht gleichmäßig in der Fläche verteilt, sondern in Abhängigkeit von der Bebauungsdichte und der Sozialstruktur der Bewohner in bestimmten Nachbarschaften gehäuft auftreten. Deshalb waren vor der Befragung an die Auswertung der Ergebnisse in der Gliederung nach der Lage der Auswahlbezirke hohe Erwartungen geknüpft worden. Im Vergleich der Akzeptanz der Fragen mit der Lage der Auswahlbezirke zeigte sich aber, dass in dieser Untergliederung die zur eindeutigen Wi-

derlegung der Effizienz-These erforderliche Trennschärfe der Häufigkeitsverteilungen nicht erreicht wurde. Gemessen an dem Anteil der geschätzten Interviews mit kritischen bis ablehnenden Reaktionen der Haushalte war zwar überwiegend das erwartete Stadt-Land-Gefälle sichtbar, aber das Maximum der Ablehnung der Frage zum Familienstand¹²⁾ (6,9 %) war nicht in den Großstädten, sondern in Kleinstädten anzutreffen. Auch bei den Fragen zur Branche und zum Minijob war die Reihenfolge umgekehrt. Ein minimaler Widerstand bei diesen Fragen wurde von Interviewer(inne)n gemeldet, die schwerpunktmäßig Auswahlbezirke in Großstädten betreuen. Der größte Widerstand gegen die Frage zur Branche (3,8 %) entfiel ebenfalls auf Kleinstädte. Die Minijob-Frage wurde mit einem Anteil von 6,1 % in Mischgebieten am häufigsten abgelehnt. Diese Beispiele zeigen, dass eine differenziertere Abfrage der Struktur der Auswahlbezirke erforderlich gewesen wäre, um die Abhängigkeit der Interviewverläufe vom Verhalten der Interviewer/-innen entscheidend relativieren zu können. Der Sachverhalt, dass die Interviewerinnen und Interviewer nicht

12) Einige Auffälligkeiten der Frage zum Familienstand, die ursprünglich als relativ problemfreie Frage nur der Justierung der Struktur der übrigen Beispielfragen dienen sollte, könnten darauf zurückzuführen sein, dass diese Frage im Jahr 2006 um die eingetragene Lebensgemeinschaft und im Jahr 2007 um verstorbene Lebenspartner und aufgehobene Lebensgemeinschaften erweitert wurde.

aus freiem Willen, sondern durch die Reaktionen der Haushalte gezwungen von der klassischen Befragung abweichen, ist trotzdem nicht zu übersehen.

In Abbildung 3 sind die Beispielfragen in der Reihenfolge der Gesamtakzeptanz, wie sie sich durch zunehmende Anteile an direkten und unverzögerten Antworten der Haushalte den Interviewerinnen und Interviewern darstellt, aufsteigend geordnet. Davon ausgehend, dass bei Erhebungswiderständen das reine Vorlesen der Frage nicht mehr genügt und durch individualisierte Formen der Befragung ersetzt werden muss, sollte sich die Rangfolge der in Abbildung 2 enthaltenen individualisierten Befragung in einer ähnlichen Rangfolge bewegen. Die größte Abweichung von der Rangfolge zeigt sich bei der Minijob-Frage. Hinsichtlich des Befragungsaufwands nahm diese Frage den 4. Rang, in Bezug auf fehlende Akzeptanz dagegen nur den 7. Rang ein. Kleinere Rangdifferenzen der gleichen Richtung zeigten sich bei der Art der Schule (+2), der Branche (+1), der Suche nach einer anderen Tätigkeit (+1) und dem Hinzuverdienst (+1). Während bei der Art der Schule und der Suche nach einer anderen Tätigkeit die Änderung der Rangfolge auf der Basis geringer Akzeptanzprobleme wenig aussagt, bedeutet diese Änderung bezüglich der Fragen zur Erwerbstätigkeit als wichtigste Maßstäbe für die Effizienz-These, dass diese Fragen von den Interviewerinnen



nen und Interviewern stärker in der Befragung betont wurden, als dies bei einer Frage geschehen wäre, die im Einvernehmen mit den befragten Haushalten die Reduktion von Folgefragen zum Ziel gehabt hätte.

Auch für den zweiten Teil der Effizienz-These, nach der die Haushalte ein Interesse an der Verkürzung des Interviews haben, sind Zweifel angebracht. Die Frage zum Hinzuverdienst als Auslöser weiterer Fragen weist zwar eine vergleichsweise geringe Akzeptanz bei den befragten Haushalten auf, aber die Beispielfragen zur Branche und zum Nettoeinkommen zeigen, dass die Motive für eine geringere Akzeptanz der Fragestellung eher im Schwierigkeitsgrad oder in der beigemessenen Sensibilität zu suchen sind, als in dem Wunsch, das Interview zu verkürzen. Obwohl – schon wegen der langen Dauer des Mikrozensus-Interviews – von der Existenz dieses Wunsches der befragten Haushalte auszugehen ist, gibt es aus dem Blickwinkel der Akzeptanz noch keine für die Interviewerin oder den Interviewer sichtbaren Anhaltspunkte dafür, dass die Befragten versuchen, dieses Ziel durch bewusst falsche Antworten auf die entscheidenden Filterfragen zu erreichen.

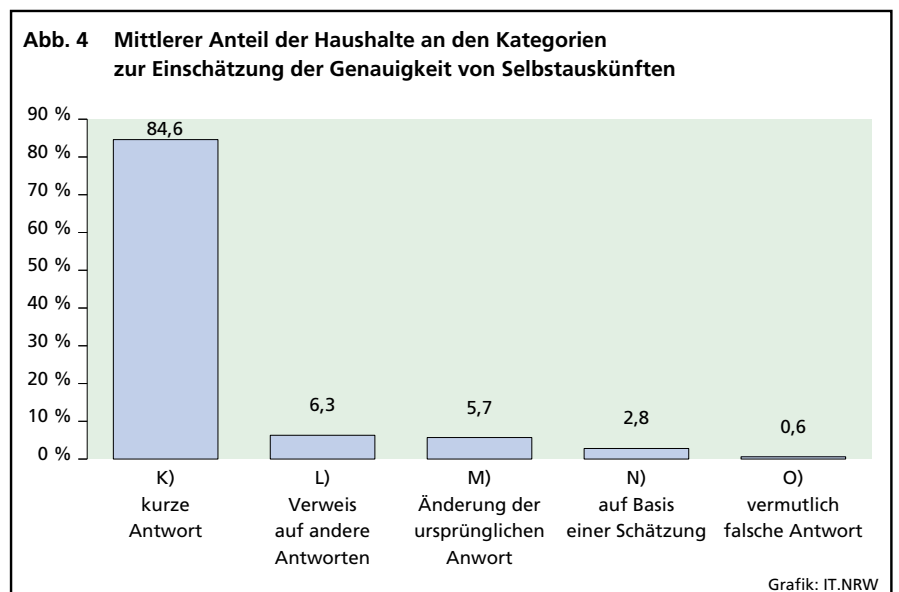
Ergebnisse zur Genauigkeit der Selbstauskünfte

