

Stichprobenziehung von Ehepaaren

Sampling Couples

*Ein Praxisbericht zur
Erhebung von seltenen
Paarkombinationen*

*An Approach to Identifying
Rare Couple Combinations*

David Glowsky

Zusammenfassung

Sozialwissenschaftliche Fragestellungen beziehen sich häufig nicht auf Einzelpersonen, sondern auf Kombinationen von Personen im Familienverband. Für seltene Kombinationen, wie gleichgeschlechtliche Paare oder Familien mit Zwillingen erweist sich die Ziehung einer Zufallsstichprobe jedoch oft als problematisch, da herkömmliche Screeningverfahren hier an ihre Grenzen stoßen. In diesen Fällen ist es wirtschaftlicher und geht es schneller, entsprechende Personenkonstellationen im Melderegister zu identifizieren, indem man Daten des Familienverbandes verknüpft und entsprechende Kombinationen direkt aus dem Melderegister herausfiltert. Der vorliegende Beitrag stellt diesen Ansatz exemplarisch anhand einer Befragung von binationalen Ehepaaren vor. In einem ersten Teil werden Schwierigkeiten bei der Ziehung seltener Paarkombinationen diskutiert. Zweitens wird die Möglichkeit erläutert, Daten von Familienmitgliedern im Melderegister zu verknüpfen. Die Genauigkeit dieses Ansatzes wird anhand der Berliner Paarstudie ausgewer-

Abstract

Research questions in the social sciences often do not focus on individuals, but rather on combinations of persons within a family. For rare combinations, such as same sex couples and families with twins, it often proves difficult to draw random samples, since conventional screening methods reach their limits when the target population is too small. In such cases, another method promises to be faster and more cost-effective: combinations of family members can be identified in the population register by linking single entries, which enables us to filter certain combinations on the basis of attributes of several family members. The article presents this approach, using a survey of binational couples as an example. In a first step, difficulties of sampling rare couple combinations are discussed. Secondly, the paper presents the alternative approach, in which data of several family members are combined within the population register. The accuracy of this approach is evaluated, using the Berlin Couple Survey, in which name-based tech-



tet, bei der zusätzlich eine onomastische Filterung zum Einsatz kam. Als Ergebnis zeigt sich, dass sich die meisten Kombinationen von Ehepartnern deutscher und ausländischer Herkunft treffsicher ermitteln lassen, während Beifang vor allem durch Einbürgerungen und deutsche Namensformen bedingt ist. Abschließend werden weitere Anwendungsmöglichkeiten des Verfahrens diskutiert.

niques supplemented the selection process. The results show that most combinations of couples with an ethnic German partner and a foreign-born partner can be identified accurately, with false identification being mainly due to naturalization and Germanized names. Finally, further possible applications of this sampling approach are discussed.

1 Einleitung¹

Quantitative Sozialforschung lebt von verallgemeinerungsfähigen und unverzerrten Datensätzen. Sofern zu einer Fragestellung keine amtlichen Daten vorliegen, werden die benötigten Informationen meist mittels einer Bevölkerungsbefragung erhoben. Dazu gilt es zunächst, eine Zufallsstichprobe aus der Zielpopulation zu gewinnen. Für den Großteil der Forschungsfragen lassen sich relativ leicht Zufallsstichproben gewinnen, da sie sich auf die Gesamtbevölkerung eines bestimmten Territoriums beziehen. Eine Zufallsauswahl aus der Gesamtbevölkerung lässt sich beispielweise über eine zufällige Auswahl von Telefonnummern, das Random-Route-Verfahren oder eine einfache Melderegisterstichprobe realisieren. Die Kunst besteht in diesen Fällen nicht darin, eine besonders große Stichprobe zu erreichen, sondern eine möglichst kleine Auswahl von wenigen tausend Befragten so zusammensetzen, dass sie stellvertretend für die Gesamtbevölkerung stehen kann. Für die Bearbeitung allgemeiner Fragestellungen, die sich auf die Gesamtbevölkerung beziehen, ist es häufig auch nicht notwendig, jeweils eine neue Befragung durchzuführen. Stattdessen kann man oft auf die existierende Dateninfrastruktur zurückgreifen, die umfangreiche Befragungen der Bevölkerung bereithält, wie das Sozio-oekonomische Panel (Wagner/Frick et al. 2007), den Mikrozensus (Lüttinger/Riede 1997) oder den ALLBUS (Terwey 2000).

Mit zunehmender Ausdifferenzierung der Sozialwissenschaften wächst aber auch die Zahl differenzierter Fragestellungen, die sich nicht auf die Gesamtbevöl-

1 Dieser Artikel ist aus einem Projekt entstanden, das von der Fritz Thyssen Stiftung und der Studienstiftung des deutschen Volkes finanziert wurde. Teile des Beitrags wurden Kapitel 4 in Glowsky (2011) entnommen. Ich danke Jürgen Schupp, Lars Breuer, den Herausgebern der MDA und zwei anonymen Gutachtern für hilfreiche Kommentare. Sabine Hüßgen, Lisette Schneider und Teresa Zeckau haben das Projekt als Hilfskräfte unterstützt.

kerung, sondern auf bestimmte Teilgruppen beziehen. Abhängig davon, wie klein diese Teilgruppe ist, kann sich die Ziehung einer Zufallsauswahl als problematisch erweisen. Dies ist beispielsweise dann der Fall, wenn die Zielpopulation nicht aus Einzelpersonen, sondern aus seltenen Kombinationen von Personen besteht. In der Familiensoziologie kommen häufiger Fragestellungen auf, die sich auf Personenkonstellationen innerhalb eines Familienverbandes beziehen. So interessiert sich die Scheidungsforschung für Paare, die sich getrennt haben oder bei denen eine Trennung bevorsteht. Für Studien zur Fertilität sucht man möglicherweise ältere kinderlose Paare oder kinderreiche Familien. Aber auch andere Forschungszweige haben Bedarf an Stichproben mit seltenen Personenkonstellationen. So dienen beispielsweise gleichgeschlechtliche Paare als Forschungsobjekte der Genderforschung, die Jugendforschung kann hingegen Adressen von Familien mit Stiefeltern suchen. Die Zwillingsforschung wiederum benötigt Stichproben von Familien mit Kindern, die am gleichen Tag geboren wurden.

Diese Populationen sind häufig so klein, dass Screening-Verfahren an ihre Grenzen stoßen. In diesen Fällen ist es wirtschaftlicher und geht es vermutlich auch schneller, entsprechende Personenkonstellationen im Melderegister zu identifizieren, indem man die Daten des Familienverbandes verknüpft. Auf diese Weise können Eigenschaften von Partnern bzw. Eltern und Kindern im Familienverband bei der Stichprobenziehung gemeinsam einbezogen und direkt aus dem Melderegister herausgefiltert werden.

Der vorliegende Beitrag präsentiert diesen Ansatz exemplarisch anhand der Berliner Paarstudie, einer Befragung von binationalen Ehepaaren. Der folgende Abschnitt (2) diskutiert Schwierigkeiten bei der Stichprobenziehung von seltenen Populationen. Der Hauptteil (3) beschreibt zunächst das übliche Vorgehen bei einer Melderegisterstichprobe und erläutert dann die Möglichkeit, Daten im Familienverband zu verknüpfen. Daran anschließend geht der Beitrag näher auf methodische Aspekte bei der Stichprobenziehung der betreffenden Studie ein. Ein Fazit (4) beurteilt schließlich den Mehrwert des hier präsentierten Vorgehens und wirft einen Ausblick auf weitere Anwendungsmöglichkeiten.

2 Schwierigkeiten der Stichprobenziehung bei seltenen Paarkombinationen

Die Grundlage für diesen Bericht bildet ein von der Fritz Thyssen Stiftung gefördertes Projekt, das von 2006 bis 2010 am Institut für Soziologie der Freien Universität Berlin durchgeführt wurde. Das Projekt war im Forschungsfeld über Heiratsmigra-

tion angesiedelt. Ziel war es, Mechanismen der Partnerwahl bei Eheschließungen zwischen deutschen Männern und Heiratsmigrantinnen aus dem ärmeren Ausland zu untersuchen. Eine Sekundäranalyse eines existierenden Datensatzes war ausgeschlossen, da in Deutschland keine Datensätze mit ausreichender Fallzahl und partnerwahlspezifischen Informationen existierten. Für das Projekt war es deshalb notwendig, einen neuen Datensatz zu erstellen. Da eine mündliche interviewergestützte Befragung aus Kostengründen ausschied, sollte stattdessen eine schriftliche Befragung von zusammenlebenden Ehepaaren durchgeführt werden, in denen der Mann in Deutschland geboren war oder seit Geburt die deutsche Staatsbürgerschaft innehatte. Ferner sollten der Mann und mindestens ein Elternteil Deutsch als Muttersprache haben. In der deutschen Unterstichprobe sollte die Frau in Deutschland geboren sein oder seit Geburt die deutsche Staatsbürgerschaft haben. In den anderen Gruppen sollte die Frau im Ausland geboren sein und eine ausländische Staatsbürgerschaft innehaben. Um diese Befragung durchführen zu können, sollte möglichst eine Zufallsstichprobe dieser Grundgesamtheit gezogen werden.

