Die Analyse sozialer Ungleichheit

Konzeptuelle Überlegungen und empirische Erkenntnisse

Dissertation zur Erlangung des akademischen Grades des
Doktors der Sozialwissenschaften (Dr. rer. soc.) an der

Universität Konstanz
Fachbereich Geschichte und Soziologie

vorgelegt von
Katrin Auspurg

Tag der mündlichen Prüfung: 22. Juli 2010

Referenten:
Prof. Dr. Thomas Hinz
Prof. Dr. Werner Georg
Prof. Dr. Monika Jungbauer-Gans

Konstanzer Online-Publikations-System (KOPS)
URN: http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-124716
URL: http://kops.ub.uni-konstanz.de/volltexte/2010/12471/

Vorwort

Die vorliegende Dissertation wurde im Zeitraum Juli 2006 bis Juni 2010 am Arbeitsbereich „Empirische Sozialforschung“ (Prof. Dr. Thomas Hinz) des Fachbereichs Geschichte und Soziologie an der Universität Konstanz verfasst. Kapitel 4 sowie einige Anschlussarbeiten, auf die in der Einleitung verwiesen wird, sind im Kontext des von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projektes „Der Faktorielle Survey als Instrument zur Einstellungsmessung in Bevölkerungsumfragen“ entstanden.

Viele Personen haben zur erfolgreichen Erstellung dieser Dissertation beigetragen. An erster Stelle möchte ich meinem Doktorvater Thomas Hinz danken, der mich die ganze Zeit über außerordentlich gefördert, motiviert und inspiriert hat. Thomas Hinz hat mein wissenschaftliches Vorankommen vor allem auch durch den stetigen Glauben an das Potenzial der sehr vielen (gemeinsamen) Forschungsprojekte vorangetrieben. Ihm ist überdies für ein äußerst freies, abwechslungsreiches und vertrauensvolles Arbeitsklima zu danken, welches mir den Weg durch die Wissenschaft geebnet und mir zugleich manch „Kleinstadt-Koller“ erspart hat.

Prof. Dr. Martin Abraham hat wesentlich dazu beigetragen, dass ich das Unternehmen einer Promotion gut gerüstet gewagt habe und die Neugierde für spannende Forschungsfragen entwickeln konnte. Prof. Dr. Monika Jungbauer-Gans danke ich für die spontane Bereitschaft, eines der Referate zu dieser Dissertation zu übernehmen. Am Promotionsverfahren mitgewirkt haben zudem Prof. Dr. Werner Georg und Prof. Dr. Sabine Sonnentag – auch ihnen sei dafür herzlich gedankt! Die Möglichkeit einer in verschiedenster Hinsicht einmaligen Datenerhebung parallel zu dieser Dissertation verdanke ich der Projektgruppe zum Faktoriellen Survey – namentlich zu danken ist hier neben Thomas Hinz auch Carsten Sauer und Stefan Wehrli für die umfangreiche Projektarbeit sowie Prof. Dr. Stefan Liebig für inspirierende Diskussionen. Nicht persönlich attribuierbar sind die vielfältigen Ideen und Anregungen, die ich bei den jährlichen Seminaren in Venedig und bei den Forschungskolloquien in Konstanz gewonnen habe.

Bei der Bewältigung von vielerlei Tücken des Alltags – von streikenden Kopierern über die rätselhafte Logik mancher Computerprogramme bis hin zur Herausforderung, sich trotz ständig wandelnder Formatierungswünsche von Kooperationspartnern bei Laune zu halten – haben mich viele Kollegen begleitet. Für umfangreiche praktische und moralische Hilfen ist besonders Sophia Prinz zu danken sowie den unmittelbaren Büronachbarn Thomas Wöhler, Ina Findeisen und Eva Amorelli. Ihnen und den (überregionalen) Kollegen Jochen Groß, Christiane Gross, Andreas Huchler, Gaia di Luzio, Marion Wölki sowie Ines Marold und Alban Knecht gilt zudem mein Dank für wertvolle Ratschläge, inspirierende Diskussionen und motivierende Worte. Mit größter Sorgfalt hat Tonie Kurtz Ausschnitte dieser Dissertation lektoriert und mir durch ihre ehrliche und konstruktive Kritik sehr geholfen.

Hervorzuheben ist überdies die Leistung von zahlreichen studentischen Hilfskräften, die mir mit viel Geduld bei der täglichen Kleinarbeit zur Seite standen, mich durch ihr gründliches Korrekturlesen vor manchem Fehler bewahrt oder auch durch ihr Interesse und Humor für eine erfreuliche Arbeitsatmosphäre gesorgt haben. Ohne die Unterstützung dieser „Forschungssklaven“ (Münch 2007) wären die umfangeichen Datenerhebungen, die dieser Dissertation zugrunde liegen, nicht möglich gewesen. Der Dank geht hier vor allem an Anja Joos, Ksenija Ursulenko, Cornelius Groß, Judith Tonner und Mirte Scholten.

Schließlich möchte ich meiner Familie meinen ausdrücklichen Dank aussprechen – für die Unterstützung meines gesamten Ausbildungsweges und für das Verständnis, dass die Zeit einer Wissenschaftlerin stets knapp bemessen ist. Zuletzt bleibt all denjenigen zu danken, die hier nicht namentlich genannt werden konnten: Danksagungen sind ein erstes Beispiel dafür, dass knappe (Platz-)Ressourcen soziale Selektionen erfordern.

Konstanz, September 2010

Katrin Auspurg

Inhaltsverzeichnis

1 Einführung und Überblick: Die Analyse sozialer Ungleichheit 8
(Katrin Auspurg)

1.1 Einführung 9

1.2 Programm einer analytischen Soziologie 12

1.2.1 Anforderungen an Erklärungen 12

1.2.2 Anforderungen an empirische Untersuchungen 16

1.2.3 Stand der Ungleichheitssoziologie 19

1.2.4 Zwischenfazit: Ein Abschied von Klassen und Schichten? 26

1.3 Inhalte und Erkenntnisleistungen der eigenen Anwendungsbeispiele 30

1.3.1 Herausbildung einer akademischen Elite? 30

1.3.2 Migration Decisions Within Dual-Earner Partnerships 34

1.3.3 Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey 37

1.3.4 Bedeutung der Publikationsorgane und eigene Leistung 40

1.4 Fazit und Ausblick 43

Literatur 48

2 Herausbildung einer akademischen Elite? 58
(Katrin Auspurg, Thomas Hinz und Jürgen Güdler)

2.1 Einleitung 60

2.2 Argumentation bei Münch, Forschungsstand und Hypothesen 62

2.3 Datengrundlage und Methodik 68

2.3.1 Empirische Auswertungen bei Münch: Vorgehen und Kritik 68

2.3.2 Eigene Datengrundlage und Methodik 69

2.4 Ergebnisse 72

2.4.1 Einfluss kontextueller Merkmale auf die Bewilligungschancen 72

2.4.2 Einfluss von Personalbeständen auf die Antragsaktivität 79

2.4.3 Ungleichheitsverteilung von Anträgen und Bewilligungen auf Universitäten im Zeitverlauf 82

2.5 Diskussion und Ausblick 84

Literatur 87

Methodischer Anhang 90

3 Migration Decisions Within Dual-Earner Partnerships 92
(Martin Abraham, Katrin Auspurg, and Thomas Hinz)

3.1 Introduction 93
3.2 Theoretical Framework 94
3.2.1 Bargaining Models 94
3.2.2 The Decision to Move the Household 95
3.2.3 The Influence of Gender 96
3.3 Hypotheses 97
3.4 Method 99
3.5 Results 102
3.5.1 Descriptive Results and Sample Description 102
3.5.2 Test of the Hypotheses 105
3.6 Discussion 110

References **113**

4 Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey 116
(Katrin Auspurg, Thomas Hinz und Stefan Liebig)

4.1 Einleitung 118
4.2 Faktorieller Survey: Aufbau, Motivation und Probleme 119
4.3 Forschungsstand und Hypothesen 122
4.3.1 Komplexität von Vignetten: Anzahl der Dimensionen 122
4.3.2 Lern- und Ermüdungseffekte 124
4.3.3 Behandlung unlogischer Fälle 125
4.4 Methodik und Datengrundlage 126
4.5 Ergebnisse 130
4.5.1 Deskriptive Befunde 130
4.5.2 Multivariate Analysen 131
4.6 Zusammenfassung und Schlussfolgerungen 140

Literatur 144
Methodischer Anhang 147

5 Zusammenfassung 150
(Katrin Auspurg)

Anhang
Literaturverzeichnis 156

Tabellenverzeichnis

Kapitel 1: Einführung und Überblick: Die Analyse sozialer Ungleichheit

Tabelle 1.1: Erscheinungsorte im Überblick ... 41

Kapitel 2: Herausbildung einer akademischen Elite?

Tabelle 2.1: Übersicht über die in den Regressionsanalysen verwendeten Variablen ... 71
Tabelle 2.2: Logistische Regressionen der Bewilligungschancen (Random Intercept Modelle; 1993 bis 2004 gesamt) ... 74
Tabelle 2.3: Logistische Regressionen der Bewilligungschancen (Random Intercept Modelle; 1996 bzw. 2000 bis 2004 gesamt) ... 77
Tabelle 2.4: Negativ-Binomial Regressionen der Antragszahlen bzw. Bewilligungen ... 80

Kapitel 3: Migration Decisions Within Dual-Earner Partnerships

Table 3.1: Sample Description ... 104
Table 3.2: Summary of Linear Regression Analyses for Variables Predicting the Willingness to Move (Random Intercept Models) ... 106
Table 3.3: Summary of Linear Regression Analyses for Variables Predicting the Conflict Potential (Random Intercept Models) ... 107
Table 3.4: Summary of Linear Regression Analyses for Variables Predicting the Willingness to Move of Men and Women (Random Intercept Models) ... 110

Kapitel 4: Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey

Tabelle 4.1: Realisierte Fallzahlen für Vignettenurteile und Befragte ... 130
Tabelle 4.2: Deskriptive Übersicht über die Vignettenurteile ... 131
Tabelle 4.3: OLS-Regressionen der Vignettenurteile ... 133
Tabelle 4.4: OLS-Regressionen der Vignettenurteile in Abhängigkeit vom Auftreten unplausibler Fälle ... 138
Tabelle 4.5: OLS-Regressionen der quadrierten Residuen und Bearbeitungszeiten pro Vignette ... 140
Tabelle 4.A1: Übersicht über die Korrelationen der Dimensionen in den Splits mit 5 und 12 Dimensionen ... 148
Tabelle 4.A2: Übersicht über die Korrelationen der Dimensionen in den Splits mit 7 und 10 Vignetten pro Befragten ... 148
Tabelle 4.A3: Übersicht über die Korrelationen aller zwölf Vignettendimensionen (nur Split mit 12 Dimensionen) ... 149

Abbildungsverzeichnis

Kapitel 1: Einführung und Überblick: Die Analyse sozialer Ungleichheit

Abb. 1.1: Erklärungsschritte, Ebenen der Beobachtung und Erklärung 14

Kapitel 2: Herausbildung einer akademischen Elite?

Abb. 2.1: Professuren und geschätzte Anzahlen bewilligte Anträge in den Geistes- und Sozialwissenschaften 81

Abb. 2.2: Professuren und geschätzte Anzahlen bewilligte Anträge in den Naturwissenschaften 81

Abb. 2.3: Gini-Koeffizienten der bewilligten Anträge in den Geistes- und Sozialwissenschaften 83

Abb. 2.4: Gini-Koeffizienten der bewilligten Anträge in den Lebenswissenschaften 83

Abb. 2.5: Gini-Koeffizienten der bewilligten Anträge in den Naturwissenschaften 84

Abb. 2.6: Gini-Koeffizienten der bewilligten Anträge in den Ingenieurwissenschaften 84

Kapitel 3: Migration Decisions Within Dual-Earner Partnerships

Figure 3.1: Example of a Vignette (Male Perspective, Own Job Offer) 100

Figure 3.2: Willingness to Move of Ego and Alter 103

Figure 3.3: Conflict Potential (Willingness to Move of Ego Minus Willingness to Move of Alter) 103

Kapitel 4: Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey

Abb. 4.1: Beispielvignette mit zwölf Dimensionen 128

Abb. 4.2: R^2-Werte und Bearbeitungszeiten pro Vignette in Abhängigkeit von der Position der Vignette und Anzahl ihrer Dimensionen 135

1

Einführung und Überblick: Die Analyse sozialer Ungleichheit

(Katrin Auspurg)

1 Einführung und Überblick: Die Analyse sozialer Ungleichheit

1.1 Einführung

Soziale Ungleichheit ist nach verbreiteter Auffassung *das* Identität stiftende Thema der Soziologie. Nach Dahrendorf (1970: 4) markieren etwa die Fragen, warum es Ungleichheit gibt und was ihre Ursachen sind, den historischen Beginn der Disziplin. Andere Autoren halten soziale Ungleichheit für das „bevorzugte Forschungsproblem“ und den „ureigensten Gegenstand“ (Müller/Schmid 2003: 6) der Soziologie. Die hohe Relevanz von sozialen Ungleichheiten wird ebenso unterstrichen, wenn sie als die „gesellschaftstheoretische Schlüsselfrage“ (Kreckel 2004: 21) oder die „Hauptachse gesellschaftlicher Differenzierung“ (Mayer 1987: 377) hervorgehoben werden. Angesichts dessen erstaunt, dass die Leistungen auf dem Gebiet der Ungleichheitsforschung – trotz vielfacher und zunehmender Beiträge (Verwiebe 2010 für aktuelle Statistiken zu deutschsprachigen Zeitschriften) – nach wie vor als unbefriedigend gelten (z. B. J. Berger 2004: 360; Blalock 1991: ix ff.; Reskin 2003: 2; Schwinn 2007). Zugespitzt bewegt sich die Kritik zwischen den scheinbar unvereinbaren Gegensätzen, dass die vorgeschlagenen Konzepte entweder zu abstrakt sind, um die zunehmend ausdifferenzierten und komplexen Ungleichheitsstrukturen adäquat abzubilden; oder sie zwar die bestehenden Verteilungen sehr exakt einfangen, dann aber gerade aufgrund dieser Detailtreue unzureichend in der Lage sind, über punktuelle Gesellschaften und Zeiten hinausreichende, *generelle Erklärungen* für deren Entstehung zu liefern. Mehrfache Versuche der Neukonzeption haben bislang nicht zu einer weithin anerkannten sowie in der Erklärung und Prognose bewährten Theorie geführt. Drohende Gefahren einer fehlenden Übereinkunft über ein bewährtes Forschungsprogramms sind, dass die Soziologie in ihrer gesellschaftspolitischen Geltung und als Informant für andere Disziplinen an Bedeutung verliert (Braun 2008).

Anliegen der vorliegenden Dissertation ist es, wenig beachtete Gründe für diesen schleppend verlaufenden Erkenntnisfortschritt aufzuzeigen und ein mögliches Forschungsprogramm zu seiner Überwindung anzuregen. Die Kernthese ist, dass sich die Soziologie sozialer Ungleichheit zu sehr mit der Suche nach allumfassenden Gesellschaftstheorien oder möglichst realitätsnahen Beschreibungen aufhält, anstatt den aussichtsreicheren Mittelweg der Weiterentwicklung und empirischen Prüfung Ungleichheit generierender *Mechanismen* zu gehen. Es erscheint vordringlicher, sich nicht auf die Entdeckung ständig „neuer“ Ungleichheiten zu konzentrieren und nach adäquaten Konzepten zu ihrer Fassung zu suchen (nun etwa um transnationale Lagen erweitert), sondern stattdessen die Ungleichheit *generierenden Prozesse* besser zu verstehen, unabhängig davon, auf welche Ressourcenverteilungen, Lebensstilmuster oder Gruppenkonstellationen sich diese im Einzelfall beziehen. Zu diesem Forschungsprogramm gehört auch, das Spektrum an empirischen Methoden stärker für die Ungleichheitssoziologie auszuschöpfen und weiterzuentwickeln.

Anspruch vorliegender Abhandlung ist es somit nicht, eine erneute Theorie sozialer Ungleichheit vorzuschlagen. Im Gegenteil soll gezeigt werden, dass die Ungleichheitsforschung vorerst mehr davon profitiert, das andernorts bereits gut ausgearbeitete Programm einer erklärenden bzw. analytischen Soziologie umzusetzen. „Das Hauptproblem der Soziologie heute besteht darin, dass sie die falschen Fragen stellt" – diese Ansicht von Ulrich Beck (2008: 42) wird geteilt, die richtigen Fragen werden aber hier abweichend nicht in *Was*- und *Wer*- (U. Beck 2008: 17), sondern in *Wie*- und *Warum*-Fragen gesehen. Die Erkenntnisgewinne, die sich durch einen stärkeren Fokus auf erklärende Mechanismen erzielen lassen, werden dann auch speziell anhand von drei empirischen Anwendungsbeispielen illustriert.

Zunächst aber zum Gegenstand – was sind soziale Ungleichheiten? Nach allen gängigen Definitionen sind dies solche Unterschiede, die nicht rein natürlich bedingt sind und zugleich eine Ungleich*wertigkeit* implizieren (Bolte 1983: 393; Burzan 2007: 7; Kreckel 2004: 15ff.; Solga et al. 2009: 14f.). Es geht also nicht um beliebige Andersartigkeiten, sondern um ungleiche Zugangschancen zu begehrten Ressourcen. Da sich die Vorstellungen vom „Wünschenswerten" und die zu seiner Erlangung erforderlichen Ressourcen wandeln, sind ebenso die Erscheinungsformen sozialer Ungleichheit eine Funktion der jeweiligen Zeit und Gesellschaft. Welche Unterschiede im historischen Einzelfall eine soziale Ungleichheit bedeuten, ist somit bereits ein erster wichtiger Forschungsgegenstand (Burzan 2007: 7).[1] Bezeichnend für die soziologische Perspektive ist aber vor allem, sich für die *Ursachen* sozialer Ungleichheit zu interessieren. Der Bruch mit dem Weltbild vormoderner Gesellschaften, Lebenschancen als natürlich und gottgewollt zu sehen, kennzeichnet nicht nur die Geburtsstunde der Soziologie, sondern liefert zugleich die Einsicht, dass soziale Ungleichheiten wandelbar und legitimationsbedürftig sind. Welche Verteilungsprinzipien und -ergebnisse anzustreben sind, fällt als normative Frage vornehmlich in den Zuständigkeitsbereich der Philosophie und Politik. Beitrag der Soziologie ist es hingegen, die real bestehenden Ungleichheitsstrukturen durch geeignete Konzepte fassbar zu machen sowie ihre Ursachen zu erforschen. Nur mit diesem Wissen sind die Verteilungen begehrter Ressourcen den gesellschaftspolitischen Idealen anzunähern. Es geht also in der Soziologie nicht nur um die Beschreibung, sondern vor allem auch um die Aufdeckung der ursächlichen Prozesse von Ungleichheitsstrukturen und um die Erklärung ihres Wandels.

In einer engeren Definition – welche auch der vorliegenden Arbeit zugrunde gelegt wird – bilden den *hauptsächlichen* Gegenstand der Soziologie solche soziale Ungleichheiten, die nicht durch Leistungs-, Begabungs- oder Qualifikationsunterschiede bedingt sind und somit dem in modernen Gesellschaften anerkannten Prinzip der Leistungsgerechtigkeit bzw. *Meritokratie* zuwiderlaufen

[1] In modernen Gesellschaften sind beispielsweise Bildungsabschlüsse für den Zugang zu herausgehobenen Stellungen wichtig, während in vormodernen Gesellschaften askriptive Merkmale wie Adelstitel und das Geschlecht einflussreicher waren. Dass es sich bei sozialen Ungleichheiten um *soziale Konstrukte* handelt, wird besonders gut an körperlichen Merkmalen offensichtlich, welche von Verschiedenartigkeiten ohne wesentliche Bedeutung für Lebenschancen zu deren zentralen Determinanten avancieren können. So vollzieht sich etwa in jüngster Zeit eine derartige Transformation in Bezug auf das Körpergewicht – Übergewichtige haben zunehmend schlechte Zugangschancen zu begehrten Arbeitsmarktpositionen (Kreckel 2004: 15f.; Solga et al. 2009: 16).

(J. Berger 2004; Meulemann 2004). Das Aufgreifen dieses Legitimationsprinzips ist weniger als kritikloser Bruch mit dem Postulat der Werturteilsfreiheit zu verstehen, sondern vielmehr als pragmatische Fokussierung auf das gesellschaftlich besonders Relevante (Schmidt 2004). Denn begrenzte Ressourcen und Anfangsverteilungen machen Prioritätensetzungen erforderlich. Zudem bestehen unvermeidliche Zielkonflikte zwischen den drei Gerechtigkeitsprinzipien Bedarf, Leistung und Gleichheit (Liebig/Schupp 2008; Schmidt 2004: 88). Die Soziologie kann jedoch als Korrektiv für ideologische Einschläge dienen, indem sie die Durchsetzung von Legitimationsprinzipien und deren Wandel analysiert (was im Schnittfeld zur politischen und Wissenssoziologie liegt). Ein weiteres Randthema bilden Gerechtigkeitsvorstellungen und damit die Frage, inwiefern normative gesellschaftliche Standards (wie etwa das Verbot einer Diskriminierung) in den (unbewussten) Einstellungen und Handlungen der Bevölkerung verankert sind (Liebig et al. 2004; Neckel et al. 2004).[2] Unabhängig von einer weit oder eng gefassten Definition gilt jedenfalls, dass eine Einsicht in die Entstehungsweisen von sozialer Ungleichheit unerlässlich ist, um letztere nicht nur zu dokumentieren, sondern auch gezielt zu verändern. Will die Soziologie bei der Gestaltung von Gesellschaften mithelfen und ihrem aufklärerischen Anspruch gerecht werden, kommt sie um die Frage nach den *Ursachen* nicht herum.

In diesem ersten, konzeptuellen Kapitel wird die These eines Erkenntnisfortschritts durch eine stärker auf ursächliche Mechanismen fokussierende, analytische Ungleichheitssoziologie weiter ausgeführt. Dabei wird mitunter die Notwendigkeit einer Weiterentwicklung adäquater empirischer Methoden begründet. Die sich daran anschließenden empirischen drei Kapitel illustrieren die praktische Umsetzung dieser Prinzipien dann anhand eines bewusst breiten Spektrums von Inhalten und Methoden. Die dort präsentierten Studien beschäftigen sich im Einzelnen mit (a) der Allokation von Forschungsdrittmitteln an Universitäten (Kapitel 2) sowie (b) der Entstehung von sozialer Ungleichheit durch Karriereentscheidungen in Doppelverdiener-Partnerschaften (Kapitel 3). Gehaltvolle Analysen setzen adäquate empirische Methoden voraus und sollten mögliche „Fallstricke“ bei der Prüfung der angenommenen Mechanismen beachten. Um auch diesen Aspekt empirisch aufzugreifen, schließt sich (c) ein Kapitel mit methodischem Fokus an. Exemplarisch wird anhand einer Prüfung von Auswirkungen verschiedener Designelemente eines Faktoriellen Surveys (wie der Komplexität und Plausibilität der verwendeten Fallbeispiele) gezeigt, wie sich für die Ungleichheitssoziologie innovative, damit aber in aller Regel noch wenig methodisch geprüfte Verfahren gewinnbringend weiterentwickeln lassen (Kapitel 4). Der inhaltliche Fokus liegt in diesem Kapitel auf dem ebenfalls einschlägigen Thema der Einkommensgerechtigkeit.

[2] Gerechtigkeit und Ungleichheit sind zwar zu trennende, aber eng miteinander zusammenhängende Aspekte. So gelten beispielsweise Ungleichheitsstrukturen prinzipiell als umso stabiler, als je gerechter sie von der Bevölkerung wahrgenommen werden (Hadjar 2008: 9; Schimank 1998). Zudem stellt das Gerechtigkeitsempfinden hinsichtlich Erwerbseinkommen in vielen Theorien eine wesentliche Ursache für Aufstiegsorientierungen und Leistungsmotivationen dar (Jaime-Castillo 2008; Jasso/Webster 1997: 67). Beide Aspekte sind wiederum zentrale Bedienungsfaktoren dafür, wie hoch die soziale Mobilität in Gesellschaften ausfällt bzw. wie beharrlich ihre Ungleichheitsgefüge sind.

Diese Studien stehen jeweils für sich und sind inzwischen alle in *peer-reviewed* Zeitschriften erschienen (oder angenommen). Sie stellen praktische Anwendungen einer analytischen Ungleichheitssoziologie dar, ohne deren Prämissen explizit herauszustellen. Aufgabe der folgenden Einleitung ist es daher, zunächst allgemein die Prinzipien analytischer Soziologie (Abschnitt 1.2.1) und ihrer empirischen Umsetzung (Abschnitt 1.2.2) zu verdeutlichen. Diese werden anschließend mit dem Forschungsstand zur Ungleichheitssoziologie kontrastiert (Abschnitt 1.2.3). Nach einem Zwischenfazit (Abschnitt 1.2.4) fasst Abschnitt 1.3 die besonderen Erkenntnisleistungen der eigenen empirischen Anwendungsbeispiele in den Kapiteln 2 bis 4 zusammen. Das abschließende Resümee in Abschnitt 1.4 verweist auf offen gebliebene Fragen und Anknüpfungspunkte für die weitere Forschungsarbeit. Am Schluss der kompletten Dissertation findet sich zudem eine Zusammenfassung *aller* zentralen Inhalte und Erkenntnisse (Kapitel 5).

1.2 Programm einer analytischen Soziologie

1.2.1 Anforderungen an Erklärungen

Die erklärende, analytische oder auch Soziologie mechanistischer Erklärungen (diese Begriffe werden im Folgenden synonym verwendet) hat aktuell noch nicht den Stand eines einheitlichen Forschungsprogramms erreicht. Es handelt sich vielmehr um Strömungen, die trotz unterschiedlicher Schwerpunktsetzungen einige zentrale Prämissen teilen.[3] Zu diesen zählt zunächst das *übergeordnete Ziel, ursächliche Mechanismen* und damit *Antworten auf Wie- und Warum-Fragen* zu finden. Zudem sind folgende vier Prämissen an *Erklärungen* bzw. *Theorien* grundlegend (vgl. N. Gross 2009; Hedström 2005; Hedström/Bearman 2009a; Schmid 2006):

Erstens ist eine präzise und widerspruchsfreie Argumentation unabdingbar. Anliegen jeglicher Theoriebildung sollte es sein, das „Komplexe und scheinbar Obskure“ verständlich zu machen (Hedström 2005: 4). Dies erfordert zunächst eine genaue Klärung der Frage, auf welche Phänomene sich die Theorie beziehen soll und wie sich diese wechselseitig bedingen, oder mit anderen Worten: eine präzise Darlegung der unterstellten kausalen Mechanismen. Mit Mechanismen sind dabei Prozesse gemeint, die regelmäßig (wenngleich nicht zwingend) zu den interessierenden Phänomenen führen. Es kann sich um reale Prozesse oder rein analytische Konstrukte handeln. Eine besonders hohe Präzision bieten mathematische Formalisierungen, mit denen sich neben der Richtung ebenso die funktionale Form und Stärke von Zusammenhängen in sparsamer Weise spezifizieren lassen (etwa durch Aussagen zu Parametern in Diffusionsmodellen, vgl. hierzu Braun 2008; Diekmann 2007;

[3] Klassische Vertreter der analytischen Soziologie sind John Elster, James Coleman, Raymond Boudon und Richard Merton; zu den neueren Vertretern zählen vor allem die Soziologen um Peter Hedström sowie Anwender von Rational-Choice-Theorien (s. Diekmann et al. 2008; Diekmann/Voss 2004; Hedström/Bearman 2009b für aktuelle Sammelbände). Überblicke über Autoren und ihre unterschiedlichen Konzepte von erklärenden „Mechanismen“ bieten zudem N. Gross 2009; Hedström/Bearman 2009a sowie Mahony 2001.

Diekmann/Voss 2004; Manzo 2007).[4] Vorrangig ist allerdings das Kriterium möglichst exakter Ausführungen. Schwache Erklärungen gewinnen nichts durch komplexe Ausdrucksweisen oder ausgetüftelte mathematische Modelle, abgesehen von der – einer guten Forschungspraxis direkt zuwider laufenden – Möglichkeit einer Verschleierung von versteckten Annahmen oder gar von „Trivialität" in Form eines so genannten *Modellplatonismus* (Diekmann 2007: 144; Sokal/Bricmont 1997 für eine lesenswerte Zusammenstellung „postmoderner" Negativbeispiele).

Zweitens setzen Erklärungen ein geringeres Aggregationsniveau als das des zu Erklärenden voraus. In der Soziologie bedeutet dies, die Makroebene interessierender sozialer Phänomene (hier ungleiche Allokationen) zu verlassen, um diese mittels der Handlungswahl von Akteuren auf der Mikroebene begreiflich zu machen (Prinzip des *methodischen Individualismus*). *Kern der Erklärungen bilden damit Handlungstheorien.* Wenngleich sich die analytische Soziologie nicht auf Rational-Choice-Theorien festlegen will (s. z. B. Hedström 2005), belaufen sich die Mikroerklärungen bislang hierauf oder auf ihre Erweiterungen in Form von Spieltheorien.[5] Letztere zielen speziell darauf ab, das interdependente Verhalten von Akteuren in Entscheidungssituationen mit *strategischer Interdependenz* zu modellieren.[6] Damit werden sie dem Anspruch analytischer Soziologie besonders gut gerecht, nicht nur das Handeln von isolierten Individuen abzubilden, sondern deren soziale Interaktionen. Vereinfacht ausgedrückt besteht das Erklärungsprinzip analytischer Soziologie jedenfalls darin, dass Akteure Ziele haben, über begrenzte Ressourcen verfügen und einer Entscheidungsregel folgen. Akteure handeln verständlich in dem Sinn, als sie sich unter den gegebenen Restriktionen für diejenige Handlungsoption entscheiden, die ihnen den größten Nutzen verspricht (Diekmann/Voss 2004).

Die Unterstellung einer rationalen Handlungswahl ist dabei durchaus mit den realistischeren Annahmen unvollständiger Information, begrenzter kognitiver Bearbeitungskapazitäten (*bounded rationality*) oder einem lediglich auf Optimierung, nicht aber Nutzenmaximierung abstellenden Verhalten (*satisficing*) vereinbar (Blossfeld/Müller 1996; Kunz 2004: 155ff.). So bilden etwa Heuristiken eine sinnvolle Auslegung von Rationalität in komplexen Situationen (Gigerenzer/Todd 1999; Goldstein 2009 für Details). Abstraktionen von solchen begrenzten Informations- und Bearbeitungskapazitäten können jedoch in frühen Forschungsstadien hilfreiche Vereinfachungen bieten (Braun 2008: 391). Handlungstheorien sind ebenso keineswegs mit einer Gleichsetzung von Ursachen mit

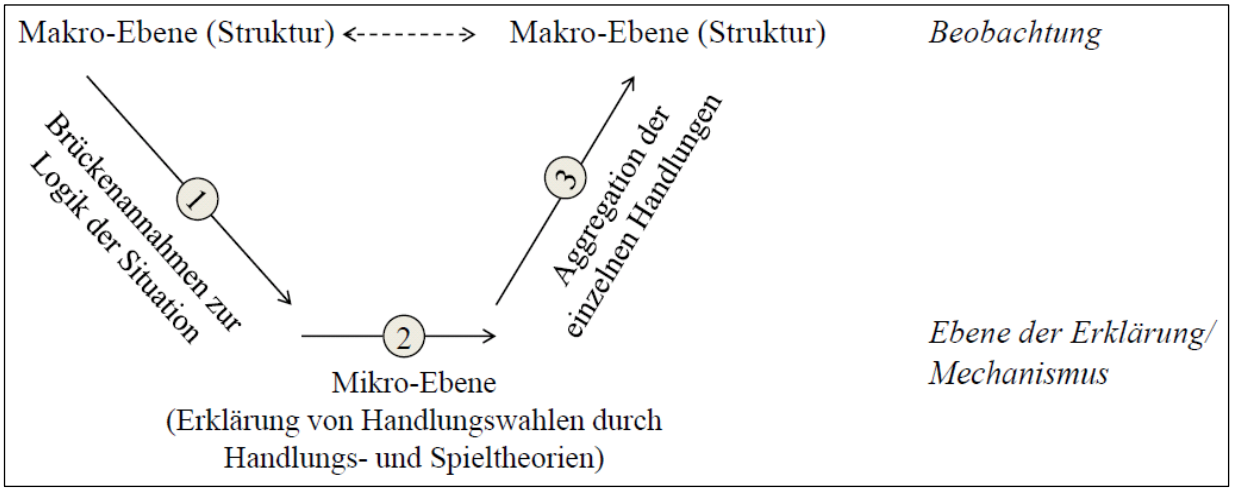
[4] Präzision ist mitunter deshalb wichtig, weil bereits kleine Unterschiede enorme Auswirkungen haben können, man denke etwa an die *Tipping*-Points bei Diffusions- und Segregationsprozessen (zu diesen Granovetter 1978; Schelling 1971).

[5] Vereinzelt werden ebenso Ansätze, die weder Handlungs- noch Spieltheorien beinhalten, zur analytischen Soziologie gezählt. Zu diesen gehören etwa Arbeiten von Aage Sørensen oder von Anhängern des kritischen Rationalismus (vgl. N. Gross 2009). Da aber gerade dieser Aspekt fehlender handlungstheoretischer Fundierung von einschlägigen Vertretern kritisiert wird (z. B. Hedström 2004), zählen sie nicht zur analytischen Soziologie im engeren Sinne und werden hier entsprechend nur am Rande berücksichtigt.

[6] Strategische Interdependenz meint, dass der Erfolg von Handlungen nicht nur vom Verhalten einzelner Akteure abhängt, sondern ebenso vom Verhalten ihrer Interaktionspartner. Die Akteure müssen in solchen Handlungssituationen ihre Entscheidungen an ihren wechselseitigen Erwartungen über das Verhalten der jeweils anderen ausrichten (s. Diekmann 2009 für eine Einführung in die Spieltheorie).

Präferenzen zu verwechseln.[7] Derartige Reduktionismen werden vermieden, indem der handlungstheoretisch begründete Mechanismus lediglich einen von drei Schritten in einem mehrere Ebenen umfassenden Erklärungsprogramm darstellt (vgl. Abbildung 1.1; grundlegend Coleman 1990): (1) Die „Logik der Situation" modelliert die strukturellen Bedingungen des sozialen Handelns und verbindet die übergeordnete Makrobene der Beobachtung mit der tiefergreifenden Ebene der Erklärung. Konkret beinhaltet dieser Erklärungsschritt die Ableitung von Kontextbedingungen für das individuelle Handeln in Form von Brückenhypothesen, also etwa Spezifikationen der für bestimmte Gruppen relevanten Ressourcen und Handlungsrestriktionen. Dem schließt sich (2) die auf der Mikroebene angesiedelte Handlungs- oder Spieltheorie an, die wiederum *erklärt*, welche Handlungsalternative unter diesen (bzw. den kognitiv wahrgenommenen) Kontextbedingungen von Akteuren mit (begrenztem) Streben nach Nutzenmaximierung am ehesten gewählt wird. Komplettiert wird die Erklärung dann (3) durch den letzten Schritt, welcher die Mikroebene der Erklärung wieder mit der Makroebene der Beobachtung verbindet. Erst die Aggregation der einzelnen individuellen Handlungen (durch einfaches Aufsummieren oder komplexe Transformationsregeln, s. dazu Coleman 1990; Esser 1999) plausibilisiert, wie es zu den eigentlich interessierenden sozialen Phänomenen (hier: soziale Ungleichheiten) kommt.

Abb. 1.1: Erklärungsschritte, Ebenen der Beobachtung und Erklärung



Quelle: Eigene Darstellung in Anlehnung an Coleman (1990: Kap. 1) sowie Manzo (2007).

Das Verständnis sozialer Ungleichheit wird also nicht ausschließlich durch die Handlungstheorie gewonnen, sondern erst durch die simultane Berücksichtigung von Kontextbedingungen, welche beispielsweise Karriereopportunitäten vorkanalisieren (Hedström/Bearman 2009a: 9). Mechanismen implizieren damit die Wahl *sozial* vorstrukturierter Handlungsalternativen. Sie sind nicht auf die

[7] Solche Erklärungsversuche werden im Gegenteil von einschlägigen Vertretern für ihre Nähe zu tautologischen Scheinerklärungen verurteilt (s. z. B. Diekmann/Voss 2004: 16). Ein häufiger Irrtum liegt zudem darin, rationale mit egoistischen Handlungen gleichzusetzen. Handlungstheorien decken prinzipiell ebenso altruistische Handlungen ab (Braun 1998, 2008).

Mikroebene beschränkt, postulieren aber – wenn überhaupt – Kausalität nur auf dieser Ebene handelnder und miteinander interagierender Individuen (Goldthorpe 2001).

Aufgrund der Verkettung einzelner Handlungen und der strategischen Interdependenz der Akteure schließen rationale Handlungen dabei keineswegs ungewollte Handlungsfolgen aus. Im Gegenteil besteht der Gewinn analytischer Soziologie gerade in dem Vermögen, überraschende und neue Hypothesen abzuleiten. Prominente Beispiele sind Schellings Segregationsmodell, das mit äußerst sparsamen Annahmen erklärt, wie es trotz Wunsch nach lediglich geringfügigen Separationen zur vollständigen ethnischen Wohnsegregation kommen kann (Schelling 1971); oder Wipplers Erklärung, warum selbst Organisationen mit ausdrücklich demokratischen Entscheidungsprinzipien zur Oligarchie tendieren (Wippler 1985). Mit der Mehrstufigkeit ihres Erklärungsprogramms verbleiben Rational-Choice-Ansätze nicht auf der Oberfläche rein makrosoziologischer Erklärungen und entgehen sie zugleich dem ihnen – oftmals zu Unrecht – gemachten Vorwurf des Reduktionismus oder gar dem Verdacht der „Banalität" (Haller 2003: Kap. 4; Kunz 2004: Kap. 5 für gängige Kritikpunkte).

Drittens sind Präzision und ein geringeres Aggregationsniveau als das des zu Erklärenden nicht damit zu verwechseln, möglichst *alle* für ein bestimmtes Phänomen relevanten Faktoren äußerst detailgetreu abzubilden, sondern ist im Gegenteil die *Abstraktion von weniger zentralen Elementen* (was zwangsläufig eine Simplifizierung der Realität bedeutet) unabdingbare Voraussetzung dafür, die *zentralen Mechanismen kenntlich zu machen*. Die Mechanismen sollten wenigstens so allgemeingültig sein, dass sie sich nicht nur auf ein einziges (historisches) Phänomen beziehen. Erklärungspotenzial haben sie insbesondere dann, wenn sie strukturelle Ähnlichkeiten zwischen anscheinend disparaten Prozessen aufzeigen (etwa Verteilungen von Forschungsdrittmitteln und Weiterbildungen ebenso erklären können wie die von Arbeitslosigkeitsrisiken; ein Beispiel hierfür stellt der in Kapitel 2 betrachtete *Matthäus*-Effekt dar). Gute Theorien erklären mit sparsamen Annahmen eine möglichst breite Palette an komplexen Phänomenen (Braun 1998: 155, 2008).

Schließlich kann *viertens* erst dann von einer Theorie oder Erklärung gesprochen werden, wenn diese *mindestens eine empirisch prüfbare Hypothese* impliziert. Allein mit Theorien ist es möglich, empirisch brauchbare Erklärungen und Prognosen zu entwickeln; umgekehrt gelingt es nur mit Empirie, die Brauchbarkeit von Theorien zu evaluieren. In Erfahrungswissenschaften wie der Soziologie sind empirische Prüfungen zur Entscheidung über die Beibehaltung, Veränderung oder Aufgabe von Theorien unerlässlich (Braun 2009: 225).

Dieses Programm einer erklärenden oder analytischen Soziologie grenzt sich damit insbesondere von zwei Polen ab: zum einen von einer reinen *Variablensoziologie* (Esser 1996), die ohne Anbindung an Theorien „freihändig" nach Zusammenhängen sucht; zum anderen von *all*umfassenden Gesellschaftstheorien, welche alle möglichen Aspekte zu erklären beanspruchen, damit aber zu oberflächlich bleiben, um empirisch prüfbar zu sein oder brauchbare Prognosen zu generieren: „A focus on explanatory mechanisms helps sociology to avoid the trap of mindless empiricism on the one hand, and conventional and empty theorizing on the other" (Hedström/Swedberg 1996: 299). Anliegen

ist stattdessen die Spezifikation von mehreren Mechanismen *mittlerer Reichweite* (im Sinne von Merton 1967), die sich für die Analyse ähnlich gelagerter Problemfelder bewähren, ohne aber der Anmaßung zu verfallen, *alle* soziologischen Gegenstände gleichermaßen zu erklären (Schmid 2009). Im Idealfall vermitteln diese Mechanismen auch Wissen über die Komponenten anderer Erklärungen und lassen sich somit im Forschungsverlauf zunehmend miteinander verzahnen (Hedström 2005: 26). Zugleich setzt sich das Programm einer analytischen Soziologie von deduktiv-nomologischen Erklärungen (Hempel/Oppenheim 1948) ab, weil diese aus Sicht der analytischen Soziologie zu anspruchsvoll sind (Hedström 2005: 15ff.; Schneider 2009: 356). Schließlich waren universelle Gesetze in den Sozialwissenschaften bislang kaum auffindbar. Mit kausalen Mechanismen sind daher in der analytischen Soziologie in aller Regel *probabilistische und nicht deterministische Zusammenhänge* gemeint, also Aussagen über in bestimmten sozialen Kontexten wahrscheinliche, aber nicht zwangsläufig eintretende Handlungen.

Innerhalb dieses Rahmens gibt es, wie bereits angedeutet, unterschiedliche Varianten. So herrscht etwa geteilte Meinung darüber, ob Mechanismen idealerweise direkt (durch Abfrage des Nutzens von Handlungsalternativen) oder indirekt (durch Prüfung von Modellimplikationen) untersucht werden (s. dazu auch den nachfolgenden Abschnitt 1.2.2). Ähnlich wird nach wie vor über das adäquate Vorgehen zur Bildung von Brückenhypothesen debattiert (Kelle/Lüdemann 1995; Lindenberg 1996; Opp 2009; Opp/Friedrichs 1996; Stachura 2009). Weiter ist umstritten, inwieweit die Gewinnung von „eleganten“ (sparsamen) Vorhersagen einen Verzicht auf den Realitätsgehalt der Modelle rechtfertigt (N. Gross 2009). Hierbei handelt es sich jedoch nur um Spielarten, die vermutlich dann einen besonders hohen Erklärungsgehalt haben, wenn sie nicht als sich ausschließende, sondern als sich ergänzende Modi der Erkenntnisgewinnung verstanden werden.[8] Keine Kompromissbereitschaft besteht jedoch bezüglich des Anspruchs, *präzise, handlungsbasierte, abstrakte und empirisch prüfbare Erklärungen* (Hedström 2005: 9) zu entwickeln.

1.2.2 Anforderungen an empirische Untersuchungen

Eine explizite Prämisse für Erklärungen ist, dass mindestens eine ihrer Annahmen empirisch prüfbar ist. Zugleich grenzt sich die analytische Soziologie von theorieloser Empirie ab. Im Kern geht es darum, Theorie und Empirie möglichst eng zu verzahnen, um die als kausal attribuierten Mechanismen fortwährend zu justieren. Dies impliziert einige weitere Anforderungen an ideale Forschungsprogramme. Kausale Schlüsse – von welchen Vertreter der analytischen Soziologie nun mal nicht absehen können (Brückner 2009: 679) – setzen zunächst folgende drei Prinzipien voraus:

Erstens ist eine möglichst exakte Prüfung des postulierten Mechanismus unerlässlich. Dieses Kriterium impliziert, dass präzise Hypothesen zu den Wirkungszusammenhängen formuliert werden.

[8] Zudem sollten sich unter diesen Verfahrensweisen die besonders geeigneten Kandidaten langfristig im empirischen Wettbewerb herauskristallisieren (Braun 2009: 229).

Deren empirische Testung verlangt dann nicht unbedingt, die postulierten Mechanismen direkt zu beobachten. Oftmals ist dies gar nicht möglich, etwa wenn es sich um nichtintendierte Effekte oder unbewusste Motive handelt. In solchen Fällen kann sich die Testung auf generelle Modellimplikationen beziehen (etwa Prüfung, ob die Kontextbedingungen wie angenommen mit Handlungswahlen kovariieren, ohne die Handlungsmotive direkt zu erfragen). Besonders stichhaltig sind Evaluationen, die sich nicht nur auf Variablen*zusammenhänge*, sondern auf zugrunde liegende *Prozesse* beziehen, indem sie etwa Thesen zu Modellparametern prüfen, welche wiederum bestimmte Verläufe von Diffusionsprozessen implizieren (Goldthorpe 2001; Manzo 2007).

Zweitens sind möglichst viele alternative Mechanismen auszuschließen. Diese Anforderung profitiert zunächst von einer möglichst direkten Operationalisierung der relevanten Konzepte. Zudem sollten die Analysen auf der richtigen Ebene ansetzen. Aufgrund der Bezugnahme auf (interdependente) Entscheidungen von Akteuren sind dies Individualdaten. Zusammenhangsanalysen, die ausschließlich auf der Makroebene verbleiben, bergen ein zu großes Risiko für ökologische Fehlschlüsse im Sinne von Robinson (1950).[9]

Drittens implizieren Ursachen einen zeitlichen Vorsprung vor ihrer Wirkung, daher werden im Idealfall *Beobachtungen mit zeitlicher Reihenfolge* (Längsschnittdaten oder experimentelle Daten) herangezogen.[10] Die richtige Reihenfolge ist allerdings nur eine *notwendige*, nicht aber *hinreichende* Bedingung für kausale Beziehungen. Stichhaltige Prüfungen setzen ebenso den Gegenbeweis, das Ausbleiben der Folge bei Fehlen ihrer postulierten Ursache, voraus.[11] Querschnittdaten bieten speziell dann wenig Aufschluss, wenn gegenläufige Mechanismen im Spiel sind oder es sich um dynamische Prozesse (z. B. sich selbst erfüllende Prophezeiungen) handelt. Ähnlich entgehen viele Erklärungen durch Präferenzen erst mittels der zeitlichen Entschlüsselung von Ursache und Wirkung dem Verdacht einer Tautologie oder *post-hoc*-Rationalisierung (Braun/Franzen 1995: 233; Kunz 2004: 146ff.).[12]

---

[9] Ein ökologischer Fehlschluss liegt vor, wenn von einem Zusammenhang auf der Makroebene fälschlicherweise auf einen entsprechenden Zusammenhang auf der Mikroebene geschlossen wird (Diekmann 2007). Beispielsweise bedeutet eine Korrelation von Männeranteilen und Weiterbildungsquoten in Betrieben noch nicht zwangsläufig, dass Männer stärker an Weiterbildungsmaßnahmen partizipieren als Frauen.

[10] Feinsinnige Kritiker vermögen hier einzuwenden, dass menschliche Ereignisse oftmals ihre Schatten voraus werfen: Weihnachtseinkäufe werden im vorweihnachtlichen Advent getätigt. Dies ist aber noch kein Widerspruch zur postulierten zeitlichen Logik, liegt doch die Ursache hier aus handlungstheoretischer Sicht gerade nicht im abstrakten „Weihnachten", sondern in der Antizipation dieses Ereignisses und der dadurch definierten Verhaltenserwartungen an die Akteure (also in der Logik der Situation).

[11] Man denke wiederum an die Weihnachtseinkäufe. Ähnlich plastisch lässt sich der unzureichende Beweis von Kausalität durch das fortwährende Aufeinandertreffen von Phänomenen anhand der Tatsache veranschaulichen, dass es vermutlich nicht die ansteigende Nachfrage nach Badehosen ist, die kausal den Sommereintritt herbeiführt. Beide Beispiele entstammen Lehrveranstaltungen von Norman Braun an der Ludwig-Maximilians-Universität (LMU) München.

[12] Versierte Verfahren der Längsschnittanalyse bieten zudem verbesserte Möglichkeiten, (unbekannte) Drittvariablen als Erklärungen auszuschließen. Das Prinzip dieser *First-Difference*-Modelle besteht vereinfacht darin, die Untersuchungsobjekte als ihre jeweils eigenen Kontrollgruppen aufzufassen. Wirkungen zeitlich variierender Effekte lassen sich dadurch von konstanten (unbekannten) Drittvariableneinflüssen der einzelnen Untersuchungseinheiten trennen (s. z. B. Wooldridge 2002 für statistische Details). Vergleichbares Potenzial haben in

Das Leitbild möglichst sparsamer, zugleich weitreichender Erklärungen führt *viertens* zu der Prämisse, *generelle Mechanismen zu modellieren* anstatt sich auf eine möglichst detailgetreue Abbildung der Realität zu konzentrieren. Zu exakte Modellierungen bergen die Gefahr des *Overfittings*, also der Abbildung von Besonderheiten der konkreten Stichprobe, die keine Entsprechung in der allgemeinen Population haben und somit keinen *generellen* Mechanismus darstellen. Praktisch bedeutet dies, dass der Forscher seine Empirie an der Theorie (Test des unterstellten Mechanismus) und nicht an einer Maximierung der Varianzaufklärung ausrichten sollte. Ausschließlich statistisch motivierte Analysen sind für die Feststellung von Kausalität unzureichend, denn es bedarf hierzu immer einer Theorie, d. h. *Vorab*-Spezifikation vermuteter ursächlicher Prozesse (Goldthorpe 2001).

Die Einlösung der genannten Prämissen ist in der Praxis kaum vollständig möglich, dennoch dienen sie der Orientierung. So sind in den Sozialwissenschaften Längsschnittdaten immer noch selten und liegen speziell bei Sekundärauswertungen viele gewünschte Informationen nicht vor. Ergebnisse sollten daher stets durch Replikationen abgesichert und zudem die Validität methodischer Verfahren durch systematische Forschung ausgelotet werden. Die von Rational-Choice-Theoretikern mit ihrem heuristischen Wert gerechtfertigten Abstraktionen von empirischen Regularien (wie unvollständigen Informationen und beschränkten Bearbeitungskapazitäten) sind nicht mehr tragbar, wenn es darum geht, interessierende Verhaltensweisen möglichst valide zu erheben (Rashotte et al. 2005: 165). Nur wenn sichergestellt ist, dass Surveyfragen oder experimentelle Stimuli standardisiert und wie vom Forscher intendiert verstanden werden, lassen sich die mit ihnen gewonnenen Daten verlässlich deuten. Fester Bestandteil analytischer Soziologie sollte daher Methodenforschung sein, dieser substanzielle Aspekt wird bislang weitgehend übersehen (Brückner 2009 für eine Ausnahme).

Im konkreten Einzelfall bestimmt sich die ideale Auswahl an Methoden durch die zu prüfenden Hypothesen. Gleichwohl lassen sich einige allgemeine Empfehlungen aussprechen. Generell weisen Experimente einen besonders starken Zuschnitt auf die genannten Prämissen auf (dazu Diekmann 2007, 2008). Die als ursächlich angenommenen Faktoren werden durch den Forscher manipuliert, gehen damit zeitlich den Reaktionen der Probanden voraus, und mögliche alternative Mechanismen bzw. Drittvariablen werden durch die *Randomisierung* (Zufallszuweisung der Stimuli auf die Probanden) neutralisiert. In der analytischen Soziologie werden Experimente daher zunehmend angeraten (z. B. Hedström 2005: 151). Allerdings weisen gerade Laborexperimente ebenso Nachteile auf: So wird ihre Validität durch die hohe Künstlichkeit und die zumeist sehr selektive Auswahl an Teilnehmenden (fast immer Studierende) beschnitten (Brückner 2009: 679; Brüderl 2004: 165; de Vaus 2006). Feldexperimente scheiden dagegen für viele soziologische Fragestellungen schon wegen praktischer oder ethischer Gründe aus (beispielsweise können Forschungsmittel nicht allein nach experimentellen Gesichtspunkten verteilt werden). Oftmals ist daher die Prüfung von kausalen Zusammenhängen in einem so genannten *ex-post-facto*-Design (üblicherweise Auswertung von Surveydaten

---

Hinblick auf die Trennung der Effekte von Gruppen- versus Individualmerkmalen Verfahren der Mehrebenenanalyse.

mit Regressionsanalysen) die einzige oder wenigstens eine sinnvolle zusätzliche Variante. Von vielen Vertretern analytischer Soziologie wird diese Verfahrensweise zu pauschal mit einer Variablensoziologie gleichgesetzt. Nicht anerkannt wird dabei, dass inzwischen äußerst elaborierte statistische Verfahren (wie z. B. Mehrebenen- und Panelregressionen, Matchingverfahren) vorliegen, mit denen sich sehr exakte Prüfungen von Theorien umsetzen lassen. Statt allzu viel Energie auf die pauschale Kritik zu verwenden, sollten sich die Anstrengungen besser darauf konzentrieren, die Surveyforschung an den genannten Prämissen auszurichten (ähnlich Blossfeld 1996: 151; Brückner 2009; Ultee 1996).

Zudem existieren in der Soziologie bislang kaum wahrgenommene Möglichkeiten, die Vorteile von Experimenten und Umfragen in Form von *Conjoint-Analysen*, *Faktoriellen Surveys* oder *Choice-Experimenten* zu kombinieren (Klein 2002 bzw. M. Beck/Opp 2001; Jasso 2006 bzw. Auspurg/ Liebe 2010 für Einführungen). Die gemeinsame Idee dieser Verfahren besteht darin, mittels systematisch variierter Fallbeispiele (*Profilkarten* bzw. *Vignetten* und *Choice-Sets*) den exakten Einfluss von Merkmalen (*Attributen* bzw. *Dimensionen*) auf interessierende Einstellungen, Normen oder Entscheidungen zu bestimmen, die hierfür in der Realität oftmals zu stark miteinander korreliert sind.[13] Derartige Kombinationen aus Umfragen und Experimenten werden bisher fast ausschließlich in der Einstellungsforschung oder außerhalb der Sozialwissenschaften genutzt (Auspurg/Liebe 2010; Wallander 2009 für Überblicke). Wie noch näher erläutert und praktisch demonstriert (Abschnitt 1.3.2 bzw. Kapitel 2), eröffnen diese Herangehensweisen aber auch der analytischen (Ungleichheits-)Soziologie erhebliche Potenziale. Der Hauptgrund ist, dass diese quasi-experimentellen Methoden einen gegenüber anderen Verfahren oftmals weitaus trennschärferen Aufschluss über die Erklärungsleistung von verschiedenen alternativen Mechanismen bieten.

1.2.3 Stand der Ungleichheitssoziologie

Warum bedarf es angesichts der vielfältigen Literatur zur erklärenden und ebenso Ungleichheitssoziologie noch dieser Ausführungen? Ein erster Grund ist, dass der Erkenntnisstand der Ungleichheitssoziologie nach verbreiteter Auffassung als unbefriedigend gilt (z. B. J. Berger 2003, 2004; Burzan 2007; Müller/Schmid 2003; Schwinn 2007; Reskin 2003). Ein zweiter, noch gewichtigerer Grund besteht darin, dass bislang – trotz mehrfacher Versuche der Neuausrichtung – kaum erkannt wird, dass dies mit einer mangelnden Orientierung an den Prinzipien erklärender Soziologie zu tun haben könnte. Vorliegende Überblickswerke kommen zwar oftmals gleichermaßen zu der Diagnose, dass die Ungleichheitssoziologie unergiebig zwischen den beiden Polen zu hoher Abstraktion (Suche nach allgemeinen Gesetzen) oder sich im Detail verlierender Beschreibungen (somit geringer Erklärungstiefe) schwankt, bieten dann aber oftmals keinerlei Therapie an, und schon gar nicht in Form der

[13] Ein Beispiel für solche Merkmale stellen das Geschlecht und die beruflichen Tätigkeiten von Akteuren dar, die etwa hinsichtlich ihres Einflusses auf familiale Karriereentscheidungen schwer zu separieren sind – auf dieses Beispiel wird in Abschnitt 1.3.2 noch näher eingegangen.

hier vertretenen These, dass mechanistische Erklärungen mittlerer Reichweite einen Ausweg bieten könnten.

Dies soll mit einem skizzenhaften Überblick über zentrale Strömungen der Ungleichheitssoziologie und Sozialstrukturanalyse belegt werden (vgl. J. Berger 2004; P. Berger 1987; P. Berger/Hradil 1990; Burzan 2007; Geißler 1996; Grusky 2001; Solga et al. 2009 zum Folgenden). Bereits die Klassen- und Schichtkonzepte der ersten Generation machen beide Extreme deutlich. Die Klassentheorie von Karl Marx und Friedrich Engels repräsentiert etwa einen klassischen Ansatz mit allgemeinem Erklärungsanspruch für gesellschaftliche Entwicklungen. Sie wird den Prämissen erklärender Soziologie zwar durch die Benennung eines expliziten Mechanismus für soziale Ungleichheiten (gesellschaftliche Organisation der Produktions- und Besitzverhältnisse) und Erklärung ihres Wandels (wachsender Antagonismus bzw. Entstehung von Klassenbewusstsein) gerecht, ist als generelle Gesellschaftstheorie aber deutlich zu abstrakt angelegt, um empirisch präzise (sowie historisch bewährte) Vorhersagen zu generieren. Ähnliches lässt sich gegen die gleichfalls zur ersten Generation universeller Gesellschaftstheorien zählenden funktionalistischen Ansätze von Talcott Parson sowie Kingsley Davis und Wilbert E. Moore vorbringen (s. Collins 1971; Mayntz 1961; Tumin 1953 für klassische Kritiken). Funktionalistische Theorien stehen zwar mit ihrer Frage nach den Funktionen sozialer Phänomene einer Ursachenforschung nahe und lassen sich auch formal als Spezialfall kausaler Erklärungen fassen (Hempel/Oppenheim 1948), jedoch verdrehen sie die zeitliche Logik: Die Folgen des Schichtgefüges sollen hier seine Ursache „erklären". Zudem ist nochmals darauf hinzuweisen, dass eine fortwährende Koinzidenz von Phänomenen (wie der sozialen Schichtung und dem Überleben von Gesellschaften) für Kausalität nicht hinreichend ist (vgl. Abschnitt 1.2.2), denn es fehlt der – bei diesem Gegenstand *per se* nicht zu erbringende – Gegenbeweis.[14] Hierdurch, und weil die Vorhersagen dafür zu abstrakt sind, werden funktionalistische Theorien dem Kriterium der empirischen Prüfbarkeit wenig gerecht. Hinzu kommt, dass funktionale – wie alle teleologischen (von einem übergeordneten Ziel her argumentierenden) oder sich auf einen ungesteuerten Selektionsmechanismus (Evolution) beziehenden Ansätze – nicht in der Lage sind, alternative Erklärungen bzw. *funktionale Äquivalente* auszuschließen. Speziell die funktionalistische Schichtungstheorie von Davis und Moore (1945) birgt zudem ein weiteres Problem: Ihre Kernidee ist bekanntlich, dass Gesellschaften nur überleben, wenn ungleiche Entlohnungen die Besetzung funktional wichtiger Positionen mit sehr leistungsfähigen Akteuren sicherstellen. Dies impliziert aber die Prämisse, dass soziale Ungleichheiten im Wesentlichen Unterschiede im Talent und in der Leistung reflektieren. Damit hat das Konzept *per se* nur soziale Ungleichheit im weiten Sinne im Blick, während die gesellschafts-

[14] Von Hartmut Esser (1999: 362) wird das Problem folgendermaßen auf den Punkt gebracht: Wie will man empirisch den „Tod" einer Gesellschaft feststellen? Daraus, dass soziale Schichtungen universeller Bestandteil aller existierenden Gesellschaften sind, kann noch nicht eindeutig auf ihre *kausale* Relevanz für deren Bestehen geschlossen werden (s. dazu Abschnitt 1.2.2 sowie das noch erläuterte Argument funktionaler Äquivalente).

politisch besonders relevanten Ungleichheiten im engen Sinne (Verstöße gegen das Prinzip der Meritokratie) als Explanandum ausgeklammert werden.[15]

Die in Reaktion auf das zu hohe Abstraktionsniveau der ersten Klassentheorien entwickelten, verfeinerten Typologien von Klassen, Ständen und Schichten, für welche Max Weber als Wegbereiter gilt, erkaufen ihre Präzision dann durch den weitgehenden Verzicht auf eine *Erklärung*, wie diese Sozialstrukturmuster zustande kommen.[16] Gerade dies ist das Dilemma auch neuerer (neomarxistischer) Klassentheorien: Entweder halten sie am kausalen Mechanismus eines Antagonismus von Großgruppen fest und sind damit zu abstrakt, um die inzwischen stärker ausdifferenzierten Erwerbskonstellationen einzufangen und für die Akteure selbst bedeutsame Gruppen darzustellen; oder sie greifen stärker auf den Mechanismus subjektiv geteilter Bewusstseinslagen zurück, sind dann aber darauf angewiesen, ein hierfür hinreichend feines Raster zu finden. Die letzere Aufgabe bedingt in den stark heterogenen Arbeitsmärkten westlicher Gesellschaften eine zunehmend „mühsame" Anbindung an die für Klassenkonzepte essenziellen sozioökonomischen Lagen.[17] Praktisch wird die Bestimmung dieser Klassen mitunter deshalb vorwiegend induktiv (durch empirische Ableitung) gelöst. Damit kommt es jedoch unweigerlich zu einer Annäherung an die lediglich beschreibenden Schichtkonzepte. So ist bereits der Versuch von Erik Olin Wright, „Mittelklassen" einzuführen, durch einen Verlust an theoretischer Stringenz gekennzeichnet, wenngleich hier mit Bezug auf das Ressourcenkonzept von John Roemer (z. B. Roemer 1985) die Ursachenanalyse noch eine zentrale Rolle spielt (ähnlich Hradil 2001: 67). Deutlicher wird das Problem bei noch stärker ausdifferenzierten Ansätzen wie dem von Kim Weeden und David Grusky (2005): Mit nicht weniger als 126 Mikroklassen für die derzeitige US-amerikanische Gesellschaft wird in diesem Konzept die für Erklärungen notwendige Abstraktion sicher der Realitätstreue geopfert (Goldthorpe 2000: Kap. 6).[18] Im Gegensatz zu den eher beschreibend angelegten Schichtkonzepten fokussieren Klassenkonzepte jedoch in aller Regel noch explizit auf die *Erklärung ungleicher Lebenschancen* und stehen zudem mit ihren Argumentationen über Mehrwertaneignungen, sozialen Schließungen und gemeinsamen Interessenslagen handlungstheoretisch basierten

[15] Verstöße gegen das Prinzip der Leistungsgerechtigkeit sind zwar nicht gänzlich unvereinbar mit funktionalistischen Theorien, sie bilden aber in diesen Ansätzen maximal am Rande erwähnte, nicht näher ursächlich interessierende Dysfunktionalitäten, welche in funktionierenden Gesellschaften ein bestimmtes Maß nicht überschreiten dürfen (ähnlich Grusky 2001: 14).

[16] Am ehesten werden noch im Konzept von Max Weber selbst ursächliche Mechanismen für die Entstehung sozialer Ungleichheiten angeführt (Arbeitsteilung und Chancen der Martkverwertung von Gütern und Leistungen für die Bildung von Besitz- und Erwerbsklassen; soziale Schließung und Monopolisierungen für die Bildung von sozialen Klassen). Allerdings zielt auch sein Ansatz mehr auf die Entwicklung eines konzeptionellen Rahmens statt auf die Ausarbeitung einer *Theorie* für die *Ursachen* verschiedener Formen sozialer Ungleichheit.

[17] Etwa bewirkt die „mehrdeutige Position der Angestellten und Beamtenschaft als ‚Zwischen-', ‚Mittelklasse' oder ‚Nicht-Klasse'" (P. Berger 1987: 69) eine vergleichsweise hohe berufliche Mobilität, welche wiederum heterogene Lebensstile und Wertorientierungen und somit geringe Chancen der Herausbildung eines Klassenbewusstseins impliziert.

[18] Der Ansatz konserviert durch die Gleichsetzung von Klassen mit Berufsgruppen vor allem die Idee gemeinsamer Interessenslagen aus älteren Klassenkonzepten. Da sich Sozialisations- und Selektionsprozesse sowie Prägungen durch gemeinsame Arbeitsbedingungen und Kontaktwahrscheinlichkeiten vor allem entlang von Berufsgruppen strukturieren würden, werden diese mit Klassen gleichgesetzt.

Mechanismen nahe. Speziell neuere Konzepte wie die von Aage Sørensen (2000) oder John Goldthorpe (2000) weisen eine große Schnittmenge zur analytischen Soziologie auf. Auf diesen Aspekt und auf die dennoch bestehenden Unterschiede wird noch näher eingegangen (Abschnitt 1.2.4). Vorweggenommen werden soll hier aber schon einmal die These, dass diese Ansätze den Klassenbegriff umso stärker obsolet machen, je besser sich die mit ihm verbundenen Erklärungsmechanismen bewähren. Eine unvermeidliche Schwäche von Klassenkonzepten ist jedenfalls, dass sie mit der Fokussierung auf Erwerbstätige *per se* einen substanziellen Teil der Bevölkerung (Arbeitslose, Rentner, Hausfrauen) mit den für diese Personen relevanten Ungleichheiten ausblenden.

In der Ungleichheitssoziologie führte die Kritik an der zu starken Erwerbszentrierung dieser Ansätze mitunter zur Suche nach neuen Lösungen. Die in den 1980er Jahren versuchten Neukonzeptionen in Form von „Sozialen Lagen", „Lebensstilen" oder „Milieus" erweitern die Perspektive um horizontale (nicht einer klaren Hierarchie- oder Prestigeordnung unterliegende Ungleichheiten), oftmals einschließlich der realen lebensweltlichen Erfahrungen und kulturellen Präferenzen der Akteure (s. etwa die programmatischen Beiträge in dem von Reinhard Kreckel im Jahr 1983 herausgegebenen Sonderband „Soziale Ungleichheiten" der Zeitschrift „Soziale Welt"). Das vorrangige Ziel wird nun in einer möglichst detailgetreuen Beschreibung der Sozialstruktur gesehen. *Systematische* Zusammenhänge zwischen Beobachtungen auf der Makroebene und den sie begründenden Handlungen auf der Mikroebene werden dabei vielfach gar nicht mehr spezifiziert (was dem Verzicht auf ursächliche Mechanismen gleichkommt), und mit dem Bemühen, alle möglichen Verschiedenheiten zu erfassen, geht die Fokussierung auf *soziale Ungleichheiten* im engeren Sinne verloren (ungleiche Zugangschancen zu begehrten Ressourcen, vgl. Abschnitt 1.1). Pointierter kommt es in den Worten Rainer Geißlers zum Ausdruck: „Ungleichheitsforschung ufert zur Vielfaltsforschung aus" (1996: 322). Viele der in dieser Zeit entwickelten Ansätze sind geradezu Paradebeispiele für die theoriearme Variablensoziologie (Esser 1996). So wird etwa über die Erklärungsleistung konkurrierender Ansätze vorrangig anhand des Kriteriums der Varianzaufklärung entschieden. Dass die induktiv anhand von Faktoren-, Cluster- und Korrespondenzanalysen gewonnenen Typologien in der Tat eine hohe Beliebigkeit aufweisen, demonstriert ihre mangelnde Reproduzierbarkeit. So zählt etwa Otte (2005) allein für Deutschland 17 unterschiedliche Lebensstiltypologien. Anstatt in der Konsequenz die Notwendigkeit einer stärkeren Abstraktion anzuerkennen, besteht die Reaktion auf die konstatierte „Unübersichtlichkeit" (Kreckel 1983: 11) oftmals in einer vollständigen Aufgabe des soziologischen Erklärungsanspruchs.[19]

Vollständig von der Agenda fällt die Erklärung sozialer Ungleichheit dann mit der Individualisierungsthese, welche im Kern die Auflösung von Klassen, Schichten und allgemein allen Großgruppen behauptet, die das Handeln und die subjektive Lebensweise strukturieren (U. Beck 1983, 1986;

[19] Als vertretbar werden nur noch kontingente historische Aussagen gesehen. Beispielhaft sei hier folgendes Zitat von Martin Bolte angeführt: „Die in einer Gesellschaft wirkenden ungleichheitsrelevanten Mechanismen können zwar einzeln aufgedeckt und benannt werden, ihre Bedeutung im Verhältnis zueinander kann aber nur am konkreten empirischen bzw. historischen Fall bestimmt werden" (1983: 401).

kritisch: Geißler 1996). Hier wird wiederum eine universelle Gesellschaftstheorie vorgeschlagen, die das Problem aller „großen Würfe“ teilt: Zentrale Bestandteile sind zu unpräzise, um sie empirisch „auf die Probe“ stellen zu können (Burzan 2007: 162). Konsequenterweise legt das Modell auch gar kein Konzept mehr vor, um Ungleichheitsgefüge empirisch zu fassen oder gar zu erklären. Diesem Anspruch werden in dieser Zeit noch am ehesten die Arbeiten zur sozialen Reproduktion von Pierre Bourdieu gerecht, allerdings mangelt es ihnen aus analytischer Sicht ebenfalls an Exaktheit (das Habituskonzept von Bourdieu gilt in dieser Hinsicht geradezu als Negativbeispiel, vgl. Hedström 2005: 4). Angezweifelt wird zudem, ob die in Anlehnung an Klassentheorien entwickelten Theorieelemente, wie speziell das des „kulturellen Kapitals“, überhaupt einen eigenständigen Mechanismus implizieren. Die betrachteten Phänomene (wie beispielsweise die Wahl ungleicher Bildungswege) lassen sich womöglich eleganter durch eine Beschränkung auf herkömmliche Handlungstheorien erklären (Goldthorpe 2007).[20] Freier formuliert zielt diese Kritik in die gleiche Richtung wie die im nachfolgenden Zwischenfazit (Abschnitt 1.2.4) ausgeführte These, dass das Festhalten an Klassenkonzepten im Hinblick auf stringente Erklärungen oftmals eher einen Umweg statt einen Gewinn darstellt. Gerade an Bourdieu anknüpfende Arbeiten verkennen zudem oftmals, dass die Konvertierbarkeit von Kapitalsorten allein noch keine Erklärung darstellt (Schwinn 2008: 36). Ein weiterer kritischer Punkt ist die starke Zirkularität – Klassen werden reproduziert, und sozialer Wandel ist lediglich durch externe Schocks oder Revolutionen möglich. Anzuerkennen ist aber, dass die Abhandlungen Bourdieus einen vergleichsweise starken analytischen Fokus aufweisen und inzwischen auch eine ganze Reihe von Arbeiten angestoßen haben, die theoriegeleitete Hypothesen über Ungleichheit generierende Mechanismen mit sehr ausgefeilten empirischen Methoden prüfen (s. exemplarisch den von Werner Georg im Jahr 2006 herausgegebenen Sammelband „Soziale Ungleichheit im Bildungssystem“). Diese Arbeiten sind wegweisend, würden aber aus der Perspektive einer analytischen Soziologie noch von einer etwas präziseren und stringenteren Herleitung der unterstellten Mechanismen profitieren.[21]

In der aktuellen Literatur ist das Fortbestehen gruppenspezifischer Ungleichheiten nun wieder weitestgehend anerkannt, zu ihrer Erfassung treten beide Strategien – Bemühungen um möglichst detailgetreue Abbildungen sowie große Theorieentwürfe – gleichermaßen auf. So wird in jüngster Zeit wieder nach adäquaten *Beschreibungs*dimensionen für die neuerlichen Ausdifferenzierungen und Transformationen gesucht. Zeugnisse hierfür geben Titel und Untertitel aktueller Monographien, die

[20] John Goldthorpe unterscheidet dabei zwischen zwei Auslegungen von Bourdieus Theorie sozialer Reproduktion: In einer engen Fassung habe sie zu wenig Neues gegenüber handlungstheoretischen Erklärungen zu bieten, wie sie schon in den 1960er Jahren vorgeschlagen wurden; in einer weiteren Fassung sei sie wenig plausibel und vor allem durch die Empirie widerlegt. Oder in Goldthorpes eigenen Worten: „the difficulty is that what is sound is not original and what is original is not sound.“(2007: 78).