Auch für den Interviewer oder die Interviewerin ist eine falsche Antwort erst dann als solche erkennbar, wenn sie im direkten Widerspruch zu anderen Antworten steht. Anders als bei postalischen oder elektronischen Befragungen besteht im persönlichen Interview (face-to-face) aber die Möglichkeit, diese Fehler noch vor der Speicherung der Antworten zu korrigieren, indem die befragte Person durch eine Nachfrage oder zusätzliche Erläuterung dazu bewegt wird, die ursprünglich gegebene Antwort zu verändern. Auch wenn danach die falsche Antwort nicht mehr in den Daten in Erscheinung tritt, ist die Häufigkeit dieser Situation (M in Übersicht 1) von großem Interesse, weil sie nicht Erhebungswiderstände, sondern Unkenntnis bei

den Befragten oder – was ebenfalls nicht aus dem Blickfeld rücken sollte – Probleme in der Fragestellung signalisiert. Im Vergleich mit einer sofort verstandenen und kurz beantworteten Frage ist die Änderung der ursprünglich gegebenen Antwort ähnlich zu bewerten wie der Hinweis auf eine bereits gegebene Antwort. Nur die Gründe für die Einschätzung einer geringeren Genauigkeit der Antworten sind unterschiedlich: In der ersten Situation sind Einflüsse der Interviewerinnen und Interviewer auf die Antwort wirksam. Die zweite Situation führt erst dann zu Fehlern, wenn die vom Haushalt und der Befragungsperson wahrgenommene Redundanz nur scheinbar vorhanden ist. Größere Einschränkungen der Genauigkeit sind zu erwarten, wenn die Befragungsperson den Haushalt ermuntern muss, eine Schätzung abzugeben, weil auf anderem Wege keine Antwort zu bekommen ist. Aber auch hier kann man nicht davon ausgehen, dass alle Schätzergebnisse falsch sind. Die Aufforderung zu einer Schätzung kann z. B. dann notwendig werden, wenn der Befragte selbst zu hohe Anforderungen an die Genauigkeit seiner Antwort stellt, um nicht antworten zu müssen. Für die Abbildung 4 wurden die gewichteten Ergebnisse für jede Beispielfrage addiert und durch die Zahl der Fragen geteilt. Ungeachtet der möglicherweise nur bedingt vorhandenen Repräsentativität der Beispielfragen für das gesamte Frageprogramm des

Mikrozensus wird die dargestellte Verteilung in den folgenden Detailbeschreibungen als den Durchschnitt bezeichnenden Maßstab genutzt werden. Die Hinnahme einer vermutlich falschen Antwort wies nach dieser Nivellierung nur noch einen Anteil von 0,6 % auf. Eine Unterschätzung dieser Quote könnte damit zusammenhängen, dass die befragten Interviewerinnen und Interviewer die Antwortvorgaben nicht nur als Indikator für die Einschätzung der Genauigkeit, sondern – was nicht beabsichtigt war – auch als implizite Abfrage ihrer Bemühungen um eine richtige Antwort interpretiert haben.

Der Verweis auf bereits gegebene Antworten zu einer ähnlichen Frage erreichte mit jeweils 7,4 % der Reaktionen der Haushalte überdurchschnittlich hohe Anteile beim Familienstand und dem Wirtschaftszweig. Zu beiden Beispielfragen gibt es im Frageprogramm mit dem Verwandtschaftszusammenhang oder dem ausgeübten Beruf zwar eng verbundene andere Fragen, aber ein Blick auf die Quoten der auf Betreiben der Interviewerinnen und Interviewer geänderten Antworten zeigt, dass dieser Antworttyp vor einem völlig anderem Hintergrund zu interpretieren ist. Während beim Familienstand nur 3,7 % der ursprünglich gegebenen Antworten durch Nachfragen oder zusätzliche Erläuterungen modifiziert wurden, erreichte diese Quote mit 11,7 % beim Wirt-



schaftszweig sogar den höchsten Wert aller Beispielfragen. Das bedeutet, dass der Verweis auf andere Antworten für sich allein genommen noch keine wesentliche Einschränkung der Genauigkeit signalisiert. Erst in Verbindung mit den übrigen Einschätzungen wird erkennbar, wie die Genauigkeit der so gewonnenen Antworten zu bewerten ist.

Unter diesem Blickwinkel ist die Frage nach dem Arbeitswunsch ohne Arbeitssuche (Stille Reserve) ähnlich zu sehen wie die Frage nach dem Familienstand. Bei beiden Fragen findet nur der Verweis auf andere Antworten überdurchschnittlich oft statt, während alle übrigen Indikatoren für eine eingeschränkte Genauigkeit unterdurchschnittlich ausgeprägt sind. Bei beiden Fragen ergibt sich aus den hohen Anteilen der beiläufigen Befragung außerdem, dass in der Einschätzung einer gewissen Redundanz dieser Fragen Befragte und Befragungspersonen übereinstimmen. Völlig anders liegen die Verhältnisse bei den Fragen zum Minijob und zum Hinzuverdienst. Obwohl bei diesen Fragen mit rund 10 % am häufigsten statt einer direkten Antwort ein Verweis auf andere Antworten erfolgte, wurden diese Reaktionen von den Interviewerinnen und Interviewern nicht mitgetragen, denn die beiläufige Befragung wies bei diesen Fragen sehr geringe Anteile auf. Zudem wurden 6,9 % der Antworten zum Hinzuverdienst sowie 6,3 % der Antworten zum Minijob noch während der Interviews verändert. Die besondere Problematik der Frage zum Hinzuverdienst ist auch daran erkennbar, dass nur bei dieser Frage und beim Nettoeinkommen überdurchschnittlich hohe Anteile an vermutlich falschen Antworten angenommen werden mussten.

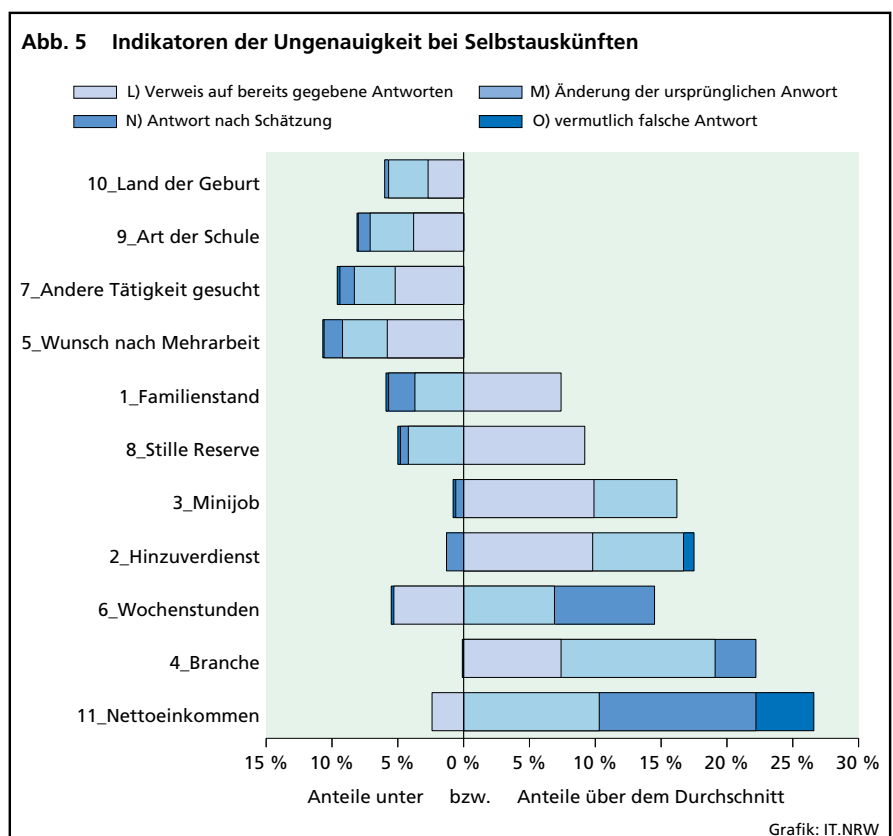
Wie in Abbildung 5 zu sehen ist, nimmt die Frage zum Hinzuverdienst schon bei Selbstauskünften hinsichtlich der Summe aller Indikatoren der Ungenauigkeit den 4. Rang hinter dem Nettoeinkommen, der Branche und der Zahl der Wochenstunden ein. Dabei ist nicht zu übersehen, dass bei den drei zuletzt genannten

Fragen überdurchschnittlich oft geschätzte Antworten berücksichtigt werden mussten. Bei einer einfachen Alternative wie der Frage nach dem Hinzuverdienst fehlt dagegen weitestgehend der Anlass für eine Schätzung. Deshalb sind die Ursachen für die geringere Genauigkeit nicht in der besonderen Schwierigkeit der Frage, sondern in der geringeren Akzeptanz zu suchen. Diese wurde in den Abfragen zur Akzeptanz nur deshalb nicht ganz so deutlich, weil sie von den Akzeptanzproblemen der Fragen zum Nettoeinkommen und zur Branche überragt wurde. Die Tendenz zur Untererfassung marginaler Beschäftigungsverhältnisse im Mikrozensus wird damit auch aus diesem Blickwinkel teilweise bestätigt, aber gleichzeitig zeigen die Ergebnisse, dass dieser Effekt nicht durch die Feldarbeit selbst verstärkt worden ist. Stattdessen ist anzunehmen, dass auch weitere Erläuterungen zum Zweck der Frage und die Intensivierung ähnlicher Bemühungen die Erhebungswiderstände gegen diese Frage nicht mehr nennenswert verringern können.

