

Krug, Gerhard

**Book**

## Paradoxe Wirkung von Niedriglohnzuschüssen: Zum Einfluss der Reziprozitätsnorm auf die Stabilität finanziell geförderter Beschäftigung

IAB-Bibliothek, No. 319

**Provided in Cooperation with:**

Institute for Employment Research (IAB)

*Suggested Citation:* Krug, Gerhard (2009) : Paradoxe Wirkung von Niedriglohnzuschüssen: Zum Einfluss der Reziprozitätsnorm auf die Stabilität finanziell geförderter Beschäftigung, IAB-Bibliothek, No. 319, ISBN 978-3-7639-4007-3, W. Bertelsmann Verlag (wbv), Bielefeld, <https://doi.org/10.3278/300674w>

This Version is available at:

<https://hdl.handle.net/10419/280172>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-sa/4.0/>

Institut für Arbeitsmarkt-  
und Berufsforschung

Die Forschungseinrichtung der  
Bundesagentur für Arbeit

IAB

# IAB-Bibliothek

319

Die Buchreihe des Instituts für Arbeitsmarkt- und Berufsforschung

## Paradoxe Wirkung von Niedriglohnzuschüssen

Zum Einfluss der Reziprozitätsnorm auf  
die Stabilität finanziell geförderter Beschäftigung

Gerhard Krug

Dissertationen



Institut für Arbeitsmarkt-  
und Berufsforschung

Die Forschungseinrichtung der  
Bundesagentur für Arbeit

IAB

# IAB-Bibliothek

319

Die Buchreihe des Instituts für Arbeitsmarkt- und Berufsforschung

## Paradoxe Wirkung von Niedriglohnzuschüssen

Zum Einfluss der Reziprozitätsnorm auf  
die Stabilität finanziell geförderter Beschäftigung

Gerhard Krug

Dissertationen



#### Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Dissertation der Otto-Friedrich-Universität Bamberg

Erstgutachter: Prof. Dr. Gerhard Schulze

Zweitgutachterin: Prof. Dr. Susanne Rässler

Tag der mündlichen Prüfung: 12.01.2009

Dieses E-Book ist auf dem Grünen Weg Open Access erschienen. Es ist lizenziert unter der CC-BY-SA-Lizenz.



**Herausgeber der Reihe IAB-Bibliothek:** Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit (IAB), Regensburger Straße 104, 90478 Nürnberg, Telefon (09 11) 179-0  
■ **Redaktion:** Martina Dorsch, Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit, 90327 Nürnberg, Telefon (09 11) 179-32 06, E-Mail: [martina.dorsch@iab.de](mailto:martina.dorsch@iab.de) ■ **Umschlaggestaltung:** Petra Wagler ■ **Gesamtherstellung:** W. Bertelsmann Verlag, Bielefeld ([www.wbv.de](http://www.wbv.de))  
■ **Rechte:** Kein Teil dieses Werkes darf ohne vorherige Genehmigung des IAB in irgendeiner Form (unter Verwendung elektronischer Systeme oder als Ausdruck, Fotokopie oder Nutzung eines anderen Vervielfältigungsverfahrens) über den persönlichen Gebrauch hinaus verarbeitet oder verbreitet werden.

© 2009 Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg/  
W. Bertelsmann Verlag GmbH & Co. KG, Bielefeld

In der „IAB-Bibliothek“ werden umfangreiche Einzelarbeiten aus dem IAB oder im Auftrag des IAB oder der BA durchgeführte Untersuchungen veröffentlicht. Beiträge, die mit dem Namen des Verfassers gekennzeichnet sind, geben nicht unbedingt die Meinung des IAB bzw. der Bundesagentur für Arbeit wieder.

ISBN 978-3-7639-4006-6 (Print)

ISBN 978-3-7639-4007-3 (E-Book)

ISSN 1865-4096

Best.-Nr. 300674

[www.iabshop.de](http://www.iabshop.de)

[www.iab.de](http://www.iab.de)

# Inhalt

Dank .....	7
<b>1 Einleitung und zusammenfassender Überblick.....</b>	<b>9</b>
<b>2 Problemstellung: Kombilohn und Förderabbrüche.....</b>	<b>15</b>
2.1 Wirtschaftstheoretische Grundlagen der Arbeitsmarktpolitik ....	17
2.2 Kombilohn als Maßnahme der aktiven Arbeitsmarktpolitik .....	22
2.3 Die Kombilohnförderung „Mainzer Modell“ .....	27
2.4 Kombilohn und dauerhafte Wiedereingliederung in Beschäftigung .....	32
2.5 Abbrüche von Kombilohnmaßnahmen als erklärungs- bedürftiger Sachverhalt .....	34
2.6 Zusammenfassung .....	37
<b>3 Eine Erklärung des Abbruchverhaltens .....</b>	<b>39</b>
3.1 Boudons kognitivistisches Erklärungsmodell .....	39
3.2 Die Reziprozitätsnorm in Beschäftigungsverhältnissen.....	45
3.2.1 Das Reziprozitätsprinzip als Handlungsorientierung in Tauschbeziehungen .....	46
3.2.2 Reziprozität in Beschäftigungsverhältnissen.....	54
3.2.3 Konsequenzen verletzter Reziprozität in Beschäftigungs- verhältnissen .....	55
3.2.4 Subjektive Aspekte der Reziprozität.....	58
3.3 Negative Wirkung finanzieller Anreize durch subjektiv verletzte Reziprozität.....	62
3.4 Zusammenfassung: Erklärungsmechanismus und Hypothesen ..	65
<b>4 Messung kausaler Effekte in Rubins Kausalmodell .....</b>	<b>71</b>
4.1 Rubins Kausalmodell .....	71
4.1.1 Das Ideal des Zufallsexperimentes.....	71
4.1.2 Potenzielle Ergebnisse .....	75
4.1.3 Zuweisungsmechanismus und Conditional Independence Assumption .....	80

4.1.4	Matching von Geförderten und Ungeförderten .....	84
4.2	Das RKM und die kausale Analyse von sozialen Handlungen.....	89
4.2.1	„No causation without manipulation“?.....	89
4.2.2	Soziologische Kritik am RKM und ihre soziologische Entkräftung.....	92
4.3	Alternativen zur Messung kausaler Effekte mit dem RKM .....	97
4.3.1	Das Problem fehlender Variablen in der OLS- und der ML-Schätzung .....	97
4.3.2	Ansätze für unbeobachtete Maßnahmeselektivität.....	102
4.3.2.1	Die Regression mit Kontrollfunktion als Gegenentwurf zum RKM.....	103
4.3.2.2	Instrumentvariablen- oder „Two Stage Least Square“- Schätzung .....	110
4.3.2.3	Der Preis unbeobachteter Maßnahmeselektivität .....	112
4.4	Zusammenfassung.....	115
<b>5</b>	<b>Selektivität des Kombilohnbezuges.....</b>	<b>117</b>
5.1	Mögliche Strategien zur Identifikation von Maßnahme- selektivitäten.....	117
5.1.1	Formal-mathematische Modelle .....	118
5.1.2	Ansätze einer allgemeinen Theorie der Maßnahmeselektivität..	122
5.1.3	Empirisch-qualitative Studien.....	123
5.1.4	Untersuchungen zu ähnlichen Maßnahmen .....	125
5.2	Identifikation der Maßnahmeselektivität im Mainzer Modell ....	126
5.2.1	Vorbemerkung.....	127
5.2.2	An der Maßnahmeselektivität beteiligte Akteure.....	128
5.2.3	„Kombilohnförderung“ als hypothetische Manipulation im RKM	129
5.2.4	Selektivität der Kombilohnmaßnahme „Mainzer Modell“ .....	133
5.2.4.1	Selbstselektion der Teilnehmer .....	134
5.2.4.2	Fremdselektion durch die Arbeitsämter.....	135
5.2.4.3	Fremdselektion durch die Betriebe .....	138
<b>6</b>	<b>Der kausale Effekt des Kombilohnes auf Lohnzufriedenheit und Beschäftigungsstabilität.....</b>	<b>139</b>
6.1	Verwendete Datenbasis.....	139
6.2	Messung von Selektionseinflüssen und Zielvariable.....	142

6.3	Propensity Score Matching von Kombilohngeförderten und Ungeförderten.....	144
6.4	Prüfung der Hypothesen .....	153
6.5	Prüfung zentraler Verfahrensannahmen .....	163
6.5.1	Die bedingte Unabhängigkeit des Kombilohnzuganges (CIA) .....	164
6.5.1.1	Rosenbaum-Bounds .....	164
6.5.1.2	Nutzung alternativer Zielvariablen.....	168
6.5.1.3	Verfahren mit Kontrolle unbeobachteter Selektivität .....	172
6.5.2	Die Stable Unit Treatment Value Assumption (SUTVA).....	177
6.5.2.1	Wechselwirkungen zwischen Geförderten und Ungeförderten.....	177
6.5.2.2	Versteckte Zuweisungsmechanismen.....	181
6.6	Fehlender job-worker-match als alternative Erklärung?.....	186
6.7	Reichweite der Ergebnisse.....	187
7	Zusammenfassung der Ergebnisse und Ausblick .....	189
	Abbildungsverzeichnis.....	193
	Tabellenverzeichnis.....	195
	Literatur .....	197
	Anhang.....	219
	Notation.....	219
	Wichtige Abkürzungen .....	219
	Tabellenanhang.....	220
	Kurzfassung.....	223





## Dank

Danken will ich vor allem den Gutachtern der Dissertation Gerhard Schulze und Susanne Rässler sowie Johannes Schwarze. Bei der Arbeit an der Dissertation habe ich wesentlich von ihrer Diskussionsbereitschaft, ihren Anregungen und ihrer Motivationsfähigkeit profitiert. Ich bedanke mich auch beim Institut für Arbeitsmarkt- und Berufsforschung für die Förderung der Dissertation mit einem Stipendium im Rahmen des IAB/WiSo-Graduiertenprogramms. Zum (subjektiven) Gelingen der Arbeit haben außerdem aufmunternder Zuspruch, kritischer Widerspruch, Korrekturlesen und Diskussions- und Kaffeetrinkbereitschaft von Uwe Blien, Stefan Fuchs, Eva Geithner, Barbara Hofmann, Christian Hohendanner, Hans Kiesl, Roman Lutz, Michael Moritz, Helmut Rudolph, Norbert Schanne, Kerstin Scheiblich, Julia Schneider, Gesine Stephan, Ulrich Wenzel, Joachim Wolf und Christoph Wunder sowie die Mitglieder und Mitwirkenden des IAB/WiSo-Graduiertenprogramms beigetragen.

Nürnberg, Juli 2009

Gerhard Krug



# 1 Einleitung und zusammenfassender Überblick

Im Zuge der Reformierung der Arbeitsvermittlung durch die sogenannten Hartz-Gesetze tauchten in den letzten Jahren Schlagworte wie „Kombilohn“, „Ein-Euro-Jobs“, „Ich-AG“ in der öffentlichen Debatte um die wirksame Bekämpfung der Arbeitslosigkeit auf. Dabei handelt es sich um Maßnahmen der aktiven Arbeitsmarktpolitik, die zu den wesentlichen politischen Instrumenten zur Bekämpfung von Arbeitslosigkeit gehören. Der Einsatz solcher Maßnahmen erfolgt oft unter hohem administrativen und finanziellen Aufwand, und so hat sich die Notwendigkeit der Prüfung ihrer Wirksamkeit als eine zentrale Aufgabe der wissenschaftlichen Politikberatung etabliert. Durch die Übernahme des Arbeitsförderungsgesetzes (AFG) in das Sozialgesetzbuch III (SGB III) im Jahr 1998 und die Einführung des Sozialgesetzbuches II (SGB II) ist diese Aufgabe weiter festgeschrieben und konkretisiert worden. Die von der Arbeitsverwaltung durchgeführten Maßnahmen sollen im Hinblick auf die Verbesserung der Wiedereingliederungschancen und Beschäftigungsfähigkeit der Teilnehmer (individuelle Ebene), die Effizienz der Maßnahmen (Kosten-Nutzenbetrachtung) und die volkswirtschaftlichen Nettoeffekte (Makroebene) evaluiert werden (§ 282 Abs. 3 SGB III).

Eine solche Wirkungsforschung versteht sich als Politikberatung in dem Sinne, dass sie im direkten Auftrag oder in Voraussicht der Notwendigkeit für die politischen Akteure prüft, ob die durchgeführten Maßnahmen das intendierte Ziel erreichen. Es zeigt sich, dass die Forschungspraxis der Maßnahmeevaluation – vor allem was die in dieser Arbeit relevante Perspektive der individuellen Ebene betrifft, die auch den Kern der Evaluationsforschung ausmacht – im Wesentlichen den politischen Vorgaben hinsichtlich der zu evaluierenden Ziele folgt. Eine Durchsicht der von 2000 bis 2006 in deutschen Fachzeitschriften und Diskussionspapierreihen veröffentlichten Analysen auf der individuellen Ebene zeigt, dass sich ihr überwiegender Anteil mit der Wirkung der Maßnahmen auf die Wiedereingliederung befasst und ein geringer Teil mit anderen Zielgrößen, wovon wiederum der Hauptanteil auf die Analyse der Wirkung auf den individuellen Lohn entfällt.

Von politischem Interesse ist oft nur, ob eine Maßnahme ihre Wirkung entfalten konnte oder nicht, ob sich für bestimmte Teilgruppen die Wiedereingliederungschancen erhöht haben oder nicht. Das „Warum“ scheint hier oft zweitrangig. In diesem Sinne lässt sich auch beobachten, dass die meisten Evaluationen von Maßnahmen auf die Messung der Richtung und Signifikanz von Maßnahmeeffekten abzielen und nicht der systematischen Erklärung erfolgter oder ausgebliebener Maßnahmewirkungen. Knuth (2004: 192; 194) stellt in diesem Zusammenhang fest: „Aussagen und implizite Annahmen zur Wirkungsweise aktiver Arbeitsförderung beruhen auf Alltagstheorien und Ad-hoc-Hypo-

thesen; es fehlt sowohl an theoretischer Verknüpfung als auch an empirischer Evidenz (...) Bedarf besteht für eine sozialwissenschaftliche Verhaltenstheorie am Arbeitsmarkt.“

In diesem Sinne versteht sich die vorliegende Arbeit auch als ein Beitrag dazu, Maßnahmen der aktiven Arbeitsmarktpolitik als Forschungsgegenstand ernst zu nehmen, nicht nur was deren Wirkung auf die politisch intendierten Ziele, sondern auch was die Bedingungen ausbleibender oder eintreffender Wirkungen betrifft oder auch die Frage, welche Nebenwirkungen sie entfalten, welche die intendierten Effekte eventuell konterkarieren oder verringern.

Die vorliegende Dissertation befasst sich mit einer konkreten Maßnahme der aktiven Arbeitsmarktpolitik, der Kombilohnförderung, und der Erklärung ihrer ausbleibenden Wirkung auf die dauerhafte Integration in Beschäftigung als Resultat einer gestörten Wahrnehmung der Reziprozität von Lohn und Arbeitsleistung im Beschäftigungsverhältnis. Der Kombilohn dient der Aktivierung von (mutmaßlich oder tatsächlich) arbeitsunwilligen Arbeitslosen zur Aufnahme einer Beschäftigung im Niedriglohnbereich durch die Zahlung eines finanziellen Zuschusses. Zielgruppe dieser Maßnahme sind vor allem Geringqualifizierte und Langzeitarbeitslose, bei denen man davon ausgeht, dass der Lohn, zu dem sie bereit sind eine Beschäftigung aufzunehmen, weit über dem liegt, was Betriebe zu zahlen bereit sind. Als Grund dafür wird u. a. die Höhe der staatlichen Unterstützungsleistungen für Arbeitslose identifiziert, die oft über dem Lohn liegt, den Arbeitsuchende mit niedrigen Qualifikationen erwarten können. Das Prinzip des Kombilohnes besteht nun darin, durch einen finanziellen Zuschuss an die Arbeitnehmer die Aufnahme von Beschäftigungsverhältnissen attraktiv zu machen, deren eigentlicher Lohn unterhalb der sozialstaatlichen Unterstützungsleistungen liegt. Diese Kombination aus vom Arbeitgeber gezahltem Marktlohn und staatlich gezahltem Zuschuss soll die Aufnahme von Beschäftigungsverhältnissen gegenüber dem Verbleib in Arbeitslosigkeit attraktiver machen. Ziel ist dabei aber nicht nur die einmalige Aktivierung zur Arbeitsaufnahme, sondern die dauerhafte Integration in Beschäftigung. Allerdings sind Kombilöhne sehr oft als befristete Förderung angelegt, mit deren Wegfall auch der ursprüngliche Grund zur Aufnahme der Beschäftigung entfällt. Folgt man Richard Layard (1997), dann führt jedoch bereits eine befristete Zahlung des Zuschusses zur dauerhaften Integration in Beschäftigung. Layard geht davon aus, dass die ehemals Geringqualifizierten nach erfolgter Arbeitsaufnahme Berufserfahrung sammeln und dadurch an Produktivität und also Lohnhöhe gewinnen. Daher kann nach einiger Zeit der Zuschuss wegfallen, ohne dass die positive Differenz zur Höhe der Sozialleistungen mit dem Wegfall des staatlichen Zuschusses ebenfalls entfällt.

Allerdings lässt sich aus den Ergebnissen der Begleitforschung verschiedener Varianten des befristeten Kombilohnes schließen, dass eine Vielzahl der geförder-

ten Beschäftigungsverhältnisse bereits vor Auslaufen des Zuschusses aufgelöst wird. Damit ergibt sich jedoch eine, wie man im Anschluss an Raymond Boudon formulieren kann, Rationalitätsparadoxie. Geförderte geben mit dem Abbruch einer bezuschussten Beschäftigung auch einen finanziellen Vorteil auf und verhalten sich daher auf den ersten Blick nicht rational. Diese Paradoxie besteht jedoch nur so lange, wie man davon ausgeht, dass die Beschäftigten an der Maximierung ihres Einkommens interessiert sind. In dieser Arbeit wird demgegenüber eine Erklärung für das Abbruchverhalten in Kombilohnmaßnahmen versucht, die sich am Reziprozitätsbegriff orientiert. Im Rahmen eines Beschäftigungsverhältnisses besteht die wechselseitige Erwartung von Betrieb und Beschäftigten, dass auf eine angemessene Arbeitsleistung ein angemessener oder fairer Lohn gezahlt wird und umgekehrt für einen angemessenen Lohn eine entsprechende Arbeitsleistung erbracht wird. Wird diese Erwartung enttäuscht, dann gefährdet dies die Stabilität des Beschäftigungsverhältnisses. Statt der Nutzenmaximierung als Bezugspunkt des Handelns rückt damit das ausgeglichene oder faire Verhältnis von Lohn und Arbeitsleistung ins Zentrum der Analyse.

Unter Reziprozität wird hier das Verhältnis von Gabe und Gegengabe im Allgemeinen verstanden bzw. im Kontext betrieblicher Arbeitsbeziehungen das Verhältnis zwischen Arbeitslohn und Arbeitsleistung. Der Begriff gewinnt zwar aktuell sowohl in der Soziologie als auch in Ökonomie und Sozialpsychologie an Bedeutung, bisher hat sich allerdings keine allgemeingültige Definition des Begriffes durchgesetzt. Der gemeinsame Kern der verschiedenen Verwendungsweisen besteht jedoch in folgenden drei Elementen. Erstens bezeichnet der Begriff der Reziprozität eine Norm, an der sich das Handeln der Beschäftigten orientiert. Die Norm legt fest, dass auf eine faire Lohnzahlung auch eine faire Arbeitsleistung zu erfolgen hat (konstruktive Reziprozität) bzw. bei einem unfairen Lohn auch das Recht besteht, die Arbeitsleistung zu verringern (destruktive Reziprozität). Zweitens wirkt eine Verletzung der Reziprozitätsnorm destabilisierend auf das Beschäftigungsverhältnis, da damit auch das gegenseitige Vertrauen zwischen Beschäftigten und Betrieb gestört wird. Drittens schließlich ist die Beantwortung der Frage, ob in einem konkreten Fall eine Reziprozitätsverletzung vorliegt, in hohem Maße eine subjektiv geprägte Interpretation der Beteiligten.

Auf Basis dieses Reziprozitätsbegriffes werden in der vorliegenden Arbeit folgende Hypothesen aufgestellt. Durch die staatliche Zuzahlung zum Arbeitslohn entsteht bei den geförderten Beschäftigten der Eindruck, dass der Arbeitgeber nicht bereit ist, der Norm der Reziprozität von Lohn und Arbeitsleistung zu genügen. Aus der Perspektive der Beschäftigten zahlt der Betrieb einen unfairen, zu niedrigen Lohn. Die „objektive“ Bestätigung dieses Urteils besteht darin, dass der Staat bereit ist, den vom Betrieb gezahlten Lohn mit einem Zuschuss aufzustocken. Subjektiv

sollte sich eine solche wahrgenommene Reziprozitätsverletzung in einer geringeren Wertschätzung desselben Lohnes auswirken. Die Überprüfung der Hypothese erfolgt mit Befragungsdaten zur Kombilohnmaßnahme „Mainzer Modell“ und zeigt, dass zum einen die Kombilohnförderung einer Beschäftigung im Vergleich zu regulären Beschäftigungsverhältnissen tatsächlich die Wertschätzung des gleichen Lohnes signifikant verringert, je nach verwendetem Schätzverfahren zwischen 14 und 30 Prozentpunkten.