[21] Die Erwiderung hierauf ist, dass die Komplexität des Gegenstandes keine simplere Sprache zulasse (s. z. B. die unter dem bezeichnenden Titel „So einfach ist es nicht“ im Jahr 2009 verfasste Kritik von Frank Hillebrandt an Hedströms Abhandlung „Anatomie des Sozialen“). Der Graben zur analytischen Soziologie verläuft hier somit vor allem entlang des Ausmaßes notwendiger Abstraktion und Akkuratheit kausaler Mechanismen.

in der deutschsprachigen Soziologie etwa „Achsen der Ungleichheit“ (Klinger 2007), „Neuvermessung der Ungleichheit unter den Menschen“ (U. Beck 2008), „Intersektionalität“ (Winker/Degele 2009), „Transnationalisierung sozialer Ungleichheiten“ (P. Berger 2008), oder „Transnationale Ungleichheitsforschung“ (Bayer 2008) lauten. Die Innovation wird nun in der Erfassung mehrdimensionaler, transnationaler und weltgesellschaftlicher Ungleichheiten gesehen; wiederum zollt die Horizonterweiterung aber der Erklärungstiefe ihren Tribut.[22] Dass man mit einem derart feinen Raster wie den bereits erwähnten über 100 Mikroklassen von Weeden und Grusky (2005) viel Varianz in *aktuell* vorliegenden Einstellungen und Verhaltensweisen „aufklären“ kann, erstaunt wenig – ein tieferes Verständnis der *generellen*, über punktuelle Gesellschaften und Zeiten noch gültigen Mechanismen ist damit aber noch nicht unbedingt gewonnen, und aufgrund des hier nahe liegende Problems des *Overfittings* wohl auch gar nicht möglich (ähnlich Rössel 2009: 126).[23] Andererseits gibt es derzeit wieder Versuche, durch die Integration großer Theorieströmungen (Theorien sozialer Ungleichheit und Theorien sozialer Differenzierung) doch noch zu *der* einheitlichen Gesellschaftstheorie zu kommen (z. B. Schimank 1998; Schwinn 2007). Wie diese empirisch umgesetzt werden soll oder sich konkrete Allokationen damit begründen lassen, dazu finden sich (aus Sicht einer analytischen Soziologie wenig überraschend) nur vage Andeutungen (s. z. B. Schimank 1998: 77f.). Dies gilt gleichfalls für sehr pointierte und lediglich eklektisch durch empirische Befunde abgesicherte Gegenwartsdiagnosen. Exemplarisch sei hier die Behauptung von Ulrich Beck (2008: 21) angeführt, dass „zum wichtigsten Einflussfaktor, der über die Position in der Ungleichheitshierarchie im globalen Zeitalter entscheidet, […] die Möglichkeiten oder Chancen zu grenzüberschreitender Interaktion und Mobilität“ geworden sind. *Wie* und *Warum* wird nicht näher erläutert; ebenso bleibt die zur Erfassung dieser neuen Ungleichheiten angemahnte „kosmopolitische Perspektive“ unklar (Solga et al. 2009: 37; Verwiebe 2010 für ähnliche Einwände). Erste empirische Prüfungen zeigen überdies, dass die Bedeutung transnationaler Klassen oder Eliten aktuell (noch?) *deutlich* überschätzt wird. Ungleichheits- und Karrieremuster sind nach wie vor zu einem *weit* überwiegenden Teil nationalstaatlich und nicht inter- oder transnational organisiert (Goldthorpe 2002; Hartmann 2009 für einschlägige Studien).

[22] Diese Stoßrichtung wird auch an dem für das Jahr 2010 geplanten „Jubiläumskongress“ der Deutschen Gesellschaft für Soziologie (DGS) ersichtlich. Dem Themenaufruf zufolge stellt sich aufgrund der zunehmenden Diversität zeitgenössischer Gesellschaften im 21. Jahrhundert die „Frage nach sozialen Ungleichheitsstrukturen völlig anders“ (DGS 2010: 3) – wobei dann konkret vor allem wieder *Was*- und *Welche*-, nicht aber *Wie*- und *Warum*-Fragen auf der Agenda stehen. Die im Themenaufruf genannten Problemstellungen lauten beispielsweise, ob die Ungleichheit nun mehr individualisiert sei als in früheren Zeiten; ob sich neue Gruppen oder Kollektive herausbilden; wie sich in einem transnationalen Kontext Aussagen über prekäre Lebenslagen treffen lassen und welche Distinktionsmöglichkeiten strukturbildend sind (DGS 2010).

[23] Zudem überrascht, für welch unterschiedliche Explananda die Klassenkategorien bei Weeden und Grusky (2005) eine „Aufklärung“ liefern sollen. Die Konzepte haben sich etwa auch an Lebenszufriedenheiten und Einstellungen gegenüber Abtreibungen zu bemessen, also an Phänomenen, die zumindest die traditionellen Klassenkonzepte keineswegs im Blick hatten. Sehr salopp aber treffend bringt Robert Erikson die problematische Erweiterung (oder Ausuferung) des Erklärungsgegenstandes auf den Punkt: „Der Klassenbegriff ist kein Mädchen für alles“ (zitiert nach Goldthorpe 2000: 280).

Insgesamt lösen damit weder die älteren noch die neueren Ansätze den Anspruch ein (oftmals verfolgen sie ihn erst gar nicht), ein Programm vorzulegen, das ungleiche Allokationen durch einen kausalen Mechanismus erklärt, dabei abstrakt genug ist, um sich nicht nur auf kontingente, historische Begebenheiten zu beziehen und zugleich hinreichend präzise, um sich empirisch prüfen zu lassen. Viele Arbeiten lesen sich geradezu als Kontrastprogramm zu einer analytischen Soziologie. Gleichwohl gibt es Musterbeispiele dafür, dass genau diese analytische Soziologie zur Erklärung sozialer Ungleichheiten beitragen kann. So liegen etliche empirische Studien im Bereich der Arbeitsmarkt-, Organisations- und Diskriminierungsforschung sowie Bildungs- und Familiensoziologie vor, die den genannten Prinzipien sehr gut gerecht werden (s. z. B. die Überblicke in dem von Peter Hedström und Peter Bearman im Jahr 2009 herausgegebenen Handbuch analytischer Soziologie sowie die Hinweise in Blossfeld/Müller 1996; Grusky 2001: 30f.; Hout/DiPrete 2004; Manzo 2007: 52ff.). Bei den geprüften Mechanismen handelt es sich um Signal- und Reputationseffekte, (statistische) Diskriminierungen, Homophilie- und Konkurrenzmechanismen oder auch um kumulative Verstärkungen anfänglich kleiner Ungleichheiten in Form eines Matthäus-Effekts (mehr zu diesem in Kapitel 2). Ebenso stehen – wie bereits angedeutet – einige der neueren Klassenansätze, wie etwa die Arbeiten von Aage Sørensen (2000) oder John H. Goldthorpe (2000), einer analytisch ausgerichteten Soziologie nahe.

Diese Anwendungen werden jedoch in einschlägigen Lexika, Lehr- und Handbüchern kaum rezipiert. Lehrbücher erwähnen oftmals gar keine empirischen Arbeiten, sondern erschöpfen sich stattdessen in einer Aneinanderreihung der genannten (überholten) Dogmen (vgl. Braun 2008; für eine Ausnahme Rössel 2009). Die fehlende Berücksichtigung analytischer Studien mag daran liegen, dass in diesen (bewusst) keine Theorien vorgeschlagen werden, unter die sich alle möglichen (transnationalen) sozialen Ungleichheiten subsumieren lassen. Nachdem ein dafür tauglicher Ansatz aber keineswegs in Sicht ist, sollte der bessere Weg dann nicht in der Weiterentwicklung von zwar bereichsspezifischen, dafür aber erklärenden Konzepten bestehen? Warum sollte es angesichts der sehr unterschiedlichen Bereiche und Felder, in denen Ressourcen verschiedenster Art verteilt werden, überhaupt gelingen, ihre Ungleichverteilung mittels lediglich einer einzigen Theorie zu erklären? Genaue Beschreibungen und Typologien sind sicher verdienstvoll und werden auch von der analytischen Soziologie gefordert. Die „Analysen“ sollten dann aber nicht an diesem Punkt stehen bleiben. Beide Extreme – Soziologie einerseits als Versuch, möglichst zeitlose und universelle Gesellschaftstheorien zu formulieren, anderseits als Praxis, mit möglichst detailgetreuen, dafür aber theorielosen Auswertungen ein Abbild der Realität zu erstellen – haben sich für die Erklärung sozialer Ungleichheiten bislang wenig bewährt.

1.2.4 Zwischenfazit: Ein Abschied von Klassen und Schichten?

In seiner Eröffnungsrede zum 34. Kongress der Deutschen Gesellschaft für Soziologie (DGS), stattfindend im Jahr 2008 unter dem Motto „Unsichere Zeiten", fordert Ulrich Beck einen „neuen soziologischen Blick, eine neue soziologische Grammatik für die beschreibende Theorie der Phänomene sozialer Ungleichheit" (2008: 52). Der Aufruf erinnert stark an die in den 1980er Jahren vorherrschenden Plädoyers für eine Neukonturierung der Ungleichheitssoziologie (denen etwa der von Reinhard Kreckel im Jahr 1983 editierte Sonderband der „Sozialen Welt" gewidmet ist) und steht damit beispielhaft für den in der deutschen Nachkriegsgeschichte immer häufiger konstatierten Bedeutungswandel sozialer Ungleichheit (dazu P. Berger 1987: 66). Findet wirklich alle paar Jahrzehnte ein derartiger Epochenwandel statt, dass bestehende Ansätze obsolet sind? Oder leben wir aktuell nur in einer besonders „unsicheren" Zeit?

Aus Sicht der vorangegangenen Überlegungen ist das rasche „Verfallsdatum" der Konzepte vor allem ein Anzeichen ihrer abwegigen Ausrichtung. Die Soziologie sozialer Ungleichheit scheint wesentlich von der Vorstellung gekennzeichnet zu sein, dass bereits ein höherer Differenzierungsgrad von (Klassen-)Ansätzen zu mehr Erklärungskraft verhilft. Das notorische Dilemma, nicht zwischen weitreichenden, abstrakten Erklärungen einerseits (wie sie etwa die älteren Klassentheorien repräsentieren) und Lebens- bzw. Realitätsnähe andererseits (auf welche insbesondere Lebensstil- und Schichtkonzepte abzielen) vermitteln zu können, wird damit gleichwohl nicht gelöst. Zugewinne an Lebensnähe werden durch die Aufgabe genereller Erklärungsmechanismen erkauft, zu allgemeine Theorien scheitern hingegen an der Realität.

Zur Überwindung dieser scheinbar unlösbaren Widersprüche scheint ein Perspektivenwechsel ganz anderer Art gefragt. Die hier vertretene These ist, dass der Fokus von *konkreten* Verteilungs*mustern* stärker auf die *generellen* Verteilungs*mechanismen* gelenkt werden sollte. Akzeptiert man die Vorstellung, dass die konkreten Ausdrucksformen eines „guten Lebens" und damit die zu seiner Realisierung erforderlichen Ressourcen historisch wandelbar sind (denn sie sind Ergebnis kontingenter Aushandlungsprozesse, kultureller Modeströmungen oder Folge einer kontinuierlichen Abschottung von Eliten; s. für solche Mechanismen Coleman 1990: Kap. 9, 10 bzw. Bourdieu 1981), überrascht kaum mehr, dass sich die verwendeten Klassifikationen mit der Zeit überholen.[24] Ein Axiom erklärender Soziologie ist aber, dass es einige *generelle* Mechanismen gibt, die Ungleichheit hervorbringen, und dies unabhängig davon, welche Ressourcen im konkreten Einzelfall zu verteilen sind. Als grundlegendste Komponente ist diesen Mechanismen gemein, dass sie auf (interdependente) Entscheidungen von Akteuren zurückgehen. Beispielsweise sind Allokationen von Forschungsdrittmitteln und Weiterbildungsmaßnahmen Folge der Entscheidungen von Forschenden und Arbeitnehmern, entsprechende Anträge zu stellen, sowie der Entscheidungen von Gutachtenden und Arbeitgebern, diese

[24] So kommt etwa Aspekten wie einem sicheren Grundeinkommen, viel Freizeit oder transnationalen Kontakten historisch ein jeweils anderer Wert für den symbolischen Ausdruck eines gelungenen Lebens zu (ähnlich Burzan 2007: Kap. 1).

Anträge positiv zu bewerten. Diese Handlungen sind wiederum wesentlich sozial vorstrukturiert, also etwa aufgrund der bestehenden Ressourcenverteilungen, sozialen Netzwerke und Erwartungen *verständlich*. Das Interesse verlagert sich damit auf die Erklärung, warum Akteure handeln wie sie handeln; und hierbei speziell auf die Kontextbedingungen, die begünstigen oder verhindern, dass Handlungen und Entscheidungen zu Ungleichheiten führen, also beispielsweise bewirken, dass Frauen und Männer trotz gleicher Leistungen unterschiedlichen Zugang zu Forschungsmitteln oder Weiterbildungen erhalten.

Dabei dürfte es zumindest in modernen Gesellschaften nicht nur einen *einzigen* relevanten Mechanismus geben. Kontextbedingungen wie die Organisation des Wissenschaftssystems, die Ausdifferenzierung von Weiterbildungsmärkten oder das geltende Recht schaffen wichtige, aber veränderliche Voraussetzungen für das Greifen Ungleichheit generierender Mechanismen. Aus ähnlichen Gründen ist davon auszugehen, dass sich die vorherrschenden Mechanismen zwischen den Bereichen der Ressourcenverteilung unterscheiden, also etwa im Wissenschaftssystem andere Verteilungsmechanismen vorrangig sind als in privaten Haushalten und Familien (ähnlich Rössel 2009: 51). Die Suche nach einer *einzigen* Globaltheorie, die alle möglichen Ungleichheiten dauerhaft zu erklären vermag, erscheint vor diesem Hintergrund von vornherein zum Scheitern verurteilt.

Dennoch werden Erklärungen mittels Mechanismen in mehrfacher Hinsicht dem Anspruch einer hohen Reichweite gerecht: Zunächst, weil sie auf Verteilungs*prozesse* fokussieren, die unabhängig von den zu verteilenden Ressourcen wirken und daher im Idealfall für mehrere ähnlich gelagerte Phänomene ursächlich sind (also beispielsweise ebenso für die Kumulation von Forschungsmitteln wie für die von Weiterbildungen; s. dazu speziell DiPrete/Eirich 2006); ferner, weil ihre Erklärungslogik nicht auf bestimmte historische Zeiten oder Orte beschränkt ist, sondern sie unabhängig davon anwendbar ist, wer konkret in welchem Kontext mit wem agiert. Und schließlich kommt der erklärenden Soziologie auch deshalb eine hohe Allgemeingültigkeit zu, weil sich mit ihr prinzipiell jegliche Allokationen in den Blick nehmen lassen, die in irgendeiner Form auf soziales Handeln zurückgehen.

Die noch am ehesten mit dem Anspruch der *Erklärung* verbundenen Klassenkonzepte sind demgegenüber unvermeidlich an die Stellung von Akteuren im Erwerbsprozess oder gemeinsame (Klassen-)Identitäten gebunden. Die hierüber definierten Klassen können, müssen aber nicht mit zentralen Achsen der Ungleichheit zusammenfallen. So decken sich Erwerbslagen nicht unbedingt mit Gruppenidentitäten und stehen beide unter Umständen quer zu Milieus und Lebensstilmustern. Das Zusammenfallen in einheitlichen Klassen ist also eher die zu erklärende Ausnahme als Regel (ähnlich Pakulski 2005). Dies wird gerade in jüngster Zeit offenkundig, denn in einer Feststellung stimmen alle Gegenwartsdiagnosen überein: in der Ausdifferenzierung der Sozialstruktur.

Das unbedingte Festhalten an Klassen-, Schicht- und sonstigen Sozialstrukturkonzepten anstatt der Konzentration auf die Erforschung der zugrunde liegenden Mechanismen erscheint daher zumindest in fortgeschrittenen Forschungsstadien nachrangig. Wozu die Anstrengungen, gefundene Erklä-

rungen mit bestehenden Typologien zu versöhnen? „Rückübersetzungen“ der Grundkonzepte der Sozialstrukturanalyse in handlungstheoretische Modelle (s. Kreckel 2004: Kap. II.3; Rössel 2005: 10 für derartige Versuche) oder auch die Suche nach den Fragen, die durch Klassentheorien noch beantwortet werden könnten (Wright 2005), scheinen primär von akademischem statt praktischem Nutzen. Jedenfalls drohen derartige Bemühungen den heuristischen Wert der Konzepte zu überschätzen. Die zwischen Berufsgruppen, Besitzklassen oder Milieus verlaufenden Grenzen sind schließlich keine *per se* verhaltensrelevanten Muster (klammert man die ursprünglichen, aber zu Recht für ihre hohe Abstraktion kritisierten und inzwischen historisch überholten antagonistischen Großgruppen des industriellen Kapitalismus aus). Ein praktizierter Weg zur Wahrung des Bedeutungsgehalts dieser Modelle besteht darin, je nach Erklärungsziel von anderen Klassenkonzepten auszugehen (s. z. B. Wright 2005). Dies läuft im Wesentlichen aber genau auf das hier propagierte Forschungsprogramm hinaus, wonach unterschiedliche Ungleichheit generierende Mechanismen anzunehmen sind. Jedoch nicht ganz – denn Klassentheorien stehen dann zusätzlich immer noch vor der Aufgabe, eine Verbindung zu den – für sie letztlich essenziellen – Erwerbsklassen herzustellen. Der alternative Weg zur Rettung von Klassentheorien in Form einer *empirisch* angeleiteten Disaggregation (um die Idee *subjektiv* bedeutsamer Gruppenidentitäten zu wahren) droht besonders irreleitend zu sein, verlaufen die hierauf getrimmten Klassifikationen doch womöglich quer zu den für andere Mechanismen relevanten Kontextbedingungen. Versuche, makrosoziologische Konzepte wie Klassen und Schichten handlungstheoretisch zu fundieren, gehen zudem deshalb verkehrt herum vor, weil einzelne Mechanismen gegenläufige Effekte hervorbringen können und sich somit womöglich auf der Makroebene gegenseitig konterkarieren. Statt allzu viel Energie auf die Überzeugung Anderer von der „wahren“ Bedeutung dieser Konzepte zu verwenden, scheint es wichtiger, mit der erklärenden Arbeit fortzufahren (Hedström 2005: 154).

Dazu kann inzwischen durchaus auf Errungenschaften aufgebaut werden. Theoretisch stringente und empirisch bewährte Erklärungsbausteine für die Genese sozialer Ungleichheiten sind etwa:

(1) Die sich selbst verstärkende Wirkung von hohem Status und Reputation (über deren positive Signalwirkung für weitere Ressourceneinwerbungen (s. z. B. Merton 1988; Podolny/ Lynn 2009 für Beispiele).

(2) Macht und soziales Kapital in Form einer unterschiedlich guten Verfügbarkeit von alternativen Tauschpartnern, damit Unabhängigkeit und Verhandlungsstärke bei Entscheidungen über Ressourcenaufteilungen (exemplarisch Molm et al. 1999; Rössel 2004).

(3) Soziale Schließungen, Segmentierungen in (berufliche) Teilmärkte sowie andere Beschränkungen des freien Wettbewerbs, welche die Aneignung eines nicht durch Leistung gerechtfertigten Mehrwertes (*rent seeking)* ermöglichen (z. B. J. Berger 2004; Madden 1973).

(4) Anwendungen von Heuristiken zur Komplexitätsreduktion (wie die Heranziehung von Gruppenmittelwerten und anderen Signalen zur Abschätzung der Vertrauenswürdigkeit oder Leistungsfähigkeit von Interaktionspartnern), was im Ergebnis ebenfalls zu leistungs*ungerechten* Verteilungen bzw. statistischen Diskriminierungen führen kann (s. z. B. Arrow 1973; Pager/ Shepherd 2008).

Dabei ist davon auszugehen, dass diese Mechanismen weniger nebeneinander bestehen, als vielmehr ineinander greifen. So dürfte eine hohe Verhandlungsmacht soziale Schließungen und Wettbewerbsbeschränkungen (Monopolisierungen) erleichtern und sollten Signal- und Reputationsmechanismen insbesondere in solch komplexen Situationen zur Wirkung kommen, in denen die Anwendung von Heuristiken rational erscheint. Die Trennung dieser Mechanismen ist also eine heuristische. Für den Erkenntnisfortschritt erscheint es dennoch hilfreich, zunächst von getrennten Prozessen auszugehen und deren Parameter durch empirische Prüfungen zu justieren, um erst in einem weiteren Schritt über ihre schlüssige Verkettung nachzudenken (Hedström 2005: Kap. 6).

Der Ansatz der vorliegenden Arbeit zur Fortentwicklung der Ungleichheitssoziologie ist es also, die Forschungsbemühungen stärker auf die Erforschung solcher *generellen Mechanismen* zu konzentrieren, anstatt die Suche nach einem möglichst genauen Raster für ihre – stets durch zusätzliche Faktoren beeinflussten und damit kontingenten – Ergebnisse auf der Makroebene fortzusetzen. Es erscheint beispielsweise sinnvoller, allgemeine Prozesse der Erzielung von Vertrauen, Reputation und weiteren Ressourcen mittels Prestige und Status zu untersuchen, als sich vorrangig darum zu bemühen, die Prestige- und Statuswerte einzelner Berufsgruppen möglichst akkurat zu bestimmen. Ähnlich sollten vorrangig *generelle* Bedingungen untersucht werden, unter denen objektive Lagen subjektiv als solche erfahren werden und in deren Folge soziale Schließungen und Protest wahrscheinlicher werden, anstatt lediglich Milieus und Lebensstile äußerst exakt zu beschreiben. Weiter sind die Mechanismen von Wettbewerbsbeschränkungen und Diskriminierungen tiefgründiger zu erforschen, unabhängig davon, entlang welcher (transnationaler) (Klassen-)Grenzen diese im konkreten Einzelfall verlaufen. Es geht also darum, Mechanismen der Ungleichheitsgenerierung und -tolerierung unabhängig von ihren konkreten Inhalten herauszuarbeiten: „The advancement of sociology could then be seen in a growing body of knowledge of ever-more mechanisms rather than ever-better theories“ (Blossfeld 1996: 201). Im Idealfall führt dies dazu, dass Klassen- und Schichtkonzepte irgendwann als *Erklärung* obsolet werden, da die hinter ihnen stehenden Mechanismen entschlüsselt sind (Breen/Rotman 1995).

Dies ist nicht misszuverstehen als die radikale Forderung nach einer vollständigen Verabschiedung von Klassen-, Schicht- und anderen Sozialstrukturkonzepten. Sozialstrukturmuster können in einem ersten Schritt hilfreiche Hinweise dazu liefern, welche Faktoren an der Entstehung sozialer

Ungleichheit beteiligt sind.[25] So helfen Klassen- und Schichtkonzepte speziell bei der Spezifikation der Logik der Situation, also der Feststellung, innerhalb welcher Kontexte die Akteure ihre Handlungswahlen treffen. Ebenso schaden detaillierte Beschreibungen von Milieus und Lebensstilen sicher nicht, wenn es darum geht, relevante Rahmenbedingungen des Handelns zu erkennen (Schmid 2009). Nach dem hier vertretenen Soziologieverständnis sind damit aber nur erste Grundlagen für die gewichtigere Aufgabe geschaffen, nämlich die *Ursachen* sozialer Ungleichheiten zu verstehen. Die Signifikanz von Typologien wie Klassen und Schichten ist in dieser Perspektive vor allem ein Hinweis dafür, dass die eigentlich interessierenden kausalen Mechanismen noch nicht gefunden sind (Hedström 2005: 153; Hedström/Swedberg 1996: 289; Reskin 2003: 14). Denn will man erklären, warum – geht man wiederum von Klassenkonzepten aus – gerade Erwerbsklassen für die Generierung sozialer Ungleichheit relevant sind, wird man nicht umhin kommen, auf die mit diesen Erwerbsklassen verbundenen Ressourcenausstattungen und Netzwerke zu rekurrieren, auf differente Möglichkeiten der sozialen Schließung, Mehrwertaneignung (*rent seeking*) oder Durchsetzung gemeinschaftlicher Interessen mittels Gruppensolidarität – oder mit anderen Worten: man wird nicht umhin kommen, erklärende Mechanismen anzuführen.

1.3 Inhalte und Erkenntnisleistungen der eigenen Anwendungsbeispiele

Dass es gangbare Wege gibt, Erkenntnisfortschritte bei der Erklärung sozialer Ungleichheiten zu erzielen, soll in den nachfolgenden Kapiteln anhand von eigenen empirischen Studien mit bewusst sehr unterschiedlichen Inhalten und Methoden veranschaulicht werden. Die drei Beiträge fokussieren auf verschiedene Lücken der Ungleichheitssoziologie und demonstrieren zugleich, wie sich wesentliche Aspekte der in den Abschnitten 1.2.1 und 1.2.2 genannten Prämissen erklärender Soziologie umsetzen lassen. Nachfolgend werden jeweils ihre Motivation, Inhalte und Erkenntnisleistungen zusammengefasst.

1.3.1 Herausbildung einer akademischen Elite?

Der Bereich, der sich am ausdrücklichsten meritokratischen Normen verschrieben hat, dürfte das Bildungs- und darunter das Wissenschaftssystem sein. Allein die Leistung von Personen sollte über die Vergabe von Bildungstiteln und Forschungsressourcen sowie über wissenschaftliche Karriereerfolge entscheiden. Dafür, dass dieses Prinzip nicht *vollständig* praktiziert wird, gibt es inzwischen verlässliche Belege. Kaum wegzudiskutieren ist beispielsweise das mittels (natürlicher) Experimente

[25] Besonders anschaulich sind empirische Dekompositionen von Ungleichheiten in ihren innerhalb (*within*) und zwischen (*between*) Gruppen bestehenden Anteil. Einen aufschlussreichen Spezialfall davon stellen *kontrafaktische Analysen* dar, mit denen sich beispielsweise Lohnunterschiede zwischen Frauen und Männern in eine Komponente zerlegen lassen, die auf ungleiche Humankapitalausstattungen zurückgeht, versus eine Komponente, die durch ungleiche Entlohnung gleicher Ausstattungen bedingt ist (Blinder 1973; Oaxaca 1973 für Details).

gewonnene Ergebnis, dass die Annahmequote wissenschaftlicher Artikel bei einem *double blind peer review* (weder Gutachtende noch Autoren sind füreinander erkennbar) weniger von askriptiven (qua Geburt zugeschriebenen) Merkmalen der Autoren wie ihrem Geschlecht beeinflusst ist als bei nichtanonymen Begutachtungsprozessen (z. B. Blank 1991). Auf der Makroebene beobachtbare Ungleichheiten – wie beispielsweise die geringere Wahrscheinlichkeit von Frauen gegenüber Männern, eine Professur zu erlangen, oder auch die ungleichgewichtige Verteilung von Forschungsdrittmitteln auf Forschungsinstitutionen (CEWS 2009 bzw. DFG 2009 für Statistiken) – können jedoch ebenso durch meritokratische Prozesse bedingt sein. In Frage kommende Ursachen für die genannten Aspekte sind etwa unterschiedliche Neigungen von Wissenschaftlerinnen und Wissenschaftlern, zugunsten von familiären Aufgaben auf wissenschaftliche Karrieren zu verzichten; oder rechtschaffene Unterschiede in der Leistungsfähigkeit von Forschenden, die an verschiedenen Institutionen tätig sind.[26] Knapp resümiert sind nach vorliegendem Erkenntnisstand soziale Ungleichheiten im Wissenschaftssystem unbestritten, ihr Ausmaß und die sie bedingenden Mechanismen hingegen unzulänglich entschlüsselt (s. z. B. Bornmann et al. 2007 für eine Metaanalyse zum „Geschlechterbias" beim *peer-review*; C. Gross/Jungbauer-Gans 2007 für einen generellen Überblick).[27]

Als Forschungsobjekt ist das Wissenschaftssystem für die Soziologie sozialer Ungleichheit zudem deshalb interessant, weil sich dank der hier besonders expliziten Betonung von meritokratischen Normen und der in der Folge vergleichsweise starken Standardisierung von Entscheidungsverfahren das Einhalten legitimierter Verteilungsprozesse direkter beobachten lässt als in anderen Bereichen. So wäre beispielsweise beim generellen Arbeitsmarkt eine größere Heterogenität an Institutionen, Karriereleitern und (Tarif-)Regelungen zu berücksichtigen. Gleichwohl handelt es sich beim Wissenschaftssystem um ein komplexes Feld, welches ein ausgefeiltes theoretisches wie empirisches Instrumentarium voraussetzt, um keine falschen kausalen Schlüsse zu ziehen. Für die vorliegende Abhandlung bietet es damit einen idealen Ausgangspunkt, um den Analysegewinn erklärender Soziologie *empirisch* zu veranschaulichen.

Konkret bildet den Gegenstand der zusammen mit den Koautoren Thomas Hinz und Jürgen Güdler verfassten Studie in Kapitel 2 die Allokation von Forschungsdrittmitteln der Deutschen

[26] Plausibel sind zudem Selektionseffekte mit sich selbst verstärkender Wirkung: Hochrangige Forschungsinstitutionen mit guten Ressourcenausstattungen können die hochkarätigeren Wissenschaftler anwerben. Dieser *Match* führt zu besonderen Synergien im Hinblick auf die wissenschaftliche Leistungsfähigkeit und damit wiederum zu überdurchschnittlichen Anwerbungen von weiteren begehrten (Forschungs-)Ressourcen.

[27] Die gesellschaftspolitische Notwendigkeit, diesbezüglich verlässlicheres Wissen zu generieren, ergibt sich ebenso aus dem Verbot einer Diskriminierung wie aus dem Anliegen heraus, „Begabungsressourcen" möglichst effizient, das heißt leistungsgerecht auszuschöpfen. Dass diesbezüglich Handlungsbedarf besteht, manifestiert sich in Deutschland besonders eindrucksvoll an dem – trotz einem inzwischen überdurchschnittlichen Frauenanteil unter den Hochschulabsolventen – immer noch weit unterdurchschnittlichen Frauenanteil bei Professuren (16 Prozent im Jahr 2007; vgl. CEWS 2009). Aktualitätsbezug hat das Thema in Deutschland zudem durch die Exzellenzinitiative: Gerade über ihre Verfahren wird spekuliert, ob sie die angestrebte Stimulierung eines meritokratischen Wettbewerbs anstoßen oder im Gegenteil dessen Verzerrung bewirken, indem etwa das Renommee von Exzellenzuniversitäten den dort beschäftigten Wissenschaftlern *per se* zu mehr Ansehen und damit Karriere- und Publikationschancen verhilft (z. B. C. Gross/Jungbauer-Gans 2007; Münch 2007; Wagner 2007).

Forschungsgemeinschaft (DFG) in den Jahren 1991 bis 2004. Die untersuchte *DFG-Einzelförderung* stellt das – zumindest gemessen am Finanzvolumen – wichtigste Verfahren der Drittmittelförderung im deutschen Wissenschaftssystem dar (DFG 2009). Entsprechend dürfte die Verteilung dieser Mittel besonders ausschlaggebend für das Gelingen individueller wissenschaftlicher Karrieren und auch für die Konkurrenzfähigkeit von Universitäten im (internationalen) Wettbewerb sein. Die zentrale Frage ist nun, inwieweit die Verteilung dieser Forschungsmittel leistungsgerecht erfolgt oder aber Größen- und Reputationseffekte eine Rolle spielen, die letztlich zu einer immer stärkeren Anhäufung von Forschungsressourcen an einer exklusiven Auswahl von immer erfolgreicheren Universitäten führen müssten. Mit dieser Fragestellung ist die Analyse in Kapitel 2 zugleich als Replikation einer Aufmerksamkeit erregenden Studie von Richard Münch (2007) angelegt – und somit als Replikation einer Studie, die allenfalls eklektisch auf erklärende Mechanismen zurückgreift. Gerade diese Kontrastierung unterschiedlicher Herangehensweisen an demselben Gegenstand dürfte den Erkenntnisgewinn durch eine stärker analytisch ausgerichtete Soziologie aufzeigen.

Zum besseren Verständnis sind zunächst einige Ausführungen zum Vorgehen Münchs erforderlich. Nach Münch werden Forschungsmittel immer weniger in einem offenen, meritokratischen Wettbewerb verteilt, sondern konzentrieren sie sich stattdessen zunehmend auf eine „akademische Elite“ von einigen wenigen, hoch renommierten Universitäten. Wie diese Monopolisierung vonstatten geht, bleibt allerdings vage. Der von Münch vorgenommene Verweis auf Kapital-Umwandlungen, auf Begriffe wie „Machtkartelle“ und „Kennziffernregimes“ und ebenso die losen Bezugnahmen auf klassische Theorien von Karl Marx, Norbert Elias, Pierre Bourdieu oder Michel Foucault genügen den Anforderungen präziser Erklärungen allesamt kaum, zumal Argumente oft überspitzt und zum Teil gar widersprüchlich sind (Hornbostel 2009; Kaube 2007; Krücken 2009; Mayer 2009 für inhaltliche Kritiken). Zwar werden verschiedentlich kausale Mechanismen unterstellt (wie Reputationseffekte und eine Vetternwirtschaft), diese werden dann aber nicht im Detail ausgeführt und keineswegs in exakte Hypothesen überführt. Weitere Unzulänglichkeiten aus Sicht einer analytischen Soziologie bestehen darin, dass die empirisch verwendeten Korrelations- und Regressionsanalysen zumeist auf der Makroebene verharren, aufgrund des Querschnittcharakters keine zeitliche Entschlüsselung leisten und zudem keine exakte Modellierung der zumindest implizit unterstellten (nichtlinearen) Zusammenhänge umfassen.

Die in Kapitel 2 präsentierten Analysen beruhen auf sehr ähnlichem Datenmaterial (prozessproduzierten Daten der DFG), fokussieren aber theoretisch auf die Handlungslogik einzelner Gutachter und Antragsteller und greifen methodisch auf Individualdaten zu einzelnen Forschungsanträgen zurück. Sie beinhalten zunächst eine Übersetzung der Annahmen Münchs in präzisere theoretische Mechanismen. Hierzu kann von vorliegenden Forschungsarbeiten profitiert werden: Der Gegenstand fällt in das Anwendungsgebiet des in der analytischen Soziologie geradezu als Idealbeispiel für einen Mechanismus geltenden Matthäus-Effekts (sich selbst verstärkende Wirkung von Forschungsressourcen gemäß dem Prinzip: „wer hat, dem wird gegeben“; DiPrete/Eirich 2006;

Hedström/Udehn 2009; Merton 1968, 1988). Aus verschiedenen Teilmechanismen (insbesondere Signal-, Selektions- und Humankapitaleffekte; statistische Diskriminierung) werden insgesamt zehn Hypothesen zur Relevanz der Größe und Reputation von Universitäten für die Drittmitteleinwerbung abgeleitet. Die empirischen Prüfungen orientieren sich gleichfalls stärker als bei Münch an Idealen analytischer Soziologie. So setzen sie überwiegend auf der Individualebene einzelner Drittmittelanträge (statt auf der Makroebene ganzer Universitäten und Wissenschaftsbereiche) an. Weiter bieten sie zeitliche Entschlüsselungen der postulierten Ursachen und Wirkungen und werden lange Zeitreihen betrachtet, um Zufallsschwankungen auszuschließen. Schließlich werden die postulierten Mechanismen möglichst direkt in statistische Modelle übersetzt, indem etwa Polynome zur Modellierung nichtlinearer Zusammenhänge sowie multivariate Mehrebenen- und Zähldatenmodelle herangezogen werden, die besonders eng mit der Datenstruktur korrespondieren.

Dass sich diese verfeinerten Analysen auszahlen, zeigen die zu Münch konträren Ergebnisse. Fast keine seiner Thesen hält der genaueren Prüfung Stand.[28] So spielen die Größe und Reputation von Universitäten für die individuellen Bewilligungschancen insgesamt kaum eine Rolle und sind lediglich zwei Resultate mögliche Hinweise auf die von Münch angenommene Verletzung meritokratischer Normen: Anträge aus Traditionsuniversitäten weisen etwas höhere Bewilligungschancen auf als Anträge aus anderen Universitäten; und Antragsteller aus ostdeutschen Universitäten haben etwas weniger Erfolg als ihre westdeutschen Kollegen. Der Beitrag in Kapitel 2 demonstriert damit in der Tat einen deutlichen Erkenntnisgewinn durch die Einhaltung von Prämissen erklärender Soziologie. Inhaltlich besteht dieser Zugewinn vor allem in dem – international beachtlichen – Ergebnis, dass der Matthäus-Effekt im Hinblick auf die Drittmitteleinwerbung von Forschungsinstitutionen geringer ist als allgemein angenommen (s. z. B. Moed 2005; Wagner 2007). Umfangreiche Ressourcen und eine hohe Reputation von Forschungseinrichtungen verhelfen den an diesen Einrichtungen beschäftigten Forschenden nicht *per se* zu besseren Erfolgschancen, dies lässt sich zumindest für die DFG-Einzelförderung konstatieren. Ob nicht doch Verstöße gegen die Leistungsgerechtigkeit vorliegen, ist allerdings mit dem hier verwendeten Untersuchungsdesign (und allgemein mit *ex-post-facto*-Analysen) nicht auszuschließen, schließlich sind ebenso verdeckte Diskriminierungen möglich (einzelne Forschende müssen für den gleichen Erfolg höhere Leistungen erbringen, s. die wegweisende Untersuchung von Wennerås/Wold 1997).[29] Die in Kapitel 3 beschriebene und im folgenden Abschnitt 1.3.2 zusammengefasste Studie illustriert daher, wie sich eine Testung von kausalen Zusammenhängen alternativerweise mit einem weitaus stärker auf eine Theorienprüfung zugeschnittenen, *experimentellen* Verfahren umsetzen lässt.

[28] Allerdings wird nur ein Ausschnitt von Münchs Thesen untersucht, siehe dazu auch die Kritik von Münch selbst (Münch 2009).

[29] Wennerås und Wold fanden heraus, dass weibliche Bewerber um Forschungsstipendien beim *Swedish Medical Council* für den gleichen Erfolg höhere Publikationsleistungen vorweisen mussten als männliche.

1.3.2 Migration Decisions Within Dual-Earner Partnerships

Die wechselseitige Verschränkung von Karrieren in einem gemeinsamen Haushalt gilt als einer der großen „weißen Flecke“ der Ungleichheitssoziologie (P. Berger/Hradil 1990: 10f.; Drobnič/Blossfeld 2004; McLanahan/Percheski 2008; Rössel 2009: 56; Solga/Wimbauer 2005). Haushalte wurden traditionell in der Ungleichheits- und Familiensoziologie lediglich als ganze Einheiten berücksichtigt oder ihre Mitglieder als separate Individuen aufgefasst, um Haushaltsmerkmale als reine Kontextvariablen betrachten zu können. Die *Partnerschaft als Ungleichheit generierende Handlungseinheit* blieb dagegen lange Zeit ausgeblendet (Drobnič/Blossfeld 2004). Ihre hohe Relevanz wird aber beispielsweise an der gut belegten Vorrangigkeit von männlichen gegenüber weiblichen Berufskarrieren in familiären Entscheidungssituationen offensichtlich. Diese Priorität männlicher Karrieren gilt wiederum als wichtige Stellgröße für geschlechtsspezifische Lohnunterschiede und Statusdifferenzen. Eine bessere Abbildung (strategischer) Interdependenz von Akteuren ist zugleich ein wesentliches Forschungsdesiderat der analytischen Soziologie (Hedström 2005: 109ff.). Mit der Spieltheorie liegen für diese inzwischen sehr elaborierte Modelle vor, deren Analysetiefe war jedoch in empirischen Studien in aller Regel bislang nur ansatzweise einzulösen.[30]

Die zusammen mit Martin Abraham und Thomas Hinz verfasste Studie in Kapitel 3 demonstriert beispielhaft, dass die in der analytischen Soziologie noch wenig genutzte Methodik des Faktoriellen Survey einen Ausweg aus diesen bestehenden Forschungsproblemen bieten kann. Die Grundidee des Verfahrens besteht darin, Befragte im Rahmen einer Umfrage fiktive Fallbeispiele (*Vignetten*) im Hinblick auf deren Gerechtigkeit, eine anzuratende Handlung oder eigene Handlungspräferenz beurteilen zu lassen (M. Beck/Opp 2001; Jasso 2006 für Einführungen). In diesen Fallbeispielen werden einzelne Merkmale (*Dimensionen*) experimentell variiert. Damit bietet die Methodik eine Kombination der Vorteile von Experiment und Umfrage: Wie in jedem Experiment sind aufgrund der gezielten Variation der Stimuli verlässliche Attribuierungen kausaler Einflüsse möglich;[31] durch die Integration in eine Befragung sowie die gängige Praxis, einzelnen Probanden mehrere Fallbeispiele vorzulegen, sind zugleich hohe Fallzahlen mit einem gegenüber Laborexperimenten vergleichsweise geringeren Aufwand realisierbar, und dies selbst mit sehr heterogenen Befragtenstichproben (s. Auspurg et al. i. E.b und 2009c für eine ausführliche Diskussion der Vor- und Nachteile). Der damit

[30] Mit den klassischerweise herangezogenen Laborexperimenten sind mehrdimensionale Versuchsdesigns, die zur Prüfung von komplexen Fragestellungen und Wechselwirkungen zwischen kausalen Faktoren unabdingbar sind, nur mit hohen Anzahlen an Probanden und somit erheblichem Aufwand umzusetzen. Zudem beschneiden die hohe Künstlichkeit und gängige Beschränkung auf einfach erreichbare Teilnehmende (Studierende) ihre Validität (vgl. Abschnitt 1.2.2). Die ideale Methodik, um breite Bevölkerungsschichten zu erreichen und zugleich hohe Fallzahlen sowie „alltagsnahe“ Beobachtungen zu realisieren – die Durchführung von Umfragen – wird dagegen für ihre vermeintliche Unzulänglichkeit kritisiert, die strategische Interdependenz von Akteuren erfassen zu können (s. z. B. Hedström 2005: Kap. 5).

[31] Die Methodik, einzelnen Befragten *kontrafaktische* (sich nur in einzelnen Merkmalen unterscheidende) Fallbeispiele vorzulegen, umgeht weitestgehend das fundamentale Problem kausaler Inferenz (dazu Holland 1986: 947), nämlich dass dasselbe Untersuchungsobjekt (hier der einzelne Befragte) unmöglich verschiedenen Bedingungen gleichzeitig ausgesetzt sein kann.

eröffnete Weg einer besonders direkten Prüfung theoretischer Zusammenhänge wurde bislang schwerpunktmäßig in der Einstellungs- und Gerechtigkeitsforschung genutzt (Wallander 2009 für einen aktuellen *Review*). Kapitel 3 demonstriert anhand der beschriebenen Forschungslücke der Interdependenz im Haushaltskontext, dass sich das Verfahren ebenso gewinnbringend für die analytische Soziologie einsetzen lässt.

Konkret besteht das Anliegen der Studie darin, die *bis dato* unzufrieden gelöste Entschlüsselung zweier alternativer Erklärungen für die Entstehung geschlechtsspezifischer Ungleichheiten im Haushaltskontext voranzubringen. Dies sind einerseits austausch- und verhandlungstheoretische Ansätze, die – ausgehend von der Annahme individueller Nutzenmaximierung – die Verhandlungsmacht der Akteure (konzipiert als Ausmaß ihrer Abhängigkeit vom Partner bzw. ihrer Verfügbarkeit von Alternativen zur Partnerschaft) als einen geschlechtsneutralen Koordinationsmechanismus spezifizieren, um das Entscheidungsverhalten in Partnerschaften zu prognostizieren. Andererseits kommen sozialisations- und identitätstheoretische Erklärungen in Betracht, welche die zentrale Handlungslogik in der Erfüllung von Geschlechterrollen bzw. der Herstellung von Geschlechtsidentitäten (*doing gender*) sehen. Da beide Ansätze gleichermaßen einen *logisch* einleuchtenden Mechanismus für die Genese geschlechtsspezifischer Ungleichheiten im Haushaltskontext vorschlagen, lässt sich ihre tatsächliche Erklärungsleistung nur mittels empirischer Untersuchungen feststellen. Dieses Unterfangen ist mit *vorliegenden* Daten jedoch nur sehr bedingt zu realisieren. Hauptgründe sind die starke Korrelation der für die Verhandlungsmacht zentralen Erwerbsoptionen mit dem für die Rollentheorien substanziellen Geschlecht der Akteure sowie Verzerrungen durch einen *sample selection bias* (Personen mit guten Erwerbsoptionen bzw. hoher Verhandlungsmacht stellen keine zufällige Auswahl aus allen Erwerbspersonen dar).[32]

Wie der Beitrag in Kapitel 3 darlegt, sind diese Probleme mit dem Design eines Faktoriellen Surveys zu bewältigen. Die eigens für dieses Forschungsziel konzipierte Studie nutzt konkret hypothetische Beschreibungen von beruflichen Umzugsanreizen, um die Stellgrößen für die Priorität der einen über die andere Karriere herauszuarbeiten. Ein methodisches Novum ist überdies, jeweils beide Partner in Doppelverdiener-Haushalten über „spiegelbildliche“ Fallbeispiele zu befragen (ein Partner wird in die Position des Umziehenden, der oder die Andere in die Position des Mitziehenden versetzt; s. Auspurg et al. 2009a für Details). Dies erlaubt es, ihre Antworten direkt aufeinander zu beziehen und somit einem zentralen Anspruch analytischer Soziologie gerecht zu werden, nämlich anstelle von isolierten Entscheidungen die Interdependenz von Akteuren zu erfassen. Zugleich wird mit diesem Erhebungsdesign das in der Ungleichheitssoziologie und Familienforschung vielfach angemahnte

[32] Am Beispiel der als Anwendungsfall dienenden beruflichen Umzugsentscheidungen erläutert bedeutet dies: Beruflich mobile Personen sind generell karriereorientierter und beruflich erfolgreicher und würden somit vermutlich auch ohne einen beruflichen Umzug überdurchschnittliche Karrierefortschritte erzielen. Daher ist mit Umfragedaten kaum zu beziffern, wie hoch der Einkommensgewinn aufgrund des Umzugs ausfällt, zudem sind es vor allem Männer, die Beschäftigungen in überregional organisierten Arbeitsmärkten ausüben. Dies erschwert wiederum die Beurteilung, inwieweit familiale Umzugsentscheidungen orientiert an Einkommensgewinnen oder Geschlechtsidentitäten erfolgen (s. z. B. Jürges 2005).

Desiderat (z. B. Brückner 2009; Huinink 2006) erfüllt, tatsächlich eine *Paar*befragung durchzuführen, anstatt sich immer nur auf die (vielfach invaliden) Proxyangaben zu den Partnern zu verlassen.

Im Ergebnis erweist sich das hypothetische Entscheidungsverhalten als weitgehend geschlechtsneutral und werden einige weitere Kernannahmen der Verhandlungstheorie gestützt. Daneben finden sich schwache Evidenzen für geschlechtsspezifische Ansätze. Dies legt es – vorbehaltlich weiterer Studien, die in diesem Bereich unzweifelhaft erforderlich sind – nahe, die beiden Mechanismen zugunsten eines höheren Erklärungsgehalts miteinander zu verzahnen (s. z. B. Blossfeld/Müller 1996 für einen solchen Vorschlag). Insbesondere die Erkenntnis, dass das Geschlecht der Akteure eine lediglich untergeordnete Rolle spielt, ist in der Forschungslandschaft sehr bemerkenswert, da hier zwar gerne die schwierige Trennbarkeit der beiden Theorierichtungen (austausch- und verhandlungstheoretische versus rollen- und sozialisationstheoretische Ansätze) konstatiert wird, dann aber oftmals die Ergebnisse doch recht einseitig (und nach den hier vorliegenden Befunden zu vorschnell) als Evidenzen geschlechtsspezifischer Normen gedeutet werden (z. B. Jürges 2005). Normen und Rollenerwartungen drohen damit zu Konzepten zu verkümmern, mit welchen sich pauschal eine verbleibende Restvarianz (nach Geschlecht) plausibilisieren lässt. Nach den hier erzielten Ergebnissen sollte genauer geprüft werden, inwieweit tatsächlich differente normative Erwartungen an die Akteure handlungsrelevant sind oder aber geschlechtsspezifische Anreizstrukturen vorliegen, welche selbst bei völlig geschlechts*neutralen* Entscheidungsmechanismen zu geschlechts*spezifischen* Resultaten führen können.[33]

Die Aussagekraft dieser Ergebnisse wird zum Abschluss des Beitrags diskutiert. Bemerkenswert ist diesbezüglich vor allem, dass ein – in einer separaten Studie (Nisic/Auspurg 2009) angestellter – Vergleich eine sehr hohe Übereinstimmung zwischen den hier präsentierten fiktiven Umzugs-

[33] Neben diesem inhaltlichen Beitrag zu einer Forschungslücke in der Familien- und Ungleichheitssoziologie kann die Studie in Kapitel 3 auch als Beispiel dafür gelesen werden, dass es Mittelwege in der analytischen Soziologie gibt zwischen den beiden bislang in einer Art „Lagermentalität“ verteidigten Strategien einer direkten versus indirekten Prüfung von Handlungstheorien (dazu Braun 1998: 157ff.; Brüderl 2004). Bei der direkten Variante werden Präferenzen, Wahrscheinlichkeiten und Kosten von Handlungsalternativen von den Akteuren erfragt und anschließend getestet, ob sie – eingesetzt in das Modell der Handlungstheorie – zu den beobachteten Verteilungen führen (s. z. B. R. Becker/Hecken 2007 für ein bildungssoziologisches Beispiel). Hieran wird zu Recht kritisiert, dass den Akteuren ihre Entscheidungsmotive nicht bewusst sein müssen (Braun/Franzen 1995; Hedström/Bearman 2009a) oder sie diese – beispielsweise aufgrund sozialer Erwünschtheit – nicht unbedingt äußern wollen (Esser 1990). Hinzu kommt in Befragungen das Problem von *post-hoc*-Rationalisierungen (Brüderl/Preisendörfer 1995: 72). Viele Strömungen von Rational-Choice-Theorien gehen ohnehin davon aus, dass die Akteure sich lediglich verhalten, „als ob“ sie ein explizites Nutzenkalkül durchführen würden (Braun 1998: 157). Die Verfechter der indirekten Variante versuchen daher erst gar nicht, die Entscheidungslogiken und Präferenzen von Akteuren direkt zu erfragen, sondern prüfen stattdessen Hypothesen lediglich indirekt über Modellimplikationen (Testung der mit unterschiedlichen Rahmenbedingungen einhergehenden Handlungsmuster; exemplarisch Hillmert/Jacob 2003 für eine bildungssoziologische Studie). Ein berechtigter Kritikpunkt an dieser Verfahrensweise ist, dass sich mit dieser eben nur indirekten Prüfstrategie alternative Erklärungen schlechter ausschließen lassen. Zudem sei bei einer Nichtbestätigung der Theorie unklar, inwieweit die unterstellten Mechanismen falsch sind oder lediglich die Brückenhypothesen (Brüderl 2004). Faktorielle Surveys bieten eine Art Mittelweg zwischen beiden Varianten: Sie ermöglichen es dem Forscher, Entscheidungsfaktoren direkt zu testen, ohne aber den Akteuren deren Evaluation im Einzelnen abzuverlangen. Erhoben werden anhand von Fallbeispielen stimulierte Entscheidungen, aus denen sich dann indirekt aber gleichwohl sehr unmittelbar auf die (den Akteuren möglicherweise unbewussten) Entscheidungsregeln schließen lässt.

bereitschaften und den anhand von Daten des Sozio-oekonomischen Panels (SOEP) beobachteten tatsächlichen Umzügen von Doppelverdiener-Paaren hinsichtlich der grundsätzlichen Einflussfaktoren (wie z. B. bestehendem Wohneigentum und vorliegenden Erwerbskonstellationen) ergibt. Dies spricht für die Validität der Ergebnisse des Faktoriellen Surveys. Gleichwohl sind weitere Methodenstudien angezeigt. Ein Beitrag hierzu schließt sich an (Abschnitt 1.3.3 bzw. Kapitel 4), zusätzliche Anregungen werden im Fazit des vorliegenden Kapitels gegeben (Abschnitt 1.4).

1.3.3 Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey

Wie eingangs (Abschnitt 1.1) angedeutet, sind nicht nur manifeste Verteilungen, sondern ebenso Einstellungen zur Verteilungs*gerechtigkeit* (etwa Gerechtigkeit verschiedener Einkommenshöhen für Berufsgruppen oder Geschlechter) für die Soziologie sozialer Ungleichheit relevant (s. Liebig et al. 2004 sowie die Beiträge in P. Berger/Schmidt 2004 für ausführliche Erläuterungen). Speziell interessiert, inwieweit das Prinzip der Meritokratie im normativen Selbstverständnis der Bevölkerung verankert ist, denn davon abweichende Einstellungen sind nach allen gängigen Theorien eine notwendige (wenngleich nicht hinreichende) Bedingung für das Auftreten von Diskriminierungen als einem zentralen Mechanismus für die Erzeugung sozialer Ungleichheit.[34] Daneben bieten Gerechtigkeitsvorstellungen Anknüpfungspunkte an Mechanismen sozialer Schließung, welche Bestandteil von Klassentheorien sind und ebenso für sich genommen einen wichtigen Erklärungsbaustein für die Genese sozialer Ungleichheit bilden. Der grundlegende Gedanke ist hier, dass das Eintreten für höhere Entlohnungen speziell der *eigenen* (Beschäftigungs- oder Berufs-)Gruppe – und sei dies nur auf der Einstellungsebene – ein erstes empirisches Indiz für die Abschottung von Gruppen mit Versuchen der Mehrwertaneignung darstellt. Beide Aspekte dürften allerdings bei einer direkten Abfrage deutlichen Verzerrungen durch soziale Erwünschtheit unterliegen. Diskriminierungen sind allein schon durch das Grundgesetz verboten; und eine Bevorteilung der eigenen Gruppe läuft den Geboten von Bescheidenheit und Altruismus zuwider.

Die Einstellungsforschung ist damit zugleich ein ideales Gebiet, um den Gewinn durch innovative Verfahren, inklusive der für ihre Entwicklung unerlässlichen Methodenforschung, zu veranschaulichen. Die Erhebung von Normen und Einstellungen fällt in das klassische Einsatzgebiet

[34] Bei *präferenzbasierter* Diskriminierung bilden den Erklärungskern nicht weiter begründete Vorurteile gegenüber Mitgliedern der diskriminierten Gruppe *(tastes for discrimination*). Diskriminierende Arbeitgeber beschäftigen aufgrund dieser Abneigung Angehörige der betroffenen Gruppe nur zu geringeren Löhnen. Der Lohnunterschied spiegelt dabei das Ausmaß ihrer Abneigung wider (G. S. Becker 1971). Bei der *statistischen* Diskriminierung (z. B. Arrow 1973) besteht der Mechanismus dagegen in der Unterstellung einer geringeren Leistungsfähigkeit aufgrund unbeobachteter Merkmale, bei welchen die diskriminierte Gruppe im Mittel (statistisch) schlechter abschneidet (etwa unterdurchschnittliche Betriebsbindung von Frauen aufgrund ihrer höheren Wahrscheinlichkeit, Erwerbstätigkeiten familienbedingt zu unterbrechen). Unabhängig vom Erklärungsansatz bilden Gerechtigkeitseinstellungen auch deshalb einen wichtigen Bedingungsfaktor für das langfristige Bestehen von Diskriminierungen, weil diskriminierende Akteure sich umso besser auf dem (Arbeits-)Markt halten können, je weniger ihr Verhalten missbilligt und entsprechend durch einen Boykott seitens der Kunden oder Arbeitskräfte sanktioniert wird (G. S. Becker 1971).

Faktorieller Surveys (Wallander 2009 für einen Überblick). Der generelle Wert dieser Methodik für die analytische Soziologie wurde bereits im vorherigen Abschnitt veranschaulicht; als „Synthese“ aus Experiment und Umfrage wird sie den genannten Anforderungen an empirische Studien (vgl. Abschnitt 1.2.2) besonders gut gerecht. Für die Erhebung von solch sensiblen Aspekten wie den soeben erwähnten Diskriminierungen kommt der spezielle Vorteil hinzu, dass das indirekte Erhebungsverfahren mittels Vignetten als weitaus weniger anfällig für Effekte sozialer Erwünschtheit gilt als alternative direkte Itemabfragen (Alexander/Becker 1978; Jann 2003; Jasso 1996: 282f.). Zudem sind Faktorielle Surveys für die Erhebung gruppenspezifischer Urteilsregeln und Einstellungen prädestiniert, wie sie etwa von Klassentheorien angenommen werden.

Als innovatives Verfahren ist die Vignettentechnik selbst allerdings wenig erforscht. Die Ungleichheitssoziologie greift vorwiegend auf bestehende Instrumente zurück oder verfolgt die Weiterentwicklung empirischer Methoden primär mit dem Fokus auf exakte *Beschreibungen*. Vertreter analytischer Soziologie wiederum fassen die gezielte Weiterentwicklung empirischer Methoden nur am Rande als ihr Aufgabengebiet auf (Brückner 2009; Goldthorpe 1996; Manzo 2007 für Ausnahmen).[35] Dies überrascht, setzt das Prinzip einer gegenseitigen Stimulierung von Theorie und Empirie doch ebenso eine Fortentwicklung des empirischen Werkzeugs voraus. Um dieser Aufgabe mehr Nachdruck zu verleihen, schließt die vorliegende Dissertation bewusst mit einem vorrangig der Methodenforschung gewidmeten Kapitel. Es handelt sich um einen Ausschnitt aus einem umfassenderen Forschungsprojekt, in welchem der inhaltliche Schwerpunkt auf dem Thema der Einkommensgerechtigkeit liegt und der methodische Fokus auf den Auswirkungen verschiedener Designaspekte von Faktoriellen Surveys. Erste inhaltliche Befunde können andernorts nachgelesen werden (z. B. Sauer et al. 2009), Kapitel 4 zeigt beispielhaft eine Studie aus dem Bereich der geleisteten Methodenforschung (s. Auspurg et al. i. E.b, 2009d für weitere einschlägige Beiträge).

Die Motivation für eine bessere methodische Erforschung Faktorieller Surveys ergibt sich mitunter aus gängigen Zweifeln an der Validität der mit ihnen erhobenen Daten. Befürchtet wird, dass die Antworten der Befragten neben den eigentlich interessierenden Mechanismen ebenso methodische Artefakte, wie beispielsweise kognitive Überforderungen, Verärgerungen über unplausible Fallbeispiele oder *Demand*-Effekte (Versuche der Befragten, den Erwartungen der Forschenden möglichst gut gerecht zu werden) widerspiegeln (s. z. B. Faia 1980; Eifler 2007). Drohende Fehlschlüsse lassen sich anschaulicher an Beispielen erläutern: Unterschiede in den Antwortmustern von jüngeren und älteren Befragten könnten statt *inhaltlich* verschiedenen Einstellungen und Entscheidungslogiken ebenso differente Fähigkeiten reflektieren, die in den Fallbeispielen vorgegebene Informationsmenge kognitiv zu verarbeiten. Ähnlich ist es denkbar, dass unerwartete Antwortmuster kein Versagen des unterstellten Mechanismus anzeigen, sondern stattdessen eine Folge dessen sind, dass die Fallbeispiele

[35] Die Ausblendung methodischer Aspekte zeigt sich etwa daran, dass von Rational-Choice-Theoretikern selten darauf hingewiesen wird, dass das Argument, reichhaltige Schlussfolgerungen würden Abstraktionen von den empirischen Regularien unvollständiger Information und unvollkommener Bearbeitungskapazitäten rechtfertigen (z. B. Braun 2008: 391; Braun/Franzen 1995: 246), keineswegs auf Datenerhebungen anzuwenden ist.

oder Urteilsaufgaben missverstanden wurden. Das bislang durch Faktorielle Surveys ermittelte, sehr unterschiedliche Ausmaß der Einkommensdiskriminierung von Frauen (Auspurg et al. 2009d; Jann 2003; Jasso/Webster 1997, 1999) ist also womöglich gar nicht durch Einstellungsunterschiede zwischen den befragten Gruppen (Studierende vs. Allgemeinbevölkerung; verschiedene Geburtskohorten und Länder) bedingt, sondern (auch) eine Folge der differenten methodischen Anlage der Studien.[36]

Die von Kritikern hervorgehobene mangelnde *externe Validität* (fehlende Übereinstimmung der hypothetischen mit realen Entscheidungen oder gar Handlungen; s. Eifler 2007; Groß/ Börensen 2009 für Faktorielle Surveys) erscheint dabei für die in der analytischen Soziologie angestrebte Theorienprüfung gar nicht so zentral. Abweichungen zwischen hypothetischen und realen Verhaltensweisen sind allein schon aufgrund der in realen gegenüber hypothetischen Entscheidungssituationen hinzutretenden Informationen zu erwarten. Sofern keine Wechselwirkungen mit außer Acht gelassenen Faktoren bestehen, sollten sich die interessierenden *grundsätzlichen* Entscheidungs*logiken* gleichwohl valide erheben lassen (Ajzen 1988, 1991; Nisic/Auspurg 2009 für Details). Es ist ja gerade die Stärke von Experimenten, einzelne Faktoren zu isolieren bzw. von real hinzutretenden Elementen zu abstrahieren. Ihr Zweck ist es nicht, das absolute Niveau von Verhaltensweisen vorherzusagen, sondern stattdessen Verhaltensänderungen infolge der Veränderung *einzelner* Modellparameter zu prüfen (Diekmann/Voss 2004: 19). Wenngleich auch hinsichtlich der externen Validität von Faktoriellen Surveys noch Forschungsbedarf besteht, erscheint die Testung ihrer *internen Validität* (Ausschluss alternativer Ergebnisinterpretationen) mindestens ebenso dringlich.

Auf eine solche Prüfung konzentriert sich die zusammen mit Thomas Hinz und Stefan Liebig verfasste Studie in Kapitel 4. Inhaltlich greift sie auf den „Klassiker" unter den Faktoriellen Surveys zurück: die Bewertung der Einkommensgerechtigkeit fiktiver Erwerbspersonen. Besonderheit der Studie ist ihr methodischer Fokus. Erstmalig wird in einem Faktoriellen Survey die Anzahl der variablen Dimensionen – und damit die von den Befragten zu verarbeitende Informationsmenge – experimentell variiert, um die Auswirkung der Dimensionszahl auf das Antwortverhalten zu prüfen. Als einen weiteren Aspekt der Komplexität adressiert die Studie die ebenfalls seit Langem angemahnte Forschungslücke (z. B. Rossi/Anderson 1982: 59f.) im Bereich von Lern- und Ermüdungseffekten mittels einer Untersuchung des Antwortverhaltens im Bearbeitungsverlauf. Ein weiteres Forschungsinteresse bildet schließlich die theoretisch stark umstrittene, aber empirisch bis dato unklare Wirkung von Fallbeispielen mit real selten bis gar nicht vorkommenden, „unplausiblen" Merkmalskombinationen.[37]

---

[36] Die Studien unterscheiden sich etwa hinsichtlich der verwendeten Anzahlen an variierten Dimensionen und Vignetten sowie Antwortskalen, damit allesamt methodischen Aspekten, deren Wirkung bislang empirisch völlig ungeklärt ist.

[37] Von einigen Anwendern werden unplausible Fälle als besondere Stärke des Verfahrens gelobt, weil sie den Bewertungshorizont über das real gegebene erweitern würden (z. B. Alves/Rossi 1978: Rossi/Alves 1980). Kritiker wenden dagegen ein, dass die durch sie hervorgerufenen Zweifel seitens der Befragten an der Ernsthaftigkeit der experimentellen Stimuli zu teuer mit flüchtigen bis „unsinnigen" Antworten und somit

Im Einzelnen werden dazu in Kapitel 4 nach einem Überblick über den Forschungsstand Hypothesen zu den genannten methodischen Aspekten aus kognitionspsychologischen Theorien sowie aus vorliegenden Erkenntnissen zu verwandten Methoden (Conjoint-Analysen und Choice-Experimente) abgeleitet. Eine besondere Leistung des Aufsatzes ist dann die Erarbeitung eines geeigneten Designs für deren Überprüfung. Nur mit einer bewussten Zusammenstellung der Vignettenstichprobe (fraktionalisiertes Design), Zufallsreihenfolge pro Befragten und einer homogenen Probandengruppe lassen sich inhaltliche Effekte sauber von methodischen trennen. Diese zumindest in soziologischen Anwendungen bislang vernachlässigten Aspekte (Dülmer 2007; Steiner/Atzmüller 2006 für Ausnahmen) werden näher erläutert und anschließend in Form einer eigens konzipierten Online-Befragung mit gut 400 teilnehmenden Studierenden umgesetzt. Im Ergebnis zeigt sich, dass das methodische Design für das Antwortverhalten und damit die Validität Faktorieller Surveys durchaus bedeutsam ist. Etwa erweist sich die Einflussstärke von Dimensionen als abhängig von der Informationsmenge in den Vignetten und deren Plausibilität. Lern- und Ermüdungseffekte sind hingegen bei der verwendeten Anzahl von maximal zehn Fallbeispielen pro Befragten kaum zu beobachten. Das Kapitel veranschaulicht damit insgesamt, wie unabdingbar es für die Vermeidung von Fehlinterpretationen ist, die Wirkung von (experimentellen) Stimuli zu hinterfragen und im Idealfall durch Methodenexperimente wie durch das vorliegende auszuloten. Inhaltlich ist der interessanteste Befund, dass sich mit dem hier verwendeten studentischen Sample keinerlei Evidenzen für eine ungleiche Bewertung von Einkommenshöhen männlicher und weiblicher Erwerbstätiger ergeben. Eine in der Anlage sehr ähnliche Bevölkerungsumfrage findet hingegen heraus, dass die Befragten eine geringere Entlohnung von Frauen gegenüber vergleichbar qualifizierten und leistungsfähigen Männern für gerecht halten – und somit zumindest auf der Einstellungsebene eine Einkommsdiskriminierung von Frauen unterstützen (Auspurg et al. 2009d). An dieser Stelle muss offen bleiben, inwieweit dieser Unterschied in den Ergebnissen einen Alters-, Kohorten- oder methodischen Effekt (unterschiedliche kognitive Leistungsfähigkeit der Probanden bzw. Anfälligkeit für soziale Erwünschtheit) repräsentiert, was nochmals eindrücklich die Notwendigkeit gezielter Methodenforschung unterstreicht.

1.3.4 Bedeutung der Publikationsorgane und eigene Leistung

Alle drei Beiträge sind inzwischen in *peer-reviewed* Zeitschriften erschienen oder zur Publikation angenommen. Tabelle 1.1 zeigt die Publikationsorgane mit ihren *impact*-Faktoren im Jahr 2008. Dass das hier vertretene Konzept einer analytischen Ungleichheitssoziologie – trotz Annahme der Beiträge in herausgehobenen Zeitschriften – noch nicht auf allgemeine Zustimmung stößt und somit die hier stehenden Ausführungen durchaus ihre Relevanz haben, zeigen besonders deutlich die kritischen Diskussionen, die der erste Aufsatz zur „akademischen Elite“ inzwischen hervorgerufen hat (Jansen et

invaliden Ergebnissen bezahlt würden (Faia 1980; Rashotte et al. 2005: 177f.). Empirische Prüfungen stehen zumindest für Faktorielle Surveys noch aus.

al. 2009; Münch 2009; für eine Replik: Auspurg et al. 2009b).[38] Speziell die Einwände von Richard Münch richten sich pauschal gegen viele der in den Abschnitten 1.2.1 und 1.2.2 genannten Prämissen, ohne dass er jedoch die Gültigkeit zentraler Ergebnisse anzweifeln würde.

Tabelle 1.1: Erscheinungsorte im Überblick

Kapitel	Literaturangabe	*Impact*-Faktor im Jahr 2008[a]	5-Jahres-*Impact*-Faktor[a]
2	Auspurg, Katrin/Hinz, Thomas/Güdler, Jürgen (2008): Herausbildung einer akademischen Elite? Zum Einfluss der Größe und Reputation von Universitäten auf Forschungsförderung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie 60*, 653-685.	1,188	0,867
3	Abraham, Martin/Auspurg, Katrin/Hinz, Thomas: Migration Decisions Within Dual-Earner Partnerships: A Test of Bargaining Theory. Erscheint in: *Journal of Marriage and Family*.	1,639	2,848
4	Auspurg, Katrin/Hinz, Thomas/Liebig, Stefan (2009c): Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey. *Methoden, Daten, Analysen 3*, 59-96. (Zugleich in englischer Fassung: Complexity, Learning Effects, and Plausibility of Vignettes in Factorial Surveys. *Paper presented at the ASA-Conference 2009*).	-[b]	-[b]

[a] Quelle: Thomson Reuters (2010).
[b] Aufgrund des jungen Gründungsdatums der Zeitschrift (2007) lassen sich diese Werte, die mit Daten zu den mindestens zwei vorhergehenden Jahrgängen berechnet werden müssten, derzeit noch nicht angeben.

Der Disput kann im Detail in den genannten Quellen nachgelesen werden und soll daher im Folgenden lediglich fokussiert auf Aspekte erklärender Soziologie zusammengefasst werden. Der zentrale Punkt ist sicherlich die Auswahl der adäquaten Analyseebene. Nach Münch (2009) sind Institutionen, nicht Personen, die zu analysierenden Einheiten. Die in unserem Beitrag vorgenommene Ansiedelung von Erklärungen auf der Handlungsebene individueller Akteure, wie sie dem von der analytischen Soziologie propagierten methodologischen Individualismus entspricht, wird ausdrücklich kritisiert.[39] Der von Münch konstatierte *individualistische Fehlschluss* (in Form einer Verschiebung von Hypothesen von der institutionellen Ebene der Universitäten auf die Ebene individueller Entscheidungen)

[38] Gerade die in dieser Studie praktizierte Umsetzung von Prinzipien analytischer Soziologie hat inzwischen allerdings eine offizielle Würdigung erfahren: In der Begründung für die Verleihung des 2. Preises der Thyssen-Stiftung für die besten sozialwissenschaftlichen Aufsätze des Jahres 2008 (an dieser Stelle sei der Stiftung dafür herzlich gedankt) wird von der Jury insbesondere die eingehend begründete Kritik an den Argumenten Münchs und die gute Verständlichkeit bei gleichzeitig methodisch hohem Niveau gelobt (s. Herausgeber der Kölner Zeitschrift für Soziologie und Sozialpsychologie 2009).
[39] Etwa in Passagen wie der folgenden: „Die entsprechende Strategie ist die Disaggregation der Daten, mit der man den Blick von den institutionellen Positionskämpfen ablenkt. Man sieht dann nicht mehr, was auf dieser Ebene passiert“ (Münch 2009: 455).

ist schlicht unzutreffend.[40] Aus Sicht einer analytischen Soziologie sind jedoch vor allem die deutlichen Vorbehalte gegenüber einer die Makroebene verlassenden Analyse bedenklich, unterbindet doch gerade der Verbleib auf dieser Ebene ein tiefergehendes Verständnis ursächlicher Prozesse (Hedström 2005) und lassen im Gegenteil reine Aggregatanalysen Fehlschlüsse erwarten. Weitere Differenzen werden hinsichtlich der Reichweite vorzuschlagender Theorien und der Rolle von Empirie im Forschungsprozess offensichtlich. Münch spricht sich für eine groß angelegte und „auf das gesamte System bezogene Gesellschaftsanalyse" aus, die zumindest anfänglich nur in Form einer „vorläufigen und unfertigen Analyse" anhand von „breit gefächerten Daten" vorzunehmen sei (2009: 454f.). Damit wird Position für umfassende Gesellschaftstheorien bezogen (statt für Erklärungen mittlerer Reichweite) und sich gegen die Dringlichkeit ausgesprochen, jegliche theoretischen Annahmen empirisch abzusichern. Aus Sicht einer analytischen Soziologie verwundert zudem, dass gerade die kompetente Anwendung einschlägiger empirischer Verfahren kritisiert wird,[41] stellt doch diese im erklärenden Paradigma gerade eine unabdingbare Prämisse der Erkenntnisgewinnung dar. In einer Hinsicht scheint es auf den ersten Blick Konsens zu geben – *in puncto* der Notwendigkeit, die zentralen Mechanismen sichtbar zu machen. Ist bei Münch die diesbezügliche Methode der Wahl die „idealtypische Zuspitzung und Überzeichnung" (2009: 455), wird anders in der analytischen Soziologie die nüchterne Abstraktion von nebensächlichen Details empfohlen.[42] Wie diese (Gespenster-)Debatte mit Richard Münch exemplarisch zeigt, ist das Forschungsprogramm einer erklärenden Soziologie also längst noch nicht überall akzeptiert.

Bei allen in den Kapiteln 2 bis 4 wiedergegebenen Aufsätzen handelt es sich um Gemeinschaftsarbeiten, für welche das Ausmaß meiner eigenen Leistungen klarzustellen ist. Für den „Zündfunken" zu Aufsatz 1 (Kapitel 2) ist wohl vor allem Richard Münch für die unbewusst provozierende Vorlage zu danken, die Idee der Replikation entstand gemeinsam mit Thomas Hinz. Da es sich bei den letzten beiden Aufsätzen (Kapitel 3 und 4) um Arbeiten handelt, die auf längerfristigen Kooperationen und thematisch engen Forschungszusammenhängen aufbauen (s. dazu z. B. Auspurg/Abraham 2007 sowie Hinz/Liebig 2006), lassen sich die Ideen und Theorieteile bei diesen beiden Studien nicht exakt einem Autor zuordnen, sondern sind diese Abschnitte eher als Gemeinschaftsprodukte fruchtbarer Zusammenarbeit zu verstehen.[43] Die Abschnitte zur Referierung des Forschungsstandes und die Ableitung von Hypothesen gehen im ersten Aufsatz auf mich zurück, im dritten überwiegend, und beim zweiten Aufsatz sind diese Inhalte wiederum aufgrund der engen Kooperation nicht klar zu attribuieren (der

[40] Ein individualistischer Fehlschluss wäre die Übertragung eines Zusammenhangs, der sich für Individualdaten (Wissenschaftler) findet, auf die Ebene von Aggregaten (Universitäten). Eine solche Übertragung findet nicht statt.