Allerdings umfassen Paare mit deutsch-ausländischer Herkunft nur einen sehr kleinen Teil der deutschen Bevölkerung. Im Mikrozensus 2003 finden sich 3106 Männer, die in Deutschland geboren sind und die deutsche Staatsangehörigkeit innehaben und deren Frau im Ausland geboren ist. Diese 3106 Männer entsprechen 0,62 % der Gesamtpopulation. Im umgekehrten Fall sind es sogar nur 2893 Frauen (0,58 % der Grundgesamtheit), die mit einem im Ausland geborenen Mann verheiratet sind. Bei einem derart geringen Anteil der Zielgruppe an der Bevölkerung handelt es sich um eine *seltene Population* (Schnell/Hill et al. 2008: 294). Seltene Populationen sind oft nur mit großem Aufwand zu ermitteln, insbesondere Zufallsstichproben gestalten sich hierbei schwierig. Im glücklichsten Fall existiert eine vollständige Liste der Zielpopulation, aus der sich dann eine Zufallsstichprobe ziehen lässt.

Alternativ kann man auf Screeningverfahren zurückgreifen, die in der Gesamtpopulation systematisch nach Angehörigen der Zielpopulation suchen und eine Zufallsstichprobe gewährleisten. Ein Telefonscreening mit dem Random-Digit-Dial-Verfahren (Schnell/Hill et al. 2008: 294ff.; Gabler/Häder 1997; Häder/Gabler 1998) sowie das Random-Route-Verfahren (Hoffmeyer-Zlotnik 2003; 1997; Alt/Bien et al. 1991) kamen für die Berliner Paarstudie nicht infrage. Bei diesen Verfahren wäre im vorliegenden Fall eine sehr hohe Zahl an Kontaktversuchen notwendig gewesen, um ausreichend Paare mit deutschem Mann und ausländischer Frau zu erreichen. Der dabei entstehende Zeit- und Kostenaufwand hätte die zur Verfügung stehenden Mittel überschritten. Eine zeit- und kostensparende Methode, die insbesondere für die Stichprobenziehung bei Migrantenforschungen

eingesetzt wird, ist das onomastische Verfahren (Humpert/Schneiderheinze 2000; 2002; Salentin 2002). Bei diesem Verfahren werden typische Vor- und Nachnamen in einem öffentlichen Verzeichnis – z. B. dem Telefonbuch – gesucht. Dieses Verfahren ließ sich für die Berliner Paarstudie jedoch nicht anwenden, da Paare mit einem deutschen Ehemann gesucht wurden. Ehefrauen nehmen auch heute noch größtenteils den Nachnamen des Mannes an. Eine Analyse des Melderegisters der Stadt Berlin zeigt, dass 95 % der Ehefrauen thailändischer Nationalität, 30 % der brasilianischen Ehefrauen, 79 % der polnischen Ehefrauen und 68 % der russischen Ehefrauen den gleichen Namen tragen wie der Ehemann. Da die Frauen fast ausschließlich den deutschen Namen des Mannes annehmen, würden zu viele Paare nicht erfasst, weshalb auch dieses Verfahren ausgeschlossen wurde.²

Zwei weitere Alternativen zur Stichprobenziehung bei seltenen Populationen sind das Schneeballverfahren (Gabler 1992; Diekmann 2008: 400; Schnell/Hill et al. 2008: 300) und die Suche in speziellen Kontexten. Beide Verfahren führen jedoch zu erheblichen Verzerrungen. Beim Schneeballverfahren besteht die Stichprobe schließlich aus einem oder mehreren persönlichen Netzwerken. Da Menschen in der Regel soziale Kontakte mit Anderen pflegen, die über ähnliche Eigenschaften verfügen, kommt es in der Folge zu einer systematischen Überrepräsentation bestimmter Merkmale (Salentin 1999: 124). Zudem lassen sich seltene Populationen häufig in speziellen Kontexten antreffen. Der Rückgriff auf derartige *foci of activity* (Feld 1981) geht jedoch in aller Regel mit einer Verzerrung mehrerer Merkmale einher. Niesner (2000) befragt lediglich ausländische Ehefrauen, die in Beratungszentren Hilfe gesucht haben; Paare, die gut miteinander auskommen, fallen aus dem Raster heraus.

Um eine Stichprobe aus der hier anvisierten seltenen Population zu gewinnen, sind die genannten Verfahren ineffektiv bzw. ungeeignet. Es bleibt deshalb zu fragen, ob es Listen gibt, in denen diese Paare verzeichnet sind. Im Ausländerzentralregister (AZR) sind alle dauerhaft in Deutschland lebenden Personen verzeichnet, die nicht die deutsche Staatsangehörigkeit besitzen (Babka von Gostomski/Pupeter 2008). Das Verzeichnis speichert Angaben zu Geschlecht, Staatsangehörigkeiten und Familienstand, womit verheiratete Frauen bestimmter Staatsangehörigkeiten ermittelt werden könnten. Allerdings sind im AZR keine Informationen zum Ehegatten gespeichert, weshalb sich aus dem Verzeichnis weder ermitteln lässt, ob der Ehegatte noch lebt, noch ob beide Partner zusammenleben, noch welche

2 Die Fallzahl beträgt 1.000 Paare pro Gruppe, mit Ausnahme der brasilianischen Frauen (451). Fast ausschließlich übernehmen die Frauen den Namen des Mannes. In den ausländischen Gruppen ließen sich nur zwei Fälle ermitteln, in denen ein Mann den Namen der Frau angenommen hat.

Staatsangehörigkeit der Ehepartner hat.³ Einer Befragung müsste also wieder ein aufwendiges Screening vorhergehen.

Es wurde deshalb entschieden, eine Melderegisterstichprobe zu ziehen. Die Melderegister der Gemeinden bilden eine solide Datenbasis für die hier anvisierte Stichprobe, da sie über die wichtigsten Merkmale für die Identifizierung verfügen und eine Suche nach Kombinationen von Ehepartnern ermöglichen. Der folgende Abschnitt stellt ein Verfahren vor, das mit einer Stichprobe aus dem Melderegister treffsicher Anschriften von binationalen Ehepaaren ermittelt.

3 Stichprobe aus dem Melderegister

3.1 Eine Einführung in Melderegisterstichproben

Die Melderegister sind aufgrund der gesetzlichen Meldepflicht das umfassendste Verzeichnis der deutschen Wohnbevölkerung mit relativ hoher Aktualität. Theoretisch sind im Melderegister alle in der Bundesrepublik lebenden Personen erfasst. Aus mehreren Gründen stimmen die Melderegister nicht vollständig mit der Wohnbevölkerung überein. Sie beinhalten einerseits Anschriften, die nicht mehr gültig sind (u. a. wegen Tod, Umzug), andererseits sind nicht alle Personen an ihrem tatsächlichen Wohnort registriert. Der Zensus 2001 (Statistische Ämter des Bundes und der Länder 2004) hat die Qualität der Melderegister der Gemeinden geprüft. Es zeigte sich, dass bundesweit 4,1 % der in den Melderegistern verzeichneten Haushalte nicht mehr existieren; die Zahl der Haushalte, die umgekehrt nicht im Melderegister enthalten sind, liegt bei 1,7 %. Die Zahl der Karteileichen kann um den Anteil „temporärer“ Karteileichen bereinigt werden, indem die Registerdaten zu einem späteren Zeitpunkt noch einmal abgeglichen werden. Durch dieses Verfahren können Fälle identifiziert werden, in denen z. B. eine Person zum Stichtag verstorben oder während eines Umzuges für kurze Zeit an zwei Orten mit Hauptwohnsitz gemeldet ist. Die bundesweite Karteileichenrate lag nach der Bereinigung bei 2,3 %. Sachsen (0,9 %), Thüringen (1,1 %) und Mecklenburg-Vorpommern (1,2 %) weisen die niedrigsten Raten an verbleibenden Karteileichen auf, in Berlin (5,8 %), dem Saarland (6,3 %) und Hamburg (5,4 %) liegen diese Raten am höchsten. Stichprobenziehungen aus dem Melderegister führen prinzipiell zu höheren Fehlerquoten in großen Städten, während sie in kleinen Gemeinden eine bessere Datenqualität erwarten lassen. Das Melderegister weicht damit vom Ideal eines

3 Telefonat mit dem Ansprechpartner für das Ausländerzentralregister beim Bundesamt für Migration und Flüchtlinge, 30.06.2011.

Vollverzeichnisses aller in Deutschland lebenden Personen ab, weist aber insgesamt eine hohe Übereinstimmung mit den tatsächlich existierenden Haushalten auf.⁴

Die rechtliche Grundlage für Stichprobenziehungen bildet das Melderechtsrahmengesetz (MRRG) des Bundes, auf dem die jeweiligen Meldegesetze der 16 Bundesländer beruhen. In der Fassung vom 19.04.2002 ermöglicht § 21 Abs. 3 eine sogenannte „Gruppenauskunft“: „Melderegisterauskunft über eine Vielzahl nicht namentlich bezeichneter Einwohner (Gruppenauskunft) darf nur erteilt werden, wenn sie im öffentlichen Interesse liegt.“ Dieselbe Formulierung findet sich auch in den Meldegesetzen der Länder. Universitäten und andere öffentliche Forschungseinrichtungen erfüllen die Anforderung des öffentlichen Interesses. Ferner sind Universitäten öffentliche Stellen, denen noch etwas weitere Auskunftsrechte eingeräumt werden.