Eine Reziprozitätsverletzung kann bei Beschäftigten unterschiedliche Reaktionen hervorrufen. Neben der Akzeptanz der Sichtweise des Betriebes können diese zu Verhaltensweisen der Arbeitnehmer führen, welche den Fortbestand der Beschäftigung direkt oder indirekt gefährden. Die direkte Gefährdung besteht darin, dass Beschäftigte dazu neigen, ein als ungerecht empfundenes Arbeitsverhältnis zu kündigen, auch unter Hinnahme finanzieller Einbußen. Indirekt gefährden Reziprozitätsverletzungen im Bereich des Arbeitslohnes die Beschäftigungsstabilität dadurch, dass die Partei, die davon ausgeht, zu wenig Gegenleistung zu erhalten, ihre Leistungsbereitschaft einschränkt. Dies kann in Beschäftigungsverhältnissen sowohl offen als demonstrativer Protest gegen die Reziprozitätsverletzung erfolgen als auch verdeckt als heimlicher Ausgleich oder sogar als bewusste Schädigung des Arbeitgebers. In beiden Fällen setzen sich die Beschäftigten jedoch dem Risiko der arbeitgeberseitigen Kündigung aus. Dies hat etwas mit der Subjektivität der Reziprozitätsverletzung zu tun. Das subjektive Element dieser Einschätzung der Reziprozität besteht vor allem darin, dass diese Sichtweise vom Arbeitgeber und seinen Vertretern im Betrieb nicht unbedingt geteilt werden muss. Dieser zahlt einen den Qualifikationen der Geförderten entsprechenden Lohn, einen Marktlohn, und kann daher aus seiner Sicht mit einer angemessenen Erfüllung der Arbeitsaufgaben rechnen. Eventuelle Leistungszurückhaltungen als Reaktionen der Geförderten auf die empfundene Reziprozitätsverletzung werden daher vom Betrieb eher als eine arbeitnehmerseitige Verletzung der Reziprozitätsnorm betrachtet.

Diese Überlegungen sprechen somit dafür, dass sich ein negatives Urteil bezüglich der Reziprozität von Lohn und Leistung nicht nur in einer geringeren Wertschätzung des Lohnes auswirkt, sondern überdies den Erfolg der Kombilohnmaßnahme auf die dauerhafte Reintegration in Beschäftigung negativ beeinflusst. Die empirischen Analysen dieses Zusammenhanges führen jedoch zu keinem eindeutigen Ergebnis. Zwar zeigt sich die untersuchte Kombilohnmaßnahme als wirkungslos in Bezug auf die Stabilisierung der Reintegration von Arbeitslosen in Beschäftigung. Doch lässt sich dieser Misserfolg nicht eindeutig bzw. einzig auf die nicht intendierte Nebenwirkung des Kombilohnes auf die Einschätzung der Reziprozität von Lohn und Leistung durch die Beschäftigten zurückführen.

Die vorliegende Arbeit ist wie folgt aufgebaut. Kapitel 2 beschreibt mit dem europäischen Arbeitsmarktmodell die wirtschaftstheoretischen Grundlagen der aktuellen Arbeitsmarktpolitik, stellt Kombilöhne im Allgemeinen und das Mainzer Modell im Besonderen als Maßnahme der aktiven Arbeitsmarktpolitik vor und führt in die Problemstellung des Zusammenhanges von Kombilohn und Beschäftigungsabbrüchen ein. In Kapitel 3 wird ein Erklärungsversuch für das Abbruchverhalten in Kombilohnmaßnahmen entwickelt. Dieser nimmt zum einen Bezug auf Boudons kognitivistisches Erklärungsmodell, das neben zweckrationalen Handlungsorientierungen auch die Relevanz wertrationaler Handlungen hervorhebt sowie die Tatsache betont, dass nicht nur objektiv wahre, sondern auch subjektiv empfundene Sachverhalte zur Erklärung sozialen Handelns herangezogen werden müssen. Zum anderen erfolgt eine Auseinandersetzung mit dem Reziprozitätsbegriff und darauf aufbauend die Formulierung eines hypothetischen Erklärungsmechanismus für Kombilohnabbrüche. Kapitel 4 stellt Rubins Kausalmodell (RKM) als Bezugsrahmen für die empirische Überprüfung der aufgestellten Hypothese vor. Dieses dient dazu, den kausalen Effekt von Maßnahmen auf bestimmte Zielvariablen zu schätzen. Es wird gezeigt, wie dieses Konzept am Ideal des Zufallsexperimentes orientiert ist, und es wird auf seine Umsetzung im Rahmen des statistischen Verfahrens „Propensity Score Matching“ sowie auf die Regression mit Kontrollfunktion und die Instrumentvariablenschätzung als alternative Verfahren eingegangen. Da das RKM bisher selten in soziologischen Analysen Anwendung gefunden hat und zum Teil auf Kritik gestoßen ist, wird in dem Kapitel auch seine Verwendbarkeit zur Analyse soziologisch interessanter Fragestellungen diskutiert.

Ein wesentlicher Aspekt des RKM ist die Auseinandersetzung mit der Selektivität der Teilnahme an arbeitsmarktpolitischen Maßnahmen. Daher befasst sich Kapitel 5 mit den gängigen Strategien zur Identifikation solcher Selektivitäten und enthält schließlich in kritischer Auseinandersetzung damit eine eigene Analyse der Maßnahmeselektivität beim Mainzer Modell. Dabei wird besonders auf die an der Entstehung der Selektivität beteiligten Akteure eingegangen. In Kapitel 6 wird das RKM dazu verwendet, um den kausalen Effekt der Kombilohnförderung einer Beschäftigung im Vergleich zu regulären Beschäftigungsverhältnissen auf die durch die Arbeitnehmer wahrgenommene Reziprozität von Lohn und Arbeitsleistung zu bestimmen. Um die Validität des erzielten Untersuchungsergebnisses zu prüfen, werden darüber hinaus eine Reihe unterschiedlicher Tests durchgeführt. Kapitel 7 beschließt die Untersuchung mit einer Zusammenfassung und einem Ausblick auf offene Forschungsfragen.





## 2 Problemstellung: Kombilohn und Förderabbrüche

Seit den 70er-Jahren des 20. Jahrhunderts ist in Deutschland wie in den meisten westlichen Industrienationen ein Phänomen zu beobachten, das als verfestigte Arbeitslosigkeit bezeichnet wird (z. B. Sachverständigenrat 2005/06: 132 ff.; Sesselmeier/Blauermeil 1997: 20 ff.). Verfestigte Arbeitslosigkeit oder auch Hysterese ist dadurch gekennzeichnet, dass in Rezessionen die Arbeitslosigkeit zunimmt, in konjunkturellen Aufschwungphasen jedoch nur leicht abnimmt. Die Arbeitslosigkeit kehrt damit nach einem Anstieg nicht mehr zu ihrem Ausgangsniveau zurück, sondern verbleibt auf einem immer höher werdenden Sockel, was sich in einem über die Zeit treppenartig ansteigenden Verlauf zeigt (vgl. Abbildung 1).

Abbildung 1: Verlauf der Arbeitslosenquote in Deutschland

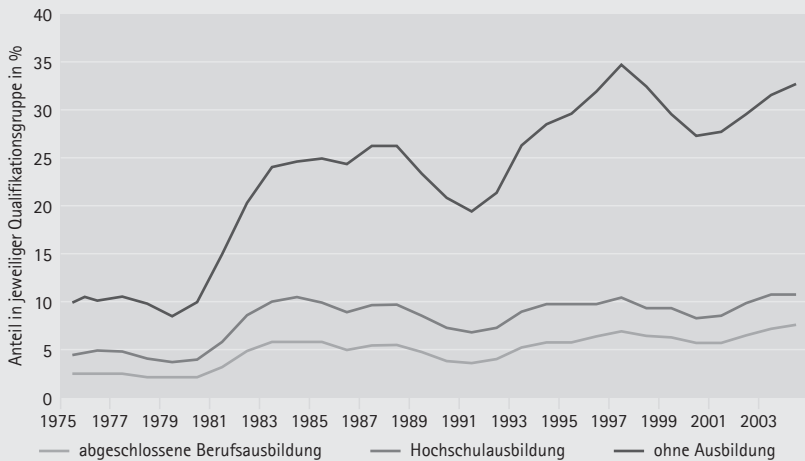


Quelle: Bundesagentur für Arbeit.

Die Betrachtung einer nach Qualifikationen unterschiedenen Arbeitslosenquote (Abbildung 2) zeigt, dass diese Entwicklung im Wesentlichen auf die veränderten Arbeitsmarktbedingungen von sogenannten Geringqualifizierten zurückgeht. Während die Quote der Arbeitslosen mit Berufsausbildung sowie Hochschulausbildung im Zeitverlauf nur gering steigt, zeigt sich in der Gruppe der Personen ohne Berufsabschluss die für die gesamte Arbeitslosigkeitsentwicklung typische Verfestigung.

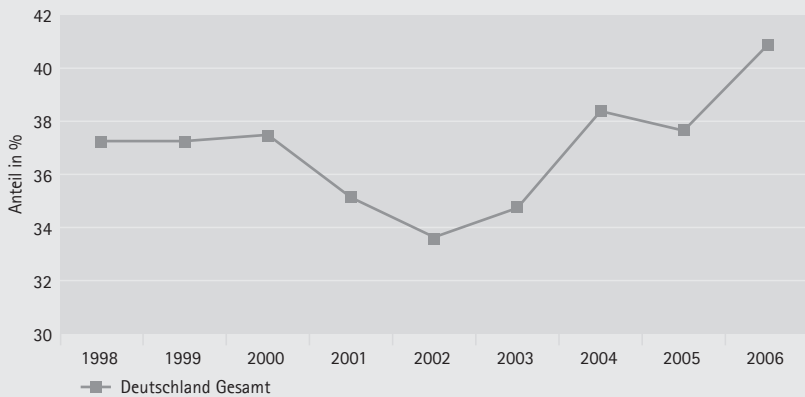
Diese schubweise Erhöhung und Verfestigung der Arbeitslosigkeit resultiert nicht nur in einer gestiegenen Eintrittswahrscheinlichkeit von Geringqualifizierten in die Arbeitslosigkeit, sondern auch in einer steigenden Verbleibsdauer. Dabei nimmt auch das Phänomen der Langzeitarbeitslosigkeit zu (vgl. Abbildung 3; leider ist nur eine eher kurze Zeitreihe verfügbar). Als langzeitarbeitslos werden Personen bezeichnet, die seit mehr als einem Jahr arbeitslos sind.

Abbildung 2: Qualifikationsspezifische Arbeitslosenquoten



Quelle: Reinberg/Hummel 2007.

Abbildung 3: Anteil der Langzeitarbeitslosen an allen Arbeitslosen im Zeitverlauf



Quelle: Bundesagentur für Arbeit.

In der auf die Erwerbsarbeit ausgerichteten Gesellschaft der Gegenwart hat Arbeitslosigkeit sowohl für die betroffenen Individuen als auch für die Gesellschaft als Ganzes erhebliche Konsequenzen. Für die Betroffenen stellt Arbeitslosigkeit nicht nur den Verlust der Haupteinkunftsquelle dar, sondern hat darüber hinaus mittelbare und unmittelbare Auswirkungen auf die verschiedensten Lebensbereiche. So gehen für viele Menschen mit dem Verlust der Erwerbsarbeit auch ein Großteil ihrer sozialen Kontakte außerhalb der Familie verloren, es kommt zu einem Verlust sozialer Anerkennung und zu Identitätskrisen, oft geht mit Arbeitslosigkeit auch ein erhöhtes Risiko für allgemein gesundheitliche und vor allem psychische Probleme

einher (Warr 1987; Morgenroth 1990, 2003; Jahoda 1981; Elkeles/Seifert 1992). Untersuchungen zeigen zudem, dass Arbeitslosigkeit auch deutliche Konsequenzen für die Rentenansprüche hat (Wunder 2005). Zu den individuellen Konsequenzen kommen gesellschaftliche hinzu, da eine steigende Arbeitslosenquote zu einer starken Belastung der Staatskassen führt, wodurch einerseits auf der Einnahmenseite immer weniger Einnahmen von Steuern und Sozialabgaben anfallen und andererseits auf der Ausgabenseite die Ausgaben für Sozialleistungen steigen.

Es verwundert daher nicht, dass die Politik erheblichen Aufwand betreibt, um das Problem der persistenten Arbeitslosigkeit zu bekämpfen. Um dies erfolgreich tun zu können, braucht die Politik zunächst eine Vorstellung von den Ursachen des Problems. Diese liefert ihr die moderne Wirtschaftstheorie (z. B. Snower 1997). Gegenwärtig beruht ein Großteil der Arbeitsmarktpolitik auf einem wirtschaftstheoretischen Konzept, das auch als das neue „Konsens-Modell für Europa“ (Stephan 1996) oder „Europäisches Arbeitsmarktmodell“ bezeichnet wird. Da es auch die Grundlage des Einsatzes von Kombilohnmodellen bildet, soll es im folgenden Abschnitt kurz dargestellt werden.

## 2.1 Wirtschaftstheoretische Grundlagen der Arbeitsmarktpolitik

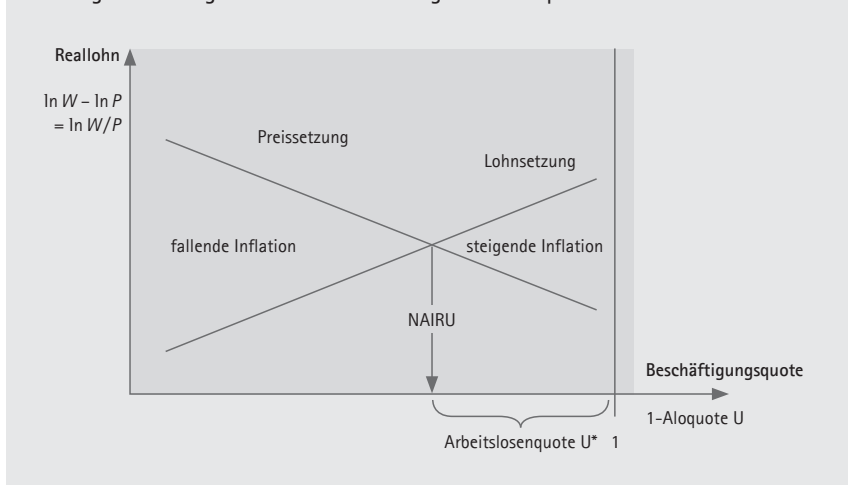
Warum geht die Zahl der Arbeitslosen selbst dann nicht wieder zurück, wenn sich die Wirtschaft wieder in einer konjunkturellen Aufschwungphase befindet und eigentlich wieder Arbeitskräfte gesucht werden? In der Ökonomie wird dies häufig im Rahmen des Arbeitsmarktmodells von Layard, Nickel und Jackman (1995) erklärt. Dieses Modell liegt weiten Teilen der aktuellen Wirtschaftspolitik explizit oder implizit zugrunde, z. B. auch den Überlegungen des Sachverständigenrates zur Begutachtung der gesamtwirtschaftlichen Entwicklung (Sachverständigenrat 2005/06), den Empfehlungen der Hartz-Kommission (Schmid 2003) und den Analysen der OECD (OECD 1993), und trägt daher zum Verständnis der Rahmenbedingungen bei, unter denen der Kombilohn als Instrument zur Lösung des Problems der Arbeitslosigkeit und speziell der Arbeitslosigkeit von Geringqualifizierten betrachtet wird.

Das Modell berücksichtigt dabei im Gegensatz zu anderen ökonomischen Arbeitsmarktmodellen die Existenz unfreiwilliger Arbeitslosigkeit und führt dies auf das Verhalten der um das Sozialprodukt konkurrierenden Arbeitnehmer und Arbeitgeber zurück. Dies wird mit der Abbildung 4 verdeutlicht.

Auf der Ordinate ist der Reallohn in logarithmierter Form abgetragen, auf der Abszisse die Beschäftigtenquote. Die Lohnsetzungskurve bildet das Lohnsetzungsverhalten der Beschäftigten bzw. der Gewerkschaften (aber auch der Unternehmen, siehe unten) ab. Ihr steigender Verlauf drückt aus, dass Arbeitskräfte umso leichter ihre Lohnansprüche durchsetzen können, je niedriger die Arbeitslosigkeit ist. Mit

steigenden Löhnen steigt auch der Reallohn. Dahinter steht folgende Argumentation (Layard/Nickel/Jackman 1995: 33 ff.): Bei niedriger Arbeitslosigkeit ist das Risiko des Verbleibes in Arbeitslosigkeit eher gering, da in einer solchen Situation eher Arbeitskräftemangel besteht. Dies stärkt zum einen die Verhandlungsmacht der Arbeitnehmer bei Lohnforderungen bzw. in deren Vertretung die Verhandlungsmacht der Gewerkschaften. Zum anderen haben auch die Unternehmen aus unterschiedlichen Gründen bei niedriger Arbeitslosigkeit ein Interesse an höheren Löhnen. Dem effizienzlohntheoretischen Shirking-Ansatz (Shapiro/Stiglitz 1984) zufolge gehen Beschäftigte bei niedriger Arbeitslosigkeit eher das Risiko ein, durch Leistungszurückhaltung (shirking, „Bummelei“) ihren Arbeitsplatz zu verlieren, weil es relativ leicht ist, einen neuen zu finden. Hohe Löhne sollen hier verhindern, dass Beschäftigte weniger Arbeitsleistung erbringen, als ihnen eigentlich möglich wäre. Ähnlich argumentiert die Labour-Turnover-Theorie (Sesselmeier 1999: 111), dass Effizienzlöhne stabile Beschäftigungsverhältnisse schaffen, weil sie eine allzu leichte Aufgabe des Beschäftigungsverhältnisses durch den Arbeitnehmer verhindern. Dadurch vermeiden Betriebe Fluktuationskosten, die etwa für Stellenausschreibungen, Einarbeitung der neuen Beschäftigten usw. anfallen. Ein weiterer Grund für erhöhte Lohnzahlungen ist nach der Theorie der adversen Selektion (Sesselmeier 1999: 113), dass Unternehmen damit Stellenbewerber mit höheren Qualifikationen anwerben können, da sich auf hoch dotierte Stellen auch eher hochqualifizierte Bewerber melden. Im Gift-Exchange-Ansatz (Akerlof 1982) wird schließlich darauf verwiesen, dass Mehrzahlungen auch dazu dienen können, sich die Loyalität der Beschäftigten zu sichern.

Abbildung 4: Erklärung der Höhe der Arbeitslosigkeit im europäischen Arbeitsmarktmodell



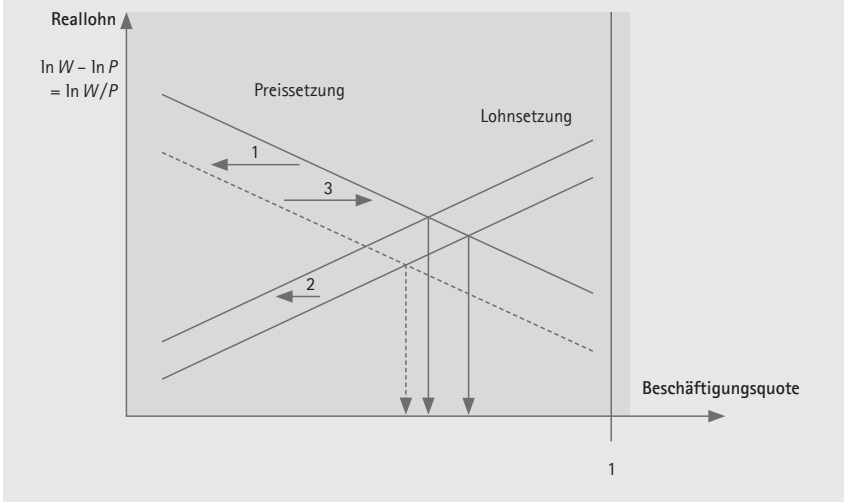
Formal lässt sich der Verlauf der Lohnsetzungskurve so beschreiben:  $\ln W = \beta_0 + \ln P^e - \beta_1 U$  wobei  $\beta_0$  ein Basisniveau des Lohnanspruches darstellt,  $P^e$  die erwarteten Preise und  $\beta_1$  die beschriebene Auswirkung der Arbeitslosigkeit  $U$  abbildet ( $\beta_0, \beta_1 > 0$ ).

Die sinkende Preissetzungskurve bezeichnet das Preissetzungsverhalten der Unternehmen. Sie erhält ihren Verlauf dadurch, dass Unternehmen in einer Volkswirtschaft mit steigendem Beschäftigungsniveau höhere Preise setzen können, wodurch der Reallohn fällt. Formal gilt entsprechend für den Verlauf der Preissetzungskurve:  $\ln P = \gamma_0 + \ln W^e - \gamma_1 U$ , mit  $\gamma_0$  als Basisniveau des Preisanspruches,  $W^e$  als den erwarteten Löhnen und  $\gamma_1$  für die Auswirkungen der Arbeitslosigkeit ( $\gamma_0, \gamma_1 > 0$ ).