[41] Etwa in folgender Passage: „Schachzug 1: Verschiebung des Untersuchungsgegenstandes in den eigenen Kompetenzbereich […]. Dabei gewinnt man die Oberhoheit über die anzulegenden Kriterien und macht den Disputationsgegner zunächst einmal mit methodischem Tadel mürbe" (Münch 2009: 454).

[42] Aus gutem Grund, denn bei zu starker (polemischer) Übertreibung steht die Ernsthaftigkeit von Argumenten auf dem Spiel. Eine Gefahr, die ähnlich von Hornbostel (2009: 17ff.) sowie Mayer (2009: 8) gesehen wird.

[43] Dies äußert sich bei dem in Kapitel 3 beschriebenen Projekt auch in einer Rotation der Autorenreihenfolge (s. z. B. das inhaltlich verwandte Buchkapitel von Auspurg et al. 2009a).

Theorieteil wurde hier vorwiegend von Martin Abraham verfasst, allerdings in Anlehnung an frühere Gemeinschaftsarbeiten). In allen drei Aufsätzen wurden die Auswertungen, damit der Ergebnisteil, komplett von mir geleistet und die Erhebungen gänzlich oder zu sehr wesentlichen Teilen durch mich konzipiert und organisiert. Das Fazit zu den letzten beiden Studien entstammt „meiner Feder", im ersten Beitrag wurde es dagegen von Thomas Hinz entworfen.

Knapp zusammengefasst sind also die Ausarbeitungen der Argumentationslinien (auch was die Überarbeitungen nach den Rückmeldungen der Gutachtenden betrifft) sowie die Organisation und Konzeption der Erhebungen zu wesentlichen Teilen mir zuzurechnen und die Datenaufbereitungen und -auswertungen gänzlich. Durchweg alle Inhalte haben dabei stets von den kritischen Rückmeldungen sowie *Supervisionen* meiner Koautoren profitiert, denen ich an dieser Stelle dafür herzlich danken möchte.[44]

1.4 Fazit und Ausblick

Die einleitende These lautete, dass eine stärkere Orientierung an den Prämissen *erklärender* Soziologie den oftmals beklagten Erkenntnismangel in der Ungleichheitssoziologie überwinden helfen könnte. Anhand eines Überblicks über die Prämissen analytischer Soziologie und den Forschungsstand der Ungleichheitssoziologie wurde dann deutlich, dass der Anspruch, *präzise, handlungsbasierte, abstrakte* und *empirisch prüfbare Mechanismen* vorzuschlagen und diese mittels möglichst adäquater empirischer Methoden zu testen (*direkte Modellierungen der angenommenen Prozesse, im Idealfall anhand von experimentellen oder Längsschnittdaten*) in der Ungleichheitssoziologie noch wenig eingelöst ist. Die vorfindbaren Konzepte schwanken mehrheitlich zwischen den beiden Extremen empiriearmer Großtheorien oder theorieferner Empirie, anstatt den Mittelweg von Erklärungen mittlerer Reichweite zu gehen.

Anliegen der vorliegenden Dissertation ist es, in Reaktion hierauf, anhand von eigenen empirischen Anwendungen den *Analysegewinn erklärender* Ungleichheitssoziologie zu illustrieren und zugleich die Notwendigkeit einer Weiterentwicklung korrespondierender empirischer Verfahren aufzuzeigen. Jeder einzelne der in Kapitel 2 bis 4 folgenden Aufsätze zeigt, dass die Orientierung an den Prinzipien erklärender Soziologie nicht nur genaue Einsichten in Verteilungsmuster (von Forschungsmitteln, beruflichen Karrieren oder Gerechtigkeitseinstellungen) bietet, sondern daneben zu substanziellen Erkenntnisfortschritten im Hinblick auf deren ursächliche Mechanismen verhilft (kumulative Ressourcenanreicherungen versus Selektionseffekte; Verhandlungsmacht versus geschlechtsspezifische Rollenerwartungen; Gerechtigkeitseinstellungen versus Anwendungen von Heuristiken bzw. Vermeidungen kognitiver Überforderungen). Gerade die sehr unterschiedlichen Inhalte und Methoden

[44] Auch dies könnte man als einen weiteren Grundsatz analytischer Soziologie formulieren: statt dem Leitbild der „Einzelforschung einsamer Wölfe" (Münch 2007: 231) sollte das Ideal besser in einer Präzisierung von Theorien und Anreicherung des Gehalts empirischer Prüfungen durch kritikoffene Forschungskooperationen gesehen werden.

sollten zusammen mit einem der Methodenforschung gewidmeten Kapitel das Prinzip einer theoriegeleiteten Spezifikation von Hypothesen sowie anschließenden möglichst direkten empirischen Prüfung verdeutlichen.

Der Anspruch, Erklärungen kontinuierlich zu justieren, bedeutet allerdings auch für die hier präsentierten Studien, dass das Ende der Untersuchungen zugleich den Beginn daran anknüpfender Forschungsarbeiten markiert. Alle Beiträge weisen Limitationen auf, lassen Fragen offen und generieren neue. Detaillierte Hinweise dazu finden sich jeweils am Ende der einzelnen Kapitel, abschließend sollen hier einige ergänzende Anregungen gegeben sowie aufgezeigt werden, welche Fortschritte seit dem Erscheinen der Publikationen bereits erzielt werden konnten.

Was die in Kapitel 2 analysierte Allokation von Forschungsmitteln betrifft, besteht das überraschende Hauptergebnis in der lediglich schwachen Evidenz für einen kumulativen, sich selbst verstärkenden Mechanismus bei der Einwerbung von Forschungsressourcen auf der *institutionellen* Ebene von Universitäten. Dies schließt nicht aus, dass der Mechanismus gleichwohl für *individuelle* Karrieren von Wissenschaftlern greift (DiPrete/Eirich 2006 für einen Überblick). Aufgrund möglicher gegenläufiger Effekte, wie den häufigen Universitätswechseln von Wissenschaftlern, sind auf der Individualebene wirkende Mechanismen nicht unbedingt mit solchen auf der Makroebene gleichzusetzen – erneut ist hier an die Wichtigkeit der richtigen Analyseebene zu erinnern. Um zu prüfen, ob der Matthäus-Effekt bei *individuellen* Karrieren in der Wissenschaft wirkt (wie dies bereits der klassische Aufsatz von Merton 1968 vorrangig diskutiert), sind Längsschnittanalysen zum individuellen Karriereverlauf von Wissenschaftlerinnen und Wissenschaftlern geplant (zunächst umgesetzt mittels einer Modellierung der „Antragshistorie“ von DFG-Drittmitteleinwerbungen).[45] Sollten sich hierbei Kumulationen der Antragserfolge finden, wäre die besonders interessante Anschlussfrage, inwieweit diese Erfolge durch individuelle Leistungen gerechtfertigt sind (leistungsfähigere Wissenschaftlerinnen und Wissenschaftler werben mehr Drittmittel ein als andere, was ihr weiteres Leistungsvermögen überproportional steigert, usw.) oder ob sie Verletzungen der Meritokratie repräsentieren (etwa aufgrund von statistischen Diskriminierungen). Um hierauf Antworten zu finden, werden im Idealfall direktere Indikatoren der Leistungsfähigkeit der Antragstellenden verwendet als dies in Kapitel 2 möglich war (wie z. B. Messungen ihres Publikationsoutputs). Zudem empfiehlt es sich, den mehrstufigen Entscheidungsprozess über die Bewilligung versus Ablehnung von Forschungsanträgen genauer in seinem Ablauf und in der Interaktionsstruktur der beteiligten Akteure zu untersuchen. So wäre es aufschlussreich, Parameter für die inhaltliche Nähe (Forschungsgebiete) und formale Distanz (Statusunterschiede) von Gutachtenden und Antragstellenden in die Analysen einfließen zu lassen, um Hinweise für mögliche soziale Schließungen oder Konkurrenzmechanismen zu

[45] Bei den Analysen in Kapitel 2 standen diesem Forschungsziel viele links- und rechtszensierte Beobachtungen entgegen. Anhand der verwendeten Daten war nicht eruierbar, welche Antragsbiografie die Antragstellenden vor und nach dem 14-jährigen Beobachtungsfenster aufweisen. Dieses Problem wäre zwischenzeitlich durch einen längeren Beobachtungszeitraum und grundsätzlich durch einschlägige Verfahren der Ereignisdatenanalyse abzuschwächen.

gewinnen. Weiter könnte die Kritik (Jansen et al. 2009; Münch 2009) an der – in der vorliegenden Studie allerdings bewusst vorgenommenen – Beschränkung auf die DFG-Einzelförderung konstruktiv aufgegriffen werden. Mittels einer Erweiterung der Analyse um Forschungsanträge in der Verbundforschung ließe sich der weit verbreiteten Vermutung nachgehen, dass speziell die Programme der Exzellenzinitiative Reputationseffekte hervorrufen (s. z. B. C. Gross/Jungbauer-Gans 2007; Münch 2007; Wagner 2007).[46]

Das Hauptergebnis der in Kapitel 3 beschriebenen Untersuchung von Karriereentscheidungen in Doppelverdiener-Partnerschaften ist, dass sich die Verhandlungstheorie für die Erklärung von Karriereentscheidungen im Haushaltskontext bewährt. Daneben finden sich schwache Hinweise für die Relevanz geschlechtsspezifischer Ansätze. Inhaltlich könnte an diese Theorienprüfung angeknüpft werden, indem sie mit noch homogeneren Samples fortgesetzt wird – etwa mit gleichgeschlechtlichen Paaren, um so gänzlich von der Wirkung *geschlechtsspezifischer* Normen und Erwerbsoptionen abstrahieren zu können (Schürmann 2005 für ein Anwendungsbeispiel zur Hausarbeitsteilung). In praktischer Hinsicht interessiert vor allem, inwiefern sich die „privaten" Abstimmungsprozesse durch institutionelle Maßnahmen beeinflussen lassen. Dies ist speziell in professionellen Berufsfeldern von Belang, in denen die Eingebundenheit in Partnerschaften nach wie vor als einer der wesentlichen Faktoren für das Scheitern weiblicher Karrieren und damit für die Unterrepräsentanz von Frauen in Führungspositionen gilt, ohne dass die zugrunde liegenden Mechanismen hinlänglich entschlüsselt wären. Diskutiert werden in der Literatur neben den in Kapitel 3 betrachteten innerpartnerschaftlichen Ressourcen- und Machtungleichheiten ebenso direkte und indirekte (statistische) Diskriminierungen seitens der Arbeitgeber (Lind 2004; Rusconi 2002; Rusconi/Solga 2007, 2008). Um besseren Aufschluss über die Wirkung von institutionellen Rahmenbedingungen zu erlangen, wurde zwischenzeitlich eine zur Analyse in Kapitel 3 analoge Vignettenbefragung von Doppelkarriere-Paaren in der Wissenschaft initiiert (Auspurg et al. i. E.a für Details).[47] Eine wesentliche Neuerung ist dort, neben dem speziellen Befragtensample, dass in den Fallbeispielen Merkmale und Angebote von Hochschulen variiert werden, um deren Relevanz für die Aufrechterhaltung von Karrieren zu evaluieren.[48] Zudem wird das Projekt zu weiteren Erkenntnissen über *generelle* Mechanismen der partnerschaftsinternen Karriereabstimmung verhelfen.

---

[46] Leistung der analytischen Soziologie wäre es bei einer solchen Untersuchung insbesondere auch, nicht-intendierte Nebenfolgen aufzudecken. Etwa lassen sich verschiedene Argumente dafür anführen, warum die verstärkten Bemühungen um eine Erhöhung des Frauenanteils in den Programmen der Verbundforschung die Chancen von Frauen auf Drittmitteleinwerbungen in der Einzelförderung paradoxerweise herabsetzen (Auspurg/Hinz i. E.). Anti-Diskriminierungsmaßnahmen können das Ausmaß von Diskriminierung überaschenderweise befördern (s. dazu auch Kaas 2009).

[47] Projektleiter ist Thomas Hinz, Eva Amorelli bearbeitet das Projekt mit dem Titel „Räumliche und institutionelle Koordination von Doppelkarrieren" als wissenschaftliche Mitarbeiterin. Für nähere Informationen: http://www.soziologie.uni-konstanz.de/hinz/forschung/forschungsprojekte/dcc/.

[48] Etwa werden sich – die in Kapitel 4 untersuchte Validität von Faktoriellen Surveys vorausgesetzt – Hinweise dazu gewinnen lassen, inwieweit Jobangebote an mitziehende Partner, Kinderbetreuungsmaßnahmen an den Hochschulen oder gute Lehr- und Forschungsbedingungen die zielführenderen Maßnahmen sind, um die Aufrechterhaltung dualer Karrieren zu erleichtern, was wiederum insbesondere den Verbleib von Frauen in der

Gerade wenn praktische Empfehlungen angestrebt sind, ist jedoch Skepsis gegenüber Verfahren angebracht, die lediglich mit hypothetischen Entscheidungen arbeiten. Ohne Methodenforschung sind Fehlschlüsse nicht auszuschließen. Für die Untersuchung der internen Validität von Faktoriellen Surveys liefert Kapitel 4 ein erstes Beispiel. Zwischenzeitlich konnte von der Projektgruppe sehr umfangreiches Datenmaterial gesammelt werden, um weitere methodische Effekte zu untersuchen und praktische Handreichungen zu generieren (Auspurg et al. i. E.b; Hinz/Liebig 2006). Als einen generellen Impuls für derartige Methodenstudien empfiehlt sich aus der Warte einer analytischen Soziologie, diese noch stärker als bisher theoretisch zu motivieren – auch das Antwortverhalten von Befragten lässt sich mit Handlungs- und Spieltheorien modellieren, gegebenenfalls erweitert um kognitionspsychologische Theorien (s. exemplarisch Esser 1986, 1990; Krumpal/Voss 2009). Für die analytische Soziologie regen die Befunde in Kapitel 4 hingegen an, den Bereich der Methodenforschung stärker als ihr Gebiet aufzufassen. Das Anliegen, alternative Interpretationen weitestgehend auszuschließen, erfordert auch, die mit jeglichen Verfahren verbundenen Validitätsprobleme besser auszuloten. Statt in „Lagermentalität“ die Schwächen einzelner Verfahren zu überzeichnen (etwa Surveyforschung mit Variablensoziologie gleichzusetzen), dürfte die Zukunft in einer Triangulation verschiedener Methoden und in deren kontinuierlichen Weiterentwicklung liegen.[49] Einen sinnvollen Ansatz hierzu bieten *kontinuierlich* in Umfragen und Experimente integrierte methodische Splits (Brückner 2009). Die durch solche Methodenexperimente gewonnen Erkenntnisse lassen sich im Idealfall dann bereits für die Eindämmung von bestehenden Validitätsproblemen (und nicht erst für die Verbesserung künftiger Erhebungen) nutzen (West 2006: 127).[50]

Zur Prüfung der externen Validität von Faktoriellen Surveys sind vor allem Längsschnittstudien anzuraten. Die eben erwähnte Studie zu Doppelkarriere-Paaren ist auch in dieser Hinsicht wegweisend, da das Forschungsprojekt eine wiederholte Befragung *derselben* Probanden mit exakt *denselben* Vignetten nach einem Zeitraum von ein bis zwei Jahren vorsieht. Ein Abgleich der Antwortmuster pro Person wird Aufschluss über die Stabilität der erhobenen Entscheidungsregeln geben. Aufgrund der hohen Umzugs- und Karriere-Mobilität von Wissenschaftlern ist zudem davon auszugehen, dass zumindest ein Subsample der Paare zwischenzeitlich Umzugs- und Karriereentscheidungen fällt, die zu den abgefragten hypothetischen Situationen ähnlich sind. Dies wird es erstmalig erlauben, die externe Validität der anhand von Vignetten ermittelten Entscheidungsprinzipien auf der *individuellen* Ebene *derselben* Befragungspersonen zu untersuchen. Derzeit

Wissenschaft fördern würde. Zusätzlich zu den Fallbeispielen wird eine Varianz in institutionellen Rahmenbedingungen durch die Befragung von Wissenschaftlerinnen und Wissenschaftlern an unterschiedlichen deutschen und schweizerischen Hochschulen erreicht.

[49] Dies gilt insbesondere auch aufgrund der zwischen den einzelnen Formen von Validität unvermeidlich bestehenden *Trade-offs* (West 2006: 137).

[50] Vorbilder für ein derartiges Vorgehen finden sich im Bereich der mit Faktoriellen Surveys verwandten Choice-Experimente, in denen methodische Effekte nicht nur festgestellt, sondern durch gezielte statistische Modellierungen (heteroskedastische Logit-Modelle) in ihrem verzerrenden Einfluss abgemildert bis eliminiert werden (s. z. B. Caussade et al. 2005 und DeShazo/Fermo 2002 für derartige Anwendungsbeispiele).

gibt es noch keine Faktoriellen Surveys mit einem Paneldesign,[51] und vorliegende Prüfungen der externen Validität (Eifler 2007; Groß/Börensen 2009; Nisic/Auspurg 2009) leiden allesamt an derselben Limitation, dass das Außenkriterium (Verhaltensbeobachtung) sich auf eine andere Probandenstichprobe stützt als der zu evaluierende Faktorielle Survey.

Für eine noch gezieltere Überprüfung externer Validität ist zudem angedacht, eine Längsschnittstudie umzusetzen, bei der die in den Vignetten abgefragten Entscheidungen in ähnlicher Form real bei *allen* Befragten anstehen. Ein solches Anwendungsgebiet könnten Bildungsentscheidungen sein (z. B. Auswahlentscheidungen zwischen verschiedenen Studiengängen). Zugleich ist geplant, die mit den Faktoriellen Surveys eng verwandte Methodik der Choice-Experimente stärker zu erproben (im Idealfall mittels eines Befragungssplits zwischen Vignetten und analogen Choice-Sets). Dieses in den Sozialwissenschaften noch unbekanntere Verfahren erscheint deshalb besonders attraktiv, weil es gegenüber Faktoriellen Surveys noch direkter auf die Erhebung von Entscheidungen anstatt von Einstellungen zugeschnitten ist und zugleich durch seine explizite Herleitung aus der *Random Utility Theory* eine deutlich engere Korrespondenz zu (soziologischen) Handlungstheorien aufweist als die eher forschungspragmatisch motivierten Vignettenanalysen. In einer ersten familiensoziologischen Pilotstudie (Auspurg/Liebe 2010) haben sich diese Eigenschaften bereits sehr gut bewährt.

Abschließend soll noch eine generelle Anregung für die Weiterentwicklung analytischer Ungleichheitssoziologie ausgesprochen werden. Vorliegende Anwendungsbeispiele machen die Umsetzung ihrer Prinzipien in aller Regel nicht explizit kenntlich und werden vermutlich auch deswegen nicht als einschlägige Beiträge rezipiert. Um den gemeinsamen Kern von bewährten Erklärungen deutlicher herauszuarbeiten und zugleich eine effizientere Verzahnung der entdeckten Mechanismen zu erreichen, erscheint es ratsam, den Wissens- und Erkenntnisstand in einem Überblickswerk zu bündeln, im Idealfall sortiert nach einzelnen Mechanismen. Konkret könnte dies in einem Handbuch analytischer Ungleichheitssoziologie umgesetzt werden. Die vielen vorliegenden Analysebausteine dürften in ihrer Zusammenschau eine erhebliche Aufklärung Ungleichheit generierender Mechanismen leisten – und damit eine unerlässliche Voraussetzung dafür schaffen, dass wir die Verteilung begehrter Ressourcen mit unseren gesellschaftspolitischen Idealen in Einklang bringen.

[51] Einzige Ausnahmen sind meines Wissens zwei methodische Studien (Auspurg et al. i. E.b; Liebig et al. 2006), die allerdings auf die Stabilität von Einstellungsmessungen fokussieren und darum zwischenzeitliche Veränderungen in den Handlungskontexten der Befragten geradezu ausschließen wollen (etwa durch kurze Befragungsabstände). Im Bereich der Choice-Experimente gibt es bereits wesentlich mehr Methodenforschung zur externen Validität. Die diesbezüglichen Ergebnisse sind heterogen, zeigen insgesamt aber eine erstaunlich hohe Übereinstimmung zwischen den mit Choice-Sets erhobenen „virtuellen“ und anhand von Daten zu realem Verhalten ermittelten tatsächlichen Entscheidungsprinzipien (z. B. hinsichtlich der Wahl von Verkehrsmitteln oder der Zahlungsbereitschaft für Umweltgüter; s. Auspurg/Liebe 2010 für Literaturhinweise).

Literatur

Abraham, Martin/Auspurg, Katrin/Hinz, Thomas (i. E.): Migration Decisions Within Dual-Earner Partnerships: A Test of Bargaining Theory. Akzeptiert für: Journal of Marriage and Family.

Ajzen, Icek (1991): The Theory of Planned Behavior. Organizational Behavior and Human Decision Processes 50, 179-211.

Ajzen, Icek (1988): Attitudes, Personality and Behavior. Chicago: Dorsey Press.

Alexander, Cheryl S./Becker, Henry J. (1978): The Use of Vignettes in Survey Research. Public Opinion Quarterly 42, 93-104.

Alves, Wayne M./Rossi, Peter H. (1978): Who Should Get What? Fairness Judgments of the Distribution of Earnings. American Journal of Sociology 84, 541-564.

Arrow, Kenneth J. (1973): The Theory of Discrimination. In: Ashenfelter, Orley/Rees, Albert (Hrsg.): Discrimination in Labor Markets. Princeton: Princeton University Press, 193-216.

Auspurg, Katrin/Abraham, Martin (2007): Die Umzugsentscheidung von Paaren als Verhandlungsproblem. Eine quasiexperimentelle Überprüfung des Bargaining-Modells. Kölner Zeitschrift für Soziologie und Sozialpsychologie 59, 271-293.

Auspurg, Katrin/Abraham, Martin/Hinz, Thomas (2009a): Wenig Fälle, viele Informationen: Die Methodik des Faktoriellen Surveys als Paarbefragung. In: Kriwy, Peter/Gross, Christiane (Hrsg.): Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen. Wiesbaden: VS Verlag für Sozialwissenschaften, 179-210.

Auspurg, Katrin/Hinz, Thomas (i. E.): Antragsaktivitäten und Förderchancen von Wissenschaftlerinnen bei Einzelanträgen auf DFG-Einzelförderung 2005-2008. Working Paper der DFG. Bonn: Deutsche Forschungsgemeinschaft.

Auspurg, Katrin/Hinz, Thomas/Amorelli, Eva (i. E.a): Räumliche und institutionelle Koordination von Doppelkarrieren in der Wissenschaft – eine Forschungsskizze. Erscheint in: Funk, Julia/ Gramespacher, Elke/Rothäusler, Iris (Hrsg.): Dual Career Couples an Hochschulen. Zwischen Wissenschaft, Praxis und Theorie. Opladen: Barbara Budrich.

Auspurg, Katrin/Hinz, Thomas/Güdler, Jürgen (2009b): Gespensterdebatte oder was ist soziologische Aufklärung? Replik auf die Kommentare von Richard Münch sowie Dorothea Jansen, Richard Heidler und Regina von Goertz. Kölner Zeitschrift für Soziologie und Sozialpsychologie 61, 469-474.

Auspurg, Katrin/Hinz, Thomas/Güdler, Jürgen (2008): Herausbildung einer akademischen Elite? Zum Einfluss der Größe und Reputation von Universitäten auf Forschungsförderung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 60, 653-685.

Auspurg, Katrin/Liebe, Ulf (2010): Choice-Experimente in der Soziologie. Unveröffentlichtes Arbeitspapier. Konstanz/Leipzig: Universitäten Konstanz und Leipzig.

Auspurg, Katrin/Hinz, Thomas/Liebig, Stefan (2009c): Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey. Methoden, Daten, Analysen 3, 59-96.

Auspurg, Katrin/Hinz, Thomas/Liebig, Stefan/Sauer, Carsten (2009d): Auf das Design kommt es an. Experimentelle Befunde zu komplexen Settings in Faktoriellen Surveys. SoFid Methoden und Instrumente der Sozialwissenschaften 2009/2, 23-39.

Auspurg, Katrin/Hinz, Thomas/Liebig, Stefan/Sauer, Carsten (i. E.b): Wie unplausibel darf es sein? Der Einfluss von Designmerkmalen auf das Antwortverhalten in Faktoriellen Surveys. Erscheint in: Deutsche Gesellschaft für Soziologie (Hrsg.): Kongressband des 34. Kongresses der DGS. Wiesbaden: VS Verlag für Sozialwissenschaften.

Bayer, Michael (2008): Transnationale Ungleichheitsforschung: Eine neue Herausforderung für die Soziologie. Frankfurt a. M.: Campus.

Beck, Michael/Opp, Karl-Dieter (2001): Der faktorielle Survey und die Messung von Normen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 53, 283-306.

Beck, Ulrich (2008): Die Neuvermessung der Ungleichheit unter den Menschen. Eröffnungsvortrag zum Soziologentag „Unsichere Zeiten“ am 6. Oktober 2008 in Jena. Frankfurt a. M.: edition suhrkamp.

Beck, Ulrich (1986): Risikogesellschaft. Auf dem Weg in eine andere Moderne. Frankfurt a. M.: edition suhrkamp.

Beck, Ulrich (1983): Jenseits von Stand und Klasse? Soziale Ungleichheit, gesellschaftliche Individualisierungstendenzen und die Entstehung neuer sozialer Formationen und Identitäten. In: Kreckel, Reinhard (Hrsg.): Soziale Ungleichheiten. Soziale Welt, Sonderband 2. Göttingen: Otto Schwartz, 35-74.

Becker, Gary S. (1971): The Economics of Discrimination. Chicago: University of Chicago Press.

Becker, Rolf/Hecken, Anna Etta (2007): Studium oder Berufsausbildung? Eine empirische Überprüfung der Modelle zur Erklärung von Bildungsungsentscheidungen von Esser sowie Breen und Goldthorpe. Zeitschrift für Soziologie 36, 100-117.

Berger, Johannes (2004): Über den Ursprung der Ungleichheit unter den Menschen. Zeitschrift für Soziologie 33, 354-374.

Berger, Johannes (2003): Sind Märkte gerecht? Zeitschrift für Soziologie 32, 462-473.

Berger, Peter A. (2008): Transnationalisierung sozialer Ungleichheit. Wiesbaden: VS Verlag für Sozialwissenschaften.

Berger, Peter A. (2003): Kontinuitäten und Brüche. Herausforderungen für die Sozialstruktur- und Ungleichheitsforschung im 21. Jahrhundert. In: Orth, Barbara/Schwietring, Thomas/Weiß, Johannes (Hrsg.): Soziologische Forschung. Stand und Perspektiven. Opladen: Leske + Budrich, 473-490.

Berger, Peter A. (1987): Klassen und Klassifikationen. Zur „neuen Unübersichtlichkeit“ in der soziologischen Ungleichheitsdiskussion. Kölner Zeitschrift für Soziologie und Sozialpsychologie 39, 59-85.

Berger, Peter A./Hradil, Stefan (1990): Die Modernisierung sozialer Ungleichheit und die neuen Konturen ihrer Erforschung. In: Berger, Peter A./Hradil, Stefan (Hrsg.): Lebenslagen, Lebensläufe, Lebensstile. Sozialen Welt., Sonderband 7. Göttingen: Otto Schwartz, 3-25.

Berger, Peter A./Schmidt, Volker H. (2004): Welche Gleichheit, welche Ungleichheit? Grundlagen der Ungleichheitsforschung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Blalock, Hubert M. (1991): Understanding Social Inequality. Modeling Allocation Processes. Newbury Park u. a.: Sage.

Blank, Rebecca M. (1991): The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from The American Economic Review. The American Economic Review 81, 1041-1067.

Blinder, Alan S. (1973): Wage Discrimination: Reduced Form and Structural Estimates. The Journal of Human Resources 8, 436-455.

Blossfeld, Hans-Peter (1996): Macro-sociology, Rational Choice Theory, and Time. A Theoretical Perspective on the Empirical Analysis of Social Processes. European Sociological Review 12, 181-206.

Blossfeld, Hans-Peter/Müller, Rolf (1996): Sozialstrukturanalyse, Rational Choice Theorie und die Rolle der Zeit. Ein Versuch zur dynamischen Integration zweier Theorieperspektiven. Soziale Welt 47, 382-410.

Bolte, Martin (1983): Anmerkungen zu Aspekten und Problemen der Erforschung sozialer Ungleichheit. In: Kreckel, Reinhard (Hrsg.): Soziale Ungleichheiten. Soziale Welt, Sonderband 2. Göttingen: Otto Schwartz, 391-407.

Bornmann, Lutz/Mutz, Rüdiger/Daniel, Hans-Dieter (2007): Gender differences in grant peer review: A meta-analysis. Journal of Infometrics 1, 226-238.

Bourdieu, Pierre (1981): Titel und Stelle. Über die Reproduktion sozialer Macht. Frankfurt a. M.: Europäische Verlagsanstalt.

Braun, Norman (2009): Eine Antwort auf meine Kommentatoren. Soziale Welt 60, 223-232.

Braun, Norman (2008): Theorie in der Soziologie. Soziale Welt 59, 373-395.

Braun, Norman (1998): Der Rational-choice-Ansatz in der Soziologie. In: Pies, Ingo/Leschke, Martin (Hrsg.): Gary Beckers ökonomischer Imperialismus. Tübingen: Mohr Siebeck, 147-173.

Braun, Norman/Franzen, Axel (1995): Umweltverhalten und Rationalität. Kölner Zeitschrift für Soziologie und Sozialpsychologie 47, 231-248.

Breen, Richard/Rotman, David (1995): Class Analysis and Class Theory. Sociology 29, 453-473.

Brückner, Hannah (2009): Surveys. In: Hedström, Peter/Bearman, Peter (Hrsg.): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press, 666-687.

Brüderl, Josef (2004): Die Überprüfung von Rational-Choice-Modellen mit Umfragedaten. In: Diekmann, Andreas/Voss, Thomas (Hrsg.): Rational-Choice-Theorie in den Sozialwissenschaften. Anwendungen und Probleme. München: Oldenbourg, 163-182.

Brüderl, Josef/Preisendörfer, Peter (1995): Der Weg zum Arbeitsplatz: Eine empirische Untersuchung zur Verkehrsmittelwahl. In: Diekmann, Andreas/Franzen, Axel (Hrsg.): Kooperatives Umwelthandeln. Modelle, Erfahrungen, Massnahmen. Chur: Ruegger, 69-88.

Burzan, Nicole (2007): Soziale Ungleichheit. Eine Einführung in die zentralen Theorien. Wiesbaden: VS Verlag für Sozialwissenschaften.

Caussade, Sebastian/Ortúzar, Juan de Dios/Rizzi, Luis I./Hensher, David A. (2005): Assessing the influence of design dimensions on stated choice experiment estimates. Transportation Research Part B 39, 621-640.

Center of Excellence Women and Science (CEWS) (2009): Hochschulranking nach Gleichstellungsaspekten – Dritte Fortschreibung (cews.publik.no13). Bonn: Kompetenzzentrum Frauen in Wissenschaft und Forschung CEWS.

Coleman, James (1990): Foundations of Social Theory. Cambridge, MA: Harvard University Press.

Collins, Randall (1971): Functional and Conflict Theories of Educational Stratification. American Sociological Review 36, 1002-1019.

Dahrendorf, Ralf (1970): On the Origin of Inequality Among Men. In: Laumann, Edward/Siegel, Paul M./Hodge, Robert W. (Hrsg.): The Logic of Social Hierarchies. Chicago: Markham, 3-30.

Davis, Kingsley/Moore, Wilbert E. (1945): Some Principles of Stratification. American Sociological Review 10, 242-249.

DeShazo, J. R./Fermo, German (2002): Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. Journal of Environmental Economics and Management 44, 123-143.

Deutsche Forschungsgemeinschaft (DFG) (2009): Förder-Ranking 2009. Institutionen – Regionen – Netzwerke. DFG-Bewilligungen und weitere Basisdaten öffentlich geförderter Forschung. Bonn: Deutsche Forschungsgemeinschaft.

Deutsche Gesellschaft für Soziologie (DGS) (2010): Transnationale Vergesellschaftungen. Themenpapier zum 35. Kongress der Deutschen Gesellschaft für Soziologie vom 11. bis 15. Oktober 2010 in Frankfurt a. M.. (Verfügbar unter: http://dgs2010.de/wp-content/uploads/DGS2010_TransnationaleVergesellschaftungen.pdf; letzter Zugriff am 20.03.2010).

de Vaus, David (2006): Editor's Introduction. Research Design – A Review. In: de Vaus, David (Hrsg.): Research Design. Volume I. London u. a.: Sage, IX-LV.

Diekmann, Andreas (2009): Spieltheorie. Einführung, Beispiele, Experimente. Reinbek bei Hamburg: Rowohlt.

Diekmann, Andreas (2008): Soziologie und Ökonomie: Der Beitrag experimenteller Wirtschaftsforschung zur Sozialtheorie. Kölner Zeitschrift für Soziologie und Sozialpsychologie 60, 528-550.

Diekmann, Andreas (2007): Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Reinbek bei Hamburg: Rowohlt.

Diekmann, Andreas/Voss, Thomas (2004): Die Theorie rationalen Handelns. Stand und Perspektiven. In: Diekmann, Andreas/Voss, Thomas (Hrsg.): Rational-Choice-Theorie in den Sozialwissenschaften. Anwendungen und Probleme. München: Oldenbourg, 13-29.

Diekmann, Andreas/Eichner, Klaus/Schmidt, Peter/Voss, Thomas (Hrsg.) (2008): Rational Choice: Theoretische Analysen und empirische Resultate. Wiesbaden: VS Verlag für Sozialwissenschaften.

DiPrete, Thomas A./Eirich, Gregory M. (2006): Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments. Annual Review of Sociology 32, 271-297.

Drobnič, Sonja/Blossfeld, Hans-Peter (2004): Career Patterns over the Life Course: Gender, Class, and Linked Lives. Research in Social Stratification and Mobility 21 (special edition: Inequality: Structures, Dynamics and Mechanisms: Essays in Honor of Aage B. Sørensen), 139-164.

Dülmer, Hermann (2007): Experimental Plans in Factorial Surveys: Random or Quota Design? Sociological Methods & Research 35, 382-409.

Eifler, Stefanie (2007): Evaluating the Validity of Self-Reported Deviant Behavior Using Vignette Analyses. Quality & Quantity 41, 303-318.

Esser, Hartmut (1999): Soziologie. Allgemeine Grundlagen. Frankfurt a. M./New York: Campus.

Esser, Hartmut (1996): What Is Wrong with 'Variable-Sociology'? European Sociological Review 12, 159-166.

Esser, Hartmut (1990): „Habits", „Frames" und „Rational Choice". Die Reichweite von Theorien der rationalen Wahl (am Beispiel des Befragtenverhaltens). In: Zeitschrift für Soziologie 19, 231-247.

Esser, Hartmut (1986): Können Befragte lügen? Zum Konzept des „wahren Wertes" im Rahmen der handlungstheoretischen Erklärung von Situationseinflüssen bei der Befragung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 38, 314-336.

Faia, Michael A. (1980): The Vagaries of the Vignette World: A Comment on Alves and Rossi. American Journal of Sociology 85, 951-954.

Geißler, Rainer (1996): Kein Abschied von Klasse und Schicht. Ideologische Gefahren der deutschen Sozialstrukturanalyse. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48, 319-338.

Georg, Werner (2006): Soziale Ungleichheit im Bildungssystem. Eine empirisch-theoretische Bestandsaufnahme. Konstanz: UVK Verlagsgesellschaft.

Gigerenzer, Gerd/Todd, Peter M. (1999): Simple Heuristics That Make Us Smart. New York u. a.: Oxford University Press.

Goldstein, Daniel G. (2009): Heuristics. In: Hedström, Peter/Bearman, Peter (Hrsg.): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press, 140-167.

Goldthorpe, John H. (2007): "Cultural capital": some critical observations. In: Scherer, Stefani/Pollak, Reinhard/Otte, Gunnar/Gangl, Markus (Hrsg.): From Origin to Destination. Trends and Mechanisms in Social Stratification Research. Frankfurt a. M./New York: Campus, 78-101.

Goldthorpe, John H. (2002): Globalisation and social class. West European Politics 25, 1-28.

Goldthorpe, John H. (2001): Causation, Statistics, and Sociology. European Sociological Review 17, 1-20.

Goldthorpe, John H. (2000): On Sociology. Volume Two: Illustration and Retrospect. Stanford: Stanford University Press.

Goldthorpe, John H. (1996): The Quantitative Analysis of Large-Scale Data-Sets and Rational Action Theory: For a Sociological Alliance. European Sociological Review 12, 109-126.

Gross, Christiane/Jungbauer-Gans, Monika (2007): Erfolg durch Leistung? Ein Forschungsüberblick zum Thema Wissenschaftskarrieren. Soziale Welt 58, 453-471.

Groß, Jochen/Börensen, Christina (2009): Wie valide sind Verhaltensmessungen mittels Vignetten? Ein methodischer Vergleich von faktoriellem Survey und Verhaltensbeobachtung. In: Kriwy, Peter/Gross, Christiane (Hrsg.): Klein aber fein! Quantiative empirische Sozialforschung mit kleinen Fallzahlen. Wiesbaden: VS Verlag für Sozialwissenschaften, 149-178.

Gross, Neil (2009): A Pragmatist Theory of Social Mechanisms. American Sociological Review 74, 358-379.

Granovetter, Mark S. (1978): Threshold Models for Collective Behavior. American Journal of Sociology 83, 1420-1443.

Grusky, David B. (2001): The Past, Present and Future of Social Inequality. In: Grusky, David B. (Hrsg.): Social Stratification. Class, Race and Gender in Sociological Perspective. Boulder, Colo. u. a.: Westview Press, 1-51.

Hadjar, Andreas (2008): Meritokratie als Legitimationsprinzip. Die Entwicklung der Akzeptanz sozialer Ungleichheit im Zuge der Bildungsexpansion. Wiesbaden: VS Verlag für Sozialwissenschaften.

Haller, Max (2003): Soziologische Theorie im systematisch-kritischen Vergleich. Opladen: Leske + Budrich.

Hartmann, Michael (2009): Die transnationale Klasse – Mythos oder Realität? Soziale Welt 60, 285-303.

Hedström, Peter (2005): Dissecting the Social. On the Principles of Analytical Sociology. Cambridge u. a.: Cambridge University Press.

Hedström, Peter (2004): Generative Models and Explanatory Research: On the Sociology of Aage B. Sørensen. Inequality 21, 13-25.

Hedström, Peter/Bearman, Peter (2009a): What is Analytical Sociology all about? An Introductory Essay. In: Hedström, Peter/Bearman, Peter (Hrsg.): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press, 4-24.

Hedström, Peter/Bearman, Peter (Hrsg.) (2009b): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press.

Hedström, Peter/Udehn, Lars (2009): Analytical Sociology and Theories of the Middle Range. In: Hedström, Peter/Bearman, Peter (Hrsg.): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press, 25-47.

Hedström, Peter/Swedberg, Richard (1996): Social Mechanisms. Acta Sociologica 39, 281-308.

Hempel, Carl G./Oppenheim, Paul (1948): Studies in the Logic of Explanations. Philosophy of Science 15, 135-175.

Herausgeber der Kölner Zeitschrift für Soziologie und Sozialpsychologie (2009): Nachrichten und Mitteilungen. Preis der Fritz Thyssen Stiftung für Sozialwissenschaftliche Aufsätze. Kölner Zeitschrift für Soziologie und Sozialpsychologie 61, 705-709.

Hillebrandt, Frank (2009): So einfach ist es nicht. Analytische Soziologie auf Abwegen. Symposiumsbeitrag zum Buch: Hedström, Peter: Anatomie des Sozialen. Soziologische Revue 32, 363-370.

Hillmert, Steffen/Jacob, Marita (2003): Social inequality in higher education: is vocational training a pathway leading to or away from university? European Sociological Review 19, 319-334.

Hinz, Thomas/Liebig, Stefan (2006): Der Faktorielle Survey als Instrument zur Messung von Einstellungen in Bevölkerungsumfragen. Projektantrag an die DFG. Konstanz und Duisburg: Universitäten Konstanz und Duisburg.

Holland, Paul W. (1986): Statistics and Causal Inference. Journal of the American Statistical Association 81, 945-960.

Hornbostel, Stefan (2009): Alles eine Frage der Perspektive. Münchs Kritik am Deutschen Wissenschaftssystem. Symposiumsbeitrag zum Buch: Münch, Richard: Die akademische Elite. Soziologische Revue 32, 15-20.

Hornbostel, Stefan (2001): Third party funding of German universities. An indicator of research activity? Scientometrics 50, 523-537.

Hout, Michael/DiPrete, Thomas A. (2004): What We Have Learned: RC28's Contributions to Knowledge About Social Stratification. Working Paper of the Survey Research Center. Berkeley: University of Berkeley.

Hradil, Stefan (2001): Soziale Ungleichheit in Deutschland. Opladen: Leske + Budrich.

Jaime-Castillo, Antonio M. (2008): Inequality of Income, Expectations of Social Mobility and the Demand for Redistribution. Vortrag am 6. September 2008 auf dem Forum der International Sociological Association (ISA) in Barcelona.

Jann, Ben (2003): Lohngerechtigkeit und Geschlechterdiskriminierung: Experimentelle Evidenz. Unveröffentlichtes Manuskript an der Eidgenössischen Technischen Hochschule Zürich. Zürich: Eidgenössische Technische Hochschule Zürich.

Jansen, Dorothea/Heidler, Richard/von Görtz, Regina (2009): Ungleiche Chancen im Wissenschaftssystem: Artefakt oder Realität? Kölner Zeitschrift für Soziologie und Sozialpsychologie 61, 463-467.

Jasso, Guillermina (2006): Factorial Survey Methods for Studying Beliefs and Judgments. Sociological Methods and Research 34, 334-423.

Jasso, Guillermina (1996): Exploring the Reciprocal Relations between Theoretical and Empirical Work: The Case of the Justice Evaluation Function (Paper in Honor of Robert K. Merton). Sociological Methods Research 24, 253-303.

Jasso, Guillermina/Webster, Murray (1999): Assessing the Gender Gap in Just Earnings and its Underlying Mechanisms. Social Psychology Quarterly 62, 367-380.

Jasso, Guillermina/Webster, Murray (1997): Double Standards in Just Earnings for Male and Female Workers. Social Psychology Quarterly 60, 66-78.

Jürges, Hendrik (2005): Gender ideology, division of housework, and the geographic mobility of families. Review of Economics of the Household 4, 299-323.

Kaas, Leo (2009): Does Equal Pay Reduce Labour Market Inequality? Scandinavian Journal of Economics 111, 51-71.

Kaube, Jürgen (2007): Abrechnung ohne Strom. Richard Münchs Polemik gegen unser Forschungssystem. Rezension in der Frankfurter Allgemeinen Zeitung vom 17.08.2007.

Kelle, Udo/Lüdemann, Christian (1995): „Grau, teurer Freund ist alle Theorie…“. Rational Choice und das Problem der Brückenannahmen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 47, 249-267.

Klein, Markus (2002): Die Conjoint-Analyse. Eine Einführung in das Verfahren mit einem Ausblick auf mögliche sozialwissenschaftliche Anwendungen. ZA-Information 50, 7-45.

Klinger, Cornelia (2007): Achsen der Ungleichheit. Zum Verhältnis von Klasse, Geschlecht und Ethnizität. Frankfurt a. M.: Campus.

Kreckel, Reinhard (2004): Politische Soziologie der sozialen Ungleichheit. Frankfurt a. M.: Campus.

Kreckel, Reinhard (1983): Theorie sozialer Ungleichheiten im Übergang. In: Kreckel, Reinhard (Hrsg.): Soziale Ungleichheiten. Soziale Welt, Sonderband 2. Göttingen: Otto Schwartz, 3-14.

Krücken, Georg (2009): Strukturwandel der Forschung. Symposiumsbeitrag zum Buch: Münch, Richard: Die akademische Elite. Soziologische Revue 32, 9-14.

Krumpal, Ivar/Voss, Thomas (2009): Explaining Respondents' Behavior in RRT Surveys: A Game Theoretic Approach. Arbeitspapier an der Universität Leipzig. Leipzig: Universität Leipzig.

Kunz, Volker (2004): Rational Choice. Frankfurt a. M./New York: Campus.

Liebig, Stefan/Meyermann, Alexia/Schulze, Andrea (2006): Temporal stability of justice evaluations. Paper presented at the 11th Conference of the International Society for Justice Research. Berlin: Humboldt-Universität.

Liebig, Stefan/Lengfeld, Holger/Mau, Steffen (2004): Gesellschaftliche Verteilungsprobleme und der Beitrag der soziologischen Gerechtigkeitsforschung. In: Liebig, Stefan/Lengfeld, Holger/Mau, Steffen (Hrsg.): Verteilungsprobleme und Gerechtigkeit in modernen Gesellschaften. Frankfurt a. M./New York: Campus, 7-28.

Liebig, Stefan/Schupp, Jürgen (2008): Leistungs- oder Bedarfsgerechtigkeit? Soziale Welt 59, 7-30.

Lind, Imken (2004): Aufstieg oder Ausstieg? Karrierewege von Wissenschaftlerinnen. Ein Forschungsüberblick. (cews.Beiträge Frauen in Wissenschaft und Forschung no. 2). Bielefeld: Kleine Verlag.

Lindenberg, Siegwart (1996): Die Relevanz theoriereicher Brückenannahmen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48, 126-140.

Madden, Janice F. (1973): The Economics of Sex Discrimination. Lexington, Mass.: Lexington Books.

Mahony, James (2001): Beyond Correlational Analysis: Recent Innovations in Theory and Method. Sociological Forum 16, 575-593.

Manzo, Gianluca (2007): Variables, Mechanisms and Simulations: Can the Three Methods be Synthesized? A Critical Analysis of the Literature. Revue française de sociologie 48, Supplement, 35-71.

Mayer, Karl-Ulrich (2009): Der Geist hat die Universität verlassen. Symposiumsbeitrag zum Buch: Münch, Richard: Die akademische Elite. Soziologische Revue 32, 3-8.

Mayer, Karl-Ulrich (1987): Zum Verhältnis von Theorie und empirischer Forschung zur sozialen Ungleichheit. In: Giesen, Bernhard/Haferkamp, Hans (Hrsg.): Soziologie der sozialen Ungleichheit. Opladen: Westdeutscher Verlag, 370-392.

Mayntz, Renate (1961): Kritische Bemerkungen zur funktionalistischen Schichtungstheorie. Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 5 (Soziale Schichtung und soziale Mobilität), 10-28.

McLanahan, Sara/Percheski, Christine (2008): Family Structure and the Reproduction of Inequalities. Annual Review of Sociology 34, 257-276.

Merton, Robert K. (1988): The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property. ISIS 79, 606-623.

Merton, Robert K. (1968): The Matthew Effect in Science. Science 57, 68-72.

Merton, Robert K. (1967): On Sociological Theory of the Middle Range. In: Merton, Robert K. (Hrsg.): On Theoretical Sociology. New York: Free Press, 39-72.

Meulemann, Heiner (2004): Sozialstruktur, soziale Ungleichheit und die Bewertung der ungleichen Verteilung von Ressourcen. In: Berger, Peter A./Schmidt, Volker H. (Hrsg.): Welche Gleichheit, welche Ungleichheit? Grundlagen der Ungleichheitsforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, 115-136.

Moed, Henk F. (2005): Citation Analysis in Research Evaluation. Dordrecht: Springer.

Molm, Linda D./Peterson, Gretchen/Takahashi, Nobuyuki (1999): Power in Negotiated and Reciprocal Exchange. American Sociological Review 64, 876-890.

Müller, Hans-Peter/Schmid, Michael (2003): Vorwort. In: Müller, Hans-Peter/Schmid, Michael (Hrsg.): Hauptwerke der Ungleichheitsforschung. Wiesbaden: Westdeutscher Verlag, 5-10.

Münch, Richard (2009): Kein Kartell, kein Monopol, keine Oligarchie? Wie die Soziologie zur Affirmationswissenschaft wird. Kölner Zeitschrift für Soziologie und Sozialpsychologie 61, 453-461.

Münch, Richard (2007): Die akademische Elite. Zur sozialen Konstruktion wissenschaftlicher Exzellenz. Frankfurt a. M.: edition suhrkamp.

Neckel, Sighard/Dröge, Kai/Somm, Irene (2004): Welche Leistung, welche Leistungsgerechtigkeit? Soziologische Konzepte, normative Fragen und einige empirische Befunde. In: Berger, Peter A./Schmidt, Volker H. (Hrsg.): Welche Gleichheit, welche Ungleichheit? Grundlagen der Ungleichheitsforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, 137-164.

Nisic, Natascha/Auspurg, Katrin (2009): Faktorieller Survey und klassische Bevölkerungsumfrage im Vergleich – Validität, Grenzen und Möglichkeiten beider Ansätze. In: Kriwy, Peter/Gross, Christiane (Hrsg.): Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen. Wiesbaden: VS Verlag für Sozialwissenschaften, 211-246.

Oaxaca, Ronald (1973): Male-Female Wage Differentials in Urban Labor Markets. International Economic Review 14, 693-709.

Opp, Karl-Dieter (2009): Das individualistische Erklärungsprogramm in der Soziologie. Entwicklung, Stand und Probleme. Zeitschrift für Soziologie 38, 26-47.

Opp, Karl-Dieter/Friedrichs, Jürgen (1996): Brückenannahmen, Produktionsfunktionen und die Messung von Präferenzen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48, 546-559.

Otte, Gunnar (2005): Entwicklung und Test einer integrativen Typologie der Lebensführung für die Bundesrepublik Deutschland. Zeitschrift für Soziologie 34, 442-467.

Pager, Devah/Shepherd, Hana (2008): The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets. Annual Review of Sociology 34, 181-209.

Pakulski, Paul (2005): Foundations of a post-class analysis. In: Wright, Olin E. (Hrsg.): Approaches to Class Analysis. Cambridge: Cambridge University Press, 152-179.

Podolny, Joel/Lynn, Freda (2009): Status. In: Hedström, Peter/Bearman, Peter (Hrsg.): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press, 544-565.

Rashotte, Lisa S./Webster, Murray/Whitmeyer, Joseph M. (2005): Pretesting Experimental Instructions. Sociological Methodology 35, 163-187.

Reskin, Barbara F. (2003): Including Mechanisms in our Models of Ascriptive Inequality. American Sociological Review 68, 1-21.

Robinson, William S. (1950): Ecological Correlations and the Behavior of Individuals. American Sociological Review 15, 351-357.

Roemer, John E. (1985): Should Marxists be Interested in Exploitation? Philosophy and Public Affairs 14, 30-65.

Rössel, Jörg (2009): Sozialstrukturanalyse. Eine kompakte Einführung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Rössel, Jörg (2005): Plurale Sozialstrukturanalyse. Eine handlungstheoretische Rekonstruktion der Grundbegriffe der Soziologie. Wiesbaden: VS Verlag für Sozialwissenschaften.

Rössel, Jörg (2004): Macht als zentrale Dimension der Sozialstrukturanalyse. Eine handlungstheoretische Begründung. In: Berger, Peter A./Schmidt, Volker H. (Hrsg.): Welche Gleichheit, welche Ungleichheit? Grundlagen der Ungleichheitsforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, 221-240.

Rossi, Peter H./Anderson, Andy B. (1982): The Factorial Survey Approach: An Introduction. In: Rossi, Peter H./Nock, Steven L. (Hrsg.): Measuring Social Judgments. The Factorial Survey Approach. Beverly Hills: Sage, 15-67.

Rossi, Peter H./Alves, Wayne M. (1980): Rejoinder to Faia. American Journal of Sociology 85, 954-955.

Rusconi, Alessandra (2002): Dual Careers – Insights from Social Research in Western Europe and Germany. Arbeitsbericht der Arbeitsgruppe „Wissenschaftspolitik“ der Jungen Akademie. Berlin: Die Junge Akademie.

Rusconi, Alessandra/Solga, Heike (2008): A Systematic Reflection Upon Dual Career Couples. WZB Discussion Paper SP I 2008-505. Berlin: Wissenschaftszentrum Berlin für Sozialforschung.

Rusconi, Alessandra/Solga, Heike (2007): Determinants of and Obstacles to Dual Careers in Germany. Zeitschrift für Familienforschung 19, 311-336.

Sauer, Carsten/Liebig, Stefan/Auspurg, Katrin/Hinz, Thomas/Donaubauer, Andy (2009): A Factorial Survey on the Justice of Earnings within the GSOEP-Pretest 2008. IZA Discussion Paper No. 4663. Bonn: Forschungsinstitut zur Zukunft der Arbeit (IZA).

Schelling, Thomas C. (1971): Dynamic Models of Segregation. Journal of Mathematical Sociology 1, 143-186.

Schimank, Uwe (1998): Funktionale Differenzierung und soziale Ungleichheit: die zwei Gesellschaftstheorien und ihre konflikttheoretische Verknüpfung. In: Giegel, Hans-Joachim (Hrsg.): Konflikt in modernen Gesellschaften. Frankfurt a. M.: Suhrkamp.

Schmid, Michael (2009): Theoriebildung und Theoriepolitik in der Soziologie. Ein Kommentar zu Norman Braun: „Theorie in der Soziologie“. Soziale Welt 60, 199-213.

Schmid, Michael (2006): Die Forschungsheuristik mechanismischer Erklärungen. In: Schmid, Michael (Hrsg.): Die Logik mechanismischer Erklärungen. Wiesbaden: VS Verlag für Sozialwissenschaften, 139-176.

Schmidt, Volker H. (2004): Ungleichgewichtige Ungleichheiten. In: Berger, Peter A./Schmidt, Volker H. (Hrsg.): Welche Gleichheit, welche Ungleichheit? Grundlagen der Ungleichheitsforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, 73-92.

Schneider, Wolfgang Ludwig (2009): Handlung oder Interaktion? Zur Bestimmung des primären Grundbegriffs in Hedströms Analytischer Soziologie. Symposiumsbeitrag zum Buch: Hedström, Peter: Anatomie des Sozialen. Soziologische Revue 32, 355-362.

Schürmann, Lena (2005): Die Konstruktion von „Hausarbeit" in gleichgeschlechtlichen Paarbeziehungen. In: Solga, Heike/Wimbauer, Christine (Hrsg.): „Wenn zwei das Gleiche tun…". Ideal und Realität sozialer (Un-)Gleichheit in Dual Career Couples. Opladen: Barbara Budrich, 141-162.

Schwinn, Thomas (2008): Zur Analyse multidimensionaler Ungleichheitsverhältnisse. Österreichische Zeitschrift für Soziolgoie 33, 20-42.

Schwinn, Thomas (2007): Soziale Ungleichheit. Bielefeld: transcript.

Sokal, Alan D./Bricmont, Jean (1997): Impostures intellectuelles. Paris: Jacob.

Solga, Heike/Powell, Justin/Berger, Peter A. (2009): Soziale Ungleichheit – Kein Schnee von gestern! Eine Einführung. In: Solga, Heike/Powell, Justin/Berger, Peter A. (Hrsg.): Soziale Ungleichheit. Klassische Texte zur Sozialstrukturanalyse. Frankfurt a. M./New York: Campus, 11-46.

Solga, Heike/Wimbauer, Christine (2005): „Wenn zwei das Gleiche tun…" – Ideal und Realität sozialer (Un-)Gleichheit in Dual Career Couples. Eine Einleitung. In: Solga, Heike/Wimbauer, Christine (Hrsg.): „Wenn zwei das Gleiche tun…". Ideal und Realität sozialer (Un-)Gleichheit in Dual Career Couples. Opladen: Barbara Budrich, 9-26.

Sørensen, Aage B. (2000): Towards a Sounder Basis for Class Analysis. American Journal for Sociology 105, 1523-1558.

Stachura, Mateusz (2009): Brückenhypothesen – Kritik der ökonomischen Theorie der Ziele. Zeitschrift für Soziologie 38, 152-169.

Steiner, Peter M./Atzmüller, Christiane (2006): Experimentelle Vignettendesigns in Faktoriellen Surveys. Kölner Zeitschrift für Soziologie und Sozialpsychologie 58, 117-146.

Thomson Reuters (2010): Journal Citation Reports. (Verfügbar unter: http://emedia1.bsb-muenchen.de/han/WOS/admin-apps.isiknowledge.com/JCR/JCR; letzter Zugriff am 26.01.2010).

Tumin, Melvin M. (1953): Some Principles of Stratification: A Critical Analysis. American Sociological Review 18, 387-394.

Ultee, Wout C. (1996): Do Rational Choice Approaches Have Problems? European Sociological Review 12, 167-179.

Verwiebe, Roland (2010): Soziale Ungleichheit – Neue Befunde aus der Gesundheits-, Armuts- und Transnationalisierungsforschung. Soziologische Revue 33, 23-33.

Wagner, Gerhard (2007): Does excellence matter? Eine wissenschaftssoziologische Perspektive. Soziologie 36, 7-20.

Wallander, Lisa (2009): 25 years of factorial surveys in sociology: a review. Social Science Research 38, 505-520.

Weeden, Kim A./Grusky, David B. (2005): The Case for a New Class Map. American Journal of Sociology 111, 141-212.

Wennerås, Christine/Wold, Agnes (1997): Nepotism and Sexism in Peer-Review. Nature 387, 341-343.

West, S. G. (2006): Validity. In: de Vaus, David (Hrsg.): Research Design. Volume I. London u. a.: Sage, 121-142.

Winker, Gabriele/Degele, Nina (2009): Intersektionalität. Zur Analyse sozialer Ungleichheiten. Bielefeld: transcript.

Wippler, Reinhardt (1985): Die Entstehung oligarchischer Strukturen in demokratisch verfassten Organisationen. In: Büschges, Günter/Raub, Werner (Hrsg.): Soziale Bedingungen – Individuelles Handeln – Soziale Konsequenzen. Frankfurt a. M.: Lang, 23-48.

Wright, Olin E. (2005): Conclusion: If „class“ is the answer, what is the question? In: Wright, Olin E. (Hrsg.): Approaches to Class Analysis. Cambridge: Cambridge University Press, 180-192.

Wooldridge, Jeffrey M. (2002): Econometric Analysis of Cross Section and Panel Data. Cambridge, Mass.: MIT Press.

2

Herausbildung einer akademischen Elite? Zum Einfluss der Größe und Reputation von Universitäten auf Forschungsförderung

(Katrin Auspurg, Thomas Hinz und Jürgen Güdler)

2 Herausbildung einer akademischen Elite?

Zum Einfluss der Größe und Reputation von Universitäten auf Forschungsförderung

Zusammenfassung

Mit der verstärkten Drittmittelfinanzierung von Forschung gewinnen die Verteilungsprinzipien von Forschungsgeldern für den Erfolg einzelner Wissenschaftler sowie ganzer Forschungsinstitutionen an Bedeutung. Vorliegender Beitrag zielt auf die empirische Überprüfung von Hypothesen zum Einfluss der Größe und Reputation von Universitäten auf die Bewilligungschancen der bei ihnen beschäftigten Wissenschaftler. Mit prozessproduzierten Daten zu den im Zeitraum 1991 bis 2004 bei der Deutschen Forschungsgemeinschaft eingereichten Anträgen auf Einzelförderung (sog. „Normalverfahren“) werden multivariate Analysen der individuellen Bewilligungschancen sowie der Drittmitteleinwerbungen universitärer Fachgebiete angestellt. Im Ergebnis sind Kontext- und Konzentrationseffekte kaum nachzuweisen: Lediglich für Wissenschaftler an westdeutschen Traditionsuniversitäten finden sich leicht erhöhte Bewilligungschancen und erst bei weit überdurchschnittlichen Größen setzen sich höhere Personalbestände nicht mehr in vermehrte Antragstellungen und Bewilligungen um. Ebenso ist auf der Makroebene einzelner Universitäten kein Trend einer zunehmenden Ungleichheitsverteilung auszumachen. Abschließend werden diese Befunde in die Debatte um die Exzellenzinitiative eingeordnet, mit den einschlägigen Thesen von Richard Münch kontrastiert und ihre Aussagekraft diskutiert.

Abstract

Given a tendency to higher third-party funding, the mechanisms of distribution of research money become more and more relevant for individual academic careers as well as for scientific institutions at large. This article focuses on an empirical test of hypotheses on the impact of universities’ size and reputation on the chances of grant approval. Using multivariate analysis of register data provided by the German Research Foundation (DFG) for all applications for single grants from 1991 to 2004, individual chances of grant approval as well as success rates in departmental grant acquisition are estimated. The analyses detect neither strong context effects on individual chances of grant approval nor a clear tendency towards a higher concentration of research funding on fewer universities. Only scientists working in West German universities with a long standing tradition have a slightly better chance to get research funding. At the level of university departments, higher personnel resources translate into a higher number of applications and approvals only for very large institutions. Regarding funding of single grants, there is no trend of a growing inequality among the universities. Finally, these results are discussed with a specific reference to the recently launched “Initiative for Excellence” – a program designed in order to foster high level research in Germany – and are contrasted to some arguments of Richard Münch.

2.1 Einleitung

„Wer hat, dem wird gegeben." Spätestens seit den Arbeiten von Robert K. Merton (1968, 1988) wird dieses „Matthäus-Prinzip" auch für Karrieren und Organisationen im Wissenschaftssystem diskutiert. Für die bedeutendste Institution zur Forschungsförderung in Deutschland, die Deutsche Forschungsgemeinschaft (DFG), postuliert entsprechende Größen- und Reputationseffekte jüngst Richard Münch in seinem Buch „Die akademische Elite" (Münch 2007). Anhand von Daten zu den Drittmittelausstattungen, Personalbeständen und Publikationen versucht er seine Hauptthese zu belegen, dass die Verteilung von Forschungsgeldern statt dem Leistungsvermögen der Antragstellenden mehr der Größe und Reputation von Institutionen folgt.

Das Anliegen von Münch ist hehr und verdienstvoll. Die Umsetzung greift jedoch in verschiedener Hinsicht zu kurz (für eine inhaltliche Rezension der Tagespresse: Kaube 2007). Viele Thesen sind dennoch von wissenschaftlichem wie öffentlichem Interesse: Angesichts stagnierender oder sinkender Grundmittel (dazu DFG 2006: 19f.; Statistisches Bundesamt 2005a, 2005b sowie Wissenschaftsrat 2000) bestimmen die Organisation und Verteilung von Drittmitteln mehr denn je die Entwicklung der Wissenschaft insgesamt, ihrer Einrichtungen und ebenso die individuellen Karriereverläufe von Wissenschaftlern und wissenschaftlichem Nachwuchs. In vielen Disziplinen gilt die Drittmittelfinanzierung inzwischen selbst für bescheidene Forschungsvorhaben als unerlässlich (zu entsprechenden Befragungsergebnissen bei Hochschullehrern Schimank 1992). Determinanten der Drittmitteleinwerbung sind daher allein schon aus dem praktisch-politischen Anliegen heraus, begabte und innovative Wissenschaftler möglichst effizient zu fördern, von großem Interesse.

Der eingangs zitierte Matthäus-Effekt fehlt zwar in fast keiner wissenschaftssoziologischen Abhandlung, blieb aber zumindest für die Drittmitteleinwerbungen und Institutionenebene – wohl aufgrund des erschwerten Datenzugangs – wenig untersucht (Ausnahmen stellen z. B. die Arbeiten von Carayol/Matt 2004 sowie Moed et al. 2005 dar; für die deutsche Forschungsförderung: Allmendinger/Hinz 2002; Neidhardt 1988). Auch die von Münch durchgeführten Analysen vermögen die bestehende Forschungslücke kaum zu schließen: Seine oftmals auf Aggregatebene ganzer Universitäten angesiedelten, und lediglich globale Bewilligungssummen statt individuelle Bewilligungschancen berücksichtigenden Auswertungen sind nicht vor Fehlschlüssen gefeit, zumal auf Drittvariablenkontrollen weitestgehend verzichtet wird. Höhere Bewilligungssummen müssen etwa nicht unbedingt höhere Bewilligungsquoten zum Ausdruck bringen – aufgrund des stark schwankenden Mittelbedarfs zwischen einzelnen Forschungsgebieten können sie allein unterschiedliche fachliche Profile widerspiegeln (man vergleiche etwa die durchschnittlichen Kosten von ingenieur- und literaturwissenschaftlichen Forschungsprojekten; zu entsprechenden Statistiken: Hinz et al. 2008: 51ff.; Hornbostel 1997: Kap. 5.3.1). Ähnlich lassen die stark fächerspezifische Nutzung und Verfügbarkeit von Drittmittelgebern (vgl. etwa DFG 2006: 19f.; Hornbostel 1997: 222-224) bei Berechnungen, die über *gesamte* Universitäten aggregiert werden, verzerrte Ergebnisse erwarten. Angesichts dessen scheinen die Schlüsse nachgewiesener „Kartell- und Monopolstrukturen im akademischen Feld"

(Münch 2007: 263) oder auch der Vorwurf, das System der Forschungsförderung bringe „paradoxerweise eine Privilegiengesellschaft [hervor], die systematisch gegen die Prinzipien von Chancengleichheit, offenen Wettbewerb und Leistungsgerechtigkeit verstößt“ (Münch 2007: 91), eindeutig zu voreilig. Gleichwohl sind sie nicht belanglos: Dass auch falsche oder verzerrte Datenauswertungen ihre Folgen haben, ist soziologisches Allgemeingut und – wie es die Ironie will – eine der Grundthesen bei Münch.

Der vorliegende Beitrag nimmt die von Münch (wieder) angeregte Diskussion zum Anlass, detaillierte Analysen zu den Stellgrößen der Forschungsförderung anzustellen. Ermöglicht wird dies durch prozessproduzierte Daten zu den Forschungsanträgen und Bewilligungen im Zeitraum 1992 bis 2004, die für eine Auftragsstudie von der DFG bereitgestellt wurden und nun für diese Sekundäranalyse genutzt werden können.[1] Mit der DFG beziehen wir uns auf die im Fördervolumen weitaus größte, damit wichtigste Institution zur Forschungsförderung in Deutschland. Ihre Finanzierung erfolgt durch öffentliche Mittel, die Bewilligungsentscheidungen werden mithilfe von *peer-review*-Verfahren gefällt. Die Begutachtung ist nicht anonym, sondern den Gutachtenden ist bekannt, wer sich mit welchem institutionellen und Qualifikationshintergrund um Forschungsgelder bewirbt.[2] Wir berücksichtigen im Folgenden ausschließlich Anträge aus Universitäten auf Sachbeihilfen im Normalverfahren.[3] Dieses auch als „Einzelförderung“ bekannte Antragsverfahren unterliegt keinerlei inhaltlichen Beschränkungen und stellt die geläufigste und meist genutzte Möglichkeit der Forschungsförderung durch die DFG dar (vgl. DFG 2005).[4] Seine vorrangige Bedeutung wird mitunter an den Fördersummen deutlich. In den Jahren 2002 bis 2004 entfielen Gesamtausgaben von über 1,3 Milliarden Euro und damit der größte Teil der DFG-Fördermittel (36 Prozent des Gesamtvolumens) auf diese Förderform (DFG 2006: 23; entsprechendes gilt für die Vorjahre, vgl. DFG 2003: 28). Das Verfahren ist streng dem so genannten *bottom-up*-Prinzip verpflichtet, d. h. die Förderungen haben entsprechende Antragstellungen zur Voraussetzung (eine Initiierung seitens der DFG ist ausgeschlossen) und sollten allein die Qualität der Anträge würdigen. Mit den uns vorliegenden „Echtdaten“ lassen sich eingereichte Anträge und Bewilligungen gegenüberstellen und damit diese Verteilungsprinzipien weitaus

[1] Diese Datenbasis wurde uns freundlicherweise von der DFG für die in ihrem Auftrag erstellte Studie „Wissenschaftlerinnen in der DFG. Förderprogramme, Förderchancen und Funktionen (1991–2004) zur Verfügung gestellt (Hinz et al. 2008). Für hilfreiche Erläuterungen zum Datenmaterial sowie aufwändige Arbeiten bei der Datenerstellung bedanken wir uns herzlich bei Thomas Rahlf, Daniel Bovelet, Alexis-Michel Mugabushaka und Anke Reinhardt. Ungenauigkeiten und Fehler liegen allein in der Verantwortung der Autoren.

[2] Die Antragstellenden wissen dagegen nicht, wer die Begutachtung vornimmt.

[3] DFG-Anträge in Koordinierten Programmen wie insbesondere Schwerpunktprogrammen bleiben aufgrund ihrer gesonderten Verfahrenslogik und differenten Förderanliegen bewusst ausgeklammert – sie sollten den Gegenstand separater Analysen bilden. Ebenso werden Anträge aus nicht-universitären Einrichtungen ausgeschlossen.

[4] Es können Mittel für Personal, wissenschaftliche Geräte, Verbrauchsmaterial, Reisen, Publikationen und die meisten anderen Erfordernisse eines Forschungsvorhabens beantragt werden. Die Grundausstattung ist allerdings von der Institution des Antragsstellers zur Verfügung zu stellen. Die übliche Förderdauer eines Projektes umfasst bei einem Neuantrag zwei bis drei Jahre, bei Bedarf kann ein Fortsetzungsantrag gestellt werden (zu weiteren Details: DFG 2005).

direkter beobachten, als dies in den Analysen von Münch der Fall ist (denen lediglich Bewilligungen auf Hochschul- und Fächerebene zugrunde liegen). Wir greifen zusätzlich auf Rahmendaten der Hochschulen (wie deren Personalgrößen) und die DFG-Förder-Rankings (DFG 1997, 2000, 2003) zurück.[5] Mit den Jahren 1991 bis 2004 deckt unser Untersuchungszeitraum ein hinreichend langes Zeitfenster ab, um Fragen einer zunehmenden Ungleichheitsverteilung von Drittmitteln auf Universitäten (nach Münch der Entstehung einer „akademischen Elite") nachzugehen: Haben Wissenschaftler an größeren Universitäten aufgrund deren besseren Ausstattung (mit Personal) und infolge dessen auch höherer Forschungsreputation bessere Bewilligungschancen? Kommt es entsprechend im Zeitverlauf zu einer immer stärkeren Konzentration von (erfolgreichen) DFG-Antragstellungen auf wenige Hochschulstandorte? Oder sind im Gegenteil ab einer bestimmten Größe Ineffizienzen im Sinne eines abnehmenden Grenznutzens festzustellen, die sich in sinkenden Antragsaktivitäten und -chancen niederschlagen?

Im nächsten Abschnitt werden ausgehend von der Arbeit Münchs und unter Einbezug weiterer einschlägiger Literatur die zu untersuchenden Hypothesen abgeleitet. Abschnitt 2.3 erläutert die Datengrundlage und stellt das methodische Vorgehen vor. Die in Abschnitt 2.4 stehenden Ergebnisse unterteilen sich in eine multivariate Analyse der Bewilligungschancen, eine Untersuchung von Größeneffekten (Personalbestände) auf die Antragsaktivitäten und Bewilligungen, sowie drittens eine Analyse der Ungleichverteilung von Drittmitteln auf einzelne Hochschulen im Zeitverlauf. Wir schließen mit einem Resümee und Ausblick.

2.2 Argumentation bei Münch, Forschungsstand und Hypothesen

Richard Münch lässt in seinem Buch „Die akademische Elite" fast keinen soziologischen Ansatz aus, um seine Kritik an der derzeitigen Förderpolitik geltend zu machen. Viele Klassiker von Rang und Namen werden zur Unterstützung der Argumentation herangezogen. Etwa dient die Zivilisationstheorie von Elias der Begründung von Monopolisierungstendenzen, die Herausbildung und Abschottung einer Elite werden aber gleichfalls mit Kapitalbegriffen in Erweiterung der Feldtheorie von Bourdieu, den „Dispositiven der Macht" nach Foucault oder Begriffen der Klassentheorie von Marx begründet. Den Inhalt im Gesamten nachzuverfolgen ist hier nicht möglich und erforderlich (für eine inhaltliche Rezension s. etwa Kaube 2007). Stattdessen sollen hier nur diejenigen Annahmen aufgegriffen werden, die von einem grundsätzlichen ungleichheits- oder wissenssoziologischen Interesse sind, damit eine gewisse theoretische Plausibilität für sich verbuchen und zudem von unserer Datengrundlage abgedeckt werden. Diese bei Münch oftmals lediglich implizit zu findenden Behauptungen werden im

[5] Mit den Förder-Rankings der DFG beziehen wir das zentrale Medium des Forschungsmonitoring in Deutschland ein, das damit besonders für Reputationseffekte und als Auslöser für die sich selbst verstärkende Wirkung von Drittmitteleinwerbungen „verdächtig" ist (vgl. Münch 2007: Kap. I.3; für allgemeine Diskussionen Laudel 2005). Es werden nicht nur die Bewilligungssummen der DFG beziffert, sondern daneben auch Kriterien wie die Drittmittelvolumen insgesamt, die Internationalität der Forschung und Einbindung in hochschulübergreifende Kooperationsnetzwerke (vgl. DFG 2003: 15f., 2006: 13f.).

Folgenden, unter Einbezug weiterer Literatur, begründet und in Hypothesenform auf den Punkt gebracht.