Die Melderegister sind auf der Gemeinde- bzw. Landkreisebene angesiedelt. Um eine Stichprobe aus einem Melderegister zu ziehen, muss jede Einwohnermeldebehörde jeweils einzeln kontaktiert werden. Dies führt zu einem erheblichen Administrationsaufwand auf der Seite der Forscher. Zudem entstehen bei jeder Ziehung neue Kosten. Der Aufwand und die Kosten lassen sich erheblich verringern, indem nur eine kleine Zahl von Einwohnermeldeämtern in die Stichprobenziehung einbezogen wird. Umfragen, die sich auf die gesamte Bundesbevölkerung beziehen, verwenden deshalb in der Regel zweistufige, geschichtete Stichproben (Albers 1997). In der ersten Stufe wird zunächst eine Zufallsstichprobe von Gemeinden gezogen.⁵ In der zweiten Stufe wird dann aus den Registern dieser Gemeinden eine Zufallsstichprobe von Personen gezogen.

Der ALLBUS 2008 z. B. wählte in der ersten Stufe bundesweit 162 „Sample-Points“ aus (diese entfielen auf 148 Gemeinden, wobei Städte wie Berlin und Ham-

- 4 Personen mit ausländischer Staatsbürgerschaft werden überproportional als Karteileichen in den Melderegistern geführt. So kann z. B. Granato (1999: 58) 93 % der Adressen der deutschen Stichprobe verifizieren, doch nur 78 % der türkischen Stichprobe. Es ist zu vermuten, dass dies auch für die hier untersuchte Population gilt, jedoch in deutlich geringerem Maße, da die Ehefrauen zwar ausländische Staatsangehörigkeit haben, die Ehemänner jedoch in Deutschland geboren sind und deutsche Staatsangehörigkeit haben. Sie weisen vermutlich ein sehr ähnliches Meldeverhalten wie die allgemeine deutsche Bevölkerung auf.
- 5 Mit der Verringerung der Gemeinden geht eine Erhöhung des Klumpeneffektes einher. Die Einwohner einer Gemeinde (Klumpen/Cluster) neigen dazu, einander ähnlich zu sein und sich von den Einwohnern anderer Gemeinden zu unterscheiden. Je homogener die Cluster und je größer die Unterschiede zwischen den Clustern sind, umso größer ist der Genauigkeitsverlust (Kish 1965). Zudem steigt der Designeffekt mit der Zahl der Elemente pro Cluster. Bei einer sehr kleinen Anzahl von Gemeinden ist eine Aussage über die Grundgesamtheit „Bevölkerung der Bundesrepublik Deutschland“ äußerst unpräzise.

burg aufgrund ihrer relativen Größe mit mehreren Sample-Points vertreten waren⁶) und zog in der zweiten Stufe aus jedem Sample-Point 51 Personen (Wasmer/Scholz et al. 2010). Eine Stichprobenziehung in diesem Umfang ist allerdings sehr kostspielig. In der Praxis entscheiden sich Forscher deshalb häufig dafür, die Ziehung auf deutlich weniger Gemeinden zu beschränken. Dies gilt für Pilotstudien (z. B. Granato 1999: Mannheim; Diehl 2007: Essen und München; Brüderl/Castiglioni et al. 2007: Bremen, Mannheim, München, Chemnitz) sowie für kleinere Projekte mit eingeschränkter finanzieller Förderung (z. B. Reuschke 2010: München, Stuttgart, Düsseldorf, Berlin). Wenn die Stichprobe nur wenige Gemeinden enthalten soll oder kann, werden bevorzugt große Städte gewählt, wofür sich generell drei Gründe nennen lassen. Erstens lässt sich so bei der kleinstmöglichen Gemeindezahl die größtmögliche Zahl von Adressen realisieren, was kostensparend wirkt (Diehl 2007). Zweitens kann die Beschränkung auf Großstädte theoriegeleitet sein, z. B. wenn sich die Fragestellung auf Berufspendler in Metropolen bezieht (Reuschke 2010). Drittens handelt es sich häufig um Städte, in denen die Forscher selbst lokalisiert sind, was die Rekrutierung und Kontrolle von Interviewern erleichtert (Brüderl/Castiglioni et al. 2007). Bei diesem Vorgehen wird gleichzeitig ein deutlicher Nachteil hingenommen: Die Stichprobe hat nur beschränkte Aussagekraft, da sie nur Rückschlüsse auf die Bevölkerung in großen Städten zulässt. Über andere Siedlungsgebiete, z. B. ländliche Regionen, sind damit keine Aussagen möglich.

3.2 Verknüpfung von Daten im Familienverband

Die Datenstruktur der Melderegister basiert auf der Erfassung von Einzelpersonen. Diese Datenstruktur reicht für eine Zufallsstichprobe der gesamten Bundesbevölkerung aus. Entsprechend dokumentieren Erfahrungsberichte zur Nutzung des Melderegisters in aller Regel die Stichprobenziehung von Einzelpersonen, ohne Informationen über weitere Familienmitglieder zu berücksichtigen (vgl. u. a. Albers 1997; Brüderl/Castiglioni et al. 2007; Granato 1999; Diehl 2007; Reuschke 2010).

Für die hier anvisierte Stichprobe sind jedoch mehrere Informationen über *beide Ehepartner* notwendig. Im Melderegistereintrag einer Person sind neben den persönlichen Daten auch Daten eventueller Ehegatten und Lebenspartner (Vor- und Familiennamen, Doktorgrad, Tag der Geburt, Anschrift und Sterbetag)

6 Die mehrfache Ziehung von Großstädten in der ersten Stufe ging auf die Gewichtung der Gemeinden nach ihrer Bevölkerungszahl zurück. Diese bestimmte das Gewicht einer Gemeinde, woraus sich z. B. ein Gewicht von 0,33 für Pinneberg (ein oder kein Sample-Point) und ein Gewicht von 2,16 für München (zwei oder drei Sample-Points) ergab (Wasmer/Scholz et al. 2010).

sowie die Daten minderjähriger Kinder (Vor- und Familiennamen, Tag der Geburt und Sterbetag) gespeichert (vgl. § 2, Melderechtsrahmengesetz). Abhängig von der verwendeten Datenbanksoftware sind die Melderegister deshalb in der Lage, Verknüpfungen zwischen Ehegatten und gemeinsam wohnenden Lebenspartnern sowie minderjährigen Kindern herzustellen. Ist diese Verknüpfung hergestellt, so lassen sich die in den einzelnen Einträgen vorhandenen Informationen gemeinsam für die Auswahl verwenden. Auf diese Weise ist es also möglich, bestimmte Kombinationen von Ehegatten und ihren Kindern aus dem Melderegister zu filtern. Die Vorteile dieses Vorgehens liegen auf der Hand. Visiert eine Befragung eine spezielle Population an, die sich aus bestimmten Kombinationen von Ehegatten oder Kindern zusammensetzt, können mittels einer Verknüpfung diejenigen Fälle ausgewählt werden, die auch tatsächlich in die Grundgesamtheit fallen. Unnötiger Beifang kann so vermieden werden.

Dieser Vorteil lässt sich anhand der Anforderungen der Berliner Paarstudie illustrieren: Die Studie zielte auf eine sehr seltene Population ab, die aus Kombinationen von Ehepartnern bestand. Eine Ziehung von Einzelpersonen aus dem Melderegister hätte zu großem Beifang geführt. Zwar wäre in einer Stichprobe mit verheirateten Frauen einer bestimmten Staatsangehörigkeit vermutlich ein großer Prozentsatz „richtiger“ Fälle enthalten gewesen. In Deutschland lebende verheiratete Frauen mit brasilianischer oder thailändischer Staatsbürgerschaft haben vermutlich häufiger einen deutschen Ehemann als einen Mann mit der gleichen Staatsbürgerschaft. Doch neben der Staatsbürgerschaft des Mannes bliebe bei dieser Methode auch die Frage ungeklärt, ob das Ehepaar zusammenlebt. Der Anteil „falscher“ Fälle wäre also nicht unerheblich.