Bei erwarteten gleich tatsächlichen Löhnen bzw. Preisen liegt im Schnittpunkt der beiden Kurven eine Gleichgewichtssituation vor (vgl. Gleichung 2.1), bei der die Ansprüche von Beschäftigten und Unternehmen kompatibel sind. Allerdings besteht in dieser Situation ein bestimmtes Maß an Arbeitslosigkeit, nämlich die NAIRU<sup>1</sup> („non accelerating inflation rate of unemployment“). Unterhalb der NAIRU fällt die Inflation, oberhalb davon steigt sie an.

$$NAIRU = U^* = \frac{\gamma_0 + \beta_0}{\gamma_1 + \beta_1} \tag{2.1}$$

Abbildung 5: Erklärung des hysteretischen Verlaufs der Arbeitslosigkeit



1 Zu einer kritischen Auseinandersetzung mit der Möglichkeit der empirischen Bestimmung der NAIRU vgl. Franz 2005.

Die Höhe der Arbeitslosigkeit wird in diesem Modell auf die institutionellen Rahmenbedingungen des nationalen Wirtschaftssystems zurückgeführt, die sich in Niveauparameter und Steigung der Lohn- und Preissetzungskurve ausdrücken. Aus dem Modell geht hervor: Je kleiner die Steigungsparameter und je kleiner die Niveauparameter, desto höher ist die Arbeitslosigkeit.<sup>2</sup> Der hysteretische Verlauf der Arbeitslosenquote wird in diesem Modell auf den externen Schock am Gütermarkt zurückgeführt, z. B. aufgrund gestiegener Rohstoffpreise (die Preissetzungskurve verschiebt sich nach links, vgl. Abbildung 5). Mit steigenden Preisen steigt die Arbeitslosenquote. Nachdem die Preise wieder auf ihr Ursprungsniveau zurückgehen, wird die Arbeitslosigkeit trotzdem in etwa auf dem nun höheren Niveau verbleiben, da die Arbeitnehmer bereits aufgrund der gestiegenen Preise höhere Löhne (Lohnsetzungskurve nach links) durchgesetzt haben (Stephan 1996).

Eingangs des Kapitels wurde beschrieben, dass von dieser Entwicklung gerade die Geringqualifizierten besonders stark betroffen sind. Ein Hauptproblem wird darin gesehen, dass gerade bei Geringqualifizierten eine Divergenz zwischen dem Marktlohn und ihrem Reservations- oder Anspruchslohn besteht. Der Marktlohn ist der für einen Arbeitsuchenden auf dem Arbeitsmarkt mögliche Lohn. Er ist aus Sicht der Arbeitsuchenden nicht beeinflussbar und abhängig von bestimmten individuellen Merkmalen wie der Schulbildung, der Berufserfahrung und der Qualifikation sowie von der regionalen und berufsspezifischen Arbeitsmarktsituation (Franz 1999: 26). Der Anspruchslohn dagegen ist der Lohn, ab dem ein Arbeitsuchender bereit ist, eine Beschäftigung aufzunehmen (Cahuc/Zylberberg 2004: 8 f.). Auch dieser ist von einer Reihe von Einflussfaktoren abhängig. Als zentrale Determinante gelten die staatlich gewährte Arbeitslosenunterstützung bzw. sonstige Leistungen des Sozialstaates an Arbeitslose (z. B. die ehemalige Sozialhilfe). Damit werden Arbeitslose nur dann eine Beschäftigung aufnehmen, wenn der erzielbare Lohn über dem Niveau der Sozialleistungen liegt. Gerade für Personen mit geringer Qualifikation liegt der erzielbare Marktlohn jedoch meist deutlich darunter, sodass sie sich meist gegen den Ausstieg aus der Arbeitslosigkeit entscheiden. „Denn eine halbwegs befriedigende Einkommenshöhe ohne jeglichen Arbeitsaufwand ist eine finanziell attraktive Option für viele Erwerbsfähige, die auf Grund von Werdegang und Qualifikation ohnehin keine Aussicht auf hohe Lohnsätze und hohe Erwerbseinkommen haben“ (Vierling 2002: 379). Je höher der Anspruchslohn, desto geringer ist die Wahrscheinlichkeit, dass ein Arbeitsloser ein Angebot erhält, das er bereit ist anzunehmen. Durch den längeren

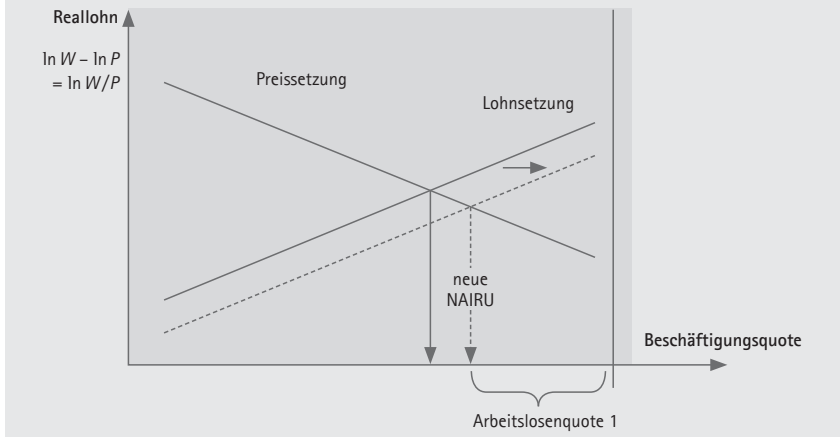
2  $\ln P + \ln W = \gamma_0 + \beta_0 + \ln P^e + \ln W^e - (\gamma_1 + \beta_1)U$

$$U(\gamma_1 + \beta_1) = \gamma_0 + \beta_0 - (\ln P + \ln W) + (\ln P^e + \ln W^e)$$

$$U = \frac{\gamma_0 + \beta_0}{\gamma_1 + \beta_1} - \frac{(\ln W - \ln W^e) + (\ln P - \ln P^e)}{\gamma_1 + \beta_1}$$

Verbleib in Arbeitslosigkeit erhöht sich aber auch das Risiko, dass die bisher erworbenen formalen oder nicht formalen Qualifikationen verloren gehen und damit der erzielbare Lohn weiter sinkt. Dieses Problem wird oft als Lohnabstandsproblem oder Anspruchslohnproblem bezeichnet, vor den institutionellen Neuregelungen im Zuge der Hartz-Gesetze auch als Sozialhilfefalle.

Abbildung 6: Lohnsetzung als Ansatzpunkt der Arbeitsmarktpolitik



Lohnabstandsprobleme in Verbindung mit dem europäischen Arbeitsmarktmodell verweisen darauf, dass der Ansatzpunkt zur Senkung der hohen Arbeitslosigkeit in einer Veränderung des Lohnsetzungsverhaltens besteht (Sachverständigenrat 2005/06: 158 ff.). Ziel der Arbeitsmarktpolitik sollte nach diesem Modell sein, die Lohnsetzungskurve nach rechts zu verschieben (Abbildung 6), das heißt, dafür zu sorgen, dass Arbeitslose bei gleichen sonstigen Bedingungen bereit sind, für weniger Lohn zu arbeiten. Bei unterstellter Gültigkeit des Modells führte dies zu einer Verschiebung der NAIU auf ein niedrigeres Niveau, also zu einer Verringerung der Arbeitslosigkeit.

Damit ist zumindest in dem hier relevanten Teilgebiet der Rahmen der wirtschaftstheoretischen Überlegungen abgesteckt, auf dem aktuell die deutsche Arbeitsmarktpolitik basiert. Diese Überlegungen sind sowohl innerhalb der Wirtschaftswissenschaften als auch in anderen Sozialwissenschaften nicht ohne kritische Gegenstimmen. Die Kritik setzt an mehreren Punkten an. Zum einen wird etwa bestritten, dass der hysteretische Prozess auf die Angebotsseite des Arbeitsmarktes zurückgeht, sondern eher auf einem Mangel an Beschäftigungsmöglichkeiten beruht, also nachfrageseitig bedingt ist (Walwei 2002: 84). Auch das den Überlegungen zum Lohnabstandsproblem zugrunde liegende Handlungsmodell eines rational kalkulierenden Einkommensmaximierers (z. B. Sesselmeier/Blauermeil 1997) und die

fehlende Berücksichtigung nicht finanzieller Aspekte der Arbeit (Blau 1991) wird kritisiert. Auf eine weitere kritische Auseinandersetzung, insbesondere mit den aus der neoklassischen Ökonomie abgeleiteten Verhaltensannahmen, die im europäischen Arbeitsmarktmodell enthalten sind, soll an dieser Stelle verzichtet werden. Stattdessen sei auf die Abschnitte 3.2.1 und 5.1.1 dieser Arbeit verwiesen.

## 2.2 Kombilohn als Maßnahme der aktiven Arbeitsmarktpolitik

Grundsätzlich bieten sich für staatliche Arbeitsmarktpolitik mehrere Ansatzpunkte (Bonß/Ludwig-Mayerhofer 2000). Auf der Makroebene tritt Arbeitsmarktpolitik in Form von Wirtschaftssteuerung auf. Ihren Höhepunkt hatte diese Form der Beschäftigungspolitik in den 70er-Jahren in der keynesianischen Nachfragepolitik. Die aktuelle Politik ist allerdings eher vom Versuch des Abbaus der steuernden Eingriffe gekennzeichnet, sodass direkte Arbeitsmarktpolitik auf der Makroebene eine immer geringere Rolle spielt. Von höherer aktueller Bedeutung ist die Mesoebene. Hier besteht die staatliche Arbeitsmarktpolitik in der Regulierung der Arbeitsbeziehungen, z. B. in der (Nicht-)Regelung der Beziehungen zwischen Arbeitgebervertretungen und Gewerkschaften, in der Regelung der Mitbestimmung der Arbeitnehmer in Betrieben und der Ausgestaltung des Kündigungsschutzes. Hier ergeben sich eine Reihe von Ansatzpunkten im Rahmen des europäischen Arbeitsmarktmodells, auf die in dieser Arbeit jedoch nicht weiter eingegangen werden soll. Stattdessen befasst sich die vorliegende Untersuchung mit staatlichen Eingriffen auf der Mikroebene in Form arbeitsmarktpolitischer Maßnahmen, zu denen auch der hier untersuchte Kombilohn gehört.

Arbeitsmarktpolitische Maßnahmen setzen auf der Mikroebene an, also am Individuum. Dabei kann man zwei unterschiedliche Maßnahmenteilen unterscheiden (Schmidt 2000). Maßnahmen der passiven Arbeitsmarktpolitik zielen vor allem auf die Kompensation der durch Arbeitslosigkeit entstandenen Einkommensausfälle ab. Sie werden als passiv bezeichnet, weil die Arbeitsverwaltung bei der Frage, ob und in welchem Umfang eine Person bestimmte Leistungen erhält (z. B. Arbeitslosengeld), keinen Gestaltungsspielraum hat. Der rechtliche Anspruch ist vom Gesetzgeber vollständig geregelt. Als Maßnahmen der aktiven Arbeitsmarktpolitik werden hingegen die Maßnahmen der Arbeitsverwaltung bezeichnet, über deren Vergabe diese eigene Entscheidungen treffen kann. Die aktive Arbeitsmarktpolitik wird in Deutschland von der Bundesagentur für Arbeit (BA) umgesetzt und umfasst eine Vielzahl von Einzelmaßnahmen (zum Zeitpunkt der Dissertation je nach Klassifikation ca. 80–200).

Aufgrund der Tatsache, dass Geringqualifizierte überproportional von Arbeitslosigkeit betroffen sind und dabei einen Großteil der Arbeitslosen insgesamt ausmachen, bilden sie eine der wichtigsten Zielgruppen der Maßnahmen aktiver Ar-



beitsmarktpolitik. Solche Maßnahmen sollen gerade den „wettbewerbsschwachen Arbeitsanbietern“ den Weg aus der Arbeitslosigkeit zurück in dauerhafte Beschäftigungsverhältnisse erleichtern, zu denen neben den Geringqualifizierten hauptsächlich Jugendliche, ältere Arbeitnehmer und behinderte Personen gezählt werden (Robinson 2000: 17). Zur Förderung der Arbeitsmarktchancen Wettbewerbsschwacher durch solche Maßnahmen werden unterschiedliche Ansätze verfolgt, die in drei Klassen eingeteilt werden können. Im Rahmen von Qualifizierungsmaßnahmen wird versucht, auf das Humankapital der Arbeitslosen Einfluss zu nehmen. Qualifizierungsmaßnahmen (z. B. Dieckhoff 2007) wie Weiterbildungsmaßnahmen, Umschulungen oder Trainingsmaßnahmen werden eingesetzt, um den durch die Langzeitarbeitslosigkeit bedingten Verlust an Humankapital zu kompensieren oder Personen ohne formalen Berufsabschluss mit einem Mindestmaß an beruflichen Qualifikationen auszustatten. Für Geringqualifizierte bedeutet dies vor allem, dass sich durch eine formale (Re-)Qualifizierung auch ihre Produktivität erhöht und damit auch ihre Chancen auf eine Beschäftigung steigen, deren Entlohnung ihrem Anspruchslohn entspricht.

Zu ähnlichen Zwecken werden auch Maßnahmen der Schaffung von Arbeitsplätzen auf einem zweiten Arbeitsmarkt eingesetzt. Es handelt sich dabei um einen zweiten Arbeitsmarkt, weil solche Arbeitsplätze dem Prinzip der Zusätzlichkeit unterliegen, also einerseits im öffentlichen Interesse liegen und andererseits die Leistung nicht auf dem ersten Arbeitsmarkt einzuholen sein darf. Eine typische Maßnahme in diesem Bereich sind die Arbeitsgelegenheiten, sog. Ein-Euro-Jobs (Hohendanner 2007). Solche Tätigkeiten sollen Teilnehmern die Chance geben, Qualifikationen durch konkrete Tätigkeit, also on-the-job aufzubauen, bzw. es soll verhindert werden, dass vorhandene formale oder sogenannte weiche Qualifikationen durch zu lange Zeit in Arbeitslosigkeit verloren gehen.

Schließlich können Maßnahmen der aktiven Arbeitsmarktpolitik auch darin bestehen, dass reguläre Beschäftigung auf dem ersten Arbeitsmarkt durch Zahlung von Zuschüssen an Arbeitgeber und/oder Arbeitnehmer subventioniert wird. Dies hat den Zweck, Arbeitsuchenden Anreize zur Aufnahme und/oder Betrieben zur Schaffung von gering entlohnter Beschäftigung ohne Einkommenseinbußen zu geben. Je nach Empfänger der Subvention lassen sich der an die Unternehmen gezahlte Lohnkostenzuschuss (z. B. Hujer/Caliendo 2003) und der an die Beschäftigten gezahlte Kombilohn (Dietz/Koch/Walwei 2006; Vierling 2002; Walwei 2002; Schneider 2002; Kaltenborn 2001; Scharpf 1994) unterscheiden.<sup>3</sup> Während Qualifizierungsmaßnah-

3 Auf einem vollständig flexiblen Markt, wie ihn die neoklassische ökonomische Theorie voraussetzt, wären die Auswirkungen von Subventionen unabhängig davon, ob sie an Unternehmen oder Beschäftigte gezahlt werden. In beiden Fällen sinkt der vom Unternehmen zu zahlende Lohn (Knabe/Schöb/Weimann 2006). Allerdings wäre auf einem solchen Markt auch kein Kombilohn nötig (für diesen Hinweis danke ich Martin Dietz).

men und Arbeitsbeschaffung auf dem zweiten Arbeitsmarkt bereits seit längerem existierende Maßnahmenteilen darstellen, gehören Kombilöhne erst seit kurzem zum Instrumentarium der deutschen Arbeitsmarktpolitik (Werner 1999).

Sowohl wissenschaftlich als auch arbeitsmarktpolitisch besteht kein Konsens darüber, was der Begriff des Kombilohnes bezeichnet und was nicht. So wird mit Kombilohn auch oft jede Subventionierung von Beschäftigung bezeichnet, unabhängig davon, ob diese an den Arbeitgeber oder Arbeitnehmer ausgezahlt wird. Unter Kombilöhnen soll hier, der Definition von Dietz, Koch und Walwei (2006) folgend, eine staatliche Transferzahlung an Arbeitnehmer verstanden werden, die an die Aufnahme einer abhängigen, sozialversicherungspflichtigen Beschäftigung unterhalb einer bestimmten Lohnhöhe gebunden ist. Ebenfalls ist umstritten, welche Ziele mit dem Kombilohn erreicht werden können bzw. sollen. Die Arbeitsmarktpolitik erhofft sich von Kombilöhnen unterschiedliche, zum Teil miteinander verbundene und zum Teil auch einander widersprechende Ziele. Im Vordergrund steht die im Rahmen des europäischen Arbeitsmarktmodells beschriebene Senkung der Arbeitslosigkeit. Sie soll vor allem durch die Aktivierung arbeitsunwilliger Leistungsbezieher und die Schaffung neuer Arbeitsplätze geschehen. Zusätzlich erwartet man sich von Kombilöhnen aber auch die Vermeidung von Armut trotz Erwerbstätigkeit und eine Reduzierung von Sozialausgaben. Wie diese Ziele erreicht werden sollen und in welcher Beziehung sie zueinander stehen, soll nun beschrieben werden.

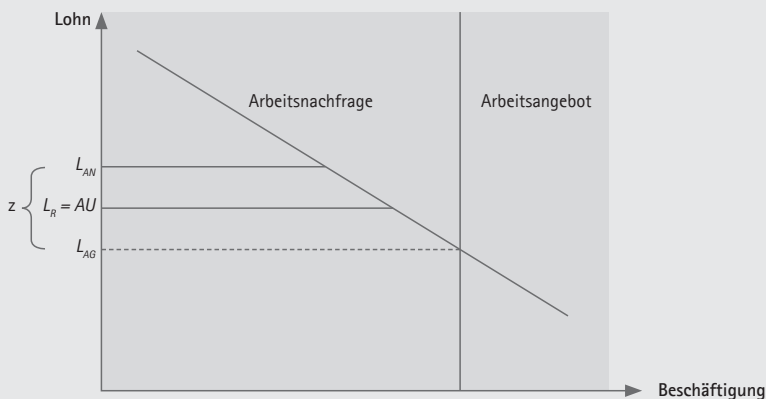
Das Hauptziel<sup>4</sup> beim Einsatz von Kombilöhnen ist die Aktivierung arbeitsunwilliger Leistungsbezieher zur Aufnahme von Beschäftigung im Niedriglohnbereich. Ausgangspunkt hierfür ist die Theorie des Anspruchslohnes (z. B. Vierling 2002; Bonin/Kempe/Schneider 2002; Spermann 2002). Dies soll Abbildung 7 verdeutlichen, die den Lohn in Bezug setzt zur Beschäftigtenzahl. Die Nachfrage nach Arbeitskräften ist dabei als eine fallende Gerade berücksichtigt, das Arbeitsangebot als eine Parallele zur y-Achse. Es wird erwartet, dass Arbeitssuchende nur dann eine Beschäftigung aufnehmen werden, wenn das aus einer Erwerbstätigkeit erzielbare Einkommen über dem Niveau der Unterstützungsleistung  $AU$  liegt. Diese bildet den Reservationslohn  $L_R = AU$ . Aufgrund deren geringer Qualifikation und/oder der erfolgten Entwertung vorhandener Qualifikationen sind Arbeitgeber aber oft nur bereit, (Langzeit-)Arbeitslose auf Arbeitsplätzen einzusetzen, deren Entlohnung ( $L_{AG}$ ) unterhalb des sozialstaatlichen Leistungsniveaus liegt. Solche Tätigkeiten sollen für Leistungsbezieher durch einen finanziellen Zuschuss attraktiv gemacht werden. Dieser hebt den insgesamt durch den Arbeitnehmer erzielbaren Lohn  $L_{AN} = L_{AG} + Z$

4 Dies gilt zumindest für Kontinentaleuropa, während in angelsächsischen Ländern eher die Vermeidung von Armut trotz Erwerbstätigkeit („working poor“) (Eichhorst 2006) im Mittelpunkt steht.

über das Unterstützungsniveau hinaus und schafft damit einen Anreiz zum Ausstieg aus der Arbeitslosigkeit.

Ein zu dieser Aktivierung paralleles Ziel ist die Schaffung neuer Arbeitsplätze. Aufgrund der hohen Unterstützungsleistungen können Unternehmen Arbeitsplätze im Niedriglohnbereich nicht besetzen und müssen sie gegebenenfalls ins Ausland verlagern. Der Kombilohnzuschuss soll dazu führen, dass solche Arbeitsplätze nun doch entstehen können, weil sich Personen finden, die bereit sind, diese Tätigkeit anzunehmen. Da der Zuschuss im Kombilohn geringer ausfallen muss als die Arbeitslosenunterstützung, erwartet man sich durch den Wiedereinstieg der Arbeitslosen in Beschäftigung auch eine Verringerung der Sozialausgaben. Dies ist aber nur dann der Fall, wenn keine Mitnahmeeffekte bestehen. Als solche werden Beschäftigungsaufnahmen bezeichnet, bei denen keine tatsächliche Aktivierung stattfand und die auch ohne Förderung vom Unternehmen angeboten und vom Arbeitslosen zum Lohn  $L_{AG}$  aufgenommen worden wäre.

Abbildung 7: Senkung des Anspruchslohnes durch Kombilohn



Quelle: Eigene Darstellung in Anlehnung an Schöb/Weimann (2006: 103).