Eine der Grundthesen Münchs lautet, dass Wissenschaftler an großen und bekannten Einrichtungen bessere Bewilligungschancen haben und dies nicht allein durch ihre höhere Leistungsfähigkeit gerechtfertigt ist: „Von entscheidender Bedeutung ist dabei die Tatsache, dass angesichts der in nahezu allen Anträgen auf Forschungsförderung gefundenen Mängel sekundäre Merkmale wie Reputation und Standortgröße eine überragende Bedeutung bekommen“ (Münch 2007: 226). Zur Begründung dient die erst mit einer gewissen Größe gegebene „kritische Masse“ für umfangreichere Projekte, die Notwendigkeit der Unsicherheitsreduktion über den zu erwartenden Erfolg und daneben, in Anlehnung an Bourdieu, die wechselseitige Verstärkung von ökonomischem, symbolischem, kulturellem und sozialem Kapital.

Stringenter als bei Münch finden sich theoretische Ableitungen von Größen- und Reputationseffekten ebenso in der Literatur. Auf der Institutionenebene ist etwa das bessere Vermögen von Einrichtungen mit guter Expertise und Ressourcenausstattung zur Anwerbung leistungsfähiger Wissenschaftler zu nennen – was dann wiederum zu einer höheren Produktivität und Drittmitteleinwerbung dieser Einrichtungen führt (für einen empirischen Nachweis dieser *beiden* Effekte: Allison/Long 1990; Long/McGinnis 1981). Ebenso sind positive Rückkoppelungen auf individueller Ebene zu erwarten: Erfolgreiche Antragstellungen führen zu einem höheren Forschungsoutput in Form von Publikationen und Patenten, und damit zur Sichtbarkeit der Wissenschaftler, was alles wiederum die Grundlage für weitere erfolgreiche Antragstellungen bildet. Ein höheres Budget bietet zudem mehr Flexibilität hinsichtlich des Mitteleinsatzes, was mitunter die Vorbereitung neuer Anträge erleichtert (vgl. Laudel 2005: 29ff.). Selbstverstärkungen sind aber allein schon deshalb zu erwarten, weil erfolgreiche Drittmitteleinwerbungen immer mehr als Nachweis wissenschaftlicher Produktivität gelten: Drittmittel spielen bei der leistungsbezogenen Budgetierung innerhalb von Hochschulen als auch im Reputationsvergleich zwischen Hochschulen eine zunehmende Rolle (vgl. Hornbostel 2001: 524; Jaeger 2006). Die Reputation eines Standortes strahlt dabei möglicherweise auch auf die dort beschäftigten Wissenschaftler aus, da Gutachtende Anträgen aus prestigeträchtigen Hochschulen eine höhere Qualität und Machbarkeit attestieren. Hinzu kommt die insbesondere an Großstandorten gute Ausgangslage für vitale regionale Forschungsumfelder. Institute der Forschungsverbünde (z. B. Max-Planck- und Fraunhofer-Gesellschaften) und *Spin-offs* der Hochschulen prägen gemeinsam das Forschungsprestige einer Region. Der *Silicon Valley*-Effekt, der sich aus solchen Agglomerationen ergibt, ist ein Standortfaktor ganz eigener Art. Verbesserte Kooperationsmöglichkeiten vor Ort führen zur Entstehung so genannter *Forschungs-Cluster* und verhelfen mit ihren nationalen und internationalen Forschungsnetzwerken zu einem Bild moderner Wissenschaft (s. am Beispiel der deutschen Medienforschung sowie der Soziologie Güdler 1996, 2003). Diese und weitere Argumente begründen den Matthäus-Effekt in der Wissenschaft – wer einmal erfolgreich ist, erhält überproportionalen Kredit (Fox 1983: 296). Und dies womöglich unabhängig davon, ob der Anfangserfolg gerechtfertigt ist.

Gleichwohl wie sie entstanden sind, übersetzen sich anfängliche Unterschiede im Sinne einer *self-fulfilling-prophecy* in zunehmende Ungleichheiten, so die Annahme und Kritik (s. Hornbostel 2001; Gillett 1991; Laudel 2005 für Diskussionen im Hinblick auf Drittmitteleinwerbungen).

Für den Forschungsoutput in Form von Publikationen und Patenten liegen bereits diverse Evidenzen vor (z. B. Allison et al. 1982; Bakanic et al. 1987; Carayol/Matt 2004; eine Übersicht über den diesbezüglichen Forschungsstand bieten zudem DiPrete/Eirich 2006 sowie Fox 1983). Untersuchungen für den Drittmittelinput sind bislang selten und sprechen nur zum Teil für einen (selbstverstärkenden) Einfluss institutioneller Merkmale (J. Cole/S. Cole 1979; Moed et al. 1998; für Bewilligungschancen bei der deutschen Forschungsgemeinschaft: Allmendinger/Hinz 2002). Zunächst ist daher die folgende Hypothese zu testen:

H_1: *Wissenschaftler aus größeren und besser reputierten Einrichtungen haben bessere Bewilligungschancen für weitere Anträge als Wissenschaftler aus kleineren Einrichtungen mit geringerer Reputation.*

Diese globale Annahme lässt sich entsprechend ihrer Indikatoren in mehrere Unterhypothesen zergliedern. Zunächst ist eine förderliche Wirkung hoher Personalkapazitäten anzunehmen. Mit größeren (und möglicherweise bereits drittmittelfinanzierten) Personalressourcen lassen sich mehr Anträge und diese zudem akribischer vorbereiten – so auch ein Hauptargument Münchs für den „Selbstläufer“ Drittmitteleinwerbungen. Hohe Personalbestände können zudem als ein Indikator für Forschungshochburgen und wissenschaftliche Ballungszentren angesehen werden, welche aufgrund ihrer besseren Grundausstattung oder Diskussionsnetzwerke für die Antragsqualitäten förderlich sein oder auch nur bei Gutachtenden das Vertrauen in die Förderungswürdigkeit bestärken können:

H_{1a}: *Je größer der Personalbestand einer Institution bereits ist, desto höher sind die Bewilligungschancen weiterer Anträge.*

Münch sieht zudem die westdeutschen (einschließlich Berlin) und Traditionsuniversitäten mit einer *von Haus aus* höheren Reputation und Sichtbarkeit verbunden (vgl. etwa 2007: 208, 2008: 74). Als vielleicht plausiblere Begründungen lassen sich deren zumeist zentralere geographische Lage in attraktiven Ballungszentren und die höheren Gehaltsniveaus in den alten Bundesländern (vgl. Domen 2007) anführen, was beides zu einem Wettbewerbsvorteil im Hinblick auf die Anwerbung besonders leistungsstarker Wissenschaftler führt: [6]

[6] Argumentieren lässt sich zudem mit geringeren Antragserfahrungen von Wissenschaftlern an neu gegründeten und, zumindest für die Jahre nach der Wende, an ostdeutschen Universitäten. Allerdings wurde Letzterem durch zeitweise spezielle Antragsregelungen für ostdeutsche Wissenschaftler entgegenzutreten versucht (Laudel 2005: 34, Anm.3).

H_{1b}: *Anträge aus westdeutschen (Traditions-)Universitäten haben höhere Bewilligungschancen als Anträge aus ostdeutschen und spät gegründeten Universitäten.*

Vor allem aber ist ein Institutioneneffekt über die sich selbst verstärkende Wirkung der Forschungsreputation anzunehmen. Drittmitteleinwerbungen gewinnen als Indikator für Forschungsleistungen zunehmende Bedeutung (Hornbostel 2001; Jaeger 2006). Nimmt man ein paralleles Ranking von Institutionen und Personen an, ist von höheren Bewilligungschancen bei Wissenschaftlern an hoch reputierten Institutionen auszugehen (Broder 1993; vgl. auch Allmendinger/Hinz 2002: 282; Münch 2008: 72). Hinzu kommt ein Selektionseffekt im Sinne der Ansiedelung besonders leistungsfähiger Wissenschaftler an prestigereichen und forschungsstarken Institutionen, wobei bereits die alleinige Unterstellung eines solchen Effekts durch Gutachtende im Sinne einer *self-fulfilling-prophecy* zu höheren Erfolgschancen der dort beschäftigten Wissenschaftler und in der Folge wiederum höheren Forschungsreputationen dieser Institutionen führen würde. Für Deutschland lässt insbesondere das von der DFG seit 1997 veröffentlichte Förder-Ranking (DFG 1997, 2000, 2003, 2006) eine derartige Signalwirkung erwarten:

H_{1c}: *Je höher die Drittmitteleinwerbung einer Institution (gemessen am DFG-Förder-Ranking) bereits ist, desto höher sind die Bewilligungschancen der dort beschäftigten Wissenschaftler.*

Weiter ist zu erwarten, dass ältere Wissenschaftler bessere Bewilligungschancen aufweisen als jüngere. Dies ist neben der erst im Karriereverlauf anzusammelnden Vorleistungen und Reputation ebenso mit ihrer höheren Erfahrung und damit einem Humankapitalvorteil zu begründen. Wesentliche Bewertungskriterien für Forschungsanträge bilden die Durchführbarkeit des geplanten Projekts und die angemessene Kalkulation zeitlicher wie finanzieller Mittel. Diesbezüglich ist anzunehmen, dass „antragserfahrene“ Wissenschaftler adäquatere Vorstellungen haben und zudem besser Bescheid wissen, wie das eigene Vorhaben geschickt in einen Antrag „verpackt“ werden kann. Da zudem Gutachtende mehrheitlich einen Professorenstatus und ein relativ hohes Durchschnittsalter aufweisen (s. Hinz et al. 2008: 91f.; Koch 2006: 31), schreiben sie möglicherweise den älteren, statusgleichen Professoren mehr Kompetenzen zu und haben sie höhere Hemmungen, Anträge anerkannter *peers* statt wissenschaftlicher Neulinge abzulehnen (ähnlich Allmendinger/Hinz 2002: 282; Neidhardt 1988; Zuckermann/Merton 1971: 82). Missallokationen und Verstöße gegen die Universalismusnorm würden vorliegen, wenn Gutachtende ein zu hohes Gewicht auf das Alter als Signal für die künftige Produktivität legen, damit begabtem Nachwuchs den Zugang zu aussichtsreichen Karrieren versperren: Alter wird möglicherweise zum *ungerechtfertigten* Ersatzindikator für Qualität (ähnlich Münch 2007: 231). Unabhängig davon, wie eventuelle Alterseffekte letztlich zu interpretieren sind, soll folgende Hypothese getestet werden:

H_2: *Mit dem Alter der Antragstellenden steigen die Bewilligungschancen.*

Münch vermutet zudem eine Wechselwirkung mit der institutionellen Reputation: Insbesondere junge Wissenschaftler, die noch über keine eigene Reputation verfügen, sollten auf das Renommee ihrer Einrichtung angewiesen sein.[7] Die dahinter stehende Annahme ist, dass den Gutachtenden bei etablierten Wissenschaftlern ausreichend Wissen vorliegt, um deren Expertise einzuschätzen. Bei jungen und daher noch unbekannten Antragstellenden greifen sie dagegen möglicherweise stärker auf das Image ihrer Herkunftsuniversität, das Prestige ihrer Mentoren und die von ihnen durchlaufene „Schule" zurück, um das in ihren Anträgen steckende Potenzial zu begutachten:

H_3: *Insbesondere für junge Wissenschaftler steigen die Chancen der Drittmitteleinwerbung mit der Reputation und Größe ihrer Institution.*

Wenden wir uns noch speziell den Wirkungen des sozialen Kapitals zu. Für die outputsteigernde Wirkung von Netzwerken aufgrund der Möglichkeit zum Informationsaustausch, effizienten Arbeitsteilungen oder auch gegenseitigen Motivationen und Anregungen liegen bereits Evidenzen vor (s. z. B. Fox 1983 für einen Überblick). Münch hat jedoch wiederum mehr die – den meritokratischen Wettbewerb begrenzenden – Reputationseffekte im Blick, hohe Gutachtertätigkeiten an einer Institution würden *allein* schon aufgrund der damit verbundenen „Symbolkraft" die Drittmittelchancen dieser Institution erhöhen (vgl. etwa Münch 2007: 256). Ist eine „Vetternwirtschaft" in Form gegenseitiger Begutachtungen von Institutskollegen durch das DFG-Begutachtungsverfahren ausgeschlossen (DFG 2003: 75), lässt sich der Vorteil einer institutionellen Nähe dennoch mit Einblicken in den Begutachtungsprozess begründen; den Wissenschaftlern ist besser bekannt, welche Informationen DFG-Gutachter erwarten. Zudem sind Effekte im Sinne einer Reziprozität denkbar: Anträge von Personen, die für ihre hohe Gutachteraktivität bekannt sind, werden weniger abgelehnt. Nähe und Ähnlichkeit zwischen Gutachtenden und Begutachtenden könnten darüber hinaus zu „Vorschusslorbeeren" führen (ähnlich Allmendinger/Hinz 2002: 289). In diese Richtung weisende Befunde liegen bereits für Begutachtungen durch den Forschungsrat eines kleinen westeuropäischen Landes vor (Moed 2005: 247-258; Münch 2007: 224).[8] Als alternative Erklärung, so auch die Argumentation der DFG, kommt wieder ein Selektionseffekt in Frage, d. h. Wissenschaftler an Standorten mit höherer Forschungsexpertise und damit höherer „Bewilligungsstärke" werden von der DFG eher als Gutachter eingesetzt (DFG 2003: 89):

[7] Hierin wird von ihm zugleich wieder ein Verstoß gegen die Meritokratie gesehen: „Der Anfangserfolg ist deshalb schon in erheblichem Maße eine soziale Konstruktion. Entscheidend ist dann in der Tat die institutionelle Herkunft der Forscher, insbesondere die Kooperation mit schon reputierten Forschern" (2007: 234). Dem wäre insoweit zuzustimmen, als sich keine plausiblen Argumente dafür finden lassen, warum Selektionseffekte (prestigereiche und wohlhabende Universitäten weisen im Kampf um die besten Köpfe einen Wettbewerbsvorteil auf) speziell für junge Wissenschaftler gelten sollten.

[8] Hatten (Mit-)Antragsteller einen Sitz in einem Ausschuss des Forschungsrats, wurden höhere Bewilligungsquoten erzielt, speziell wenn es sich um den entscheidenden Ausschuss handelte (Moed 2005: 249-253).

H_4: *Die Chancen auf Drittmitteleinwerbung sind umso höher, je mehr DFG-Gutachter der Institution des Antragstellers angehören.*

Für die bislang genannten Effekte ist zudem jeweils eine Verstärkung durch die steigende Bedeutung der Drittmittelforschung anzunehmen (s. Hornbostel 1997: Kap. 5.3.1, 2001 zu entsprechenden Statistiken). Bei Münch wird dies eher lapidar mit der erforderlichen Unsicherheitsreduktion begründet, der Effekt lässt sich stringenter aus der *Signaling*-Theorie herleiten (für deren allgemeine Darstellung z. B. Spence 1973), denn die Bewertung von Anträgen ist für Gutachtende arbeits- und zeitintensiv. Mit zunehmender Nachfrage nach Drittmitteln und damit der Notwendigkeit, auch besser qualifizierte Anträge abzulehnen, wird es zunehmend aufwändiger wenn nicht unmöglich, die „Spreu vom Weizen zu trennen". In der Folge sollten Begutachtungsleistungen früherer Kollegen, bereits eingeworbene Forschungsmittel und ganz allgemein die Reputation von Wissenschaftlern und Institutionen als einfach (damit zeiteffizient) zu überblickende *Screening*-Merkmale an Attraktivität gewinnen. Dies führt zur folgenden Hypothese:

H_5: *Die positiven Effekte von Größe und Gutachteranzahl der Institution auf die Bewilligungschancen nehmen mit der Konkurrenz um Forschungsmittel zu.*

Bislang wurden nur positive Größeneffekte von Institutionen hervorgehoben. Aus einer „produktionstechnischen" Sicht lässt sich jedoch ebenso ein negativer Effekt erwarten (vgl. Jansen et al. 2007 für eine ähnliche Argumentation im Hinblick auf den Publikationsoutput). Mit der Größe von Arbeitsgruppen nehmen der Bedarf an Koordination und Verwaltung zu, oder theoretisch gesprochen, steigen die Transaktionskosten. Überwiegen die Kosten den Mehrgewinn, sollten sich jenseits eines kritischen Niveaus die positiven Effekte hoher Personalbestände in einen negativen Effekt verkehren. Für diese These *abnehmenden Grenznutzens* liegen für den Output in Form von Publikationen und Patenten schon bestätigende Befunde vor, allerdings scheinen die negativen Effekte erst bei weit überdurchschnittlichen Gruppengrößen durchzuschlagen (für Deutschland: Jansen et al. 2007: 137).[9] Für den Input finden sich bislang nur wenige und widersprüchliche Ergebnisse. Schimank (1992: 31) findet keinen Zusammenhang zwischen haushaltsmäßiger Personalausstattung und drittmittelfinanzierten Stellen; Münch (2007) stellt dagegen eine positive Korrelation zwischen Mitarbeiterzahlen und Drittmittelvolumen fest. Die zu untersuchende These eines abnehmenden Grenznutzens hoher Personalbestände lässt sich in Anlehnung an Münch (vgl. 2007: 289) folgendermaßen formulieren:

H_6: *Die Drittmitteleinwerbungen (Antragszahlen und Bewilligungen) nehmen nur bis zu einer mittleren Größe mit den Personalbeständen zu, danach sinken sie mit jedem weiterem Personal.*

[9] Münch findet abweichend bereits ab durchschnittlichen Personalbeständen negative Effekte, die Analysen scheinen jedoch wenig belastbar; näheres dazu unten in Abschnitt 2.3.

Abschließend wollen wir noch die Makroebene in den Blick nehmen. Werden einzelne Wissenschaftler im Laufe ihrer Karriere immer erfolgreicher, zeigt sich dies nicht zwangsläufig ebenso auf der Ebene von Instituten und Universitäten – allein die hohe Mobilität von Wissenschaftlern vermag dem entgegenzuwirken. Eine zusätzliche Validierung auf der Aggregatebene ist zudem im Hinblick darauf von Interesse, welcher Nettoeffekt letztlich aus den beiden gegenläufigen Größeneffekten (Matthäus-Effekt und abnehmender Grenznutzen) verbleibt. Unsere letzte Hypothese lautet daher:

H_7: *Langfristig kommt es zu einer immer stärkeren Konzentration von Anträgen und Bewilligungen auf wenige Großstandorte.*

Diese Annahmen wurden, so unsere Behauptung, von Münch allesamt zu wenig stichhaltig untersucht und sind allgemein für die (deutsche) Forschungsförderung empirisch vernachlässigt. Ein Erkenntnisgewinn unserer Arbeit besteht sicher in der detaillierten Analyse von Kontexteffekten auf die Drittmitteleinwerbung. Die methodische Kritik an Münch und entsprechende Weiterführung unserer Analysen werden im Folgenden näher erläutert.

2.3 Datengrundlage und Methodik

Bevor wir unser eigenes Vorgehen vorstellen, präzisiert der folgende Einschub die bislang nur angeschnittenen Kritikpunkte an den Münch`schen Analysen. Diese bilden schließlich einen Referenzpunkt für die hier vorgestellten Auswertungen und Ergebnisse.

2.3.1 Empirische Auswertungen bei Münch: Vorgehen und Kritik

Als Material dienen Münch Sekundärdaten zu den Mitarbeiterzahlen, eingeworbenen Drittmittelsummen und Publikationen von Hochschulen, die aus verschiedenen Quellen wie den Förder-Rankings der DFG, den Hochschul-Rankings des Centrums für Hochschulentwicklung (CHE) oder Veröffentlichungen des Statistischen Bundesamts zusammengetragen werden. Seine Innovation sieht Münch insbesondere in der Berechnung *relativer* Kennzahlen: Input wie Output werden mit den Personalbeständen der Universitäten gewichtet. Dieses Vorgehen ist sinnvoll und dem absoluten Kennziffernwesen überlegen. Problematisch ist jedoch, dass es sich um Vergleiche globaler Mittelwerte handelt. Fehlende fachliche Differenzierungen dürften speziell bei der gewählten abhängigen Variablen, dem Drittmittelvolumen (statt Bewilligungsquoten oder Anzahl von Drittmittelprojekten) zu Verzerrungen führen. Zwischen und innerhalb von Fachgebieten bestehen schließlich erhebliche Unterschiede im Mittelbedarf für Forschungsvorhaben (zu entsprechenden Statistiken: DFG 2006; Hornbostel 1997: Kap.5.3.2). Man denke etwa an die hohen Kosten der Laborausstattung natur-

wissenschaftlicher Projekte im Vergleich zum typischen Bedarf eines philosophischen Projekts.[10] Ohne entsprechende Disaggregationen können unterschiedliche Bewilligungssummen daher allein differente fachliche Ausrichtungen von Standorten widerspiegeln und können höhere Mittel pro Kopf oder pro Institution durchaus mit weniger Drittmittelprojekten und niedrigeren Bewilligungschancen einhergehen; gerade hieran entzündet sich ja die Hauptkritik an Drittmitteln als Leistungsindikator (Hornbostel 1997, 2001; Laudel 2005).

Die von Münch auf Fachgebietsebene angesiedelten Auswertungen bleiben dann bei bivariaten Analysen stehen. Dies ist im Hinblick auf nicht auszuschließende Drittvariableneffekte kritisierbar. So liegen beispielsweise bereits Evidenzen für einen positiven Zusammenhang zwischen dem Karrierealter von Forschern und ihren Personalressourcen vor (für entsprechende Literaturhinweise: Carayol/ Matt 2004: 1082). Die Korrelationen zwischen Personalbeständen und Bewilligungssummen stellen somit womöglich statt einem Größen- einen Humankapitaleffekt dar (Belohnung der größeren Forschungs- und Antragserfahrung älterer Wissenschaftler). Die multivariaten Analysen lassen andererseits wiederum eine hinreichende Kontrolle für Fachgebiete missen und sind zudem aufgrund des Querschnittdesigns nicht in der Lage, die Wirkungsrichtung von Effekten zu bestimmen: Führt Größe zu mehr Mitteln oder ist umgekehrt Größe das Ergebnis erfolgreicherer Bewilligungen? Die zur Prüfung vieler Hypothesen eigentlich erforderlichen Interaktionseffekte sind ebenso wenig zu finden wie eine Modellierung der postulieren kurvilinearen Effekte (abnehmender Grenznutzen).[11] Hinzu kommen fehlerhafte Interpretationen von Kennwerten; so werden Determinationskoeffizienten über unterschiedliche Fallzahlen und abhängige Variablen hinweg verglichen (z. B. Münch 2007: 288). Trendanalysen werden von Münch erst gar nicht angestellt. Dass Auswertungsebene und Wahl eines geeigneten Auswertungsverfahrens durchaus von Relevanz sind, belegen nicht zuletzt unsere nachfolgend dargestellten, zum Teil doch deutlich abweichenden Ergebnisse.

2.3.2 Eigene Datengrundlage und Methodik

Unsere Berechungsgrundlage bilden gleichfalls prozessproduzierte Sekundärdaten. Diese umfassen für jeden im Zeitraum 1991 bis 2004 entschiedenen Antrag auf Sachbeihilfe Informationen zu dem Ergebnis der Förderentscheidung, zur institutionellen Einbindung (Name der Hochschule) und wenigen weiteren Merkmalen der Antragstellenden (wie ihrem Geburtsjahr und Geschlecht). Ab dem Jahr 1999 umfasst die Datenbasis zudem Angaben zu den Gutachtenden und deren institutionellen Herkunft.[12]

[10] Hinzu kommen unterschiedlich starke Antragsaktivitäten und die fachgebiets- wie themenspezifische Verfügbarkeit von weiteren Drittmittelgebern neben der DFG (Hornbostel 1997: 214f.; Laudel 2005).

[11] Zwar wird von Interaktionseffekten gesprochen (etwa Münch 2007: 293), dabei scheint es sich aber fälschlicherweise um Drittvariableneffekte zu handeln.

[12] Differenzierungen zwischen Sonder- und Fachgutachtenden werden dabei nicht getroffen. Die ehrenamtlichen Sondergutachterinnen und -gutachter werden von der DFG-Geschäftsstelle aufgrund ihrer fachlichen Expertise ausgewählt. Die bis zur Reform des Begutachtungssystems und der Einführung des Systems der Fachkollegien im Jahr 2004 ebenfalls noch tätigen Fachgutachter wurden dagegen auf Vorschlag der Fachverbände alle vier

Wir begrenzen unsere Analysen auf die an Universitäten angesiedelten Antragstellenden und Gutachtenden und berücksichtigen lediglich Neu-, keine Fortsetzungsanträge.

Die Größe der Institution der Antragsteller operationalisieren wir analog zu Münch mit ihrem Personalbestand, genauer ihren Anzahlen an wissenschaftlichen Mitarbeitenden und Dozenten. Da die Mehrzahl der Anträge von Inhabern einer Professur erfolgt, stellen wir zudem jeweils alternative Berechnungen mit einer Eingrenzung auf diesen Personenkreis an.[13] Diese Informationen wurden mittels Sonderauswertungen von Daten des Statistischen Bundesamts zugespielt, welche ab dem Jahr 1992 die jährliche Anzahl von Mitarbeitenden unterschiedlicher Statusebenen an den Fachbereichen der einzelnen Hochschulen enthalten. Die fachlichen Klassifikationen von DFG und Statistischem Bundesamt wurden mittels eines eigens erstellten Zuordnungsschlüssels angeglichen. Sinnvolle Zuordnungen lassen sich dabei überwiegend nur auf der Ebene der DFG-Fachgebiete vornehmen, sodass sich die von uns herangezogenen Kontextmerkmale (wie eben der Personalbestand) als kleinste Einheit auf das jeweilige DFG-Fachgebiet[14] und Jahr der betreffenden Hochschule beziehen. Da die fachlichen Klassifikationen bei der DFG nicht der institutionellen Herkunft der Antragstellenden, sondern den Fachkollegien folgen, welche die Förderempfehlung entwickeln (bzw. im hier betrachteten Zeitraum: den Fachausschüssen), sind dennoch vereinzelte Unschärfen nicht zu vermeiden. Die hiervon primär betroffenen Fachgebiete Medizin und Agrar-/Forstwissenschaften/Gartenbau/Tiermedizin werden von unseren Analysen vollständig ausgeschlossen.[15] Diese unweigerlichen Nachteile prozessproduzierter Daten werden in mancher Hinsicht durch die Vorteile einer Vollerhebung und zudem „Echtdaten“ wieder aufgewogen.

Jahre von der wissenschaftlichen *Community* gewählt und waren hauptamtlich für Begutachtungen zuständig (für Einzelheiten zum Begutachtungsverfahren s. DFG 2003 sowie Koch 2006, dort insbesondere für die mit der Umstellung auf Fachkollegien verbundenen Neuerungen).

[13] Der Status der Antragstellenden ist uns leider nicht bekannt. Nach einer Sonderauswertung der DFG wurden im Jahr 2006 aber 57 Prozent der Anträge von Inhabern einer Professur eingereicht (Hinz et al. 2008: 32). Da uns keine Angaben zum akademischen Grad der Mitarbeitenden vorliegen, ist eine Eingrenzung auf das antragsberechtigte Personal (mit zumindest abgeschlossener Promotion, vgl. DFG 2005) nicht möglich.

[14] Genauer legen wir die DFG-Fachsystematik zugrunde, wie sie seit der Reform des Begutachtungsprozesses im Jahr 2004 (dazu Koch 2006) im Einsatz ist. Diese Klassifikation folgt der operativen Struktur der Antragsbearbeitung und unterscheidet vier Wissenschaftsbereiche sowie insgesamt 14 Fachgebiete. Vor der Reform eingereichte Anträge wurden von der DFG rückwirkend an diese Fächerstruktur angepasst. Dies ist insofern erwähnenswert, als ihr damit in mancher Hinsicht ein „fiktiver“ Charakter zukommt: Fachliche Zuordnungen waren vorab zum Teil anders geregelt.

[15] Es ist etwa möglich, dass Anträge von Mitarbeitern eines biologischen Instituts von einem der Chemie zugeordneten Fachkollegium entschieden werden (vgl. DFG 2006: 14). Ebenso werden in den Fachkollegien der Medizin Anträge von Chemikern, Biologen oder auch Ingenieuren beurteilt. Die letztendliche Förderentscheidung trifft dann der Hauptausschuss der DFG. Zu den Einzelheiten der Verfahren und Gremien, wie etwa auch zur Umstellung von Fachausschüssen auf Fachkollegien, siehe www.dfg.de > Förderung > Förderung auf einen Blick. Nach Ausschluss der genannten beiden Fachgebiete reduzieren sich die fehlenden Werte bei den Personalbeständen antragstellender Institutionen deutlich zu insgesamt nur mehr 2 Prozent.

Als „Traditionsuniversitäten“ werten wir alle in Westdeutschland gelegenen Universitäten mit einem Gründungsdatum vor 1945.[16] Da die Entwicklung der Forschungstradition durch die DDR zumindest zeitweise einen Einschnitt erfahren hat, vergeben wir dieses „Label“ lediglich für Westuniversitäten. Bei der West-/Ost-Einteilung rechnen wir analog zu Münch *alle* Berliner Universitäten (also auch die Humboldt-Universität) zu Westdeutschland. Aufgrund ihrer, selbst während den DDR-Zeiten gegebenen, Ausnahmestellung im Sinne einer hohen Reputation und Internationalität erscheint uns dieses Vorgehen gerechtfertigt.

Als Indikator für die Forschungsreputation ziehen wir (wie Münch) die anhand der DFG-Förder-Rankings (DFG 1997, 2000, 2003) ermittelten Forschungsstärken heran. Diese leiten sich aus dem absoluten Bewilligungsvolumen ab, das die Hochschulen in den Vorjahren bei der DFG eingeworben haben. Das Ranking erfolgt spezifisch für die vier von der DFG unterschiedenen Wissenschaftsbereiche und erscheint aufgrund seiner zentralen Bedeutung für das Forschungsmonitoring in Deutschland besonders geeignet, unsere Thesen zur Reputationswirkung zu operationalisieren.

Für alle Institutionenmerkmale verwenden wir jeweils die Daten der Vorjahre, um eine bessere kausale Aufschlüsselung zu gewährleisten (also etwa für die Analysen der im Jahr 2003 entschiedenen Anträge den Personalbestand im Jahr 2002; damit kann dieser Personalbestand lediglich Ursache, nicht aber Folge der im Jahr 2003 erfolgten Drittmitteleinwerbungen sein). Eine deskriptive Übersicht zu den verwendeten Variablen und Fallzahlen leistet Tabelle 2.1.

Tabelle 2.1: Übersicht über die in den Regressionsanalysen verwendeten Variablen

	N	Min	Max	Mittelw.	SD	Median
Anzahl Professuren pro Universität, Fachgebiet und Jahr	8.158	0	342	25,5	29,3	17
Anzahl wissenschaftlicher Mitarbeiter pro Universität, Fachgebiet und Jahr	8.158	0	1084	1.22,1	123,6	89
Anzahl Gutachter pro Universität, Fachgebiet und Jahr	3.621	0	33	3,8	4,9	2
Anzahl gestellte Anträge pro Universität, Fachgebiet und Jahr	8.578	0	80	5,0	6,7	3
Anzahl bewilligte Anträge pro Universität, Fachgebiet und Jahr	8.578	0	38	2,6	3,7	1
Förderentscheidung (1 = bewilligt)	43.187	0	1	0,5		
Antrag aus westdeutscher Universität (1 = ja)	43.187	0	1	0,8		
Antrag aus Traditionsuniversität (1 = ja)	43.187	0	1	0,6		
Antrag aus Universitäten nach deren Forschungsstärke (Ref.: Mittelfeld)						
- Spitzengruppe (1 = ja)	33.987	0	1	0,5		
- Schlussgruppe (1 = ja)	33.987	0	1	0,1		
Geschlecht des Antragstellenden (1 = weiblich)	43.186	0	1	0,1		
Alter des Antragstellenden	41.967	25	93	49,8	9,7	50

[16] Die oftmals bereits in den 1960er Jahren gegründeten Reformuniversitäten erscheinen uns kaum mit dem Bild „altehrwürdiger“ und damit hochreputierter Universitäten assoziiert. Wir weichen an dieser Stelle daher bewusst von dem von Münch gewählten Stichdatum (Gründungen vor 1970) ab.

Als Auswertungsverfahren greifen wir auf multivariate Analysen der Antragschancen und -zahlen zurück. Aufgrund unserer besonderen Datenstruktur verwenden wir Mehrebenen- bzw. Panelverfahren (Random Intercept Modelle). Mit diesen wird vereinfacht gesprochen der unbeobachteten Heterogenität zwischen einzelnen Hochschulstandorten Rechnung getragen: Mittels institutionenspezifischen Fehlertermen werden die nicht explizit berücksichtigten Merkmale der Institutionen (hier also der Fachgebiete an einzelnen Universitäten) modelliert, was unverzerrtere Schätzungen der eigentlich interessierenden Effekte ermöglicht (vgl. Greene 2003). Einzelheiten zu diesen Verfahren wie auch zu den für die Analyse der Antragszahlen eingesetzten Zähldatenmodellen (Poission- und Negativ-Binomial Modelle) sind dem methodischen Anhang dieses Kapitels zu entnehmen.[17]

Für die Frage der Ungleichheitsverteilung von Antragstellungen (wie Bewilligungen) auf der Aggregatsebene, also der Ebene von Fachgebieten einzelner Hochschulen, ziehen wir als Ungleichheitsmaß den Gini-Koeffizienten heran, exakter seine auch für kleine Fallzahlen auf den Wertebereich von Null bis Eins normierte Variante (sog. „Lorenz-Münzner-Koeffzient"). Es interessieren Veränderungen im Zeitverlauf (gemäß H_7 ist mit einem Anstieg der Ungleichverteilung von Anträgen auf Hochschulen über die Jahre zu rechnen), sodass das Ausmaß der Ungleichverteilung unabhängig von den im Untersuchungszeitraum stattfindenden *generellen* Zu- und Abnahmen gemessen werden sollte. Insbesondere diese Eigenschaft der *Skaleninvarianz* spricht für die Wahl des Gini-Koeffizienten; sein Wert bleibt unverändert, wenn sich die Antragszahlen an allen Universitäten um den *gleichen* Faktor erhöhen (für Diskussionen speziell im Hinblick auf die Messung des Matthäus-Effekts in der Wissenschaft: Allison et al. 1982: 616f.; Okrasa 1986: 223).

2.4 Ergebnisse

Für unseren Untersuchungszeitraum (1991 bis 2004) liegen Daten zu insgesamt 49.498 entschiedenen Anträgen auf Sachbeihilfe aus 84 verschiedenen Universitäten vor (unter Ausschluss der Fachgebiete Medizin und Agrar-/Forstwissenschaften/Gartenbau/Tiermedizin). Wir berichten zunächst Regressionsschätzungen zum Einfluss kontextueller Faktoren auf die individuellen Bewilligungschancen (Abschnitt 2.4.1). Anhand der Antrags- und Bewilligungszahlen analysieren wir dann die These eines abnehmenden Grenznutzens größerer Personalbestände (Abschnitt 2.4.2). Zuletzt gehen wir der Frage (zunehmender) Ungleichheitsverteilung auf der Ebene der Universitäten nach (Abschnitt 2.4.3).

2.4.1 Einfluss kontextueller Merkmale auf die Bewilligungschancen

Angaben zu den kontextuellen Merkmalen der Antragstellungen liegen uns für die Jahre 1993 bis 2004 vor. Die Förderquote beträgt über diesen gesamten Zeitraum gemittelt 51,7 Prozent, sie sinkt im Zeit-

[17] Obgleich die prozessproduzierten Daten jeweils für sich den Charakter einer Vollerhebung tragen, gilt dies nicht mehr für unsere gesamte Datengrundlage. Vereinzelt fehlen Eintragungen, zusätzlich treten durch das *Matching* der unterschiedlichen Datenquellen *Missings* auf. Wir gehen von zufälligen Ausfallprozessen aus und berichten daher Signifikanzschwellen (s. Behnke 2005 zum Nutzen von Signifikanztests bei Vollerhebungen).

verlauf und insbesondere ab dem Jahr 2000 deutlich. Während 1993 noch 61,5 Prozent aller Anträge gefördert wurden, sind es 2004 nur mehr 39,6 Prozent. Zugleich steigen die jährlichen Antragszahlen (1993: 2.756; 2004: 4.326). Im Zeitverlauf nimmt die Konkurrenzsituation um Forschungsgelder also deutlich zu.[18]

Umso mehr interessiert, ob Anträge aus großen und hoch reputierten Einrichtungen bessere Bewilligungschancen aufweisen (H_1) und ob sich dieser Effekt mit der steigenden Konkurrenz verschärft (H_5). Die Antwort ist Tabelle 2.2 zu entnehmen, in der die Ergebnisse von logistischen Regressionen der Bewilligungschancen (bewilligt versus abgelehnt) präsentiert werden.[19] Tabelliert sind dort statt den Koeffizientenwerten die anschaulicher zu interpretierenden marginalen Wahrscheinlichkeitseffekte bzw. *discrete changes* (Long/Freese 2006: Kap. 4.6.6). Diese geben jeweils an, um wie viele Prozentpunkte sich die vom Modell vorhergesagte Wahrscheinlichkeit einer Bewilligung *im Durchschnitt* verändert, wenn die Ausprägung einer Dummy-Variablen vom Wert ‚Null' auf den Wert ‚Eins' wechselt oder sich metrische Variablen marginal (um etwa eine Einheit) erhöhen.

Um greifbarere Werte zu erhalten, wird das Personal in 100-Personen-Einheiten und das Alter der Antragstellenden in 10-Jahres-Schritten gemessen. Der für das Personal in Modell 1 ausgewiesene Effekt von 0,53 bedeutet also beispielsweise, dass sich die Chance einer Bewilligung im Mittel um 0,5 Prozentpunkte erhöht, wenn in einem Fachgebiet einer Hochschule 100 Mitarbeiter mehr beschäftigt sind. Dieser Effekt weist in die vorhergesagte Richtung, die Signifikanzschwelle von 5 Prozent wird jedoch nicht erreicht. Die These höherer Bewilligungschancen bei größeren Personalbeständen (H_{1a}) ist somit zurückzuweisen.

Bestätigung findet dagegen die Vermutung, dass Antragstellende an westdeutschen (einschließlich Berlin) sowie an Traditions-Universitäten bessere Bewilligungschancen aufweisen als ihre Kollegen an ostdeutschen und an nach 1945 gegründeten Universitäten (H_{1b}). Zumindest der West-Ost-Effekt erweist sich auch in seiner Effektstärke nicht als marginal: Anträge aus westdeutschen Universitäten haben gut 7 Prozentpunkte bessere Bewilligungschancen als solche aus ostdeutschen Universitäten.[20] Inwieweit es sich hierbei um einen reinen Selektionseffekt handelt – westdeutsche (Traditions-) Universitäten beschäftigen leistungsfähigere Wissenschaftler, oder ebenso das Renommee der westdeutschen (Traditions-)Universitäten die Bewilligungschancen unabhängig von der Antragsgüte erhöht, bleibt an dieser Stelle spekulativ. Weiterführende Analysen, in denen zusätzliche Merkmale der Anträge, Antragstellenden oder deren Universitäten (wie etwa deren Grundausstattung) kontrolliert werden, sind sicherlich angezeigt.

---

[18] Das Budget ist zwar ebenfalls stark gestiegen, aber eben nicht im gleichem Umfang wie die Nachfrage (für Statistiken: DFG 1997, 2000, 2003, 2006; Hinz et al. 2008).

[19] Die in der Datenbank vereinzelt ebenfalls zu findende Kategorie „anderweitig erledigter Anträge" wird von uns, der bei der DFG üblichen Berichtspraxis folgend, zu den nicht bewilligten Anträgen gezählt.

[20] Dieser Effekt bleibt fast unverändert bestehen, wenn die Berliner Universitäten als separate Kategorie gefasst werden (Anträge aus westdeutschen Universitäten weisen dann immer noch um 6,5 Prozentpunkte höhere Bewillligungschancen auf als solche aus ostdeutschen Universitäten).

Tabelle 2.2: Logistische Regressionen der Bewilligungschancen (Random Intercept Modelle; 1993 bis 2004 gesamt)[a]

	Modell 1	Modell 2	Modell 3	Modell 4
Kontextmerkmale				
Anzahl Mitarbeiter[b] [100]	0,53	0,54	0,53	0,00
Anzahl Mitarbeiter[b] [100] x Antragstellung nach 1999				1,08***
Westdeutsche Universität	7,14***	7,96***	7,19***	7,08***
Westdeutsche Universität x junger Antragsteller		-2,83**		
Traditionsuniversität	2,94**	2,95**	3,61***	2,96**
Traditionsuniversität x junger Antragsteller			-2,49*	
Stadtgröße [100.000 Einwohner]	-0,10*	-0,10*	-0,10*	-0,09
Fachgebiet (Ref.: Geisteswissenschaften)				
- Sozialwissenschaften	-2,37	-2,37	-2,34	-2,32
- Biologie	3,95*	4,06*	3,97*	3,90*
- Chemie	8,77***	8,88***	8,85***	8,79***
- Physik	9,25***	9,34***	9,27***	9,21***
- Mathematik	7,58***	7,66***	7,58***	7,50***
- Geowissenschaften (inkl. Geographie)	4,43**	4,49**	4,41**	4,39**
- Elektro-/Systemtechnik, Informatik	-7,67***	-7,64***	-7,68***	-7,67***
- Bauwesen, Architektur	-6,21**	-6,16**	-6,19**	-6,25**
- Fachgebiete des Maschinenbaus	-3,49*	-3,48*	-3,48*	-3,31*
Individualmerkmale				
Alter [10 Jahre]	-22,28***	-26,96***	-25,14***	-22,42***
Alter [10 Jahre] quadriert	0,95***	1,00***	1,00***	0,96***
Geschlecht (1 = Frau)	-2,61**	-2,62**	-2,64**	-2,59**
Antragsjahr (Ref.: 1993)				
- 1994	1,99	2,03	2,01	2,00
- 1995	1,95	1,97	1,94	1,92
- 1996	-6,20***	-6,16***	-6,18***	-6,20***
- 1997	-12,74***	-12,69***	-12,72***	-12,76***
- 1998	-13,66***	-13,61***	-13,63***	-13,64***
- 1999	-2,84*	-2,77*	-2,79*	-2,83*
- 2000	-6,10***	-6,02***	-6,06***	-8,65***
- 2001	-16,67***	-16,58***	-16,62***	-18,93***
- 2002	-20,12***	-20,06***	-20,07***	-22,32***
- 2003	-24,96***	-24,91***	-24,95***	-27,04***
- 2004	-22,42***	-22,37***	-22,40***	-24,64***
Beobachtungen	41.194	41.194	41.194	41.194
(Fachgebiete)	(594)	(594)	(594)	(594)
Varianz(anteil)[c]:				
- σ_v	0,25	0,25	0,25	0,25
- rho	0,02***	0,02***	0,02***	0,02***
Wald Chi²	1.894,11***	1.899,86***	1.899,08***	1.908,54***

[a] Tabelliert sind jeweils *discrete changes* (durchschnittliche Veränderung der Wahrscheinlichkeit einer Bewilligung in Prozentpunkten bei Veränderung der unabhängigen Variablen um eine Einheit).
[b] Im Vorjahr des Fachgebiets an der jeweiligen Universität.
[c] Im Nullmodell betragen diese Größen: $\sigma_v = 0{,}41$; rho = 0,05.
*** p<0,001; ** p<0,01; * p<0,05.

Wertet man das Lebensalter als grobe Annäherung, bestätigt sich die These der Honorierung von individuellen Vorerfahrungen und Reputationen (H_2) *nicht*. Das Alter der Antragstellenden übt im Gegenteil einen *negativen* Einfluss auf die Bewilligungschance aus. Der signifikante (positive) Effekt des quadrierten Alters zeigt allerdings an, dass dies nur bis zu einem Alter von Mitte 50 gilt, danach steigen die Chancen wieder leicht an (grafisch handelt es sich um einen u-förmigen Effekt). Hier nicht dargestellte, nach Antragsjahrgängen separierende Analysen ergeben, dass es sich dabei in der Tat um einen Alters- statt Kohorten- oder Periodeneffekt handelt (einzusehen in Hinz et al. 2008). Über die Erklärung dieses überraschenden Effekts, der insbesondere auf die Ingenieurwissenschaften zurückgeht (vgl. ebenfalls Hinz et al. 2008), kann an dieser Stelle nur spekuliert werden. Möglicherweise handelt es sich bei den jungen Antragstellern um eine Positivselektion besonders engagierter und innovativer Wissenschaftler oder sind Gutachter bei Anträgen von jüngeren Kollegen weniger kritisch, wie dies bereits für die Begutachtung von Zeitschriftenbeiträgen gefunden wurde (Zuckermann/ Merton 1971: 93). Zur Überprüfung wären ebenfalls gesonderte Auswertungen anzustellen. Festzuhalten ist aber, dass ein Mindest-(Karriere-)Alter zumindest unseren Modellschätzungen zufolge keine Bedingung für eine erhöhte Bewilligungschance bei der DFG darstellt.

Die weiteren in Modell 1 enthaltenen Variablen erfüllen Kontrollfunktionen und sind daher lediglich kurz zu berichten. Durch die Kontrolle der Einwohnerzahlen an den Hochschulstandorten ist wenigstens annäherungsweise sichergestellt, dass sich die höheren Bewilligungschancen von Wissenschaftlern an westdeutschen sowie Traditionsuniversitäten nicht ausschließlich der vergleichsweise häufigeren Großstadtlage von Universitäten in Westdeutschland verdanken (es scheint sich also hinter dem Effekt mehr zu verbergen als der alleinige Standortvorteil großer, damit in der Regel kulturell attraktiver und verkehrstechnisch gut erreichbarer Städte für die Anwerbung von Spitzenwissenschaftlern). Die Werte für die einzelnen Fachgebiete spiegeln die in der (Fach-)Öffentlichkeit viel diskutierten höheren Drittmitteleinwerbungen naturwissenschaftlicher Disziplinen wider.[21] Die Schätzungen für die Antragsjahre belegen die bereits angesprochene Abnahme der Bewilligungschancen im Zeitverlauf. Anträge von Wissenschaftlerinnen weisen geringfügig schlechtere Bewilligungschancen auf als Anträge ihrer männlichen Kollegen – dieser Befund ist für die hier interessierenden Hypothesentestungen ebenfalls sekundär, für eine vertiefende Analyse verweisen wir auf Hinz et al. 2008.

Es interessiert allerdings, ob junge (hier operationalisiert mit dem untersten, bis einschließlich 42 Jahre reichenden Altersquartil), und damit vermutlich in der *Community* noch weitgehend unbekannte Antragstellende speziell von der Größe und Reputation ihrer Institution profitieren (H_3). Ein Blick auf die entsprechenden (negativ ausfallenden) Interaktionseffekte in Modell 2 (mit Westdeutschland) und Modell 3 (mit Traditionsuniversität) lässt uns diese Annahme zurückweisen: Es sind im Gegenteil

[21] Ihre zum Teil deutlichen Effekte auf die Bewilligungschance weisen die Fachgebiete als eine wichtige Kontrollvariable aus. Bei alleiniger Kontrolle für die vier Wissenschaftsbereiche findet sich ein signifikanter positiver Effekt der Mitarbeiterzahlen – der fälschlicherweise für eine Bestätigung der *Größenhypothese* gedeutet werden könnte. Die von Münch vorgelegten multivariaten Analysen (s. Münch 2007: 446f.) greifen in dieser Hinsicht eindeutig zu kurz.

Antragstellende mittleren und höheren Alters, die bei Beschäftigung an einer westdeutschen (Traditions-)Universität signifikant höhere Bewilligungschancen aufweisen.[22] Zur Prüfung der Annahme, dass speziell junge Wissenschaftler von großen Institutionen profitieren (ebenfalls H_3), haben wir alternativ zu den Interaktionen in Modell 2 und 3 eine solche zwischen einer Antragstellung in jungem Alter (bis einschließlich 42 Jahre) und dem Personalbestand gebildet. Die hier nicht dargestellte Modellschätzung weist diesen Interaktionseffekt als nicht signifikant aus. Demnach ist die Größe der Universität für die *altersspezifische* Bewilligungschance unerheblich und findet die Annahme einer besonders förderlichen Wirkung *sichtbarer* Institutionen für den wissenschaftlichen Nachwuchs keine Bestätigung.

Der Vermutung, dass sich die Wirkung kontextueller Merkmale insbesondere unter Bedingungen verschärfter Konkurrenz entfaltet (H_5), gehen wir ebenfalls mit Interaktionstermen nach. Die Bewilligungsquoten erfahren speziell mit dem Jahr 2000 einen Einbruch (dazu Hinz et al. 2008), weshalb wir Wechselwirkungen mit einer Antragstellung nach 1999 prüfen. Modell 4 beinhaltet die einzig signifikante positive Interaktion mit dem Personalbestand (geprüft haben wir ebenso Interaktionen mit den Merkmalen „Westdeutschland“ und „Traditionsuniversität“). Inhaltlich bedeutet die Interaktion, dass große Personalbestände an der Universität der Antragstellenden die Bewilligungschancen nach 1999 stärker begünstigen als es bis zu diesem Jahr der Fall war.[23] Verglichen mit Merkmalen wie den Fachgebieten oder Antragsjahren ist dieser Effekt allerdings sehr schwach.

Unseren Indikator für Reputation, die in den DFG-Förder-Rankings ausgewiesenen Drittmitteleinwerbungen bei der DFG (DFG 1997, 2000, 2003), haben wir bislang ausgespart. Um unabhängige und abhängige Variablen zeitlich (und damit kausal) separieren zu können, verwenden wir jeweils Rankings zeitlich vorgelagerter Perioden. Das früheste Förder-Ranking der DFG basiert auf den Drittmitteleinwerbungen in den Jahren 1991 bis 1995 und dient uns für die Operationalisierung der Forschungsreputation ab dem Jahr 1996; die Forschungsreputation ab dem Jahr 1999 wird anhand des zweiten (auf den Einwerbungen der Jahre 1996 bis 1998 basierenden) Förder-Rankings gemessen und die Reputation ab dem Jahr 2002 mittels des dritten (die Einwerbungen in den Jahren 1999 bis 2001 berücksichtigenden) Rankings. Wir betrachten somit nur mehr Antragstellungen ab dem Jahr 1996. Modell 5 in Tabelle 2.3 wiederholt zunächst die Modellschätzung 1 für diesen verkürzten Untersuchungszeitraum. Die bislang gefundenen Effekte erweisen sich dabei als weitgehend stabil.

[22] Was es nahe legt, primär von Selektionseffekten auszugehen, denn die in ihren Gehaltsstrukturen attraktiveren westdeutschen (Traditions-)Universitäten (Domen 2007) weisen bei Berufungsverhandlungen um besonders profilierte Wissenschaftler einen Wettbewerbsvorteil auf. Allerdings bleiben diese Interpretationen mit unserer Datenbasis spekulativ.

[23] Bis 1999 steigt die Bewilligungschance pro 100 Mitarbeiter mehr um durchschnittlich 0,20 Prozentpunkte; nach 1999 sind es dagegen 0,88 Prozentpunkte. Nur der Effekt nach 1999 ist zum 5-Prozent-Niveau signifikant. Ebenfalls schwach signifikant, aber mit einem negativen Vorzeichen versehen, ist die nicht dargestellte Interaktion des Antragsjahrs mit Traditionsuniversitäten. Der *Traditionseffekt* schwächt sich somit konträr zu H_5 mit der im Zeitverlauf steigenden Konkurrenz ab.

Tabelle 2.3: Logistische Regressionen der Bewilligungschancen (Random Intercept Modelle; 1996 bzw. 2000 bis 2004 gesamt)[a]

	Modell 5	Modell 6	Modell 7	Modell 8
Kontextmerkmale				
Anzahl Mitarbeiter[b] [100]	0,80*	0,36	0,40	-0,07
Westdeutsche Universität	7,28***	6,96***	7,04***	6,84***
Traditionsuniversität	2,48*	1,84	0,15	-0,16
Forschungsstärke[b] (Ref.: Mittelfeld)				
- Spitzengruppe		1,76	2,56*	2,05
- Schlussgruppe		-3,63*	-2,09	-1,87
Anzahl Gutachter[b] [10]				1,97
Stadtgröße [100.000 Einwohner]	-0,13*	-0,14**	-0,09	-0,09
Fachgebiet (Ref.: Geisteswissenschaften)				
- Sozialwissenschaften	-2,78	-2,45	-1,17	-0,43
- Biologie	2,82	2,30	3,99	3,26
- Chemie	8,58***	8,30***	6,55**	6,69**
- Physik	8,42***	8,05***	6,39**	6,56**
- Mathematik	5,50*	4,92*	2,03	2,53
- Geowissenschaften (inkl. Geographie)	4,06*	3,34	4,87*	4,67*
- Elektro-/Systemtechnik, Informatik	-6,91***	-7,25***	-4,65*	-4,53*
- Bauwesen, Architektur	-6,27*	-7,09**	-8,05**	-7,31*
- Fachgebiete des Maschinenbaus	-3,20	-3,57*	-1,93	-2,21
Individualmerkmale				
Alter [10 Jahre]	-21,96***	-21,76***	-22,02***	-21,97***
Alter [10 Jahre] quadriert	1,10***	1,11***	1,02**	1,03**
Geschlecht (1 = Frau)	-3,09**	-3,08**	-2,33	-2,32
Antragsjahr (Ref.: 1996 bzw. 2000)				
- 1997	-6,65***	-6,63***		
- 1998	-7,65***	-7,66***		
- 1999	3,46*	3,51**		
- 2000	0,99	0,14		
- 2001	-10,67***	-10,63***	-10,53***	-10,62***
- 2002	-14,26***	-14,32***	-13,93***	-14,03***
- 2003	-19,35***	-19,36***	-19,01***	-19,16***
- 2004	-16,73***	-16,75***	-16,63***	-17,06***
Beobachtungen	32.791	32.791	18.517	18.517
(Fachgebiete)	(585)	(585)	(572)	(572)
Varianz(anteil):				
- σ_v	0,25	0,24	0,26	0,26
- rho	0,02***	0,02***	0,02***	0,02
Wald Chi²	1.148,89***	1.170,76***	531,72***	535,09***

[a] Tabelliert sind jeweils *discrete changes* (durchschnittliche Veränderung der Wahrscheinlichkeit einer Bewilligung in Prozentpunkten bei Veränderung der unabhängigen Variablen um eine Einheit).
[b] Im Vorjahr des Fachgebiets an der jeweiligen Universität.
*** p<0,001; ** p<0,01; * p<0,05.

Modell 6 bezieht nun zusätzlich die für die jeweiligen Vorperioden ermittelte Forschungsstärke der Hochschulen ein. Anträge aus Hochschulen mit einer Bewertung im oberen (unteren) Quartil des je-

weiligen Wissenschaftsbereichs werden von uns als Antrag aus einer forschungsstarken (-schwachen) Hochschule gewertet. Die verbleibenden Anträge aus dem „Mittelfeld“ dienen als Referenzgruppe. Da speziell von den leicht rezipierbaren DFG-Rankingtabellen Reputationseffekte ausgehen sollten, verwenden wir analog zu diesen die absoluten (statt relativen, mit dem Personalbestand gewichteten) Bewilligungssummen. Das Ergebnis fällt nur teilweise erwartungskonform aus: Kein „Bonus“ für Universitäten des Spitzenfeldes, aber ein geringfügiger „Malus“ für Anträge aus der Schlussgruppe. Ob ersteres einer stärker disaggregierten Analyse Stand hält (etwa auf Ebene der Fachgebiete) und letzteres einen Selektions- oder Reputationseffekt darstellt, bleibt künftigen Untersuchungen vorbehalten. Hier kann lediglich noch berichtet werden, dass entgegen von H_3 und H_5 keine Wechselwirkungen mit dem Alter der Antragsstellenden oder Antragsjahr bestehen. Zur Prüfung haben wir separate Schätzungen von Modell 6 mit zusätzlicher Aufnahme eines Interaktionsterms zwischen der Schlussgruppe und einer Antragstellung in jungem Alter bzw. einer Antragstellung nach 1999 durchgeführt. Beide Interaktionen erwiesen sich als nicht signifikant, was inhaltlich bedeutet, dass der Einfluss der Forschungsreputation altersunspezifisch ist und sich auch nicht mit den Antragsjahren (bzw. der wachsenden Konkurrenz um Forschungsgelder) verstärkt.

Aus Überlegungen zur positiven Wirkung sozialen und symbolischen Kapitals wurde eine weitere Hypothese abgeleitet, die umso höhere Chancen der Drittmitteleinwerbung postuliert, je mehr DFG-Gutachtende der Universität angehören (H_4). Daten zu den Gutachtenden und ihrer institutionellen Herkunft liegen uns erst ab dem Jahr 1999 vor. Modell 7 unterscheidet sich von Modell 6 lediglich durch die stärkere zeitliche Einschränkung (Antragstellungen in den Jahren 2000 bis 2004 statt 1996 bis 2004, damit auch kleinere Fallzahlen). Für die in Modell 8 zusätzlich berücksichtigte Anzahl von DFG-Gutachtern (im jeweiligen Fachgebiet und Vorjahr der Hochschule) findet sich kein Effekt auf die Bewilligungschance. Ebenso lässt sich entgegen H_5 keine Verstärkung des Effektes der Gutachteranzahl an der Hochschule mit zunehmender Konkurrenz feststellen (separate, hier nicht abgebildete Modellschätzung mit einer zusätzlichen Interaktion von Gutachteranzahl und Antragsjahr; die Interaktion erweist sich als nicht signifikant).

Analoge Analysen mit einer Eingrenzung des Personalbestands auf Professuren führen zu keinen anderen inhaltlichen Ergebnissen (aus Platzgründen nicht dargestellt). Ziehen wir ein Zwischenfazit: Hinsichtlich der Bewilligungschancen bestehen vereinzelt Evidenzen für Selektions- oder Reputationseffekte. Anträge aus westdeutschen und Traditionsuniversitäten weisen leicht erhöhte Bewilligungschancen auf. Dies gilt speziell für Anträge älterer Wissenschaftler, was eher auf einen Selektions- statt Reputationseffekt schließen lässt. Allgemein sinken die Bewilligungschancen jedoch zunächst mit dem Alter und steigen sie erst ab einem Alter von Mitte 50 wieder leicht an. Die Größe von Einrichtungen, gemessen an ihren Personalbeständen, eine hohe Forschungsreputation sowie die Anzahl ihrer DFG-Gutachtenden erweisen sich dagegen als neutral für die Bewilligungschancen. Dass diese Institutionenmerkmale allesamt wenig zur Aufklärung unterschiedlicher Bewilligungschancen beitragen, wird auch an den in den unteren Zeilen aufgeführten Varianzanteilen deutlich – der Anteil

unerklärter Varianz, der auf die Institutionenebene zurückgeht (angegeben als „rho“) bleibt über alle Modelle hinweg stabil bei knapp 2 Prozent. Auch nach diesem Parameter sind also die Ursachen für unterschiedliche Bewilligungschancen primär in individuellen Merkmalen der Anträge und der Antragsteller zu suchen, und nicht in den Eigenschaften der Institutionen.

2.4.2 Einfluss von Personalbeständen auf die Antragsaktivität

Zur Beantwortung der Frage, bis zu welcher Einrichtungsgröße eine Zunahme des Personals für Drittmitteleinwerbungen vorteilhaft ist (H_6), ziehen wir Zähldatenmodelle heran (Tabelle 2.4). Bei diesen bilden die (erfolgreichen) Antragstellungen pro Jahr und Fachgebiet einer Universität die abhängige Variable. Statt den Koeffizientenwerten präsentieren wir anschaulicher zu interpretierende Prozenteffekte. Die tabellierten Werte geben jeweils an, um wieviel Prozent sich die abhängige Variable verändert, wenn sich die entsprechende unabhängige Variable um eine Einheit erhöht (bzw. bei Dummy-Variablen die Einser-Kategorie statt der Referenzkategorie vorliegt; s. Wooldridge 2003: 187-189 zu diesen so genannten „Ertragsraten“).

Die Modelle 1 und 2 beziehen sich auf alle Antragstellungen, die Modelle 3 und 4 nur auf die Bewilligungen. Sie unterscheiden sich in der Operationalisierung des Personalbestandes mit ausschließlich Professuren (Modelle 1 und 3) oder allen Mitarbeitenden (Modelle 2 und 4). Stets liegt ein positiver Effekt des Personals aber negativer des quadrierten Personals vor. Inhaltlich bedeutet dies, dass die Anzahlen von (gestellten wie bewilligten) Forschungsanträgen mit dem Personalbestand wachsen, aber mit abnehmender Rate. Jenseits eines kritischen Wertes wirken zusätzliche Mitarbeiter nicht mehr positiv, sondern negativ auf die Antragszahlen. Entgegen der Vermutung Münchs werden diese kritischen Schwellen jedoch erst bei einer weit überdurchschnittlichen Personalausstattung erreicht. Dies ist Tabelle 2.5 zu entnehmen, in der neben den Wendepunkten[24] die mittleren Personalbestände aufgeführt sind sowie die Anteile der Einrichtungen (Fachgebiete pro Jahr und Universität), welche diese bereits überschritten haben. Nur in seltenen Ausnahmen von jeweils unter 1 Prozent der Fälle haben Einrichtungen bereits eine Personalgröße erreicht, bei denen unser Modell negative Effekte weiterer Personalzuwächse auf die Drittmitteleinwerbung prognostiziert. (Inhaltlich könnten diese auch als fehlende Notwendigkeit dieser Einheiten interpretiert werden, durch Drittmittel weitere Stellen für Projektmitarbeitende zu schaffen).

[24] Die Wendepunkte errechnen sich durch Ableiten und Nullsetzen der Schätzformel (zu dieser siehe den Methodenanhang). In unserem Fall ergeben sie sich als: $x = -\beta_{\text{Anzahl Personal}}/2\beta_{\text{Anzal Personal quadriert}}$.

Tabelle 2.4: Negativ-Binomial Regressionen der Antragszahlen bzw. Bewilligungen (1993 bis 2004 gesamt)[a]

	Gestellte Anträge		Bewilligte Anträge	
	Modell 1	**Modell 2**	**Modell 3**	**Modell 4**
Anzahl Professuren[b] [10]	24,55***		32,97***	
Anzahl Professuren[b] [10] quadriert	-0,65***		-0,93***	
Anzahl Mitarbeiter[b] [100]		58,95***		65,81***
Anzahl Mitarbeiter[b] [100] quadriert		-3,01***		-3,04***
Westdeutsche Universität	-30, 76***	-25,44**	-23,83**	-17,41*
Traditionsuniversität	101, 56***	82,04***	106,58***	90,48***
Stadtgröße [100.000 Einwohner]	0,00*	0,00*	0,00*	0,00*
Fachgebiet (Ref.: Geisteswissenschaften)				
- Sozialwissenschaften	7,98	-5,74	1,64	-11,26
- Biologie	257,99***	112,67***	359,18***	148,69***
- Chemie	196,93***	63,38***	287,77***	93,63***
- Physik	89,56***	19,10	186,52***	61,59***
- Mathematik	-36,54**	-48,76***	-20,76	-41,50***
- Geowissenschaften (inkl. Geographie)	110,21***	54,10***	208,07***	98,91***
- Elektro-/Systemtechnik, Informatik	117,26***	23,72	107,87***	6,10
- Bauwesen, Architektur	-10,19	-37,87**	0,35	-38,81**
- Fachgebiete des Maschinenbaus	192,84***	70,53***	231,27***	68,00***
Antragsjahr (Ref.: 1993)				
- 1994	10,29**	10,37**	14,32***	14,46***
- 1995	17,72***	17,82***	21,58***	21,57***
- 1996	19,33***	17,32***	11,00**	8,68*
- 1997	35,51***	31,06***	11,23**	7,10
- 1998	44,63***	39,68***	14,81***	10,42**
- 1999	28,42***	23,06***	26,24***	20,15***
- 2000	31,96***	27,66***	23,04***	18,02***
- 2001	31,40***	26,78***	-0,98	-5,10
- 2002	34,76***	28,66***	-5,29	-10,30**
- 2003	44,44***	35,86***	-11,41**	-17,53***
- 2004	57,87***	48,11***	3,06	-4,62
Beobachtungen	7.234	7.234	7.234	7.234
(Fachgebiete)	(653)	(653)	(653)	(653)
Wald Chi²	1.106,01***	1.138,82***	956,55***	971,46***

[a] Tabelliert sind jeweils Prozenteffekte (Veränderungen der Antragszahlen in Prozent bei Änderung der unabhängigen Variablen um eine Einheit).
[b] Im Vorjahr des Fachgebiets an der jeweiligen Universität.
*** p<0,001; ** p<0,01; * p<0.05.

Tabelle 2.5: Wendepunkte für den Zusammenhang von Personalbeständen und (bewilligten) Anträgen, mittlere Personalbestände und Fachgebiete über dem Wendepunkt

	Wendepunkte[a]		Mittlerer Personalbestand pro Fachgebiet einer Univ.		Fachgebiete größer als Wendepunkt gestellte Anträge		Fachgebiete größer als Wendepunkt bewilligte Anträge	
	Gestellte Anträge	**Bewilligte Anträge**	**Mittelw.**	**SD**	**Anteil**	**Anzahl**	**Anteil**	**Anzahl**
Professuren	169,04	152,59	25,46	29,33	0,53	43	0,67	55
Mitarbeiter	757,54	819,03	122,12	123,55	0,37	30	0,27	22

[a]Jeweils basierend auf Negativ-Binomial Regressionsschätzungen mit Random Intercepts.

Diese Befunde halten im Wesentlichen auch einer Disaggregation nach Fachgebieten stand (hier nicht dargestellte Schätzungen der Modelle 1 bis 4 pro Fachgebiet): Hinsichtlich der Professuren liegen nach diesen fachgebietsspezifischen Berechnungen maximal in der Mathematik 8,6 Prozent der Fachgebiete jenseits des kritischen Wendepunkts für bewilligte Anträge, in der Mehrzahl der Fachgebiete sind es deutlich unter 5 Prozent. In allen Fachgebieten liegen die Wendepunkte weit über den mittleren Personalbeständen. Exemplarisch ist dies in den Abbildungen 2.1 und 2.2 für die bewilligten Anträge in den Fachgebieten der Geistes- und Sozialwissenschaften sowie Naturwissenschaften dargestellt. Die senkrechten Linien kennzeichnen dort die fachgebietsspezifischen Durchschnittswerte, die Endpunkte der Kurven die jeweiligen maximalen Bestände an Professuren. Wie leicht ersichtlich ist, liegen die Wendepunkte der Kurven (und damit die Personalbestände, bei denen sich ein positiver Effekt von Personalzuwächsen auf die Bewilligungen in einen negativen verkehrt) weit über den Durchschnittswerten. Die geringen Überlappungen zwischen den Kurven einzelner Fachgebiete weisen zudem auf deutliche disziplinäre Unterschiede hin, was globale Betrachtungen ohne fachliche Differenzierungen kritisch erscheinen lässt.

Abb.2.1: Professuren und geschätzte Anzahlen bewilligte Anträge in den Geistes- und Sozialwissenschaften

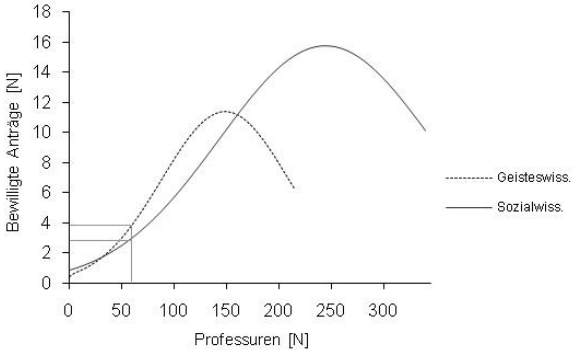
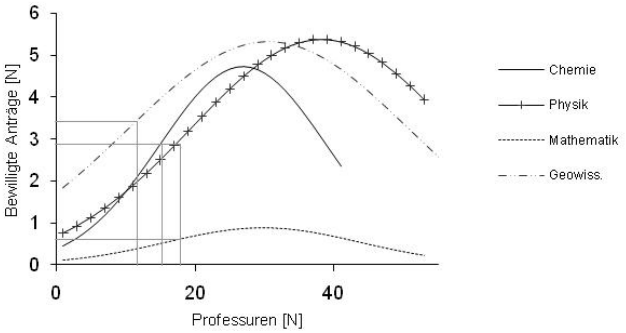


Abb. 2.2: Professuren und geschätzte Anzahlen bewilligte Anträge in den Naturwissenschaften



Bei Berechnungen auf Basis aller Mitarbeitenden sind die fachgebietsspezifischen Wendepunkte ähnlich weit von den Durchschnittswerten entfernt. Maximal haben in der Biologie 18,1 Prozent der Fachgebiete den Wendepunkt für bewilligte Anträge überschritten, bei der Hälfte der Fachgebiete sind

es unter 6 Prozent. Wenngleich also Evidenzen für einen abnehmenden Grenznutzen bestehen, treten negative Effekte erst bei weit überdurchschnittlichen Institutionengrößen auf.[25]

2.4.3 Ungleichheitsverteilung von Anträgen und Bewilligungen auf Universitäten im Zeitverlauf

Für die Chancen von Anträgen haben wir vereinzelt kontextuelle Einflüsse gefunden, die zu einer zunehmenden Ansammlung von Drittmitteln an einzelnen Universitäten führen könnten. Etwa werden aus Traditionsuniversitäten durchschnittlich mehr Anträge als aus anderen Universitäten eingereicht und weisen diese zugleich etwas bessere Bewilligungschancen auf. Die Mobilität von Wissenschaftlern und weitere, bislang nicht kontrollierbare Einflussfaktoren könnten einer zunehmenden Drittmittelkonzentration aber entgegenwirken oder sie im Gegenteil noch verstärken (s. etwa Münch 2008: 76 zu entsprechenden Argumenten). Abschließend soll daher der Frage einer wachsenden Ungleichverteilung auf der Aggregatebene von Universitäten nachgegangen werden.[26] Kristallisieren sich in der deutschen Hochschullandschaft zunehmend einzelne „Forschungsriesen" heraus und werden Mitbewerber aus dem Drittmittelmarkt gedrängt?

Unsere Messeinheiten bilden wiederum die zehn verschiedenen Fachgebiete in den einzelnen Universitäten und Jahren der Antragstellung. Anhand des Gini-Koeffizienten lässt sich das Ausmaß der Ungleichverteilung von Anträgen (bzw. Bewilligungen) auf diese Forschungseinheiten beziffern. Der Wert von Null würde eine völlige Gleichverteilung anzeigen: Es werden etwa 25 Prozent der Anträge von 25 Prozent der Universitäten eingereicht; beim Wert von Eins entfallen dagegen alle Antragstellungen auf eine einzige Universität. Mehr als das Niveau interessieren jedoch die Veränderungen der Ungleichverteilungen im Zeitverlauf. In den Abbildungen 2.3 bis 2.6 sind daher pro Fachgebiet die Gini-Koeffzienten der bewilligten Anträge für die einzelnen Jahre unseres Untersuchungszeitraums abgetragen.

Keinesfalls lässt sich eine eindeutige Tendenz hin zu einer zunehmenden (oder abnehmenden) Ungleichverteilung ausmachen. Lediglich für einzelne Fachgebiete zeichnet sich für die späteren Jahre des Untersuchungszeitraums eine leichte Zunahme der Ungleichheit ab (etwa in denen der Geistes- und Sozialwissenschaften oder in der Biologie ab 2002). Gerade angesichts der wellenförmigen Gesamtverläufe handelt es sich möglicherweise aber auch hierbei nur um kurzfristige (zufällige) Schwankungen. Für die Annahme eines Trends ist es selbst in diesen Fachgebieten verfrüht. Die Ungleichverteilung reicht im Jahr 2004 überhaupt nur in der Hälfte der Fachgebiete (fünf von zehn)

[25] So werden auch bei überdurchschnittlichen Personalbeständen in der Regel noch Elastizitäten über dem Wert von Eins erreicht. Prozentuale Zuwächse an Mitarbeitern setzen sich also selbst bei überdurchschnittlichen Fachgebietsgrößen noch in mindestens ebenso hohe prozentuale Zuwächse an (bewilligten) Anträgen um. (Die entsprechenden Berechnungen sind aus Platzgründen nicht dargestellt.)