Für die Berliner Paarstudie ließ sich hingegen die Möglichkeit nutzen, eine Stichprobe nach kombinierten Eigenschaften der Ehepartner zu ziehen. Dies ermöglichte eine deutlich genauere Eingrenzung der Fälle. Auf diese Weise konnten *beide* Partner nach ihrer Staatsangehörigkeit bestimmt werden, was die Ermittlung aller denkbaren Kombinationen erlaubt. Weiterhin war es möglich, nur jene Paare zu berücksichtigen, die unter der gleichen Anschrift gemeldet waren. Getrennt lebende Paare wurden so vermieden.⁷ Abschnitt 3.4 geht auf die Vorgehensweise

7 Eine noch bessere Eingrenzung der Stichprobe ist theoretisch über die Filterung nach dem Geburtsort beider Partner möglich. So könnte man eine Stichprobe auf Ehepaare begrenzen, bei denen ein Partner in Deutschland geboren ist und die deutsche Staatsbürgerschaft besitzt, der andere Partner dagegen in einem anderen Land geboren ist und dessen Staatsangehörigkeit trägt. Bei diesem Verfahren ist es jedoch notwendig, möglichst vollständige landesspezifische Ortslisten anzulegen, die alle Schreibvarianten der Orte beinhalten und danach im Melderegister zu suchen (Salentin 2007). Dieses Verfahren ist mit großem Codieraufwand verbunden und erhöht so die Kosten der

genauer ein. Zuvor diskutiert Abschnitt 3.3 jedoch mögliche Ungenauigkeiten, die auftreten, wenn deutsch-ausländische Ehepaare aus dem Melderegister gezogen werden.

3.3 Mögliche Untererfassung der Zielpopulation

Mit einer Melderegisterstichprobe von Ehepaaren gehen mehrere Nachteile einher, die zu einer möglichen Untererfassung der Zielpopulation führen können (Groves/Fowler et al. 2004: 70ff.; Diekmann 2008: 376ff.; Schnell/Hill et al. 2008: 271ff.):

Erstens werden keine Paare erfasst, die im Ausland leben und keinen Wohnsitz in Deutschland gemeldet haben. Es ist möglich, dass sich diese Paare systematisch von Paaren unterscheiden, die sich nach der Heirat entscheiden, in Deutschland zu wohnen. Vermutlich sind deutsche Partner, die der Kultur des ausländischen Partners gegenüber aufgeschlossener sind, auch eher bereit, in dessen Herkunftsland zu leben. Dies verzerrt die Stichprobe auch für andere Merkmale, die mit kultureller Offenheit korrelieren.

Zweitens sind bei diesem Verfahren Paare mit eingebürgerten Partnern ausgeschlossen. Auch hier kann es zu einer Verzerrung kommen, wenn ein Zusammenhang zwischen der Bereitschaft zur Einbürgerung und hypothesenbezogenen Merkmalen besteht.

Drittens sind geschiedene Paare im Melderegister nicht verzeichnet. Die Scheidungswahrscheinlichkeit hängt vom Grad der kulturellen Differenz zwischen den Partnern ab (Kalmijn/Graaf et al. 2005; Roloff 1998). Paare mit besonders großen Differenzen in Werteinstellungen dürften deshalb zwischenzeitlich häufiger geschieden und in der Melderegisterstichprobe unterrepräsentiert sein.

Wie problematisch diese Einschränkungen für die Befragungsergebnisse tatsächlich sind, hängt von der Fragestellung der Studie ab. Wenn sich die Fragestellung z. B. nur auf verheiratete, in Deutschland lebende Paare mit unterschiedlicher Staatsbürgerschaft bezieht, führen die genannten Einschränkungen nicht zu Undercoverage. Wenn die Studie jedoch untersuchen möchte, welche deutschen Staatsbürger Partner aus dem Ausland heiraten, dann sollte man auch Paare befragen, die im Ausland leben, bei denen der ausländische Partner eingebürgert ist oder

Fortsetzung Fußnote 7

Stichprobenziehung erheblich. Personen mit Geburtsorten, die in der Ortsliste nicht aufgeführt sind, werden bei diesem Verfahren zudem fälschlicherweise nicht einbezogen, was bei einer so seltenen Stichprobe wie den hier untersuchten deutsch-ausländischen Ehen unbedingt vermieden werden sollte.

die bereits geschieden sind – diese Paare sind allerdings nur mit sehr großer Mühe bzw. nicht mittels einer Zufallsstichprobe zu ermitteln.

3.4 Stichprobenziehung der Berliner Paarstudie

Die Berliner Paarstudie beschränkte sich auf eine Stichprobe aus dem Berliner Melderegister. Das geht vor allem auf die Kosten für eine Gruppenauskunft zurück, die in vielen Städten relativ hoch sind. Salentin (2007: 37) schätzt die durchschnittlichen Gebühren auf 200 bis 1000 Euro je Kommune, mit Ausreißern von über 10000 Euro. Da die Studie mit einem eingeschränkten Budget auskommen musste, war es leider notwendig, die Stichprobe auf ein Melderegister zu beschränken. Die Wahl fiel auf Berlin, da es sich hierbei mit 3,4 Mio. Einwohnern (Amt für Statistik Berlin Brandenburg 2009) um das größte Verzeichnis in Deutschland handelt. Zudem liegt aufgrund des hohen Ausländeranteils eine hohe Dichte der gesuchten Ehepaare vor. Darüber hinaus hat Berlin den Vorteil, dass hier langjährige BRD-Bürger und ehemalige DDR-Bürger gemeinsam anzutreffen sind.

Die Stichprobenziehung in Berlin hat zur Folge, dass direkte Schlussfolgerungen aus dem Datensatz nur für die Berliner Bevölkerung möglich sind. Verallgemeinerungen auf andere Teile der Bundesbevölkerung müssen von einer starken Ähnlichkeit mit der Berliner Bevölkerung ausgehen. Dies dürfte für andere Großstädte z. B. viel eher der Fall sein als für ländliche Regionen.

Die Berliner Paarstudie zog eine Stichprobe von Ehepaaren mit deutschen Männern und Frauen aus Polen, Russland, Thailand, den Philippinen und Brasilien. Ehepaare mit deutschen Frauen dienten als Kontrollgruppe.⁸ Es war vorgesehen, für die sechs genannten Herkunftsländer jeweils 1000 Frauen mit deutschen Ehemännern anzuschreiben. Dazu war ein mehrstufiger Auswahlprozess notwendig. Im Oktober 2007 wurde eine nach Herkunftsland der Ehefrauen geschichtete Stich-

8 Diese Länderauswahl orientierte sich an empirischen und theoretischen Kriterien. Polen ist seit zwei Jahrzehnten das wichtigste Herkunftsland ausländischer Ehefrauen. Russland und Thailand nahmen im gleichen Zeitraum alternierend den zweiten und dritten Platz ein. Die Philippinen waren in den 1980er Jahren das Hauptherkunftsland ausländischer Ehefrauen außerhalb Europas. Brasilien ist in den letzten Jahren als einziges südamerikanisches Herkunftsland unter den zehn wichtigsten Herkunftsländern vertreten gewesen. Zudem weisen alle diese Länder einen niedrigeren Human Development Index auf als Deutschland. Die Türkei ist seit den 1970er Jahren regelmäßig unter den zehn wichtigsten Herkunftsländern ausländischer Ehefrauen zu finden, allerdings geht der Großteil dieser Heiraten auf türkischstämmige Deutsche zurück (vgl. Straßburger 2000: 29f.; 2003).

probe aus dem Melderegister Berlins gezogen.⁹ Zunächst wurde eine Zufallsauswahl von jeweils 2.000 Adressen von Ehepaaren mit Männern deutscher Staatsangehörigkeit und Frauen deutscher, thailändischer, polnischer und russischer Staatsangehörigkeit vorgenommen. Von deutsch-brasilianischen Paaren lagen im Melderegister nur 682 Adressen vor, von den deutsch-philippinischen Paaren nur 377 Adressen, die jeweils vollständig übermittelt wurden. Technisch handelte es sich dabei um eine Ziehung von Einzelpersonen: verheiratete Frauen, die eine der genannten Staatsangehörigkeiten innehatten. Diese wurden jedoch nur dann ausgewählt, wenn ihr Ehemann ausschließlich die deutsche Staatsangehörigkeit besaß. Die übermittelten Daten umfassten Name, Geburtsjahr und Anschrift der Frau sowie des Mannes.

Im ersten Schritt wurde die Anschrift der Partner von der Meldebehörde noch nicht automatisch für die Auswahl einbezogen. Deshalb mussten im zweiten Schritt aus den übermittelten Daten jene Paare gefiltert werden, die unter der gleichen Adresse gemeldet waren. Dabei stellten sich große Unterschiede zwischen den Unterstichproben heraus (Tabelle 1).

Deutsche Frauen sind am häufigsten unter der gleichen Adresse wie ihre Ehemänner gemeldet (89,0 %). Unter den Paaren mit ausländischen Ehefrauen zeigt sich eine Zweiteilung: Von den Frauen aus Thailand, den Philippinen und Brasilien wohnen etwa 80 % mit ihren Männern zusammen. Bei den Frauen aus Osteuropa – Polen und Russland – sind es nur weniger als 70 %. Aus der Tabelle geht ebenfalls hervor, dass die Zahl der fehlenden Angaben teilweise erheblich ist. Bei Ehefrauen aus Polen und Russland fehlt in bis zu 20,4 % bzw. 16,3 % der Fälle die Anschrift des Mannes.