Ein Ziel, welches auch unabhängig von der Theorie des Anspruchslohnes mit dem Kombilohn verbunden ist, ist die Bekämpfung der Armut. Geht man davon aus, dass in vielen Fällen die sozialstaatlichen Unterstützungsleistungen gerade bei Langzeitarbeitslosen (wie die frühere Sozialhilfe und das aktuelle Arbeitslosengeld II) nur das Existenzminimum abdecken, ist es für Geringqualifizierte nicht unbedingt möglich, durch eine Arbeitsaufnahme ein Einkommen zu erzielen, das aus der Armut führt. In diesem Fall spricht man von Armut trotz Erwerbstätigkeit oder „working poor“ (z. B. Bäcker 2002). Für diesen Fall ist mit dem Kombilohn auch die Bekämpfung solcher Armutslagen verbunden. Dient der Kombilohn ausschließlich der Erhöhung des erzielbaren Einkommens von Geringqualifizierten, ohne dass ein

Aktivierungseffekt besteht, ist damit allerdings keine Verringerung, sondern eine Erhöhung der Sozialausgaben verbunden.

Wie und in welchem Umfang der Kombilohn sein(e) Ziel(e) erreicht, ist auch vom konkreten institutionellen Gefüge einer Wirtschaft abhängig, in die sich der Kombilohn einpassen muss. Darunter fallen vor allem die Dauer und Höhe der Unterstützungsleistung sowie die bestehenden Hinzuverdienstmöglichkeiten von Arbeitslosen und die Transferentzugsraten von Unterstützungsleistungen bei Arbeitsaufnahme. Die Transferentzugsrate bezeichnet den Anteil der Unterstützungsleistungen, die dem Arbeitslosen für jeden hinzuverdienten Euro entzogen werden. Eine Entzugsrate von 90 % bedeutet etwa, dass bei einem Hinzuverdienst von 100 Euro nur ein Mehreinkommen gegenüber der Unterstützungsleistung von 10 Euro verbleibt. Kombiniert werden Transferentzugsraten oft mit Freibeträgen, sodass der Entzug nur für den Mehrverdienst geschieht, der über dem Freibetrag liegt.

Im Einzelnen müssen zur Ausgestaltung einer konkreten Kombilohnmaßnahme eine Vielzahl von Entscheidungen getroffen werden (Blundell 2000: 28; Dietz/Koch/Walwei 2006), von denen hier nur einige genannt werden sollen:

- Sollte der Kombilohn befristet oder unbefristet gezahlt werden?
- Soll der Zuschuss allen Arbeitslosen ausgezahlt werden oder nur einer bestimmten Teilgruppe, z. B. den Langzeitarbeitslosen?
- Wie hoch ist der Zuschuss?
- Wie erfolgt der Zuschuss, als unmittelbare Zahlung, als Sozialversicherungszuschuss, als Lohnsteuerbefreiung etc.?
- Wird der Kombilohn nur an Bedürftige gezahlt?
- Orientiert sich die Prüfung der Bedürftigkeit am individuellen Lohn, am gesamten individuellen Einkommen oder am Haushaltseinkommen?
- Wenn am individuellen Lohn, dann am Monats- oder am Stundenlohn?
- Sollte auch der Arbeitgeber einen Zuschuss erhalten?

Angesichts dieser Vielzahl von Möglichkeiten der Ausgestaltung einer Kombilohnfördermaßnahme standen Politik und Wissenschaft vor der Frage, welche Variante am besten geeignet ist, das Problem der hohen Arbeitslosigkeit bei Geringqualifizierten zu lösen. Daher wurde vor einigen Jahren damit begonnen, im Rahmen von Modellversuchen verschiedene Varianten des Kombilohnes zu testen. Zu den Modellen, die Fallzahlen im nennenswerten Umfang realisieren konnten (vgl. Kaltenborn/Wielage 2005), gehören vor allem der „Plus-Lohn Duisburg“, das „Hamburger Modell“ (Gerhardt/Meyer Larsen 2005), das „Einstiegsgehalt Baden-Württemberg“ (Dann et al. 2002: 49) und das „Mainzer Modell“.

Letzteres wurde nach der Modellerprobungsphase sogar bundesweit eingeführt und erlangte dadurch ein hohes Maß an öffentlicher Aufmerksamkeit, auch

was sein Scheitern betrifft. Daher gilt es oftmals als prototypisches Beispiel für die Unwirksamkeit von Kombilohnfördermaßnahmen.<sup>5</sup> Aufgrund der Tatsache, dass das Mainzer Modell einer wissenschaftlichen Begleitforschung unterlag, liegen zu ihm umfassende Informationen zur Maßnahmeimplementation sowie eine standardisierte Befragung mit für die Arbeitsmarktforschung ungewöhnlich breitem Variablenpektrum vor. Da diese der vorliegenden Untersuchung als Datenbasis dient, soll im nächsten Abschnitt näher auf das Mainzer Modell eingegangen werden.

### 2.3 Die Kombilohnförderung „Mainzer Modell“

Das Mainzer Modell geht in seiner Konzeption wesentlich auf Florian Gerster und Ingolf Deubel zurück (Gerster/Deubel 2001). Es wurde Mitte 2000 im Rahmen des Sonderprogramms „CAST – Chancen und Anreize zur Aufnahme sozialversicherungspflichtiger Tätigkeiten“ in jeweils zwei Arbeitsamtsbezirken der Bundesländer Rheinland-Pfalz und Brandenburg erprobt und am 1. März 2002 auf das gesamte Bundesgebiet ausgedehnt. Die Zielsetzung des Mainzer Modells wurde in den Förderrichtlinien wie folgt definiert:

„Ziele sind die Schaffung zusätzlicher Beschäftigung für gering verdienende Arbeitnehmer, allein Erziehende und Klein-Verdiener-Familien mit Kindern unter Vermeidung der sogenannten Sozialhilfefalle sowie die Steigerung der Attraktivität von Teilzeitarbeit. Hierdurch sollen die Chancen dieser Personengruppen zur Eingliederung in den ersten Arbeitsmarkt nachhaltig verbessert werden“ (CAST-Richtlinien vom 29. Juni 2000, zuletzt geändert am 6. Februar 2002).

Mit der Einführung einer Gleitzone für die Sozialabgaben der Beschäftigten (Mini- bzw. Midi-Jobs) wurde das Mainzer Modell dann zum 1. April 2003 vorzeitig eingestellt.

Die konkrete Problemsituation, vor dessen Hintergrund sich das Mainzer Modell als Lösungsvorschlag versteht, bildet die Situation des deutschen Arbeitsmarktes und dessen institutionelle Reglementierung vor den umfassenden Reformen im Rahmen der „Gesetze für moderne Dienstleistungen auf dem Arbeitsmarkt“ (die sogenannten Hartz-Gesetze). Dies muss vor allem bei der Beschreibung der Implementationsbedingungen berücksichtigt werden. Zu diesem Zeitpunkt handelte es sich z. B. noch um die Bundesanstalt für Arbeit und um Arbeitsämter. Aber auch nach der Umsetzung dieser Reformen hat das Thema Kombilohn nicht an arbeitsmarktpolitischer Relevanz verloren. So sieht der Koalitionsvertrag auf der Bundes-

<sup>5</sup> Z. B. jüngst im Rahmen der Diskussion um den Koalitionsvertrag von Union und SPD (Berliner Zeitung 2006; Nürnberger Zeitung 2006; Frankfurter Allgemeine Zeitung 2006; Frankfurter Rundschau 2006).

ebene für die 16. Legislaturperiode zwischen CDU, CSU und SPD (2005) vor, die Einführung eines Kombilohnes insbesondere zugunsten Geringqualifizierter zu prüfen.

Ausgangspunkt des Mainzer Modells ist wie bei den meisten Konzeptionen für Kombilohnmaßnahmen in Deutschland (z. B. Sinn et al. 2006) die Annahme, dass Unternehmen zwar Arbeitskräfte für geringqualifizierte Tätigkeiten suchen, es ihnen jedoch aus unterschiedlichen Gründen nicht gelingt, diese „Einfachjobs“ zu besetzen (Gerster/Deubel 2001: 64 f.). Konkret werden dafür folgende Sachverhalte verantwortlich gemacht:

- Inflexibilität der Institutionen des Arbeitsmarktes (z. B. Kündigungsschutz)
- Mängel bei personalen und sozialen Qualifikationen
- Mangelhafte Zusammenarbeit von Arbeits- und Sozialämtern
- Finanzielle Fehlanreize durch Sozialleistungen
- Fehlende Teilzeitangebote.

Das Mainzer Modell setzte dabei schwerpunktmäßig an der Beseitigung der drei letztgenannten Hürden für die Entstehung bzw. Aufnahme von Beschäftigung im Niedriglohnbereich an. Die mangelhafte Zusammenarbeit von Arbeits- und Sozialämtern stellte einen wichtigen Ansatzpunkt für die dauerhafte Wiedereingliederung Arbeitsloser in Beschäftigung dar, weil sie zu einer ineffizienten Verteilung der potenziellen Arbeitskräfte auf die Arbeits- und Sozialämter führt.<sup>6</sup> Personen, die bereits im Erwerbsleben gestanden haben und die daher Leistungen (Arbeitslosengeld oder -hilfe) von der Bundesagentur für Arbeit beziehen, werden von den Arbeitsämtern betreut. Dagegen sind die Sozialämter für die Auszahlung von Sozialhilfe an die Arbeitsfähigen zuständig, welche die formalen Voraussetzungen für solche Leistungen nicht erfüllen. Diese Aufgabenverteilung führt oft dazu, dass die Bemühungen der Ämter nicht an der bestmöglichen Vermittlungsleistung ausgerichtet sind, sondern an der internen Rationalität und den internen Ressourcen. Auch wenn Nichtleistungsbezieher ebenfalls einen Anspruch auf die Beratungsleistungen der Arbeitsämter haben, konzentrierten sich die Vermittlungsbemühungen oft auf den Personenkreis, der Anspruch auf Lohnersatzleistungen hat. Bei den Sozialämtern der Kommunen war hingegen die Strategie beobachtet worden, arbeitsfähige Empfänger von Sozialhilfe für ein Jahr mit gemeinnützigen Tätigkeiten zu beschäftigen, wodurch diese nach Ablauf des Jahres einen Anspruch auf Arbeitslosenhilfe erwerben konnten (Gerster/Deubel 2001: 65 f.). Daher war die intensive Zusammenarbeit von Sozial- und Arbeitsämtern ein wesentlicher Bestandteil des Mainzer Modells (Brinkmann et al. 2004: 44 f.).

---

6 Diese ist inzwischen im Rahmen des vierten Gesetzes für moderne Dienstleistungen am Arbeitsmarkt (Hartz IV) aufgehoben worden.

Den zweiten Ansatzpunkt des Mainzer Modells bildete das Lohnabstandsproblem (oder auch als „Sozialhilfefalle“ bezeichnet), das bereits oben angesprochen wurde. Es wird davon ausgegangen, dass gerade im Bereich niedriger Entlohnung, also dort, wo geringfügig Qualifizierte hauptsächlich beschäftigt sind, der Marktlohn unter dem Niveau der Sozialleistungen Sozialhilfe, Arbeitslosengeld bzw. Arbeitslosenhilfe liegt. Verstärkt wird dieses Problem dadurch, dass die Transferentzugsrate für Arbeitslose, die eine Beschäftigung aufnehmen, unter den damaligen Bedingungen besonders hoch war. Vierling (2002: 382) nennt für Familien in der damaligen Sozialhilfe Entzugsraten von 85 %, wobei 144 Euro anrechnungsfrei bleiben. Als besonders gravierend wurde das Lohnabstandsproblem bei den Alleinerziehenden betrachtet, aber auch in Haushalten mit Kindern, da sich bei diesen die Sozialhilfesätze am tatsächlichen finanziellen Aufwand für die Kinder orientieren, während bei Erwerbstätigkeit lediglich das pauschale Kindergeld bzw. der Kinderfreibetrag zur finanziellen Entlastung beiträgt (Walwei 2002: 86).

Im Zusammenhang mit den Alleinerziehenden steht auch der dritte Ansatzpunkt des Mainzer Modells zur Schaffung von Mehrbeschäftigung. Während Alleinerziehende aufgrund des Betreuungsaufwandes für ihre Kinder hauptsächlich Teilzeitstellen suchen, werden diese im Niedriglohnbereich jedoch weder in nennenswertem Umfang angeboten, noch bieten sie den Alleinerziehenden eine Möglichkeit, ihren Lebensunterhalt unabhängig von der Sozialhilfe zu bestreiten. Die Ursache dafür ist, dass nach Überschreiten der Geringfügigkeitsgrenze von 325 Euro in voller Höhe Sozialabgaben auf das Einkommen erhoben werden (bzw. damals wurden). Damit lohnt es sich erst ab einer bestimmten Einkommenshöhe, Arbeit aufzunehmen, Einkommenshöhen, die bei der Kombination Niedriglohn und Teilzeitarbeit schwer zu erreichen sind.

Vor diesem Problemhintergrund bietet das Mainzer Modell eine Lösung an, um die arbeitgeberseitige Schaffung von Arbeitsplätzen im Bereich einfacher und daher niedrig entlohnter Tätigkeiten zu ermöglichen und gleichzeitig das Problem der hohen Arbeitslosigkeit unter Geringqualifizierten zu bekämpfen. Im Juli 2000 startete das Mainzer Modell als Modellversuch in je zwei Arbeitsamtsbezirken der Bundesländer Rheinland-Pfalz und Brandenburg im Rahmen des Sonderprogramms „CAST – Chancen und Anreize zur Aufnahme sozialversicherungspflichtiger Tätigkeiten“.

Die finanzielle Förderung war im Mainzer Modell wie folgt gestaltet: Die Förderung bestand in einem arbeitnehmerseitigen Zuschuss zu den Sozialversicherungsbeiträgen und wurde für Eltern noch mit einem Kindergeldzuschlag ergänzt. Während der Erprobungsphase war das Mainzer Modell als Kombilohnmodell konzipiert, d. h., es wurden ausschließlich die Beschäftigten gefördert und nicht die einstellenden Betriebe. Im Zuge der bundesweiten Ausdehnung wurde bei Einarbeitung nach dem SGB III eine parallele Förderung der Betriebe mit einem Eingliederungs-

zuschuss möglich, deren maximale Förderdauer sechs Monate betrug, allerdings selten genutzt wurde.



Die Kombilohnförderung war an die Aufnahme einer sozialversicherungspflichtigen Beschäftigung im Niedriglohnbereich gekoppelt und stellte außer gleichzeitiger Bedürftigkeit der Geförderten keine weiteren Anforderungen, etwa an das Qualifikationsniveau oder die Arbeitslosigkeitsdauer. Die Beschäftigung musste tariflich bzw. ortsüblich vergütet werden, wobei ein Monatsentgelt von 325 Euro und eine wöchentliche Arbeitszeit von 15 Stunden nicht unterschritten werden durften. Kriterium der Bedürftigkeit war das Einkommen, sowohl des Antragstellers als auch eventueller Ehe- oder Lebenspartner. Zum Einkommen zählte dabei das Arbeitsentgelt aus der aufzunehmenden Beschäftigung, Einkünfte aus selbstständiger Tätigkeit, Beamten-, Renten- und Versorgungsbezüge sowie Arbeitslosen- und Unterhaltsgeld; eine Werbungskostenpauschale (87 Euro monatlich je Förderperson) wurde abgezogen. Das Vorliegen von Bedürftigkeit der potenziellen Maßnahmeteilnehmer wurde vor Ort im Arbeitsamt anhand des Einkommens des Antragstellers und eventueller Ehe- oder Lebenspartner geprüft.

Die Zuschüsse waren je nach Haushaltskonstellation gestaffelt. Alleinstehende ohne Kinder erhielten bei positivem Förderbescheid einen degressiven Zuschuss zu den Sozialversicherungsbeiträgen (bis zu 67 Euro), der bis zu einem Monatseinkommen von maximal 810 Euro gezahlt wurde. Für Alleinerziehende und Paare lagen Förderhöchstbetrag und Einkommensgrenze doppelt so hoch. Bis zu einem Monatseinkommen von 1.740 Euro wurde ein Kindergeldzuschlag von höchstens



75 Euro monatlich je minderjähriges Kind gezahlt, zuzüglich 205 Euro jeweils für das zweite bis fünfte Kind. Abbildung 8 macht am Beispiel der Alleinstehenden deutlich, wie durch den Zuschuss die von Einkommen des Arbeitnehmers abzuziehenden Sozialversicherungsbeiträge sinken, wodurch sich umgekehrt das faktische Nettoeinkommen aus einer Beschäftigung erhöht. Das Mainzer Modell wurde schließlich vorzeitig eingestellt und durch eine Neuregelung der geringfügigen Beschäftigungen (Mini-Jobs) und der Einführung einer Gleitzone für die Sozialabgaben der Beschäftigten (Midi-Jobs) zum 1. April 2003 ersetzt. Ab diesem Zeitpunkt fanden keine Neueintritte in die Förderung statt. Laufende Förderungen wurden jedoch wie vorgesehen weitergeführt und endeten spätestens im März 2006.

Allgemein lässt sich aus den Ergebnissen der Begleitforschung ableiten, dass das Mainzer Modell als gescheitert gelten kann, auch wenn an der mehrdimensionalen Zielformulierung des Mainzer Modells bereits zu erkennen ist, dass eine Evaluation der Zielerreichung nicht ganz einfach ist und ein Misserfolg in einer Zieldimension nicht zwangsläufig einen Misserfolg bezüglich der anderen Ziele nach sich zieht.<sup>7</sup>

Schon recht bald zeigte sich, dass die erhoffte Beschäftigungswirkung ausblieb: „Trotz größter Anstrengungen blieben die Bewilligungen hinter den großen Erwartungen zurück“ (Hollederer/Rudolph 2001: 15; siehe auch dies.: 2002). Die Gründe, die dafür angeführt werden, sind unterschiedlicher Art. So wird etwa argumentiert, dass „die Prämisse, dass die Arbeitsaufnahme von dem Abstand des Lohnes zum Sozialhilfeniveau („Lohnabstandsgebot“) abhängt, an den realen Problemlagen vieler Arbeitsloser vorbeigeht“ (Ziegler 2002: 531). Entsprechend kommen Hieming et al. (2005) in einer Studie zum Stellenbesetzungsprozess im Bereich sog. einfacher Dienstleistungen zum Schluss, dass dieser meist nicht am Bewerbermangel scheitert, sondern zum großen Teil daran, dass auch für einfache Tätigkeiten Qualifikationen erforderlich sind.

Als weitere Ursache wurden einerseits der hohe administrative Aufwand und mit dem Mainzer Modell konkurrierende Förderinstrumente identifiziert und andererseits die geringe Neigung der Betriebe, förderfähige Stellen anzubieten (Hollederer/Rudolph 2001). Allerdings war der Bekanntheitsgrad der Förderung unter den Betrieben eher gering, sogar wenn sie eine nach dem Mainzer Modell geförderte Arbeitskraft beschäftigen, sodass weder positive noch negative Beschäftigungseffekte im Sinne von Substitutions- oder Mitnahmeeffekten auftreten konnten (Gewiese 2004; Kaltenborn et al. 2005).

Im Gegensatz dazu konnte das Mainzer Modell in seiner eher sozialpolitischen Komponente zumindest Teilerfolge verbuchen. So spricht einiges dafür, dass es tatsächlich gelungen ist, Teilzeitarbeit attraktiver zu machen, und auch die beab-

<sup>7</sup> Eine differenzierte Bewertung geben auch Czommer, Kalina und Vanselow (2003) ab.

sichtigte finanzielle Unterstützung von Alleinerziehenden und Frauen mit Kindern gelungen ist. So konnte gerade das Einkommen aus Teilzeitarbeit solcher Personengruppen aufgestockt werden, denen es aufgrund von Betreuungspflichten nicht ohne Weiteres möglich war, eine Vollzeitarbeit aufzunehmen, und bei denen der Lohnabstand zur Sozialhilfe am geringsten ausfiel (Krug/Rudolph 2004). Darüber konnte das Mainzer Modell wenn auch nicht zur Schaffung von Mehrbeschäftigung, so doch zumindest zur Wiedereingliederung spezieller Problemgruppen des Arbeitsmarktes beitragen, wenn diese auch nur einen geringen Teil der gesamten Förderfälle ausmachten (Czommer/Kalina/Vanselow 2003).

## 2.4 Kombilohn und dauerhafte Wiedereingliederung in Beschäftigung

Egal in welcher Variante, der Kombilohn kann nur dann als wirksame Maßnahme zur Wiedereingliederung Geringqualifizierter und/oder Langzeitarbeitsloser betrachtet werden, wenn sowohl eine Aktivierung des Arbeitsangebotes und der Arbeitsnachfrage erfolgt als auch die entstandenen Beschäftigungsverhältnisse *dauerhaft* sind. Sind sie es nicht, entstehen Drehtüreffekte, bei denen Arbeitslose eine geförderte Beschäftigung aufnehmen und nach Ablauf der Förderung in die Arbeitslosigkeit zurückkehren. Allerdings, wie kann ein solcher Drehtüreffekt verhindert werden? Schließlich kommt es bei befristeten Kombilöhnen – und alles andere scheint für den Staat finanziell nicht tragbar – mit einem Auslaufen der Förderung auch zu einem Wegfall des Grundes für die Attraktivität dieser Beschäftigung, der staatlichen Lohnzahlung.