[26] Ob es zu einer zunehmenden Ungleichverteilung kommt, ist nicht zuletzt im Hinblick auf die soeben gefundenen Wendepunkte optimaler Personalgrößen von Interesse. Starke Konzentrationstendenzen könnten in Zukunft zu höheren Anteilen überoptimaler Einrichtungsgrößen führen, während zugleich andere Institute an zu wenig „Masse" für größere Vorhaben „kranken" (ähnlich Jansen et al. 2007: 144).

und selbst dort nur geringfügig über das in 1991 zu findende Maß hinaus. Sehr ähnlich sind die Verläufe bei den Antragstellungen, es finden sich ebenfalls keine Anzeichen einer *stetigen* Zunahme der Ungleichverteilung, weshalb wir von eigenen Abbildungen absehen. Ein Matthäus-Effekt zunehmender Selbstverstärkung von Drittmitteln ist also auf der Aggregatebene von Universitäten und für das Normalverfahren der DFG kaum auszumachen, vielmehr ist ein *konstant* sehr hohes Ausmaß der Ungleichverteilung zwischen Institutionen zu beobachten. Die Analysen der vorherigen Abschnitte und die hier nicht dargestellte, ähnlich ungleiche Verteilung der Antragstellungen sprechen dafür, dass diese Niveauunterschiede primär Unterschieden in der Antragsaktivität (und nicht institutionenspezifischen Bewilligungschancen) geschuldet sind. Die Analyse möglicher Gründe bleibt einem umfassenderen Datenmaterial vorbehalten, sie könnten etwa in differenten Grundausstattungen und Verfügbarkeiten alternativer Drittmittelgeber zu suchen sein.

Abb. 2.3: Gini-Koeffizienten der bewilligten Anträge in den Geistes- und Sozialwissenschaften (1991 bis 2004, nur Universitäten mit mindestens einem bewilligten Antrag im jeweiligen Fachgebiet und Jahr)

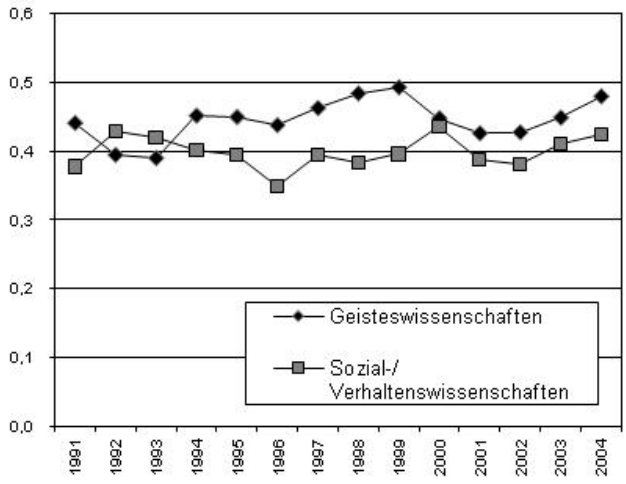


Abb. 2.4: Gini-Koeffizienten der bewilligten Anträge in den Lebenswissenschaften (1991 bis 2004, nur Universitäten mit mindestens einem bewilligten Antrag im jeweiligen Fachgebiet und Jahr)

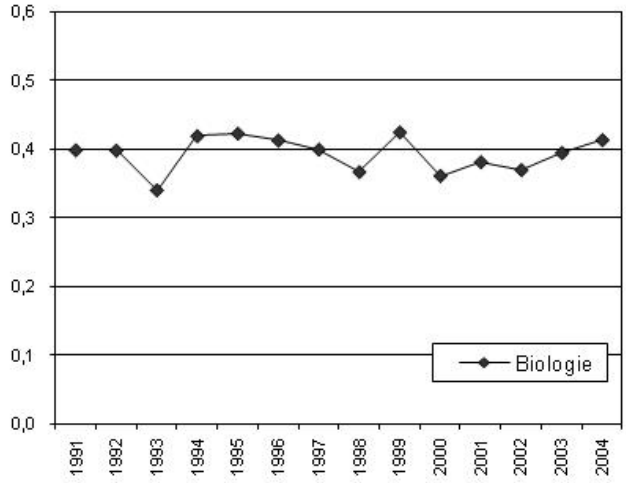


Abb. 2.5: Gini-Koeffizienten der bewilligten Anträge in den Naturwissenschaften (1991 bis 2004, nur Universitäten mit mindestens einem bewilligten Antrag im jeweiligen Fachgebiet und Jahr)

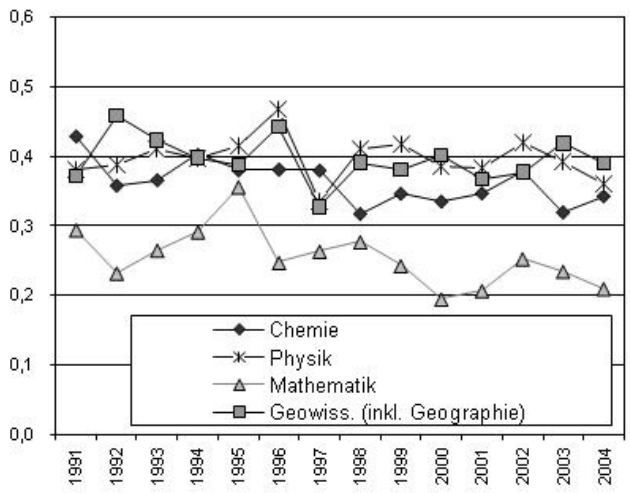
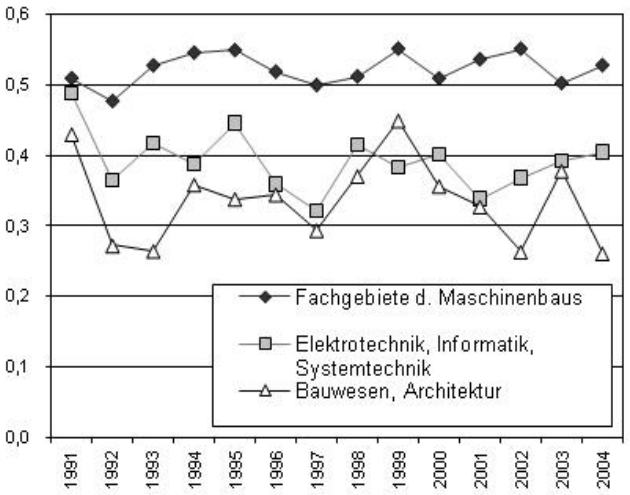


Abb. 2.6: Gini-Koeffizienten der bewilligten Anträge in den Ingenieurwissenschaften (1991 bis 2004, nur Universitäten mit mindestens einem bewilligten Antrag im jeweiligen Fachgebiet und Jahr)



2.5 Diskussion und Ausblick

Die Thesen von Richard Münch zur Herausbildung einer akademischen Elite haben, wie hier nur angedeutet werden konnte, eine sehr lebhafte Diskussion im Wissenschaftssystem und in der Öffentlichkeit ausgelöst. Darin liegt sicher ein unzweifelhafter Verdienst, schließlich beeinflussen die Verfahren der Exzellenzuweisung die Universitätslandschaft nachhaltig. Gleichzeitig fordern die Thesen geradezu Widerspruch heraus. Viele der bei Münch angedeuteten Mechanismen laufen auf sich selbst verstärkende Wege der Exzellenzzuweisung hinaus, die mit der wissenschaftlichen Leistung nichts mehr zu tun haben. Die hier vorgenommenen Analysen versuchen die Thesen von Münch auf das in Anteilswerten und Fördersummen wichtigste DFG-Programm, das so genannte Normalverfahren zu beziehen. Dabei werden prozessproduzierte Antragsdaten, wie sie in der Antragsbearbeitung entstehen, herangezogen. Diese Anträge werden von Wissenschaftlerinnen und Wissenschaftlern an Universitäten und Forschungsinstituten verfasst und eingereicht. Beurteilt werden sie in einem *peer-review*-Verfahren, das, wenn Münchs Überlegungen zutreffen, anfällig ist für die von ihm beschriebene Ausdifferenzierung einer akademischen Elite von wenigen drittmittelstarken, aber nicht *per se*

leistungsfähigeren Universitäten. Damit ist aber auch klar, was die hier angestellten Analysen nicht leisten können und wollen: eine Auseinandersetzung mit den Programmen der Exzellenzinitiative oder anderer Verbundforschungsprojekte, die im medialen Echo der Aufhänger für die Diskussionen um Münchs Buch waren. Die Untersuchung von Münch nutzt im empirischen Teil jedoch gleiche oder ähnliche Datenbestände wie wir. Wichtige Aussagen von Münch werden hinsichtlich der Anträge im Normalverfahren rekonstruiert und in einige prüfbare Hypothesen auf Individual- und Aggregatebene überführt. Insbesondere untersuchen wir den von Münch behaupteten Zusammenhang von Kontextfaktoren wie Standortgröße und Reputation von Universitäten mit einer zentralen Exzellenzzuweisung durch das Wissenschaftssystem. Sicher sind die Auswirkungen der neuen Art von Verbundförderung sowie die der bisherigen Programmförderung (etwa von Sonderforschungsbereichen) ebenfalls von größtem wissenschaftssoziologischem Interesse. Diese Auseinandersetzungen müssen aber an anderer Stelle geführt werden. Worum es hier geht ist Folgendes: Sind zentrale Hypothesen der Münch`schen Argumentation im Bereich der DFG-Normalverfahren haltbar?

Die empirische Lage ist eindeutig. Für die meisten rekonstruierten Hypothesen zur akademischen Elite fand sich keine oder nur eine schwache Evidenz. Dies gilt vor allem für die bei Münch herausgestellten Zusammenhänge von Standortgröße und Reputation. Dabei werden jeweils einfache Variablen zu deren Messung verwendet: die Anzahl von Mitarbeitenden bzw. Professuren in den Fachgebieten der Universitäten und die Forschungsstärke nach dem DFG-Ranking in den Vorjahren der Antragstellung. Die von Münch thematisierten, umgekehrt u-förmigen Effekte von Standortgröße auf die Leistungsfähigkeit zeigen sich für die Zahl von Antragstellungen und Bewilligungen nur für exorbitant große Einrichtungen bzw. Fachgebiete an Universitäten. An dieser Stelle können wir aus Gründen der Anonymität der Antragsverfahren keine Outputanalysen (Auswirkungen auf Veröffentlichungen und Patente) anstellen. Diese wären zur weiteren Prüfung der Münch`schen Thesen wünschenswert und zumindest für einige Fachgebiete, bei denen Forschungsresultate in Aufsatzform veröffentlicht und in entsprechende Datenbanken aufgenommen werden, auch durchführbar. Festzuhalten ist aber vor allem, dass auch bei den individuellen Bewilligungschancen die Kontextfaktoren Größe und Reputation keine bedeutende Rolle spielen.

Von den individuellen Eigenschaften der Antragstellenden weist das Lebensalter einen unerwarteten Effekt auf. Jüngere Wissenschaftler haben tendenziell höhere Bewilligungschancen für ihre Anträge als erfahrene Kollegen. Vor allem zeigen die Auswertungen, dass es zwischen Alter und Reputation nicht die von Münch erwartete Wechselwirkung gibt. Es ist wiederum umgekehrt: Es sind die älteren und erfahrenen Wissenschaftler, welche an größeren und angesehenen Universitäten erhöhte Bewilligungschancen haben. Jüngere Wissenschaftler profitieren von Größe und Ansehen der Universitäten nicht. Das Gleiche gilt für den vermuteten Interaktionseffekt von Reputation und Jahr der Antragstellung. Bei höherer Konkurrenz zeigt sich kein stärkerer Einfluss der Reputation. Dieses Ergebnis deutet darauf hin, dass auch bei deutlich gestiegenen Antragszahlen im Verfahren nicht auf einfache Signale wie das Renommee von Universitäten zurückgegriffen wird. Die DFG-Antrags-

bewertung erscheint gegenüber dem durch das Ranking erzeugten Reputationsmechanismus unempfindlich.

Es gibt durchaus einige Resultate, die weiterer Aufmerksamkeit bedürfen. Ostdeutsche Universitäten bieten ein Umfeld, in dem DFG-Anträge im Mittel weniger erfolgreich sind als in westdeutschen Universitäten. Für diesen Kontexteffekt können verschiedene Ursachen verantwortlich sein, die in tiefergehenden Analysen zu untersuchen wären. Zunächst könnten Unterschiede in der durchschnittlichen wissenschaftlichen Qualität der Anträge der beschäftigten Wissenschaftler bestehen. An verschiedenen Stellen haben wir auf mögliche Selektionseffekte hingewiesen, die letztlich auch in der Tradition der Merton-These stehen. Westdeutsche Universitäten könnten die im Durchschnitt produktiveren Wissenschaftler beschäftigen, möglicherweise auch wegen eines allgemein günstigeren Umfeldes (z. B. Gehaltsstrukturen). Eine alternative, aber damit eng zusammenhängende Überlegung würde auf eine geringere universitäre Grundausstattung in Ostdeutschland Bezug nehmen. Die laufenden Grundmittel liegen in den neuen Bundesländern für Universitäten unterhalb denen der alten Bundesländer (pro wissenschaftlicher Personalstelle, Professur oder Studierendem berechnet; vgl. Statistisches Bundesamt 2007: insb. 66). Eine geringere Bewilligungschance in Ostdeutschland ist möglicherweise auch diesen verhältnismäßig ungleichen Startchancen geschuldet.

Sehr deutlich fallen die Ergebnisse auf der Aggregatebene aus: Eine Konzentration von Antragstellungen und Bewilligungen auf einzelne Universitäten im Zeitverlauf ist nicht erkennbar. Zwar besteht eine deutliche, durch weitere Studien zu analysierende, Ungleichverteilung (erfolgreicher) Antragstellungen auf einzelne Universitäten, ihr Niveau bleibt aber in dem von uns betrachteten 14-Jahreszeitraum unverändert. Diesen Befund deuten wir als klaren empirischen Beleg gegen den Generalverdacht, wonach Verfahren der Exzellenzzuweisung zunehmend von „Machtkartellen“ geprägt sind. Das hier betrachtete Normalverfahren ist nicht durch Vetternwirtschaft und Machtkartelle gesteuert, sondern scheint weitgehend den vom Wissenschaftssystem propagierten Normen zu folgen.

Literatur

Allison, Paul D./Long, J. Scott (1990): Departmental Effects of Scientific Productivity. American Sociological Review 55, 469-478.

Allison Paul D./Long, J. Scott/Krauze, Tad K. (1982): Cumulative Advantage and Inequality in Science. American Sociological Review 47, 615-625.

Allmendinger, Jutta/Hinz, Thomas (2002): Programmierte (Un-)Gleichheit? Geschlechtsspezifische Chancen bei der Bewilligung von Forschungsanträgen. Zeitschrift für Soziologie 31, 275-293.

Bakanic, Von/Mc.Phail, Clark /Simon, Rita J. (1987): The Manuscript Review and Decision-Making Process. American Sociological Review 52, 631-642.

Behnke, Joachim (2005): Lassen sich Signifikanztests auf Vollerhebungen anwenden? Einige essayistische Anmerkungen. Politische Vierteljahreszeitschrift 46, 1-15.

Broder, Ivy E. (1993): Review of NSF Economics Proposals: Gender and Institutional Patterns. American Economic Review 83, 964-970.

Cameron, Colin A./Triverdi, Pravin K. (1986): Econometric models based on count data: Comparisons and applications of some estimators and tests. Journal of Applied Econometrics 1, 29-53.

Carayol, Nicolas/Matt, Mireille (2004): Does research organization influence academic productivity? Laboratory level evidence from a large European university. Research Policy 33, 1081-1102.

Cole, Jonathan R./Cole, Stephen (1979): Which researcher will get the grant? Nature 279, 575-576.

Deutsche Forschungsgemeinschaft (DFG) (2006): Förder-Ranking 2006. Institutionen – Regionen – Netzwerke. DFG-Bewilligungen und weitere Basisdaten öffentlich geförderter Forschung. Bonn: Deutsche Forschungsgemeinschaft.

Deutsche Forschungsgemeinschaft (DFG) (2005): Aufbau und Aufgaben. Bonn: Deutsche Forschungsgemeinschaft.

Deutsche Forschungsgemeinschaft (DFG) (2003): Förder-Ranking 2003. Institutionen – Regionen – Netzwerke. DFG-Bewilligungen und weitere Basisdaten öffentlich geförderter Forschung. Bonn: Deutsche Forschungsgemeinschaft.

Deutsche Forschungsgemeinschaft (DFG) (2000): DFG-Bewilligungen an Hochschulen und außeruniversitäre Forschungseinrichtungen 1996 bis 1998. Bonn: Deutsche Forschungsgemeinschaft.

Deutsche Forschungsgemeinschaft (DFG) (1997): Bewilligungen an Hochschulen 1991 bis 1995. Bonn: Deutsche Forschungsgemeinschaft.

DiPrete, Thomas A./Eirich, Gregory M. (2006): Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments. Annual Review of Sociology 32, 271-297.

Domen, Dieter (2007): Perspektiven für den Hochschulraum Ost. Zusammenfassende Thesen eines Projekts im Auftrag des Bundesministeriums für Verkehr, Bau und Stadtentwicklung. FiBS-Forum Nr. 35. Berlin: Forschungsinstitut für Bildungs- und Sozialökonomie.

Fahrmeir, Ludwig/Künstler, Rita/Pigeot, Iris/Tutz, Gerhard (2003): Statistik. Der Weg zur Datenanalyse. Berlin u. a.: Springer.

Fox, Mary Frank (1983): Publication Productivity among Scientists: A Critical Review. Social Studies of Science 13, 285-305.

Gillett, Raphael (1991): Pitfalls in Assessing Research Performance by Grant Income. Scientometrics 22, 253-263.

Greene, William H. (2003): Econometric Analysis. Upper Saddle River, NJ: Prentice Hall.

Güdler, Jürgen (2003): Kooperationsnetzwerke in der Forschung – Entstehung, Struktur und Wirkung am Beispiel der Soziologie. Forschungsberichte, Band 5. Bonn: Informationszentrum Sozialwissenschaften.

Güdler, Jürgen (1996): Dynamik der Medienforschung – Eine szientometrische Analyse auf der Grundlage sozialwissenschaftlicher Fachdatenbanken. Forschungsberichte, Band 1. Bonn: Informationszentrum Sozialwissenschaften.

Hausman, Jerry/Brownwyn, H. Hall/Griliches, Zvi (1984): Econometric Models for Count Data with an Application to the Patents-R&D Relationship. Econometrica 52, 909-938.

Hinz, Thomas/Findeisen, Ina/Auspurg, Katrin (2008): Wissenschaftlerinnen in der DFG. Förderprogramme, Förderchancen und Funktionen (1991–2004). Bonn/New York: Wiley.

Hornbostel, Stefan (2001): Third party funding of German universities. An indicator of research activity? Scientometrics 50, 523-537.

Hornbostel, Stefan (1997): Wissenschaftsindikatoren. Bewertungen in der Wissenschaft. Opladen: Westdeutscher Verlag.

Jaeger, Michael (2006): Leistungsbezogene Budgetierung an deutschen Universitäten. Wissenschaftsmanagement 3, 30-36.

Jansen, Dorothea/Wald, Andreas/Franke, Karola/Schmoch, Ulrich/Schubert, Torben (2007): Drittmittel als Performanzindikator der wissenschaftlichen Forschung. Zum Einfluss von Rahmenbedingungen auf Forschungsleistung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 59, 125-149.

Kaube, Jürgen (2007): Abrechnung ohne Strom. Richard Münchs Polemik gegen unser Forschungssystem. Rezension in der Frankfurter Allgemeinen Zeitung vom 17.08.2007.

Koch, Stefan (2006): Die Deutsche Forschungsgemeinschaft und die Reform ihres Begutachtungssystems: Zur Einführung der Fachkollegien. Wissenschaftsrecht 39, 25-47.

Laudel, Grit (2005): Quality-only assumption. Is external research funding a valid indicator for research performance? Research Evaluation 14, 27-34.

Long, Scott J./Freese, Jeremy (2006): Regression Models for Categorical Dependent Variables Using Stata. Second Edition. Texas: Stata Corp.

Long, Scott J./McGinnis, Robert (1981): Organizational Context and Scientific Productivity. American Sociological Review 46, 422-442.

Merton, Robert K. (1988): The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property. ISIS 79, 606-623.

Merton, Robert K. (1968): The Matthew Effect in Science. Science 57, 68-72.

Moed, Henk F. (2005): Citation Analysis in Research Evaluation. Dordrecht: Springer.

Moed, Henk F./Luwel, Marc/Houben, Josee A./Spruyt Eric/Van Den Berghe, Herman (1998): The Effects of Changes in the Funding Structure of the Flemish Universities on their Research Capacity, Productivity and Impact during the 1980's and early 1990's. Scientometrics 43, 231-255.

Münch, Richard (2008): Stratifikation durch Evaluation: Mechanismen der Konstruktion von Statushierarchien in der Forschung. Zeitschrift für Soziologie 37, 60-80.

Münch, Richard (2007): Die akademische Elite. Zur sozialen Konstruktion wissenschaftlicher Exzellenz. Frankfurt a. M.: edition suhrkamp.

Neidhardt, Friedhelm (1988): Selbststeuerung in der Forschungsförderung. Das Gutachterwesen der DFG. Opladen: Westdeutscher Verlag.

Okrasa, Wlodzimierz (1986): Differences in Scientific Productivity of Research Units: Measurement and Analysis of Output Inequality. Scientometrics 12, 221-239.

Schimank, Uwe (1992): Forschungsbedingungen der Professoren an den westdeutschen Hochschulen. Daten aus einer Befragung im Wintersemester 1990/91. MPIFG Discussion Paper 92/2. Köln: MPI für Gesellschaftsforschung.

Spence, Michael (1973): Job Market Signaling. The Quarterly Journal of Economics 87, 355-374.

Statistisches Bundesamt (2007): Monetäre hochschulstatistische Kennzahlen. Fachserie 11, Reihe 4.3.2. Wiesbaden: Statistisches Bundesamt.

Statistisches Bundesamt (2005a): Monetäre hochschulstatistische Kennzahlen. Fachserie 11, Reihe 4.3.2. Wiesbaden: Statistisches Bundesamt.

Statistisches Bundesamt (2005b): Finanzen der Hochschulen. Fachserie 11, Reihe 4.5. Wiesbaden: Statistisches Bundesamt.

Wissenschaftsrat (2000): Drittmittel und Grundmittel der Hochschulen 1993 bis 1998 (Drs. 4717/00). Köln: Geschäftsstelle des Wissenschaftsrats.

Wooldridge, Jeffrey M. (2003): Introductory Econometrics. A Modern Approach. Mason, Ohio: Thomson.

Zuckermann, Harriet/Merton, Robert K. (1971): Patterns of Evaluation in Science. Minerva 9, 66-100.

Methodischer Anhang

Bei der Analyse der Bewilligungschancen (Abschnitt 2.4.1) tritt das Problem auf, dass in der Regel aus einzelnen universitären Fachgebieten jeweils mehrere Anträge eingereicht werden, diese somit nicht unabhängig voneinander sind: Unkontrollierte institutionelle Einflussfaktoren auf die Bewilligungschancen (wie z. B. unterschiedliche universitäre Grundausstattungen) führen bei gewöhnlichen Regressionsschätzungen zu einer Korrelation der Fehlerterme, damit inkorrekten Standardfehlern (vgl. Greene 2003; Wooldridge 2003: 469ff.). Der von uns gewählte Ausweg besteht in der gezielten Modellierung der vorliegenden Mehrebenenstruktur in Form von Random Intercept Modellen. Bei diesen wird für jede Einrichtung (hier Fachgebiet einer Hochschule) ein spezifischer Term aufgenommen. Die Schätzgleichung lautet:

$$LOGIT_{ij} = \ln [P_{ij}/(1-P_{ij})] = \beta_0 + \beta X_{ij} + \gamma Z_j + \nu_j$$

P_{ij} ist die Wahrscheinlichkeit der Bewilligung des Antrags i aus der Einrichtung j, der $LOGIT_{ij}$ das logarithmierte Verhältnis der Wahrscheinlichkeit einer Bewilligung zur Wahrscheinlichkeit einer Ablehnung. X_{ij} stellt einen Vektor mit den individuellen Merkmalen des Antragsstellenden (Alter und Geschlecht) dar, Z_j ein Vektor mit den Merkmalen der Einrichtung (etwa deren Personalbestand). ν_j ist der einrichtungsspezifische Term, der hier im Gegensatz zu den verwandten *Fixed Effects* Modellen als Zufallsvariable und damit als ein einrichtungsspezifischer Fehlerterm aufgefasst wird. Er modelliert vereinfacht gesprochen die unbeobachtete Heterogenität zwischen Einrichtungen, was verlässlichere Schätzungen des Einflusses erklärender Variablen erlaubt. Die einzelnen Antragsjahre lassen sich zusätzlich als Dummy-Variablen (damit fixe Effekte) in dieses Modell integrieren, um die zeitliche Veränderung der Bewilligungschancen abzubilden (so genanntes *mixed model* – da es sich um eine Vollerhebung aller Antragsjahre im betrachteten Zeitraum handelt, erscheint hier eine Modellierung als *feste* Effekte der Datenstruktur angemessener). Wir berichten anstatt der Koeffizientenwerte anschaulichere marginale Prozenteffekte bzw. *discrete changes*: Diese geben die durchschnittliche prognostizierte Veränderung der Bewilligungswahrscheinlichkeit wieder, die aus einer Erhöhung der jeweiligen erklärenden Variablen um eine Einheit resultiert (Berechnung mithilfe des *ado* "margeff" des Statistik-Programms Stata 10).

Bei den in Abschnitt 2.3.2 betrachteten Antragszahlen und Bewilligungen pro Fachgebiet handelt es sich um ganzzahlige positive Größen, damit einen Paradefall für Zähldatenmodelle. Die hier einschlägigen und in ihrer „Schlichtheit" attraktiven Poisson-Regressionen setzen das Vorliegen von *Äqui-Dispersion* voraus (proportionaler Anstieg der Varianz zum Anstieg der Mittelwerte) – eine Annahme, die in unseren Modellen oftmals verletzt ist (entsprechende Likelihood-Ratio-Tests zeigen hochsignifikant das Vorliegen von *Over-Dispersion* an; s. Long/Freese 2006: 376f. zu den verwendeten Teststatistiken). Die bei solchen Annahmeverletzungen von uns alternativ verwendeten Negativ-Binomial Modelle unterscheiden sich durch die zusätzliche Aufnahme eines Dispersionsparameters α

(für Einzelheiten: Long/Freese 2006: 372-380). Die prognostizierte Anzahl μ_{jk} der Anträge (bzw. Bewilligungen) der Einrichtung j im Jahr k und ihre Varianz ergeben sich dann als:

$$\mu_{jk} = \exp(\beta_0 + \gamma Z_{jk} + \alpha_{jk}); \quad Var(y_{jk} \mid x_{jk}) = \mu_{jk} + \alpha_{jk}\, \mu_{jk}^2$$

Bei den berichteten Werten handelt es sich um einfache Umrechnungen der Form $(\exp(\gamma)-1)\cdot 100$. Sie messen die erwartete Veränderung der Antragszahlen in Prozent, wenn sich die jeweilige erklärende Variable um eine Einheit erhöht. Wir führen wiederum Maximum-Likelihood-Schätzungen mit Random Intercepts durch, um der Tatsache Rechnung zu tragen, dass einzelne Einrichtungen mehrfach (nämlich jeweils pro einzelnem Antragsjahr) beobachtet werden (dazu Greene 2003: 747-749; Hausman et al. 1984; allgemein zur Modellwahl bei Zähldaten: Cameron/Triverdi 1986).[27]

Bei den im letzten Abschnitt des Ergebniskapitels (2.4.3) herangezogenen Gini-Koeffizienten handelt es sich um ein klassisches Konzentrationsmaß. Ausgangspunkt für seine Berechnung ist die so genannte „Lorenzkurve“. Sie zeigt an, auf welchen Anteil der Einrichtungen jeweils welcher Anteil der Gesamtsumme an Anträgen und Bewilligungen entfällt. Der Gini-Koeffzient erfasst dann die Abweichung von der eine Gleichverteilung indizierenden Winkelhalbierenden. Beim Wert ‚Null‘ liegen alle Werte auf der Winkelhalbierenden, es besteht eine völlig Gleichverteilung von Anträgen bzw. Bewilligungen auf die Universitäten. Beim Wert ‚Eins‘ konzentrieren sich dagegen sämtliche Anträge bzw. Bewilligungen auf eine einzige Universität, die Lorenzkurve ist maximal ausgebuchtet und von der Winkelhalbierenden entfernt. Wir verwenden den für kleine Fallzahlen auf den Wertebereich [0;1] normierten Gini-Koeffizienten (so genannter „Lorenz-Münzner-Koeffizient“; s. dazu Fahrmeir et al. 2001: 82). Insbesondere die Eigenschaft der Skaleninvarianz spricht für die Wahl des Gini-Koeffizienten: Die jahresweise Zunahme der absoluten Antragszahlen beeinträchtigt den Zeitvergleich nicht.

[27] Unsere Daten weisen die Struktur von Paneldaten auf. Random Intercept Modelle entsprechen methodisch aber exakt Paneldaten-Modellen, bei denen für jede über die Jahre beobachtete Einheit (hier Fachgebiete der einzelnen Universitäten) ein zufälliger Term mit aufgenommen wird. Der Fehlerterm setzt sich also genau genommen aus $(\alpha_{jk} + \nu_j)$ zusammen.

3

Migration Decisions Within Dual-Earner Partnerships: A Test of Bargaining Theory

(Martin Abraham, Katrin Auspurg, and Thomas Hinz)

3 Migration Decisions Within Dual-Earner Partnerships: A Test of Bargaining Theory

Abstract

This paper focuses on the problems faced by dual-earner partnerships arising from the regional coordination of the partners' careers. It is still unknown whether the fact that couples are less mobile than singles is caused by homogenous preferences within couples or by a process of balancing conflicting interests. Consequently, we analyze the potential conflicts provoked by work-related migration incentives. Hypotheses derived from bargaining theory are tested using quasi-experimental data from a factorial survey of nearly 280 European couples. Our results support the bargaining approach and confirm that the potential for conflict is driven by asymmetrical shifts in bargaining power. Women's willingness to move is generally lower than men's, but the impact of employment prospects differs only slightly by gender.

3.1 Introduction

It is well known that in modern societies, changing employment patterns affect couples in various ways. In particular, the increasing labor market participation of women leads to new challenges for couples living together. One of these challenges arises from the necessity to coordinate two careers in time and space. For example, the probability of migration drops significantly when people form a couple (McHugh et al. 1990; Quigley/Weinberg 1977). The migration literature offers a wide range of empirical findings as well as theoretical explanations regarding the determinants of household migration (Bartel 1982; W. Bielby/D. Bielby 1992; Shauman/Noonan 2007). These findings, however, focus only on actual migration and not on the underlying considerations for or against a household move. As a consequence, there is inadequate knowledge regarding bargaining and conflicts within couples. Moreover, migration research based on *classic* survey data is usually biased towards the traditional case of men as the main breadwinners. Women who receive a job offer with a considerably higher salary than their male spouses are still rare (Cooke 2006: 467 for Germany and the U.S.). Therefore it is difficult to disentangle gender and structural effects on migration behavior.

Looking at household migration from a theoretical point of view, most researchers stress the necessity of considering the conflicting interests of the two partners. Especially in dual-earner couples, new job opportunities for one partner may correspond with decreasing opportunities for his or her mate. In line with this argument, one of the dominant theoretical models for analyzing dual-earner households is bargaining theory. Within this framework, the partners have to bargain over: (a) the decision for or against migration, and (b) the arrangements that may follow after the household move. Although this theoretical model seems appropriate for the research problem at hand, there are astonishingly few empirical tests of it. One reason is the lack of paired data for couples that are

necessary to observe the diverging interests of partners. Another limitation arises from focusing only on the results of the bargaining process. That is, theorists have tried to derive hypotheses regarding the shift of resources allocated to each partner and the consequences of that shift (Bernasco/Giesen 2000; Klasen 1998; Lundberg/Pollak 1996). One of the main problems within this approach is the difficulty of empirically measuring the complete allocation of resources including immaterial resources such as love, emotional care and recreational time.

In this paper, we aim to fill these gaps by applying factorial survey analysis, a research design for combining survey research with an experiment. We requested couples in Germany and Switzerland to evaluate hypothetical descriptions of better employment opportunities in a distant location for one partner. This design gave us the opportunity to analyze to what extent moving preferences conflict within a partnership. By varying the hypothetical situations in an experimental manner, we were able to identify factors that influence the extent of potential conflicts arising from differing migration incentives. The aim was to complement the existing empirical knowledge on willingness to migrate within couples and to provide a direct test of bargaining models in family research.

3.2 Theoretical Framework

3.2.1 Bargaining Models

In family research, bargaining models are often used to analyze allocation problems among household members and especially among spouses or cohabitation partners (Bitman et al. 2003; Cooke 2006; Lundberg/Pollak, 1996). At the core of this theoretical framework is the situation in which two actors have to decide on the allocation of resources. These can refer to material resources, like household income, and to immaterial resources, such as the right to make decisions, enjoy leisure time, and avoid unpleasant household tasks. The result of this allocation decision depends on the relative bargaining power of each actor within the partnership.

This conceptual core can be found in different theoretical frameworks like classical exchange theory (Blood/Wolfe 1960), social psychology (Thibaut/Kelley 1959), and rational choice theory (Manser/Brown 1980). The crucial question within all these approaches is the definition of relative power. Although there are some differences, most researchers agree that, to a large extent, the available alternatives to the current partnership determine each actor's bargaining power (England/Farkas 1986: 56). The partner with better *outside options* has more power in the relationship because he or she can credibly threaten to leave the relationship. These outside options are also called *threat points*. Note that the bargaining process itself can stay implicit, that is, without any discussion between the partners, but it is assumed that both partners are aware of the power imbalance.

Although available alternatives may include the presence of other attractive partners (Bergstrom, 1996: 1929-1930), in modern societies, a person's economic standing is probably the most decisive determinant of structural bargaining power in relationships. Labor market participation and an inde-

pendent source of income make it possible for one partner to dissolve an unsatisfactory partnership. The next best alternative is often not the move into another relationship, but rather into a single household. Moreover, the higher the persons' income, the more they should be able to find an attractive alternative partner. Thus, most research focused on labor market opportunities for measuring bargaining power in partnerships (Blau et al. 2001: 43; Blood/Wolfe 1960; Ott, 1992).

3.2.2 The Decision to Move the Household

Within this theoretical framework, household moves can be seen as events that are likely to be affected by the relative bargaining power of exchange partners and vice versa (Lundberg/Pollak 2003). As Mincer (1978) explains, it is unlikely that both partners will receive optimal job offers at the same time and within the same region. Consequently, a couple has to decide whether the migration incentive for one partner (the *mover*) is sufficient for accepting, at least temporarily, a worse employment situation for the other.

Mincer coined the term *tied mover* for a partner who experiences individual disadvantages by a move that is generally beneficial for the household. In contrast, a *tied stayer* foregoes the better job because his or her individual gains will not be sufficient to compensate for the overall household loss. In Mincer's classical model the actors will accept any individual loss as long as the household utility is maximized. Hence this model fails to recognize the conflicts that result from the clash of individual interests between the partners.

From a dynamic bargaining point of view, even a higher household utility may not be sufficient for a potential tied mover to accept the migration. This is because within the bargaining theoretical framework, it is expected that the actors rationally anticipate future allocations. More concretely, the actors are assumed to consider how a new situation affects their relative bargaining power and thereby their individual benefit of a move. Under the assumption that the bargaining power is mainly influenced by the employment and income possibilities of the actors, by definition, a tied mover will experience a loss of relative power. If this loss is too high, the tied mover may get even a lower level of household resources at the new destination than he or she had before the move, although the overall household resources are increased by the move (Lundberg/Pollak 2003). Thus, potential tied movers may reject the household move whereas tied stayers may use their job offer to renegotiate the allocation of resources at the present location. The job offer can be seen as a new outside option because, theoretically, the tied stayer has an attractive alternative to the relationship consisting of the possibility of moving without the partner.

Bargaining models were most often used to predict the allocation of resources within a relationship (see Lundberg/Pollak 1996 for a review). In our opinion, however, measuring only the results of the bargaining process (such as the shares of family income used for *private* consumptions or the distribution of leisure times and cumbersome household tasks) is not the most promising strategy to test this kind of theory. First, the actual allocation of resources and duties may be influenced by

additional factors that are difficult to control for (e.g., individual preferences). Second, and more importantly, actors are trying to solve these kinds of cooperation problems using various mechanisms. These can be joint investments in the relationship such as children, formal marriage contracts and further commitment that settles mutual rights and obligations (Bernasco/Giesen 2000; Ott 1992). Relation-specific investments enhance the interdependence of the partners by increasing the sunk-costs of the current relationship and thereby promote further commitment that stabilizes the (trust in the) stability of the partnership (Brines/Joyner 1999). Non-binding agreements, in contrast, are considered insufficient. In intimate relationships, partners may (and often do) promise that, because of love, even if there is a shift in bargaining power, no renegotiation of resources will take place. But it is doubtful whether this promise is credible, as we all know that love is no guarantee against conflict in marriage or cohabitation. These mechanisms are not captured when only the results of bargaining are considered.

3.2.3 The Influence of Gender

Bargaining models are, in principal, gender neutral and this is the point that mostly arouses criticism (W. Bielby/D. Bielby 1992; Jürges 2005 for the literature on migrations). The bargaining process is influenced by different income options, and of course, the income opportunities differ by gender. But after controlling for those structural factors and for the outside options in general, the willingness to move is expected to be no longer gender-dependent. A gendered perspective in contrast suggests that the bargaining process also contributes to the production of gender, in other words, leads the individuals to affirm their categorical status of *male* or *female* (the process of *doing gender*, Fenstermaker 2002: 110; West/Zimmerman: 1987). Even if doing gender clearly means more than "to live up to normative conceptions of femininity or masculinity" (West/Zimmerman 1987: 136), gender-appropriate behavior should tend to, at least in general, be in line with culturally approved standards. Therefore the decision-making process is expected to differ by gender especially if the actor's structural positions are incongruent with the gender norms (Hook/Chalsani 2008: 979). Because these institutionalized norms still assign the main responsibility for the household income to the male partner (Huppé/Cyr 1997; Stickney/Konrad 2007; Szinovacz/Harpster 1993), the male employment options are more likely to take priority in the decisions compared to the female ones.

Empirically, the relative impact of the two alternative explanations – bargaining vs. subscription to gender norms – is still an unresolved issue (Blackburn 2006; Lichter 1983; Maxwell 1988; Mincer 1978; Spitze 1984). By reason of data limitations, it was difficult to decide whether gender disparities in household decision-making resulted from employment variables that were not controlled for and therefore corresponded with a gender-neutral bargaining process, or if they reflected normative expectations. The classifications of occupations typically used were too broad to capture the specific job and labor market characteristics that were unevenly distributed by gender (such as the geographic ubiquity and length of career ladders, see, e.g., Shauman/Noonan 2007: 1739). Thus, most results were in line

with both theoretical approaches. Our experimental design provided men and women with exactly the same employment opportunities, as will be explained in more detail below. This may help gain insight into the gender influence on migration decisions.

3.3 Hypotheses

The bargaining model allows us to derive several hypotheses regarding how actors make migration decisions in dual-earner relationships. When evaluating a migration opportunity, the partners will consider the expected changes in relative bargaining power. If the actors do indeed anticipate the shifts in bargaining power and if there is any interest in preserving the relationship, they will be less willing to move if one of them will be worse off at the new destination. This leads to our first hypothesis:

H_{1a}: *Individual willingness to move will depend not only on the changes in one's own situation, but also on the changes in the partner's situation.*

Results in favor of the first hypothesis constitute, of course, only weak evidence for the bargaining model because this phenomenon could also be explained by altruism. Altruistic actors care about the future situation of their partners; hence, they will take potential disadvantages for their partners into account. The bargaining approach, however, additionally suggests that every actor will consider his or her own situation more important than his or her partner's situation. Any consideration of one's partner's interests will be in vain if the partnership fails in the future. The risk of a future separation and the interest in receiving a favorable position in the resource allocation lead to a primary interest in preserving one's own bargaining power. Therefore, we can hypothesize that no matter who has the better option, the change in one's own bargaining power is a stronger predictor of the willingness to move than is the respective change in the partner's power.

H_{1b}: *Actors' willingness to move is influenced more by their own options than those of their partners.*

The general hypothesis regarding a couple's conflict potential is straightforward. All unexpected options to improve the relative bargaining power of one partner will question the negotiated arrangements and evoke clashing interests. The partner with the potential improvement is interested to move whereas the other one will reject the move to preserve the actual power balance. We term this clashing interest *conflict potential* because the partners may be able to actually resolve the conflict (see also Miller et al. 2007: 361 et seq.).

H_2: *An actor's potential increase in relative bargaining power due to a possible migration will lead to a higher conflict potential in the partnership.*

Note that, to our knowledge, this is a new hypothesis derived from bargaining theory that has yet to be tested. As outlined above, there are many reasons why couples may differ in the ways they use to resolve the conflict. These may result for example from certain psychological characteristics of the partners. As sociologists, we are more interested in structural factors. One type of solution arises from the strength of *each* actor's commitment to the partnership (Nock 1995). A credible commitment in this sense increases the threshold for leaving the partnership and therefore decreases an actor's bargaining power by reducing the utility of outside options. In families and relationships, there are three main types of *mutual* commitment: marriage, children, and common property (especially home ownership). We hypothesize that all three types of commitment lead to a lower level of conflict potential.

H_{3a}: *The conflict potential will be lower for couples who are married, have children, or hold real estate property than for couples without these characteristics.*

Couples will also differ in respect to the time they have had to learn about each other and the relationship. Learning about each other's characteristics and preferences allows partners to assess each other's future behavior. Consequently, the joint history represents a proxy for mutual commitment that is not easily captured such as learning to participate in and appreciate each others` hobbies. Additionally, the duration of partnership might reflect a positive selection process: Partnerships in which mutual expectations were more often fulfilled and in which therefore a higher trust has been established will have a smaller risk of breakdown. This grown knowledge of each other and amount of commitment might cause a belief that, all else being equal, no renegotiation of resources after the move will take place.

H_{3b}: *The conflict potential declines with the duration of the partnership.*

Finally, we consider the effect of gender on our research question. Within the bargaining framework, gender should matter only to the extent that one gender tends to have less bargaining power because of structurally worse outside options. A gendered perspective in contrast suggests that even when female partners have strong incentives to move, they may favor their male partners' careers to conform to normative gender ideologies. That is, we suppose an interaction effect of gender and labor market options on the willingness to move.

H_4: *Female partners are less willing to move in response to their own labor market options as compared to their male partners.*

3.4 Method

The insufficient empirical research on the bargaining model in the field of families and couples is mainly caused by a lack of adequate data. Existing survey data provided little or no information on the long-term career prospects and labor market situations at potential destinations (Shauman/Noonan 2007). Because of the low numbers of moving households, it was difficult to observe the shifts in resources allocations caused by household moves. Another problem is the high selectivity of migration: Persons with job-related migrations are generally more career-oriented and professionally successful than immobile persons, so they probably would have realized similar gains in income and relative bargaining power even without a move (Antel 1980; DaVanzo/Hosek 1981). Moreover, in general, only realized migration and employment opportunities were recorded and therefore the changes in threat points or the losses of household gains due to forgone opportunities are unknown. Conflicts and reasons for separation are difficult to assess with ordinary survey data and require paired information from both partners that is often not provided. An additional problem arises from the high correlation between employment characteristics and gender. Observed asymmetric decisions and outcomes may have arised only because the labor market conditions created better returns for the male partner and not because his career concerns were fostered to a greater extent (Jacobson/Levin 2000: 294).

For these reasons, we adopted a research approach that uses a factorial survey design (also known as *vignette analysis*). The key idea in this design is that the respondents react to hypothetical descriptions of situations or objects (*vignettes*) instead of answering single item questions (see Jasso 2006; Rossi/Anderson 1982). By independently varying a restricted number of dimensions of the vignettes in an experimental design, the exact impact of each dimension on respondents' judgments or decisions can be estimated. Therefore, the method is able to isolate the weight of single factors that are often confounded in reality. It is possible to separate the impact of gender and employment prospects by offering the same mobility incentives to male and female partners. The random allocation of vignettes leads to an overall independence of migration incentives and thereby overcomes the selection biases that limit the results of usual migration research.

In our study, we used vignettes to describe various incentives for a household move. As an innovation to the vignette design we interviewed both partners in dual-earner couples about inverse situations. Both partners were presented with the same set of situations, but one (*ego*) was given the role of a mover, the other (*alter*) was put in the position of a tied mover. The constant stimulus consisted of a job offer for ego in another geographical region that provided the couple similar living conditions. The varying dimensions simulated differently attractive employment prospects in this new location for the two partners. This design allowed us to study the reactions to changing threat points in the relationship by varying the respective employment options. More concretely, we systematically varied (a) the increase in income and career prospects for the mover, and (b) the employment and income prospects for the tied mover. By conducting computer assisted personal interviews (CAPI), we were able to

simulate the income prospects as a percentage gain of the actual salary, which made the situations as plausible as possible. In order to test our hypotheses, we assumed that ego's increase in income and career prospects improved his or her bargaining power in the relationship, whereas alter's employment prospects in the new destination affected the anticipated change in the bargaining power of ego's partner.

Concerning this assumption, two points are noteworthy: First, we manipulated the threat points of each partner independently, so it was possible that *both* actors' outside options would increase due to the move and the relative bargaining in the partnership would not change. Thereby, our design had the advantage that alter may profit from the move and is not restricted to the role of the tied mover. Second, we assumed that bargaining power is determined by the employment situations of *both* actors. Especially if one actor is engaged in highly specialized paid work, whereas the other is mostly engaged in housework, this assumption may be problematic. To address this problem, we restricted our sample to dual-earner couples. Figure 3.1 gives an example of the vignettes and shows the answer scale employed in this paper. The vignettes also described the commuting time and the commuting mode (train or car) to the new destination in order to make the situations as realistic as possible. We used these variables as controls. Within a couple, both partners were confronted with exactly the same hypothetical situations. The vignettes were completely identical, except one of the partners (ego) responded as the partner who received a job offer, whereas the other (alter) responded as the partner who did not receive the offer.

Fig. 3.1: Example of a Vignette (Male Perspective, Own Job Offer)

Assume, …

You are offered a net salary of EUR 2,400 at the new location. In the long run, the new job will provide you with no advancement opportunities. If you don't move, commuting to the new job will take 1 ½ hours each way. It is only possible to get there by car.

Your partner's chances of finding a job at the new location are small and her income prospects are higher there compared to the actual labor market.

How much would you like to take the job and move to the new location?

Not at all	1	2	3	4	5	6	7	8	9	10	11	Very much
	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	

Note: The inverse situation for the female partner in this example would read: Assume, your partner is offered a net salary of EUR 2,400 at the new location. […]. Your own chances of finding […]. How much would you like to move to the new location?

Our dependent variables were each partner's willingness to move and the conflict potential in a partnership that resulted from the migration incentive. The willingness to move scale scored from 1 = *not at all* to 11 = *very much*. The mirror-inverted design of our study allowed us to directly compare the alter's and ego's willingness to move for each vignette and to measure the conflict potential by taking the difference between ego's and alter's willingness to move. The resulting variable ranged from -10 = *alter would very much like to move but ego does not want to move at all* to +10 = *ego would very much like to move but alter does not want to move at all*. Zero indicated that both partners shared the same tendency for a move regardless of the strength of that preference. We treated all dependent variables as continuous. Several alternative statistical models for non-continuous variables (e.g., ordinal regressions, Poisson regressions) did not yield different results; therefore, we utilized the more straightforward interpretation of an Ordinary Least Square (OLS) model. Because each person evaluated up to ten different vignettes, the models had to be corrected for correlated observations, as is the case with any repeated measures in within-subject designs (Hox et al. 1991). We applied random intercept models (Raudenbush/Bryk 2001; Snijders/Bosker 1999) that accounted for dependent observations by estimating one joint random intercept for all observations from one single respondent or couple. The characteristics of the vignettes varied on level 1, and either individual or couple characteristics were modeled as level 2 variables.

The used sample of 200 vignettes was a fraction of all possible combinations of the six variable dimensions (maximization of the *D-Efficiency* with orthogonalization of all main effects and first-order-interactions in a so-called 'resolution IV' according to Kuhfeld 2005; Kuhfeld et al. 1994). Ten vignettes were presented to each person; that is, we used 20 different decks of vignettes and 80 different questionnaires (20 decks x 2 versions for mover vs. tied mover x 2 versions for men vs. women).

All interviews were conducted with couples living in Germany and Switzerland. Both countries represent conservative welfare regimes with gendered labor market participation rates, high occupational sex segregations and gender pay gaps. For instance, in 2007 in Germany only 64 percent of women between 15 and 64 years old were employed compared to 75 percent of men; the overall gender pay gap was 23 percent (European Commission 2009). In both countries, traditional gender role beliefs are persistent (Stickney/Konrad 2007: 807).

3.5 Results

3.5.1 Descriptive Results and Sample Description

First, we provide a brief overview of some descriptive results for our nearly 280 interviewed couples. Only heterosexual couples were included in the sample to address the hypotheses about gender and to reach a higher standardization of the sample. Two universities (Konstanz and Bern) collaboratively conducted the research project using the CAPI method. The sample was selected by the network contacts of our 60 interviewers. For the research target at hand, a convenient sample seemed adequate because the central requirement is a random allocation of the experimental stimuli to the respondents. This procedure guaranteed the independence of the vignettes' and respondents' characteristics that is crucial for a strict test of our hypotheses. For example, it ensured the independence of migration incentives from the respondents' attributes and therefore prevented the selectivity bias of non-experimental survey data on migrations. To restrict interviewer effects, the recruitment was restricted to an average of seven couples per interviewer. All interviews were conducted between June 2007 and February 2008. Independently, both partners in a couple responded to the vignettes and also provided some basic information about themselves and the partnership. This was guaranteed throughout by the presence of an interviewer who prevented mutual consultations. We over-sampled couples without children living in their household (couples without children represented nearly 90 percent of our sample) in order to gain adequate variance in the willingness to move; children are often seen as a strict obstacle to migration. Additionally, the sample was restricted to salaried employees living in one joint household on a daily basis (no commuting or long-distance relationships) because of the well-known fact that self-employed persons and older people are especially immobile (see Jürges 2005 for evidence on Germany). Moreover, we wanted to simulate a real decision making dilemma by eliminating the desire to establish a joint household as an incentive to move.

An interesting observation is that 32 percent of all egos reported zero willingness to move, although the minimum gain was a 30 percent higher income (see Figure 3.2). Obviously, income prospects were not the only determinant of migration decisions. This finding is in line with the general low tendency to move, as reported in the migration literature (Taylor 2007: 815). A further result worth noting is that in about 39 percent of the vignettes, alter was more willing to move than ego (the partner who actually got the job offer, see Figure 3.3).

Fig. 3.2: Willingness to Move of Ego and Alter

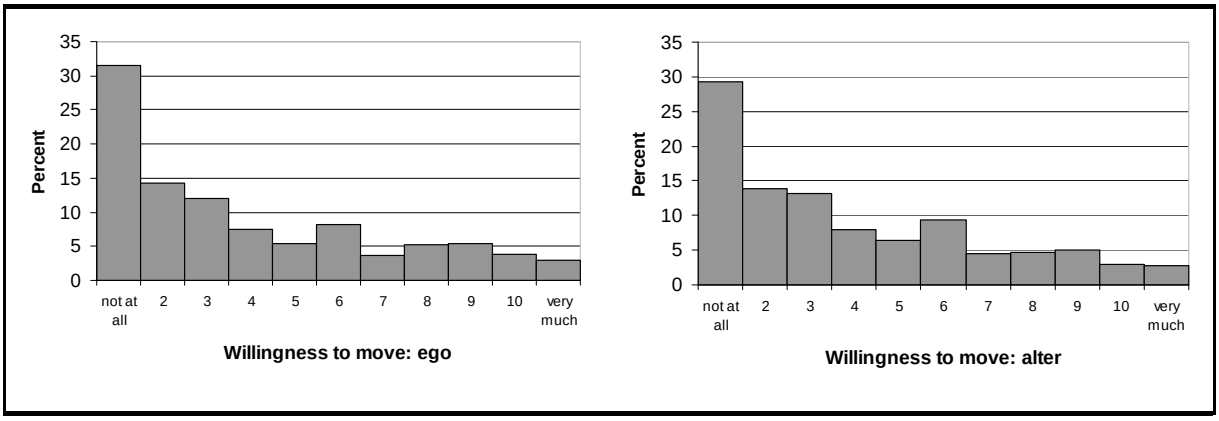
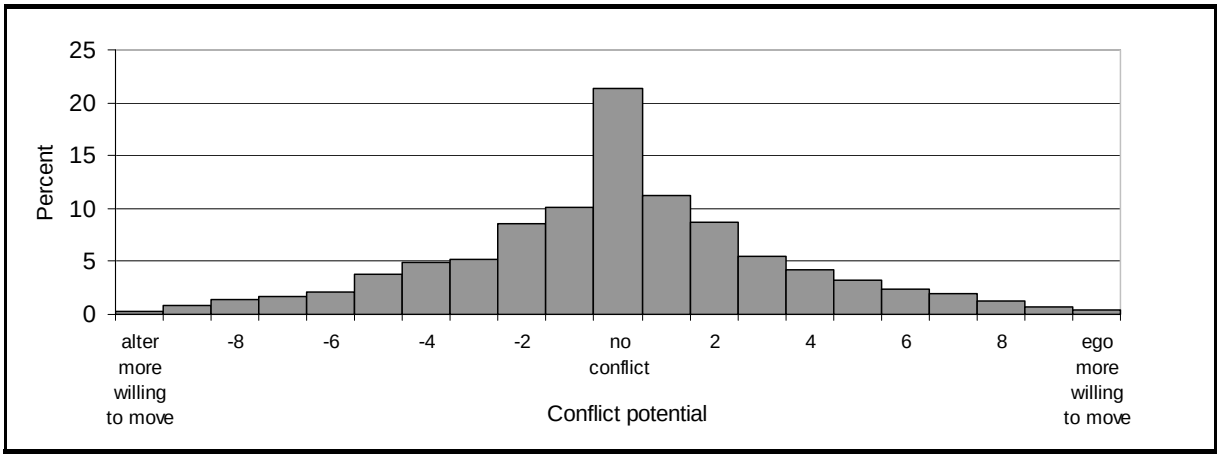


Fig. 3.3: Conflict Potential (Willingness to Move of Ego Minus Willingness to Move of Alter)



Note: In the regression analysis, we employed the absolute values of the conflict potential.

Table 3.1 presents descriptive data for the dependent variable and the sample. Because the vignettes' characteristics were all a direct result of our experimental plan, they always occurred with about the same frequency. The vast majority of our respondents were between 20 and 45 years old. All respondents were employed for at least 19 hours per week. The vast majority was employed full-time, only 13 percent were employed less than 35 hours per week.

Table 3.1: Sample Description

	N	Range	Mean	SD	Median
Dependent variables					
Willingness to move					
- Ego	2,772	1 – 11	3.90	3.04	3
- Alter	2,770	1 – 11	3.91	2.93	3
Conflict Potential (absolute value)	2,762	0 – 10	2.55	2.38	2
Vignettes' characteristics					
Gain of income for ego [percent]	2,780	30 – 70	49.82	14.07	50
Career prospects for ego (ref.: none)					
- Some	2,780	0 – 1	0.34	0.47	
- Many	2,780	0 – 1	0.32	0.47	
Commuting time [hours]	2,780	0.75 – 3.0	1.62	0.79	1.5
Only reachable by car	2,780	0 – 1	0.48	0.50	
Employment prospects for alter (ref.: little)					
- Moderate	2,780	0 – 1	0.32	0.47	
- Good	2,780	0 – 1	0.33	0.47	
Income prospects for alter at destination (ref.: smaller)					
- Equal	2,780	0 – 1	0.33	0.47	
- Better	2,780	0 – 1	0.33	0.47	
Respondents' characteristics					
Female respondent	556	0 – 1	0.50		
Age	547	20 – 50	30.93	4.97	30
Duration of employment [months]	551	0 – 360	61.96	58.03	45
Duration of residence [years]	552	0 – 45	20.33	11.70	24
Net income [euros][a]	555	360 – 5,014	1,935.61	865.32	1,750
Propensity to lose the actual job[b]	555	1 – 11	3.15	2.49	2
University graduate	556	0 – 1	0.44	0.50	
Households' characteristics					
Living in Switzerland	278	0 – 1	0.23	0.42	
Real estate property	278	0 – 1	0.25	0.43	
Married couple	278	0 – 1	0.35	0.48	
Duration of partnership [months]	278	2 – 325	88.03	56.79	78
Children living in household	277	0 – 1	0.12	0.33	
Married, long duration of partnership and children in household	277	0 – 1	0.08	0.27	
Education[c]					
- Male partner has higher education	278	0 – 1	0.32	0.47	
- Female partner has higher education	278	0 – 1	0.29	0.45	
Income[d]					
- Male partner has higher income	277	0 – 1	0.77	0.42	
- Female partner has higher income	277	0 – 1	0.17	0.38	
Age[e]					
- Male partner is older	269	0 – 1	0.44	0.50	
- Female partner is older	269	0 – 1	0.06	0.23	

[a] Net income: adjusted for the different purchasing power in Switzerland and Germany. [b] Propensity to lose the actual job: scored from 1 = *very unlikely* to 11 = *very likely*. [c] Education: 0 = *spouses have the same education*, 1 = *male or female spouse has higher education (at least one year more schooling).* [d] Income: 0 = *the monthly net income of the two spouses differ by less than 100 Euros*, 1 = *male or female partner has at least 100 Euros more income.* [e] Age: 0 = *the age difference is less than three years*, 1 = *male or female partner is at least three years older than his or her spouse.*

3.5.2 Test of the Hypotheses

Our first two hypotheses ($H1_a$ and $H1_b$) were tested using the analyses of ego's and alter's individual willingness to move. Table 3.2 shows the results of a random intercept model that controls for the individual variables. Negative coefficients indicate a lower willingness to migrate as the independent variable increases. The effects found in Models 1 and 2 implicate that both ego and alter showed a higher tendency to move as the employment conditions at the destination improved. Moreover, the data support our first hypothesis: Actors considered their partner's incentives in addition to their own incentives. Ego was more willing to move as alter's employment and income prospects improved at the new destination; and vice versa, alter showed a higher willingness to move as ego's migration gain increased. Because of the experimental, standardized design, we could compare the coefficients of alter and ego directly. An incentive to move for either partner was more relevant for the person who would profit from it than was the same incentive for his or her partner. For example, many career prospects for ego (compared to none) increased ego's willingness to move by an amount of 0.91 but alter's willingness to move only by an amount of 0.76. On the other hand, alter's willingness to move was influenced more strongly by his or her own better income prospects (1.60) as compared to ego's willingness to move (0.87). Both ego and alter took their partners' situation into account but gave their own hypothetical gains more weight than their partners' throughout.

One interesting finding is that female egos displayed a lower willingness to move as compared to male egos (see Baldridge et al. 2006; W. Bielby/D. Bielby 1992 for similar results). Without controlling for the absolute income level of male and female respondents, this difference was even higher (results not displayed). A reason for this observation may be the fact that women prefer a shorter distance from their home to their place of employment (McLafferty/Preston 1997) and rely on social support from relatives and other networks when they have children or anticipate their family obligations in the future. As a consequence, they may be more reluctant to leave a carefully constructed work arrangement (Kalleberg/Rosenfield 1990). Employment opportunities for better pay may be less beneficial for women because it may be difficult to achieve the same kind of balance between household responsibilities and employment responsibilities in a new job. We return to this when we present models differentiated by gender.

As for the other individual determinants, our results are mostly in line with previous findings from survey research on migration. People tended to be less mobile when they owned real estate property (Deane 1990: 72) or as the time lived in one place increased (Baldridge et al. 2006; Jürges 2005). The respondents' actual employment characteristics, education level, and current network embeddedness, however, did not have a significant influence on the willingness to move. This may be the result of effects working in opposite directions: A higher education level may, for example, lead to better earning-prospects at the actual destination and at the same time increase the potential gains of migration. Because these variables served only as controls, we refrained from deeper analyses.

Table 3.2: Summary of Linear Regression Analyses for Variables Predicting the Willingness to Move (Random Intercept Models)

	Model 1 Own job offer (Ego)		Model 2 Partner gets job offer (Alter)	
	β	SE β	β	SE β
Vignettes' characteristics				
Gain of income for ego [10 %]	0.30 ***	0.03	0.23***	0.03
Career prospects for ego (ref. none)				
- Some	0.74 ***	0.10	0.45***	0.10
- Many	0.91 ***	0.10	0.76***	0.10
Commuting time [hours]	-0.03	0.05	-0.17**	0.05
Only reachable by car (ref.: also by train)	0.00	0.08	-0.07	0.08
Employment prospects for alter at destination (ref.: little)				
- Moderate	0.45 ***	0.10	0.84***	0.10
- Good	1.03 ***	0.10	1.98***	0.10
Income prospects for alter at destination (ref.: smaller in comparison with the actual destination)				
- Equal	0.52 ***	0.10	0.67***	0.10
- Better	0.87 ***	0.10	1.60***	0.10
Respondents' characteristics				
Interview conducted in Switzerland	-0.01	0.38	0.02	0.32
Female respondent	-0.80 **	0.29	0.11	0.25
Age	-0.06	0.03	-0.04	0.03
Duration of residence	-0.02	0.01	-0.03*	0.01
Real estate property	-1.13 **	0.33	-0.64*	0.29
University graduate	0.41	0.31	0.23	0.28
Duration of employment [10 years]	0.02	0.03	-0.01	0.02
Net income [1,000 Euros][a]	0.28	0.20	-0.10	0.18
Propensity to lose the actual job	0.14 **	0.05	0.11*	0.05
Constant	2.56 *	1.01	2.72**	0.89
Observations	2,704		2,671	
(Interviewees)	(271)		(268)	
Variance				
- σν	2.02 ***	0.10	1.69 ***	0.08
- σε	1.95	0.03	1.97	0.03

Note: Gray highlighted fields mark vignette variables directly concerning the employment prospects of the respondent (instead of those of his or her partner).
[a] Adjusted for the different purchasing power in Switzerland and Germany.
*$p < .05$, **$p < .01$, ***$p < .001$.

Taken together, the outcomes from the individual models were mainly consistent with previous results based on migration research and provided strong evidence for our first hypothesis. We next examined our hypothesis regarding the conflict potential. It was measured as difference between ego's and alter's willingness to move, and hence represented a variable at the couple level. Table 3.3 presents the results of a random intercept model that includes partnership-level variables at level 2. For ease of interpretation, we used the absolute values of the conflict potential as dependent variable; higher values always indicated a greater disagreement in interests, whoever is more willing to move. Therefore, a positive (negative) regression coefficient indicated an increase (decrease) of the conflict potential as the respective variable increased.

Table 3.3: Summary of Linear Regression Analyses for Variables Predicting the Conflict Potential (Random Intercept Models)

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	β	SE β	β	SE β	β	SE β	β	SE β	β	SE β	β	SE β
Vignettes' characteristics												
Gain of income for ego [10%]	0.14***	0.03	0.14***	0.03	0.14***	0.03	0.14***	0.03	0.15***	0.03	0.15***	0.03
Career prospects for ego (ref.: none)												
- Some	0.11	0.10	0.12	0.10	0.11	0.10	0.11	0.10	0.12	0.10	0.12	0.10
- Many	0.32**	0.10	0.31**	0.10	0.31**	0.10	0.31**	0.10	0.34**	0.10	0.34**	0.10
Commuting time [hours]	-0.06	0.05	-0.06	0.05	-0.07	0.05	-0.07	0.05	-0.07	0.05	-0.07	0.05
Only reachable by car (ref.: also by train)	0.05	0.08	0.05	0.08	0.06	0.08	0.06	0.08	0.05	0.08	0.05	0.08
Employment prospects for alter at destination (ref.: little)												
- Moderate	0.10	0.10	0.11	0.10	0.10	0.10	0.10	0.10	0.12	0.10	0.13	0.10
- Good	0.51***	0.10	0.52***	0.10	0.51***	0.10	0.51***	0.10	0.52***	0.10	0.51***	0.10
Income prospects for alter (ref.: smaller in comp. with actual location)												
- Equal	0.23*	0.10	0.24*	0.10	0.23*	0.10	0.23*	0.10	0.24*	0.10	0.24*	0.10
- Better	0.64***	0.10	0.65***	0.10	0.64***	0.10	0.64***	0.10	0.65***	0.10	0.66***	0.10
Couples' characteristics												
Married couple	-0.31	0.17							0.03	0.21	-0.00	0.20
Children living in household			-0.48	0.25					-0.49	0.28		
Real estate property					-0.51**	0.19			-0.42*	0.21	-0.49*	0.20
Duration of partnership [10 months]							-0.04**	0.01	-0.02	0.02		
Married, long duration of partnership and children in household[a]											-0.76*	0.33
Education (ref.: same education)												
- Male partner has higher education									-0.19	0.20	-0.14	0.20
- Female partner has higher education									-0.25	0.20	-0.22	0.20
Income (ref.: same income)[b]												
- Male partner has higher income									0.31	0.37	0.30	0.37
- Female partner has higher income									0.49	0.41	0.47	0.41
Age (ref.: same age)[c]												
- Male partner is older									0.34	0.18	0.35*	0.17
- Female partner is older									0.31	0.38	0.31	0.37
Interview conducted in Switzerland									-0.24	0.21	-0.25	0.21
Constant	1.35***	0.25	1.28***	0.24	1.37***	0.25	1.58***	0.27	1.45**	0.53	1.32*	0.52
Observations (couples)	2,762 (278)		2,752 (277)		2,762 (278)		2,762 (278)		2,658 (267)		2,658 (267)	
Variance												
- σν	1.21***	0.07	1.21***	0.07	1.20***	0.07	1.20***	0.07	1.17***	0.07	1.17***	0.07
- σε	2.00	0.03	2.00	0.03	2.00	0.03	2.00	0.03	2.01	0.03	2.01	0.03

[a] 'Long duration of partnership' is defined by more than 6.5 years. [b] Cases are counted as 'same income' if the monthly net income of the two partners differs by less than 100 Euros. [c] Cases are counted as 'same age' if the age difference is less than three years.
*$p < .05$, **$p < .01$, ***$p < .001$

Starting with ego's mobility incentives, an interesting observation is that as income and employment prospects improved, the potential for conflicts also increased. This result is consistent with our second hypothesis: A one-sided increase in bargaining power would raise the difference in interests and, as a result, the conflict potential. Because of the characteristics of a multivariate regression analysis, the effect in question is indeed one-sided because we controlled for alter's bargaining power in the multivariate analysis. It is also exactly this aspect that makes sense of alter's results that appear to be similarly counter-intuitive. Alter's improving income and employment prospects increased the conflict potential. Because the indicators of the partner's bargaining power were always kept constant by the multiple regression models, we could interpret *any* actor's changes in employment prospects as one-sided. Furthermore, bargaining theory is very clear on this point. If alter's (ego's) incentive to move increases and ego's (alter's) incentive stays constant, we should observe a higher conflict, and this pattern was sustained by the findings. Consequently, our core hypothesis, H_2, was supported by these data.

For the third hypothesis, we have to look at partnership characteristics. We assumed that knowledge about each other and the level of commitment might be variables that counteract the conflict potential. Models 1 through 4 in Table 3.3 present the individual effects of the four relevant variables: marriage, children, and estate property as indicators of commitment; and duration of the partnership as an indicator of learning as well as commitment. The results mainly supported our hypotheses. The conflict potential decreased if a couple was married, had children, or owned real estate and decreased as the length of the relationship increased. The existence of a marriage and children showed the expected negative influence on the conflict potential but both effects failed to reach statistical significance. In the full Model 5, the effects of these variables were all still negative, but no longer significant. A possible explanation for this may be the relatively homogenous sample with high correlations of these partnership attributes (e.g., marriage and partnership were correlated at 0.45). Even with a random sample and more cases, it would be difficult to disentangle these effects because of an underlying mechanism: Marriage and children make a partnership more stable, and couples with a longer common history are more likely to marry and to become parents. To avoid potential multicollinearity, we compared an *extreme group* (couples that were married, had a long partnership duration, and children) with the rest of our sample. *Long duration* was defined by more than 6.5 years and applied to about half of our sample. The results presented in Model 6 supported our hypothesis that these determinants, especially when in combination with each other, were associated with less conflict potential in a partnership.

The variables indicating homogeneity of education, income, and age in Models 5 and 6 served as controls for different compositions of relationships. In the bargaining literature, the actors who possess more of these resources (education, income, and age) as compared to their partners are often considered to be more powerful. For example, a higher age indicates, in general, a higher advancement of occupational careers (Halleröd 2005). Job-related migrations are, however, assumed to be more

financially rewarding for younger people because of the longer time span left for labor market payoffs (Becker 1993: 85-89). Hence, any increase in the age difference between partners should increase the conflict potential regarding migrations. We found a significantly higher conflict potential only in Model 6 and in those partnerships where the male partner was at least three years older than his female partner. That we did not observe corresponding results for the partnerships in which the female partner was older may be a consequence of the small number of those couples in our sample; only 6 percent fell into this category (see again Table 3.1). Similarly, the high homogeneity of our sample and the high tendency of 'positive assortative mating' led to small variances in the education and income differences between the partners and may explain why these variables did not have strong effects.

Additionally, we tested several interaction effects of the career and employment prospects of the two partners to find out if the conflict potential was dampened if *both* partner had good opportunities at the destination. The regression estimates, which are not displayed but are available on request, revealed that the interaction terms mainly showed the expected negative effects on the conflict potential, but failed to reach significance.

Finally, we focused on possible gender effects. In order to test H_4, which states that male employment characteristics have a stronger effect on the willingness to move than female employment characteristics, we estimated separate models for male and female respondents. The results shown in Table 3.4 did not provide very strong support for gendered evaluations of the vignettes. Comparing the results for men with the results for women in Model 1 (ego) and Model 2 (alter), the gender gap was, in general, in line with the hypothesis, but not statistically significant. We analyzed overall differences between men and women by using Chow-tests (Wooldridge 2003) and the difference of single variables by interaction terms with gender, estimating in each case a joint model for male and female respondents. There was only one significant difference in the Model for alters (Model 2): The coefficient for ego's career prospects was significantly higher in the case of female alters than in the case of male alters. This means that women in the position of alter valued the career options related to their partners' job options higher than did men. Further analyses (not presented here) demonstrated that the overall results were similar when we controlled for differences in age, education, and income within partnerships. The evaluation, therefore, seemed to not be influenced by actors' actual relative power within the partnership. Thus, the evidence regarding H_4 was mixed overall: A female ego's willingness to move in general was lower (see Table 3.2), but the impact of incentives did not vary much by gender. Only alter's willingness to move was more influenced by ego's employment opportunities when alter was the female instead of male partner.

Table 3.4: Summary of Linear Regression Analyses for Variables Predicting the Willingness to Move of Men and Women (Random Intercept Models)

	Model 1 Own job offer (Ego)				**Model 2 Partner gets job offer (Alter)**			
	Men		**Women**		**Men**		**Women**	
	β	**SE β**	**β**	**SE β**	**β**	**SE β**	**β**	**SE β**
Vignettes' characteristics								
Gain of income for ego [10%]	0.35***	0.04	0.25***	0.04	0.22***	0.04	0.27***	0.04
Career prospects for ego (ref. none)								
- Some	0.85***	0.14	0.57***	0.14	**0.20**	0.14	**0.66*****	0.15
- Many	0.90***	0.14	0.84***	0.14	0.67***	0.14	0.85***	0.14
Commuting time [hours]	0.08	0.07	-0.14*	0.07	-0.21**	0.07	-0.15*	0.07
Only reachable by car (ref.: also by train)	0.08	0.11	-0.12	0.11	-0.02	0.11	-0.09	0.12
Employment prosp.for alter at destination (ref.: little)								
- Moderate	0.51***	0.14	0.33*	0.14	0.82***	0.13	0.78***	0.14
- Good	1.09***	0.14	0.86***	0.14	1.83***	0.13	2.06***	0.14
Income prospects for alter (ref.: smaller in comp. with actual location)								
- Equal	0.53***	0.13	0.55***	0.13	0.62***	0.13	0.62***	0.14
- Better	0.92***	0.13	0.90***	0.14	1.61***	0.13	1.51***	0.14
Constant	0.68	0.37	1.35***	0.37	1.17**	0.36	0.92*	0.36
Observations	1,438		1,334		1,336		1,434	
(Interviewees)	(144)		(134)		(134)		(144)	
Variance								
- σν	2.36***	0.15	1.99***	0.13	1.82***	0.12	1.84***	0.12
- σε	1.99	0.04	1.92	0.04	1.90	0.04	2.05	0.04

Note: In this regression analysis only vignettes' characteristics were included. Gray highlighted fields mark variables concerning the male partner. Coefficients in bold indicate significant differences between men and women.
*$p < .05$, **$p < .01$, ***$p < .001$.

3.6 Discussion

The goal of this paper was twofold. We attempted to shed light on the influence of both partners' job opportunities in migration decisions within dual-earner households, and we proposed a new way to test bargaining theory in family research. Based on a factorial survey of about 280 European couples, we demonstrated that there is a substantial potential for conflicts in a partnership when work-related incentives for a move arise. As predicted, the conflict potential within a couple varied according to each partner's change in relative bargaining power. Moreover, we found clear evidence that couples were able to minimize this conflict through mutual commitment and knowledge on each other (measured by the duration of partnership). Hence, the bargaining model is supported by our data. Additionally, we found little evidence for the relevance of gender regarding the influence of bargaining power on the individual willingness to move. This result has to be treated cautiously yet

because couples without children are oversampled in our data and couples with children tend to develop a more gendered division of labor (Baxter et al. 2008). And even though we found only slight evidence for gendered decisions, these might in some constellations already be sufficient to trigger a continuously widening of the career gap between men and women, as – from a bargaining point of view – subsequent career and migration decisions will promote especially the career of the more successful and therefore powerful spouse.