Es ist davon auszugehen, dass die Wahrscheinlichkeit für eine richtige

Antwort im Ergebnis aus einem Verweis auf eine andere Antwort und insbesondere nach Änderung der Antwort noch vergleichsweise groß ist. Auf der Gegenseite ist damit zu rechnen, dass im Einzelfall besondere Reaktionen auf eine Frage nicht wahrgenommen wurden. Deshalb ist der in der grafischen Darstellung vernachlässigte Anteil derjenigen Haushalte, die in der Selbstauskunft die Fragen klar und glaubhaft beantwortet haben, als sichere Untergrenze des Anteils objektiv richtiger Antworten anzusehen. Beim Nettoeinkommen beträgt dieser aus der Gewichtung errechnete Anteil in der Selbstauskunft nur 71,1 %. Da die Interviewerinnen und Interviewer bei dieser Frage die Antwort am häufigsten (10,1 %) erst nach Aufforderung zu einer Schätzung erhalten haben, scheint die besonders kritische Bewertung der Einkommensfrage bei beiden Beteiligten nicht zu berücksichtigen, dass es in den meisten Fällen für die Zuordnung zur richtigen Einkommensklasse nicht auf den einzelnen Euro ankommt.

Nach der in diesem Beitrag bereits am Anfang vorgenommenen Relati-



vierung der Zuordnung zur Selbstauskunft, der ebenfalls notwendigen Relativierung der Quoten hinsichtlich ihrer Bezugsgrößen und dem Hinweis, dass die zu den einzelnen Aspekten abgefragten Situationen weder überschneidungsfrei noch additiv sind, macht sich an dieser Stelle zusätzlich bemerkbar, dass bei den Genauigkeitsbetrachtungen auch die Beurteilungsmaßstäbe für die Einschätzungen von Frage zu Frage variieren. Eine gewisse Überzeichnung des Ausmaßes der Ungenauigkeit kann sich auch aus der Fragestellung selbst ergeben. Dies ist z. B. bei den Antworten zur Frage, ob man sich Arbeit wünscht, auch wenn man im Augenblick nicht danach sucht, der Fall. Bei dieser Frage dürften ohne die Anlegung besonderer Maßstäbe Schätzungen und falsche Antworten eigentlich überhaupt nicht vorkommen. Aber auch die mit dieser Frage verbundene Zuordnung von Personen zur Stillen Reserve erfolgt – vermutlich nachdem die Befragungsperson den Befragten zu einer Entscheidung ermuntert hat – nicht ohne Schätzanteil und einer in gleicher Weise unterdurchschnittlichen Zahl an Änderungen der zunächst gegebenen Antwort. Das Beispiel zeigt, dass die anschließend betrachtete Genauigkeit der Fremdauskünfte unbedingt einen direkten Vergleich mit den Selbstauskünften erfordert.

Ergebnisse zur Genauigkeit der Fremdauskünfte

Aus theoretischer Sicht ist es naheliegend und schon mehrfach versucht worden, die Effekte von Fremdworten auf die Genauigkeit der Ergebnisse durch Testerhebungen bzw. Nachbefragungen zu ermitteln. Da im Mikrozensus aber jeder Datensatz eine Kennzeichnung der Art der Auskunftserteilung enthält, ist mit den Ergebnissen selbst bereits eine zusätzliche Datenquelle für die Ermittlung von Proxy-Effekten verfügbar. Unter anderem wurde anhand der aus mehreren Berichtsjahren zusammengeführten Daten des Mikrozensus ermittelt, welche Auswirkungen Proxy-Antworten auf die Qualität

der Ergebnisse haben¹³⁾. Diese Analyse kommt zu dem Schluss, dass Fremdauskünfte die Wahrscheinlichkeit für eine falsche Antwort zwar leicht erhöhen, dass aber die Stärke dieses Effektes vergleichsweise gering ist. Auch Nachbefragungen führten bisher fast immer zu ähnlichen Ergebnissen.

Zum Stellenwert von Fremdauskünften und den noch folgenden Einschätzungen durch die Interviewerinnen und Interviewer ist zunächst festzustellen, dass im Berichtsjahr 2007 insgesamt nur 80,9 % der hochgerechneten¹⁴⁾ Bevölkerung im Alter von 15 und mehr Jahren von Interviewer(inne)n befragt wurden. Abgesehen von wenigen telefonischen Interviews (1,4 % der erwachsenen Bevölkerung) entfiel der Rest auf sogenannte „Selbstauffüller“, die einen ausgefüllten Fragebogen an das Landesamt zurückgeschickt hatten. Die Selbstauffüller umfassten mit nur 17,7 % den geringsten Anteil an Fremdauskünften¹⁵⁾. Bei den vor Ort durchgeführten Laptop-Interviews lag diese Quote bei 28,5 % und umfasste 3,5 von 12,4 Millionen hochgerechneter Antworten. Der Anteil der Personen, die ihre Antwort nicht selbst gegeben hatten, umfasste in der Altersgruppe der 15- bis unter 30-Jährigen 56 %. Der höchste Anteil an Fremdauskünften (81 %) ergab sich in der Altersgruppe von 15 bis unter 20 Jahren. Unter anderem bewirkt bereits dieser altersspezifische Schwerpunkt der Fremdauskünfte, dass die Effekte von Fremdauskünften mit wenigen Ausnahmen – wie z. B. bei der Art der besuchten Schule – bei anderen Merkmalen auch in den Einschätzungen der Interviewerinnen und Interviewer nicht so deutlich werden können, wie dies bei einer zufälligen Verteilung der

13) Vgl. hierzu: Sylvia Zühlke „Auswirkungen von Proxy-Interviews auf die Datenqualität des Mikrozensus“ in Statistische Analysen und Studien, Band 53; LDS NRW. – 14) Fremdauskünfte beziehen sich häufig auf schwer erreichbare Personen. Damit verbundene Antwortausfälle werden erst durch das Hochrechnungsverfahren ausgeglichen. Deshalb wird auf die alternative Darstellung der Stichprobenbesetzungen verzichtet. – 15) Während die Einstufung zur Art der Beteiligung der einzelnen Personen an der Erhebung am Schluss des Laptop-Interviews vom Interviewer vorgenommen wird, beruht diese Angabe bei den Selbstauffüllern auf einer Selbsteinschätzung.

Fremdauskünfte möglich gewesen wäre.