Bei einer Betrachtung des Forschungsstandes fällt auf, dass kaum theoretisch fundierte Mechanismen aufgezeigt werden, welche die Hoffnung einer Beschäftigung über die Förderdauer hinaus begründen könnten. Eine Möglichkeit ist, dass der Kombilohn diesen Mechanismus in Form einer Requalifizierung on-the-job in sich trägt. Der Mechanismus der Requalifizierung durch subventionierte Beschäftigung wird am deutlichsten von Layard (1997) im Rahmen eines Konzeptes zur Verhinderung von Langzeitarbeitslosigkeit herausgearbeitet.

„Unemployed people would have re-acquired work habits plus the ability to prove their working capacity. They would have a regular employer who could provide a reference – or (even better) retail the individual on a permanent basis. The main justification for the proposal is not that it employs people on a subsidised basis but that, by doing so, it restores them to the universe of employable people. This is an investment in human capital" (Layard 1997: 336).

Ausgangspunkt ist Layards Problemdiagnose, dass Langzeitarbeitslosigkeit im Wesentlichen auf die Berechtigungsdauer des Bezugs von Arbeitslosenunterstüt-

zung zurückgeht. Die Möglichkeit, diese Dauer streng zu begrenzen, verwirft Layard, da dies dazu führt, dass Arbeitslose auch äußerst niedrig entlohnte Tätigkeiten annehmen müssen. Er schlägt dagegen eine Strategie vor, die die Beschäftigungsfähigkeit von Arbeitslosen und damit auch den auf dem Arbeitsmarkt erzielbaren Lohn erhöhen soll. Beschäftigungsfähigkeit (employability) ist dabei ein mehrdimensionales Konzept aus Qualifikation, Lohnanspruch sowie Suchintensität.<sup>8</sup>

„People clearly differ along a wide spectrum of employability. Near one end is A: a skilled worker who is willing to take any job and searches every day. Near the other is B: unskilled worker with an excessive reservation wage who only samples the job market once a month“ (Layard 1997: 340).

Layard schlägt nun vor, dass Arbeitslose nur ein Jahr lang Arbeitslosenunterstützung erhalten sollen und dann verpflichtet werden, für mindestens sechs Monate eine Beschäftigung aufzunehmen. Dass sie einen Arbeitsplatz erhalten, soll dadurch ermöglicht werden, dass für diese sechs Monate der Arbeitgeber den Betrag erhält, den sonst der Arbeitslose als Unterstützung bezogen hätte. Im Gegensatz zum Verbleib in der Arbeitslosigkeit soll eine solche Tätigkeit auf dem ersten Arbeitsmarkt dazu führen, dass vorher abhandengekommene Beschäftigungsfähigkeit durch Aufbau von Humankapital on-the-job wiederhergestellt wird.

Obwohl Layard vorschlägt, den Zuschuss an den Arbeitgeber zu zahlen, lässt sich die Argumentation auch auf Kombilöhne übertragen. Sofern der Zuschuss nach einer festgelegten Zeit ausläuft, können dauerhafte Effekte des Kombilohnes nur dann erzielt werden, wenn „im Zuge der Einarbeitung und Bewährung der Stundenlohn wächst“ (Dietz/Koch/Walwei 2006: 3; vgl. auch Walwei 2002: 91; Czommer/Kalina/Vanselow 2003: 38) und damit der Wegfall der Förderung ausgeglichen wird. Das wird genau dann der Fall sein, wenn auch die Produktivität zunimmt, die letztendlich auch vom Humankapital abhängt. Dies setzt wiederum voraus, dass eine Requalifizierung der ehemals Geringqualifizierten und/oder Langzeitarbeitslosen erfolgt ist (Vierling 2002: 383 f.). Ob ein solcher Anstieg des Lohnes tatsächlich erfolgt, ist eine empirische Frage. Generell ist zu beobachten, dass es Beschäftigten im deutschen Niedriglohnsektor immer seltener gelingt, Einkommenszuwächse zu erzielen, die sie aus dem Niedriglohnsektor herausführen (Rhein/Gartner/Krug 2005; Bosch/Kalina 2005).

Empirische Ergebnisse zur Frage der *dauerhaften* Wiedereingliederung Arbeitsloser durch Kombilohnmaßnahmen liegen bisher für Deutschland kaum vor. Für das Hamburger Modell zeigt sich zwar ein positiver Effekt auf die Beschäftigungsdauer (Jirjahn/Pfeiffer/Tsertsvadze 2006), doch ist dieses kein reines Kombilohnmodell,

8 Beschäftigungsfähigkeit ist auch als Ziel der deutschen Arbeitsmarktpolitik verankert (siehe SGB III § 1 Satz 2). Zu den Schwierigkeiten dieses Begriffes vgl. Deeke/Kruppe 2003.

sondern ergänzt durch einen Lohnkostenzuschuss an den Arbeitgeber. Ereignisanalysen im Rahmen der Begleitforschung zum Mainzer Modell, einer reinen Kombilohnförderung, kommen dagegen zum Ergebnis, dass in geförderten Beschäftigungsverhältnissen kürzere Beschäftigungsdauern herrschen als bei ungeforderten Beschäftigungsverhältnissen (Hartmann/Krug 2004; Kaltenborn et al. 2005; Eichhorst et al. 2005).

## 2.5 Abbrüche von Kombilohnmaßnahmen als erklärungsbedürftiger Sachverhalt

Die Beendigung von Beschäftigungsverhältnissen *nach* Auslaufen der Förderung lässt sich noch im Rahmen der bestehenden Theorieansätze begründen, nämlich durch das Wegfallen des ursprünglichen Vorteils einer geförderten Beschäftigung (ob diese Begründung empirisch zutrifft, ist eine davon zu unterscheidende Frage). Beim Beschäftigungsende *während* der Laufzeit der Förderung ist dies schwieriger. Eine solche Beendigung der bezuschussten Beschäftigung kann auf zwei Akteure zurückgehen, sowohl auf den Arbeitgeber als auch auf den Arbeitnehmer selbst. Für das Verhalten beider Akteure liefert die ökonomische Effizienzlohntheorie gerade umgekehrte Prognosen der Beschäftigungsstabilität. Einerseits führen höhere Löhne aus Sicht der Effizienzlohntheorie zu einer höheren Neigung der Beschäftigten, im Betrieb zu verbleiben (Sesselmeier/Blauermel 1997: 163 f.; Franz 1999: 315). Im konkreten Fall des Kombilohnzuschusses lässt sich dies so begründen, dass Geförderte in der aufgenommenen Beschäftigung einen höheren Verdienst erzielen, als dies in alternativen Arbeitsmarktzuständen der Fall wäre. Zum einen steigen bei geförderten Beschäftigungsverhältnissen die Opportunitätskosten für die „Freizeit“ (d. h. für die Rückkehr in die Arbeitslosigkeit) an, ein Beschäftigter würde durch die Kündigung auf mehr Lohn verzichten als ohne Förderung. Zum anderen ist auch der Wechsel in ein anderes Beschäftigungsverhältnis unwahrscheinlicher, weil dort Einkommenseinbußen drohen, weil nicht gewährleistet ist, dass ebenfalls eine Bezuschussung erfolgt, da diese keine Pflicht-, sondern eine Ermessensleistung des Arbeitsamtes darstellt.

Andererseits liefert die Effizienztheorie auch eine Begründung dafür, warum Arbeitgeber an geförderten Beschäftigten stärker festhalten sollten als an ungeforderten, auch wenn die Förderung allein arbeitnehmerseitig ausgezahlt wird. Unternehmen sind mit dem Problem konfrontiert, dass Arbeitsverträge nicht alle Arbeitspflichten genau festlegen können bzw. eine umfassende Kontrolle der Vertragseinhaltung unrentabel hohe Kosten erzeugen würde. Um zu verhindern, dass Beschäftigte sich mit ihrer Arbeitsleistung zurückhalten, zahlen Betriebe laut dieser Variante der Effizienzlohntheorie höhere Löhne als auf dem Arbeitsmarkt üb-

lich. Der höhere Lohn ist es, der „die Beschäftigten nicht nur zu höheren Leistungen anspornt, sondern auch die Strafkosten für solche Arbeitskräfte heraufsetzt, die dann noch beim Bummeln erwischt werden, weil ihnen dann gekündigt wird und sie mit Arbeitslosigkeit und deshalb – je nach Höhe der Arbeitslosenunterstützung – mit erheblichen Einkommenseinbußen konfrontiert sind“ (Franz 1999: 310). Im Falle des Kombilohnes wird ein solcher Anreiz zur Mehrleistung bzw. Furcht vor Entdeckung der Leistungszurückhaltung geschaffen, ohne dass der Arbeitgeber einen höheren Lohn zahlen müsste, sodass auch aus dieser Perspektive eine stabilere Beschäftigung von geförderten Arbeitnehmern zu erwarten wäre.

Zu ähnlichen Schlüssen kommt man aus Sicht der Motivationspsychologie. Die Arbeitsmotivation ist eine variable Größe und hat Auswirkungen auf verschiedene Bereiche der Beschäftigung (Kleinbeck 1996: 79 ff.). Mit der Arbeitsmotivation steigt zum einen die Arbeitsproduktivität der Beschäftigten und damit auch das Interesse der Arbeitgeber, diese Beschäftigten im Betrieb zu halten. Zum anderen sinkt die Wahrscheinlichkeit für krankheitsbedingte Ausfälle, sei es aufgrund tatsächlicher Erkrankungen oder aufgrund der Entscheidung des Beschäftigten, seiner Arbeitsstelle ohne das Vorliegen einer realen Erkrankung fernzubleiben (ebd. 88 f.). Auch in diesem Fall fördert eine hohe Arbeitsmotivation indirekt die Arbeitgeberentscheidung zur Beibehaltung des Beschäftigungsverhältnisses. Schließlich bindet eine hohe Motivation auch die Beschäftigten selbst an den Betrieb und verringert die Kündigungsbereitschaft (ebd. 79). Eine zentrale Determinante der Arbeitsmotivation und damit der Leistungsbereitschaft und des Wunsches, im Betrieb zu verbleiben, sind Belohnungen. Diese können in externale und internale Belohnungsformen unterschieden werden (entsprechend wird auch von extrinsischer und intrinsischer Motivation gesprochen). Externale Belohnung besteht vor allem in der Bezahlung für die geleistete Tätigkeit in Form des Arbeitslohnes, kann aber auch in Form von Beförderungen oder Anerkennung der Arbeitsleistung erfolgen. Weitere Belohnungsarten sind besondere Sozialleistungen, Alterszuschläge, flexible Arbeitszeiten, Firmenwagen etc. Im Unterschied dazu sind internale Belohnungen eher emotional orientiert, wie der Stolz auf die eigene Leistung und die Freude über Erfolgserlebnisse (ebd. 86). Aus dieser Perspektive erhöhen Kombilöhne durch die höhere Lohnzahlung die extrinsische Arbeitsmotivation einer Beschäftigung und führen daher tendenziell zu höherem Nutzen des Beschäftigten durch erhöhte Leistungsbereitschaft und geringere Fehlzeiten sowie zu geringerer Bereitschaft des Beschäftigten, den Betrieb zu verlassen.<sup>9</sup>

9 Der Wegfall des Kombilohnes nach Auslaufen der Förderung könnte in diesem Zusammenhang die Wirkung einer Strafe haben.

Abbildung 9: Abbruchquoten in verschiedenen Kombilohnmaßnahmen

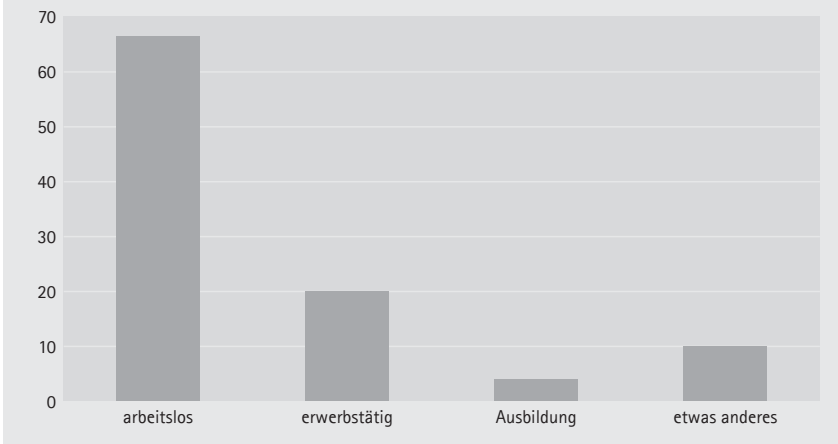


Im Gegensatz zu diesen Überlegungen ist allerdings zu beobachten, dass Kombilohnmodelle tendenziell eine eher hohe Abbruchquote der Beschäftigungen nicht nur nach, sondern vor allem während der Förderdauer aufweisen (vgl. Abbildung 9). So betrug die Quote der frühzeitigen Beendigungen der Förderung bei einer Höchstdauer der Förderung von 10 Monaten im Hamburger Modell 56 % (Gerhardt/Meyer Larsen 2005: 12; vgl. auch Jirjahn/Pfeiffer/Tsertsvadze 2006: 27; Gerhardt/Prang 2003: 22). Im Mainzer Modell mit einer maximalen Förderdauer von 36 Monaten (nach der bundesweiten Ausdehnung) brachen in Westdeutschland innerhalb der ersten 12 Fördermonate 40 % (Ostdeutschland 50 %) und innerhalb der ersten 18 Monate 50 % die geförderte Beschäftigung ab (Kaltenborn et al. 2005: 38). Ähnlich beim Einstiegsgeld in Baden-Württemberg, dort brachen ca. 7 % die Förderung frühzeitig ab (Dann et al. 2002: 49). Obwohl rational handelnde Individuen durch den Zuschuss zur Beibehaltung der Beschäftigung zumindest bis zu dessen Auslaufen motiviert wären, sind Kombilohnmaßnahmen über die verschiedenen Maßnahmevarianten hinweg von hohen Abbruchquoten gekennzeichnet. Man kann daher von einer Lohn- oder Rationalitäts-Paradoxie (vgl. dazu Abschnitt 3.3) sprechen.

Diese Paradoxie könnte sich dadurch auflösen, wenn mit dem Abbruch der Kombilohnmaßnahme für die Betroffenen eine Einkommensverbesserung verbunden wäre. Zumindest für die Abbrecher im Mainzer Modell lässt sich jedoch zeigen, dass dies für den überwiegenden Anteil nicht der Fall ist (Abbildung 10). Ganze zwei Drittel der befragten ehemals kombilohngefördert Beschäftigten sind zum Befragungszeitpunkt wieder arbeitslos, und nur 20 % sind in eine andere Beschäftigung gewechselt. Von diesen geben zwar 61 % (das entspricht 12,6 % der Geförderten) an, sie hätten sich durch die neue Erwerbstätigkeit im Vergleich zur geförderten Tätigkeit verbes-

sert. Allerdings lässt die Art der Fragestellung weder den Schluss zu, worauf diese Verbesserung beruht (z. B. Einkommen, Arbeitsplatzsicherheit etc.), noch ob ein Verbesserungswunsch der Grund für das Ende der geförderten Beschäftigung war.<sup>10</sup> Die Paradoxie bleibt damit zunächst bestehen, die überwiegende Zahl der Abbrüche von Kombilohnmaßnahmen (zumindest im Mainzer Modell) ist nicht durch den Wechsel in einen finanziell und auch sonst vorteilhafteren Erwerbsstatus motiviert.

Abbildung 10: Beschäftigungsstatus nach Abbruch der geförderten Beschäftigung im Mainzer Modell



## 2.6 Zusammenfassung

In diesem Abschnitt wurde gezeigt, auf welcher handlungstheoretischen Basis Kombilohnmaßnahmen beruhen. Es wird im mehr oder weniger stringenten Anschluss an die neoklassische Ökonomische Theorie davon ausgegangen, dass eine Mehrzahlung Arbeitslose mit hohen Reservationslöhnen dazu motiviert, Beschäftigungsverhältnisse unterhalb dieses Anspruchsniveaus einzugehen. Implizit wird dabei auch eine positive Wirkung des Anreizes über die Arbeitsaufnahme hinaus auf den Verbleib in Beschäftigung angenommen, da erst dann von einer erfolgreichen Integration in den Arbeitsmarkt gesprochen werden kann, wenn ein dauerhaftes Beschäftigungsverhältnis entsteht. Diese positive Wirkung lässt sich mit den durch den finanziellen Zuschuss gestiegenen Opportunitätskosten für die Freizeit bzw. den Zustand der Arbeitslosigkeit begründen. Entgegen der Annahme der Stabilisierung von Beschäftigungsverhältnissen durch den Kombilohn wird jedoch über

<sup>10</sup> Die Frage lautet konkret: „Alles in allem: Haben Sie sich durch Ihre heutige Erwerbstätigkeit im Vergleich zur vorherigen verbessert, sind die beiden etwa gleichwertig oder haben Sie sich verschlechtert?“ und ist Teil der in Abschnitt 6.1 beschriebenen Erhebung.

die verschiedenen Modellvarianten hinweg eine hohe Abbruchquote beklagt. Damit stellt sich die Frage, wie eine solche als Lohnparadoxie charakterisierte Reaktion der Geförderten auf den finanziellen Zuschuss zu erklären ist.

Bisher ist das Lohnparadox im Abbruch von Kombilohnmaßnahmen nicht zur Kenntnis genommen, geschweige denn untersucht worden. Die wenigen Hinweise auf eine Auseinandersetzung mit der Abbruchquote in Kombilohnfördermaßnahmen betrachten diese entweder als Ausdruck einer generell hohen Fluktuation auf Arbeitsplätzen im Niedriglohnbereich (Kaltenborn et al. 2005: 38). Oder sie gestehen eine im Vergleich zu regulären Beschäftigungsverhältnissen überhöhte Abbruchquote zu, interpretieren diese aber positiv. Daraus, dass 61 % der Abbrüche auf einer arbeitgeberseitigen Beendigung des Beschäftigungsverhältnisses beruhen, wird etwa im Rahmen des Hamburger Modells geschlossen, „dass für die Arbeitgeber die Einstellung vormals Arbeitsloser ein Experiment darstellt, das sie ohne die Förderung nicht gewagt hätten“ (Gerhardt/Meyer Larsen 2005: 13). Eine empirische Überprüfung dieser eher als Ad-hoc-Hypothesen (vgl. Knuth 2004: 192) denn als Erklärungen fungierenden Begründungen findet allerdings nicht statt. Im Folgenden soll eine andere Erklärung für das Abbruchverhalten entwickelt werden. Sie besteht darin, dass der Kombilohnbezug von den geförderten Personen nicht als zusätzlicher Lohn empfunden wird, sondern als ein Signal interpretiert wird, dass der Arbeitgeber nicht bereit ist, sich an die Norm der Reziprozität zu halten und einen fairen Lohn zu zahlen. Abbrüche werden dann als Resultat unterschiedlichster Reaktionen auf einen subjektiv als unfair betrachteten Lohn verständlich.



### 3 Eine Erklärung des Abbruchverhaltens

Die neoklassische Fundierung des Kombilohnzuschusses legt es nahe, sich mit Erklärungen auseinanderzusetzen, die über die Maximierung des finanziellen Nutzens noch weitere Handlungsorientierungen zulassen. Hierzu wird im Abschnitt 3.1 Boudons kognitivistisches Erklärungsmodell vorgestellt, das explizit in der Auseinandersetzung mit dem neoklassischen Verständnis rationaler Handlungen als nutzenmaximierend entwickelt wurde. In Abschnitt 3.2 wird dann die Relevanz der Reziprozitätsnorm in Beschäftigungsverhältnissen herausgearbeitet, und in Abschnitt 3.3 wird dargelegt, inwiefern diese Relevanz dazu führen kann, dass finanzielle Anreize nicht die intendierte Wirkung entfalten. Abschnitt 3.4 leitet daraus eine Erklärung für das Abbruchverhalten in Kombilohnmaßnahmen ab und formuliert die entsprechenden Hypothesen.

#### 3.1 Boudons kognitivistisches Erklärungsmodell

Die neoklassische Theorie der Nutzenmaximierung besteht in ihrem Kern aus drei zentralen Hypothesen (Opp 2004): der Motivationshypothese, der Restriktionshypothese und der Hypothese der Nutzenmaximierung. Die Motivationshypothese besagt, dass Menschen gemäß ihren Präferenzen (Zielen, Wünschen, Motiven) handeln, die Restriktionshypothese, dass bei der Verwirklichung der Handlungsziele bestimmte Handlungsbeschränkungen eine Rolle spielen, und die Hypothese der Nutzenmaximierung besagt, dass Individuen solche Handlungen ausführen, die ihre Ziele im höchsten Maße realisieren (Opp 2004: 44 f.). Die neoklassische Handlungstheorie ähnelt dabei soziologischen Rational-Choice-(RC-)Ansätzen. Ihnen ist gemeinsam, dass sie Handlungserklärungen allein auf Basis der Betrachtung der Akteure selbst leisten (methodologischer Individualismus). Zwischen ökonomischen und soziologischen RC-Theorien bestehen tendenziell einige Unterschiede (vgl. Tabelle 1).

Tabelle 1: Gegenüberstellung: ökonomische und soziologische RC-Theorie

Ökonomische RC-Theorie	Soziologische RC-Theorie
Nur egoistische Präferenzen erklären eine Handlung	Alle Arten von Präferenzen kommen zur Erklärung in Betracht
Nur materielle Restriktionen sind handlungswirksam	Alle Arten von Restriktionen können Handeln beeinflussen
Menschen sind vollständig informiert	Menschen sind nicht vollständig informiert
Nur objektive Restriktionen erklären eine Handlung	Sowohl subjektive als auch objektive Restriktionen sind handlungserklärend
Nur Restriktionen erklären eine Handlung	Präferenzen und/oder Restriktionen erklären eine Handlung

Tabelle nach Opp (2004: 46); leicht verändert.