Es sollten weiterhin nur jene Paare angeschrieben werden, bei denen der Mann in Deutschland geboren war oder von Geburt an die deutsche Staatsbürgerschaft innehatte, zudem sollten er und mindestens ein Elternteil Deutsch als Muttersprache haben. Über den Zeitpunkt des Erwerbs der deutschen Staatsbürgerschaft und Sprachkenntnisse liegen im Melderegister jedoch keine Angaben vor; das Geburtsland ließe sich über den Geburtsort nachvollziehen, dieses Verfahren ist jedoch sehr aufwendig und kostenintensiv (Salentin 2007). Die Entscheidung fiel deshalb auf ein onomastisches Verfahren, bei dem alle Männer ausgeschlossen wurden, deren Kombination des Vor- und Nachnamens auf eine ausländische Herkunft schließen ließ. Jeder Name wurde von einer Hilfskraft gelesen und als

9 Dankenswerterweise zeigte sich das zuständige Berliner Landesamt für Bürger- und Ordnungsangelegenheiten (LABO) trotz des notwendigen Programmieraufwands sehr kooperativ.

Tabelle 1 Initialstichprobe aus dem Berliner Melderegister (%)

	Deutschland	Thailand	Philippinen	Brasilien	Polen	Russland
n	2.000	2.000	377	682	5.000	3.440
Wohnung						
Zusammenlebend	89,0	81,6	81,2	80,2	65,7	69,1
Getrennt lebend	9,5	14,0	10,3	14,5	13,9	14,6
Anschrift Mann fehlt/ unvollständig	1,5	4,4	8,5	5,3	20,4	16,3
Mann						
Deutscher Name	87,8	86,3	73,8	79,9	42,4	49,3
Ausländischer Name	11,7	10,4	23,3	18,0	48,9	42,8
Name fehlt	0,5	3,3	2,9	2,1	8,7	7,9
Zusammenlebend, Mann trägt deutschen Namen	87,4 (1.756)	73,8 (1.476)	64,0 (241)	66,1 (451)	33,7 (1685)	36,6 (1260)

Quelle: Melderegister der Stadt Berlin (2007), eigene Berechnungen. Stichprobe von 2.000 Paaren für deutsche und thailändische Gruppe, Stichprobe von 5000 Paaren für polnische Gruppe. Vollerhebung für philippinische, brasilianische und russische Gruppe.¹⁰

deutsch oder ausländisch eingestuft.¹¹ Für Paare mit Frauen aus Deutschland, Thailand, den Philippinen und Brasilien war diese Aufgabe leicht zu lösen, da die Namen der Männer in den meisten Fällen eindeutig einer deutschen oder ausländischen Herkunft zuzuordnen sind. Eine Auswertung der rückgesendeten Fragebögen bestätigt dies: In den genannten Gruppen waren 95,2 % bis 99,0 % der angeschriebenen Männer in Deutschland geboren (Tabelle 3). Deutlich schwieriger gestaltete sich die Namensfilterung für Männer, deren Frauen die polnische oder russische Staatsbürgerschaft hatten. Polnische Familiennamen enden u. a. häufig auf ek, czyk, ski und ihren orthografischen Varianten. Diese Namen sind aber auch in Deutschland bei Familien geläufig, die seit mehreren Generationen in Deutschland leben, deutsch sprechen und keine Verbindung zum Herkunftsland ihrer Ahnen mehr haben. Um dies zu umgehen, gingen in diesen Fällen verstärkt

10 Der Anteil der zusammenlebenden Paare, bei denen der Name des Mannes als deutsch eingestuft wurde, lag bei der polnischen und russischen Unterstichprobe bei jeweils nur etwa einem Drittel, weshalb die Stichprobe aus dem Melderegister von 2.000 Paaren nicht ausreichte, um 1.000 Paare anzuschreiben. Deshalb wurde erneut eine Zufallsstichprobe von 5.000 Paaren mit polnischer Ehefrau gezogen; bei den Paaren mit russischer Ehefrau wurden alle 3.440 gemeldeten Paare aus dem Melderegister zur Verfügung gestellt.

11 Um die Intercoderreliabilität einschätzen zu können, ist es meist üblich, derartige Ratings von zwei oder mehr Ratern vornehmen zu lassen. Aus Kapazitätsgründen musste hierauf jedoch verzichtet werden.

die Vornamen in die Entscheidung mit ein.¹² Lediglich 42,4 % der Ehemänner polnischer Frauen wurden anhand ihres Namens als deutschstämmig eingestuft. In dieser Gruppe ist die Trefferquote etwas schlechter. Nur 87,6 % der angeschriebenen Männer mit polnischen Frauen sind in Deutschland geboren (Tabelle 3). Die größten Schwierigkeiten bereiteten jedoch die Namen von Männern mit russischen Frauen. Hierbei handelt es sich offenbar in vielen Fällen um Aussiedler, die seit der Einwanderung über die deutsche Staatsangehörigkeit verfügen. Sie nehmen häufig auch deutsch klingende Namen an, was dazu führt, dass sie anhand dieser Kriterien nicht mehr als russischstämmig zu erkennen sind. Immerhin konnten 42,8 % der Männer mit russischen Frauen anhand ihres Namens mit hoher Wahrscheinlichkeit als Migranten identifiziert werden, doch in dieser Gruppe liegt die Trefferquote am niedrigsten: Lediglich 67,7 % der angeschriebenen Männer russischer Frauen sind in Deutschland geboren, die übrigen 32,3 % stammen aus Russland oder einem anderen Staat der ehemaligen Sowjetunion. Sie waren über ihre Namen nicht identifizierbar gewesen.

Nach dieser Filterung blieben 1.756 Paare mit deutscher Frau, 1.476 Paare mit thailändischer Frau, 241 Paare mit philippinischer Frau, 451 Paare mit brasilianischer Frau, 1.685 Paare mit polnischer Frau und 1260 Paare mit russischer Frau übrig. Die philippinische Unterstichprobe wurde wegen einer zu niedrigen Fallzahl verworfen, wohingegen die brasilianische Gruppe vollständig in die Befragung einging. In den anderen Gruppen wurde aus den verbleibenden Fällen eine Zufallsauswahl von jeweils 1.000 Paaren gezogen.

Die Studie strebte vorrangig einen Vergleich zwischen zwei Gruppen von Männern an: jenen, die eine Frau aus dem ärmeren Ausland geheiratet hatten und jenen, die eine deutsche Frau geheiratet hatten. Die Altersverteilung zwischen Männern mit deutschen Frauen (Mittelwert des Geburtsjahres 1949, SD = 13,4) und Männern mit ausländischen Frauen (1958, SD = 12,2) differierte jedoch stark. Um zu verhindern, dass die Altersdifferenz von knapp zehn Jahren andere Unterschiede verdeckte, wurden in der deutsch-deutschen Stichprobe die Männer aus den Geburtsjahrgängen disproportional ausgewählt. Mit der Anpassung konnte eine Halbierung der Differenz erreicht werden. In der angepassten 1.000er-Stichprobe mit deutschen Ehefrauen liegt das durchschnittliche Geburtsjahr der Männer bei 1954 (SD = 12,4).

12 Migranten vergeben auch in der zweiten Generation meist noch Vornamen, die im Herkunftsland üblich sind (Gerhards/Hans 2009). Deshalb ist davon auszugehen, dass nach diesem Kriterium auch Migranten der zweiten oder dritten Generation herausgefiltert werden.

Bei der anschließenden Befragung wurden 4.451 Paare angeschrieben: jeweils 1.000 Paare mit deutschen Männern und Frauen aus Deutschland, Thailand, Polen und Russland sowie 451 Paare mit Frauen aus Brasilien.

3.5 Rücklauf der Fragebögen und mögliche Ausfälle

Im Februar 2008 wurden die Fragebögen versendet, gefolgt von einer schriftlichen und einer telefonischen Erinnerung. Die Bruttoreücklaufquote aus den einzelnen Stichproben beträgt 31,1 % für Deutschland, 20,4 % für Thailand, 22,8 % für Brasilien, 26,7 % für Polen und 25,4 % für Russland. Insgesamt kamen 25,6 % aller versendeten Fragebögen ausgefüllt zurück. Es war leider nicht möglich, die Zahl der falschen Adressen zu ermitteln, da die Fragebögen aus Kostengründen per *Infopost* versendet wurden und bei dieser Frankierart nicht zustellbare Sendungen vernichtet werden. Nach den Ergebnissen des Zensus-tests (Statistische Ämter des Bundes und der Länder 2004) sowie den Erfahrungen von Granato (1999) und Brüderl et al. (2007) kann man aber von etwa 8 % falscher Adressen ausgehen, was einen korrigierten Rücklauf von 28 % bedeuten würde. Darüber hinaus gab es auch Fälle, bei denen die Anschrift eines Partners zwar noch korrekt war, die Partnerschaft aber seit mehreren Jahren nicht mehr bestand. Die Nettoreücklaufquote aus der Grundgesamtheit dürfte also noch etwas höher liegen.

Eine Antwortquote in dieser Größenordnung lässt nichtsdestotrotz vermuten, dass die Ausfälle zu systematischen Verzerrungen führen. Dies lässt sich überprüfen, indem man die Liste der angeschriebenen Personen mit den Eigenschaften der antwortenden Personen vergleicht. Die Melderegisterstichprobe enthält die Geburtsjahre der angeschriebenen Männer und Frauen. In Tabelle 2 ist die Altersverteilung der angeschriebenen sowie der antwortenden Personen abgetragen.