An important limitation of our study is that we observed only hypothetical decisions and no real actions. It is still unclear to what extent factorial surveys produce valid results (Eifler 2007). But fortunately a lot of research evaluating similar methods (conjoint analysis and stated choice experiments) exists. This research showed, all in all, an astonishingly high correspondence of hypothetical and experimental decisions with real observations (Blamey/Bennett 2001; Louviere et al. 2000: 354-382). Moreover, there was evidence that the willingness to move is a feasible predictor of actual relocation behavior (Brett/Reilly 1988). In addition, further analyses of migration behavior in Germany found that variables such as real estate property and occupational characteristics have a similar impact on the willingness to move reported in the analyses on hand and on real migration propensity measured with the German Socio-Economic Panel (Nisic/Auspurg 2009). Thus, we do not have any major reason to doubt the overall external validity of our results.

Another critical aspect is the occurrence of a social desirability bias. The sensitivity of factorial surveys for those effects in general is controversial. In our case, the risk seems minor because we did not ask very sensitive questions. Social desirability should, by definition, intensify the measured relevance of norms. That we observed only weak evidence for gendered norms is, therefore, a hint for the robustness of our results, even if social desirability cannot be completely excluded. Regarding the convenience sample, we encourage replications of our design with other sampling techniques.

Another direction for future research would be a test of the supposition that the actors anticipate future allocations. This could be done by additional answer scales, questioning the anticipated share of resources or even obligations (such as housework tasks) in the case where a household move has taken place. Gender dynamics underlying the decision making processes not visible through the quantitative data on hand might be addressed by combining vignette analyses with qualitative interviews (see Ganong/Coleman 2006 for an example).

In general, the results of our study underline the relevance of a couple- or household-oriented approach when labor market processes are studied. Individual career mobility is often directly linked to regional mobility, and we have shown that new individual career options for one partner create potential conflicts within the relationship. Cooperative solutions require balanced power shifts or, at the least, credible commitment to the partnership. The scope of this finding is probably not restricted to migration decisions, because the same theoretical mechanism can be assumed for other issues such as fertility decisions. Without solutions to the conflicts connected with choices that asymmetrically

change the employment prospects of the partners and therefore shift their balance of bargaining power, suboptimal levels of family welfare will arise (Jürges 2005: 302; Ott 1992).

At least in our sample of European dual-earner couples, the responses to the incentives for regional moves differed only slightly by gender. Apart from the fact that women tended to be more resistant to moving compared to their male partners, gender roles were found to be less obviously important than is often stated. The high prevalence of tied women reported in the literature seems to result primarily from their inferior labor market positions than from traditional gender roles within the household. Both male and female partners were found to take their respective partner's perspective into account, and they did it mostly in the same way. In a more general view, this means that the efforts to achieve greater gender equality in the labor market and in families go hand in hand. Better career prospects for women and more balanced (migration) decisions presuppose each other.

References

Antel, John J. (1980): Returns to Migration: Literature Review and Critique. Santa Monica: Rand Corporation.

Baldridge, David C./Eddleston, Kimberley A./Veiga, John F. (2006): Saying 'no' to being uprooted: The impact of family and gender on willingness to relocate. Journal of Occupational and Organizational Psychology 79, 131-149.

Bartel, Ann P. (1982): Wages, Nonwage Job Characteristics, and Labor Mobility. Industrial and Labor Relations Review 35, 578-589.

Baxter, Janeen/Hewitt, Belinda/Haynes, Michele (2008): Life Course Transitions and Housework: Marriage, Parenthood, and Time on Housework. Journal of Marriage and Family 70, 259-272.

Becker, Gary S. (1993): Human Capital. A Theoretical and Empirical Analysis with Special Reference to Education. Chicago, London: University of Chicago Press.

Bergstrom, Theodore C. (1996): Economics in a Family Way. Journal of Economic Literature 34, 1903-1934.

Bernasco, Wim/Giesen, Deidre (2000): A Bargaining Approach to Specialization in Couples. In: Weesie, Jeroen/Raub, Werner (Eds.): The Management of Durable Relations. Theoretical Models and Empirical Studies of Households and Organizations. Amsterdam, Thela Thesis, 42-43.

Bielby, William T./Bielby, Denise D. (1992): I Will Follow Him: Family Ties, Gender-Role Beliefs, and Reluctance to Relocate for a Better Job. American Journal of Sociology 97, 1241-1267.

Bitman, Michael/England, Paula/Folbre, Nancy/Sayer, Liana/Matheson, George (2003): When Does Gender Trump Money? Bargaining and Time in Household Work. American Journal of Sociology 109, 186-214.

Blackburn, McKinley (2006): The Impact of Internal Migration on Married Couples' Earnings in Britain, with a Comparison to the United States. ISER Working Paper 2006-24. Essex: Institute for Social and Economic Research.

Blamey, Russell/Bennett, Jeff (2001): Yea-saying and Validation of a Choice Model of Green Product Choice. In: Bennett, Jeff/Blamey, Russell (Eds.): The Choice Modeling Approach to Environmental Valuation. Cheltenham Northampton: Edward Elgar, 179-201.

Blau, Francine D./Ferber, Marianne A./Winkler, Anne E. (2001): The Economics of Women, Men, and Work. Englewood Cliffs: Prentice Hall.

Blood, Robert O./Wolfe, Donald M. (1960): Husbands & Wives. The Dynamics of Married Living. Glencoe: Free Press.

Brett, Jeanne M./Reilly, Anne H. (1988): On the road again: Predicting the job transfer decision. Journal of Applied Psychology 73, 614-620.

Brines, Julie/Joyner, Kara (1999): The Ties That Bind: Principles of Cohesion in Cohabitation and Marriage. American Sociological Review 64, 333-355.

Cooke, Lynn P. (2006): "Doing" Gender in Context: Household Bargaining and Risk of Divorce in Germany and the United States. American Journal of Sociology 112, 442-472.

DaVanzo, Julie/Hosek, James R. (1981): Does Migration Increase Wage Rates? An Analysis of Alternative Techniques for Measuring Wage Gains to Migration. Santa Monica: Rand Corporation.

Deane, Glenn D. (1990): Mobility and Adjustment: Paths to the Resolution of Residential Stress. Demography 27, 65-79.

Eifler, Stefanie (2007): Evaluating the Validity of Self-Reported Deviant Behavior Using Vignette Analyses. Quality & Quantity 41, 303-318.

England, Paula/Farkas, Georg (1986): Households, Employment, and Gender. A Social, Economic, and Demographic View. New York: Aldine.

Euorpean Commission (2009): Report on Equality between Women and Men 2009. Luxembourg: Office for Official Publications of the European Communities.

Fenstermaker, Sarah (2002): Work and Gender. In: Fenstermaker, Sarah/West, Candace (Eds.): Doing Gender, Doing Difference. Inequality, Power, and Institutional Change. New York: Routledge, 105-114.

Ganong, Lawrence H./Coleman, Marilyn (2006): Multiple Segment Factorial Vignette Design. Journal of Marriage and Family 68, 455-468.

Halleröd, Björn (2005): Sharing of Housework and Money Among Swedish Couples: Do They Behave Rationally? European Sociological Review 21, 273-288.

Hook, Jennifer L./Chalsani, Satvika (2008): Gendered Expectations? Reconsidering Single Father‘s Child-Care Time. Journal of Marriage and Family 71, 978-990.

Hox, Joop J./Kreft, Ita G. G./Hermkens, Piet L. J. (1991): The Analysis of Factorial Surveys. Sociological Methods & Research 19, 493-510.

Huppé, Micheline/Cyr, Mireille (1997): Division of Household Labor and Marital Satisfaction of Dual Income Couples According to Family Life Cycle. Canadian Journal of Counseling 31, 145-162.

Jacobson, Joyce P./Levin, Laurence M. (2000): The effects of internal migration on the relative economic status of women and men. Journal of Socio-Economics 29, 291-304.

Jasso, Guillermina (2006): Factorial Survey Methods for Studying Beliefs and Judgments. Sociological Methods and Research 34, 334-423.

Jürges, Hendrik (2005): Gender ideology, division of housework, and the geographic mobility of families. Review of Economics of the Household 4, 299-323.

Kalleberg, Arne L./Rosenfield, Rachel A. (1990): Work in the Family and in the Labor Market: A Cross-National, Reciprocal Analysis. Journal of Marriage and the Family 52, 331-346.

Klasen, Stephan (1998): Marriage, Bargaining, and Intrahousehold Resource Allocation: Excess Female Mortality among Adults during Early German Development, 1740 – 1860. Journal of Economic History 58, 432-467.

Kuhfeld, Warren F. (2005): Marketing Research Methods in SAS. Experimental Design, Choice, Conjoint and Graphical Techniques. Cary: SAS Institute.

Kuhfeld, Warren F./Randall, Tobias D./Garratt, Mark (1994): Efficient Experimental Design with Marketing Research Applications. Journal of Marketing Research 31, 545-557.

Lichter, Daniel T. (1983): Socioeconomic Returns to Migration Among Married Women. Social Forces 62, 487-503.

Louviere, Jordan J./Hensher, David A./Swait, Joffre D. (2000): Stated Choice Methods. Analysis and Application. Cambridge, UK: University Press.

Lundberg, Shelly/Pollak, Robert A. (2003): Efficiency in Marriage. Review of Economics of the Household 1, 153-167.

Lundberg, Shelly/Pollak, Robert A. (1996): Bargaining and Distribution in Marriage. Journal of Economic Perspectives 10, 139-158.

Manser, Marilyn/Brown, Murray (1980): Marriage and Household Decision-Making: A Bargaining Analysis. International Economic Review 21, 31-44.

Maxwell, Nan L. (1988): Economic Returns to Migration: Marital Status and Gender Differences. Social Science Quarterly 69, 108-121.

McHugh, Kevin E./Gober, Patricia/Reid, Neil (1990): Determinants of Short- and Long-Term Mobility. Expectations for Home Owners and Renters. Demography 27, 81-95.

McLafferty, Sara/Preston, Valerie (1997): Gender, race, and the determinants of commuting: New York in 1990. Urban Geography 18, 192-212.

Miller, Roland S./Perlman, Daniel/Brehm, Sharon S. (2007): Intimate Relationships. Boston: McGraw & Hill.

Mincer, Jacob (1978): Family Migration Decisions. Journal of Political Economy 86, 749-773.

Nisic, Natascha/Auspurg, Katrin (2009): Faktorieller Survey und klassische Bevölkerungsumfrage im Vergleich – Validität, Grenzen und Möglichkeiten beider Ansätze. In: Kriwy, Peter/Gross, Christiane (Eds.): Klein aber fein! Quantiative empirische Sozialforschung mit kleinen Fallzahlen. Wiesbaden: VS Verlag für Sozialwissenschaften, 211-246.

Nock, Steven L. (1995): Commitment and Dependency in Marriage. Journal of Marriage and the Family 57, 503-514.

Ott, Notburga (1992): Intrafamily Bargaining and Household Decisions. Berlin: Springer.

Quigley, John M./Weinberg, Daniel H. (1977): Intra-Urban Residential Mobility: A Review and Synthesis. International Regional Science Review 2, 41-66.

Raudenbush, Stephen W./Bryk, Anthony S. (2001): Hierarchical Linear Models: Applications and Data Analysis Methods. Sage: Thousand Oaks.

Rossi, Peter H./Anderson, Andy B. (1982): The Factorial Survey Approach: An Introduction. In: Rossi, Peter H./Nock, Steven L. (Eds.): Measuring Social Judgments: The Factorial Survey Approach. Beverly Hills: Sage, 15-67.

Shauman, Kimberley A./Noonan, Mary C. (2007): Family Migration and Labor Force Outcomes: Sex Differences in Occupational Context. Social Forces 85, 1735-1764.

Snijders, Tom A./Bosker, Roel J. (1999): Multilevel Analysis. An introduction to basic and advanced modeling. London: Sage.

Spitze, Glenna (1984): The Effect of Family Migration on Wives' Employment: How Long Does it Last? Social Science Quarterly 65, 21-36.

Stickney, Lisa T./Konrad, Alison M. (2007): Gender-Role Attitudes and Earnings: A Multinational Study of Married Women and Men. Sex Roles 57, 801-811.

Szinovacz, Maximiliane/Harpster, Paula (1993): Employment Status, Gender Role Attitudes, and Marital Dependence in Later Life. Journal of Marriage and the Family 55, 927-940.

Taylor, Mark P. (2007): Tied Migration and Subsequent Employment: Evidence from Couples in Britain. Oxford Bulletin of Economics and Statistics 69, 795-818.

Thibaut, John W./Kelley, Harold H. (1959): The Social Psychology of Groups. New York: Wiley.

West, Candace/Zimmerman, Don H. (1987): Doing Gender. Gender and Society 1, 125-151.

Wooldridge, Jeffrey M. (2003): Introductory Econometrics. A Modern Approach. Mahson, Ohio: Thomson.

4

Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey

(Katrin Auspurg, Thomas Hinz und Stefan Liebig)

1 Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey[1]

Zusammenfassung

Der Faktorielle Survey gilt als eine Erhebungsmethode, bei der sich die Vorteile der Umfrageforschung mit denen experimenteller Designs verbinden. Statt einzelner Items bewerten die Befragten hypothetische Objekt- oder Situationsbeschreibungen. Indem in diesen *Vignetten* einzelne Merkmalsausprägungen experimentell variiert werden, lässt sich ihr Einfluss auf die abgefragten Urteile oder Entscheidungen exakt bestimmen und somit das Gewicht von Faktoren isolieren, die in der Realität oftmals konfundiert sind. Bislang liegen allerdings nur sehr wenige Methodenstudien zur Validität der erzielten Messungen vor. Der Beitrag gibt zunächst einen knappen Überblick zum Einsatz des Faktoriellen Surveys in der sozialwissenschaftlichen Forschung und benennt anschließend bislang ungeklärte methodische Probleme. Die mit einer eigenen experimentellen Datenerhebung durchgeführten Analysen beziehen sich auf die Stabilität des Urteilsverhaltens der Befragten in Abhängigkeit von der Anzahl der in den Vignetten abgebildeten Dimensionen, möglichen Lerneffekten sowie „unplausiblen“ oder „unlogischen“ Fällen (Vignettentexte für Situationen, die in der Realität sehr selten oder gar nicht vorkommen und die Befragten daher irritieren könnten). Getestet werden verschiedene Hypothesen zur Komplexität der Erhebungssituation und der Kohärenz der Urteile. Nach unseren Ergebnissen führen eine hohe Komplexität der Vignetten und unplausible Fälle zu einem weniger Vignettendimensionen einbeziehenden Urteilsverhalten, damit geringeren Einflussstärken einzelner Vignettenmerkmale bei gleich bleibender Konsistenz. Abschließend diskutieren wir die praktischen Konsequenzen dieser Befunde.

Abstract

The factorial survey is a method of data collection that combines the advantages of survey research and the advantages of experimental designs. Respondents react to hypothetical descriptions of objects or situations (vignettes) instead of answering to single-item questions. By varying each dimension of the vignettes in an experimental design, the dimensions’ impact on respondents’ judgments or decisions can be estimated accurately. Thus, the method is able to identify the effects of single factors which are often confounded in reality. So far, only few methodological studies address questions of measurement validity when a factorial survey design is used. The article provides a brief overview of the use of the factorial design in the social sciences and points still unresolved methodological questions. Using experimental data specifically designed for this purpose our analyses consider the stability of respondents’ judgments with respect to the number of dimensions presented in the vignettes, possible learning effects and ‚implausible’ or ‚illogical’ cases (vignettes describing objects or situations which are rare or even impossible). We test several hypotheses regarding the complexity of vignettes and the consistency of judgments. According to our results, a high complexity of vignettes and implausible cases cause respondents to consider fewer dimensions in their judgments; we find smaller influences of vignette variables while the consistency of the judgments remains the same. Finally, we discuss the practical consequences of these results.

[1] Der Beitrag entstand im Rahmen des von der DFG geförderten Forschungsprojekts „Der Faktorielle Survey als Instrument zur Einstellungsmessung in Umfragen“. Projektleiter sind Thomas Hinz (Universität Konstanz) und Stefan Liebig (Universität Bielefeld). Die Autoren danken Peter Steiner sowie einem anonymen Gutachter für wertvolle Hinweise und Anmerkungen. Für die Unterstützung bei der Organisation der Feldphase bedanken wir uns bei Judith Tonner.

4.1 Einleitung

Der Faktorielle Survey ist eine in Umfragen einsetzbare experimentelle Methode, bei der den Befragten hypothetische Objekt- oder Situationsbeschreibungen (*Vignetten*) vorgelegt werden.[2] Die Vignetten unterscheiden sich nach Merkmalen (*Dimensionen*), die in ihren Ausprägungen (*Levels*) variieren. Solche hypothetischen Fälle und Szenarien, die Befragte beurteilen oder bewerten, werden heute in verschiedenen akademischen und nicht-akademischen Forschungszusammenhängen vermehrt eingesetzt, neben den Sozialwissenschaften etwa auch in den Gesundheitswissenschaften, der Rechtswissenschaft, der Psychologie und der Marktforschung. Thematisch zeigen die Studien in der Soziologie eine beachtliche Breite. In der Norm- und Werteforschung beschäftigen sie sich mit der Messung von Status und Prestige von Individuen und Haushalten (Meudell 1982; Nock 1982; Rossi 1979; Rossi et al. 1974), den Vorstellungen über ein gerechtes Erwerbseinkommen (Alves/Rossi 1978; Hermkens/Boerman 1989; Jann 2003; Jasso 1994; Jasso/Webster 1997, 1999; Shepelak/Alwin 1986), der Bewertung von Armutsdimensionen (Will 1993), den Kriterien zur Festlegung wohlfahrtsstaatlicher Unterstützungszahlungen (Liebig/Mau 2002), mit gerechten Steuersätzen (Liebig/Mau 2005) und Entlassungsverfahren (Struck et al. 2008). Ebenso liegen Arbeiten vor zur Bewertung von sexuellem Missbrauch bzw. sexueller Belästigung (Garrett 1982; O´Toole et al. 1999; Rossi/Anderson 1982), zu der Bestrafung und dem Umgang mit Straftätern (Berk/Rossi 1977; Hembroff 1987; Miller et al. 1986), unterschiedlichen Kriterien der Einbürgerung (Jasso 1988), zur Vergabe medizinischer Hilfen (Hechter et al. 1999), zur Qualität von Kinderbetreuungsmaßnahmen (Shlay et al. 2005) und zum sozialen Kontext von Normgeltung (Beck/Opp 2001; Diefenbach/Opp 2007; Horne 2003; Jasso/Opp 1997). Ferner existieren Arbeiten, die der Frage möglicher Diskriminierungen nachgehen (Jann 2003; John/Bates 1990), Effekte sozialer Einbettung analysieren (Buskens/Weesie 2000) oder familiensoziologische Theorien untersuchen (Auspurg/Abraham 2007). Angesichts des großen und vielfältigen Interesses für diese Erhebungsmethode verwundert es, dass methodische und modelltheoretische Fragen sehr selten diskutiert werden (Ausnahmen: Dülmer 2001, 2007; Dülmer/ Klein 2003; Steiner/Atzmüller 2006). Wenn sie thematisiert werden, so besteht das Anliegen meistens darin, die Vorteile dieser Befragungsmethode gegenüber itembasierten Abfragen oder den traditionellen experimentellen Vorgehensweisen zu unterstreichen (Hechter et al. 2005; Jasso 1988). Die im Verfahren angelegten methodischen Probleme waren dagegen kaum Gegenstand einer expliziten Untersuchung. Dies gilt insbesondere für Probleme, die sich aus der Anlage und Durchführung eines Faktoriellen Surveys und dem Einsatz von Vignetten in Umfragen ergeben.

Wir verfolgen daher das Ziel, drei miteinander verbundene und als besonders relevant geltende methodische Probleme zu diskutieren und anhand von empirischen Tests zu untersuchen. Dies sind (1) die Effekte der Komplexität der den Befragten geschilderten Situation und (2) die hiermit in Verbin-

[2] Ursprünglich wurde der Faktorielle Survey in den Sozialwissenschaften 1951 von Peter H. Rossi in seiner Dissertation entwickelt und zur Einschätzung des sozialen Status von Haushalten verwendet (Alves/Rossi 1978; Rossi 1979; Rossi/Nock 1982). Rossis zentrales Anliegen war es, ein Messverfahren zu entwickeln, das es ermöglicht herauszufinden, welche Objekteigenschaften in welchem Ausmaß für soziale Einstellungen relevant sind (Rossi/Anderson 1982: 15ff.; Rossi/Nock 1982: 9ff.).

dung stehenden Lerneffekte bei wiederholter Präsentation von Vignetten. Da das Risiko von unplausiblen Fällen mit der Komplexität steigt und diese zudem als ursächlich für Lerneffekte in Form von vereinfachten Entscheidungsheuristiken gelten, analysieren wir zudem (3) die Auswirkungen unplausibler Vignetten auf das Urteilsverhalten. Diese Aspekte wurden nach unserem Wissen für Vignettenstudien allesamt noch nicht gezielt untersucht. Eine Beschäftigung mit methodischen Effekten scheint für eine Verbesserung der Datenqualität jedoch dringend angeraten, auch um möglichen Fehlschlüssen gezielt vorzubeugen (die andernfalls beim Vergleich verschieden komplexer Vignettenstudien oder kognitiv unterschiedlich belastbarer Probandengruppen zu befürchten sind) – sei es durch ihren Einbezug bei der Vignettenkonstruktion, Datenauswertung oder Ergebnisinterpretation.

Die Gliederung ist wie folgt: Zunächst werden die Verfahrensweise des Faktoriellen Surveys sowie der Stand der Methodendiskussion knapp vorgestellt (Abschnitt 4.2). Dann werden ausgehend vom Forschungsstand Hypothesen zu den genannten Problemstellungen abgeleitet (Abschnitt 4.3) und auf der Grundlage einer experimentellen Online-Erhebung getestet (Abschnitte 4.4 und 4.5). Schließlich werden die Ergebnisse diskutiert und weiterer Analysebedarf aufgezeigt (Abschnitt 4.6).

4.2 Faktorieller Survey: Aufbau, Motivation und Probleme

Der Faktorielle Survey zielt darauf ab, die *relativen* Gewichte einzelner Objekt- oder Situationsmerkmale für Einstellungen, Bewertungen oder Entscheidungen zu bestimmen (Beck/Opp 2001; Jasso 2006; Rossi/Anderson 1982 für detaillierte Einführungen). Dazu sind zunächst die in den Vignetten enthaltenen Merkmalsdimensionen und ihre Ausprägungen nach theoretischen Vorüberlegungen auszuwählen. In den Befragungssituationen werden diese Ausprägungen dann experimentell variiert um zu prüfen, ob diese gezielt erzeugte Variation der Objekt- und Situationsmerkmale eine entsprechende Variation der Urteile der Befragten nach sich zieht. In den Auswertungen lassen sich dann die exakten Beziehungen zwischen den Merkmalen und den Urteilen der Befragten ermitteln.

In der Durchführung Faktorieller Surveys werden die Befragten in der Regel also mit mehreren, zufällig oder systematisch ausgewählten Vignetten konfrontiert.[3] Die Befragungsmethode hat gegenüber itembasierten Survey-Studien vier wesentliche Vorteile. *Erstens* erlaubt sie eine Konstruktion von Objekten und Situationen, bei denen eine Mehrzahl solcher Merkmale zusammentreten, die in der Realität oft stark miteinander korrelieren und deswegen keine getrennte Einschätzung ihrer Bedeutung erlauben. Im experimentellen Design des Faktoriellen Surveys lassen sich diese Faktoren isolieren, im

[3] Dies stellt auch das Standardvorgehen in der vornehmlich in der Marktforschung verwendeten und dem Faktoriellen Survey ähnlichen Conjoint-Analyse dar (Carroll/Green 1995). Hier werden den Probanden meist simulierte oder echte Produktbeschreibungen vorgelegt und anschließend die relativen Nutzenwerte je Produktmerkmal ermittelt. Die Produkte weisen wie die Vignetten ein mehrfaktorielles Merkmalsbündel auf (Klein 2002; Orme 2006). Geht es um die Ermittlung von Entscheidungen, werden dagegen zum Teil nur wenige, in manchen Fällen nur eine einzige Vignette präsentiert. Es gibt durchaus Argumente, bei randomisierter Verteilung der Vignetten auf die Befragten nur eine einzige Vignette zu präsentieren: Die Effekte sozialer Erwünschtheit werden vermieden sowie die in diesem Aufsatz thematisierten Lerneffekte ausgeschlossen. In solchen Studien muss auch kein Mehrebenendesign bemüht werden (Jann 2003 für ein Beispiel).

technischen Sinn zueinander orthogonal setzen. Die so erzeugte Unkorreliertheit der Merkmale ermöglicht eine separate Bestimmung ihres jeweiligen Einflusses auf Urteil und Entscheidung. *Zweitens* können entsprechende Forschungshypothesen im Unterschied zur klassischen Laborforschung auf der Grundlage größerer (Zufalls-)Stichproben in Bevölkerungsumfragen überprüft werden. *Drittens* eröffnen sich interessante Analysemöglichkeiten, wenn den Befragten mehrere Vignetten vorgelegt und deshalb pro Befragten mehrere Urteile erzielt werden. Es entsteht hierdurch eine hierarchische Mehrebenenstruktur, die genutzt werden kann, um zwischen *between*- und *within-subject*-Faktoren zu unterscheiden. Es ist möglich, die Kovariation des Einflusses von Vignetten- und Befragtenmerkmalen auf die Urteile zu ermitteln. *Viertens* kann mit Faktoriellen Surveys einem gewichtigen Vorwurf an die konventionelle Einstellungsmessung begegnet werden, die Analyse lediglich einzelner Item-Werte würde der komplexen Struktur von Einstellungen nicht gerecht (Jasso/Opp 1997: 949; Liebig/Mau 2002: 114-116). Im Faktoriellen Survey sind komplexe Beurteilungs- und Entscheidungsprobleme simulierbar, indem eine Vielzahl von Merkmalen gekreuzt wird. Dies gilt insbesondere für solche Objekte und Situationen, bei denen verschiedene Merkmale in unterschiedlichem Grad urteilsrelevant werden und bei denen der soziale Kontext der Entscheidungssituation eine wichtige Rolle spielt. So wird beispielsweise die Höhe eines als gerecht empfundenen Erwerbseinkommens für einen Erwerbstätigen oder das gerechte Strafmaß für einen Verurteilten an das Vorliegen verschiedener Bedingungen gekoppelt sein. Genau diese Bedingungen können im Rahmen des Faktoriellen Surveys berücksichtigt und *alltagsnah* simuliert werden. Durch eine solche Verbundmessung könne – so die Argumentation einiger Autoren (Dülmer/Klein 2003; Hechter et al. 2005; Jasso 1988) – eine validere Messung von Einstellungen erzielt werden als durch itembasierte Verfahren. Denn die Einstellungen zu den einzelnen Dimensionen werden nicht sequenziell, sondern in der Situationsbeschreibung gemeinsam erfragt. Darüber hinaus verhindere die wiederholte Bewertung einer größeren Anzahl von Objekten und Situationen, dass Befragte ein falsches oder künstliches Bild ihrer Einstellungen zeichnen (Hechter et al. 1999). Tatsächlich haben Vergleiche von item- und vignettenbasierten Messungen gezeigt, dass über Faktorielle Surveys erfasste Einstellungen weniger durch soziale Erwünschtheit verzerrt werden (Jann 2003; Liebig/Mau 2002; Smith 1986). Vor diesem Hintergrund resümieren Dülmer/Klein (2003), dass über die Vignettenanalyse eine vergleichsweise exakte Einstellungsmessung möglich sei (s. auch Hechter et al. 2005: 103; Jasso 1988).

Von Kritikern des Faktoriellen Surveys werden aber auch eine ganze Reihe von Nachteilen bzw. Unzulänglichkeiten genannt. Grundsätzliche Einwände beziehen sich zunächst auf den vergleichsweise hohen zeitlichen Befragungsaufwand und die daraus resultierenden Opportunitätskosten bezüglich der Erhebung alternativer Items (Sniderman/Grob 1996). Die Bewertung von zehn und mehr Vignetten ist zeitlich aufwändiger als eine entsprechende itembasierte Abfrage der Dimensionen (Dülmer/Klein 2003; Liebig/Mau 2002). Als problematischer wird jedoch angesehen, dass bei Vignettenstudien vergleichsweise starke Antworteffekte plausibel sind, die sich aus der Auswahl der Beispiele (z. B. Kontrasteffekte), deren Reihenfolge (*carry-over*-Effekte) oder aus der Komplexität der präsentierten Beispiele ergeben können. Mit Faktoriellen Surveys erhobene Einstellungen wären daher

höchst instabil und letztlich Artefakte. Letzteres sei insbesondere dann zu erwarten, wenn die Befragten aufgrund der hohen Komplexität der Bewertungsaufgabe überfordert seien. Sie würden mitunter solche Dimensionen in ihr Antwortverhalten einfließen lassen, denen sie „eigentlich" gar keine Bedeutung zumessen. Kritisch angemerkt wird diesbezüglich zudem die Gefahr einer zu starken oder ausschließlichen Konzentration der Befragten auf ein in sich stimmiges Antwortverhalten (Faia 1980; Seyde 2005). Ferner können mögliche Kontexteffekte durch Namen, Begriffe oder Bezeichnungen entstehen und Störeffekte hervorrufen, die aus den individuellen Erfahrungen der Befragten entstammen (welche den unterschiedlichsten Alltagssituationen inhärent sind) und daher kaum kontrollierbar sind.[4] Diese Einwände konnten bislang aufgrund fehlender Methodenstudien weder bestätigt noch entkräftet werden.

Der vorliegende Beitrag bezieht sich auf derartige Forschungslücken und entstand in der ersten Phase eines breiter ansetzenden, von der Deutschen Forschungsgemeinschaft (DFG) finanzierten Projekts der Universitäten Konstanz und Bielefeld.[5] Die hier präsentierten Analysen konzentrieren sich auf folgende drei Aspekte: (1) Zunächst geht es um die Bestimmung einer noch handhabbaren Komplexität der geschilderten Situationen (Beck/Opp 2001: 287; Rossi/Anderson 1982: 59). Diese wird anhand der Menge an variablen Dimensionen untersucht. Mögliche kognitive Über- bzw. Unterforderungen sind allerdings nicht unabhängig von Lerneffekten durch die wiederholte Bearbeitung von Vignetten zu beurteilen, weshalb wir als weiteren Aspekt (2) die Konsistenz des Urteilsverhaltens im Bearbeitungsverlauf analysieren. Schließlich adressieren wir (3) die Auswirkung von unplausiblen Fällen, die – wie noch ausführlicher begründet – ebenfalls in Wechselwirkung mit den beiden anderen Aspekten zu sehen ist.

[4] Fraglich ist schließlich auch die prognostische Validität des Verfahrens (Rooks et al. 2000), da die Befragten nur hypothetische und nicht aktuelle Entscheidungen treffen (dazu Hechter et al. 2005; für Versuche einer externen Validierung Eifler 2007; Groß/Börensen 2009; Nisic/Auspurg 2009).

[5] Im Rahmen dieses DFG-Forschungsprojekts werden vielfältige experimentelle Variationen zur Komplexität der Erhebungssituation (Anzahl der Merkmalsdimensionen und Vignetten) sowie Relevanz von möglichen Reihenfolgeeffekten und zur Bedeutung von Darstellungsformen (Bandbreite der Ausprägungen bzw. *range*-Effekte, Einflüsse verschiedener Beurteilungsskalen und Präsentationsformen) untersucht. Außerdem gilt die Aufmerksamkeit der zeitlichen Stabilität der Messungen. Umfangreiche Experimentalreihen werden mit einem Studierenden-Sample bearbeitet, es geht darauf aufbauend im Projekt aber ebenso um die Tauglichkeit der Befragungsmethode in allgemeinen Bevölkerungsumfragen. Um die Belastbarkeit der Befragten und den Zeitaufwand alters- und bildungsübergreifend einschätzen zu können, werden unterschiedlich komplexe Designs an einer bevölkerungsrepräsentativen Stichprobe in zwei Surveysituationen (*face-to-face* und schriftlich) getestet. Für nähere Informationen: http://www.soziologie.uni-konstanz.de/hinz/forschung/forschungsprojekte/fs/.

4.3 Forschungsstand und Hypothesen

Im Folgenden berichten wir den Forschungsstand zu den drei benannten methodischen Problemen und leiten daraus Hypothesen zu den Effekten auf das Antwortverhalten ab. Aufgrund der unzureichenden Forschungslage zu Faktoriellen Surveys ziehen wir mitunter Literatur zu verwandten Verfahren der Marktforschung und Umwelt- und Gesundheitsökonomie heran (Conjoint- und Choice-Experimente).

4.3.1 Komplexität von Vignetten: Anzahl der Dimensionen

Wie bereits erwähnt, ist der Faktorielle Survey insbesondere für Fragestellungen geeignet, bei denen komplexe Bewertungen vorzunehmen sind. Der Wunsch, über viele Dimensionen eine möglichst detaillierte und alltagsnahe Beschreibung zu erhalten, kollidiert allerdings mit der eingeschränkten Verarbeitungskapazität der Befragten. Die Entscheidung für eine bestimmte Anzahl von Dimensionen ist somit von weitreichender Bedeutung (Rossi/Anderson 1982). Dies gilt, weil die Anzahl der Dimensionen über die Länge der Situationsbeschreibungen und damit die Komplexität der Bewertungsaufgabe entscheidet. Eine Vielzahl von Dimensionen erzeugt für die Befragten eine möglicherweise nicht mehr oder nur schwer handhabbare Komplexität. Die Folge wäre, dass die entsprechenden Urteile – falls es nicht zum vorzeitigen Abbruch kommt – im ungünstigsten Falle nur noch Artefakte darstellen. Jasso (2006) schlägt vor, nur solche Dimensionen auszuwählen, von denen eine Relevanz für die Bewertung bekannt ist. Dies kann durch theoretische Überlegungen, vorherige Untersuchungen oder aufgrund von Alltagsbeobachtungen geschehen. In Anknüpfung an kognitionspsychologische Arbeiten argumentiert sie zudem, dass Personen nur wenige Dimensionen zur Meinungsbildung heranziehen. Rossi und Anderson (1982) empfehlen, sich auf sechs Dimensionen zu beschränken. In den bislang durchgeführten Faktoriellen Surveys reicht die Anzahl der verwendeten Dimensionen unseres Wissens von drei (Berk/Rossi 1977) bis 21 (Shlay et al. 2005). In der Mehrzahl der Studien werden fünf bis sieben Dimensionen verwendet. Man stützt sich dabei allerdings nur auf eine „Daumenregel“ aus den Informations- und Kognitionswissenschaften, wonach Menschen sieben plus/minus zwei Informationen am besten verarbeiten können (Zimbardo 1988: 275). Es zeigt sich also, dass die bisherige Forschungspraxis durch sehr unterschiedliche Vorgehensweisen bestimmt ist. Die in der Literatur zu findenden Empfehlungen gehen über allgemeine Ratschläge nicht wirklich hinaus, etwa wenn Beck und Opp (2001: 287) raten, die Ausprägungen aus Hypothesen zu generieren und nur solche zu verwenden, bei deren Variation man einen tatsächlichen Einfluss vermutet.[6]

Die zunächst nahe liegende, grundsätzliche Annahme lautet, dass die kognitive Anforderung für die Befragten mit der Anzahl der Dimensionen steigt, bis hin zu einer eventuell nicht mehr handhabbaren Komplexität (Rossi/Anderson 1982; für Choice- und Conjoint-Analysen: DeShazo/Fermo 2002; Melles 2001). Weitaus weniger klar ist, wie sich die dann zu erwartende Tendenz zur

[6] Neben der Anzahl der Dimensionen ist auch die Zahl der Ausprägungen pro Dimension relevant, weil damit die Größe des *Vignettenuniversums* festgelegt wird. Als Vignettenuniversum wird die Gesamtheit aller möglichen Varianten der Situations- bzw. Objektbeschreibung bezeichnet.

Vereinfachung äußert. Neben einem kompletten Befragungsabbruch und Item-Nonresponses kommt ebenso ein inkonsistenteres Antwortverhalten in Frage. Alternativ sind Heuristiken in Form eines vollständigen Ausblendens inhaltlich weniger relevanter (oder vergleichsweise unauffällig operationalisierter, da z. B. mit weniger Ausprägungen vorgegebener) Dimensionen erwartbar (Wason et al. 2002; für Befunde bei Choice- und Conjoint-Analysen: DeShazo/Fermo 2002; Melles 2001; Swait/Adamowicz 2001). Vertreten wird bei Choice- und Conjoint-Analysen zudem auch die Gegenhypothese eines *konsistenteren* Antwortverhaltens bei mehr Dimensionen (Sauer 2009). Die dahinter stehende Annahme ist, dass in den wenig-dimensionalen Vignetten urteilsrelevante Informationen fehlen, die daher von den Befragten selbst konstruiert werden müssen.[7] Gegenüber der expliziten Vorgabe durch den Forscher bedeutet die „Unterkomplexität" eine geringere inhaltliche Kontrolle über das Vignettenexperiment, was zumindest befragtenübergreifend eine höhere Varianz und damit geringere Präzision der Schätzungen erwarten lässt (Caussade et al. 2005: 632; DeShazo/Fermo 2002; Johnson 2006: 46f.). Ähnlich wird vermutet, dass unkontrollierte *Framing*-Effekte wahrscheinlicher werden (dazu z. B. Melles 2001: 186). Und schließlich gilt auch ein Informationsmangel als kognitiv belastend, weil es beispielsweise bei wenigen Merkmalsvorgaben schwieriger ist, Unterschiede in den Fallbeispielen zu erkennen und damit zwischen ihnen zu differenzieren (s. Hensher 2006 für dieses Argument bei Choice-Experimenten). Als ein erster Beleg für einen solchen *information-underload* können die Befunde einer Wiederholungsbefragung gewertet werden, bei der Studierende zu drei Messzeitpunkten mit jeweils denselben Vignetten befragt wurden: Die Stabilität der Urteile erwies sich bei acht Dimensionen höher als bei fünf Dimensionen (Liebig et al. 2006).

Für alle Effekte ist jedenfalls unklar, ab welcher Dimensionszahl mit ihnen zu rechnen ist. Für die vorliegende Untersuchung wird daher mit fünf versus zwölf Dimensionen bewusst ein starker Kontrast gewählt. Die – gemessen an den vorliegenden Studien mit überwiegend fünf bis neun Dimensionen – überdurchschnittliche maximale Dimensionszahl von zwölf lässt ein Durchschlagen des „Überforderungseffektes" erwarten. Es ergeben sich zwei Teilhypothesen:

H_{1a}: *Bei zwölf Dimensionen sind Befragungsabbrüche häufiger als bei fünf Dimensionen.*

H_{1b}: *Das Urteilsverhalten ist bei zwölf Dimensionen inkonsistenter als bei fünf Dimensionen.*

Alternativ ist von einer vereinfachten Urteilsstrategie in Form einer Ausblendung einzelner Merkmale auszugehen (s. zu dieser *dimensional reductions*-Strategie bei Choice-Anaylsen: Swait/Adamowicz 2001: 137):

H_{1c}: *Bei zwölf Dimensionen sind einzelne Vignettenvariablen weniger urteilsrelevant, zeigen also geringere Einflüsse auf die Urteile als bei fünf Dimensionen.*

[7] In Vignettenstudien zur Einkommensgerechtigkeit könnte ein solches Informationsdefizit z. B. in der Berufserfahrung der Einkommensbezieher bestehen.

4.3.2 Lern- und Ermüdungseffekte

In fast allen Vignettenstudien sollen die einzelnen Befragten mehrere Vignetten beurteilen. Gängig sind zehn bis 20 Vignetten, in einer Studie waren es ganze 95 Vignetten pro einzelnem Befragten (Beck/Opp 2001; Rossi et al. 1974). Die mehrfache Präsentation von Vignetten ermöglicht es, selbst bei geringen Befragtenzahlen noch ausreichend viele Urteilszahlen zur Hypothesentestung zu sammeln (Auspurg et al. 2009). Zudem erlaubt sie, befragtenspezifische Urteils- und Entscheidungsregeln (so genannte *within subjects*-Effekte) aufzudecken. Mit der wiederholten Bewertungsaufgabe sind allerdings Lerneffekte zu erwarten, die mit anderen Kennzeichen der Erhebungssituation in Wechselwirkung stehen. Sehr deutlich ist dies bei der Anzahl der Dimensionen. Bei einer höheren Dimensionszahl benötigen Lernprozesse länger, gleichzeitig könnten Ermüdungserscheinungen früher einsetzen. Lern- und Ermüdungseffekte sind wechselseitige Aspekte von Komplexität. Beim Lernen geht es um ein zunehmend konsistentes Antwortverhalten sowie um das Vermögen, mehr Dimensionen gleichzeitig in ein Urteil zu integrieren.[8] Ermüdungs- und Langeweile-Effekte schlagen sich umgekehrt in einer sinkenden Konsistenz und in einer Beachtung weniger Merkmale oder anderen vereinfachten Entscheidungsregeln nieder (für Choice-Analysen: Carson et al. 1994: 335f.).[9] Die Rolle und das Ausmaß von Lern- und Ermüdungseffekten sind für Vignettenstudien bislang unerforscht. Ebenso ist es eine noch völlig ungeklärte Frage, ab welcher Vignettenzahl mit einem Umkippen von Lern- in Ermüdungseffekte zu rechnen ist.

Als ein erster Orientierungspunkt können Erfahrungen aus den verwandten Choice-Experimenten herangezogen werden. Demnach nimmt die Urteilskonsistenz bis etwa zum zehnten Urteil zu, um danach wieder abzusinken (z. B. Bradley/Daly 1994: 180; Caussade et al. 2005: 631f.). Da selbst bei Vignettenstudien mit 50 oder mehr Vignetten bislang keine nennenswerten Probleme im Hinblick auf die Urteilsgüte berichtet werden (Jasso 2006), scheint bei der vorliegenden Fallzahl von maximal zehn Vignetten pro Befragten (dazu mehr in Abschnitt 4.4) eine Dominanz der Lerneffekte plausibel. Es ergeben sich die folgenden Annahmen:

$H_{2a:}$ *Mit der Position der Vignetten steigt die Konsistenz des Antwortverhaltens und/oder die Anzahl berücksichtigter Dimensionen.*

H_{2b}: *Diese Lerneffekte treten stärker bei zwölf als bei fünf Dimensionen auf.*

[8] Eine im Befragungsverlauf zunehmende Beachtung von Dimensionen wird zudem damit begründet, dass die Probanden die in der Realität korrelierten Merkmale zu Beginn als redundant ansehen. Erst wenn sie nach einer ganzen Reihe von präsentierten Vignetten erkennen, dass sie im experimentellen Design unabhängig voneinander variieren, schenken sie ihnen mehr Aufmerksamkeit bzw. lassen sie separat in ihr Urteil einfließen (für Conjoint-Analysen: Melles 2001: 118).

[9] Grafisch ist also ein umgekehrt u-förmiger Zusammenhang zwischen der Bearbeitungsabfolge der Vignetten und der Konsistenz bzw. Anzahl berücksichtiger Dimensionen zu erwarten.

4.3.3 Behandlung unlogischer Fälle

Bevor die tatsächlich zu bewertenden Vignetten zusammengestellt werden (also eine Auswahl aus dem Universum aller möglichen Kombinationen von Merkmalsausprägungen getroffen wird; dazu Beck/Opp 2001; Dülmer 2007; Steiner/Atzmüller 2006), ist es bisher gängige Praxis, „unlogische" und „unplausible" Fälle zu eliminieren. Es werden also solche Vignetten ausgeschlossen, die offensichtlich ungewöhnliche oder unsinnige Merkmalskombinationen enthalten. Ein Beispiel dafür wären erwerbstätige Personen ohne Schul- oder Berufsausbildung in einem Beruf, bei dem eine Ausbildung unabdingbar ist (etwa Richter, Hochschullehrer). Der Ausschluss solcher Fälle wird vor allem mit den zu erwartenden Folgen für das Antwortverhalten begründet. Offensichtlich unsinnige Fälle würden die Ernsthaftigkeit der Bewertungsaufgabe in Frage stellen und zu einem Anstieg der *Item-Non-Response*-Quote, oder gar zum völligen Befragungsabbruch führen (Faia 1980; Jasso 2006).

Dieses Argument ist durchaus einleuchtend, doch sind die Kriterien, was als unlogisch oder unsinnig zu gelten hat, sehr vage. In vielen Faktoriellen Surveys geht es darum, möglichst unabhängig von den gängigen Normen, bestehenden Gesetzen und empirischen Beobachtungen Bewertungen vornehmen zu lassen, um so auch die kontrafaktischen Meinungen und Überzeugungen der Befragten zu erheben. Die Norm eines „logischen Falles" wird durch empirische Regelmäßigkeiten und damit zusammenhängende Erwartungshaltungen geprägt. Faktorielle Surveys bieten jedoch die seltene Möglichkeit, die Probanden bewusst mit abweichenden Fällen zu konfrontierten – und gerade in der Reaktion auf solche abweichende Fälle kann ein Erkenntnisziel liegen. In dieser Hinsicht sind Eingriffe in die Merkmalskombinationen problematisch, engen sie doch die Variation der Situations- und Objektbeschreibungen *a priori* auf ein empirisch vorfindbares Maß ein (Beck/Opp 2001).

Solides methodisches Wissen besteht bislang ausschließlich im Hinblick auf die *statistischen* Folgen. Durch den gezielten Ausschluss einzelner Fälle wird die Orthogonalität der Dimensionen im Vignettenuniversum eingeschränkt, Multikollinearität wird also erzwungen (zu deren Konsequenzen für Schätzverfahren: Greene 2003: 56-59; Wooldridge 2003: 96-100). Die Relevanz des Ausschlusses von Fällen für die Balanciertheit und Unkorreliertheit von Vignettensamples ist inzwischen gut einschätzbar (Dülmer 2007: 391f.; Kuhfeld et al. 1994: 551; Steiner/Atzmüller 2006) und es liegen Algorithmen vor, welche die Einbußen an Effizienz gezielt minimieren (dazu Kuhfeld 2005). Aufgrund des andernfalls hohen Effizienzverlustes lautet daher die eindeutige Empfehlung, diese Algorithmen auch einzusetzen.

Die Auswirkungen der unplausiblen oder unlogischen Fälle auf das *Antwortverhalten* sind dagegen weitaus strittiger, was vor allem durch fehlende einschlägige Untersuchungen bedingt ist.[10] Trifft die oben angesprochene Vermutung zu, dass durch unplausible Vignetten der grundsätzliche Glaube an den Wert der Befragung und damit den Nutzen eigener Mitarbeit beeinträchtigt wird, sind Befra-

[10] Die zwischen dem Autorenteam Rossi/Alves (1980) und Faia (1980) ausgetragene Diskussion über die (Un-)Sinnigkeit und den Nutzen unplausibler Vignetten ist daher nach wie vor nicht mit empirischen Argumenten zu entscheiden.

gungsabbrüche und invalide Antworten zu erwarten (Response-Sets oder flüchtige und inkonsistente Urteile). Es ergeben sich daher zunächst die folgenden Hypothesen:

H_{3a}: *Werden den Befragten unplausible Fälle vorgelegt, sind Befragungsabbrüche häufiger als wenn dies nicht der Fall ist.*

H_{3b}: *Werden die Befragten mit unplausiblen Fällen konfrontiert, ist die Konsistenz ihres Antwortverhaltens geringer, als wenn dies nicht der Fall ist.*

Faia (1980) erwartet zudem, dass die für die Unplausibilität ursächlichen Dimension in den Vordergrund geraten – die Befragten würden ihre Aufgabe in einen reinen „Intelligenztest" zur Entlarvung von „Anomalien" uminterpretieren. Gerade dies würde die Gültigkeit der Urteile beinträchtigen und verdient daher eine Überprüfung:

H_{3c}: *Nach einer Konfrontation mit unplausiblen Fällen beziehen die Befragten primär die für die Unplausibilität verantwortlichen Dimensionen in ihre Urteile ein, gewinnen diese somit relativ zu allen anderen Dimensionen an Bedeutung.*

Als alternative Begründung hierfür lässt sich ein Lerneffekt anführen: Die Befragten bemerken erst bei einer empirisch seltenen Kombination, dass die Merkmale unabhängig voneinander variieren und somit nicht redundant sind. Ähnlich könnte sich so eine sinkende Bereitschaft zu differenzierten Urteilen manifestieren: Dimensionen verlieren durch ein Umschwenken auf ein vereinfachtes, weniger Merkmale einbeziehendes und daher kognitiv weniger belastendes Antwortverhalten an Relevanz.

Die Diskussion dieser drei Problemstellungen verdeutlicht, dass komplexe Wechselwirkungen zwischen den methodischen Aspekten von Faktoriellen Surveys zu erwarten sind. Wir können hier schon aus Platzgründen nur die besonders nahe liegenden Zusammenhänge analysieren, im genannten DFG-Projekt wird derzeit ein weitaus größeres Spektrum methodischer Effekte untersucht.

4.4 Methodik und Datengrundlage

Die drei methodischen Probleme lassen sich nicht analytisch lösen, sondern erfordern eine empirische Herangehensweise. Ideal dazu ist ein Methodenexperiment, bei dem die Bedeutung von Designelementen für das Antwortverhalten durch ihre gezielte Variation beobachtbar wird. Wichtig ist, dass die methodischen Splits zufällig auf die Befragten verteilt werden und sie zudem nicht mit einzelnen Vignetten(decks) korreliert sind – wie bei jedem Experiment erlaubt erst diese Randomisierung, unbekannte Drittvariablen zu neutralisieren und ungewünschte Konfundierungen mit den inhaltlichen Dimensionen der Vignetten zu vermeiden.[11] In den vorliegenden Experimenten wird die Komplexität

[11] Bei Choice-Experimenten werden derartige Studien unter dem Namen „*Design of Design*" geführt (s. z. B. Caussade et al. 2005; Hensher 2004, 2006). Im Prinzip handelt es sich um eine mehrfaktorielle Erweiterung des

der Vignetten über die Zahl der Dimensionen variiert: Etwa die Hälfte der Befragten bekommt durchgehend Vignetten mit fünf, die andere mit zwölf Dimensionen vorgelegt (es handelt sich also um ein reines *between-subject*-Design).[12] Zunächst wurden jedem Teilnehmer sieben Vignetten zugeteilt; aufgrund der geringen Abbruchquote wurde diese Zahl in einer zweiten (kleineren) Befragungswelle auf zehn erhöht.

Als inhaltliche Fragestellung dient der besonders gut erforschte „Klassiker" unter den Vignettenstudien – die Erhebung von Einkommensgerechtigkeit (z. B. Alves/Rossi 1978; Hermkens/Boerman 1989; Jann 2003; Jasso/Webster 1997,1999; Shepelak/Alwin 1986). Den Befragten werden jeweils fiktive Personen vorgestellt, die sich in einer Reihe von einkommensrelevanten Merkmalen unterscheiden, wie dem Geschlecht, Alter, Bildungsstand oder Beruf. Zusätzlich enthält jede Vignette das monatliche Netto-Einkommen der beschriebenen Person. Dieses soll dann auf einer elf-stufigen Ratingskala danach beurteilt werden, ob und in welchem Ausmaß es (un-)gerecht erscheint. Abbildung 4.1 zeigt eine Beispielvignette mit zwölf Dimensionen. Die Ausprägungen der Dimensionen sind darunter im Überblick aufgeführt.[13] Bei der Auswahl der Merkmale wurde darauf geachtet, dass ihre Relevanz für das Urteilsverhalten bereits belegt ist. Damit sollte sichergestellt werden, dass eine mögliche Nicht-Beachtung methodisch (kognitive Überforderung) und nicht inhaltlich (genuine Einschätzung als irrelevant) zu deuten ist (ähnlich für Choice-Experimente: Hensher 2006: 16).

Um die zufällige Variation der experimentellen Splits und Vignetten mit verhältnismäßig wenig Aufwand umsetzen zu können, fiel die Wahl auf eine Online-Befragung. Ein weiterer Grund für diesen Befragungsmodus ist die gute Erfassbarkeit von Metadaten (z. B. Beantwortungszeiten), welche zusätzlichen Aufschluss über die Bearbeitungsstrategien bieten. Bei Experimenten kommt es nicht auf eine repräsentative und zufällige Stichprobe der Probanden an, sondern es sind zumindest bei kleinen Stichproben homogene Experimentalgruppen vorteilhaft (da diese ein geringeres Risiko ungleich verteilter Drittvariablen bergen; s. z. B. Diekmann 2007: 337ff.). Ihre relativ große Homogenität und gute Erreichbarkeit sprachen für die Wahl von Studierenden verschiedener Universitäten, die über E-Mailverteiler der Fachschaften kontaktiert und mit einem Link zur Befragung um ihre Teilnahme gebeten wurden.

---

split-ballot-Designs: Es werden gleich *mehrere* Designelemente unabhängig voneinander variiert (dazu Sniderman/Grob 1996).

[12] Dies ist nicht vollständig korrekt, denn als zweiter experimenteller Faktor wurde eine der Dimensionen, das Geschlecht der Vignettenpersonen, nur bei einem Teil der Befragten zwischen den Vignetten variiert (*within*-Variation). Den anderen Befragten wurden stets nur Vignetten eines Geschlechts vorgelegt (*between*-Variation), sie bewerteten also durchgehend jeweils nur Beschreibungen mit männlichen oder weiblichen Protagonisten, womit sich für sie die Anzahl variabler Dimensionen auf vier bzw. elf Merkmale reduziert. Der Hintergrund dieses Splits ist der, dass sich damit Effekte sozialer Erwünschtheit bzw. eines bewussten versus unbewussten Urteilsverhaltens untersuchen lassen. Da dieser Faktor aber vollständig unabhängig variiert wurde, kann er an dieser Stelle und den nachfolgenden Analysen ausgeblendet werden, er verdient eine eigenständige Betrachtung.

[13] Diese Aufstellung aller Dimensionen dient hier nur als Information für den Leser; den Befragten wurde diese Übersicht nicht vorgelegt.

Abb. 4.1: Beispielvignette mit zwölf Dimensionen

Ein 45-jähriger Mann ohne Berufsabschluss arbeitet seit 28 Jahren Vollzeit als Programmierer. Er ist erst kürzlich in das Unternehmen eingetreten und erbringt dort durchschnittliche Leistungen. Das Unternehmen mit insgesamt 5 Mitarbeitern ist vom Konkurs bedroht.

Er ist gesund und hat 4 Kinder.

Sein Einkommen beträgt monatlich 1.700,- Euro (Netto).

Wie gerecht stufen Sie das Einkommen der beschriebenen Person ein? Es ist…

Viel zu niedrig					gerecht					Viel zu hoch
-5	-4	-3	-2	-1	0	1	2	3	4	5
○	○	○	○	○	○	○	○	○	○	○

Vignettendimensionen und Ausprägungen:

1) Alter: 25, 35, 45, 55 Jahre
2) Geschlecht: Mann, Frau
3) Berufsabschluss: ohne Berufsabschluss, mit abgeschlossener Berufsausbildung, mit Hochschulabschluss
4) Beruf: 10 Ausprägungen von Hilfsarbeiter/in bis Anwalt/Anwältin (Auswahl n. Dezentilen der *Magnitude*-Prestige-Skala)
5) Einkommen: 10 Ausprägungen von 250,- bis 15.000,- Euro (Netto)

6) Berufserfahrung: keine, 25%, 50%, 100% der potenziellen Erwerbszeit
7) Betriebszugehörigkeit: erst kürzlich eingetreten, schon seit langem im Unternehmen beschäftigt
8) Leistung: unterdurchschnittlich, durchschnittlich, überdurchschnittlich
9) Betriebsgröße: 5, 20, 200, 2.000 Mitarbeiter
10) Wirtschaftliche Lage des Unternehmens: vom Konkurs bedroht, ausgeglichene Bilanz, hohe Gewinne
11) Gesundheitszustand: gesund, 30% schwerbehindert
12) Kinder: 6 Ausprägungen von keine bis 5 Kinder

Bei den Vignetten handelt es sich um eine fraktionalisierte Auswahl aus dem kompletten Universum für zwölf Dimensionen, wobei auf eine Orthogonalisierung aller Haupteffekte geachtet wurde (so genanntes *resolution III-Design*, dazu Kuhfeld et al. 1994: 546). Mit dieser Anforderung sind bei der vorliegenden Spezifikation von Dimensionen und Ausprägungen etwa 100 Vignetten für eine effiziente Stichprobe hinreichend.[14] Durch den Ausschluss logisch unmöglicher Kombinationen (wie Personen ohne Berufserfahrung, die schon lange im Betrieb arbeiten), reduzierte sich das Sample weiter zu insgesamt 93 unterschiedlichen Vignetten (empirisch seltene, aber gleichwohl mögliche Fälle wurden dagegen bewusst beibehalten – mehr dazu unten). Bei dem Split mit fünf Dimensionen wurde exakt dieselbe Vignettenstichprobe eingesetzt (es wurden einfach die überflüssigen Dimensionen gelöscht). Zwar ließen sich für diese „sparsameren" Vignetten weitaus effizientere Designs bilden, gerade diese statistischen Effizienzwerte sollten aber konstant gehalten werden, um eine reine Abschätzung der *methodischen* Effekte zu ermöglichen. Nur unter Kontrolle der statistischen Effizienz lassen sich Unterschiede in den Signifikanzen von Regressionskoeffizienten tatsächlich auf das Antwortverhalten zurückführen.[15] Zugleich wird mit der Verwendung identischer Vignettensamples

[14] Es wird eine D-Effizienz von 98,2 erreicht, wobei Werte über 90 als zufrieden stellend gelten (Kuhfeld 2005). Allerdings reduziert sich die Effizienz mit dem Ausschluss unlogischer Fälle wieder.

[15] Schließlich ist für die Präzision der Schätzungen die statistische Effizienz der Vignettenstichprobe ähnlich wichtig wie die „kognitive Effizienz" der von den Befragten abgegebenen Urteile (für entsprechende Argumente in Bezug auf Choice- und Conjoint-Analysen: Louviere 2001b; Melles 2001: 109).

für die Splits mit fünf und zwölf Dimensionen einer Vermischung von inhaltlichen und Designeffekten vorgebeugt. Es lassen sich durch dieses Vorgehen auftretende Unterschiede im Antwortverhalten eindeutiger auf die differente Anzahl an Dimensionen zurückführen, anstatt auf unterschiedliche inhaltliche Kombinationen der Vignettendimensionen.

Alle Teilnehmer wurden zufällig einem der beiden methodischen Splits sowie einem Subset an Vignetten zugewiesen. Pro Befragten wurde eine eigene Zufallsziehung von Vignetten (Ziehung ohne Zurücklegen) vorgenommen. Mit dieser randomisierten Setbildung sollte eine möglichst hohe Ausschöpfung der Stichprobe von 93 Vignetten gewährleistet werden. Zudem wurde eine befragtenspezifische, zufällige Reihenfolge der Vignetten gewählt, um Kontrast- und Reihenfolge-Effekte auszuschließen: „Extreme“ Vignetten verteilen sich dann zufällig auf die Bearbeitungspositionen, womit über alle Befragten hinweg zu beobachtende Einflüsse der Reihenfolge eindeutiger als Lern- bzw. Ermüdungseffekte zu deuten sind. Die befragtenspezifische Zufallsauswahl von Vignetten hat zudem den Vorteil, dass sich automatisch weitere methodische Variationen zwischen den Befragten ergeben, etwa im Auftreten und in der Häufigkeit von unplausiblen Fällen.[16]

Die Befragung fand im Zeitraum Dezember 2007 bis März 2008 statt. Die Vignetten wurden in einen Rahmenfragebogen integriert, in dem neben soziodemografischen Merkmalen politische und soziale Einstellungen über „klassische“ Itemabfragen erhoben wurden. Den Befragungslink haben 558 Personen aufgerufen, für die Vignetten liegen 3.480 Urteile von insgesamt 460 Probanden vor.[17] Tabelle 4.1 zeigt die für die einzelnen experimentellen Varianten realisierten Fallzahlen.

[16] Bei der Alternative einer bewussten bzw. fraktionalisierten Setbildung wären zwar Konfundierungen besser kontrollierbar, aber angesichts der geringen Setgröße von sieben bzw. zehn Vignetten auch nicht vermeidbar – gerade für die komplexere Variante mit zwölf Dimensionen wären der Preis unweigerlich starke Kontexteffekte der einzelnen Sets (selbst Haupteffekte wären innerhalb der einzelnen Sets untereinander korreliert). Aus diesen Gründen ist der Einsatz einer möglichst hohen Anzahl an unterschiedlichen Sets vorzuziehen, zumal angesichts des homogenen Samples und der hohen Befragtenzahl die Gefahr der Konfundierung von Vignetten- mit Befragtenmerkmalen gering erscheint (s. Steiner/Atzmüller 2006 für eine ausführliche Diskussion der Vor- und Nachteile unterschiedlicher Setbildungen). Hinzu kommt, dass fraktionalisierte Setbildungen einem der Analyseziele zuwiderlaufen: Sie arbeiten mit einer möglichst gleichmäßigen Verteilung von Extremfällen, was impliziert, dass auch unplausible Fälle sehr regelmäßig auf die Sets bzw. Befragten verteilt werden und daher die *between*-Varianz zu gering ausfallen dürfte, um ihren Einfluss verlässlich zu prüfen. Insgesamt lassen diese Abwägungen somit bei den vorliegenden Analysezielen eine randomisierte Setbildung als vorteilhaft erscheinen. Das mit ihr verbundene Risiko einer unbalancierten Verteilung von Vignetten auf die Splits mit fünf versus zwölf Dimensionen (bzw. sieben versus zehn Vignetten) ist angesichts der hohen Set- und Befragtenzahlen gering. Problematisch wären für die angestrebten Analysen insbesondere Unterschiede in den Korrelationsstrukturen. Diese stimmen jedoch in der Tat sehr gut zwischen den einzelnen Splits überein, wie die im Anhang dieses Kapitels aufgeführten Korrelationsmatrizen belegen (Tabellen 4.A1 und 4.A2).

[17] Aufgrund der verwendeten Samplingprozedur lassen sich keine Rücklaufquoten berichten. An dieser Stelle ist nochmals zu betonen, dass wir lediglich einen experimentellen Hypothesentest, nicht aber deskriptive Aussagen zu Gerechtigkeitseinstellungen anstreben. Dafür scheint der Verzicht auf eine Zufallsstichprobe unproblematisch. Mehrfachteilnahmen wurden so gut wie möglich durch technische Vorkehrungen ausgeschlossen.

Tabelle 4.1: Realisierte Fallzahlen für Vignettenurteile und Befragte[a]

	5 Dimensionen		12 Dimensionen		Σ	
	Vignetten	**Befragte**	**Vignetten**	**Befragte**	**Vignetten**	**Befragte**
Sieben Vignetten pro Befragten	1.213	176	1.109	162	2.322	338
Zehn Vignetten pro Befragten	574	59	584	63	1.158	122
Σ	1.787	235	1.693	225	3.480	460

[a] Nur Befragte, die mindestens eine Vignette beantwortet haben.

Bei der Datenauswertung ist die Mehrebenenstruktur zu beachten. Werden einzelnen Befragten mehrere Vignetten vorgelegt, entsteht ein hierarchischer Datensatz (s. Beck/Opp 2001 für eine anschauliche Darstellung). Auf der untersten Ebene stehen die Vignettenurteile, eine zweite Analyseebene bilden die Merkmale der Befragten. Da wir nur auf die Analyse von Vignettendimensionen (der ersten Ebene) abstellen und zudem ein homogenes Befragtensample verwenden, berücksichtigen wir die Datenstruktur lediglich durch die Schätzung von robusten Standardfehlern (Wooldridge 2002, 2003: 258ff.; zur Modellwahl speziell bei Vignettenstudien: Auspurg et al. 2009; Hox et al. 1991; Jasso 2006). Befragtenspezifische Schwankungen der Urteile und ihre mögliche Erklärung interessieren hier nicht. Die für die einzelnen Hypothesen eingesetzten Analysestrategien und Operationalisierungen werden im folgenden Abschnitt erläutert.

4.5 Ergebnisse

4.5.1 Deskriptive Befunde

Bevor die Hypothesen mit multivariaten Analysen geprüft werden (Abschnitt 4.5.2), gibt ein Blick auf die deskriptiven Verteilungen und Rücklaufquoten erste Aufschlüsse über das Antwortverhalten. Insgesamt haben 124 der 558 Teilnehmer (22,2 Prozent) die Umfrage nicht beendet. Die Abbrüche konzentrieren sich zu einem sehr großen Teil auf die Begrüßungsseite oder den Rahmenfragebogen vor den Vignetten; direkt im Vignettenteil haben lediglich 23 Befragte (4,1 Prozent der Gesamt-Teilnehmerschaft) abgebrochen, im anschließenden Befragungsteil sind es weitere 19 Personen (3,4 Prozent). Eine Differenzierung der Abbrüche nach experimentellen Splits erscheint angesichts dieser geringen Fallzahlen kaum sinnvoll. Festhalten lässt sich jedenfalls, dass selbst die umfangreiche Bewertungsaufgabe bei zwölf Dimensionen (der immerhin ca. die Hälfte der Befragten ausgesetzt war) und das Auftreten ungewöhnlicher Fälle (wie Anwälten ohne Hochschulabschluss) nicht zu auffallend hohen Abbruchquoten führen. Dies gilt ähnlich für Antwortverweigerungen: Lediglich 68 Vignetten, damit 1,9 Prozent blieben unbeantwortet.[18]

[18] Diese Quote an *Missings* entspricht etwa der von „herkömmlichen“ Itemabfragen in der gleichen Erhebung. Die Befragten wurden direkt im Anschluss an die Vignetten gebeten, die Bedeutung der Vignettendimensionen für eine gerechte Entlohnung jeweils einzeln auf sieben-stufigen Itemskalen einzustufen (von sollte „überhaupt keine Bedeutung“ bis sollte „sehr große Bedeutung“ spielen). Die *Missings* bewegen sich bei diesen Items zwischen 0,9 und 2,2 Prozent, im Mittel sind es 1,2 Prozent (vorherige Befragungsabbrüche nicht mitgezählt).

Die vorangegangenen Ausführungen haben jedoch bereits gezeigt, dass sich eine mangelnde Kooperationsbereitschaft oder Überforderung ebenso in einem veränderten Antwortverhalten bei fortgesetzter Befragung niederschlagen kann – speziell dessen Verkennung wäre für die Ergebnisinterpretationen kritisch.[19] Einen ersten Hinweis auf mögliche Response-Sets liefern die Verteilungen der Vignettenurteile, wie sie in Tabelle 4.2 für die unterschiedlichen experimentellen Splits aufgeschlüsselt sind. Über alle Befragten hinweg (mittlere Spalte) als auch pro Befragten berechnet (letzte Spalte), wird eine etwas geringere Streuung (Standardabweichung) der Vignettenurteile, damit stärkere Konstanz des Antwortverhaltens bei zwölf gegenüber fünf Dimensionen offensichtlich. Allerdings verfehlt dieser Unterschied das Signifikanzniveau von 5 Prozent.[20]

Tabelle 4.2: Deskriptive Übersicht über die Vignettenurteile[a]

Experimentelle Variante	N	Mittelwert	SD	Mittlerer Mittelwert pro Befragten	Mittlere SD pro Befragten
5 Dimensionen, 7 Vignetten	1.213	5,21	3,10	5,20	2,94
5 Dimensionen, 10 Vignetten	574	5,44	3,21	5,45	3,12
12 Dimensionen, 7 Vignetten	1.109	5,51	2,96	5,51	2,87
12 Dimensionen, 10 Vignetten	584	5,36	2,98	5,35	2,86

[a] Skala von 1 ‚ungerechterweise zu niedrig' bis 11 ‚ungerechterweise zu hoch'. Der Wert ‚Sechs' kennzeichnet eine als gerecht empfundene Entlohnung.

4.5.2 Multivariate Analysen

Erwartete Folgen einer zu hohen Komplexität sind ein inkonsistenteres Antwortverhalten (H_{1b}), statistisch eine geringere erklärte Varianz bzw. höhere Fehlervarianz, und ein Ausblenden einzelner Dimensionen (H_{1c}), was sich statistisch in geringeren Einflussstärken bzw. weniger signifikanten Effekten äußert. Zur Prüfung dieser beiden Annahmen dienen die in Tabelle 4.3 aufgeführten Ordinary Least Square- (OLS-)Regressionen, die wegen der hierarchischen Datenstruktur jeweils mit robusten Standardfehlern geschätzt sind. Um die methodischen Effekte besser von möglichen Drittvariableneffekten trennen zu können, werden die Regressionen für die „Zwölfer-Vignetten" ohne (Modell 2) und mit (Modell 3) Kontrolle der zusätzlichen Dimensionen präsentiert.[21]

[19] "If tasks are too long or too difficult or lack sufficient realism and credibility, data quality will suffer in the sense of not containing the information sought. Unfortunately, respondents generally answer the questions asked and seldom go out of their way to point out problems with tasks posed" (Carson et al. 1994: 355).

[20] T-Test für die Mittelwertdifferenz der befragtenspezifischen Urteilsvarianz zwischen fünf und zwölf Dimensionen: t = 1,48; p = 0,140 bei zweiseitigem Test und Adaption für die verletzte Annahme der Varianzhomogenität (vorherige Prüfung mit *Levene's* Test).

[21] Prinzipiell sind die Vignettendimensionen bei fraktionalisierten Auswahlen zwar unkorreliert, geben sie also ihren reinen „Nettoeffekt" selbst dann wieder, wenn nicht auf Drittvariablen kontrolliert wird. Gerade hierin liegt ja eine wesentliche Stärke dieses Verfahrens. Einschränkung erfährt dieser Aspekt allerdings mit dem gezielten Ausschluss von Kombinationen, da dieser unweigerlich zu Korrelationen führt. Dies trifft auch auf das vorliegende Sample zu, da aus diesem die logisch völlig unmöglichen Fälle ausgeschlossen wurden (wie z. B. Per-

Zunächst zur inhaltlichen „Lesart“ der Ergebnisse: Bei der vorliegenden Kodierung der abhängigen Variablen bedeuten positive (negative) Koeffizientenwerte, dass das Einkommen als ungerechterweise zu hoch (niedrig) empfunden wird. Negative Effekte lassen sich somit als eine Erhöhung des als angemessen empfundenen Nettoeinkommens deuten. Nach allen drei Modellen wird beispielsweise Personen mit einem Berufsabschluss ein höheres Einkommen zugestanden als solchen ohne Abschluss. Für unser methodisches Forschungsinteresse ist aber interessanter, ob sich Unterschiede zwischen den Koeffizientenwerten der drei Modelle zeigen. Dies ist im Hinblick auf die Vorzeichen nicht der Fall, jedoch zeigen die Vignettenmerkmale bei den komplexeren, zwölfdimensionalen Varianten oftmals einen betragsmäßig schwächeren Einfluss. Ein Chow-Test bestätigt signifikante Differenzen zwischen den Modellen 1 und 2 ($F = 4{,}04$ bei $df = 7$ und 459; $p = 0{,}000$).[22] Einzeln geprüft erweisen sich die Einflussstärken des Hochschulabschlusses und des Prestiges als signifikant verschieden. Da gerade die Einflüsse dieser beiden Variablen bei Kontrolle für die weiteren Dimensionen stabil bleiben (die Koeffizienten unterschieden sich nur marginal zwischen Modell 2 und 3), ist dieser Unterschied nicht durch Drittvariableneffekte bedingt, sondern deutet er vielmehr darauf hin, dass mit höherer Komplexität tatsächlich Dimensionen tendenziell ausgeblendet werden.[23] Die Anteile erklärter Varianz (R^2-Werte), welche als Maß für die Konsistenz des Antwortverhaltens herangezogen werden können, unterscheiden sich dagegen nicht substanziell zwischen den Modellen.[24]

sonen ohne Berufserfahrung, die schon lange in einem Betrieb arbeiten, vgl. Abschnitt 4.4). Eine Übersicht über die Korrelationen zwischen den einzelnen Dimensionen findet sich in Tabelle 4.A3 im Anhang dieses Kapitels.

[22] Technisch besteht dieser Test darin, ein *gepooltes* Modell zu schätzen, in das zusätzlich eine Dummy-Variable für die zu prüfende Designvariante (hier die Anzahl der Dimensionen) sowie Interaktionsterme aller Vignettendimensionen mit dieser Designvariante aufgenommen werden. Geprüft wird dann, ob die Aufnahme dieser Variablen *insgesamt* zu einer signifikanten Modellverbesserung führt; im vorliegenden Falle einer OLS-Regression, ob es zu einem signifikanten Anstieg der erklärten Varianz kommt (s. Wooldridge 2003: 238f. für Details).