In der neoklassischen Ökonomik wird – etwa im Gegensatz zur soziologischen Rational-Choice-Theorie – typischerweise von sehr starken Rationalitätsannahmen (Goldthorpe 1998) ausgegangen: Individuen sind perfekt über die Rahmenbedingungen ihres Handelns informiert (3), kennen alle Alternativen und können den Nutzen, den sie aus den verschiedenen Handlungsalternativen erwarten, nicht nur exakt und schon im Voraus berechnen, sondern auch die Wahrscheinlichkeit des Eintritts erwünschter und unerwünschter Handlungsfolgen. Das rational handelnde Individuum „has reconciled all his conflicting partial values and synthesized them into a single utility function that orders, by his preferences for them, all these future states of the world“ (Simon 1983, zit. n. Goldthorpe 1998: 119). Ökonomisches Handeln ist im Rahmen dieser Theorietradition vor allem dadurch gekennzeichnet, dass Individuen nur egoistische Präferenzen (1) besitzen, also nur ihren eigenen Nutzen maximieren. Für die Wahl einer Handlung sind dabei ausschließlich materielle (2) und objektive (4) Restriktionen relevant. Materielle Restriktionen stellen vor allem die finanziellen Kosten und Nutzen der Handlungsalternativen dar, wobei soziale Restriktionen (Normen, kulturelle Wertvorstellungen) damit entweder ausgeschlossen werden oder in materielle Restriktionen (Strafen) umdefiniert werden müssen. Insofern Individuen perfekt über die Restriktionen informiert sind, können auch nur objektive Restriktionen als für die Handlungswahl entscheidend angenommen werden. Damit schließt die Theorie aus, dass sich Individuen aufgrund subjektiv als zutreffend empfundener, faktisch jedoch nicht vorliegender Sachverhalte für eine Handlung entscheiden (5).

Der Soziologe Raymond Boudon zeigt im Rahmen seines kognitivistischen Erklärungsmodells, dass Handlungsweisen, die aus der Perspektive der Nutzenmaximierung als paradox erscheinen, unter Bezug auf ein breiteres Handlungsmodell erklärbar sind. Als ein solches identifiziert er Max Webers Typologie des sozialen Handelns. Soziales Handeln ist dem Sinn nach auf das Handeln anderer bezogen. Im Gegensatz vor allem zu ökonomischen Handlungserklärungen basieren soziologische Erklärungen auf einem breiteren Verständnis dessen, was soziologisch als der Sinn einer Handlung bezeichnet wird (Schütz 1974). Dieser kann nach Max Weber einem von vier Idealtypen zugeordnet werden:

„Wie jedes Handeln kann auch das soziale Handeln bestimmt sein 1. zweckrational: durch Erwartungen des Verhaltens von Gegenständen der Außenwelt und von anderen Menschen und unter Benutzung dieser Erwartungen als Bedingungen oder als Mittel für rational, als Erfolg, erstrebte und abgewogene eigne Zwecke, – 2. wertrational: durch bewußten Glauben an den – ethischen, ästhetischen, religiösen oder wie immer sonst zu deutenden – unbedingten Eigenwert eines bestimmten Sichverhaltens rein als solchen und unabhängig vom Erfolg, – 3. affektiv, insbesondere emotional: durch aktuelle Affekte und Gefühlslagen, – 4. traditional: durch

eingelebte Gewohnheit" (Weber 1976: 12). Empirisch wird in Handlungssituationen nicht nur eine dieser Handlungsorientierungen eingenommen, sondern das konkrete Handeln ist stets einer Mischung aller vier Handlungsorientierungen unterworfen.

Auf die Frage: „Warum bestellt Herr X Wein zum Essen?“ gibt es demnach etwa folgende Antworten. Erstens: Er bestellt Wein, um seine Begleiter mit seiner Weinkennerschaft zu beeindrucken (zweckrational). Zweitens: Er bestellt Wein, weil er es richtig findet, die lokalen Winzer zu unterstützen (wertrational). Drittens: Herr X bestellt Wein, weil er Wein gerne trinkt (affektuell). Und schließlich viertens: Er bestellt Wein, weil er immer Wein zum Essen trinkt (traditional). Um entscheiden zu können, welche Erklärung im Einzelnen bei einer konkreten Handlung zutrifft, muss der gemeinte Sinn dieser Handlung deutend verstanden werden. In Boudons Terminologie bedeutet dies, die Gründe für das Handeln zu rekonstruieren (Boudon 1996: 126). „The Weberian motto ‚deutend verstehen‘ (...) says nothing else – that explaining a belief, an action, an attitude amounts to finding out the meaning to the social actor of the given belief, action or attitude“ (Boudon 1996: 126; vgl. auch Boudon 1997: 9).

Im Anschluss an Weber leitet Boudon daraus folgendes Vorgehen zur Erklärung von Handlungen ab. Zunächst sollte eine Erklärung davon ausgehen, dass das zu erklärende Handeln ein Mittel zu einem bestimmten Zweck darstellt. Es ist leicht zu erkennen, dass solche auf Zweckrationalität basierenden Erklärungen dem in der Ökonomie verbreiteten Schema der Erklärung eines Handelns durch individuelle Nutzenmaximierung entspricht (Ramb 1993; Opp 2004). So wird dann auch oben (vgl. Abschnitt 2.1) der Verbleib in Arbeitslosigkeit bei Geringqualifizierten auf diese Weise erklärt. Arbeitslose verfolgen mit diesem Handeln einen bestimmten Zweck. Dieser besteht darin, ein möglichst hohes Einkommen (Nutzen) bei möglichst geringer Anstrengung (Kosten) zu erzielen. Da die geringqualifizierten Arbeitslosen durch Arbeit weniger verdienen, als sie vom Arbeitsamt an Lohnersatzleistungen erhalten, entscheiden sie sich gegen die Annahme einer Arbeit.

Nicht alle Handlungen können auf eine solche instrumentelle Handlungsorientierung zurückgeführt werden. Das gilt auch dann, wenn man wie die soziologische Rational-Choice-(RC-)Erklärung in einer Erweiterung des neoklassischen Handlungsmodells neben der Nutzenmaximierung andere Ziele zulässt. Boudon deckt das Erklärungsdefizit der RC-Erklärung bzw. neoklassischer Theorien anhand einiger Beispiele auf, die er als Paradoxien bezeichnet und die eine ähnliche Struktur aufweisen wie die hier zur Erklärung stehende „Kombilohn-Paradoxie“.

So beschreibt er unter Berufung auf C. Wright Mills (1951) die Konflikte zwischen den weiblichen Angestellten eines taylorisierten Betriebes. Diese entzündeten sich häufig um Kleinigkeiten, z. B. die Nähe zu einer Licht- oder Wärmequelle. Sie weisen aber im Vergleich zur Geringfügigkeit des Streitgegenstandes ein hohes

Maß an Intensität und sogar Gewalttätigkeit auf. Für Boudon kann eine zweckrationale oder nutzenmaximierende Handlungsorientierung die Disproportionalität der Handlung als Mittel für die Beseitigung dieser kleinen Vorteile einiger gegenüber anderen Angestellten nicht erklären. Aus dieser Sicht käme ein RC-Modell zum Schluss, dass es für das beschriebene Verhalten keine Erklärung gibt, weil es vollkommen irrational ist (Boudon 1996: 144 f.).

Ein anderes Beispiel für Erklärungsdefizite der Beschränkung auf Zweckrationalität sind Reaktionen auf Bagatelldelikte wie den Diebstahl, auch wenn es sich beim Dieb um einen armen und bedürftigen Menschen handelt. Eigentlich hat Diebstahl einige Vorteile, für den Dieb und für die Gesellschaft insgesamt. Arme, die sich bestimmte Güter auf legalem Wege nicht leisten können, gelangen über diesen Umweg trotzdem daran. Außerdem entstehen durch die Möglichkeit des Diebstahls viele Arbeitsplätze in der Sicherheitsbranche. Trotzdem wird auch der kleinste Diebstahl von der Gesellschaft bestraft. Auch hier ist überdies die Reaktion des Opfers und der Gerichte oft in keiner Weise proportional zum Wert des gestohlenen Gegenstandes (Boudon 1997: 20 f.; Greve 2003: 628).

Schließlich lässt sich die Unvollständigkeit zweckrationaler Erklärungen auch an dem „Wahl-Paradox“ aufzeigen (Boudon 1998b: 198). Einerseits dürfte den Wählern aus zweckrationaler Perspektive klar sein, dass ihre eigene Stimme keinen Einfluss auf den Ausgang der Wahl hat und damit der Nutzen des Wahlgangs sehr gering ist. Dagegen verursacht der Wahlgang zum Teil sogar Kosten, z. B. die Opportunitätskosten eines Wahlganges am Sonntag. Trotzdem geht immer noch die Mehrheit der Bevölkerung in demokratischen Gesellschaften zur Wahl.

Da aufgrund solcher Paradoxien die Erklärungskraft auf Zweckrationalität beschränkter Handlungsmodelle eingeschränkt erscheint, es sei denn, es werden weite Bereiche menschlichen Handelns als irrational betrachtet, wurde Webers Typologie der Handlungsorientierungen von Boudon zu einem kognitivistischen Erklärungsmodell ausgearbeitet (Boudon 1996, 1998a; Greve 2003). Dabei sollte bis zum Beleg des Gegenteils davon ausgegangen werden, dass Akteure rational handeln.<sup>11</sup> Rationalität im Sinne Boudons bedeutet, die Akteure haben „strong reasons of believing what they believe, of doing what they do, and so forth“ (Boudon 1998b: 190).

Kernpunkte dieses Handlungsmodells ist die Erweiterung der Rational-Choice-Theorie um wertrationales oder in Boudons Terminologie axiologisches Handeln und um den kognitiven Aspekt des Handelns. „Actions can be meaningful in social actors because they are grounded on instrumental reasons (instrumental rationality). But they can also be meaningful in social actors because they are grounded

---

11 Dadurch konzentriert sich Boudons Handlungsmodell auf die Zweck- und die Wertrationalität. Erklärungen auf Basis affektuellen und traditionellen Handelns erfolgt erst dann, wenn keine rationale Erklärung zu finden ist.

on axiological reasons („axiological rationality’)" (Boudon 1997: 15). Manches Handeln kann durch den Bezug auf die Kosten-Nutzenerwägungen des Akteurs erklärt werden. Anderes Handeln ist als wertrational zu bezeichnen und kann dagegen nur durch die normativen Überzeugungen des Handelnden erklärt werden. X tat dies, weil er glaubte, dass dies zu tun fair, gut, richtig etc. ist. Eine zusätzliche Erweiterung des kognitivistischen Handlungsmodells gegenüber vielen RC-Theorien ist, dass soziales Handeln auch durch Bezug auf dessen kognitive Gründe erklärt wird. X tat dies, weil er *glaubte*, dass der Sachverhalt Y vorlag, und gute Gründe dafür hatte, dies zu glauben (Boudon 1998b: 190 f.; 1998a: 825 f.).

Vor allem durch die Integration der Wertrationalität in das Handlungsmodell lassen sich die oben beschriebenen Paradoxien leicht auflösen. Wertrationalität darf allerdings nicht mit einer durch die Internalisierung von Normen und Werten bedingtes normkonformes Handeln verwechselt werden (Boudon 1997: 4). Auch für die Orientierung ihres Handelns an Werten und Normen haben die Akteure gute Gründe und handeln daher rational. Diese erschöpfen sich auch nicht darin, aus instrumentell-rationalen Gründen nur scheinbar wertrational zu handeln, etwa um seinen guten Ruf nicht zu gefährden oder um unerwünschten Sanktionen zu entgehen (Boudon 1998b: 88). Auch wenn solches nur scheinbar wertrationales, tatsächlich aber instrumentelles Handeln ebenfalls auftritt: Axiologisches Handeln kennzeichnet, dass es dem Sinn nach im Glauben an die Richtigkeit der Handlung im Sinne einer moralischen oder normativen Wertung erfolgt, unabhängig von den unmittelbaren Konsequenzen einer Handlung für die Kosten-Nutzenbilanz des Handelnden selbst.

So können die gewalttätigen Konflikte der Angestellten im Überreaktions-Paradox oben dadurch erklärt werden, dass diese ihr Handeln an einer Gleichheitsnorm (Boudon 1998b) ausrichten. In tayloristischen Betrieben sind die Arbeitsabläufe streng festgelegt. Die Angestellten erfüllen alle dieselbe Aufgabe. Daher ist für die Akteure eine Abweichung von Leistung und Belohnung bei anderen Angestellten sofort erkennbar und wird als eine ungerechte Bevorteilung empfunden. „From an utilitarian viewpoint, it matters little to sit closer to the window. But, as soon as this advantage to the benefit of X results from a decision of the supervisor, it is perceived by Y, not as a disadvantage to himself, but as an injustice: I am here to get reward from my contribution; I am in a situation of social contract with the firm; any unjustified advantage in favour of X, however minor, is a violation of this basic contract and consequently morally intolerable, even if materially of weak relevance" (Boudon 1996: 144 f.).

Ganz ähnlich lässt sich die Bestrafung auch kleiner Diebstähle erklären. Während eine an den Kosten und Nutzen des Stehlens orientierte Erklärung keine überzeugende Antwort auf die gesellschaftliche Ächtung des Diebstahls findet, formuliert Boudon eine Erklärung auf Basis einer normativen Wertung. „Social order is

based on an adequation between retribution and contribution. (...) Now, stealing is a typical violation of these basic principles of social organization, since the thief attributes to himself unilaterally a retribution without offering any contribution as a counterpart" (Boudon 1997: 21; vgl. 1998a: 21 f.).

Auch das „Wahl-Paradox“ kann durch ein auf Wertrationalität beruhendes Verstehen des Handlungssinnes erklärt werden. Eine als gut betrachtete Regierung sollte die Interessen des Volkes verfolgen, nicht die eigenen. Durch Wahlen in einem demokratischen politischen System wird dies am besten gewährleistet. Dadurch, dass Politiker dem Risiko ausgesetzt sind, bei der nächsten Wahl nicht wiedergewählt zu werden, vermindert sich die Wahrscheinlichkeit für Korruption und die Vernachlässigung der Interessen des Volkes (Boudon 1997: 24 f.).

Neben der Erklärung von Handlungen auf Basis axiologischer und instrumenteller Rationalität führt Boudon noch Erklärungen auf Basis kognitiver Rationalität an (Boudon 1998a: 824; 1998b: 182 ff.; 1996: 125 ff.). „In some circumstances the core of some action is constituted by ‚cognitive‘ reasons: He did X because he believed Z is likely to be true, and he had strong reasons for believing so“ (Boudon 1998b: 191). Boudon führt drei Typen von „beliefs“ (Überzeugungen) auf (Boudon 1996: 125). Der erste Typ (Typ I) von Überzeugungen ist empirisch validierbar und wahr, z. B. jemand ist der Überzeugung: Delfine sind Säugetiere. Der zweite Typ (Typ II) ist ebenfalls empirisch validierbar, jedoch unwahr. Eine solche Überzeugung ist etwa, wenn jemand glaubt, Delfine seien Fische. Schließlich ist ein dritter Typ (Typ III) von Überzeugungen nicht validierbar. Hierunter fallen normative oder axiologische Überzeugungen.

Aufgrund der Komplexität sozialer Zusammenhänge kommt es nicht selten zu Handlungen, die sich auf empirisch nicht unbedingt zutreffende Gründe stützen. Trotzdem ist dies nicht gleichbedeutend mit Irrationalität. „Even though these reasons appear to us as false, they can be perceived as right and strong by the actors themselves“ (Boudon 1998b: 168). Boudon hebt hervor, dass es sich dabei nicht nur um die Überzeugungen einzelner, sondern um kollektive Überzeugungen des Typs II handeln kann. Als Beispiel nennt er ein psychologisches Experiment, in dem Ärzten ein Test beschrieben wird, der bei einer Krankheit, die bei einer von tausend Personen auftritt, mit einer Wahrscheinlichkeit von 5 % fälschlich die Krankheit diagnostiziert. Die Ärzte wurden nun gebeten, die Wahrscheinlichkeit dafür anzugeben, dass bei einer Person, bei der der Test angeschlagen hat, die Krankheit tatsächlich vorliegt. Die Mehrheit der Ärzte (56 %) gibt hierauf an, dass die Wahrscheinlichkeit für eine tatsächliche Erkrankung bei 95 % liegt, die korrekte Antwort lautet jedoch 2 % (zu den Details siehe Boudon 1996: 130).

Zusammengefasst, erweitert Boudon den Rationalitätsbegriff von der Fixierung auf die Maximierung des individuellen Nutzens auf die Befolgung sozialer

Normen einerseits und auf die Handlungsrelevanz subjektiven Für-wahr-Haltens. Eine Erklärung des Kombilohnparadoxes sollte daher, Boudon folgend, nicht nur die Wirkung des Lohnzuschusses auf das individuelle Nutzenkalkül berücksichtigen, sondern auch danach fragen, welche normativen Handlungsorientierungen Beschäftigte aufweisen und wie sie die Tatsache der zusätzlichen finanziellen Förderung subjektiv interpretieren. Vor allem in der neueren Ökonomischen Theorie, aber auch in Soziologie und Sozialpsychologie spielt der Begriff der Reziprozität genau diese Rolle. Er verweist zum einen auf Abweichungen vom Prinzip der Nutzenmaximierung auch und gerade dann, wenn es um die Frage von finanziellem Zugewinn oder Verlust geht. Zum anderen verweist er darauf, dass die Anreizwirkung des Lohnes nicht nur von seiner objektiven Höhe, sondern auch von subjektiven Interpretationen abhängig ist.

### 3.2 Die Reziprozitätsnorm in Beschäftigungsverhältnissen

Reziprozität (Gegenseitigkeit) bezeichnet die „Logik des Gebens, Nehmens und Erwiderns“ (Adloff/Mau 2005: 9). Wie Adloff und Mau feststellen, ist „Reziprozität kein einheitliches Konzept (...), sondern in verschiedenen Theorietraditionen aufgegriffen und gewendet“ worden (ebd. 11). Dies gilt umso mehr, wenn man sich nicht nur mit soziologischen (z. B. Voswinkel 2005), sondern auch sozialpsychologischen (Rousseau 1989) und ökonomischen Verwendungsweisen (Falk 2003) des Begriffes auseinandersetzt. Trotz der Unterschiede, die sich innerhalb und zwischen den Disziplinen beobachten lassen, wird in dieser Arbeit davon ausgegangen, dass es sich dabei lediglich um unterschiedliche Perspektiven auf denselben Gegenstand handelt. In diesem Abschnitt wird zunächst das Reziprozitätsprinzip als eine allgemeine Handlungsorientierung in Tauschbeziehungen vorgestellt (Abschnitt 3.2.1). Dabei wird neben den Grundlagen des Reziprozitätsbegriffes in der anthropologischen Forschung vor allem auf aktuelle Erkenntnisse aus der Experimentellen Wirtschaftsforschung eingegangen. Anschließend (Abschnitt 3.2.2) wird die Unvollständigkeit des Arbeitsvertrages als Grund für die Relevanz der Reziprozitätsnorm auch in Beschäftigungsverhältnissen aufgezeigt. Abschnitt 3.2.3 zeigt, dass die Dauerhaftigkeit des Beschäftigungsverhältnisses gefährdet wird, wenn Betriebe die Reziprozitätsnorm gegenüber dem Beschäftigten verletzen. In Abschnitt 3.2.4 wird herausgearbeitet, dass reziprokes Handeln nicht nur normatives Handeln darstellt, sondern auch einen kognitiven Aspekt aufweist. Dieser besteht darin, dass die Verletzung der Reziprozitätsnorm für den Beschäftigten nicht auf einer objektiven Gegenüberstellung von Leistung und Gegenleistung ersichtlich wird, sondern dass dabei ein hohes Maß an subjektiver Interpretation bzw. kognitiven Überzeugungen des Typs II im Sinne Boudons (vgl. Abschnitt 3.1) besteht.