Die Abweichungen zwischen der angeschriebenen und der realisierten Stichprobe liegen meist unter einem Prozentpunkt pro Zehnjahresgruppe und sind damit als niedrig einzuschätzen. Bei den Frauen fällt die leichte Unterrepräsentation von 40 bis 59-jährigen zugunsten der 20 bis 29-jährigen Frauen auf. Doch auch wenn man die Unterstichproben einzeln betrachtet, zeigt sich ein ähnliches Ergebnis: Der relative Ausfall in Bezug auf das Alter hält sich in einem durchaus annehmbaren Rahmen. Weitere Ausfallanalysen, z. B. zum Bildungsstand, lassen sich nicht realisieren, da das Melderegister hierzu keine Angaben enthält.

Tabelle 2 Relativer Ausfall nach Alter und Geschlecht (%)

Alter	Männer		Frauen	
	Angeschrieben	Rücklauf	Angeschrieben	Rücklauf
20 bis 29 Jahre	3,3	3,9	9,7	11,4
30 bis 39 Jahre	16,5	17,7	31,1	31,1
40 bis 49 Jahre	32,9	31,7	30,4	29,0
50 bis 59 Jahre	23,5	22,8	18,2	17,3
60 bis 69 Jahre	16,7	16,8	7,5	7,9
70 bis 79 Jahre	5,7	6,0	2,6	2,8
80 bis 95 Jahre	1,4	1,1	0,5	0,4
n	4.451	1.135	4.451	1.134

Quelle: Melderegister der Stadt Berlin; Berliner Paarstudie.

Ehepaare mit deutschen Männern und deutschen Frauen haben am häufigsten geantwortet, Paare mit Frauen ausländischer Staatsbürgerschaft bis zu zehn Prozentpunkte seltener. Worin ist dies begründet? Im Datensatz liegen hierzu keine Angaben vor, mehrere Vermutungen liegen jedoch nahe. Einerseits sind die Adressdaten von Paaren mit einem ausländischen Partner möglicherweise häufiger veraltet als jene von deutsch-deutschen Ehepaaren. Dies lässt sich hier nicht überprüfen, doch aus anderen Studien ist bekannt, dass die im Melderegister gespeicherten Adressdaten von Ausländern häufiger veraltet sind als jene von Personen mit deutscher Staatsangehörigkeit. Dies kann z. B. darauf zurückzuführen sein, dass Personen, die ins Ausland umziehen, sich nicht abmelden (Salentin 1999: 117f.). Das erklärt aber nicht, warum der Rücklauf bei Paaren mit Frauen aus Thailand und Brasilien niedriger ist als bei Paaren aus Polen und Russland. Es ist zu vermuten, dass der größte Bias auf die Intimität des Themas und die moralisch aufgeladene öffentliche Diskussion binationaler Ehen zurückgeht, von denen thailändisch-deutsche und brasilianisch-deutsche Paare aufgrund der offensichtlich fremden Herkunft der Frauen möglicherweise häufiger betroffen sind als Paare mit Frauen aus Polen und Russland. Eine starke Ablehnung mehrerer Items kann dann, gerade bei einer schriftlichen Befragung ohne Incentives, auch zu einem Unit-Nonresponse führen. Der Fragebogen erhebt relativ intime Informationen, aus denen sich u. a. der Body Mass Index der Befragten und der Altersabstand des Paares berechnen lässt. Personen, die einen besonders auffälligen BMI oder einen ungewöhnlichen Altersabstand aufweisen, werden möglicherweise von der Beantwortung absehen. Das Gleiche gilt für Personen, die z. B. traditionelle Geschlechterrolleneinstellungen

vertreten oder häusliche Gewalt für akzeptabel halten. Tatsächlich weisen die Teilstichproben mit den niedrigsten Rücklaufquoten empirisch die stärksten Tendenzen in diese Richtungen auf.

Die Anrede beider Partner mit vollem Namen im Anschreiben, die Zweisprachigkeit des Fragebogens für die Frau und die Intimität einiger Fragen machten viele Probanden misstrauisch. Nach Versendung der Fragebögen erreichten uns innerhalb von zwei Wochen etwa 150 Anrufe und E-Mails, wovon der Großteil aus Beschwerden und Absagen bestand. Die Beschwerden der Befragten lassen sich in zwei Punkten zusammenfassen. Erstens waren sie über mangelnden Datenschutz bei der Herausgabe ihrer Anschriftendaten und die Verwendung ihrer Antworten besorgt. Einige wendeten sich an den Datenschutzbeauftragten des Landes Berlin, der eine Anfrage an uns stellte, welche Daten weitergegeben worden waren und wann sie gelöscht würden. Wir gaben an, welche Daten aus dem Melderegister übermittelt worden waren und konnten versichern, dass sie direkt nach der Befragung gelöscht wurden. Rechtlich ist gegen das Vorgehen nichts einzuwenden, doch die Besorgnis der Befragten sollten Sozialwissenschaftler ernst nehmen.¹³ Das Anschreiben teilte bereits mit, dass die Daten aus dem Melderegister stammten und ein beigelegtes Datenschutzblatt erklärte, wie die Antworten anonymisiert werden. Bei Anfragen über die Herkunft der Daten klärten wir die Befragten ausführlicher auf. Zweitens waren einige Befragte über den Inhalt des Fragebogens verärgert. Manche fanden die gestellten Fragen zu persönlich; andere erklärten, dass sie mit der öffentlichen Diskussion des Themas binationale Ehen bereits schlechte Erfahrungen gemacht hätten und sich nicht weiter dazu äußern wollten. Damit gehen Äußerungen einher, der Fragebogen enthalte „zu viele Klischees“ und diene lediglich der „Beschreibung und Entwicklung von Stereotypen“. Aus der rein wissenschaftlichen Warte heraus sind diese Beschwerden nicht resonanzfähig. Wissenschaftler sind bestrebt, die Realität zu beschreiben. Wenn sich ein unangenehmes Klischee im Zuge einer Untersuchung als „wahr“ herausstellt, ist das ein auf legitime Weise erzieltes und berichtenswertes Ergebnis. Soziologen sollten sich aber stets bewusst machen, welche möglichen Wirkungen mit ihrer Forschung verbunden sind. Der Ethik-Kodex der Soziologie erklärt: „Soziologinnen und Soziologen [tragen] soziale Verantwortung. Ihre Empfehlungen, Entscheidungen und Aussagen

13 Neben den Datenschutzgesetzen betont auch der Ethik-Kodex der Soziologie (Deutsche Gesellschaft für Soziologie 1992) die Bedeutung des Datenschutzes. So ist die Anonymität von Untersuchten zu wahren und „von untersuchten Personen erlangte vertrauliche Informationen müssen entsprechend behandelt werden.“ Obwohl nicht explizit ausgewiesen, gilt dies selbstverständlich auch für Daten, die gesammelt werden, um Probanden zu kontaktieren.

können das Leben ihrer Mitmenschen beeinflussen. Sie sollten sich der Situation und immanenten Zwänge bewusst sein, die zu einem Missbrauch ihres Einflusses führen könnten" (Deutsche Gesellschaft für Soziologie 1992). Insbesondere bei stigmatisierten Gruppen, wie den in der Berliner Paarstudie untersuchten deutschen Männern mit Frauen aus dem deutschen Ausland, sind Soziologen in der Pflicht, bei der Veröffentlichung von Ergebnissen behutsam vorzugehen. Aussagen über stigmatisierte Gruppen sollten nur dann erfolgen, wenn sie wissenschaftlich abgesichert sind; nicht abgesicherte Spekulationen sollten weitgehend vermieden werden, um benachteiligten Randgruppen nicht noch weiter zu schaden.

3.6 Effektivität der Stichprobenfilter

Die Studie zielte auf eine sehr spezielle Population ab. Gesucht wurden zusammenlebende Ehepaare, in denen der Mann in Deutschland geboren war oder seit Geburt die deutsche Staatsbürgerschaft innehatte. Zudem sollten der Mann und mindestens ein Elternteil Deutsch als Muttersprache haben. Die ausländischen Frauen sollten im Ausland geboren sein und eine ausländische Staatsbürgerschaft besitzen. Bei der Stichprobenziehung wurde großer Wert darauf gelegt, die anvisierte Grundgesamtheit mittels verschiedener Filter bereits so weit wie möglich zu ermitteln. Aufgrund verschiedener Ungenauigkeiten war ein gewisser Beifang aber nicht zu vermeiden. Anhand des Rücklaufes läßt sich bestimmen, wie groß der Anteil der korrekt ausgewählten Paare ist (Tabelle 3).

Bei den Kriterien Geburtsland und Staatsbürgerschaft bei Geburt der Frau wurden hohe Trefferquoten erzielt. In der deutschen Unterstichprobe sind 95,2 % der Frauen in Deutschland geboren, in den polnischen, thailändischen und brasilianischen Unterstichproben sind zwischen 98,9 % und 100 % der Frauen im anvisierten Land geboren. Lediglich in der russischen Stichprobe liegt die Trefferquote niedriger, dort sind 6,9 % der Frauen in einem anderen Land geboren. Es handelt sich dabei ausschließlich um Republiken der ehemaligen Sowjetunion (Weißrussland, Usbekistan, Kasachstan etc.). Alle Frauen in den ausländischen Unterstichproben besitzen ihre ausländische Staatsangehörigkeit bereits von Geburt an. Bei den Frauen in der deutschen Stichprobe haben 3,5 % die deutsche Staatsangehörigkeit erst später angenommen.