Außerdem legte das etwas diffuse Gesamtbild der bisherigen Informationen nahe, die Gelegenheit der Interviewerbefragung zu nutzen, um zunächst die Zuordnungen zur Art der Auskunftserteilung selbst einer Prüfung auszusetzen. Die in der Interviewerbefragung nicht berücksichtigte Standardsituation für eine Fremdauskunft ist die, dass ein anderes Haushaltsmitglied für eine während des gesamten Interviews nicht in der Wohnung anwesende Person antwortet. Diese Situation kommt relativ häufig vor, ist zugleich von der Interviewerin oder vom Interviewer eindeutig zuzuordnen und erlaubt zudem Rückschlüsse auf den Kontext jeder einzelnen Antwort. In der Praxis gibt es aber zwischen der Fremdwortantwort in dieser reinen Form und der durchgängigen Selbstauskunft verschiedene Mischformen, die ohne adäquat differenzierte Vorgaben der Einstufung durch die Interviewerinnen und Interviewer oder der Selbsteinstufung der Selbstauffüller überlassen werden müssen.

Die Einschätzung der Interviewer/-innen lässt erkennen, dass in nicht wenigen Fällen keine reine Zweiteilung in Selbst- und Fremdauskünfte möglich ist. Vielmehr zeigt sich in den Verteilungen der folgenden Tabelle 2 eine differenzierte und durch Schwerpunktsetzungen auf individuelle Verhältnisse reagierende Sichtweise. Bei der Formulierung der Antwortmöglichkeiten war zunächst nur an die Anwesenheit während des Interviews und an die Kombination mit unterschiedlichen Graden der aktiven Beteiligung gedacht worden, um zu ermitteln, an welcher Stelle der einigermaßen hierarchischen Sortierung der Antworten eine Wende von der Fremdauskunft zur Selbstauskunft sichtbar wird. Da abgesehen von den fehlenden Vorgaben auch keine Möglichkeit bestand, die Art der Auskunftserteilung für jede einzelne Antwort anzugeben, hatte die/der Interviewer/-in nicht nur die Aufgabe, den Zeitraum der Anwesenheit, sondern auch den Umfang der Beteiligung am Interview individuell zu bewerten.

2. Die Art der Auskunftserteilung aus der Sicht der Interviewer/-innen				
Situation	Am Ende des Interviews geben Sie an, in welcher Form die einzelnen Haushaltsmitglieder an der Beantwortung der Fragen beteiligt waren. Was kreuzen Sie an, wenn ...	Eigene Beteiligung	Durch eine andere Person vertreten	Noch nicht vorgekommen
		Anzahl der Nennungen		
A)	die befragte Person bei einer bestimmten Frage zögert und telefonischen Kontakt zur betroffenen Person aufnimmt.	36	168	110
B)	die betroffene Person zwar zum Zeitpunkt des Interviews in der Wohnung ist, sich aber nur kurzfristig im gleichen Raum befindet.	40	237	37
C)	die betroffene Person zwar nicht selbst antwortet, aber während des gesamten Interviews im gleichen Raum ist.	198	67	49
D)	die betroffene Person nur einen geringen Teil der Fragen selbst beantwortet.	182	84	49
E)	die betroffene Person fast alle Fragen selbst beantwortet.	310	2	3

Im Ergebnis zeigt sich zunächst, dass nicht der Umfang der Beteiligung, sondern der Grad der räumlich-körperlichen Anwesenheit als vorrangiges Kriterium angesehen wird: Einer Person, die – wie in der Situation C) beschrieben – während des gesamten Interviews anwesend ist, ohne etwas zu sagen, wird in nur 67 Fällen unterstellt, dass sie von einer anderen Person vertreten wird. Auf der anderen Seite genügt bei vergleichbarer Aktivität die eingeschränkte Anwesenheit der Situation B) für eine gegenteilige Zuordnung. Interessant ist in diesem Zusammenhang auch der leichte Knick in der Einschätzung der Situation D). Dieser Knick ist darauf zurückzuführen, dass in der Beschreibung die unterstellte durchgängige Anwesenheit während des Interviews nicht ausdrücklich wiederholt wurde. Durch diese Unschärfe in der Fragestellung wurden vermutlich einige Interviewer/-innen animiert, sich bei der Zuordnung auf das bereits als zweitrangig identifizierte Kriterium der aktiven Beteiligung zu konzentrieren.

Formal betrachtet, müsste eigentlich eine Person, die zwar anwesend ist, aber keine einzige Frage selbst beantwortet, als durch eine andere Person vertreten gelten. Der überwiegenden Mehrheit der Interviewer/-innen genügte aber eine Anwesenheit, die auch als schweigende Zustimmung gedeutet werden kann. Dies erklärt zwar den überwiegenden Wechsel von der Vertretung durch eine andere Person hin zur ei-

genen Beteiligung, nicht aber, weshalb bereits in den Situationen A) und B) ein nicht zu vernachlässigender Teil der Interviewerinnen und Interviewer die alternative Zuordnung gewählt hat. Bei einer telefonischen Kontaktaufnahme ist zu vermuten, dass nicht das ganze Interview, sondern nur wenige Zweifelsfragen mit dem abwesenden Familienmitglied geklärt werden. Eine eigene Beteiligung kann in diesem Fall nur dann sinnvoll angenommen werden, wenn man unterstellt, dass bei den nicht abgestimmten Fragen kein Unterschied zwischen der Selbst- und der Fremdauskunft möglich ist. Das heißt mit anderen Worten, dass nicht nur die Intensität, sondern auch die Relevanz der Beteiligungen an den Interviews in die Zuordnungen einfließt.

Bei den folgenden Genauigkeitsvergleichen ist daher zu berücksichtigen, dass selbst bei den dann auf Einzelfragen bezogenen Vergleichen immer noch sehr komplexe Konstellationen im Hintergrund wirksam sind, die sich nicht nur auf die Einschätzung der Interviewerinnen und Interviewer zur Genauigkeit von Fremdantworten, sondern auch auf die Einordnung der Art der Auskunftserteilung selbst beziehen. Um die Trennschärfe zu erhöhen, wäre es daher wünschenswert gewesen, die vorher beschriebenen Einschätzungen für jede Beispielfrage zu ermitteln und erst auf dieser Basis die Genauigkeitsvergleiche vorzunehmen. Dies hätte aber den Rahmen der kleinen Zusatzbefragung gesprengt.

Der bei Selbstauskünften sofort verstandenen und kurz beantworteten Frage wurde in der Interviewerbefragung gegenübergestellt, ob die Antwort auch für abwesende Haushaltsmitglieder ohne Zögern gegeben wurde. Da bezüglich aller Fragen beide Situationen von der Mehrheit der Interviewerinnen und Interviewer als oft vorkommend bezeichnet waren, ergab sich erst aus der Gewichtung der Antworten, dass die Quote der in jeder Hinsicht zuverlässigen Antworten in der Fremdauskunft durchgängig kleiner war als bei den Selbstauskünften. Beim Nettoeinkommen sank diese Quote im Vergleich zu den Selbstauskünften um 23,7 Prozentpunkte auf nur noch 47,3 % der Antworten. Selbstverständlich sagt auch diese beispiellos niedrige Quote nichts über die tatsächliche Genauigkeit der gespeicherten Ergebnisse der Erhebungen aus, sondern bestätigt lediglich, dass die mit der Einkommensfrage verbundenen Akzeptanzprobleme alle übrigen Messergebnisse der Interviewerbefragung überlagern. Nach dem Eindruck der Interviewer/-innen hatten 23,8 % der anwesenden Haushaltsmitglieder nur lückenhafte Vorstellungen, wie hoch die persönlichen Einkommen der Abwesenden waren. Auch die in dieser Quote enthaltenen 10,1 % der Fälle, bei denen eine Schätzung hingenommen wurde, spiegelt nur teilweise die tatsächliche Unkenntnis der anwesenden Haushaltsmitglieder wider. Dies ergibt sich bereits daraus, dass in 9,2 % der Fälle die Beantwortung der Ein-

kommensfrage für abwesende Personen zunächst einfach abgelehnt wird. Außerdem ist nicht mit Sicherheit anzunehmen, dass bei gemeinsam wohnenden und wirtschaftenden Personen die Unkenntnis bezüglich der einzelnen Beiträge zum gemeinsamen Haushaltseinkommen stets größer ist als z. B. zur Frage, ob jemand mehr arbeiten möchte. Insofern sind auch die durch vorsichtige Worte eingeschränkten 19,5 % der Antworten zum Einkommen nur bedingt als Indikator für eine fehlende Genauigkeit der Antwort zu interpretieren.