### 3.2.1 Das Reziprozitätsprinzip als Handlungsorientierung in Tauschbeziehungen

Seinen Ursprung hat der Begriff der Reziprozität in anthropologischen und ethnologischen Untersuchungen des Tauschhandels (Mauss 1990; Firth 1959; Malinowski 2001; Sahlins 1999). Das Reziprozitätsprinzip oder die Reziprozitätsnorm „umfasst eine flexible Anzahl von Instruktionen für das Handeln, die besagen oder zumindest implizieren, dass denen, die helfen, geholfen werden soll, dass die, die helfen, nicht geschädigt werden sollen und dass denen, die schädigen, nicht geholfen werden soll“ (Gouldner 2005: 119). Mit dem Reziprozitätsprinzip wird damit eine Grundstruktur des Gabentausches in archaischen Gesellschaften bezeichnet, die maßgeblich zur Aufnahme und Stabilisierung sozialer Beziehungen beiträgt. Im Anschluss daran sieht Karl Polanyi (2002) archaische und vormoderne Gesellschaften vom Prinzip der Reziprozität und Redistribution gekennzeichnet, während er der modernen Gesellschaft das marktwirtschaftliche Prinzip des Tausches zuschreibt. Polanyi geht davon aus, dass die Produktion und Distribution von Gütern in archaischen Gesellschaften nicht wie im modernen Tauschhandel auf Gewinnstreben beruht, sondern anderen Prinzipien folgt. So beruht der „Kula“-Handel auf den kreisförmig angeordneten Trobrianderinseln stattdessen auf der Reziprozität sozialen Verhaltens. Dabei werden Armbänder aus weißen Muscheln in die eine Richtung der Inselgruppe gehandelt und Halsketten mit roten Muscheln in die andere. „Wir sprechen in diesem Zusammenhang von Handel, obwohl dabei kein Profit erzielt wird, weder in Form von Geld noch in Form von Gütern (...) Weiter handelt es sich beim Kula in der Regel um einzelne Tauschpartner, die die jeweiligen Kula-Gaben mit gleich wertvollen Armbändern und Ketten erwidern“ (Polanyi 2002: 80). Vom symmetrischen Reziprozitätsprinzip unterscheidet Polanyi das auf Zentrität beruhende Redistributionsprinzip. Hierbei werden die erwirtschafteten Güter zentral gesammelt, z. B. bei den Stammeshäuptlingen oder in komplexeren Gesellschaften wie dem ägyptischen Neuen Reich in zentral verwalteten Lagerhäusern der Pharaonen, und von dort aus verteilt. Redistribution kann auf vollkommen freiwilliger Basis stattfinden oder in das bestehende Herrschaftssystem eingebunden sein (ebd. 83 f.). Erst im 19. Jahrhundert sieht Polanyi die Prinzipien der Reziprozität und Redistribution vom Prinzip des Tauschhandels abgelöst, der auf der Existenz von Märkten beruht (ebd. 87).

Während Polanyi die Entstehung des marktförmig organisierten Tauschhandels als eine Ablösung vormoderner Prinzipien des Handelns beschreibt, zeigt die aktuelle Forschung, dass auch in modernen Marktgesellschaften Gegenseitigkeitsbeziehungen eine zentrale Rolle spielen (Adloff/Mau 2005: 18). Das Reziprozitätsprinzip wird zum einen als Alternative zur Teilnahme am Marktgeschehen betrachtet, so etwa in Form von Tauschringen (Hinz/Wagner 2006) oder in Form privater Hilfeleistungen (z. B. Nelson 2000). Zum anderen, und hier von höherer Bedeutung, wird das



Reziprozitätsprinzip als eine gleichzeitige Ergänzung des Marktprinzips betrachtet. Sowohl in der Soziologie als auch in der Sozialpsychologie und der Ökonomie beschreibt der Reziprozitätsbegriff das Wirken sozialer Erwartungen der Gegenseitigkeit auch in ökonomischen Transaktionen. Die Verwendung des Begriffes erfolgt jedoch innerhalb und erst recht zwischen den sozialwissenschaftlichen Disziplinen mit unterschiedlichen Schwerpunktsetzungen.<sup>12</sup>

In der Soziologie gibt es mit der makrosoziologisch-kollektivistisch und der mikrosoziologisch-individualistisch orientierten Theorietradition zwei Ansatzpunkte für die Einordnung des Reziprozitätsbegriffes. Im Bereich makrosoziologischer Theorien kann Reziprozität als synonym mit dem Begriff der Solidarität betrachtet werden (Adloff 2005: 43 f.). „Es scheint (...) grundsätzlich naheliegend, Solidarität auch in ihrer modernen Form in terms einer mehr oder weniger weiter gefassten Reziprozität zu fassen (Bode/Brose 1999: 181, Herv. weggelassen; vgl. auch Adloff/Mau 2005: 20 ff.). Mikrosoziologisch (vgl. Adloff/Mau 2005: 26 ff.) findet der Reziprozitätsbegriff Anschluss an Theorien des sozialen Austausches (Blau 2005; Homans 1961). Solche Theorien gehen davon aus, dass soziale Beziehungen analog zu wirtschaftlichen Transaktionen als eine strategische Beziehung des wechselseitigen Tausches verstanden werden können. Menschen gehen soziale Kontakte ein, weil sie sich davon einen Vorteil versprechen. „Will man Vorteile aus einem Kontakt ziehen, so ist man verpflichtet, sich zu revanchieren und den anderen im Gegenzug Vorteile zu gewähren. (...) Wer der Abgeltung solcher Verpflichtungen nicht nachkommt und sich für erwiesene Vorteile nicht revanchiert, beraubt den anderen des Anreizes, die Freundlichkeit ihm gegenüber fortzusetzen“ (Blau 2005: 26).

Aus austauschtheoretischer Perspektive grenzen Molm, Peterson und Takahashi (2001) den reziproken Austausch (reciprocal exchange) von einem ausgehandelten Austausch (negotiated exchange) ab. Im ausgehandelten Austausch gilt: „actors engage in a joint decision process, such as explicit bargaining, in which they reach an agreement on the terms of the exchange“ (ebd. 161). Der Austausch kommt erst zustande, wenn beide Akteure für sich in den Tauschbedingungen einen Vorteil sehen. Dagegen ist reziproker Austausch dadurch gekennzeichnet, dass einer der Akteure zum Vorteil des anderen handelt, ohne dass irgendwelche Gegenleistungen spezifiziert wären. Ob der Empfänger irgendwann eine solche Gegenleistung erbringt, bleibt unklar und zeigt sich erst im Lauf der Zeit.

Obwohl die Soziologie schon sehr früh den Begriff der Reziprozität aus der anthropologischen und ethnologischen Forschung in die soziologische Analyse

12 Der Reziprozitätsbegriff spielt in unterschiedlichen Forschungsgebieten eine Rolle, so z. B. auch in der Wohlfahrtsstaatsforschung (Lessenich/Mau 2005; Bowles/Gintis 2000) und in Generationenbeziehungen (Hollstein 2005). Der folgende Überblick beschränkt sich allerdings auf Verwendungsweisen im Rahmen marktlicher Transaktionen, vor allem im Rahmen von Beschäftigungsverhältnissen.

moderner Austauschbeziehungen übernommen hat, hat er eine nur geringe Prominenz in der soziologischen Theoriebildung erlangt (Adloff/Mau 2005: 9), wenn er auch aktuell wieder an Bedeutung gewinnt. Gegenwärtig trägt vor allem die Experimentelle Wirtschaftsforschung zur Weiterentwicklung des Konzeptes bei. Im Rahmen austauschtheoretischer Überlegungen macht der Begriff der Reziprozität das Tauschmotiv auch in nicht marktlichen Transaktionen sichtbar. Dagegen ist seine Verwendung in der Ökonomischen Theorie (z. B. Akerlof 1982) umgekehrt als Reaktion auf die geringe Erklärungskraft der neoklassischen Handlungstheorie für bestimmte Handlungen zu verstehen, etwa wenn Betriebe einerseits mehr als den Mindestlohn für eine Arbeitsleistung zahlen und Beschäftigte andererseits mehr leisten, als vom jeweiligen Betrieb als Mindestanforderung festgelegt wird, oder wenn Betriebe einerseits ihre Beschäftigten in wirtschaftlich schwierigen Situationen nicht entlassen und Beschäftigte im Gegenzug dazu bereit sind, unbezahlte Mehrarbeit zu leisten.

Vor allem die Experimentelle Wirtschaftsforschung konnte auf Basis des Reziprozitätsbegriffes einen Beitrag zur Frage liefern, warum Akteure auch im Rahmen marktlicher Austauschprozesse offensichtlich vom Prinzip der Nutzenmaximierung abweichen. Ein typisches Experiment, das marktliche Austauschprozesse simuliert, ist in diesem Zusammenhang das Ultimatumspiel (Güth/Schmittberger/Schwarze 1982; Fehr/Schmidt 1999). Dort verhandeln zwei Spieler um einen Geldbetrag, z. B. um 10 Euro, wobei der Spieler A einen Vorschlag zur Aufteilung des Betrages macht, den Spieler B dann entweder annehmen oder ablehnen kann. Bei Annahme des Angebotes durch Spieler B wird das Geld nach der vorgeschlagenen Aufteilung an beide Spieler ausgezahlt, lehnt Spieler B das Angebot ab, erhält keiner der Spieler Geld. Aus der Handlungstheorie der Nutzenmaximierung lassen sich für dieses Design spezifische Prognosen über das Handeln von Spieler A und Spieler B ableiten. Den höchsten Nutzen wird Spieler A dann erzielen, wenn er Spieler B gerade den Mindestbetrag anbietet, also (wenn nur „runde“ Beträge erlaubt sind) 1 Euro. Dieser wird dann, unter den gegebenen Restriktionen entweder 1 Euro oder nichts zu erhalten, seinen Nutzen maximieren und dem Angebot zustimmen.

Die empirischen Ergebnisse des Ultimatumspiels widersprechen diesen Prognosen allerdings regelmäßig und systematisch (Falk/Fischbacher 2006). Einerseits bewegen sich die Angebote der vorschlagenden Spieler bei 40–50 % des Betrages und sind damit nicht nur weit höher als ein Mindestangebot, sondern nahe an einer 50/50-Aufteilung. Andererseits werden Angebote von unter 30 % meist von Spieler B abgelehnt, sodass keiner der beiden etwas erhält. Die ablehnenden Spieler nehmen also einen persönlichen Verlust in Kauf. Diese Abweichungen vom Handlungsmodell des rational kalkulierenden Nutzenmaximierers sind mithilfe einer Orientierung der Spieler nicht (nur) an der Nutzenmaximierung, sondern auch an

Fairness- oder Reziprozitätsnormen erklärbar. „Reciprocity means that in response to friendly actions, people are frequently much nicer and much more cooperative than predicted by the self-interest model; conversely, in response to hostile actions they are frequently much more nasty and even brutal (Fehr/Gächter 2000: 159). Ähnlich drückt es auch der Soziologe Gouldner aus: „We may suggest that each side lives up to its obligations, not simply because of constraints imposed by the division of labor with its attendant mutual dependency, but also because the partners share the higher level *moral norm*: You *should* give benefits to those who give you benefit“ (Gouldner 1960: 170; Herv. im Orig.).<sup>13</sup>

Sobel (2005) grenzt hier zunächst intrinsische Reziprozität von instrumenteller Reziprozität ab. Das Prinzip der instrumentellen Reziprozität erklärt kurzfristige Abweichungen von der Nutzenmaximierung auf Basis eines langfristigen Kalküls. So könnte ein Betrieb deshalb nicht zu Entlassungen greifen, weil er befürchten muss, von den übrigen Beschäftigten abgestraft zu werden oder auf dem Arbeitsmarkt als potenzieller Arbeitgeber einen schlechten Ruf zu erhalten. Die Ergebnisse des Ultimatumspiels verweisen allerdings auf eine intrinsische Reziprozität (Sobel 2005: 193), da innerhalb des Spiels keine wiederholten Interaktionen der Spieler stattfinden, es handelt sich damit um einmalige Austauschbeziehungen, in denen sich trotzdem reziprokes Verhalten zeigt. Intrinsische Reziprozität ist der Nutzenmaximierung tatsächlich entgegengestellt, während instrumentelle Reziprozität lediglich ein Mittel der langfristigen Nutzenmaximierung darstellt.

Reziprokes Verhalten lässt sich weiterhin danach unterscheiden, ob positives Verhalten belohnt oder negatives bestraft wird (Sobel 2005: 397; Fehr/Gächter 2000). Ersteres kann als positive oder konstruktive Reziprozität, Letzteres als negative<sup>14</sup> oder destruktive Reziprozität bezeichnet werden. Das Ultimatumspiel zeigt einen Fall der destruktiven Reziprozität, da hier Spieler B als negativ empfundene Angebote des Spielers A bestrafen und dabei auch eigene Verluste hinnehmen. Auch konstruktive Reziprozität lässt sich in den Experimenten der Ökonomie nachweisen. In einem sogenannten „Vertrauensspiel“ erhält Spieler A einen bestimmten Geldbetrag  $x$  und kann dann einen beliebigen Anteil von null bis 100 % an Spieler B geben. Dieser Betrag  $y$  wird dann vom Leiter des Experimentes verdreifacht ( $3y$ ), sodass Spieler B tatsächlich  $3y$  erhält. Spieler B kann wiederum frei entscheiden, ob bzw. wie viel er nun an Spieler A sendet. Im Sinne der positiven Reziprozität ergibt sich in solchen Experimenten, dass sowohl Spieler A einen von null verschiedenen

<sup>13</sup> Gouldner kommt zu dieser der neueren ökonomischen Forschung entsprechenden Erkenntnis allerdings auf der Basis einer anderen theoriegeleiteten Perspektive.

<sup>14</sup> Da der Begriff negative Reziprozität weiter unten mit einem anderen Inhalt besetzt ist, wird zur Vermeidung begrifflicher Unklarheiten im Folgenden von destruktiver und konstruktiver Reziprozität gesprochen, auch wenn die Begriffe positiv und negativ gebräuchlicher sind.

Betrag an Spieler B sendet als auch Spieler B einen von null verschiedenen Betrag zurücksendet. Dieser ist dann positiv mit dem Ursprungsbetrag  $y$  korreliert (Fehr/Gächter 2000: 162). Sowohl die Ergebnisse des Ultimatumspiels als auch die des Vertrauensspiels sind nicht auf die Situation mit kleinen Geldbeträgen beschränkt, sondern bleiben auch dann bestehen, wenn es um sehr hohe Beträge geht (ebd.).

Der Reziprozitätsbegriff läuft oft Gefahr, mit dem Begriff des Altruismus vermischt zu werden (z. B. Ockenfels 1999: 15 ff.). Altruistisches Verhalten ist jedoch unabhängig von einer Gegenleistung des Gegenübers. Dies spiegelt sich in einem weiteren Experiment wider, dem Diktatorspiel (Eckel/Grossman 1996). Dort erhält ein Spieler A einen Betrag, den er nach Belieben mit einem Spieler B teilen kann, wobei er auch den gesamten Betrag für sich einbehalten kann. Spieler B hat hier, im Gegensatz zum Ultimatumspiel, kein Vetorecht. Damit besteht für Spieler A im Diktatorspiel kein Risiko, dass eventuell unfaire Offerten abgelehnt werden. Es zeigt sich jedoch, dass auch dann die Spieler A – in je nach konkreter Versuchsanordnung unterschiedlichem Ausmaß – bereit sind, einen wesentlichen Teil des Betrages an die Spieler B zu geben. Dieses Ergebnis zeigt sich auch dann, wenn die Entscheidung von Spieler A gegenüber B und dem Experimentator per Versuchsanordnung unbekannt bleibt, sodass A keine sozialen Sanktionen zu fürchten hat.

Ebenfalls mit dem Reziprozitätsbegriff wird vor allem in der Ökonomie der Begriff der Ungleichheitsaversion in Verbindung gebracht. Als Abneigung gegenüber Ungleichheiten wird eine Einstellung ökonomischer Akteure bezeichnet, die ausschließlich am Ergebnis des Austauschprozesses orientiert ist, also daran, ob das Ergebnis eines Austauschprozesses fair ist. Demgegenüber schließt der Begriff der Reziprozität Fairnesserwägungen mit ein, geht aber darüber hinaus (Falk/Fischbacher 2000). Der Reziprozitätsbegriff stellt nicht nur die faktische Ausgewogenheit der Ergebnisse im Tauschhandeln in den Mittelpunkt, sondern vor allem auch die mit dem Handeln verbundene Intention. Das heißt, dass die Belohnung oder Bestrafung einer Tauschhandlung nicht nur vom Umfang der getauschten Güter abhängt, sondern auch und vor allem von der Intention des Gegenübers, mit der er seinem Tauschpartner entgegentritt. Leichte Variationen des Ultimatumspiels zeigen dies sehr deutlich (Falk/Fischbacher 2006, 2000: 5 f.). In Variante 1 soll Spieler A wieder 10 Euro auf sich und sein Gegenüber aufteilen. Diesmal darf er aber nur zwischen zwei Aufteilungsmodi wählen. Entweder kann er 2 Euro abgeben und behält 8 Euro (8/2), oder er kann die Aufteilung 5/5 wählen. Wiederum kann Spieler B die angebotene Aufteilung akzeptieren oder ablehnen, wodurch keiner der beiden etwas erhält. Variante 2 entspricht Spielvariante 1, allerdings sind hier die beiden möglichen Aufteilungen 8/2 und 10/0. Zentrales Ergebnis des Experimentes ist, dass sich die Ablehnungsraten der Aufteilung 8/2 in beiden Spielvarianten signifikant unterscheiden. Obwohl das Ergebnis des angebotenen Tauschhandels dasselbe ist, wird das

Aufteilungsangebot in Variante 1 zu 44,4 % abgelehnt, während in Variante 2 nur (aber immerhin) 8,9 % der Spieler B ablehnen. Der Unterschied besteht allein in den Intentionen der vorschlagenden Spieler. In Variante 1 hätte Spieler A eine faire Aufteilung wählen können, was von Spieler B negativ gedeutet wird. Demgegenüber ist in Variante 2 die Aufteilung 8/2 die „fairestmögliche“ Aufteilung. Dies wird von Spieler B bei der Interpretation des Verhaltens des Spielers A berücksichtigt.

Angesichts der fast verwirrenden Vielfalt von Verwendungsmöglichkeiten des Reziprozitätsbegriffes lässt sich fragen, ob und gegebenenfalls wie sich diese zueinander in Beziehung setzen und ordnen lassen. Ohne damit eine allumfassende Lösung zu beanspruchen, wird in dieser Arbeit davon ausgegangen, dass Austauschhandlungen in einem Kontinuum der Reziprozität stattfinden. Die verschiedenen Verwendungsweisen des Reziprozitätsbegriffes werden als eine Lokalisierung auf diesem Kontinuum interpretiert. Im Anschluss an Sahlins (1999) lässt sich ein solches Kontinuum der Reziprozität zwischen dem altruistischen Handeln einerseits und individuell-nutzenmaximierendem Handeln andererseits aufspannen. „Die Form des Austausches pendelt von selbstloser Rücksichtnahme auf andere durch Gegenseitigkeit bis hin zum Eigennutz“ (Sahlins 1999: 153; vgl. aus Sicht der Ökonomischen Theorie auch Ockenfels 1999: 103 ff.; Bethwaite/Tompkinson 1996).

Sahlins Kontinuumbegriff der Reziprozität reicht von der generalisierten über die ausgleichende zur negativen Reziprozität. Diese drei Typen unterscheiden sich in der Art der sozialen Beziehungen, in die sie eingebettet sind, in der Unmittelbarkeit der Erwidern einer Gabe und der Gleichwertigkeit von Gabe und Gegengabe. *Generalisierte Reziprozität* bezeichnet eine Solidarbeziehung, in der soziale Aspekte des Tausches den materiellen Aspekt überwiegen. Gaben werden innerhalb solcher Beziehungen freiwillig und scheinbar ohne Gegenleistung erbracht. Es handelt sich dabei jedoch nur in der Extremform der generalisierten Reziprozität um echten Altruismus, denn oft enthält eine Gabe, ein Geschenk oder eine Hilfeleistung für den Empfänger implizit die Verpflichtung zur Erwidern, etwa wenn der Geber selbst Hilfe braucht. Wann eine Erwidern erfolgt, bleibt dabei aber ebenso unbestimmt wie der Wert der Gegengabe. Zeitpunkt und Wert der Gegengabe sind zudem davon abhängig, was sich der Empfänger leisten kann, sodass vom Geber auch in Kauf genommen wird, dass die Gegengabe nie erfolgt, ohne dass die Solidarbeziehung dadurch aufgelöst würde. Beispiele für solche generalisierte Reziprozität sind etwa Hilfeleistungen unter Freunden und Verwandten oder Spenden an Hilfsorganisationen.

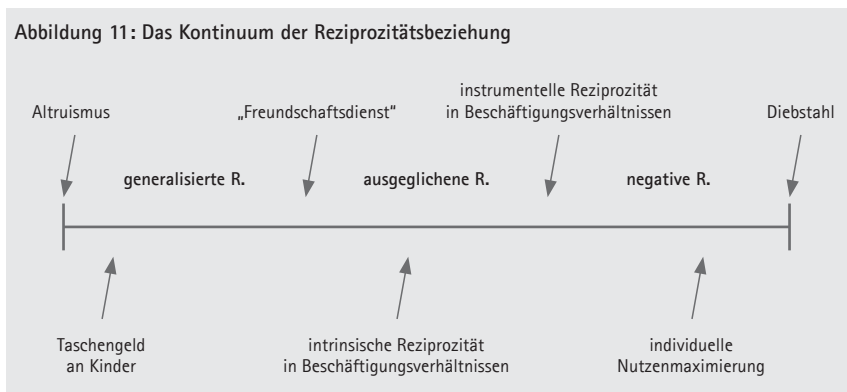
Den Gegenpol zur generalisierten Reziprozität stellt die *negative Reziprozität*<sup>15</sup> dar. Negativ reziprok ist ein Handeln, wenn versucht wird, etwas für umsonst und ungestraft zu bekommen. Ein solches Verhalten ist die unpersönlichste Art des Tau-

15 Vgl. Fußnote 14.

sches, wobei die Tauschparteien bestrebt sind, ihren Nutzen ohne Rücksicht auf die dadurch entstehenden Kosten der anderen zu maximieren. Beispiele dafür reichen vom Feilschen über den Betrug bis hin zum Diebstahl. Bei der negativen Reziprozität steht der materielle Aspekt eindeutig im Vordergrund, eine soziale Beziehung stiften negativ reziproke Tauschhandlungen nicht. Dem Bereich der negativen Solidarität lässt sich die instrumentelle Reziprozität zuordnen, aber auch die als typisch für ökonomische Transaktionen betrachtete individuelle Nutzenmaximierung.