Bei den Männern liegt die Trefferquote bezüglich des Geburtslandes deutlich niedriger. In der deutschen, thailändischen und brasilianischen Stichprobe ist sie mit 95,2 % bis 99,0 % zufriedenstellend. Doch trotz der Namensfilterung sind in der polnischen Stichprobe nur 87,6 % aller Männer in Deutschland geboren, in der russischen Stichprobe sind es sogar nur 67,7 %. In dieser Gruppe wurden 7,9 % in

Tabelle 3 Anteil der korrekt identifizierten Paare (%)

	Deutschland	Thailand	Brasilien	Polen	Russland
Basis (Rücklauf)	311	204	103	267	254
Frau					
Im anvisierten Land geboren	95,2	99,5	100,0	98,9	91,5
Staatsbürgerschaft bei Geburt Deutsch	96,5	0,0	0,0	0,0	0,0
Mann					
In Deutschland geboren	96,5	99,0	95,2	87,6	67,7
Staatsbürgerschaft bei Geburt Deutsch	97,7	99,0	97,1	88,4	76,0
Muttersprache Deutsch	98,7	99,5	97,2	89,9	79,8
Ein Elternteil Muttersprache Deutsch	95,0	93,5	93,6	89,1	81,2
Effektivität der Stichprobenfilter					
Paar erfüllt alle Kriterien	93,9	82,4	89,3	82,8	63,0
Mann verletzt Kriterien*	4,5	3,4	5,8	15,0	29,1
Frau verletzt Kriterien*	3,5	14,2	4,9	2,6	9,8

Quelle: Berliner Paarstudie. Für jedes Merkmal ist nur die gesuchte Kategorie angegeben, um die Übersichtlichkeit zu erhöhen.

* Diese beiden Kategorien schließen sich gegenseitig nicht aus. Wenn in einem Paar beide Partner die Kriterien verletzen, steigt die Gesamtprozentzahl auf über 100.

einem dritten Land geboren, wobei es sich wieder um ehemalige Sowjetrepubliken handelt (Kasachstan, Ukraine, Tadschikistan). Entsprechend ist auch die Verteilung der Staatsbürgerschaft bei Geburt: Während die Männer in der deutschen, thailändischen und brasilianischen Stichprobe überwiegend in Deutschland zur Welt gekommen sind, liegt diese Quote in den anderen beiden Gruppen entsprechend niedriger. Analog dazu stellt sich auch die Verteilung von Deutsch als Muttersprache dar.

Der Erfolg der Namensfilterung ist somit für die fünf Gruppen unterschiedlich. In der deutschen, thailändischen und brasilianischen Stichprobe wurde auf diese Weise eine sehr hohe Genauigkeit der Stichprobenziehung erzielt. In der polnischen und russischen Gruppe war der Erfolg hingegen schwächer. Zwar konnte bei der Namensfilterung bereits etwa die Hälfte der gezogenen Adressen ausgeschlossen werden, die Zahl der „falsch“ gewählten Männer ist jedoch weiterhin beträchtlich.

Abschließend kann man die einzelnen Informationen kombinieren und so diejenigen Paare ermitteln, auf die alle Kriterien zutreffen.¹⁴ An dieser Erfolgsbilanz

14 Dabei sind alle Kriterien mit *und* verknüpft, mit folgender Ausnahme: Der Mann muss in Deutschland geboren sein *oder* bei Geburt die deutsche Staatsangehörigkeit gehabt haben. Gleiches gilt für Frauen in der deutschen Gruppe.

lässt sich die Güte des Vorgehens bei der Stichprobenziehung ablesen, denn wenn alle Paare alle Kriterien erfüllten, wäre die Stichprobenziehung dieser seltenen Population perfekt verlaufen. Eine besonders gute Auswahl wurde in der deutschen Unterstichprobe mit 93,9 % erzielt. Auch in der brasilianischen Stichprobe gehören neun von zehn Paaren zu der anvisierten Population. Bei den thailändischen Paaren fällt dieser Wert etwas schlechter aus, was meist auf fehlende Angaben bei den Frauen zurückgeht. In dieser Gruppe haben viele Frauen entweder ihr Geburtsland oder ihre Staatsbürgerschaft bei Geburt nicht angegeben und konnten deshalb nicht eindeutig zugeordnet werden. In der polnischen und der russischen Gruppe erfüllen 82,8 % bzw. nur 63,0 % alle Kriterien, was meist auf die Angaben der Ehemänner zurückgeht. Sie stammen, wie bereits besprochen, teilweise aus dem Herkunftsland der Frau. In der russischen Gruppe könnte ein etwas höherer Wert erzielt werden, wenn man auch ehemalige Mitgliedsstaaten der Sowjetunion als Herkunftsland der Frau zulässt.

Insgesamt ist das Ergebnis sehr zufriedenstellend. Mit der Stichprobenziehung ist es gelungen, die anvisierte seltene Stichprobe nahezu exakt einzugrenzen. Während lediglich 0,5 % der Grundgesamtheit der deutschen Wohnbevölkerung zu der Population deutscher Männer mit im Ausland geborenen Ehefrauen gehören, ist es hier gelungen, diese Paare mit einer Genauigkeit zwischen 63,0 % und 89,3 % aus dem Melderegister der Stadt Berlin herauszufiltern. Das ist eine beachtlich genaue Stichprobenziehung, bei der sich der Beifang zwischen 10,7 % und 17,6 % verschmerzen lässt. Lediglich in der russischen Unterstichprobe liegt mit 37 % ein hoher Beifang vor.

4 Schlussbemerkungen

Ausgangspunkt dieses Beitrags war ein Forschungsvorhaben zur Partnerwahl bei binationalen Ehepaaren. Hierzu wurde erstens eine ausreichend große Stichprobe benötigt, zweitens mussten sich die abgefragten Informationen auf den Zeitpunkt der Eheschließung beziehen. Binationale Ehepaare stellen jedoch eine äußerst seltene Population dar, sodass sie nicht in ausreichend großer Zahl in bereits existierenden Datensätzen repräsentiert sind. Es galt also, für das Forschungsprojekt eine eigene Erhebung durchzuführen, wofür eine Zufallsstichprobe aus dieser seltenen Population benötigt wurde. Die Wahl fiel auf ein mehrstufiges Filterverfahren, das zuerst eine Stichprobe aus dem Melderegister mit verknüpften Partnerdaten und zweitens eine onomastische Filterung der Männer beinhaltete. Auf diese Weise konnten in den einzelnen realisierten Unterstichproben Trefferquoten von 63,0 % bis 89,3 % erzielt werden. Das Filterverfahren zeigte seine Stärken vor allem bei

Paaren, in denen die Frau aus Thailand oder Brasilien stammt, ferner auch bei Paaren mit Frauen aus Polen. Bei deutsch-russischen Paaren stellte sich das Verfahren als etwas schwächer heraus, da aus Russland stammende Männer, die vermutlich als Aussiedler die deutsche Staatsangehörigkeit angenommen haben, häufig nicht an ihren Namen erkennbar waren. Dies ließe sich mit dem von Salentin (2007) vorgeschlagenen Verfahren noch verbessern, bei dem der Geburtsort für die Auswahl verwendet wird. Doch ist dieses Vorgehen ungleich kostenintensiver, hier wäre je nach Zielgruppe zwischen Kosten und Nutzen abzuwägen.

Die Verknüpfung von Daten im Familienverband, wie sie das hier vorgestellte Projekt genutzt hat, ermöglicht schließlich noch eine Reihe von anderen Anwendungen, die für die Familiensoziologie und andere Forschungsgebiete von Interesse sein können. Spezielle Stichproben lassen sich auf diese Weise um ein Vielfaches punktgenauer ziehen und Beifang vermeiden. Für die Scheidungsforschung ist es auf diese Weise z. B. möglich, getrennt lebende Ehepaare zu ermitteln, indem man die Anschrift beider Partner vergleicht. Über die Einbeziehung des Geschlechts der Partner ließen sich auch gleichgeschlechtliche Paare finden, die nach dem Lebenspartnerschaftsgesetz verheiratet sind. Über die Verknüpfung der Daten von Eltern und Kindern wäre es u. a. auch möglich, Familien mit Zwillingen sowie kinderlose oder kinderreiche Familien auszuwählen und zudem noch die Alterskohorte der Eltern bzw. der Kinder zu bestimmen. Über die Kombination der Familiennamen der Eltern und der Kinder sowie den zeitlichen Abstand zwischen Geburtsdatum der Kinder und dem Heiratsdatum der Eltern wäre es zudem möglich, mit einer gewissen Wahrscheinlichkeit Stieffamilien zu ermitteln.¹⁵ Derzeit gibt eine Reihe von aufwendigen Studien, die sich mit diesen und ähnlichen Themen befassen, wie Pairfam und den Gender and Generations Survey. Für eventuelle Erweiterungen dieser Studien oder andere Projekte, die sich z. B. mit binationalen Partnerschaften befassen,¹⁶ ermöglicht dieses Vorgehen sehr spezielle Stichproben. Inwiefern Meldeämter diese Datenverknüpfung öffentlichen Stellen wie Universitäten ermöglichen, hängt vermutlich jeweils vom Einzelfall ab. Wenn es umsetzbar ist, kann dieses Vorgehen bei der Stichprobenziehung sehr hilfreich sein.