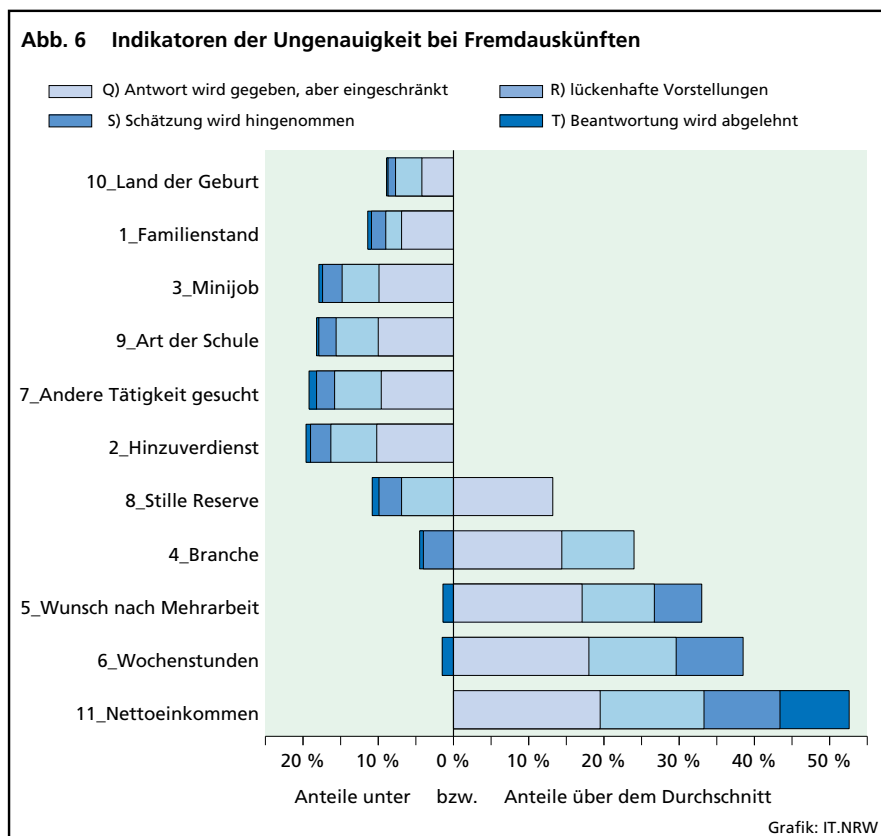
Die Frage nach dem Wunsch mehr zu arbeiten und die unmittelbar folgende Frage nach der in der Berichtswoche tatsächlich geleisteten Arbeitsstunden erweisen sich dagegen als Repräsentanten für die Proxy-Problematik deutlich besser geeignet. Der Wunsch nach Mehrarbeit wurde in den Fremdauskünften mit einem Anteil von 65,5 % häufiger problemlos beantwortet als die Einkommensfrage, aber dies ist vor dem Hintergrund zu sehen, dass zwischen dieser Quote und der Quote der problemlos Selbstauskünfte (92,4 %) eine sogar die Einkommensfrage übertreffende Differenz gemessen wurde. Eine etwas kleinere, aber alle übrigen Beispielfragen überragende Differenz (19,8 Prozentpunkte) zwischen der relativen Genauigkeit von Selbst- und Fremdauskünften ergab sich auch bei den Wochenstunden. Interessanterweise wird der objektiv überprüfbareren Frage nach den Wochenstunden eine geringere Genauigkeit der Fremdauskünfte beimessen. Den beiden Fragen ist gemeinsam, dass sie bezüglich aller Formen der Einschränkungen der Genauigkeit überdurchschnittliche Anteile aufweisen. In Abbildung 6 zu den Indikatoren bei Fremdauskünften erscheinen die rund 1,5 % Ablehnungen zu diesen Fragen zwar als unterdurchschnittliche Anteile auf der linken Seite der Skala, weil selbst der Mittelwert dieser Quoten von der Einkommensfrage dominiert wird, aber die hier vergleichsweise hohen Anteile haben weitaus mehr Gewicht unter dem Blickwinkel, dass die Ablehnung der Beantwortung

auf Unwissen beruht. Gleiches ist für die Fragen zur Arbeitssuche (1,0 % Ablehnung) und zur Frage nach dem Arbeitswunsch ohne Arbeitssuche (0,9 %) anzunehmen.

Bei der Ermittlung der Wirtschaftszweige von Erwerbstätigen aus Fremdauskünften liegt die Quote der hingenommenen Schätzungen (4,0 %) bereits unter dem Durchschnitt. Hoch ist dagegen mit 14,4 % der Anteil der vorsichtig eingeschränkten Antworten. Von den Befragungspersonen wurden zwar nur 71,1 % der Antworten als problemlos eingeschätzt, aber dieser Anteil war mit einem Abstand von nur noch 6,2 Prozentpunkten bei den Selbstauskünften ähnlich hoch. Auf den ersten Blick scheint es daher – wegen der direkten Codierung verbaler Beschreibungen durch die Befragungspersonen – keine besonderen Unterschiede zwischen der Fremd- und der Selbstauskunft zu geben. Erst mit der Beobachtung, dass bei den Selbstauskünften die Antwort auf die Branchenfrage zu 11,7 % erheblich häufiger durch Nachfragen der Interviewerinnen und Interviewer geändert wurde als bei allen anderen Beispielfragen,

zeigt sich ein etwas anderes Bild: Da bei den Fremdauskünften nicht abgefragt wurde, wie oft Änderungen der Antwort vorgenommen wurden, kann nur vermutet werden, dass der Vorteil der Selbstauskunft bei bestimmten Fragen vor allem darauf beruht, dass zwischen den Befragungspersonen und den unmittelbar betroffenen Befragten ein intensiver Dialog möglich ist.