[23] Der Vergleich zwischen fünf- und zwölfdimensionalen Vignetten ist statistisch nicht trivial. Mit einer höheren Variablenzahl steigt automatisch die Wahrscheinlichkeit von Korrelationen der Variablen untereinander oder von Konfundierungen mit Wechselwirkungen. Aufgrund der hohen Anzahl an möglichen Wechselwirkungen (bei der Variante mit zwölf zum Teil kategorialen Dimensionen liegen allein mehr als 70 mögliche Interaktionen erster Ordnung vor) sind diese nicht alle modellierbar (mitunter wird dies bereits durch die Stichprobenbildung verhindert). Damit ist nicht gänzlich auszuschließen, dass die Effekte im Falle der höher dimensionalen Vignetten leicht verzerrt geschätzt werden (*omitted variable bias*; wir danken Peter Steiner für diesen wertvollen Hinweis). Es sollten daher künftig nochmals Replikationen mit anderen Vignettenstichproben durchgeführt werden. Darauf wird in der Schlussbetrachtung (Abschnitt 4.6) zurückgekommen.

[24] Der Vergleich von R^2-Werten zwischen Modellen ist nicht unproblematisch (Wooldridge 2003). Im vorliegenden Fall scheinen die Voraussetzungen jedoch erfüllt: Die Fallzahlen sind etwa vergleichbar und ebenso bestehen nur minimale Unterschiede in der Varianz der abhängigen Variablen.

Tabelle 4.3: OLS-Regressionen der Vignettenurteile[a] (robuste Standardfehler in Klammern; signifikante Unterschiede der Koeffizienten zwischen Modell 1 und 2 hervorgehoben)[b]

	Modell 1 **5 Dimensionen**	**Modell 2** **12 Dimensionen**	**Modell 3** **12 Dimensionen**
Weibliche Vignettenperson	-0,057 (0,122)	-0,136 (0,115)	-0,105 (0,113)
Alter [Jahre]	-0,021*** (0,005)	-0,029*** (0,005)	-0,020*** (0,005)
Abschluss (Ref.: kein Abschluss)			
- Berufsabschluss	-0,654*** (0,133)	-0,472*** (0,131)	-0,429*** (0,129)
- Hochschulabschluss	-1,126*** (0,129)	-0,623*** (0,126)	-0,830*** (0,130)
Berufsprestige [10 MPS-Score]	-0,157*** (0,011)	-0,097*** (0,012)	-0,106*** (0,012)
Nettoeinkommen [100,- Euro]	0,060*** (0,002)	0,055*** (0,002)	0,058*** (0,002)
Berufserfahrung [Prozent der potenziellen Erwerbszeit]			0,066 (0,048)
Schon seit langem im Betrieb beschäftigt (Ref.: erst seit kurzem)			-0,645*** (0,131)
Leistung (Ref.: unterdurchschnittlich)			
- Durchschnittlich			-0,813*** (0,129)
- Überdurchschnittlich			-0,788*** (0,138)
Anzahl Mitarbeiter [100]			0,028*** (0,006)
Betriebssituation (Ref.: vom Konkurs bedroht)			
- Ausgeglichene Bilanz			-0,037 (0,130)
- Hohe Gewinne			-0,292** (0,122)
Zu 30% schwerbehindert (Ref.: gesund)			0,049 (0,114)
Anzahl Kinder			-0,152*** (0,029)
Konstante	6,465*** (0,280)	6,274*** (0,236)	6,820*** (0,272)
Beobachtungen (Befragte)	1.787 (235)	1.693 (225)	1.693 (225)
R^2	0,47	0,45	0,49

[a] Skala von 1 ‚ungerechterweise zu niedrig' bis 11 ‚ungerechterweise zu hoch'. Der Wert ‚Sechs' kennzeichnet eine als gerecht empfundene Entlohnung.
[b] Prüfung mittels Interaktionstermen zwischen den Vignettendimensionen und der Dimensionszahl in einem *gepoolten* Modell, Signifikanzniveau von 5 Prozent.
*** p<0,01; ** p<0,05; * p<0,1 bei zweiseitigem Test; Schätzungen mit robusten Standardfehlern.

Insgesamt wird unsere erste Hypothese (H_{1a}, H_{1b}, H_{1c}) damit nur in dem Teilaspekt H_{1c} bestätigt.[25] Die Anzeichen für eine kognitive Überforderung sind – trotz der hohen Dimensionszahl – gering. Solange nicht Wege gefunden werden, für die Komplexität von Vignetten zu kontrollieren, sollten absolute Effektstärken dennoch vorsichtig interpretiert werden (für sich genommen und beim Vergleich von Studien). An dieser Stelle ist auf eine weitere, bei anderen Autoren zu findende, problematische Interpretation zu verweisen: Oftmals werden „hohe" R^2-Werte als Beleg dafür gewertet, dass es gelungen sei, alle für die Befragten relevanten Merkmale in die Vignetten aufzunehmen (es bleibt kaum mehr etwas unerklärt, somit seien alle urteilsrelevanten Informationen berücksichtigt; s. z. B. Beck/Opp 2001: 302). Wie unsere Ergebnisse zeigen, kann dies ein Trugschluss sein, denn die hinzukommenden Merkmale in Modell 3 erweisen sich fast ausnahmslos als signifikant, ohne dass es zu einem bedeutenden Anstieg der Varianzaufklärung käme. Eine hohe Modellanpassung ist somit zwar ein Maß für ein in sich konsistentes Urteilsverhalten, damit aber noch nicht unbedingt ein Indikator dafür, dass *alle* inhaltlich relevanten Dimensionen berücksichtigt sind.[26]

Zu beachten ist ferner, dass unseren Befragten mit maximal zehn Vignetten vergleichsweise wenige Urteile abverlangt wurden. Möglicherweise fallen kognitive Überforderungen und Ermüdungen erst bei weitaus höheren Vignettenzahlen ins Gewicht, oder schwächen sich umgekehrt mit zunehmender Übung ab. Damit sind die Hypothesen $H2_a$ und $H2_b$ angesprochen, die eine mit der Beantwortungssequenz zunehmende Konsistenz des Antwortverhaltens postulieren, speziell bei den komplexeren „Zwölfer"-Vignetten. Zur Überprüfung stellen wir Regressionsschätzungen getrennt für die einzelnen Bearbeitungspositionen der Vignetten an.

In Abbildung 4.2 sind die resultierenden R^2-Werte für die beiden Designvarianten (fünf- versus zwölfdimensional) gegen die Positionen der Vignetten abgetragen (dunklere, obere Linien). Da diese ebenfalls Aufschluss über Lern- bzw. Ermüdungseffekte geben, sind zugleich die mittleren Bearbeitungszeiten pro Vignette dargestellt (untere, bzw. hellere Linien).[27] Was die Varianzaufklärung bzw. R^2-Werte betrifft, ist im Bearbeitungsverlauf ein leichter Anstieg zu erkennen. Die durchschnittliche Bearbeitungszeit pro Vignette sinkt dagegen insbesondere nach der ersten Vignette sprunghaft und mit abnehmender Rate weiter bis zur siebten Vignette. Zusammen genommen deutet dies auf einen Lerneffekt hin: Die Befragten können die Vignetten in zunehmend kürzerer Zeit beantworten, ohne dass es zu Einbußen ihrer Antwortkonsistenz käme. Entgegen unserer Erwartung (H_{2b}) gilt dies nicht

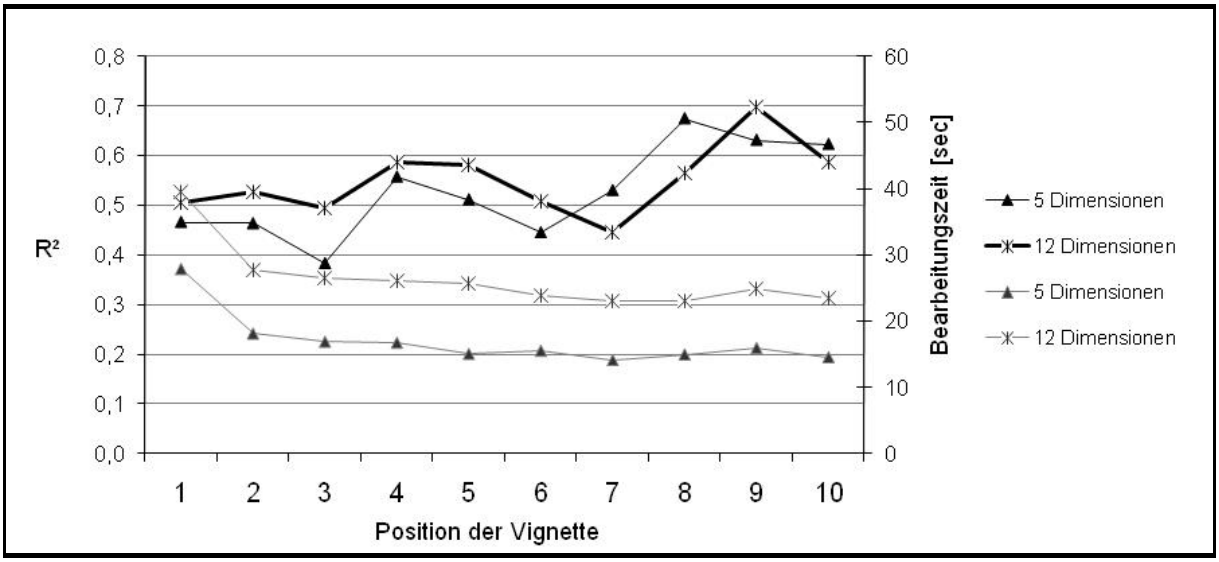
[25] Wobei sich die These H_{1a} (häufigere Befragungsabbrüche bei höherdimensionalen Vignetten) aufgrund der geringen Abbruchquoten nicht *statistisch* prüfen lässt.

[26] Womit auch Aussagen wie die folgende eine Relativierung finden: „The factorial survey method makes it possible to assess the number and identity of the characteristics a person uses in reaching a judgement" (Jasso 2006: 342).

[27] Die verwendete Online-Programmierung erlaubt es, die Bearbeitungszeit pro Vignette auf die Sekunde genau zu messen; exakter handelt es sich um die Zeit, die zwischen dem Abschicken der jeweiligen Vignettenseite und der Beendigung der vorherigen Seite verstrichen ist. Für derartige Zeitmessungen ist die bei Online-Befragungen geringe Kontrolle über das Setting nachteilig: Pausen der Befragten werden unvermeidlich mit zur Bearbeitungszeit gerechnet. Aus diesem Grunde wurde jeweils das obere 5-Prozent-Perzentil der Antwortzeiten aus den Berechnungen ausgeschlossen (s. für die grundsätzliche Empfehlung einer Bereinigung um *Outliers* bei Befragungszeiten: Mayerl et al. 2005; Urban/Mayerl 2007).

verstärkt für die komplexeren Vignetten: Die Linien für die fünf- und zwölfdimensionalen Vignetten verlaufen jeweils parallel zueinander, was bedeutet, dass die Lerneffekte für beide Versionen etwa gleich stark ausfallen. Für die vermutete Wechselwirkung zwischen Komplexitäts- und Lerneffekten findet sich also kein Beleg.

Abb. 4.2: R^2-Werte (dicker gedruckte, obere Linien) und Bearbeitungszeiten pro Vignette (schwächere, untere Linien) in Abhängigkeit von der Position der Vignette und Anzahl ihrer Dimensionen



Um die Interpretation als einen Lerneffekt abzusichern, ist zusätzlich noch zu prüfen, ob die steigende (oder zumindest gleich bleibende) Konsistenz nicht einer verstärkten Ausblendung von Dimensionen, also einer vereinfachten Entscheidungsheuristik, geschuldet ist. Um dies auszuschließen, wurden separate Regressionen mit dem ersten, zweiten und letzten Drittel der Vignetten berechnet. Die hier aus Platzgründen nicht dargestellten Modellschätzungen unterscheiden sich nicht signifikant voneinander, d. h. die Anzahl einflussreicher Dimensionen, ihre Effektstärken und allgemein das Antwortmuster bleiben in der Bearbeitungssequenz stabil.[28] Trotz der hohen Komplexität von zwölf Dimensionen führen also bereits die ersten Vignettenurteile zu sehr reliablen Urteilen – was bedeutet, dass sie nicht als „Übungsfälle“ betrachtet werden müssen, die aus den Ergebnisanalysen ausgeschlossen werden sollten (zu einer solchen Empfehlung bei Choice-Analysen: Caussade et al. 2005: 632, Anm. 6). An dieser Stelle ist aber darauf hinzuweisen, dass sich diese Aussagen nicht über die hier vorliegende, geringe Vignettenzahl und das sehr alters- und bildungshomogene Befragtensample

[28] Entsprechende Chow-Tests fallen nicht signifikant aus. Einzeln betrachtet nehmen die Dimensionen mit den Vignettenpositionen in ihren Effektstärken tendenziell zu (wiederum Vergleich des ersten mit den beiden anderen Dritteln an Vignetten), kommt es also zu einer immer stärkeren Beachtung der Dimensionen, was den Lerneffekt noch untermauert. Allerdings wird die Signifikanzschwelle von 5 Prozent keinesfalls erreicht. Auch finden sich keine signifikanten Modellunterschiede, wenn die Berechnungen getrennt für die beiden Splits mit fünf und zwölf Dimensionen wiederholt werden.

hinaus verallgemeinern lassen. Ob bei höheren Vignettenanzahlen oder anderen Befragtengruppen nicht doch Ermüdungserscheinungen durchschlagen, bleibt künftigen Untersuchungen vorbehalten.

Die Wirkung unplausibler Fälle kann ähnlich untersucht werden. Dafür ist zunächst eine Festlegung erforderlich, was überhaupt als „unplausibel" zu gelten hat (vgl. Abschnitt 4.3.3) – aufgrund der hier bestehenden, fließenden Übergänge und sicherlich auch subjektiv differierenden Einschätzungen keine triviale Aufgabe. Um möglichst objektiv vorzugehen, ziehen wir die Rückmeldungen aus ca. 60 mündlichen Pretestinterviews heran, die mit demselben Vignettensample im Herbst 2007 im Rahmen eines Forschungs-Projektseminars an der Universität Konstanz durchgeführt wurden. Für diese Interviews wurden bewusst sehr heterogene Personen der Allgemeinbevölkerung ausgewählt. Besonders häufig monierten die Befragten die Unsinnigkeit von Kombinationen der beiden Dimensionen Beruf und Ausbildungsabschluss: Speziell Anwälte ohne Ausbildung oder Hochschulabschluss sorgten oftmals geradezu für Verärgerung.[29] Wir werten entsprechend alle Vignetten als unplausibel, bei denen die geschilderten Personen nicht über einen Ausbildungs- oder Hochschulabschluss verfügen, der für ihren Beruf in Deutschland eigentlich Voraussetzung oder zumindest sehr üblich wäre.[30] Dies trifft auf insgesamt 22,5 Prozent (N = 800) unserer Vignetten zu.

Wir ziehen zudem eine alternative Operationalisierung auf Basis der Dimensionen Einkommen und Beruf heran. Speziell für Vignetten mit einem *berufsspezifisch* ungewöhnlichen Einkommen finden sich ebenfalls mehrere Pretestkommentare, welche deren Ernsthaftigkeit in Zweifel ziehen (z. B. sorgten Vollzeit arbeitende, leitende Manager mit einem monatlichen Gehalt von nur 250 Euro Netto für starke Irritationen, oder Friseure mit 15.000 Euro Netto). Überdies besteht für diese beiden Dimensionen wiederum die Möglichkeit einer weitgehend objektiven Definition. Unplausible Vignetten lassen sich durch einen Abgleich der Vignetteneinkommen mit den realen Nettoeinkommen der jeweiligen Berufsgruppen identifizieren. Dazu bestimmen wir zunächst anhand der Stichprobe des Sozio-ökonomischen Panels (SOEP) von 2007 die mittleren tatsächlichen Nettoeinkommen der zehn in den Vignetten verwendeten Berufsgruppen und berechnen dann für jede Vignette die absolute Differenz zwischen ihrem „virtuellen" Vignetteneinkommen und dem tatsächlichen mittleren Berufseinkommen nach dem SOEP. Atypisch sind dann Fälle mit einer betragsmäßig besonders großen Differenz nach oben oder unten; konkret werten wir alle Vignetten mit einer absoluten Abweichung von mindestens 3.000 Euro als unplausibel (das sind 23,9 Prozent aller Vignetten).[31]

---

[29] Es fielen Äußerungen wie „man fühle sich auf die Palme gebracht"; „wer denkt sich so einen Unsinn aus". Befragt wurden Personen unterschiedlichen Alters und Bildungsgrades, die Rekrutierung und Durchführung der Interviews geschah durch die Seminarteilnehmer – ihnen sei an dieser Stelle unser Dank ausgesprochen.

[30] Konkret sind dies: Anwälte ohne Hochschulabschluss oder nur mit Berufsabschluss; Verwaltungsfachkräfte, Elektroingenieure, Sozialarbeiter, Lokführer und leitende Manager ohne Ausbildung. Bei den übrigen (allesamt auch nicht gesetzlich geschützten) Berufsangaben (Friseure, Pförtner, Programmierer und ungelernte Arbeiter) scheinen dagegen Tätigkeiten ohne Ausbildungsabschluss plausibler.

[31] Neben dieser Definition über das Quartil haben wir Kontrollrechnungen mit dem Quintil und Dezentil durchgeführt – also mit einer noch stärkeren Eingrenzung unplausibler Fälle. Diese bestätigen unsere Befunde vollständig und werden daher nicht separat dargestellt. Zur Berechnung der Netto-Berufseinkommen wurde die generierte Einkommensvariable des SOEP 2007 verwendet (für nähere Informationen zum SOEP: Wagner et al. 2007).

Zu wählen ist noch ein geeignetes Analyseverfahren. Der zunächst nahe liegende Abgleich von Regressionsschätzungen auf Basis von plausiblen versus unplausiblen Fällen wird durch die unterschiedliche Varianz der unabhängigen Variablen in diesen beiden Gruppen beeinträchtigt: Bei den unplausiblen Fällen liegt *per se* eine andere Korrelation und Varianz der sie definierenden Dimensionen vor. Zudem würden die deutlichen Diskrepanzen in den Fallzahlen den Vergleich erschweren (die unplausiblen Fälle machen jeweils nur eine Minderheit aus). Theoretisch zu erwarten ist aber ohnehin ein Effekt, der sich nicht nur auf die unplausiblen Vignetten selbst bezieht, sondern ebenso auf die nachfolgenden: Wir gehen schließlich davon aus, dass die Konfrontation mit unrealistischen Fällen den *grundsätzlichen* Glauben an die Ernsthaftigkeit der Befragung und damit die *generelle* Kooperationsbereitschaft schmälert – d. h. genau genommen rechnen wir *ab* dem Auftreten unplausibler Vignetten mit einem weniger konsistenten (H_{3b}) oder stärker vereinfachten (d. h. weniger bzw. allein die unplausiblen Dimensionen einbeziehenden) Urteilsverhalten (H_{3c}).

Dies legt es nahe, die Antwortmuster *bis* und *ab* dem Auftreten eines ersten unplausiblen Falles miteinander zu vergleichen. Wann das erste Mal eine unplausible Vignette erscheint, und wie viele unplausible Vignetten es pro Befragten sind, variiert aufgrund der zufälligen Deckzusammenstellung zwischen den Befragten. Insbesondere die zufällige Reihenfolgeposition pro Befragten erlaubt es, die Wirkung der unplausiblen Vignetten von Lerneffekten zu trennen.[32] Tabelle 4.4 beinhaltet Regressionsschätzungen einerseits für die Definition der Unplausibilität über die Ausbildung (Modelle 1 und 2), andererseits für ihre Operationalisierung über das Einkommen (Modelle 3 und 4). Es werden jeweils Schätzungen für die Urteile *vor* (Modelle 1 und 3) sowie *ab* dem Auftreten eines ersten unplausiblen Falles (Modelle 2 und 4) gegenübergestellt. Ein Chow-Test weist die Modellunterschiede

[32] Hierzu wird in den Modellen zudem für die Bearbeitungsposition kontrolliert. Bei beiden Operationalisierungen treten unplausible Vignetten zu etwa 60 Prozent auf einer der ersten drei Rangpositionen das erste Mal auf, und es haben jeweils ungefähr ein Drittel der Befragten eine bis maximal zwei unplausible Vignetten erhalten. Maximal sind es fünf unplausible Vignetten pro einzelnem Befragten. Das Bestehen eines Lerneffekts wird überdies durch die entsprechenden Analysen im vorherigen Abschnitt entkräftet, schließlich haben sich die dort geprüften Unterschiede der Koeffizienten nach Bearbeitungsposition der Vignetten nicht als signifikant erwiesen. Um Fehlschlüsse zu vermeiden, sind dennoch ein paar weitere Überlegungen zur Vergleichbarkeit der beiden Gruppen erforderlich. Das Problem der unterschiedlichen Varianz der Dimensionen ist noch nicht völlig ausgeräumt – die Antworten ab dem Auftreten unplausibler Fälle beziehen sich unweigerlich auf stärker variierende und geringer korrelierte Vignettendimensionen, schließlich sind bei diesen Vignetten (im Gegensatz zur Vergleichsgruppe ohne unplausible Fälle) *alle* Merkmalskombinationen zulässig und tritt somit beispielsweise das berufsspezifische Einkommen in einer höheren Bandbreite auf. Diese stärkere Varianz und Unkorreliertheit bewirken aber *per se* eine höhere Schätzpräzision, damit *Power* von Signifikanztests. Finden sich stärkere Einflussstärken unplausibler Dimensionen, stellen diese daher noch nicht unbedingt ein Beleg für die von Faia (1980) vermutete Fokussierung der Befragten auf diese Merkmale dar, sie können ebenso allein mathematisch-statistisch bedingt sein. Und selbst wenn sich eine höhere kognitive Aufmerksamkeit der Befragten feststellen ließe, wäre diese dann noch nicht notwendigerweise der „Irrealität" der Fälle geschuldet, sie könnte ebenso Anzeichen eines *number-of-levels*- oder *range*-Effektes sein. Hierunter werden höhere kognitionspsychologische Aufmerksamkeiten der Befragten für stärker (bzw. in größeren Spannweiten) variierende Merkmale gefasst, unabhängig von deren Inhalten. Diese Effekte, die zumindest für die verwandten Conjoint-Verfahren bereits gut belegt sind (z. B. Louviere 2001a; Ohler et al. 2000; Perrey 1996; Wittink et al. 1982, 1989), bieten also alternative Interpretationen für die Wirkung unplausibler Fälle. Auf diese Vermischung weisen ähnlich bereits Creyer und Ross (1988) hin, mitunter durch empirische Befunde gestützt (s. dazu auch Klein 2002: 15).

bei beiden Operationalisierungen als signifikant aus (F = 2,06 bei df = 7 und 459; p = 0,046 bzw. F = 71,37 bei df = 7 und 459; p = 0,000). Bei der Operationalisierung über die Ausbildung geht dies neben einem Niveaueffekt (die Konstante ist in Modell 2 geringer als in Modell 1) auf die Ausbildungsdimension zurück. Der Einfluss des Berufsabschlusses unterscheidet sich zu einem 10-, der des Hochschulabschlusses zu einem 5-Prozent-Niveau signifikant (Prüfung durch Schätzung eines gemeinsamen Modells mit entsprechenden Interaktionstermen).[33] Dagegen erweist sich bei der Operationalisierung über das Einkommen (Modelle 3 und 4) allein der Einfluss dieser Dimension als signifikant verschieden.

Tabelle 4.4: OLS-Regressionen der Vignettenurteile[a] in Abhängigkeit vom Auftreten unplausibler Fälle (robuste Standardfehler in Klammern; signifikante Unterschiede in den Koeffizienten zwischen Modell 1 und 2 bzw. 3 und 4 hervorgehoben)[b]

	Definition über die Ausbildung		**Definition über das Einkommen**	
	Modell 1 Vor unplaus. Vignetten	**Modell 2 Ab unplaus. Vignetten**	**Modell 3 Vor unplaus. Vignetten**	**Modell 4 Ab unplaus. Vignetten**
Weibliche Vignettenperson	-0,244*	-0,023	-0,178	-0,072
	(0,130)	(0,104)	(0,117)	(0,100)
Alter [Jahre]	-0,030***	-0,025***	-0,030***	-0,022***
	(0,006)	(0,004)	(0,005)	(0,004)
Abschluss (Ref.: kein Abschluss)				
- Berufsabschluss	-0,972***	-0,571***	-0,671***	-0,670***
	(0,202)	(0,105)	(0,145)	(0,107)
- Hochschulabschluss	-1,386***	-0,784***	-1,101***	-0,962***
	(0,182)	(0,117)	(0,134)	(0,109)
Berufsprestige [10 MPS-Score]	-0,135***	-0,114***	-0,126***	-0,149***
	(0,015)	(0,010)	(0,013)	(0,010)
Nettoeinkommen [100,- Euro]	0,058***	0,057***	0,161***	0,055***
	(0,002)	(0,002)	(0,005)	(0,001)
Position der Vignette[c]	0,004	0,063***	-0,029	0,071***
	(0,032)	(0,021)	(0,029)	(0,021)
Konstante	7,152***	5,813***	5,021***	6,124***
	(0,353)	(0,241)	(0,277)	(0,252)
Beobachtungen	1.301	2.179	1.197	2.283
(Befragte)	(355)	(400)	(344)	(409)
R^2	0,44	0,48	0,56	0,52

[a] Skala von 1‘ungerechterweise zu niedrig‘ bis 11‘ungerechterweise zu hoch‘. Der Wert ‚Sechs‘ kennzeichnet eine als gerecht empfundene Entlohnung.
[b] Prüfung mittels Interaktionstermen zwischen den Vignettendimensionen und der Dimensionszahl in einem *gepoolten* Modell, Signifikanzniveau von 5 Prozent.
[c] Hier als lineare Variable ausgewiesen, da ihr ausschließlich eine Kontrollfunktion zukommt. Bei einer alternativen Modellierung mit Dummy-Variablen für die einzelnen Positionen bleiben die Ergebnisse stabil.
*** p<0,01; ** p<0,05; * p<0,1 bei zweiseitigem Test; Schätzungen mit robusten Standardfehlern.

Eine Analogie besteht auch in der Richtung der Unterschiede: Die Effekte fallen jeweils für die Modelle ab dem Auftreten eines unplausiblen Falles (also in Modell 2 statt 1 und 4 statt 3) schwächer aus. Dies entspricht nicht unserer Annahme, lässt sich aber gleichwohl als ein schlüssiger Befund deuten:

[33] Es wurden Interaktionsterme zwischen allen Vignettendimensionen und einer Dummy-Variable gebildet, welche die Vignette einordnet (vor versus ab unplausiblem Fall).

Die Befragten beziehen ab der Konfrontation mit einer unplausiblen Kombination die ursächlichen Dimensionen weniger in ihr Urteil ein – man könnte auch sagen, sie nehmen diese weniger „ernst".[34] Sollte sich dieser überraschende Befund in künftigen Untersuchungen, die aufgrund des *ad hoc*-Charakters unserer Interpretation angezeigt sind, replizieren, kann die Empfehlung nur lauten, sparsam mit solchen Fällen umzugehen; oder zumindest wie hier mit Zufallsreihenfolgen der Vignetten zu arbeiten, um Konfundierungen der Auswirkungen unplausibler Fälle (die eben erst ab ihrem Erscheinen auf den hinteren Bearbeitungspositionen auftreten können) mit inhaltlichen Effekten zu vermeiden.

Bevor wir zu einem Fazit kommen, wollen wir die Einflüsse auf die Antwortkonsistenz und die Bearbeitungszeit noch in zwei abschließenden, multivariaten Modellen zusammenfassen. Als Maß für die Konsistenz verwenden wir die unerklärt gebliebene Varianz, genauer gesagt die quadrierten Residuen und schätzen OLS-Regression mit den Designvariablen als unabhängigen Variablen. Ein negativer Effekt bedeutet dann eine geringere Fehlervarianz bzw. höhere Konsistenz des Antwortverhaltens. Wie Tabelle 4.5 zeigt, finden sich lediglich zwei derartige (und nur zum 10-Prozent-Niveau) signifikante Effekte (Modell 1): Ab dem Auftreten von unplausiblen Fällen verringert sich die Fehlervarianz bzw. erhöht sich die Antwortkonsistenz (was wie oben gezeigt einem weniger Dimensionen einbeziehenden, damit vereinfachten Antwortverhalten geschuldet ist), und ebenso ist die Antwortkonsistenz umso höher, je mehr Zeit sich die Befragten für das Beantworten der einzelnen Vignetten nehmen.[35] Die in Modell 2 betrachtete Bearbeitungszeit sinkt dagegen mit der Position der Vignette, d. h. den Befragten gelingen zunehmend zeiteffiziente Urteile. Wie an dem positiven Effekt der quadrierten Bearbeitungszeit ersichtlich ist, handelt es sich um einen Effekt mit abnehmender Rate (grafisch einen u-förmigen Effekt).[36] Zudem bestätigt sich nochmals die zeitraubendere Bearbeitung von Vignetten mit einer höheren Dimensionszahl. Unplausible Fälle führen hingegen zumindest bei der hier verwendeten Operationalisierung über die Ausbildung nicht zu einem *zeitlich* flüchtigeren Antwortverhalten.[37]

[34] Der Annahme von Faia (1980) ist dies ebenso diametral entgegengesetzt wie einem *number-of-levels*- oder *range*-Effekt.

[35] Um von grundsätzlichen Unterschieden in der individuellen „Basisgeschwindigkeit" (der vom Frageinhalt unabhängig Grundgeschwindigkeit der Befragten) zu abstrahieren, wurden die Bearbeitungszeiten der Vignetten bei den hier vorliegenden Analysen jeweils pro Befragten mit seiner Antwortzeit bei „herkömmlichen" Itembatterien gewichtet (so genannte „Latenzzeiten"). Die Geschwindigkeiten wurden dabei vorab um Ausreißer (oberes 5-Prozent-Perzentil) bereinigt (zur Empfehlung eines ähnlichen Vorgehens s. Mayerl et al. 2005; Urban/Mayerl 2007).

[36] Der Wendepunkt wäre nach der vorliegenden Modellschätzung bei der zehnten Vignette erreicht. Da für darüber hinausgehende Vignetten keine Beobachtungen vorliegen, wäre dieser Befund aber nochmals mit umfangreicheren Vignettendecks pro Befragten zu validieren.

[37] Bei einer Operationalisierung über das Einkommen findet sich allerdings eine zum 5-Prozent-Niveau signifikante Verringerung der Beantwortungszeit (um durchschnittlich 0,016 Sekunden). Diese differenten Befunde je nach Operationalisierung unplausibler Fälle fordern zu weiteren Untersuchungen heraus.

Tabelle 4.5: OLS-Regressionen der quadrierten Residuen[a] und Bearbeitungszeiten[b] pro Vignette (robuste Standardfehler in Klammern)

	Modell 1 Quadrierte Residuen[a]	**Modell 2 Bearbeitungszeit pro Vignette[b]**
Position der Vignette	-0,106	-0,064***
	(0,197)	(0,004)
Position der Vignette, quadriert	0,006	0,005***
	(0,018)	(0,000)
Zwölf Dimensionen (Ref.: fünf)	-0,123	0,119***
	(0,300)	(0,008)
Ab unplausiblem Fall[c]	-0,564*	-0,004
	(0,306)	(0,008)
Bearbeitungszeit pro Vignette[a]	-1,461*	
	(0,865)	
Konstante	6,102***	0,371***
	(0,554)	(0,010)
Beobachtungen	3.095	3.095
(Befragte)	(416)	(416)
R^2	0,00	0,26

[a] Residuen einer OLS-Regression des Vignettenurteils auf die ersten fünf Vignettendimensionen.
[b] Bearbeitungszeit in Sekunden, pro Befragten mit der Bearbeitungszeit einer inhaltlich verwandten Itembatterie gewichtet; bei beiden Bearbeitungszeiten wurde vorab das obere 5-Prozent-Perzentil ausgeschlossen. Hieraus resultieren die geringeren Fallzahlen in diesen Modellen.
[3] Hier definiert über den Ausbildungsabschluss.
*** p<0.01, ** p<0.05, * p<0.1 bei zweiseitigem Test; Schätzungen mit robusten Standardfehlern.

4.6 Zusammenfassung und Schlussfolgerungen

Der Faktorielle Survey hat sich in der soziologischen Norm- und Einstellungsforschung inzwischen als Erhebungsmethode sehr gut durchgesetzt. Aber auch in anderen soziologischen Forschungszusammenhängen (wie z. B. der Diskriminierungsforschung) wird das Verfahren in den letzten Jahren vermehrt als eine ideale und innovative Methodik entdeckt. In Diskrepanz zu dieser guten Etablierung steht die geringe Erforschung des Verfahrens selbst. Es fehlen anwendungsbezogene Kriterien für die Konzeption Faktorieller Surveys, was ihre Durchführung erschwert (Beck/Opp 2001: 283f.). Ferner bestehen substanzielle Zweifel an ihrer (internen) Validität fort. Invalide Urteile könnten etwa aus einer kognitiven Überforderung oder aus der Anwendung vereinfachter Entscheidungsstrategien (Heuristiken) resultieren. Ohne gezielte Methodenstudien ist kaum zu entscheiden, inwieweit die mit Vignetten gewonnenen Ergebnisse belastbar oder als methodische Artefakte zu interpretieren sind.

Der vorliegende Beitrag zielt daher auf eine erste Untersuchung der Stabilität und Konsistenz des Antwortverhaltens in Abhängigkeit von Designmerkmalen, konkret der Komplexität, Reihenfolge und Plausibilität von Vignetten. Mittels einer experimentellen Onlinebefragung von 460 Studierenden wurden systematische Analysen zum Einfluss der Anzahl an Dimensionen, von Lerneffekten und der Wirkung unplausibler Fällen vorgenommen. Als ein erster, übergreifender Befund lässt sich festhalten,

dass diese drei methodischen Aspekte für die Antwortmuster und damit die Interpretation der inhaltlichen Ergebnisse durchaus relevant sind.

So zeigen unsere Analysen, dass die Komplexität von Vignetten, jedenfalls gemessen an ihrer Dimensionszahl, die Urteile signifikant beeinflusst. Die Effekte einzelner Merkmale schwächen sich mit der Anzahl der Dimensionen ab – was zunächst bedeutet, dass die Interpretationen der absoluten Effektstärken nicht überstrapaziert werden sollten. Der Einfluss der einzelnen Merkmale scheint mitunter eine Funktion ihrer „Einzelständigkeit" zu sein, was zumindest beim Vergleich von unterschiedlichen Studien zu berücksichtigen ist. Anders gesagt: aussagekräftige Vergleiche von Effekten in verschiedenen Erhebungen bedürfen ähnlich komplexer Vignetten. Die bei höherer Komplexität zu beobachtenden Heuristiken führen zumindest dann zu Artefakten, wenn sie keine Entsprechung mehr zu realen Urteilen aufweisen (ähnlich Swait/Adamowicz 2001:147). Nicht signifikante Einflüsse sind daher möglicherweise nochmals mit wenigerdimensionalen Fallbeispielen zu validieren. Die gute Botschaft lautet aber, dass die Auswirkungen auf das Antwortverhalten insgesamt gering sind und die Befragten selbst die hier vorgelegte, hohe Komplexität von zwölf Dimensionen insgesamt noch gut zu bewerkstelligen scheinen. Aufgrund der zu vermutenden Wechselwirkungen mit anderen Designmerkmalen (wie der Anzahl an Ausprägungen) und Eigenschaften der Befragten (kognitive Leistungsfähigkeit) sind allerdings vertiefte Untersuchungen angebracht.

Methodenstudien zu den verwandten Conjoint- und Choice-Analysen sprechen für ein komplexes Verhältnis von Lern- und Ermüdungseffekten. Um dieses vollständig abzubilden, ist die hier verwendete Fallzahl von maximal zehn Vignetten pro Befragten zu gering. Ein interessanter Befund ist aber schon einmal, dass bis zu unserer letzten, zehnten Vignette Lerneffekte dominieren, welche sich primär in einer zunehmenden Antwortgeschwindigkeit bei gleich bleibender Konsistenz (unveränderte R^2-Werte) äußern. Die mit den ersten Vignetten gewonnenen Urteile sind in unserer Stichprobe reliabel, sie werden inhaltlich also durch die nachfolgenden bestätigt. Dies spricht für die grundsätzliche Verwertbarkeit dieser „ungeübten" ersten Urteile und damit die Validität von Studien, die mit einem reinen *between-subject*-Design arbeiten (nur eine Vignette pro Befragten). Ab welcher Anzahl an Vignetten die quantitativen Zugewinne an Urteilen mit merklichen Einbußen ihrer Datenqualität bezahlt werden, ist dagegen erst mit umfangreicheren Vignettendecks zu klären. Zudem sollte geprüft werden, ob diese Befunde auch einer anderen Komplexität von Vignetten und einem heterogeneren Befragtensample Stand halten.[38]

Von den einen als Stärke des Verfahrens gelobt, sehen Kritiker gerade durch empirisch seltene, daher besonders „virtuelle" Vignetten artifizielle Urteile herbeigeführt. Eine Skepsis, die nach unseren Analysen durchaus angebracht ist. Unplausible Merkmalskombinationen scheinen zwar zu keinen drastischen Befragungsabbrüchen oder Antwortverweigerungen zu führen (die Quote an Abbrüchen

[38] Anzunehmen ist, dass die Komplexität für die Befragten auch mit dem inhaltlichen Thema variiert, genauer gesagt mit ihrer Vertrautheit mit dem zu beurteilenden Gegenstand. Dies wurde zumindest für Choice-Experimente vereinzelt bereits untersucht (in der gesundheitsökonomischen Panelstudie von Bryan et al. 2000 war allerdings *kein* Effekt der Erfahrung mit dem Befragungsthema auf die Reliabilität der Antworten festzustellen).

und *Non-Responses* ist insgesamt sehr gering), aber sie provozieren eine geringere Berücksichtigung (bis möglicherweise vollständige Ausblendung) der für die Unplausibilität ursächlichen Dimensionen. Dieser Befund ist bei inhaltlichen Ergebnisinterpretationen zu beachten: Fehlende oder geringe Signifikanzen können statt einer genuinen Irrelevanz für das Urteilsverhalten ebenso anzeigen, dass die Dimensionen in Folge ihrer Irrealität weniger ernst genommen werden. Sollte sich dieses Ergebnis in weiteren Untersuchungen bestätigen, kann die praktische Empfehlung nur lauten, auf unplausible Fälle zu verzichten, oder zumindest sparsam mit ihnen umzugehen. Dank computerbasierter Verfahren lassen sich die durch ihren Ausschluss hervorgerufenen Einbußen an Effizienz der Vignettenstichprobe (Balanciertheit und Unkorreliertheit der Dimensionen) auf ein vertretbares Maß reduzieren.

Unsere Analysen haben zudem gezeigt, dass es zur Feststellung methodischer Effekte multipler Kriterien bedarf. Vereinfachte Entscheidungsregeln tragen tendenziell zu einer höheren Messgüte der Ergebnisse bei (bewertet an der Varianzaufklärung), die Abweichung von den tatsächlichen Einstellungen und Urteilsregeln der Befragten kann gleichwohl groß sein.[39] Die zentrale Schlussfolgerung ist hier, dass die R^2-Werte ausschließlich als ein Indikator für die Konsistenz der Urteile zu gebrauchen sind, nicht aber als ein Maß dafür, inwieweit es gelungen ist, alle urteilsrelevanten Dimensionen ausfindig zu machen. Die Befragten scheinen sich bei einer drohenden Überforderung eher auf ein weiterhin konsistentes, aber gerade darum weniger detailliertes Urteilsverhalten zu konzentrieren. Faktorielle Surveys sind demnach primär ein geeignetes Verfahren für die Feststellung der Signifikanz *einzelner* Merkmale (etwa für entsprechende Hypothesentests), und weniger für die Aufdeckung inhaltlich *erschöpfender* Urteilsregeln. Anders ausgedrückt lassen sich mit ihnen Aussagen über den Einfluss der berücksichtigten Dimensionen treffen, nicht aber über die zusätzliche (Ir-)Relevanz weiterer Merkmale. Dies verweist nochmals auf die hohe Bedeutung einer sorgfältigen Zusammenstellung der Dimensionen.

Aufgrund der Vielzahl weiterer methodischer Problemlagen und der zu erwartenden Wechselwirkungen mit anderen Designmerkmalen (wie z. B. der Anzahl an Ausprägungen und deren Variation und Bandbreite) ist mit den hier vorgelegten Untersuchungen erst ein Anfang gemacht. Empfehlenswert erscheinen zunächst Replikationen mit anderen Vignettenstichproben, etwa mit einem fraktionalisierten Design mit zusätzlicher Konfundierung aller Interaktionen erster Ordnung oder mit fraktionalisierten statt randomisierten Setbildungen. Dies erscheint angebracht, weil durch die Auswahl und Zusammenstellung von Vignetten unweigerlich die im Vignettenuniversum gegebene, vollständige Orthogonalität aller Dimensionen und ihrer Interaktionen verloren geht. In diesem Zusammenhang ist nicht *vollständig* auszuschließen, dass sich die dadurch hervorgerufene Verringerung der statistischen Effizienz unterschiedlich auf die hier gegenübergestellten Gruppen mit weniger oder mehr Vignettendimensionen bzw. auf die Gruppen von Vignetten vor und ab dem Auftreten unplausibler Fälle verteilt, was ihre Vergleichbarkeit etwas beeinträchtigen könnte. Zudem sind weitere, hier aus Platzgründen nicht angesprochene methodische Aspekte von Interesse, wie etwa die Wahl möglichst

[39] Bei Berücksichtigung lediglich eines (oder weniger) Merkmale ist eine hohe Antwortkonsistenz schließlich keine kognitive Herausforderung – als Indikator für eine valide Messung ist sie gerade darum nicht hinreichend.

geeigneter Präsentationsformen (Fließtext oder tabellarische Darstellung) und der Einsatz von unterschiedlichen Antwortskalen. Darüber hinaus wären andere statistische Auswertungsverfahren zu erproben. Der gängigen Praxis folgend wurden die Vignettenurteile als metrisch behandelt, genau genommen weisen sie lediglich ordinales Skalenniveau auf. Die Wahl von OLS Schätzungen wird zwar allgemein durch ihre hohe Robustheit und bessere Interpretierbarkeit gerechtfertigt (Winship/ Mare 1984); speziell für Vignettenstudien wurde bislang aber noch zu wenig ausgelotet, welche Analysegewinne sich mit adäquateren – besser mit der Datenstruktur korrespondierenden – (Mehr-) Ebenenverfahren erzielen lassen.[40]

Mit den Daten des DFG-geförderten Projekts „Der Faktorielle Survey als Instrument zur Einstellungsmessung in Umfragen“ können eine Vielzahl der benannten Aspekte untersucht werden. Nach den hier präsentierten, ersten Befunden verdienen sie in methodischer wie inhaltlicher Hinsicht stärkere Beachtung.

[40] Aufgrund der mehrfachen Bewertungsaufgabe sind zumindest bei den hier verwendeten, geschlossenen Antwortskalen Zensierungen der Urteile zu befürchten (wurden von den Befragten bereits extreme Urteile abgegeben, können sie ihr Antwortverhalten womöglich bei späteren Vignetten nicht mehr hinreichend abstufen). Dies kann ebenfalls zu verzerrten Ergebnissen führen und legt den Einsatz von einschlägigen Regressionsverfahren (z. B. Tobit Modellen) nahe. Wiederum sind Wechselwirkungen mit dem Vorkommen von unplausiblen Fällen zu erwarten, da diese vermehrt zu extremen Antworten motivieren dürften.

Literatur

Alves, Wayne M./Rossi, Peter H. (1978): Who Should Get What? Fairness Judgments of the Distribution of Earnings. American Journal of Sociology 84, 541-564.

Auspurg, Katrin/Abraham, Martin (2007): Die Umzugsentscheidung von Paaren als Verhandlungsproblem. Eine quasiexperimentelle Überprüfung des Bargaining-Modells. Kölner Zeitschrift für Soziologie und Sozialpsychologie 59, 271-293.

Auspurg, Katrin/Abraham, Martin/Hinz, Thomas (2009): Wenig Fälle, viele Informationen: Die Methodik des Faktoriellen Surveys als Paarbefragung. In: Kriwy, Peter/Gross, Christiane (Hrsg.): Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen. Wiesbaden: VS Verlag für Sozialwissenschaften, 179-210.

Beck, Michael/Opp, Karl-Dieter (2001): Der faktorielle Survey und die Messung von Normen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 53, 283-306.

Berk, Richard A./Rossi, Peter H. (1977): Prison reform and state elites. Cambridge, Mass.: Ballinger.

Buskens, Vincent/Weesie, Jeroen (2000): An Experiment on the Effects of Embeddedness in Trust Situations: Buying a Used Car. Rationality and Society 12, 227-253.

Bradley, Mark/Daly, Andrew (1994): Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. Transportation 21, 167-184.

Bryan, Stirlin/Gold, Lisa/Sheldon, Rob/Buxton, Martin (2000): Preference Measurement Using Conjoint Methods: An Empirical Investigation of Reliability. Health Economics 9, 385-395.

Carroll, Duglas. J./Green, Paul E. (1995): Psychometric Methods in Marketing Research: Part I, Conjoint Analysis. Journal of Marketing Research 32, 358-391.

Carson, Richard T./Louviere, Jordan J./Anderson, Donald/Arabie, Phipps/Bunch, David/Hensher, David A./Johnson, Richard M./Kuhfeld, Warren F./Steinberg, Dan/Swait, Joffre/Timmerman, Harry/Wiley, James B. (1994): Experimental Analysis of Choice. Marketing Letters 5, 351-368.

Caussade, Sebastián/Ortúzar, Juan de Dios/Rizzi, Luis I./Hensher, David A. (2005): Assessing the influence of design dimensions on stated choice experiment estimates. Transportation Research Part B 39, 621-640.

Creyer, Elisabeth/Ross, William T. (1988): The Effect of Range-Frequency Manipulations on Conjoint Importance Weight Stability. Advances in Consumer Research 15, 505-509.

DeShazo, J. R./Fermo, German (2002): Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. Journal of Environmental Economics and Management 44, 123-143.

Diefenbach, Heike/Opp, Karl-Dieter (2007): When and Why do People Think there should be a Divorce? An Application of the Factorial Survey. Rationality and Society 19, 485-517.

Diekmann, Andreas (2007): Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Reinbek b. Hamburg: Rowohlt.

Dülmer, Hermann (2007): Experimental Plans in Factorial Surveys: Random or Quota Design? Sociological Methods & Research 35, 382-409.

Dülmer, Hermann (2001): Bildung und der Einfluss von Argumenten auf das moralische Urteil: Eine empirische Analyse zur moralischen Entwicklungstheorie Kohlbergs. Kölner Zeitschrift für Soziologie und Sozialpsychologie 53, 1-27.

Dülmer, Hermann/Klein, Markus (2003): Die Messung gesellschaftlicher Wertorientierungen via Conjoint- und Vignettenanalyse: Ein Ansatz zur adäquaten Operationalisierung von Ingelharts materialistischen und postmaterialistischen Wertorientierungen. Unveröffentlichter Abschlussbericht an die Fritz-Thyssen Stiftung.

Eifler, Stefanie (2007): Evaluating the Validity of Self-Reported Deviant Behavior Using Vignette Analyses. Quality & Quantity 41, 303-318.

Faia, Michael A. (1980): The Vagaries of the Vignette World: A Comment on Alves and Rossi. American Journal of Sociology 85, 951-954.

Garrett, Karen (1982): Child Abuse: Problems of Definition. In: Rossi, Peter H./Nock, Steven L. (Hrsg.): Measuring Social Judgements. The Factorial Survey Approach. Beverly Hills u. a.: Sage, 177-204.

Greene, William H. (2003): Econometric Analysis. Upper Saddle River, NJ: Prentice Hall.

Groß, Jochen/Börensen, Christina (2009): Wie valide sind Verhaltensmessungen mittels Vignetten? Ein methodischer Vergleich von faktoriellem Survey und Verhaltensbeobachtung. In: Kriwy, Peter/Gross, Christiane (Hrsg.): Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen. Wiesbaden: VS Verlag für Sozialwissenschaften, 149-178.

Hechter, Michael/Kim, Hyojoung/Baer, Justin (2005): Prediction Versus Explanation in the Measurement of Values. European Sociological Review 21, 91-108.

Hechter, Michael/Ranger-Moore, James/Jasso, Guillermina/Horne, Christine (1999): Do Values Matter? An Analysis of Advance Directives for Medical Treatment. European Sociological Review 15, 405-430.

Hembroff, Larry A. (1987): The Seriousness of Acts and Social Contexts: A Test of Black`s Theory of the Behavior of Law. American Journal of Sociology 93, 322-347.

Hermkens, Piet L. J./Boerman, Frank (1989): Consensus with respect to the fairness of incomes: Differences between social groups. Social Justice Research 3, 201-215.

Hensher, David A. (2006): Revealing Differences in Willingness to Pay due to the Dimensionality of Stated Choice Designs: An Initial Assessment. Environmental & Resource Economics 34, 7-44.

Hensher, David A. (2004): How do Respondents Handle Stated Choice Experiments? Information processing strategies under varying information load. Working Papter DoD#5, University of Sydney: Institute of Transport Studies.

Horne, Christine (2003): The Internal Enforcement of Norms. European Sociological Review 19, 335-343.

Hox, Joop J./Kreft, Ita G. G./Hermkens, Piet L. J. (1991): The Analysis of Factorial Surveys. Sociological Methods & Research 19, 493-510.

Jann, Ben (2003): Lohngerechtigkeit und Geschlechterdiskriminierung: Experimentelle Evidenz. Unveröffentlichtes Manuskript an der Eidgenössischen Technischen Hochschule Zürich.

Jasso, Guillermina (2006): Factorial-Survey Methods for Studying Beliefs and Judgments. Sociological Methods and Research 34, 334-423.

Jasso, Guillermina (1994): Assessing Individual and Group Differences in the Sense of Justice: Framework and Application to Gender Differences in the Justice of Earnings. Social Science Research 23, 368-406.

Jasso, Guillermina (1988): Whom Shall We Welcome? Elite Judgments of the Criteria for the Selection of Immigrants. American Sociological Review 53, 919-932.

Jasso, Guillermina/Opp, Karl-Dieter (1997): Probing the Character of Norms: A Factorial Survey Analysis of the Norms of Political Action. American Sociological Review 62, 947-964.

Jasso, Guillermina/Webster, Murray (1999): Assessing the Gender Gap in Just Earnings and Its Underlying Mechanisms. Social Psychology Quarterly 62, 367-380.

Jasso, Guillermina/Webster, Murray (1997): Double Standards in Just Earnings for Male and Female Workers. Social Psychology Quarterly 60, 66-78.

John St., Craig/Bates, Nancy A. (1990): Racial Composition and Neighborhood Evaluation. Social Science Research 19, 47-61.

Johnson, Reed F. (2006): Comment on "Revealing Differences in Willingness to Pay Due to the Dimensionality of Stated Choice Designs: An Initial Assessment". Environmental & Resource Economics 34, 45-50.

Klein, Markus (2002): Die Conjoint-Analyse. Eine Einführung in das Verfahren mit einem Ausblick auf mögliche sozialwissenschaftliche Anwendungen. ZA-Information 50, 7-45.

Kuhfeld, Warren F. (2005): Marketing Research Methods in SAS. Experimental Design, Choice, Conjoint and Graphical Techniques. Cary: SAS Institute.

Kuhfeld, Warren F./Randall, Tobias D./Garratt, Mark (1994): Efficient Experimental Design with Marketing Research Applications. Journal of Marketing Research 31, 545-557.

Liebig, Stefan/Mau, Steffen (2005): Wann ist ein Steuersystem gerecht? Zeitschrift für Soziologie 34, 468-491.

Liebig, Stefan/Mau, Steffen (2002): Einstellungen zur sozialen Mindestsicherung. Ein Vorschlag zur differenzierten Erfassung normativer Urteile. Kölner Zeitschrift für Soziologie und Sozialpsychologie 54: 109-134.

Liebig, Stefan/Meyermann, Alexia/Schulze, Andrea (2006): Temporal stability of justice evaluations. Paper presented at the 11th Conference of the International Society for Justice Research. Berlin: Humboldt-Universität.

Louviere, Jordan J. (2001a): What if Consumer Experiments Impact Variances as well as Means? Response Variablility as a Behavioral Phenomen. Journal of Consumer Research 28, 506-511.

Louviere, Jordan J. (2001b): Choice Experiments: an Overview of Concepts and Issues. In: Bennett, Jeff/Blamey, Russell (Hrsg.): The Choice Modelling Approach to Environmental Valuation. Cheltenham, Northhampton: Edward Elgar, 13-36.

Mayerl, Jochen/Selke, Piet/Urban, Dieter (2005): Analyzing cognitive processes in CATI-Surveys with response latencies: An empirical evaluation of the consequences of using different baseline speed measures. Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart, SISS No. 2/2005.

Melles, Thorsten (2001): Framing-Effekte in der Conjoint-Analyse. Ein Beispiel für Probleme der Merkmalsdefinition. Berichte aus der Psychologie 6. Aachen: Shaker.

Meudell, Bonner M. (1982): Household and Social Standing: Dynamic and Static Dimensions. In: Rossi, Peter H./Nock, Steven L. (Hrsg.): Measuring Social Judgements. The Factorial Survey Approach. Beverly Hills u. a.: Sage, 69-94.

Miller, J. L./Rossi, Peter H./Simpson, Jon E. (1986): Perceptions of Justice: Race and Gender Differences in Judgments of Appropriate Prison Sentences. Law & Society Review 20, 313-334.

Nisic, Natascha/Auspurg, Katrin (2009): Faktorieller Survey und klassische Bevölkerungsumfrage im Vergleich – Validität, Grenzen und Möglichkeiten beider Ansätze. In: Kriwy, Peter/Gross, Christiane (Hrsg.): Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen. Wiesbaden: VS Verlag für Sozialwissenschaften, 211-246.

Nock, Steven L. (1982): Family Social Standing: Consensus on Characteristics. In: Rossi, Peter H./Nock, Steven L. (Hrsg.): Measuring Social Judgements. The Factorial Survey Approach. Beverly Hills u. a.: Sage, 95-118.

Ohler, Tobias/Le, Aihong /Louviere, Jordan/Swait, Joffre (2000): Attribute Range Effects in Binary Response Tasks. Marketing Letters 11, 249-260.

Orme, Bryan K. (2006): Getting started with Conjoint Analysis: Strategies for Product Design and Pricing Research. Madison, Wisconsin: Research Publishers LLC.

O`Toole, Richard/Webster, Stephen W./O`Toole, Anita W./Lucal, Betsy (1999): Teachers´ Recognition and Reporting of Child Abuse: A Factorial Survey. Child Abuse & Neglect 23, 1083-1101.

Perrey, Jesko (1996): Erhebungsdesign-Effekte bei der Conjoint-Analyse. Marketing – Zeitschrift für Forschung und Praxis 18, 105-116.

Rooks, Gerrit/Raub, Werner/Selten, Robert/Tazelaar, Frits (2000): How Inter-firm Co-operation Depends on Social Embeddedness: A Vignette Study. Acta Sociologica 43, 123-137.

Rossi, Peter H. (1979): Vignette Analysis: Uncovering the Normative Structure of Complex Judgments. In: Merton, Robert K./Coleman, James S./Rossi, Peter H.: Qualitative and Quantitative Social Research: Papers in Honour of Paul F. Lazarsfeld. New York: Free Press, 176-186.

Rossi, Peter H./Alves, Wayne M. (1980): Rejoinder to Faia. American Journal of Sociology 85, 54-955.

Rossi, Peter H./Anderson, Andy B. (1982): The Factorial Survey Approach: An Introduction. In: Rossi, Peter H./Nock, Steven L. (Hrsg.): Measuring Social Judgments. The Factorial Survey Approach. Beverly Hills: Sage, 15-67.

Rossi, Peter H./Nock, Steven L. (1982): Measuring Social Judgements. The Factorial Survey Approach. Beverly Hills u. a.: Sage.

Rossi, Peter H./Sampson, William/Bose, Christine E./Jasso, Guillermina/Passel, Jeff (1974): Measuring Household Social Standing. Social Science Research 3, 169-190.

Sauer, Carsten (2009): Methodische Probleme von Conjoint- und Vignettenanalysen – Literaturreview. Arbeitsbericht Nr. 1 des Projekts „Der faktorielle Survey als Instrument zur Einstellungsmessung in Umfragen“. Bielefeld/Konstanz: Universität Bielefeld/Universität Konstanz.

Seyde, Christian (2005): Beiträge und Sanktionen in Kollektivgutsituationen: Ein faktorieller Survey. Arbeitsbericht Nr. 42 des Instituts für Soziologie. Leipzig: Universität Leipzig.

Shepelak, Norma J./Alwin, Duane F. (1986): Beliefs about Inequality and Perceptions of Distributive Justice. American Sociological Review 51, 30-46.

Shlay, Anne B./Tran, Henry/Weinraub, Marsha/Harmon, Michelle (2005): Teasing apart the child care conundrum: A factorial survey analysis of perceptions of child care quality, fair market price and willingness to pay by low-income, African American parents. Early Childhood Research Quarterly 20, 393-416.

Smith, Tom W. (1986): A Study of Non-Response and Negative Values on the Factorial Vignettes on Welfare. GSS Methodological Report No. 44. Chicago: NORC.

Sniderman, Paul M./Grob, Douglas B. (1996): Innovations in Experimental Design in Attitude Surveys. Annual Review of Sociology 22, 377-399.

Steiner, Peter M./Atzmüller, Christiane (2006): Experimentelle Vignettendesigns in Faktoriellen Surveys. Kölner Zeitschrift für Soziologie und Sozialpsychologie 58, 117-146.

Struck, Olaf/Pfeifer, Christian/Krause, Alexandra (2008): Entlassungen: Gerechtigkeitsempfinden und Folgewirkungen. Theoretische Konzepte und empirische Ergebnisse. Kölner Zeitschrift für Soziologie und Sozialpsychologie 60, 102-122.

Swait, Joffre/Adamowicz, Wiktor (2001): The Influence of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching. Journal of Consumer Research 28, 135-148.

Urban, Dieter/Mayerl, Jochen (2007): Antwortlatenzzeiten in der survey-basierten Verhaltensforschung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 59, 692-713.

Wagner, Gert G./Frick, Joachim R./Schupp, Jürgen (2007): The German Socio-Economic Panel Study (SOEP) – Evolution, Scope and Enhancements. Schmollers Jahrbuch (Journal of Applied Social Science Studies) 127, 139-169.

Wason, Kelly D./Polonsky, Michael J./Hyman, Michael R. (2002): Designing Vignette Studies in Marketing. Australasian Marketing Journal 10, 41-58.

Will, Jeffry A. (1993): The Dimensions of Poverty: Public Perceptions of the Deserving Poor. Social Science Research 22, 312-332.

Winship, Christopher/Mare, Robert D. (1984): Regression Models with Ordinal Variables.

Wittink, Dick R. /Krishnamurthi, Lakshman/Nutter, Julia B. (1982): Comparing derived importance weights across attributes. Journal of Consumer Research 8, 471-474.

Wittink, Dick R./Krishnamurthi, Lakshman/Reibstein, David J. (1989): The Effect in Differences in the Number of Attribute Levels on Conjoint Results. Marketing Letters 1, 113-123.

Wooldridge, Jeffrey M. (2003): Introductory Econometrics. A Modern Approach. Mahson, Ohio: Thomson.

Wooldridge, Jeffrey M. (2002): Econometric Analysis of Cross Section and Panel Data. Cambridge, Mass.: MIT Press.

Zimbardo, Philip G. (1988): Psychologie. Berlin u. a.: Springer.

Methodischer Anhang

Tabelle 4.A1: Übersicht über die Korrelationen der Dimensionen in den Splits mit 5 und 12 Dimensionen[a]

Split mit 5 Dimensionen					
	Geschlecht	Alter	Abschluss	Berufsprestige	Einkommen
Geschlecht	1,000				
Alter	-0,056	1,000			
Abschluss	-0,019	0,016	1,000		
Berufsprestige	0,017	0,096	-0,023	1,000	
Einkommen	-0,027	0,008	0,042	0,148	1,000
Split mit 12 Dimensionen					
	Geschlecht	Alter	Abschluss	Berufsprestige	Einkommen
Geschlecht	1,000				
Alter	0,034	1,000			
Abschluss	-0,007	-0,022	1,000		
Berufsprestige	0,032	0,049	-0,031	1,000	
Einkommen	0,030	-0,031	0,073	0,172	1,000

[a] Korrelationskoeffizient nach Pearson bzw. bei der ordinalen Dimension „Abschluss“ Rang-Korrelationskoeffizient nach Spearman.

Tabelle 4.A2: Übersicht über die Korrelationen der Dimensionen in den Splits mit 7 und 10 Vignetten pro Befragten[a]

Split mit 7 Vignetten					
	Geschlecht	Alter	Abschluss	Berufsprestige	Einkommen
Geschlecht	1,000				
Alter	-0,033	1,000			
Abschluss	0,003	0,002	1,000		
Berufsprestige	0,039	0,070	-0,031	1,000	
Einkommen	-0,015	-0,016	0,055	0,182	1,000
Split mit 10 Vignetten					
	Geschlecht	Alter	Abschluss	Berufsprestige	Einkommen
Geschlecht	1,000				
Alter	0,032	1,000			
Abschluss	-0,050	-0,011	1,000		
Berufsprestige	-0,006	0,080	-0,016	1,000	
Einkommen	0,038	-0,004	0,059	0,115	1,000

[a] Korrelationskoeffizient nach Pearson bzw. bei der ordinalen Dimension „Abschluss“ Rang-Korrelationskoeffizient nach Spearman.

Tabelle 4.A3: Übersicht über die Korrelationen[a] aller zwölf Vignettendimensionen (nur Split mit 12 Dimensionen)

	Geschlecht	Alter	Abschluss	Berufsprest.	Einkommen	Berufserf.	Betriebszug.	Leistung	Betriebsgr.	Wirtsch. Lage	Gesundheit	Kinder
Geschlecht	1,000											
Alter	0,034	1,000										
Abschluss	-0,007	-0,022	1,000									
Berufsprestige	0,032	0,049	-0,031	1,000								
Einkommen	0,030	-0,031	0,073	0,172	1,000							
Berufserf.	0,018	0,187	-0,107	0,109	0,045	1,000						
Betriebszug.	0,024	0,279	-0,101	0,093	0,117	0,416	1,000					
Leistung	-0,015	0,051	-0,025	-0,015	0,001	0,027	0,013	1,000				
Betriebsgröße	-0,036	0,072	0,156	-0,092	-0,153	-0,012	0,212	-0,083	1,000			
Wirtsch. Lage	0,049	0,059	-0,031	0,091	-0,032	0,042	0,059	0,012	0,075	1,000		
Gesundheit	0,027	0,015	0,013	0,086	-0,165	-0,013	-0,041	-0,072	-0,014	-0,023	1,000	
Kinder	0,018	0,101	0,023	-0,051	0,001	0,088	-0,064	-0,005	0,063	0,023	-0,087	1,000

[a] Bei metrischen und binären Variablen Korrelationskoeffizient nach Pearson; bei ordinalen Variablen (Abschluss, Leistung und wirtschaftliche Lage) Rang-Korrelationskoeffizient nach Spearman.

5

Zusammenfassung

(Katrin Auspurg)

1 Zusammenfassung

Die Kernthese der vorliegenden Arbeit ist, dass die Ungleichheitssoziologie von einer stärkeren Beachtung der Prinzipien *erklärender* bzw. *analytischer* Soziologie profitieren würde. Anstatt die Suche nach allumfassenden Gesellschaftstheorien oder möglichst realitätsnahen Beschreibungen fortzusetzen, sollte der Erkenntnisfortschritt vor allem in der Weiterentwicklung und empirischen Prüfung von *erklärenden Mechanismen* liegen. Diese These wird in Kapitel 1 theoretisch begründet und anschließend in den Kapiteln 2 bis 4 durch empirische Anwendungsbeispiele illustriert. Im Folgenden werden die zentralen Inhalte und Ergebnisse zusammengefasst, um dann die Arbeit mit einem knappen Resümee zu beschließen.

Kapitel 1 definiert zunächst grundlegende Begriffe. Soziale Ungleichheiten sind Unterschiede im Zugang zu begehrten Ressourcen und Lebenschancen, die nicht rein natürlich bedingt sind. In der hier präferierten engen Fassung handelt es sich um Verletzungen des Prinzips der Chancengleichheit. Zentrale Aufgabe der Soziologie ist die Aufklärung der ursächlichen Prozesse. Gerade diesbezüglich gilt der Forschungsstand als unbefriedigend. Um die These zu begründen, dass eine stärkere Befolgung von Prinzipien analytischer bzw. erklärender Soziologie zu einem substanziellen Erkenntnisfortschritt verhelfen würde, wird zunächst das Forschungsprogramm der analytischen Soziologie referiert. Das übergeordnete Ziel der von Peter Hedström und vielen Rational-Choice-Theoretikern vertretenen analytischen Soziologie besteht darin, möglichst *präzise, handlungsbasierte, abstrakte und empirisch prüfbare Erklärungen* für soziale Phänomene (hier soziale Ungleichheiten) zu entwickeln. Im Einzelnen beinhaltet dieses Programm eine klare Präzisierung der unterstellten kausalen Mechanismen, den Einbezug von Handlungs- und Spieltheorien sowie die Abstraktion von Details, die in ihrer kausalen Wirkung nebensächlich sind. Die vorgeschlagenen Erklärungen sollten wenigstens so allgemeingültig sein, dass sie sich nicht nur auf ein einziges (historisches) Phänomen anwenden lassen. Gute Theorien erklären mit sparsamen Annahmen eine möglichst breite Palette an Phänomenen. Nicht verfolgt wird der Anspruch einer allumfassenden Gesellschaftstheorie. Vielmehr geht es um die Spezifikation von mehreren Mechanismen mittlerer Reichweite, die sich im idealen Forschungsverlauf zunehmend miteinander verzahnen lassen.

Explizite Prämisse analytischer Soziologie ist zudem die möglichst enge Verknüpfung von Theorie und Empirie. Ideale Untersuchungsdesigns nehmen eine exakte Spezifikation von Hypothesen vor und gestatten den Ausschluss möglichst vieler alternativer Mechanismen. Hinzu kommt die Anforderung, *generelle* Mechanismen zu modellieren anstatt die Realität möglichst genau abzubilden. Aussichtsreich hierfür sind experimentelle oder Längsschnittdaten sowie Kombinationen aus Experimenten und Umfragen (wie etwa Faktorielle Surveys). Eine bislang in ihrer Relevanz unterschätzte Aufgabe besteht darin, durch gezielte Methodenforschung das Repertoire an geeigneten empirischen Verfahren weiterzuentwickeln. Für valide Schlussfolgerungen darf der Forscher bei der Konzeption

und Auswertung von Datenerhebungen, anders als bei der Theoriebildung, nicht von kognitiven Beschränkungen (*bounded rationality*) oder Informationsdefiziten auf Seiten der Probanden absehen.

Der Überblick über den Forschungsstand ergibt zusammenfassend, dass diese Prinzipien in der Soziologie sozialer Ungleichheit und Sozialstrukturanalyse wenig Beachtung erfahren. Einerseits liegen mit den frühen Klassentheorien, funktionalistischen Sichtungstheorien oder Theorien der sozialen Differenzierung groß angelegte Gesellschaftstheorien vor. Diese lösen die Prämissen analytischer Soziologie nicht ein, empirisch bewährte und prüfbare Hypothesen zu generieren und hierfür Handlungstheorien zu verwenden. Das andere Extrem bilden detailgetreue Beschreibungen von Schichten, Lebensstilen, Milieus oder transnationalen Lagen, welche wiederum den Kritierien einer theoretischen Motivation empirischer Untersuchungen sowie eines hinreichenden Abstraktionsniveaus zu wenig gerecht werden, um über kontingente Phänomene hinausreichende „Erklärungen" zu bieten. Novum der vorliegenden Arbeit ist es angesichts dessen, explizit eine stärker *analytisch* ausgerichtete Ungleichheitssoziologie anzuregen. Dieser Therapievorschlag zur Überwindung des allgemein beklagten Erkenntnismangels ist – wie im Zwischenfazit in Abschnitt 1.2.4 argumentiert – nicht mit einer radikalen Verabschiedung von Sozialstrukturkonzepten gleichzusetzen. Typologien können hilfreiche Hinweise dafür bieten, welche Faktoren soziale Ungleichheiten herbeiführen. Das unbedingte Festhalten an (Klassen-)Konzepten ist aber in fortgeschrittenen Forschungsstadien hinderlich. Denn Ungleichheit generierende Mechanismen (wie Status- und Reputationseffekte, Soziale Schließungen, Heuristiken oder Diskriminierungen) müssen nicht mit den über Erwerbsklassen oder Gruppenidentitäten definierten (Klassen-)Konzepten zusammenfallen. Empirisch ist die Bedeutung von makrostrukturellen Kategorien daher vor allem ein Hinweis darauf, dass die eigentlich zu entdeckenden Mechanismen noch nicht gefunden sind.

Gleichwohl lautet das Resümee dieses ersten, konzeptuellen Kapitels nicht, dass es einer neuerlichen Theorie oder Beschreibung sozialer Ungleichheit bedarf. Der in der vorliegenden Arbeit begründete und praktizierte Ansatz besteht stattdessen darin, die Ursachen sozialer Ungleichheit durch einen verstärkten Einsatz analytischer Soziologie zu erforschen. Propagiert wird ein Perspektivenwechsel, weg von der Beschreibung konkreter Verteilungsmuster, hin zu einer stärkeren *Aufklärung* der generellen *ursächlichen Mechanismen*. Dieses Forschungsprogramm wird im zweiten Teil der Dissertation durch drei empirische Beispiele illustriert.

Die in **Kapitel 2** präsentierte und gemeinsam mit Thomas Hinz und Jürgen Güdler verfasste Studie „Herausbildung einer akademischen Elite?" untersucht Verteilungsprinzipien von Forschungsdrittmitteln der Deutschen Forschungsgemeinschaft (DFG). Das Wissenschaftssystem ist der Bereich, der sich besonders explizit meritokratischen Normen verschrieben hat. Entsprechend sind Verteilungsprozesse stärker standardisiert und damit direkter zu beobachten als in anderen Bereichen. Gleichwohl bedarf es zur Vermeidung von Fehlschlüssen eines präzisen theoretischen wie empirischen Instrumentariums. Die zentrale Fragestellung ist, inwieweit die Verteilung von Forschungsmitteln leistungsgerecht erfolgt oder Größen- und Reputationseffekte eine Rolle spielen, die letztlich zu einer immer

stärkeren Anhäufung von Forschungsressourcen an wenigen, zunehmend hoch reputierten Universitäten führen würden. Den Ausgangspunkt der Untersuchung bildet der in der analytischen Soziologie geradezu als Idealbeispiel für einen Mechanismus geltende Matthäus-Effekt (sich selbst verstärkende Wirkung von Forschungsressourcen gemäß dem Prinzip: „wer hat, dem wird gegeben"). Aus verschiedenen Teilmechanismen (Signal- und Humankapitaleffekte, statistische Diskriminierung) werden Hypothesen zum Einfluss von Größe und Reputation der Universitäten auf die Bewilligungschancen von Forschungsanträgen der bei ihnen beschäftigten Wissenschaftler abgeleitet. Diese werden anschließend mit prozessproduzierten Daten zu den im Zeitraum 1991 bis 2004 bei der Deutschen Forschungsgemeinschaft eingereichten Einzelanträgen überprüft. Durchgeführt werden multivariate Analysen der individuellen Bewilligungschancen sowie der Drittmitteleinwerbungen universitärer Fachgebiete. Im Ergebnis sind Kontext- und Konzentrationseffekte kaum nachzuweisen. Lediglich für Wissenschaftler an westdeutschen Traditionsuniversitäten finden sich, im Vergleich zu Wissenschaftlern an anderen Universitäten, leicht erhöhte Bewilligungschancen. Ebenso zeigen sich auf der Makroebene einzelner Universitäten keine Evidenzen für eine zunehmende Ungleichheitsverteilung. Damit zeigt das Kapitel sehr plastisch das Erkenntispotenzial einer analytischen Soziologie auf – die Ergebnisse sind fast durchweg konträr zu den Befunden einer mit sehr ähnlichem Datenmaterial arbeitenden, aber allenfalls eklektisch auf erklärende Mechanismen zurückgreifenden Studie von Richard Münch (2007). Insbesondere werden die in der analytischen Soziologie prophezeiten Fehlschlüsse offensichtlich, die mit Untersuchungen einhergehen, welche (wie bei Münch) vorwiegend auf der Makroebene von Institutionen (statt Mikroebene von Individuen) verharren. Unabhängig davon sind die geringen Evidenzen für einen Matthäus-Effekt auf der Ebene von Forschungsinstitutionen ein international beachtlicher Befund, besteht diesbezüglich doch eine Forschungslücke. Grenzen der Aussagekraft der vorgelegten Analysen werden in der Schlussbetrachtung des Kapitels diskutiert, ebenso erfolgt eine Einordnung in die Debatte um die Exzellenzinitiative.