15 Ein linguistisches Projekt zur Entwicklung von Zweisprachigkeit bei Babys hat kürzlich eine Stichprobe aus dem Berliner Melderegister gezogen, bei dem das Geburtsdatum des Kindes und die Staatsangehörigkeiten der Eltern kombiniert wurden („Sprachvergleichende Untersuchungen zur Entwicklung von Rhythmuspräferenzen“, <http://www.uni-potsdam.de/babylab/projects/language.html>).

16 So beispielsweise das internationale Forschungsprojekt „Toward a European Society: Single Market, Binational Marriages, and Social Group Formation in Europe (EUMARR)“ (<http://www.ibe.org>).

Literatur

- Albers, I., 1997: Einwohnermelderegister-Stichproben in der Praxis. Ein Erfahrungsbericht. S. 117-125 in: S. Gabler und J. Hoffmeyer-Zlotnik (Hg.): Stichproben in der Umfragepraxis. Opladen: Westdeutscher Verlag.
- Alt, C., W. Bien und D. Krebs, 1991: Wie zuverlässig ist die Verwirklichung von Stichprobenverfahren? Random route versus Einwohnermeldeamtsstichprobe. ZUMA-Nachrichten 28: 65-72. (PID): <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-209813>
- Amt für Statistik Berlin Brandenburg, 2009: Bevölkerungsstand. <http://www.statistik-berlin-brandenburg.de> (02.07.2009).
- Babka von Gostomski, C. und M. Pupeter, 2008: Zufallsbefragung von Ausländern auf Basis des Ausländerzentralregisters. Erfahrungen bei der Repräsentativbefragung „Ausgewählte Migrantengruppen in Deutschland 2006/2007“ (RAM). Methoden – Daten – Analysen 2(2): 149-177. (PID): <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-126673>
- Brüderl, J., L. Castiglioni, U. Krieger und K. Pforr, 2007: Design und Feldphase des Mini-Panels. http://www.mzes.uni-mannheim.de/publications/papers/Minipanel-Methodological_Report_070711.pdf (23.02.2009).
- Deutsche Gesellschaft für Soziologie, 1992: Ethik-Kodex. Fassung vom 27.11.1992. <http://www.sociologie.de/index.php?id=19> (09.04.2010).
- Diehl, C., 2007: Materialband und Endbericht zur Neuzuwandererbefragung – Pilotstudie. Erste und zweite Welle. Wiesbaden: Bundesinstitut für Bevölkerungsforschung.
- Diekmann, A., 2008: Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. 19. Auflage. Reinbek bei Hamburg: Rowohlt.
- Feld, S. L., 1981: The Focused Organization of Social Ties. *American Journal of Sociology* 86(5): 1015-1035.
- Gabler, S., 1992: Schneeballverfahren und verwandte Stichprobendesigns. ZUMA-Nachrichten 31: 47-69. (PID): <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-210848>
- Gabler, S. und S. Häder, 1997: Überlegungen zu einem Stichprobendesign für Telefonumfragen in Deutschland. ZUMA-Nachrichten 41(41): 7-18. (PID): <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-208339>
- Gerhards, J. und S. Hans, 2009: From Hasan to Herbert: Name-Giving Patterns of Immigrant Parents between Acculturation and Ethnic Maintenance. *American Journal of Sociology* 114(4): 1102-1128.
- Glowsky, D., 2011: Globale Partnerwahl. Soziale Ungleichheit als Motor transnationaler Heiratsentscheidungen. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Granato, N., 1999: Die Befragung von Arbeitsmigranten. Einwohnermeldeamt-Stichprobe und telefonische Erhebung? ZUMA-Nachrichten 23(45): 44-60. (PID): <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-208119>
- Groves, R. M., F. J. Fowler, M. P. Couper, J. M. Lepkowski et al., 2004: *Survey Methodology*. Hoboken, NJ: Wiley-Interscience.
- Häder, S. und S. Gabler, 1998: Ein neues Stichprobendesign für telefonische Umfragen in Deutschland. S. 69-88 in: S. Gabler, S. Häder und J. Hoffmeyer-Zlotnik (Hg.): *Telefonstichproben in Deutschland*. Opladen: Westdeutscher Verlag.
- Hoffmeyer-Zlotnik, J., 1997: Random-Route-Stichproben nach ADM. S. 33-42 in: S. Gabler und J. Hoffmeyer-Zlotnik (Hg.): *Stichproben in der Umfragepraxis*. Opladen: Westdeutscher Verlag.
- Hoffmeyer-Zlotnik, J., 2003: New Sampling Designs and the Quality of Data. S. 205-217 in: A. Ferligoj und A. Mrvar (Hg.): *Developments in Applied Statistics*. Ljubljana: FDV.

- Humpert, A. und K. Schneiderheinze, 2000: Stichprobenziehung für telefonische Zuwandererumfragen. Einsatzmöglichkeiten der Namensforschung. ZUMA-Nachrichten 24(47): 36-64. (PID): <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-208052>
- Humpert, A. und K. Schneiderheinze, 2002: Stichprobenziehung für telefonische Zuwandererumfragen – Praktische Erfahrungen und Erweiterung der Auswahlgrundlage. S. 187-208 in: S. Gabler und S. Häder (Hg.): Telefonstichproben – Methodische Innovationen und Anwendungen in Deutschland. Münster: Waxmann.
- Kalmijn, M., P. M. d. Graaf und J. P. G. Janssen, 2005: Inter-marriage and the Risk of Divorce in the Netherlands: The Effects of Differences in Religion and in Nationality, 1974-94. *Population Studies* 59(1): 71-85.
- Kish, L., 1965: *Survey Sampling*. New York: Wiley.
- Lüttinger, P. und T. Riede, 1997: Der Mikrozensus. Amtliche Daten für die Sozialforschung. ZUMA-Nachrichten 41: 19-44. (PID): <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-208344>
- Niesner, E., 2000: Mythos und Wirklichkeit auf einem bikulturellen Heiratsmarkt. S. 163-181 in: E. Friebe-Blum (Hg.): *Wer ist fremd? Ethnische Herkunft, Familie und Gesellschaft*. Opladen: Leske + Budrich.
- Reuschke, D., 2010: *Multilokales Wohnen: Raum-zeitliche Muster multilokaler Wohnarrangements von Shuttles und Personen in einer Fernbeziehung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Roloff, J., 1998: Eheschließungen und Ehescheidungen von und mit Ausländern in Deutschland. *Zeitschrift für Bevölkerungswissenschaft* 23(3): 319-334.
- Salentin, K., 1999: Die Stichprobenziehung bei Zuwandererbefragungen. ZUMA-Nachrichten 23(45): 115-135. (PID): <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-208159>
- Salentin, K., 2002: Zuwandererstichproben aus dem Telefonbuch: Möglichkeiten und Grenzen. S. 164-186 in: S. Gabler und S. Häder (Hg.): *Methodische Innovationen und Anwendungen in Deutschland*. Münster: Waxmann.
- Salentin, K., 2007: Die Aussiedler-Stichprobenziehung. *Methoden – Daten – Analysen* 1(1): 25-44. (PID): <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-126439>
- Schnell, R., P. B. Hill und E. Esser, 2008: *Methoden der empirischen Sozialforschung*. München/Wien: Oldenbourg.
- Statistische Ämter des Bundes und der Länder, 2004: Ergebnisse des Zensus 2004. *Wirtschaft und Statistik* 8/2004: 813-833.
- Straßburger, G., 2000: Das Heiratsverhalten von Personen ausländischer Nationalität oder Herkunft in Deutschland. Sachverständigenkommission 6. Familienbericht (Hg.): *Familien ausländischer Herkunft in Deutschland: Empirische Beiträge zur Familienentwicklung und Akkulturation*. Materialien zum 6. Familienbericht. Opladen: Leske + Budrich.
- Straßburger, G., 2003: Heiratsverhalten und Partnerwahl im Einwanderungskontext. Eheschließungen der zweiten Migrantenmigration türkischer Herkunft. Würzburg: Ergon.
- Terwey, M., 2000: ALLBUS: A German General Social Survey. *Schmollers Jahrbuch* 120(1): 151-158.
- Wagner, G. G., J. R. Frick und J. Schupp, 2007: The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. *Schmollers Jahrbuch* 127(1): 139-169.
- Wasmer, M., E. Scholz und M. Blohm, 2010: Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 2008. http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2010/TechnicalReport_10-4.pdf (27.05.2011).

Anschrift des Autors

David Glowsky
Humboldt-Universität zu Berlin
Präsidialbereich, Stabsstelle Internationalisierung
Unter den Linden 6
10099 Berlin
E-Mail: david.glowsky@hu-berlin.de