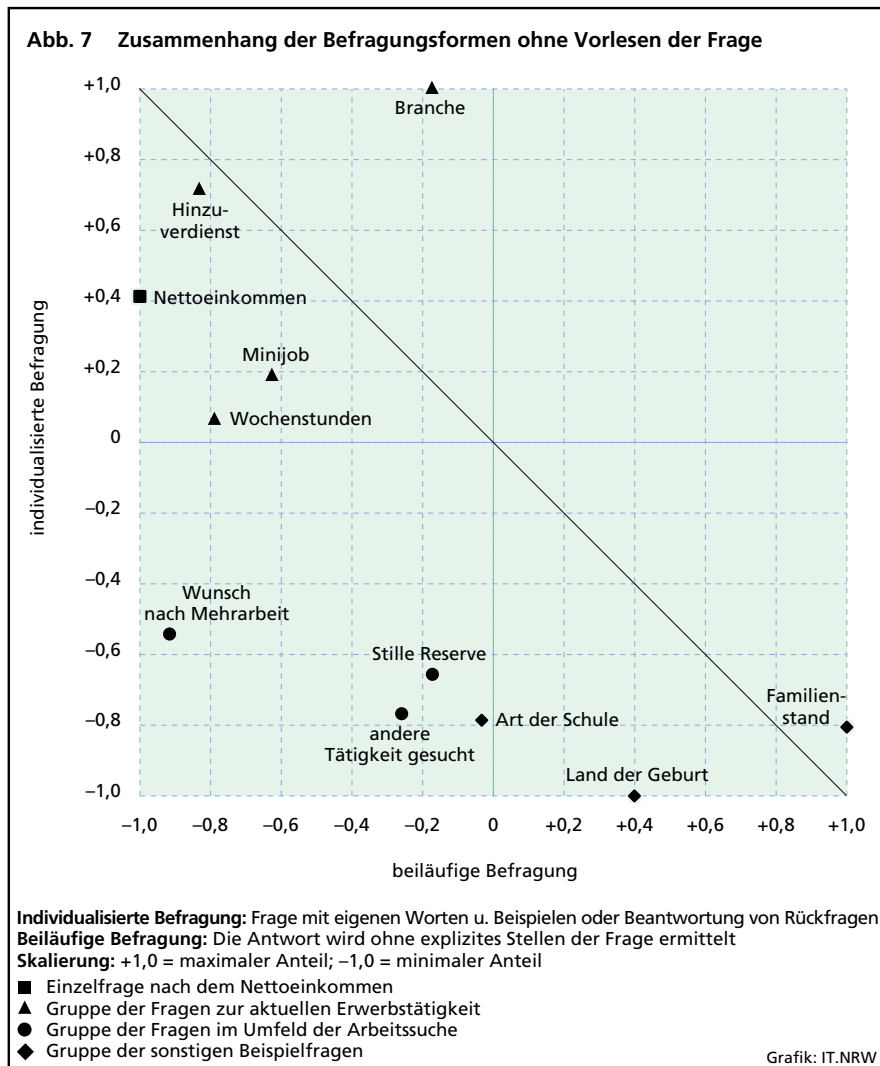
Mit nur 0,8 Prozentpunkten zeigte sich bei der Frage zum Hinzuverdienst die geringste Differenz zwischen den Anteilen der unauffälligen Selbst- und Fremdauskünfte. Bei häufiger akzeptierten Fragestellungen wie dem Familienstand (2,4 Prozentpunkte) oder dem Geburtsland (3,4 Prozentpunkte) war dagegen eine geringere Übereinstimmung zu verzeichnen. Dies deutet bereits darauf hin, dass eine geringere Akzeptanz der Frage dazu führen kann, dass weder die Vorteile der individualisierten noch der direkten Befragung zum Tragen kommen können. Deshalb kann die so gut wie nie kritisierte Frage nach der Art der Schule oder Hochschule als Normalfall der Struktur einer Auskunft über Dritte angenommen werden. Zu mehr als



80 % der Antworten unterscheidet sie sich nicht von einer Selbstauskunft. Bei weiteren 10 % der Antworten sind nur geringfügige Unsicherheiten der Auskunft gebenden Person sichtbar. Erst bei den restlichen 10 % der Antworten wird ein im Vergleich zur unmittelbar betroffenen Person geringeres Wissen oder die Ablehnung der Stellvertreterrolle (0,3 %) sichtbar. Auch aus dem Blickwinkel der Interviewerinnen und Interviewer erscheinen demnach die Unterschiede zwischen Selbst- und Fremdauskünften weniger bedeutsam zu sein, als vorher angenommen. Ein ähnliches Ergebnis war bereits in den anfangs erwähnten Analysen sichtbar. Obwohl die Interviewerbefragung hinsichtlich der Proxy-These demnach keine wirklich neuen Ergebnisse ergeben hat, wurde zumindest deutlicher sichtbar, weshalb eine direkte Messung der Unterschiede zwischen Selbst- und Fremdauskünften schwierig ist.

Zusammenfassung

Für die statistische Widerlegung der eingangs beschriebenen Thesen wäre es – über die Beobachtung der Reaktionen von Interviewer/-innen und Haushalten auf die einzelnen Fragen hinaus – erforderlich gewesen, auch die Beteiligung der Haushaltsmitglieder an der Beantwortung der Fragen, die Struktur der befragten Haushalte und weitere Aspekte fragenspezifisch zu ermitteln. Schon die fehlende Eindeutigkeit der Unterteilung der Antworten in Fremdauskünfte und Selbstauskünften zeigt aber, dass direkte Nachfragen bei den Haushalten oder bei den Interviewer(inne)n nicht mehr ergeben können, als z. B. die Angabe „oft“ beim Vorlesen der Frage, bei der uneingeschränkten Akzeptanz und bei der uneingeschränkten Genauigkeit von Selbst- und von Fremdauskünften. Erst nach der Gewichtung der Häufigkeitsnennungen im Zusammenhang wurde sichtbar, dass die ausgewählten Beispielfragen in allen betrachteten Dimensionen Unterschiede aufweisen. Diese Unterschiede zeigten sich jedoch nicht als Ergebnis einer direkten



Messung, sondern nur auf der Basis geschätzter relativer Häufigkeiten bezüglich einzelner Beispielfragen. Deshalb wird von der möglichen Darstellung berechneter Korrelationen abgesehen und stattdessen eine heuristische¹⁶⁾ Gegenüberstellung der relativen Häufigkeiten genutzt, um die wesentlichen Ergebnisse der Interviewerbefragung noch einmal zusammenzufassen.

Aus dem Blickwinkel der Ergebnisse der Interviewerbefragung ist die Effizienz-These zurückzuweisen, soweit sie sich auf den Interviewer selbst bezieht. Da die in dieser These unterstellte Absicht, das Interview durch eine falsche Antwort zu verkürzen, mit Sicherheit nicht durch das Geben von Beispielen oder der Beantwortung von Rückfragen in

den Blickpunkt der befragten Personen gerückt würde, müsste die Indikatorfrage zum Hinzuverdienst in der Abbildung 7 an einer anderen Stelle zu finden sein. Die Abbildung zeigt allgemein den insgesamt negativen Zusammenhang¹⁷⁾ zwischen der beiläufigen und der individualisierten Befragung nach Normierung¹⁸⁾ der relativen Häufigkeiten anhand der maximalen und minimalen Häufigkeiten. Im Vergleich zur Frage nach der Branche, die – wie bereits zu sehen war – der besonderen Aufmerksamkeit der Interviewerinnen und Interviewer bedarf, fiel

17) Der dargestellte Zusammenhang ist kein Artefakt der vorgenommenen Normierung. Dies ist daran zu sehen, dass der über die Beispielfragen gerechnete Rangkorrelationskoeffizient (-0,700) exakt an der Signifikanzgrenze liegt. Auch die für jede Beispielfrage zu berechnenden Korrelationen der relativen Häufigkeiten waren – z. B. beim Hinzuverdienst (-0,139) – zwar nur schwach, aber immer noch signifikant negativ. – 18) Die relativen Häufigkeiten h wurden umgerechnet in den Wert:

$$h_{Norm} = 2 * (h - h_{Min}) / (h_{Max} - h_{Min}) - 1$$

16) Als Heuristik bezeichnet man eine Methode, komplexe Probleme, die sich nicht vollständig lösen lassen, mit Hilfe einfacher Regeln und unter Zuhilfenahme nur weniger Informationen zu entwirren.

als erste Bestätigung der intensiven Bemühungen der Interviewerinnen und Interviewer um die Frage nach dem Hinzuverdienst auch der Anteil der beiläufigen Befragung fast so gering aus, wie bei den Fragen zum Nettoeinkommen oder dem Wunsch nach Mehrarbeit. Die letztgenannte Frage erweist sich aus dem Blickwinkel des Interviewereinsatzes trotz des subjektiven Charakters als völlig unproblematisch. Sie wird im Normalfall einfach vorgelesen und beantwortet. Falls ein Wunsch nach Mehrarbeit besteht, ist im Interview zwar nur die Folgefrage nach der Stundenzahl vorgesehen, aber wenn eine messbare Tendenz der Haushalte bestehen würde, der Effizienz-These zu folgen, hätte auch bei dieser Frage ein etwas anderer Gesamteindruck sichtbar werden müssen.