Kapitel 3 fokussiert auf soziale Ungleichheiten, die durch die wechselseitige Verschränkung von beruflichen Karrieren in privaten Partnerschaften erzeugt werden. Mit der Spieltheorie liegen für die Analyse derartiger Interdependenzen von Akteuren in Entscheidungssituationen inzwischen sehr elaborierte *theoretische* Modelle vor. Deren Analysetiefe lässt sich mit vorliegenden Daten aber nur sehr bedingt einholen. Damit ist der Gegenstand beruflicher Karriereentscheidungen in Partnerschaften ideal, um nicht nur eine Forschungslücke der Ungleichheitssoziolgie zu adressieren, sondern zugleich den propagierten Gewinn durch innovative empirische Methoden zu veranschaulichen.

Im Detail analysiert die mit den beiden Koautoren Martin Abraham und Thomas Hinz verfasste Studie „Migration Decisions Within Dual-Earner Partnerships: A Test of Bargaining Theory" das Problem der regionalen Koordination von beruflichen Karrieren in Doppelverdiener-Partnerschaften. Nur in seltenen Ausnahmefällen bietet ein Standort *beiden* Partnern optimale Erwerbsoptionen. Ein beruflicher Umzug zugunsten der einen Karriere bedeutet somit in aller Regel eine Verschlechterung

der anderen Karriere. Gut belegt ist, dass es sich bei der nachrangigen Karriere in aller Regel um die weibliche handelt, ohne dass die ursächlichen Mechanismen hinreichend entschlüsselt wären. Die Kernidee des Beitrags ist es nun, die mit Umzugsentscheidungen verbundenen Konfliktpotenziale heranzuziehen, um alternative Theorien zu Mechanismen der Karrierekoordination zu prüfen. Dies sind einerseits geschlechtsneutrale Verhandlungstheorien, andererseits sozialisations- und identitätstheoretische Ansätze (*doing gender)*. Abgeleitete Hypothesen werden mittels einer eigens durchgeführten Befragung mit etwa 280 Doppelverdiener-Paaren in Deutschland und der Schweiz getestet. Erstmalig wird ein Faktorieller Survey in einer Paarbefragung eingesetzt. Als Kombination aus Experiment und Umfrage bietet dieses Verfahren besonders aussichtsreiche Möglichkeiten, um die in der analytischen Soziologie geforderte direkte Prüfung von Theorien und empirische Trennung von alternativen kausalen Mechanismen umzusetzen. Das zentrale Ergebnis ist, dass Umzugsentscheidungen ein umso höheres Konfliktpotenzial bergen, mit je stärker asymmetrischen Verschiebungen der Erwerbsoptionen bzw. Verhandlungsmacht sie zwischen den Partnern einhergehen. Dies ist ein deutliches Indiz für die Gültigkeit der Verhandlungstheorie. Frauen weisen prinzipiell eine etwas geringere Umzugsneigung auf als Männer. Hingegen unterscheiden sich weibliche und männliche Befragte kaum in dem Ausmaß, zu dem sie ihre Umzugspräferenz an den eigenen Erwerbsoptionen versus solchen ihrer Partner ausrichten. Gerade diese sehr schwachen Anzeichen für geschlechtsspezifische Entscheidungsprinzipien sind innerhalb der internationalen Forschungslandschaft bemerkenswert, deuten sie doch darauf hin, dass bislang aufgefundene Geschlechtsunterschiede oft zu vorschnell als Ausdruck von Normen, Rollenleitbildern oder Geschlechtsidentitäten gedeutet werden. Die Methodik eines Faktoriellen Surveys bietet gegenüber diesen vorliegenden Arbeiten tiefere Einblicke in Entscheidungsmechanismen. Fraglich ist allerdings deren externe Validität. Diese und andere Aspekte der Aussagekraft der Ergebnisse werden zum Abschluss der Studie diskutiert.

Mit der in **Kapitel 4** beschriebenen Forschungsarbeit „Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey“ (Koautoren: Thomas Hinz und Stefan Liebig) ist bewusst ein methodischer Beitrag an das Ende der Dissertation gestellt. Es soll damit der hohen Relevanz von Methodenforschung für valide Ergebnisinterpretationen mehr Nachdruck verliehen werden, denn bislang wird diese in der analytischen und ebenso in der Ungleichheitssoziologie stark vernachlässigt. Das Ziel der Studie ist konkret die bessere Erforschung der theoretisch umstrittenen, aber bislang noch nicht gezielt geprüften internen Validität von Faktoriellen Surveys. Das mit Vignettenanalysen gegebene Potenzial einer möglichst exakten Prüfung von kausalen Mechanismen erscheint für das Programm einer analytischen Ungleichheitssoziologie sehr wegweisend (wie mitunter in dem Beitrag von Kapitel 3 demonstriert). Allerdings liegen nur sehr wenige Methodenstudien zur Validität der erzielten Messungen vor.

Die anhand einer eigens konzipierten experimentellen Datenerhebung (Online-Befragung mit 460 teilnehmenden Studierenden) durchgeführten Analysen beziehen sich im Einzelnen auf die Stabilität des Urteilsverhaltens der Befragten in Abhängigkeit von der Anzahl der in den Vignetten abgebildeten

Dimensionen, ferner auf mögliche Lerneffekte sowie auf die Wirkung von „unplausiblen“ oder „unlogischen“ Fällen (Vignettentexte für Situationen, die in der Realität sehr selten oder gar nicht vorkommen und die Befragten daher irritieren könnten). Inhaltlicher Gegenstand der Befragung ist das für die Ungleichheitssoziologie einschlägige Thema der Einkommensgerechtigkeit. Erstmalig erlaubt das Datenmaterial, verschiedene Hypothesen zur Komplexität der Erhebungssituation und der Kohärenz der Urteile zu testen. Die zentralen Ergebnisse sind, dass eine hohe Komplexität von Vignetten und ebenso unplausible Merkmalskombinationen zu einem weniger Vignettendimensionen einbeziehenden Urteilsverhalten führen, und somit zu geringeren Effektstärken einzelner Vignettenmerkmale bei gleich bleibender Konsistenz des Urteilsverhaltens. Die in der vorliegenden Arbeit vertretene These, dass valide Aufklärungen von kausalen Mechanismen unbedingt Methodenforschung voraussetzen, wird durch diese Befunde eindrücklich untermauert. Oberflächliche Blicke auf Gütekritierien (wie Konsistenzmaße) sind nicht ausreichend, um methodische Artefakte auszuschließen. Anschaulicher formuliert lässt sich ohne Methodenforschung nicht ergründen, inwieweit Gruppen von Befragten deshalb unterschiedlich urteilen, weil sie andere Einstellungen teilen, oder weil sie eine differente kognitive Aufnahmefähigkeit für die vom Forscher gestellten Fragen besitzen. Diese und weitere Konsequenzen werden am Ende des Beitrags diskutiert.

Alle drei Anwendungsbeispiele zeigen somit aus sehr unterschiedlichen inhaltlichen sowie methodischen Perspektiven den Erkenntnisgewinn auf, den die Ungleichheitssoziologie durch eine stärkere Einhaltung von Prinzipien analytischer Soziologie erzielen könnte. Die Studie in Kapitel 2 demonstriert durch ihre Anlage als Replikation einer mit sehr ähnlichen Daten arbeitenden, aber stärker auf der Makroebene verharrenden Untersuchung, dass sich durch verfeinerte Analysen Fehlschlüsse vermeiden lassen; der Beitrag in Kapitel 3 illustriert, wie mit einer möglichst engen Verzahnung von Theorie und Empirie vertiefte Einsichten in kausale Mechanismen der Ungleichheitsgenerierung erreichbar sind; und schließlich veranschaulicht die Analyse in Kapitel 4 eindrücklich, dass valide Schlussfolgerungen der Methodenforschung bedürfen. Der Innovationsgehalt und auch Anspruch der vorliegenden Dissertation besteht also nicht in der Entwicklung einer neuerlichen Theorie oder anderen Beschreibungsweise sozialer Ungleichheit. *Bewusst* nicht, sollte doch Folgendes gezeigt werden: Dass die Soziologie sozialer Ungleichheit dann vorankommt, wenn wir das Potenzial vorliegender *analytischer* Ansätze ausschöpfen.

Anhang

Literaturverzeichnis

Literaturverzeichnis

Abraham, Martin/Auspurg, Katrin/Hinz, Thomas (i. E.): Migration Decisions Within Dual-Earner Partnerships: A Test of Bargaining Theory. Akzeptiert für: Journal of Marriage and Family.

Ajzen, Icek (1991): The Theory of Planned Behavior. Organizational Behavior and Human Decision Processes 50, 179-211.

Ajzen, Icek (1988): Attitudes, Personality and Behavior. Chicago: Dorsey Press.

Alexander, Cheryl S./Becker, Henry Jay (1978): The Use of Vignettes in Survey Research. Public Opinion Quarterly 42, 93-104.

Allison, Paul D./Long, J. Scott (1990): Departmental Effects of Scientific Productivity. American Sociological Review 55, 469-478.

Allison Paul D./Long, J. Scott/Krauze, Tad K. (1982): Cumulative Advantage and Inequality in Science. American Sociological Review 47, 615-625.

Allmendinger, Jutta/Hinz, Thomas (2002): Programmierte (Un-)Gleichheit? Geschlechtsspezifische Chancen bei der Bewilligung von Forschungsanträgen. Zeitschrift für Soziologie 31, 275-293.

Alves, Wayne M./Rossi, Peter H. (1978): Who Should Get What? Fairness Judgments of the Distribution of Earnings. American Journal of Sociology 84, 541-564.

Antel, John J. (1980): Returns to Migration: Literature Review and Critique. Santa Monica: Rand Corporation.

Arrow, Kenneth J. (1973): The Theory of Discrimination. In: Ashenfelter, Orley/Rees, Albert (Hrsg.): Discrimination in Labor Markets. Princeton: Princeton University Press, 193-216.

Auspurg, Katrin/Abraham, Martin (2007): Die Umzugsentscheidung von Paaren als Verhandlungsproblem. Eine quasiexperimentelle Überprüfung des Bargaining-Modells. Kölner Zeitschrift für Soziologie und Sozialpsychologie 59, 271-293.

Auspurg, Katrin/Abraham, Martin/Hinz, Thomas (2009): Wenig Fälle, viele Informationen: Die Methodik des Faktoriellen Surveys als Paarbefragung. In: Kriwy, Peter/Gross, Christiane (Hrsg.): Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen. Wiesbaden: VS Verlag für Sozialwissenschaften, 179-210.

Auspurg, Katrin/Hinz, Thomas (i. E.): Antragsaktivitäten und Förderchancen von Wissenschaftlerinnen bei Einzelanträgen auf DFG-Einzelförderung 2005-2008. Working Paper der DFG. Bonn: Deutsche Forschungsgemeinschaft.

Auspurg, Katrin/Hinz, Thomas/Amorelli, Eva (i. E.): Räumliche und institutionelle Koordination von Doppelkarrieren in der Wissenschaft – eine Forschungsskizze. Erscheint in: Funk, Julia/ Gramespacher, Elke/Rothäusler, Iris (Hrsg.): Dual Career Couples an Hochschulen. Zwischen Wissenschaft, Praxis und Theorie. Opladen: Barbara Budrich.

Auspurg, Katrin/Hinz, Thomas/Güdler, Jürgen (2009): Gespensterdebatte oder was ist soziologische Aufklärung? Replik auf die Kommentare von Richard Münch sowie Dorothea Jansen, Richard Heidler und Regina von Goertz. Kölner Zeitschrift für Soziologie und Sozialpsychologie 61, 469-474.

Auspurg, Katrin/Hinz, Thomas/Güdler, Jürgen (2008): Herausbildung einer akademischen Elite? Zum Einfluss der Größe und Reputation von Universitäten auf Forschungsförderung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 60, 653-685.

Auspurg, Katrin/Hinz, Thomas/Liebig, Stefan (2009): Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey. Methoden, Daten, Analysen 3, 59-96.

Auspurg, Katrin/Hinz, Thomas/Liebig, Stefan/Sauer, Carsten (i. E.): Wie unplausibel darf es sein? Der Einfluss von Designmerkmalen auf das Antwortverhalten in Faktoriellen Surveys. Erscheint in: Deutsche Gesellschaft für Soziologie (Hrsg.): Kongressband des 34. Kongresses der DGS. Wiesbaden: VS Verlag für Sozialwissenschaften.

Auspurg, Katrin/Hinz, Thomas/Liebig, Stefan/Sauer, Carsten (2009): Auf das Design kommt es an. Experimentelle Befunde zu komplexen Settings in Faktoriellen Surveys. SoFid Methoden und Instrumente der Sozialwissenschaften 2009/2, 23-39.

Auspurg, Katrin/Liebe, Ulf (2010): Choice-Experimente in der Soziologie. Unveröffentlichtes Arbeitspapier. Konstanz/Leipzig: Universitäten Konstanz und Leipzig.

Bakanic, Von/Mc.Phail, Clark /Simon, Rita J. (1987): The Manuscript Review and Decision-Making Process. American Sociological Review 52, 631-642.

Baldridge, David C./Eddleston, Kimberley A./Veiga, John F. (2006): Saying 'no' to being uprooted: The impact of family and gender on willingness to relocate. Journal of Occupational and Organizational Psychology 79, 131-149.

Bartel, Ann P. (1982): Wages, Nonwage Job Characteristics, and Labor Mobility. Industrial and Labor Relations Review 35, 578-589.

Baxter, Janeen/Hewitt, Belinda/Haynes, Michele (2008): Life Course Transitions and Housework: Marriage, Parenthood, and Time on Housework. Journal of Marriage and Family 70, 259-272.

Bayer, Michael (2008): Transnationale Ungleichheitsforschung: Eine neue Herausforderung für die Soziologie. Frankfurt a. M.: Campus.

Beck, Michael/Opp, Karl-Dieter (2001): Der faktorielle Survey und die Messung von Normen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 53, 283-306.

Beck, Ulrich (2008): Die Neuvermessung der Ungleichheit unter den Menschen. Eröffnungsvortrag zum Soziologentag „Unsichere Zeiten" am 6. Oktober 2008 in Jena. Frankfurt a. M.: edition suhrkamp.

Beck, Ulrich (1986): Risikogesellschaft. Auf dem Weg in eine andere Moderne. Frankfurt a. M.: edition suhrkamp.

Beck, Ulrich (1983): Jenseits von Stand und Klasse? Soziale Ungleichheit, gesellschaftliche Individualisierungstendenzen und die Entstehung neuer sozialer Formationen und Identitäten. In: Kreckel, Reinhard (Hrsg.): Soziale Ungleichheiten. Soziale Welt, Sonderband 2. Göttingen: Otto Schwartz, 35-74.

Becker, Gary S. (1993): Human Capital. A Theoretical and Empirical Analysis with Special Reference to Education. Chicago u. a.: University of Chicago Press.

Becker, Gary S. (1971): The Economics of Discrimination. Chicago: University of Chicago Press.

Becker, Rolf/Hecken, Anna Etta (2007): Studium oder Berufsausbildung? Eine empirische Überprüfung der Modelle zur Erklärung von Bildungsungsentscheidungen von Esser sowie Breen und Goldthorpe. Zeitschrift für Soziologie 36, 100-117.

Behnke, Joachim (2005): Lassen sich Signifikanztests auf Vollerhebungen anwenden? Einige essayistische Anmerkungen. Politische Vierteljahreszeitschrift 46, 1-15.

Berger, Johannes (2004): Über den Ursprung der Ungleichheit unter den Menschen. Zeitschrift für Soziologie 33, 354-374.

Berger, Johannes (2003): Sind Märkte gerecht? Zeitschrift für Soziologie 32, 462-473.

Berger, Peter A. (2008): Transnationalisierung sozialer Ungleichheit. Wiesbaden: VS Verlag für Sozialwissenschaften.

Berger, Peter A. (2003): Kontinuitäten und Brüche. Herausforderungen für die Sozialstruktur- und Ungleichheitsforschung im 21. Jahrhundert. In: Orth, Barbara/Schwietring, Thomas/Weiß, Johannes (Hrsg.): Soziologische Forschung. Stand und Perspektiven. Opladen: Leske + Budrich, 473-490.

Berger, Peter A. (1987): Klassen und Klassifikationen. Zur „neuen Unübersichtlichkeit" in der soziologischen Ungleichheitsdiskussion. Kölner Zeitschrift für Soziologie und Sozialpsychologie 39, 59-85.

Berger, Peter A./Hradil, Stefan (1990): Die Modernisierung sozialer Ungleichheit und die neuen Konturen ihrer Erforschung. In: Berger, Peter A./Hradil, Stefan (Hrsg.): Lebenslagen, Lebensläufe, Lebensstile. Soziale Welt, Sonderband 7. Göttingen: Otto Schwartz, 3-25.

Berger, Peter A./Schmidt, Volker H. (2004): Welche Gleichheit, welche Ungleichheit? Grundlagen der Ungleichheitsforschung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Bergstrom, Theodore C. (1996): Economics in a Family Way. Journal of Economic Literature 34, 1903-1934.

Berk, Richard A./Rossi, Peter H. (1977): Prison reform and state elites. Cambridge, Mass.: Ballinger.

Bernasco, Wim/Giesen, Deidre (2000): A Bargaining Approach to Specialization in Couples. In: Weesie, Jeroen/Raub, Werner (Hrsg.): The Management of Durable Relations. Theoretical Models and Empirical Studies of Households and Organizations. Amsterdam, Thela Thesis, 42-43.

Bielby, William T./Bielby, Denise D. (1992): I Will Follow Him: Family Ties, Gender-Role Beliefs, and Reluctance to Relocate for a Better Job. American Journal of Sociology 97, 1241-1267.

Bitman, Michael/England, Paula/Folbre, Nancy/Sayer, Liana/Matheson, George (2003): When Does Gender Trump Money? Bargaining and Time in Household Work. American Journal of Sociology 109, 186-214.

Blackburn, McKinley (2006): The Impact of Internal Migration on Married Couples' Earnings in Britain, with a Comparison to the United States. ISER Working Paper 2006-24. Essex: Institute for Social and Economic Research.

Blalock, Hubert M. (1991): Understanding Social Inequality. Modeling Allocation Processes. Newbury Park u. a.: Sage.

Blank, Rebecca M. (1991): The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from The American Economic Review. The American Economic Review 81, 1041-1067.

Blamey, Russell/Bennett, Jeff (2001): Yea-saying and Validation of a Choice Model of Green Product Choice. In: Bennett, Jeff/Blamey, Russell (Hrsg.): The Choice Modeling Approach to Environmental Valuation. Cheltenham Northampton: Edward Elgar, 179-201.

Blau, Francine D./Ferber, Marianne A./Winkler, Anne E. (2001): The Economics of Women, Men, and Work. Englewood Cliffs: Prentice Hall.

Blinder, Alan S. (1973): Wage Discrimination: Reduced Form and Structural Estimates. The Journal of Human Resources 8, 436-455.

Blood, Robert O./Wolfe, Donald M. (1960): Husbands & Wives. The Dynamics of Married Living. Glencoe: Free Press.

Blossfeld, Hans-Peter (1996): Macro-sociology, Rational Choice Theory, and Time. A Theoretical Perspective on the Empirical Analysis of Social Processes. European Sociological Review 12, 181-206.

Blossfeld, Hans-Peter/Müller, Rolf (1996): Sozialstrukturanalyse, Rational Choice Theorie und die Rolle der Zeit. Ein Versuch zur dynamischen Integration zweier Theorieperspektiven. Soziale Welt 47, 382-410.

Bolte, Martin (1983): Anmerkungen zu Aspekten und Problemen der Erforschung sozialer Ungleichheit. In: Kreckel, Reinhard (Hrsg.): Soziale Ungleichheiten. Soziale Welt, Sonderband 2. Göttingen: Otto Schwartz, 391-407.

Bornmann, Lutz/Mutz, Rüdiger/Daniel, Hans-Dieter (2007): Gender differences in grant peer review: A meta-analysis. Journal of Infometrics 1, 226-238.

Bourdieu, Pierre (1981): Titel und Stelle. Über die Reproduktion sozialer Macht. Frankfurt a. M.: Europäische Verlagsanstalt.

Bradley, Mark/Daly, Andrew (1994): Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. Transportation 21, 167-184.

Braun, Norman (2009): Eine Antwort auf meine Kommentatoren. Soziale Welt 60, 223-232.

Braun, Norman (2008): Theorie in der Soziologie. Soziale Welt 59, 373-395.

Braun, Norman (1998): Der Rational-choice-Ansatz in der Soziologie. In: Pies, Ingo/Leschke, Martin (Hrsg.): Gary Beckers ökonomischer Imperialismus. Tübingen: Mohr Siebeck, 147-173.

Braun, Norman/Franzen, Axel (1995): Umweltverhalten und Rationalität. Kölner Zeitschrift für Soziologie und Sozialpsychologie 47, 231-248.

Breen, Richard/Rotman, David (1995): Class Analysis and Class Theory. Sociology 29, 453-473.

Brett, Jeanne/Reilly, Anne H. (1988): On the road again: Predicting the job transfer decision. Journal of Applied Psychology 73, 614-620.

Brines, Julie/Joyner, Kara (1999): The Ties That Bind: Principles of Cohesion in Cohabitation and Marriage. American Sociological Review 64, 333-355.

Broder, Ivy E. (1993): Review of NSF Economics Proposals: Gender and Institutional Patterns. American Economic Review 83, 964-970.

Brückner, Hannah (2009): Surveys. In: Hedström, Peter/Bearman, Peter (Hrsg.): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press, 666-687.

Brüderl, Josef (2004): Die Überprüfung von Rational-Choice-Modellen mit Umfragedaten. In: Diekmann, Andreas/Voss, Thomas (Hrsg.): Rational-Choice-Theorie in den Sozialwissenschaften. Anwendungen und Probleme. München: Oldenbourg, 163-182.

Brüderl, Josef/Preisendörfer, Peter (1995): Der Weg zum Arbeitsplatz: Eine empirische Untersuchung zur Verkehrsmittelwahl. In: Diekmann, Andreas/Franzen, Axel (Hrsg.): Kooperatives Umwelthandeln. Modelle, Erfahrungen, Massnahmen. Chur: Ruegger, 69-88.

Bryan, Stirlin/Gold, Lisa/Sheldon, Rob/Buxton, Martin (2000): Preference Measurement Using Conjoint Methods: An Empirical Investigation of Reliability. Health Economics 9, 385-395.

Burzan, Nicole (2007): Soziale Ungleichheit. Eine Einführung in die zentralen Theorien. Wiesbaden: VS Verlag für Sozialwissenschaften.

Buskens, Vincent/Weesie, Jeroen (2000): An Experiment on the Effects of Embeddedness in Trust Situations: Buying a Used Car. Rationality and Society 12, 227-253.

Cameron, Colin A. /Triverdi, Pravin K. (1986): Econometric models based on count data: Comparisons and applications of some estimators and tests. Journal of Applied Econometrics 1, 29-53.

Carayol, Nicolas/Matt, Mireille (2004): Does research organization influence academic productivity? Laboratory level evidence from a large European university. Research Policy 33, 1081-1102.

Carroll, Duglas. J./Green, Paul E. (1995): Psychometric Methods in Marketing Research: Part I, Conjoint Analysis. Journal of Marketing Research 32, 358-391.

Carson, Richard T./Louviere, Jordan J./Anderson, Donald/Arabie, Phipps/Bunch, David/Hensher, David A./Johnson, Richard M./Kuhfeld, Warren F./Steinberg, Dan/Swait, Joffre/Timmerman, Harry/Wiley, James B. (1994): Experimental Analysis of Choice. Marketing Letters 5, 351-368.

Caussade, Sebastian/Ortúzar, Juan de Dios/Rizzi, Luis I./Hensher, David A. (2005): Assessing the influence of design dimensions on stated choice experiment estimates. Transportation Research Part B 39, 621-640.

Center of Excellence Women and Science (CEWS) (2009): Hochschulranking nach Gleichstellungsaspekten – Dritte Fortschreibung (cews.publik.no13). Bonn: Kompetenzzentrum Frauen in Wissenschaft und Forschung CEWS.

Cole, Jonathan R./Cole, Stephen (1979): Which researcher will get the grant? Nature 279, 575-576.

Coleman, James (1990): Foundations of Social Theory. Cambridge, MA: Harvard University Press.

Collins, Randall (1971): Functional and Conflict Theories of Educational Stratification. American Sociological Review 36, 1002-1019.

Cooke, Lynn P. (2006): "Doing" Gender in Context: Household Bargaining and Risk of Divorce in Germany and the United States. American Journal of Sociology 112, 442-472.

Creyer, Elisabeth/Ross, William T. (1988): The Effect of Range-Frequency Manipulations on Conjoint Importance Weight Stability. Advances in Consumer Research 15, 505-509.

Dahrendorf, Ralf (1970): On the Origin of Inequality Among Men. In: Laumann, Edward/Siegel, Paul M./Hodge, Robert W. (Hrsg.): The Logic of Social Hierarchies. Chicago: Markham, 3-30.

DaVanzo, Julie/Hosek, James R. (1981): Does Migration Increase Wage Rates? An Analysis of Alternative Techniques for Measuring Wage Gains to Migration. Santa Monica: Rand Corporation.

Davis, Kingsley/Moore, Wilbert E. (1945): Some Principles of Stratification. American Sociological Review 10, 242-249.

de Vaus, David (2006): Editor's Introduction. Research Design – A Review. In: de Vaus, David (Hrsg.): Research Design. Volume I. London u. a.: Sage, IX-LV.

Deane, Glenn D. (1990): Mobility and Adjustment: Paths to the Resolution of Residential Stress. Demography 27, 65-79.

DeShazo, J. R./Fermo, German (2002): Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. Journal of Environmental Economics and Management 44, 123-143.

Deutsche Forschungsgemeinschaft (2009): Förder-Ranking 2009. Institutionen – Regionen – Netzwerke. DFG-Bewilligungen und weitere Basisdaten öffentlich geförderter Forschung. Bonn: Deutsche Forschungsgemeinschaft.

Deutsche Forschungsgemeinschaft (2006): Förder-Ranking 2006. Institutionen – Regionen – Netzwerke. DFG-Bewilligungen und weitere Basisdaten öffentlich geförderter Forschung. Bonn: Deutsche Forschungsgemeinschaft.

Deutsche Forschungsgemeinschaft (2005): Aufbau und Aufgaben. Bonn: Deutsche Forschungsgemeinschaft.

Deutsche Forschungsgemeinschaft (2003): Förder-Ranking 2003. Institutionen – Regionen – Netzwerke. DFG-Bewilligungen und weitere Basisdaten öffentlich geförderter Forschung. Bonn: Deutsche Forschungsgemeinschaft.

Deutsche Forschungsgemeinschaft (2000): DFG-Bewilligungen an Hochschulen und außeruniversitäre Forschungseinrichtungen 1996 bis 1998. Bonn: Deutsche Forschungsgemeinschaft.

Deutsche Forschungsgemeinschaft (1997): Bewilligungen an Hochschulen 1991 bis 1995. Bonn: Deutsche Forschungsgemeinschaft.

Deutsche Gesellschaft für Soziologie (2010): Transnationale Vergesellschaftungen. Themenpapier zum 35. Kongress der Deutschen Gesellschaft für Soziologie vom 11. bis 15. Oktober 2010 in Frankfurt a. M. (Verfügbar unter: http://dgs2010.de/wp-content/uploads/DGS2010_TransnationaleVergesellschaftungen.pdf; letzter Zugriff am 20.03.2010).

Diefenbach, Heike/Opp, Karl-Dieter (2007): When and Why do People Think there should be a Divorce? An Application of the Factorial Survey. Rationality and Society 19, 485-517.

Diekmann, Andreas (2009): Spieltheorie. Einführung, Beispiele, Experimente. Reinbek bei Hamburg: Rowohlt.

Diekmann, Andreas (2008): Soziologie und Ökonomie: Der Beitrag experimenteller Wirtschaftsforschung zur Sozialtheorie. Kölner Zeitschrift für Soziologie und Sozialpsychologie 60, 528-550.

Diekmann, Andreas (2007): Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Reinbek bei Hamburg: Rowohlt.

Diekmann, Andreas/Eichner, Klaus/Schmidt, Peter/Voss, Thomas (Hrsg.) (2008): Rational Choice: Theoretische Analysen und empirische Resultate. Wiesbaden: VS Verlag für Sozialwissenschaften.

Diekmann, Andreas/Voss, Thomas (2004): Die Theorie rationalen Handelns. Stand und Perspektiven. In: Diekmann, Andreas/Voss, Thomas (Hrsg.): Rational-Choice-Theorie in den Sozialwissenschaften. Anwendungen und Probleme. München: Oldenburg, 13-29.

DiPrete, Thomas A./Eirich, Gregory M. (2006): Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments. Annual Review of Sociology 32, 271-297.

Domen, Dieter (2007): Perspektiven für den Hochschulraum Ost. Zusammenfassende Thesen eines Projekts im Auftrag des Bundesministeriums für Verkehr, Bau und Stadtentwicklung. FiBS-Forum Nr. 35. Berlin: Forschungsinstitut für Bildungs- und Sozialökonomie.

Drobnič, Sonja/Blossfeld, Hans-Peter (2004): Career Patterns over the Life Course: Gender, Class, and Linked Lives. Research in Social Stratification and Mobility 21 (special edition: Inequality: Structures, Dynamics and Mechanisms: Essays in Honor of Aage B. Sørensen), 139-164.

Dülmer, Hermann (2007): Experimental Plans in Factorial Surveys: Random or Quota Design? Sociological Methods & Research 35, 382-409.

Dülmer, Hermann (2001): Bildung und der Einfluss von Argumenten auf das moralische Urteil: Eine empirische Analyse zur moralischen Entwicklungstheorie Kohlbergs. Kölner Zeitschrift für Soziologie und Sozialpsychologie 53, 1-27.

Dülmer, Hermann/Klein, Markus (2003): Die Messung gesellschaftlicher Wertorientierungen via Conjoint- und Vignettenanalyse: Ein Ansatz zur adäquaten Operationalisierung von Ingelharts materialistischen und postmaterialistischen Wertorientierungen. Unveröffentlichter Abschlussbericht an die Fritz-Thyssen Stiftung.

Eifler, Stefanie (2007): Evaluating the Validity of Self-Reported Deviant Behavior Using Vignette Analyses. Quality & Quantity 41, 303-318.

England, Paula/Farkas, Georg (1986): Households, Employment, and Gender. A Social, Economic, and Demographic View. New York: Aldine.

Esser, Hartmut (1999): Soziologie. Allgemeine Grundlagen. Frankfurt a. M./New York: Campus.

Esser, Hartmut (1996): What Is Wrong with 'Variable-Sociology'? European Sociological Review 12, 159-166.

Esser, Hartmut (1990): „Habits", „Frames" und „Rational Choice". Die Reichweite von Theorien der rationalen Wahl (am Beispiel des Befragtenverhaltens). In: Zeitschrift für Soziologie 19, 231-247.

Esser, Hartmut (1986): Können Befragte lügen? Zum Konzept des „wahren Wertes" im Rahmen der handlungstheoretischen Erklärung von Situationseinflüssen bei der Befragung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 38, 314-336.

Euorpean Commission (2009): Report on Equality between Women and Men 2009. Luxembourg: Office for Official Publications of the European Communities.

Fahrmeir, Ludwig/Künstler, Rita/Pigeot, Iris/Tutz, Gerhard (2003): Statistik. Der Weg zur Datenanalyse. Berlin u.a.: Springer.

Faia, Michael A. (1980): The Vagaries of the Vignette World: A Comment on Alves and Rossi. American Journal of Sociology 85, 951-954.

Fenstermaker, Sarah (2002): Work and Gender. In: Fenstermaker, Sarah/West, Candace (Hrsg.): Doing Gender, Doing Difference. Inequality, Power, and Institutional Change. New York: Routledge, 105-114.

Fox, Mary Frank (1983): Publication Productivity among Scientists: A Critical Review. Social Studies of Science 13, 285-305.

Ganong, Lawrence H./Coleman, Marilyn (2006): Multiple Segment Factorial Vignette Design. Journal of Marriage and Family 68, 455-468.

Garrett, Karen (1982): Child Abuse: Problems of Definition. In: Rossi, Peter H./Nock, Steven L. (Hrsg.): Measuring Social Judgements. The Factorial Survey Approach. Beverly Hills u. a.: Sage, 177-204.

Geißler, Rainer (1996): Kein Abschied von Klasse und Schicht. Ideologische Gefahren der deutschen Sozialstrukturanalyse. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48, 319-338.

Georg, Werner (2006): Soziale Ungleichheit im Bildungssystem. Eine empirisch-theoretische Bestandsaufnahme. Konstanz: UVK Verlagsgesellschaft.

Gigerenzer, Gerd/Todd, Peter M. (1999): Simple Heuristics That Make Us Smart. New York u. a.: Oxford University Press.

Gillett, Raphael (1991): Pitfalls in Assessing Research Performance by Grant Income. Scientometrics 22, 253-263.

Goldstein, Daniel G. (2009): Heuristics. In: Hedström, Peter/Bearman, Peter (Hrsg.): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press, 140-167.

Goldthorpe, John H. (2007): "Cultural capital": some critical observations. In: Scherer, Stefani/Pollak, Reinhard/Otte, Gunnar/Gangl, Markus (Hrsg.): From Origin to Destination. Trends and Mechanisms in Social Stratification Research. Frankfurt a. M./New York: Campus, 78-101.

Goldthorpe, John H. (2002): Globalisation and social class. West European Politics 25, 1-28.

Goldthorpe, John H. (2001): Causation, Statistics, and Sociology. European Sociological Review 17, 1-20.

Goldthorpe, John H. (2000): On Sociology. Volume Two: Illustration and Retrospect. Stanford: Stanford University Press.

Goldthorpe, John H. (1996): The Quantitative Analysis of Large-Scale Data-Sets and Rational Action Theory: For a Sociological Alliance. European Sociological Review 12, 109-126.

Granovetter, Mark S. (1978): Threshold Models for Collective Behavior. American Journal of Sociology 83, 1420-1443.

Greene, William H. (2003): Econometric Analysis. Upper Saddle River, NJ: Prentice Hall.

Gross, Christiane/Jungbauer-Gans, Monika (2007): Erfolg durch Leistung? Ein Forschungsüberblick zum Thema Wissenschaftskarrieren. Soziale Welt 58, 453-471.

Groß, Jochen/Börensen, Christina (2009): Wie valide sind Verhaltensmessungen mittels Vignetten? Ein methodischer Vergleich von faktoriellem Survey und Verhaltensbeobachtung. In: Kriwy, Peter/Gross, Christiane (Hrsg.): Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen. Wiesbaden: VS Verlag für Sozialwissenschaften, 149-178.

Gross, Neil (2009): A Pragmatist Theory of Social Mechanisms. American Sociological Review 74, 358-379.

Grusky, David B. (2001): The Past, Present and Future of Social Inequality. In: Grusky, David B. (Hrsg.): Social Stratification. Class, Race and Gender in Sociological Perspective. Boulder, Colo. u. a.: Westview Press, 1-51.

Güdler, Jürgen (2003): Kooperationsnetzwerke in der Forschung – Entstehung, Struktur und Wirkung am Beispiel der Soziologie. Forschungsberichte, Band 5. Bonn: Informationszentrum Sozialwissenschaften.

Güdler, Jürgen (1996): Dynamik der Medienforschung – Eine szientometrische Analyse auf der Grundlage sozialwissenschaftlicher Fachdatenbanken. Forschungsberichte, Band 1. Bonn: Informationszentrum Sozialwissenschaften.

Hadjar, Andreas (2008): Meritokratie als Legitimationsprinzip. Die Entwicklung der Akzeptanz sozialer Ungleichheit im Zuge der Bildungsexpansion. Wiesbaden: VS Verlag für Sozialwissenschaften.

Haller, Max (2003): Soziologische Theorie im systematisch-kritischen Vergleich. Opladen: Leske + Budrich.

Halleröd, Björn (2005): Sharing of Housework and Money Among Swedish Couples: Do They Behave Rationally? European Sociological Review 21, 273-288.

Hartmann, Michael (2009): Die transnationale Klasse – Mythos oder Realität? Soziale Welt 60, 285-303.

Hausman, Jerry/Brownwyn, H. Hall/Griliches, Zvi (1984): Econometric Models for Count Data with an Application to the Patents-R&D Relationship. Econometrica 52, 909-938.

Hechter, Michael/Kim, Hyojoung/Baer, Justin (2005): Prediction Versus Explanation in the Measurement of Values. European Sociological Review 21, 91-108.

Hechter, Michael/Ranger-Moore, James/Jasso, Guillermina/Horne, Christine (1999): Do Values Matter? An Analysis of Advance Directives for Medical Treatment. European Sociological Review 15, 405-430.

Hedström, Peter (2005): Dissecting the Social. On the Principles of Analytical Sociology. Cambridge u. a.: Cambridge University Press.

Hedström, Peter (2004): Generativ Models and Explanatory Research: On the Sociology of Aage B Sørensen. Inequality 21, 13-25.

Hedström, Peter/Bearman, Peter (2009a): What is Analytical Sociology all about? An Introductory Essay. In: Hedström, Peter/Bearman, Peter (Hrsg.): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press, 4-24.

Hedström, Peter/Bearman, Peter (Hrsg.) (2009b): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press.

Hedström, Peter/Swedberg, Richard (1996): Social Mechanisms. Acta Sociologica 39, 281-308.

Hedström, Peter/Udehn, Lars (2009): Analytical Sociology and Theories of the Middle Range. In: Hedström, Peter/Bearman, Peter (Hrsg.): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press, 25-47.

Hembroff, Larry A. (1987): The Seriousness of Acts and Social Contexts: A Test of Black`s Theory of the Behavior of Law. American Journal of Sociology 93, 322-347.

Hempel, Carl G./Oppenheim, Paul (1948): Studies in the Logic of Explanations. Philosophy of Science 15, 135-175.

Hensher, David A. (2006): Revealing Differences in Willingness to Pay due to the Dimensionality of Stated Choice Designs: An Initial Assessment. Environmental & Resource Economics 34, 7-44.

Hensher, David A. (2004): How do Respondents Handle Stated Choice Experiments? Information processing strategies under varying information load. Working Papter DoD#5, University of Sydney: Institute of Transport Studies.

Herausgeber der Kölner Zeitschrift für Soziologie und Sozialpsychologie (2009): Nachrichten und Mitteilungen. Preis der Fritz Thyssen Stiftung für Sozialwissenschaftliche Aufsätze. Kölner Zeitschrift für Soziologie und Sozialpsychologie 61, 705-709.

Hermkens, Piet L. J./Boerman, Frank (1989): Consensus with respect to the fairness of incomes: Differences between social groups. Social Justice Research 3, 201-215.

Hillebrandt, Frank (2009): So einfach ist es nicht. Analytische Soziologie auf Abwegen. Symposiumsbeitrag zum Buch: Hedström, Peter: Anatomie des Sozialen. Soziologische Revue 32, 363-370.

Hillmert, Steffen/Jacob, Marita (2003): Social inequality in higher education: is vocational training a pathway leading to or away from university? European Sociological Review 19, 319-334.

Hinz, Thomas/Findeisen, Ina/Auspurg, Katrin (2008): Wissenschaftlerinnen in der DFG. Förderprogramme, Förderchancen und Funktionen (1991–2004). Bonn/New York: Wiley.

Hinz, Thomas/Liebig, Stefan (2006): Der Faktorielle Survey als Instrument zur Messung von Einstellungen in Bevölkerungsumfragen. Projektantrag an die DFG. Konstanz und Duisburg: Universitäten Konstanz und Duisburg.

Holland, Paul W. (1986): Statistics and Causal Inference. Journal of the American Statistical Association 81, 945-960.

Hook, Jennifer L./Chalsani, Satvika (2008): Gendered Expectations? Reconsidering Single Father‘s Child-Care Time. Journal of Marriage and Family 71, 978-990.

Hornbostel, Stefan (2009): Alles eine Frage der Perspektive. Münchs Kritik am Deutschen Wissenschaftssystem. Symposiumsbeitrag zum Buch: Münch, Richard: Die akademische Elite. Soziologische Revue 32, 15-20.

Hornbostel, Stefan (2001): Third party funding of German universities. An indicator of research activity? Scientometrics 50, 523-537.

Hornbostel, Stefan (1997): Wissenschaftsindikatoren. Bewertungen in der Wissenschaft. Opladen: Westdeutscher Verlag.

Horne, Christine (2003): The Internal Enforcement of Norms. European Sociological Review 19, 335-343.

Hout, Michael/DiPrete, Thomas A. (2004): What We Have Learned: RC28’s Contributions to Knowledge About Social Stratification. Working Paper of the Survey Research Center. Berkley: University of Berkley.

Hox, Joop J./Kreft, Ita G. G./Hermkens, Piet L. J. (1991): The Analysis of Factorial Surveys. Sociological Methods & Research 19, 493-510.

Hradil, Stefan (2001): Soziale Ungleichheit in Deutschland. Opladen: Leske + Budrich.

Huppé, Micheline/Cyr, Mireille (1997): Division of Household Labor and Marital Satisfaction of Dual Income Couples According to Family Life Cycle. Canadian Journal of Counseling 31, 145-162.

Jacobson, Joyce P./Levin, Laurence M. (2000): The effects of internal migration on the relative economic status of women and men. Journal of Socio-Economics 29, 291-304.

Jaeger, Michael (2006): Leistungsbezogene Budgetierung an deutschen Universitäten. Wissenschaftsmanagement 3, 30-36.

Jaime-Castillo, Antonio M. (2008): Inequality of Income, Expectations of Social Mobility and the Demand for Redistribution. Vortrag am 6. September 2008 auf dem Forum der International Sociological Association (ISA) in Barcelona.

Jann, Ben (2003): Lohngerechtigkeit und Geschlechterdiskriminierung: Experimentelle Evidenz. Unveröffentlichtes Manuskript an der Eidgenössischen Technischen Hochschule Zürich.

Jansen, Dorothea/Heidler, Richard/von Görtz, Regina (2009): Ungleiche Chancen im Wissenschaftssystem: Artefakt oder Realität? Kölner Zeitschrift für Soziologie und Sozialpsychologie 61, 463-467.

Jansen, Dorothea/Wald, Andreas/Franke, Karola/Schmoch, Ulrich/Schubert, Torben (2007): Drittmittel als Performanzindikator der wissenschaftlichen Forschung. Zum Einfluss von Rahmenbedingungen auf Forschungsleistung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 59, 125-149.

Jasso, Guillermina (2006): Factorial Survey Methods for Studying Beliefs and Judgments. Sociological Methods and Research 34, 334-423.

Jasso, Guillermina (1996): Exploring the Reciprocal Relations between Theoretical and Empirical Work: The Case of the Justice Evaluation Function (Paper in Honor of Robert K. Merton). Sociological Methods Research 24, 253-303.

Jasso, Guillermina (1994): Assessing Individual and Group Differences in the Sense of Justice: Framework and Application to Gender Differences in the Justice of Earnings. Social Science Research 23, 368-406.

Jasso, Guillermina (1988): Whom Shall We Welcome? Elite Judgments of the Criteria for the Selection of Immigrants. American Sociological Review 53, 919-932.

Jasso, Guillermina/Opp, Karl-Dieter (1997): Probing the Character of Norms: A Factorial Survey Analysis of the Norms of Political Action. American Sociological Review 62, 947-964.

Jasso, Guillermina/Webster, Murray (1999): Assessing the Gender Gap in Just Earnings and Its Underlying Mechanisms. Social Psychology Quarterly 62, 367-380.

Jasso, Guillermina/Webster, Murray (1997): Double Standards in Just Earnings for Male and Female Workers. Social Psychology Quarterly 60, 66-78.

John St., Craig/Bates, Nancy A. (1990): Racial Composition and Neighborhood Evaluation. Social Science Research 19, 47-61.

Johnson, Reed F. (2006): Comment on "Revealing Differences in Willingness to Pay Due to the Dimensionality of Stated Choice Designs: An Initial Assessment". Environmental & Resource Economics 34, 45-50.

Jürges, Hendrik (2005): Gender ideology, division of housework, and the geographic mobility of families. Review of Economics of the Household 4, 299-323.

Kaas, Leo (2009): Does Equal Pay Reduce Labor Market Inequality? Scandinavian Journal of Economics 111, 51-71.

Kalleberg, Arne L./Rosenfield, Rachel A. (1990): Work in the Family and in the Labor Market: A Cross-National, Reciprocal Analysis. Journal of Marriage and the Family 52, 331-346.

Kaube, Jürgen (2007): Abrechnung ohne Strom. Richard Münchs Polemik gegen unser Forschungssystem. Rezension in der Frankfurter Allgemeinen Zeitung vom 17.08.2007.

Kelle, Udo/Lüdemann, Christian (1995): „Grau, teurer Freund ist alle Theorie…". Rational Choice und das Problem der Brückenannahmen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 47, 249-267.

Klasen, Stephan (1998): Marriage, Bargaining, and Intrahousehold Resource Allocation: Excess Female Mortality among Adults during Early German Development, 1740 – 1860. Journal of Economic History 58, 432-467.

Klein, Markus (2002): Die Conjoint-Analyse. Eine Einführung in das Verfahren mit einem Ausblick auf mögliche sozialwissenschaftliche Anwendungen. ZA-Information 50, 7-45.

Klinger, Cornelia (2007): Achsen der Ungleichheit. Zum Verhältnis von Klasse, Geschlecht und Ethnizität. Frankfurt a. M.: Campus.

Koch, Stefan (2006): Die Deutsche Forschungsgemeinschaft und die Reform ihres Begutachtungssystems: Zur Einführung der Fachkollegien. Wissenschaftsrecht 39, 25-47.

Kreckel, Reinhard (2004): Politische Soziologie der sozialen Ungleichheit. Frankfurt a. M.: Campus.

Kreckel, Reinhard (1983): Theorie sozialer Ungleichheiten im Übergang. In: Kreckel, Reinhard (Hrsg.): Soziale Ungleichheiten. Soziale Welt, Sonderband 2. Göttingen: Otto Schwartz, 3-14.

Krücken, Georg (2009): Strukturwandel der Forschung. Symposiumsbeitrag zum Buch: Münch, Richard: Die akademische Elite. Soziologische Revue 32, 9-14.

Krumpal, Ivar/Voss, Thomas (2009): Explaining Respondents' Behavior in RRT Surveys: A Game Theoretic Approach. Arbeitspapier an der Universität Leipzig. Leipzig: Universität Leipzig.

Kuhfeld, Warren F. (2005): Marketing Research Methods in SAS. Experimental Design, Choice, Conjoint and Graphical Techniques. Cary: SAS Institute.

Kuhfeld, Warren F./Randall, Tobias D./Garratt, Mark (1994): Efficient Experimental Design with Marketing Research Applications. Journal of Marketing Research 31, 545-557.

Kunz, Volker (2004): Rational Choice. Frankfurt a. M./New York: Campus.

Laudel, Grit (2005): Quality-only assumption. Is external research funding a valid indicator for research performance? Research Evaluation 14, 27-34.

Lichter, Daniel T. (1983): Socioeconomic Returns to Migration Among Married Women. Social Forces 62, 487-503.

Liebig, Stefan/Lengfeld, Holger/Mau, Steffen (2004): Gesellschaftliche Verteilungsprobleme und der Beitrag der soziologischen Gerechtigkeitsforschung. In: Liebig, Stefan/Lengfeld, Holger/Mau, Steffen (Hrsg.): Verteilungsprobleme und Gerechtigkeit in modernen Gesellschaften. Frankfurt a. M./ New York: Campus, 7-28.

Liebig, Stefan/Mau, Steffen (2005): Wann ist ein Steuersystem gerecht? Zeitschrift für Soziologie 34, 468-491.

Liebig, Stefan/Mau, Steffen (2002): Einstellungen zur sozialen Mindestsicherung. Ein Vorschlag zur differenzierten Erfassung normativer Urteile. Kölner Zeitschrift für Soziologie und Sozialpsychologie 54: 109-134.

Liebig, Stefan/Meyermann, Alexia/Schulze, Andrea (2006): Temporal stability of justice evaluations. Paper presented at the 11th Conference of the International Society for Justice Research. Berlin: Humboldt-Universität.

Liebig, Stefan/Schupp, Jürgen (2008): Leistungs- oder Bedarfsgerechtigkeit? Soziale Welt 59, 7-30.

Lind, Imken (2004): Aufstieg oder Ausstieg? Karrierewege von Wissenschaftlerinnen. Ein Forschungsüberblick. (cews.-Beiträge Frauen in Wissenschaft und Forschung no. 2). Bielefeld: Kleine Verlag.

Lindenberg, Siegwart (1996): Die Relevanz theoriereicher Brückenannahmen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48, 126-140.

Long, Scott J./Freese, Jeremy (2006): Regression Models for Categorical Dependent Variables Using Stata. Second Edition. Texas: Stata Corp.

Long, Scott J./McGinnis, Robert (1981): Organizational Context and Scientific Productivity. American Sociological Review 46, 422-442.

Louviere, Jordan J. (2001a): What if Consumer Experiments Impact Variances as well as Means? Response Variablility as a Behavioral Phenomen. Journal of Consumer Research 28, 506-511.

Louviere, Jordan J. (2001b): Choice Experiments: an Overview of Concepts and Issues. In: Bennett, Jeff/Blamey, Russell (Hrsg.): The Choice Modelling Approach to Environmental Valuation. Cheltenham, Northhampton: Edward Elgar, 13-36.

Louviere, Jordan J./Hensher, David A./Swait, Joffre D. (2000): Stated Choice Methods. Analysis and Application. Cambridge, UK: University Press.

Lundberg, Shelly/Pollak, Robert A. (2003): Efficiency in Marriage. Review of Economics of the Household 1, 153-167.

Lundberg, Shelly/Pollak, Robert A. (1996): Bargaining and Distribution in Marriage. Journal of Economic Perspectives 10, 139-158.

Madden, Janice F. (1973): The Economics of Sex Discrimination. Lexington, Mass.: Lexington Books.

Mahony, James (2001): Beyond Correlational Analysis: Recent Innovations in Theory and Method. Sociological Forum 16, 575-593.

Manser, Marilyn/Brown, Murray (1980): Marriage and Household Decision-Making: A Bargaining Analysis. International Economic Review 21, 31-44.

Manzo, Gianluca (2007): Variables, Mechanisms and Simulations: Can the Three Methods be Synthesized? A Critical Analysis of the Literature. Revue française de sociologie 48, Supplement, 35-71.

Maxwell, Nan L. (1988): Economic Returns to Migration: Marital Status and Gender Differences. Social Science Quarterly 69, 108-121.

Mayer, Karl-Ulrich (2009): Der Geist hat die Universität verlassen. Symposiumsbeitrag zum Buch: Münch, Richard: Die akademische Elite. Soziologische Revue 32, 3-8.

Mayer, Karl-Ulrich (1987): Zum Verhältnis von Theorie und empirischer Forschung zur sozialen Ungleichheit. In: Giesen, Bernhard/Haferkamp, Hans (Hrsg.): Soziologie der sozialen Ungleichheit. Opladen: Westdeutscher Verlag, 370-392.

Mayerl, Jochen/Selke, Piet/Urban, Dieter (2005): Analyzing cognitive processes in CATI-Surveys with response latencies: An empirical evaluation of the consequences of using different baseline speed measures. Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart, SISS No. 2/2005.

Mayntz, Renate (1961): Kritische Bemerkungen zur funktionalistischen Schichtungstheorie. Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 5 (Soziale Schichtung und soziale Mobilität), 10-28.

McHugh, Kevin E./Gober, Patricia/Reid, Neil (1990): Determinants of Short- and Long-Term Mobility. Expectations for Home Owners and Renters. Demography 27, 81-95.

McLafferty, Sara/Preston, Valerie (1997): Gender, race, and the determinants of commuting: New York in 1990. Urban Geography 18, 192-212.

McLanahan, Sara/Percheski, Christine (2008): Family Structure and the Reproduction of Inequalities. Annual Review of Sociology 34, 257-276.

Melles, Thorsten (2001): Framing-Effekte in der Conjoint-Analyse. Ein Beispiel für Probleme der Merkmalsdefinition. Berichte aus der Psychologie 6. Aachen: Shaker.

Merton, Robert K. (1988): The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property. ISIS 79, 606-623.

Merton, Robert K. (1968): The Matthew Effect in Science. Science 57, 68-72.

Merton, Robert K. (1967): On Sociological Theory of the Middle Range. In: Merton, Robert K. (Hrsg.): On Theoretical Sociology. New York: Free Press, 39-72.

Meudell, Bonner M. (1982): Household and Social Standing: Dynamic and Static Dimensions. In: Rossi, Peter H./Nock, Steven L. (Hrsg.): Measuring Social Judgements. The Factorial Survey Approach. Beverly Hills u. a.: Sage, 69-94.

Meulemann, Heiner (2004): Sozialstruktur, soziale Ungleichheit und die Bewertung der ungleichen Verteilung von Ressourcen. In: Berger, Peter A./Schmidt, Volker H. (Hrsg.): Welche Gleichheit, welche Ungleichheit? Grundlagen der Ungleichheitsforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, 115-136.

Miller, J. L./Rossi, Peter H./Simpson, Jon E. (1986): Perceptions of Justice: Race and Gender Differences in Judgments of Appropriate Prison Sentences. Law & Society Review 20, 313-334.

Miller, Roland S./Perlman, Daniel/Brehm, Sharon S. (2007): Intimate Relationships. Boston: McGraw & Hill.

Mincer, Jacob (1978): Family Migration Decisions. Journal of Political Economy 86, 749-773.

Moed, Henk F. (2005): Citation Analysis in Research Evaluation. Dordrecht: Springer.

Moed, Henk F./Luwel, Marc/Houben, Josee A./Spruyt Eric/Van Den Berghe, Herman (1998): The Effects of Changes in the Funding Structure of the Flemish Universities on their Research Capacity, Productivity and Impact during the 1980's and early 1990's. Scientometrics 43, 231-255.

Molm, Linda D./Peterson, Gretchen/Takahashi, Nobuyuki (1999): Power in Negotiated and Reciprocal Exchange. American Sociological Review 64, 876-890.

Müller, Hans-Peter/Schmid, Michael (2003): Vorwort. In: Müller, Hans-Peter/Schmid, Michael (Hrsg.): Hauptwerke der Ungleichheitsforschung. Wiesbaden: Westdeutscher Verlag, 5-10.

Münch, Richard (2009): Kein Kartell, kein Monopol, keine Oligarchie? Wie die Soziologie zur Affirmationswissenschaft wird. Kölner Zeitschrift für Soziologie und Sozialpsychologie 61, 453-461.

Münch, Richard (2008): Stratifikation durch Evaluation: Mechanismen der Konstruktion von Statushierarchien in der Forschung. Zeitschrift für Soziologie 37, 60-80.

Münch, Richard (2007): Die akademische Elite. Zur sozialen Konstruktion wissenschaftlicher Exzellenz. Frankfurt a. M.: edition suhrkamp.

Neckel, Sighard/Dröge, Kai/Somm, Irene (2004): Welche Leistung, welche Leistungsgerechtigkeit? Soziologische Konzepte, normative Fragen und einige empirische Befunde. In: Berger, Peter A./Schmidt, Volker H. (Hrsg.): Welche Gleichheit, welche Ungleichheit? Grundlagen der Ungleichheitsforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, 137-164.

Neidhardt, Friedhelm (1988): Selbststeuerung in der Forschungsförderung. Das Gutachterwesen der DFG. Opladen: Westdeutscher Verlag.

Nisic, Natascha/Auspurg, Katrin (2009): Faktorieller Survey und klassische Bevölkerungsumfrage im Vergleich – Validität, Grenzen und Möglichkeiten beider Ansätze. In: Kriwy, Peter/Gross, Christiane (Hrsg.): Klein aber fein! Quantiative empirische Sozialforschung mit kleinen Fallzahlen. Wiesbaden: VS Verlag für Sozialwissenschaften, 211-246.

Nock, Steven L. (1995): Commitment and Dependency in Marriage. Journal of Marriage and the Family 57, 503-514.

Nock, Steven L. (1982): Family Social Standing: Consensus on Characteristics. In: Rossi, Peter H./Nock, Steven L. (Hrsg.): Measuring Social Judgements. The Factorial Survey Approach. Beverly Hills u. a.: Sage, 95-118.

Oaxaca, Ronald (1973): Male-Female Wage Differentials in Urban Labor Markets. International Economic Review 14, 693-709.

Ohler, Tobias/Le, Aihong /Louviere, Jordan/Swait, Joffre (2000): Attribute Range Effects in Binary Response Tasks. Marketing Letters 11, 249-260.

Okrasa, Wlodzimierz (1986): Differences in Scientific Productivity of Research Units: Measurement and Analysis of Output Inequality. Scientometrics 12, 221-239.

Opp, Karl-Dieter (2009): Das individualistische Erklärungsprogramm in der Soziologie. Entwicklung, Stand und Probleme. Zeitschrift für Soziologie 38, 26-47.

Opp, Karl-Dieter/Friedrichs, Jürgen (1996): Brückenannahmen, Produktionsfunktionen und die Messung von Präferenzen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48, 546-559.

Orme, Bryan K. (2006): Getting started with Conjoint Analysis: Strategies for Product Design and Pricing Research. Madison, Wisconsin: Research Publishers LLC.

O`Toole, Richard/Webster, Stephen W./O`Toole, Anita W./Lucal, Betsy (1999): Teachers´ Recognition and Reporting of Child Abuse: A Factorial Survey. Child Abuse & Neglect 23, 1083-1101.

Ott, Notburga (1992): Intrafamily Bargaining and Household Decisions. Berlin: Springer.

Otte, Gunnar (2005): Entwicklung und Test einer integrativen Typologie der Lebensführung für die Bundesrepublik Deutschland. Zeitschrift für Soziologie 34, 442-467.

Pager, Devah/Shepherd, Hana (2008): The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets. Annual Review of Sociology 34, 181-209.

Pakulski, Paul (2005): Foundations of a post-class analysis. In: Wright, Olin E. (Hrsg.): Approaches to Class Analysis. Cambridge: Cambridge University Press, 152-179.

Perrey, Jesko (1996): Erhebungsdesign-Effekte bei der Conjoint-Analyse. Marketing – Zeitschrift für Forschung und Praxis 18, 105-116.

Podolny, Joel/Lynn, Freda (2009): Status. In: Hedström, Peter/Bearman, Peter (Hrsg.): The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press, 544-565.

Quigley, John M./Weinberg, Daniel H. (1977): Intra-Urban Residential Mobility: A Review and Synthesis. International Regional Science Review 2, 41-66.

Rashotte, Lisa S./Webster, Murray/Whitmeyer, Joseph M. (2005): Pretesting Experimental Instructions. Sociological Methodology 35, 163-187.

Raudenbush, Stephen W./Bryk, Anthony S. (2001): Hierarchical Linear Models: Applications and Data Analysis Methods. Sage: Thousand Oaks.

Reskin, Barbara F. (2003): Including Mechanisms in our Models of Ascriptive Inequality. American Sociological Review 68, 1-21.

Robinson, William S. (1950): Ecological Correlations and the Behavior of Individuals. American Sociological Review 15, 351-357.

Roemer, John E. (1985): Should Marxists be Interested in Exploitation? Philosophy and Public Affairs 14, 30-65.

Rooks, Gerrit/Raub, Werner/Selten, Robert/Tazelaar, Frits (2000): How Inter-firm Cooperation Depends on Social Embeddedness: A Vignette Study. Acta Sociologica 43, 123-137.

Rössel, Jörg (2009): Sozialstrukturanalyse. Eine kompakte Einführung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Rössel, Jörg (2005): Plurale Sozialstrukturanalyse. Eine handlungstheoretische Rekonstruktion der Grundbegriffe der Soziologie. Wiesbaden: VS Verlag für Sozialwissenschaften.

Rössel, Jörg (2004): Macht als zentrale Dimension der Sozialstrukturanalyse. Eine handlungstheoretische Begründung. In: Berger, Peter A./Schmidt, Volker H. (Hrsg.): Welche Gleichheit, welche Ungleichheit? Grundlagen der Ungleichheitsforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, 221-240.

Rossi, Peter H. (1979): Vignette Analysis: Uncovering the Normative Structure of Complex Judgments. In: Merton, Robert K./Coleman, James S./Rossi, Peter H.: Qualitative and Quantitative Social Research: Papers in Honour of Paul F. Lazarsfeld. New York: Free Press, 176-186.

Rossi, Peter H./Alves, Wayne M. (1980): Rejoinder to Faia. American Journal of Sociology 85, 954-955.

Rossi, Peter H./Anderson, Andy B. (1982): The Factorial Survey Approach: An Introduction. In: Rossi, Peter H./Nock, Steven L. (Hrsg.): Measuring Social Judgments. The Factorial Survey Approach. Beverly Hills: Sage, 15-67.

Rossi, Peter H./Nock, Steven L. (1982): Measuring Social Judgements. The Factorial Survey Approach. Beverly Hills u. a.: Sage.

Rossi, Peter H./Sampson, William/Bose, Christine E./Jasso, Guillermina/Passel, Jeff (1974): Measuring Household Social Standing. Social Science Research 3, 169-190.

Rusconi, Alessandra (2002): Dual Careers – Insights from Social Research in Western Europe and Germany. Arbeitsbericht der Arbeitsgruppe „Wissenschaftspolitik“ der Jungen Akademie. Berlin: Die Junge Akademie.

Rusconi, Alessandra/Solga, Heike (2008): A Systematic Reflection upon Dual Career Couples. WZB Discussion Paper SP I 2008-505. Berlin: Wissenschaftszentrum Berlin für Sozialforschung.

Rusconi, Alessandra/Solga, Heike (2007): Determinants of and Obstacles to Dual Careers in Germany. Zeitschrift für Familienforschung 19, 311-336.

Sauer, Carsten (2009): Methodische Probleme von Conjoint- und Vignettenanalysen – Literaturreview. Arbeitsbericht Nr. 1 des Projekts „Der faktorielle Survey als Instrument zur Einstellungsmessung in Umfragen“. Bielefeld/Konstanz: Universität Bielefeld/Universität Konstanz.

Sauer, Carsten/Liebig, Stefan/Auspurg, Katrin/Hinz, Thomas/Donaubauer, Andy (2009): A Factorial Survey on the Justice of Earnings within the GSOEP-Pretest 2008. IZA Discussion Paper No. 4663. Bonn: Forschungsinstitut zur Zukunft der Arbeit (IZA).

Schelling, Thomas C. (1971): Dynamic Models of Segregation. Journal of Mathematical Sociology 1, 143-186.

Schimank, Uwe (1998): Funktionale Differenzierung und soziale Ungleichheit: die zwei Gesellschaftstheorien und ihre konflikttheoretische Verknüpfung. In: Giegel, Hans-Joachim (Hrsg.): Konflikt in modernen Gesellschaften. Frankfurt a. M.: Suhrkamp.

Schimank, Uwe (1992): Forschungsbedingungen der Professoren an den westdeutschen Hochschulen. Daten aus einer Befragung im Wintersemester 1990/91. MPIFG Discussion Paper 92/2. Köln: MPI für Gesellschaftsforschung.

Schmid, Michael (2009): Theoriebildung und Theoriepolitik in der Soziologie. Ein Kommentar zu Norman Braun: „Theorie in der Soziologie“. Soziale Welt 60, 199-213.

Schmid, Michael (2006): Die Forschungsheuristik mechanismischer Erklärungen. In: Schmid, Michael (Hrsg.): Die Logik mechanismischer Erklärungen. Wiesbaden: VS Verlag für Sozialwissenschaften, 139-176.

Schmidt, Volker H. (2004): Ungleichgewichtige Ungleichheiten. In: Berger, Peter A./Schmidt, Volker H. (Hrsg.): Welche Gleichheit, welche Ungleichheit? Grundlagen der Ungleichheitsforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, 73-92.

Schneider, Wolfgang Ludwig (2009): Handlung oder Interaktion? Zur Bestimmung des primären Grundbegriffs in Hedströms Analytischer Soziologie. Symposiumsbeitrag zum Buch: Hedström, Peter: Anatomie des Sozialen. Soziologische Revue 32, 355-362.

Schürmann, Lena (2005): Die Konstruktion von „Hausarbeit“ in gleichgeschlechtlichen Paarbeziehungen. In: Solga, Heike/Wimbauer, Christine (Hrsg.): „Wenn zwei das Gleiche tun…“. Ideal und Realität sozialer (Un-)Gleichheit in Dual Career Couples. Opladen: Barbara Budrich, 141-162.

Schwinn, Thomas (2008): Zur Analyse multidimensionaler Ungleichheitsverhältnisse. Österreichische Zeitschrift für Soziolgoie 33, 20-42.

Schwinn, Thomas (2007): Soziale Ungleichheit. Bielefeld: transcript.

Seyde, Christian (2005): Beiträge und Sanktionen in Kollektivgutsituationen: Ein faktorieller Survey. Arbeitsbericht Nr. 42 des Instituts für Soziologie. Leipzig: Universität Leipzig.

Shauman, Kimberley A./Noonan, Mary C. (2007): Family Migration and Labor Force Outcomes: Sex Differences in Occupational Context. Social Forces 85, 1735-1764.

Shepelak, Norma J./Alwin, Duane F. (1986): Beliefs about Inequality and Perceptions of Distributive Justice. American Sociological Review 51, 30-46.

Shlay, Anne B./Tran, Henry/Weinraub, Marsha/Harmon, Michelle (2005): Teasing apart the child care conundrum: A factorial survey analysis of perceptions of child care quality, fair market price and willingness to pay by low-income, African American parents. Early Childhood Research Quarterly 20, 393-416.

Smith, Tom W. (1986): A Study of Non-Response and Negative Values on the Factorial Vignettes on Welfare. GSS Methodological Report No. 44. Chicago: NORC.

Sniderman, Paul M./Grob, Douglas B. (1996): Innovations in Experimental Design in Attitude Surveys. Annual Review of Sociology 22, 377-399.

Snijders, Tom A./Bosker, Roel J. (1999): Multilevel Analysis. An introduction to basic and advanced modeling. London: Sage.

Sokal, Alan D./Bricmont, Jean (1997): Impostures intellectuelles. Paris: Jacob.

Solga, Heike/Powell, Justin/Berger, Peter A. (2009): Soziale Ungleichheit – Kein Schnee von gestern! Eine Einführung. In: Solga, Heike/Powell, Justin/Berger, Peter A. (Hrsg.): Soziale Ungleichheit. Klassische Texte zur Sozialstrukturanalyse. Frankfurt a. M./New York: Campus, 11-46.

Solga, Heike/Wimbauer, Christine (2005): „Wenn zwei das Gleiche tun…“ – Ideal und Realität sozialer (Un-)Gleichheit in Dual Career Couples. Eine Einleitung. In: Solga, Heike/Wimbauer, Christine (Hrsg.): „Wenn zwei das Gleiche tun…“. Ideal und Realität sozialer (Un-)Gleichheit in Dual Career Couples. Opladen: Barbara Budrich, 9-26.

Sørensen, Aage B. (2000): Towards a Sounder Basis for Class Analysis. American Journal for Sociology 105, 1523-1558.

Spence, Michael (1973): Job Market Signaling. The Quarterly Journal of Economics 87, 355-374.

Spitze, Glenna (1984): The Effect of Family Migration on Wives' Employment: How Long Does it Last? Social Science Quarterly 65, 21-36.

Stachura, Mateusz (2009): Brückenhypothesen – Kritik der ökonomischen Theorie der Ziele. Zeitschrift für Soziologie 38, 152-169.

Statistisches Bundesamt (2007): Monetäre hochschulstatistische Kennzahlen. Fachserie 11, Reihe 4.3.2. Wiesbaden: Statistisches Bundesamt.

Statistisches Bundesamt (2005a): Monetäre hochschulstatistische Kennzahlen. Fachserie 11, Reihe 4.3.2. Wiesbaden: Statistisches Bundesamt.

Statistisches Bundesamt (2005b): Finanzen der Hochschulen. Fachserie 11, Reihe 4.5. Wiesbaden: Statistisches Bundesamt.

Steiner, Peter M./Atzmüller, Christiane (2006): Experimentelle Vignettendesigns in Faktoriellen Surveys. Kölner Zeitschrift für Soziologie und Sozialpsychologie 58, 117-146.

Stickney, Lisa T./Konrad, Alison M. (2007): Gender-Role Attitudes and Earnings: A Multinational Study of Married Women and Men. Sex Roles 57, 801-811.

Struck, Olaf/Pfeifer, Christian/Krause, Alexandra (2008): Entlassungen: Gerechtigkeitsempfinden und Folgewirkungen. Theoretische Konzepte und empirische Ergebnisse. Kölner Zeitschrift für Soziologie und Sozialpsychologie 60, 102-122.

Swait, Joffre/Adamowicz, Wiktor (2001): The Influence of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching. Journal of Consumer Research 28, 135-148.

Szinovacz, Maximiliane/Harpster, Paula (1993): Employment Status, Gender Role Attitudes, and Marital Dependence in Later Life. Journal of Marriage and the Family 55, 927-940.

Taylor, Mark P. (2007): Tied Migration and Subsequent Employment: Evidence from Couples in Britain. Oxford Bulletin of Economics and Statistics 69, 795-818.

Thibaut, John W./Kelley, Harold H. (1959): The Social Psychology of Groups. New York: Wiley.

Thomson Reuters (2010): Journal Citation Reports. (Verfügbar unter: http://emedia1.bsb-muenchen.de/han/WOS/admin-apps.isiknowledge.com/JCR/JCR; letzter Zugriff am 26.01.2010).

Tumin, Melvin M. (1953): Some Principles of Stratification: A Critical Analysis. American Sociological Review 18, 387-394.

Ultee, Wout C. (1996): Do Rational Choice Approaches Have Problems? European Sociological Review 12, 167-179.

Urban, Dieter/Mayerl, Jochen (2007): Antwortlatenzzeiten in der survey-basierten Verhaltensforschung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 59, 692-713.

Verwiebe, Roland (2010): Soziale Ungleichheit – Neue Befunde aus der Gesundheits-, Armuts- und Transnationalisierungsforschung. Soziologische Revue 33, 23-33.

Wagner, Gerhard (2007): Does excellence matter? Eine wissenschaftssoziologische Perspektive. Soziologie 36, 7-20.

Wagner, Gert G./Frick, Joachim R./Schupp, Jürgen (2007): The German Socio-Economic Panel Study (SOEP) – Evolution, Scope and Enhancements. Schmollers Jahrbuch (Journal of Applied Social Science Studies) 127, 139-169.

Wallander, Lisa (2009): 25 years of factorial surveys in sociology: a review. Social Science Research 38, 505-520.

Wason, Kelly D./Polonsky, Michael J./Hyman, Michael R. (2002): Designing Vignette Studies in Marketing. Australasian Marketing Journal 10, 41-58.

Weeden, Kim A./Grusky, David B. (2005): The Case for a New Class Map. American Journal of Sociology 111, 141-212.

Wennerås, Christine/Wold, Agnes (1997): Nepotism and Sexism in Peer-Review. Nature 387, 341-343.

West, Candace/Zimmerman, Don H. (1987): Doing Gender. Gender and Society 1, 125-151.

West, S. G. (2006): Validity. In: de Vaus, David (Hrsg.): Research Design. Volume I. London u. a.: Sage, 121-142.

Will, Jeffry A. (1993): The Dimensions of Poverty: Public Perceptions of the Deserving Poor. Social Science Research 22, 312-332.

Winker, Gabriele/Degele, Nina (2009): Intersektionalität. Zur Analyse sozialer Ungleichheiten. Bielefeld: transcript.

Winship, Christopher/Mare, Robert D. (1984): Regression Models with Ordinal Variables.

Wippler, Reinhardt (1985): Die Entstehung oligarchischer Strukturen in demokratisch verfassten Organisationen. In: Büschges, Günter/Raub, Werner (Hrsg.): Soziale Bedingungen – Individuelles Handeln – Soziale Konsequenzen. Frankfurt a. M.: Lang, 23-48.

Wissenschaftsrat (2000): Drittmittel und Grundmittel der Hochschulen 1993 bis 1998 (Drs. 4717/00). Köln: Geschäftsstelle des Wissenschaftsrats.

Wittink, Dick R. /Krishnamurthi, Lakshman/Nutter, Julia B. (1982): Comparing derived importance weights across attributes. Journal of Consumer Research 8, 471-474.

Wittink, Dick R./Krishnamurthi, Lakshman/Reibstein, David J. (1989): The Effect in Differences in the Number of Attribute Levels on Conjoint Results. Marketing Letters 1, 113-123.

Wooldridge, Jeffrey M. (2003): Introductory Econometrics. A Modern Approach. Mahson, Ohio: Thomson.

Wooldridge, Jeffrey M. (2002): Econometric Analysis of Cross Section and Panel Data. Cambridge, Mass.: MIT Press.

Wright, Olin E. (2005): Conclusion: If „class“ is the answer, what is the question? In: Wright, Olin E. (Hrsg.): Approaches to Class Analysis. Cambridge: Cambridge University Press, 180-192.

Zimbardo, Philip G. (1988): Psychologie. Berlin u. a.: Springer.

Zuckermann, Harriet/Merton, Robert K. (1971): Patterns of Evaluation in Science. Minerva 9, 66-100.