Die Zusatzbefragung der Interviewerinnen und Interviewer hat vor allem gezeigt, dass sowohl die konkreten Befragungsmöglichkeiten als auch die Qualität der Antworten im persönlichen Interview entscheidend davon abhängen, ob die befragten Personen die Frage verstehen und akzeptieren. Dies ergibt sich aus dem noch deutlicheren Zusammenhang zwischen der Akzeptanz und der Genauigkeit der Fragen. In dieser Gegenüberstellung zeigt sich die Sonderstellung der Frage nach dem Nettoeinkommen besonders deutlich. Einschränkungen der Akzeptanz und der Genauigkeit sind bei dieser Frage gleichermaßen am stärksten ausgeprägt. Dabei ist zu beachten, dass nicht unbedingt die geringe Akzeptanz als Ursache für vermutete oder tatsächliche Ungenauigkeiten anzusehen ist. Denkbar ist auch die umgekehrte Wirkungsrichtung, nämlich dass erst eine Überforderung der Befragten dazu führt, dass die Frage selbst kritisiert wird. Auffällige Besonderheiten in diesem Zusammenhang sind bei den Fragen zu den Wochenstunden und zur Branche zu erkennen. Durch die Ermittlung der zutreffenden Wirtschaftszweige im Dialog mit den Interviewerinnen und Interviewern – in der Abbildung 8 sind nur die Einschränkungen der Genauigkeit von Selbstauskünften dargestellt – wird zwar ein hoher

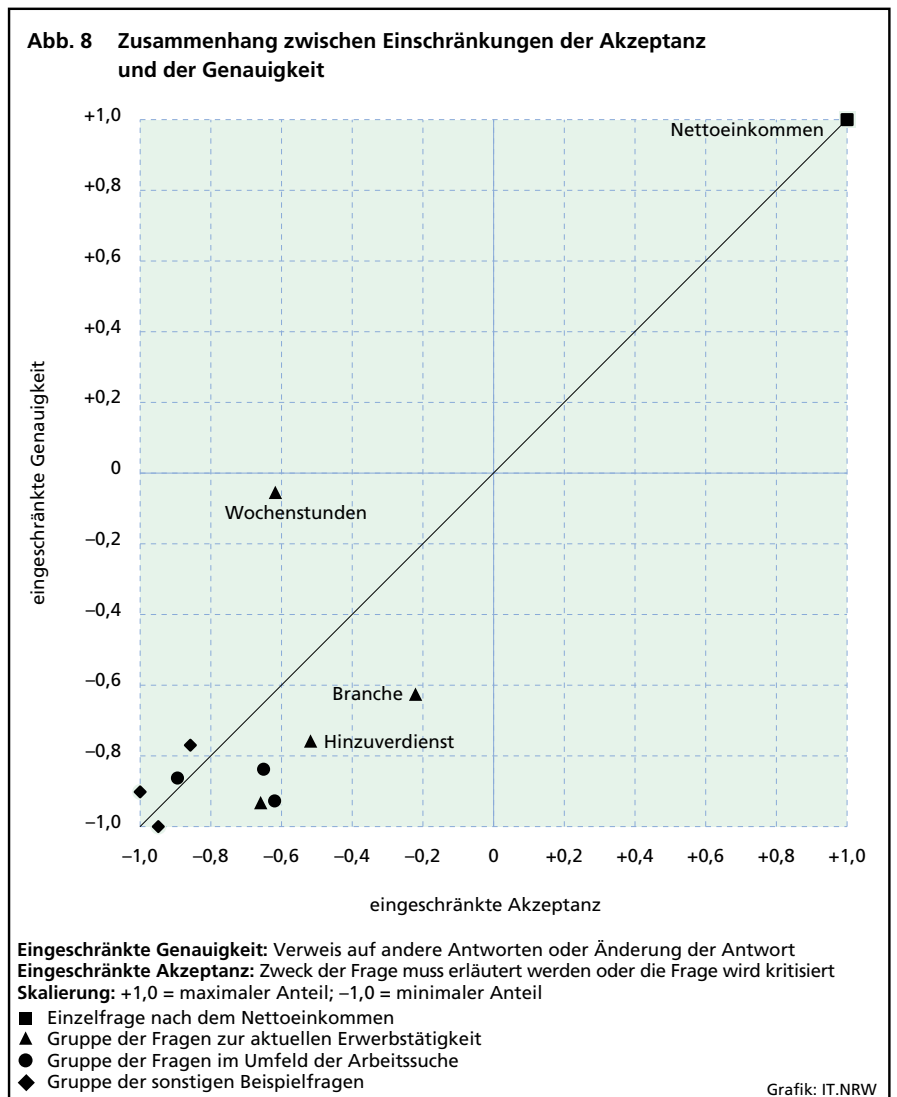
Grad an Genauigkeit erreicht, aber der möglicherweise zeitraubende Prozess der systematischen Zuordnung führt dazu, dass Einschränkungen der Akzeptanz sogar häufiger auftreten als bei den Fragen zum Hinzuverdienst oder den Wochenstunden.

Der gleichgerichtete Zusammenhang zwischen Akzeptanz und Genauigkeit wird nur durch die Frage nach der Zahl der tatsächlich gearbeiteten Stunden durchbrochen. Der Zweck der Frage ist nur selten zu erläutern und auch kritische Einwände der Haushalte treten nicht wesentlich häufiger auf als bei den übrigen Beispielfragen. Trotzdem wird die Genauigkeit durch Verweise auf die normale Zahl der Arbeitsstunden eingeschränkt oder erst durch Nachfragen der Interviewerinnen und Interviewer korrigiert. Das Beispiel zeigt, dass für die Gewinnung einer

genauen Antwort die Akzeptanz der Frage nötig ist, aber auch eine akzeptierte Frage nicht unbedingt genaue Antworten liefert.

Während der Einfluss der Akzeptanz der Fragestellung bei den Befragten auf die Genauigkeit der Antworten sehr deutlich ist, ist eine Bestätigung des Proxy-Effekts im Sinne einer die vergleichbare Selbstauskunft an Ungenauigkeit übertreffenden Fremdauskunft nur in Einzelfällen klar zu sehen. In der direkten Gegenüberstellung der relativen Häufigkeiten der genauen Antworten zeigt sich zunächst nichts anderes als eine hohe Übereinstimmung der Qualität der Selbst- und Fremdauskünfte.

Eine eindeutige Überlegenheit der Selbstauskünfte ist nur bei den Wochenstunden und beim Wunsch nach Mehrarbeit zu sehen. Bei vier der Beispielfragen ergab sich nach dem



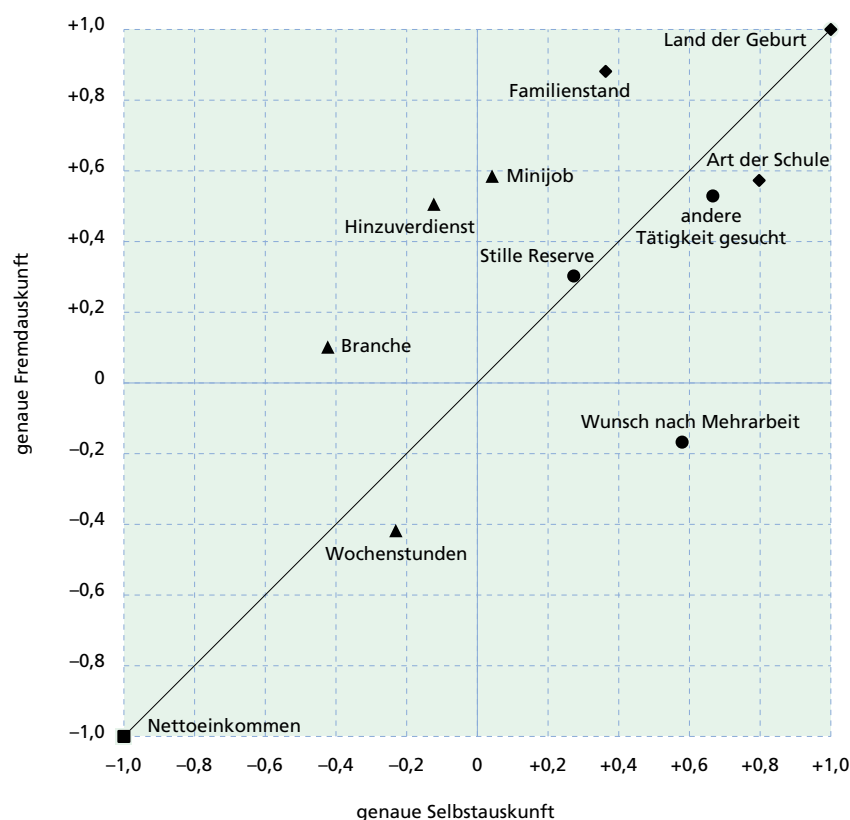
rechnerischen Abzug aller Indikatoren der Ungenauigkeit in der Abbildung 9 bei den Fremdanworten sogar ein größerer Rest an scheinbar genauen Antworten als bei den Selbstauskünften. Bezeichnenderweise taucht in diesem Zusammenhang nicht nur die Frage nach dem Wirtschaftszweig, sondern auch die Frage nach dem Familienstand wieder auf. Zum Familienstand war schon vor der Befragung der Interviewerinnen und Interviewer eine vergleichsweise hohe Stabilität der Fremdanworten festgestellt worden. In dieser Befragung war außerdem zu beobachten, dass die Probleme der Frage zum Wirtschaftszweig und anderen Sachverhalten im direkten Dialog mit den Befragten am besten gelöst werden. Insgesamt ist zu sehen, dass auch im direkten Interview grundsätzlich ein Unterschied besteht, ob die unmittelbar betroffene Person selbst antwortet oder ob die Antwort von einem anderen Haushaltsmitglied gegeben wird. Insofern ist die Proxy-These im Gegensatz zur Effizienz-These nicht völlig von der Hand zu weisen.

Hinsichtlich der Konsequenzen¹⁹⁾ ist allerdings zu beachten, dass bei vielen Fragen die Unterschiede zwischen Selbst- und Fremdauskunft nur geringfügig sind, während die Vorzüge der direkten Befragung gegenüber einer telefonischen oder postalischen Befragung aus der Ferne nicht nur in dieser Befragung der Interviewer/-innen, sondern auch in den erhobenen Daten selbst deutlich zu sehen sind. Das bedeutet für die Qualität der Ergebnisse des Mikrozensus, dass eine Ausweitung oder zumindest ein Erhalt des aktuellen Umfangs der persönlichen Befragung (face-to-face) durch Interviewerinnen und Interviewer Vorrang haben muss vor anderen Maßnahmen, die möglicherweise nur eine marginale Reduktion der Proxy-Effekte bewirken.

Um zu dem Schluss zu kommen, dass vor allen anderen Faktoren, die zur

19) In den Fachgremien des Bundes und der Länder wurde z. B. diskutiert, den bestehenden Erhebungsvordruck für Haushalte durch Fragebogen für jede Person zu ersetzen, um den Anteil der Fremdauskünfte im Mikrozensus zu senken.

Abb. 9 Zusammenhang zwischen den Genauigkeiten von Selbst- und von Fremdauskünften



Genauere Fremdauskunft: Frage wird ohne Zögern auch für abwesende Haushaltsmitglieder beantwortet
Genauere Selbstauskunft: Frage wird sofort verstanden und kurz beantwortet
Skalierung: +1,0 = maximaler Anteil; -1,0 = minimaler Anteil

- Einzelfrage nach dem Nettoeinkommen
- ▲ Gruppe der Fragen zur aktuellen Erwerbstätigkeit
- Gruppe der Fragen im Umfeld der Arbeitssuche
- ◆ Gruppe der sonstigen Beispielfragen

Grafik: IT.NRW

Genauigkeit von Befragungsergebnissen beitragen, der Akzeptanz der Fragestellung bei den Befragten eine herausragende Bedeutung zukommt, fehlt noch ein Blick auf die Hauptstatus-These. Wie bereits am Anfang erwähnt, wurde die Frage zum Hinzuverdienst nicht nur wegen der Effizienz-These als Beispielfrage ausgewählt. Durch das Bemühen, eine möglichst große Ausschöpfung der geringfügigen Beschäftigungsverhältnisse zu erreichen, hat diese Frage einen fast suggestiven Charakter angenommen. Schon durch die Formulierung „Auch wenn man eigentlich nicht ...“ wird zusätzlich eingeräumt, dass diese Frage sich auf einen Sachverhalt bezieht, der nicht unbedingt dem Selbstverständnis der befragten Bevölkerung entspricht. Die Interviewerbefragung hat ergeben, dass zusätzliche Erläuterungen und die Beantwortung von Rückfragen der Befragten durch die Befra-

gungsperson fast so häufig vorkommen wie bei der ebenfalls ohne Dialog nur schwer zu beantwortenden Frage nach dem Wirtschaftszweig. Entsprechend hoch war der Erläuterungsbedarf (14,0 %) zum Zweck der Frage, sodass insgesamt alle Einschränkungen der Akzeptanz nur noch von den Fragen zur Branche und zum Nettoeinkommen übertroffen wurden. Entsprechend selten – in nur 57,2 % der Fälle – wurde den Interviewerinnen und Interviewern die Gelegenheit eingeräumt, die Frage im Wortlaut vorzulesen.

Insgesamt konnte mit den Ergebnissen der Interviewerbefragung gezeigt werden, dass Verständnis- und Akzeptanzprobleme durch den Interviewereinsatz bis zu einem bestimmten Punkt gemildert werden können und gemildert werden. Das letzte Beispiel hat aber gezeigt, dass weitere Verbesserungen der Ergebnisse

nicht über noch genauere Vorgaben oder durch eine noch intensivere Befragung der Haushalte zu erreichen sind. Das bedeutet, dass schon bei der theoriegeleiteten Definition der Inhalte voraussehbare Verständnis- und Akzeptanzprobleme in der Bevölkerung mit berücksichtigt werden müssen. Nach der intensiven Überprüfung der Feldarbeit, in der die Ergebnisse dieser Zusatzbefragung nur einen kleinen Ausschnitt darstellen, ist es an der Zeit, die Fragestellungen selbst in die Überprüfungen einzubeziehen, um noch vorhandene Gegensätze zwischen der Theorie und der Erhebungspraxis zu überbrücken. Zur Ermittlung der Zahl der geringfügig Beschäftigten wäre es

aus befragungstechnischer Sicht z. B. denkbar, dass zunächst das Selbstverständnis der Befragten (als Rentner etc.) akzeptiert wird und erst danach und völlig losgelöst vom Erwerbsstatus die dann einfache Frage gestellt wird, ob in der Berichtswoche einer Arbeit nachgegangen wurde. Eine dem Selbstverständnis der Bevölkerung nähere Fragestellung, zu der die Auftraggeber²⁰⁾ der Statistik bereit sein müssten, besteht in

20) Die Fragestellung ist nicht nur im Mikrozensusgesetz, sondern auch in EU-Verordnungen geregelt. Änderungen der Fragestellungen erfordern daher ein intensives Engagement in den Abstimmungsprozessen zur Novellierung der Rechtsgrundlagen. Kurzfristige Änderungen sind deshalb nur über andere Formulierungen und andere Filterführungen erreichbar.

diesem Beispiel darin, dass die Definition des ILO-Erwerbsstatus nicht mehr im direkten Zusammenhang abgefragt wird, sondern sich erst aus der Kombination der Antworten zu den einfacheren Fragen nach dem Hauptstatus und einer Arbeit in der Berichtswoche ergibt.