Den Mittelpunkt zwischen negativer und generalisierter Reziprozität bildet die *ausgeglichene Reziprozität*. Diese bezeichnet eine Tauschbeziehung, in der für eine Gabe stets eine Gegengabe von entsprechendem Wert erwartet wird. Ausgeglichen-reziproke Handlungen sind stärker von persönlichen Beziehungen geprägt als die negative, aber weniger als die generalisierte Reziprozität. Die Tauschpartner verbinden hier vor allem deren komplementäre wirtschaftliche und soziale Interessen. Bei einem ausgeglichen-reziproken Austausch ist der materielle Aspekt ebenso wichtig wie der soziale. Daher folgt dem Ausbleiben der Gegengabe auch eine Auflösung der sozialen Beziehung zwischen Geber und Empfänger. Kennzeichen ausgeglichener Reziprozität ist also, dass beide Tauschpartner aus der Transaktion Nutzen ziehen.

Diese breite Definition macht zum einen deutlich, dass soziale Austauschbeziehungen meist weder auf rein egoistischen noch auf rein selbstlosen Motiven beruhen, sondern an unterschiedlichen Stellen dieses Kontinuums situiert sein können. Dies verdeutlicht die Abbildung 11 anhand einiger Beispiele.<sup>16</sup>



Die Extremform der generalisierten Reziprozität ist das altruistische Handeln, beispielsweise in Form von Spenden an Hilfsorganisationen. Austauschbeziehungen

<sup>16</sup> Vermutlich lässt sich über die genaue Einordnung der Beispiele auf dem Kontinuum streiten, zum Beispiel, ob familiäre Unterstützungsleistungen altruistisch motiviert und damit einseitig sind oder in den Bereich des reziproken Austauschs fallen. Die Relationen der Beispiele zueinander dürften jedoch weitgehend konsensfähig sein.

in Familien lassen sich relativ weit auf der Seite der generalisierten Reziprozität verorten. So ist das Zahlen von Taschengeld an die Kinder vonseiten der Eltern nur an generelle Erwartungen der Erwidern dieser Gabe geknüpft, und auch eine eventuelle Undankbarkeit der Kinder kann dabei die soziale Beziehung nicht auflösen. Immer noch als generalisierte Reziprozitätsbeziehung können gegenseitige Hilfeleistungen zwischen befreundeten Personen betrachtet werden, etwa die Hilfe bei Umzügen oder das Ausleihen des Autos etc. Dabei ist das Aufrechterhalten der sozialen Beziehung bei ausbleibenden Erwidern solcher Freundschaftsdienste weniger gesichert, als dies bei Familien meist der Fall ist.

Auf der anderen Seite ist Diebstahl der prototypische Fall der extremen negativen Reziprozität. Diebstahl ist in keine soziale Beziehung eingebettet, sondern nur an den materiellen Aspekten des „Tauschhandels“ orientiert. Bei individuell nutzenmaximierendem Handeln steht ebenfalls der materielle Aspekt im Vordergrund. Hierbei wird versucht, mit einer möglichst geringen Leistung eine möglichst große Gegenleistung zu erzielen. Da allerdings eine Gegenleistung erfolgt, kann individuelle Nutzenmaximierung zwischen negativer und ausgeglichener Reziprozität verortet werden. Das gilt noch mehr für die sogenannte instrumentelle Reziprozität, bei der Tauschpartner aus eigenem Interesse den Anschein erwecken wollen, sie handelten aufgrund ausgeglichener Reziprozität. Dies ist etwa der Fall, wenn Tauschpartner an wiederholten Interaktionen interessiert sind und vermeiden wollen, bei den aktuellen bzw. zukünftigen potenziellen Tauschpartnern in den Ruf zu kommen, andere zu übervorteilen. Intrinsische Reziprozität ist der prototypische Fall der ausgeglichenen Reziprozität. Die Tauschpartner orientieren sich, also ohne damit indirekt ihre eigenen Interessen zu verfolgen, um ihrer selbst willen an der Reziprozitätsnorm als Ausgeglichenheit von Leistung und Gegenleistung.

Die Lokalisierung bestimmten Tauschhandelns auf dem Reziprozitätskontinuum macht auch deutlich, dass sich soziale Beziehungen im Zeitverlauf auf diesem Kontinuum bewegen können, etwa wenn aus einer Geschäftsbeziehung eine Freundschaft entsteht oder kooperative Beziehungen zwischen Betrieb und Beschäftigten in Arbeitskämpfe umschlagen. Ebenso werden betriebliche Tauschbeziehungen in Beschäftigungsverhältnissen anfangs hauptsächlich dem Typ der ausgeglichenen Reziprozität zuzuordnen sein, sie beinhalten aber stets auch Elemente der Generalisierung, die im Laufe des Beschäftigungsverhältnisses durch den Aufbau von Vertrauen zunehmen.

Worauf die Orientierung am Reziprozitätsprinzip letztendlich beruht, ob auf einer genetischen Disposition, auf sozialen Lernvorgängen oder instrumentellen Erwägungen (vgl. Adloff/Mau 2005: 32 f.; Ockenfels 1999: 15 ff.), kann und muss im Folgenden unberücksichtigt bleiben.

### 3.2.2 Reziprozität in Beschäftigungsverhältnissen

Wenn von Reziprozität in Beschäftigungsverhältnissen die Rede ist, dann ist damit meist die ausgleichende Reziprozität gemeint. Weder wird ein Beschäftigter erwarten, den Arbeitslohn ohne Gegenleistung zu erhalten, noch wird er bereit sein, ohne Lohn zu arbeiten. Zum Teil wird bei der Analyse von Beschäftigungsverhältnissen allerdings auch von generalisierter Reziprozität gesprochen (z. B. Voswinkel 2005). Damit kann jedoch, das macht die hier verwendete Typologie Sahlins deutlich, lediglich eine Verschiebung in Richtung Generalisierung gemeint sein. Das Prinzip des Ausgleiches wird in Beschäftigungsverhältnissen nicht vollständig verlassen. Da es im Folgenden ausschließlich um die Reziprozität in Beschäftigungsverhältnissen geht, ist mit dem Begriff Reziprozität – außer es wird ausdrücklich darauf hingewiesen – stets die ausgeglichene Reziprozität gemeint.

In betrieblichen Beschäftigungsverhältnissen bildet die Unvollständigkeit des Arbeitsvertrages den Grund, warum diese auf Dauer nicht als rein eigeninteressierte (negativ reziproke) Handlungszusammenhänge aufrechtzuerhalten sind und der (ausgeglichenen) Reziprozität bedürfen (Bode/Brose 1999: 184). Die Unvollständigkeit oder Unbestimmtheit des Arbeitsvertrages (Hart/Moore 1999) bezeichnet die Tatsache, dass es im Gegensatz zu anderen Vertragsarten nicht möglich ist, die von beiden Seiten geschuldeten Leistungen genau zu spezifizieren. Dies ist auf die Unbestimmtheit zukünftiger Zustände und auf unvollständige Informationen zurückzuführen. Im Rahmen unvollständiger Verträge entsteht auf der Seite, die eine Leistung erbringen soll, gegenüber dem Vertragspartner ein Informationsvorsprung (asymmetrische Information), sowohl was die eigene Leistungsfähigkeit und Leistungsbereitschaft als auch was die Zurechnung des Ergebnisses betrifft. „Given the technological complexity of contemporary organizations, it is unlikely that a written contract could be drafted to cover all performance requirements of a person who will work for an employer for 30 years“ (Rousseau 1989: 123).

Dies führt zu Unbestimmtheiten bei den Vertragspartnern, ob auf eine Leistungserbringung eine Gegenleistung entsprechenden Ausmaßes erfolgt. Solche Unbestimmtheiten treten in Beschäftigungsverhältnissen in drei Dimensionen auf, in der zeitlichen, der sozialen und sachlichen Dimension (Voswinkel 2005). Sind Arbeitsverhältnisse – zumindest der Möglichkeit nach – auf Dauer angelegt, müssen sowohl Beschäftigte als auch der Betrieb in Vorleistungen eintreten, von denen sie nicht sicher sein können, wann bzw. ob sie sich auszahlen werden. Betriebe investieren etwa in das Humankapital der Beschäftigten, ohne sicher zu sein, ob diese Investition nicht später einem anderen Betrieb zugutekommt, während Beschäftigte Einkommenseinbußen in einer Einarbeitungszeit akzeptieren, ohne die Sicherheit einer Übernahme nach der Probezeit zu haben. Die soziale Dimension



einer Arbeitsbeziehung entsteht durch die Tatsache, dass der individuelle Arbeitnehmer Teil der Betriebsgemeinschaft ist. (Gegen-)Leistungen des Betriebes richten sich daher nicht nur an den Beschäftigten selbst, sondern werden etwa in Form von betrieblichen Zusatzleistungen an die Beschäftigten als Kollektiv gerichtet. In sachlicher Hinsicht schließlich ist auch die konkret erwartete Arbeitsleistung nicht exakt spezifiziert, sondern es besteht die Erwartung, dass die Beschäftigten ihre Tätigkeit im Betrieb in loyaler und mitdenkender Weise ausführen.

Durch diese Unbestimmtheit der Erwidierung von Leistungen mit entsprechenden Gegenleistungen entsteht für die Beteiligten ein Spielraum, innerhalb dessen sie ihre Leistung frei wählen können. Dieser Entscheidungsspielraum kann unter Umständen bis hin zur Leistungsverweigerung gehen, ohne dass dies vom Arbeitgeber erkannt und sanktioniert werden kann. Die Reziprozitätsnorm bildet einen Mechanismus, der dem entgegenwirkt. Die Erwartung, dass eine gute Leistung mit einer entsprechenden reziproken Gegenleistung belohnt wird, sorgt dafür, dass die Vertragspartner trotz der Möglichkeit des opportunistischen Verhaltens bereit sind, in Vorleistungen einzutreten. Reziprozität kann dabei als instrumentelle Reziprozität stattfinden, etwa indem vom Arbeitgeber mehr Lohn (der sogenannte Effizienzlohn) gezahlt wird als eigentlich üblich. Dadurch erwarten diese sich eine Mehrleistung der Beschäftigung bzw. dass Beschäftigte weniger zu opportunistischem Verhalten motiviert sind (z. B. Sesselmeier/Blauermel 1997: 153 ff.). Reziprozität in Beschäftigungsverhältnissen kann und wird jedoch in weiten Teilen intrinsisch motiviert sein, wie es in den ökonomischen Ultimatumexperimenten zum Vorschein kommt. Damit stellt Reziprozität eine im Sinne Boudons axiologische Handlungsorientierung dar.

Während in Beschäftigungsverhältnissen Arbeitgeber und Arbeitnehmer sowohl Erbringer als auch Empfänger von Leistungen sind, liegt der Fokus dieser Untersuchung und damit auch der folgenden Ausführungen auf den Beschäftigten. Diese Entscheidung ist wie folgt begründet: Sowohl im Fall des Kombilohnes „Mainzer Modell“ (vgl. Abschnitt 5.2.4.3) als auch in anderen empirischen Analysen (Whitehouse 1996: 132) ist die Arbeitgeberseite insofern unproblematisch, als sie nicht über die Kombilohnförderung der Beschäftigten informiert war. Folglich ist die Ursache für die hohe Abbruchrate eher bei den Beschäftigten selbst zu suchen.

### 3.2.3 Konsequenzen verletzter Reziprozität in Beschäftigungsverhältnissen

Aus den oben ausgeführten ökonomischen Experimenten wurde deutlich, dass die Verletzung der Reziprozitätsnorm zu sanktionierenden Handlungen führt. In diesem Abschnitt wird der Frage nachgegangen, welche Konsequenzen eine Verletzung der Reziprozitätsnorm außerhalb des engen Möglichkeitsraumes von Laborexperimen-

ten hat. Wie im Folgenden zu sehen ist, kann die Verletzung der Reziprozität in Beschäftigungsverhältnissen unterschiedliche Konsequenzen haben.

Die unmittelbarste Folge ist sicherlich ein negativer Einfluss auf die Bewertung des erhaltenen Einkommens. In seiner Theorie distributiver Gerechtigkeit<sup>17</sup> schlägt Jasso (1978, 1980) mit  $f = \ln(I/I^f)$  eine mathematische Repräsentation des kognitiven Bewertungsprozesses des erhaltenen Einkommens vor. Dabei ist  $f$  das Gerechtigkeits- oder Fairnessurteil des Beschäftigten in Bezug auf das erhaltene Einkommen. Dieses bestimmt sich aus dem Lohn  $I$  im Verhältnis zu einem für die ausgeführte Tätigkeit als fair betrachteten Lohn  $I^f$ . Im Kontext der betrieblichen Austauschbeziehungen kann der Bruch  $I/I^f$  folgendermaßen interpretiert werden: Er bezeichnet die Reziprozität zwischen dem vom Betrieb gezahlten Lohn und dem Lohn, den der Beschäftigte für seinen Arbeitseinsatz erwartet. Im Falle  $I=I^f$  herrscht vollständig ausgeglichene Reziprozität. Die Logarithmierung führt zu positiven Werten, wenn sich die Reziprozität (aus Sicht des Beschäftigten) in Richtung generalisierte Reziprozität verschiebt ( $I^f < I$ ), und umgekehrt zu negativen Fairnessurteilen, wenn der erhaltene Lohn das Reziprozitätsprinzip (aus Sicht des Beschäftigten) in Richtung negative Reziprozität verletzt ( $I^f > I$ ). Letzteres kann als Verletzung der Reziprozitätsnorm interpretiert werden. Die Logarithmierung reflektiert dabei zwei weitere von Jasso (Jasso 1978: 1414 f.) als wichtige Eigenschaften von Fairnessurteilen identifizierte Faktoren. Einerseits führt ein Lohn, der um die Hälfte kleiner ist als der faire Lohn, zu einer Negativverschiebung des Urteils um denselben Wert, wie eine Verdopplung des Lohnes zu einer Erhöhung des Fairnessurteils führt. Andererseits aber hat – bei konstantem fairem Lohn – eine negative Einkommensveränderung einen stärkeren Einfluss auf  $f$  als eine positive Einkommensveränderung. Dies entspricht der „common human experience that deficiency is felt more keenly than excess“ (Jasso 1978: 1415).

Während Jassos Theorie des Fairnessurteils dazu beiträgt, einstellungsrelevante Folgen von Reziprozitätsverletzungen herauszuarbeiten, kann die Fair-wage-effort-Hypothese von Akerlof und Yellen (1990) als eine Darstellung der handlungsrelevanten Folgen betrachtet werden. Akerlof und Yellen orientieren sich an soziologischen und sozialpsychologischen Austauschtheorien (Adams 1963; Homans 1961). Um die handlungsrelevanten Konsequenzen verletzter Reziprozität herauszustellen, soll das Modell zunächst in einer vereinfachten Form dargestellt werden. Konkret betrachtet die Fair-wage-effort-Hypothese den Einfluss der Reziprozität auf die Arbeitsleistung  $e = \min(I/I^f, 1)$ . Mit  $e = 1$  wird der „normale“ Arbeitseinsatz bezeichnet (Akerlof/Yel-

17 Die Begriffe „Reziprozität von Lohn und Arbeitsleistung“ sowie „distributive Gerechtigkeit“ und „Fairness“ werden im Rahmen dieser Arbeit als größtenteils auf dasselbe empirische Phänomen bezogen betrachtet und daher als synonym behandelt. Eine einheitliche Abgrenzung der Begriffe ist auch in der Fachliteratur bisher nicht erfolgt und stellt eine Aufgabe dar, die über die Möglichkeiten der vorliegenden Arbeit hinausgeht.

len 1990: 255), der sich im Falle vollständig ausgeglichener Reziprozität ergibt. Im Falle einer verletzten Reziprozität reduziert ein Beschäftigter seine Arbeitsleistung so lange, bis ein neuer Ausgleich hergestellt wurde (vgl. auch Robinson 1995; Lester et al. 2002: 42). Im Gegensatz zu Jassos Fairness-Modell sehen Akerlof/Yellen jedoch keinen zusätzlichen Effekt von Mehrzahlungen über den fairen Lohn hinaus auf die Arbeitsleistung, repräsentiert durch die Minimumbedingung.

Eine Reduktion der Arbeitsleistung ist allerdings nicht die einzige Handlungsoption für Beschäftigte. In vielerlei Zusammenhängen wird darauf hingewiesen, dass die ausgeglichene Reziprozität die Grundlage stabiler sozialer Beziehungen darstellt (Gouldner 1960: 169; Blau 2005: 126; Bode/Brose 1999: 181). Daher kommt es bei Verletzungen des Reziprozitätsprinzips zu einer Gefährdung der sozialen Beziehung insgesamt (Sahlins 1999: 154 f.). Dies gilt auch für Beschäftigungsverhältnisse. Beschäftigte, welche die Erfahrung machen, dass der Betrieb seinen Verpflichtungen zur Reziprozität in einem bestimmten Punkt nicht nachgekommen ist, gehen davon aus, dass auch im weiteren Verlauf der betrieblichen Arbeitsbeziehung keineswegs gesichert ist, dass auf die eigenen Arbeitsleistungen mit entsprechenden Gegenleistungen reagiert wird. Es erscheint ihnen das ausgeglichene reziproke Verhältnis von Leistung und Gegenleistung in der Beziehung verletzt und in Richtung negative Reziprozität verschoben. In Anlehnung an Sheppard, Lewicki und Minton (1992) Erweiterung des Exit-Voice-Ansatz (Hirschmann 1970) können vier Möglichkeiten der Reaktion auf ungerechte Verhaltensweisen, wie die Verletzung des Reziprozitätsprinzips, abgeleitet werden. Diese lassen sich als Kombinationen der Verhaltensdimensionen proaktiv/reaktiv und konstruktiv/destruktiv darstellen (Tabelle 2). Welche der Verhaltensweisen gewählt wird, hängt vom Ausmaß der wahrgenommenen Ungerechtigkeit ab. Exit (das Verlassen des Betriebes) und Voice (Widerspruch gegen die Ungerechtigkeit) sind proaktive Verhaltensweisen, die auf eine Veränderung der Situation abzielen. Beim Exit-Verhalten wählt ein Beschäftigter auf eine Verletzung der Reziprozität die Kündigung. Die Voice-Strategie könnte dagegen so aussehen, dass ein Arbeitnehmer seine Vorgesetzten direkt anspricht und sich beschwert oder auch den Betriebsrat einschaltet. Das Verlassen des Betriebes wird als destruktiv charakterisiert, da es das Beschäftigungsverhältnis auflöst. Dagegen wird der Widerspruch als konstruktives Verhalten klassifiziert. Die Beschäftigten haben aber noch andere Möglichkeiten, die Sheppard, Lewicki und Minton als reaktiv bezeichnen. Destruktiv-reaktives Verhalten äußert sich im Rückzugsverhalten, in der „inneren Kündigung“. So kann ein Beschäftigter sich dafür entscheiden, jegliche Eigeninitiative zu unterlassen und nur noch „Dienst nach Vorschrift“ durchzuführen. Loyalität bezeichnet demgegenüber das konstruktiv-reaktive Verhalten des „rationalizing“, bei dem der Beschäftigte etwa eine Umdeutung der ihm widerfahrenen Ungerechtigkeit entwickelt (Liebig 1997: 237 f.).

Tabelle 2: Mögliche Reaktionen auf Reziprozitätsverletzungen in Beschäftigungsverhältnissen

	proaktiv	reaktiv
konstruktiv	<i>Voice behavior</i> Change behavior	<i>Loyalty behavior</i> Rationalizing the injustice
destruktiv	<i>Exit behavior</i> Leave the scene	<i>Withdrawal behavior</i> Change behavior

Quelle: Sheppard/Lewicki/Minton 1992; Abbildung nach Liebig 1997: 238.

Der dem Reziprozitätsbegriff nahe stehende Begriff der Anerkennung (Voswinkel 2005; Kotthoff 2000) verweist darauf, dass es in Beschäftigungsverhältnissen bei der Reziprozität von Leistung und Gegenleistung nicht nur um rational-kalkulierende Anpassungsreaktionen der Beschäftigten an empfundene Verletzungen geht, sondern Emotionen eine wichtige Rolle spielen. „Die Gefühlsreaktion auf das Nichteinhalten einer Reziprozitätsnorm ist (...) persönliche Kränkung, Beleidigtsein, Empörung, Wut. In diesen Gefühlen drückt sich das Verlangen nach Sanktionen aus. Diese ist ein separater Aspekt der Schädigung des Normverletzers. Sie hat die Bedeutung eines Denkkzettels oder eines Appells, die eigene Würde wieder zu respektieren“ (Kotthoff 2000: 33; vgl. auch Honneth 2003: 222; Rousseau 1989: 128).

### 3.2.4 Subjektive Aspekte der Reziprozität

Die Reaktion des Beschäftigten auf eine Verletzung der Reziprozitätsnorm ist jedoch nicht nur vom objektiven Verhältnis zwischen dem gezahlten Lohn und einem als Maßstab herangezogenen fairen Lohn abhängig. In der Terminologie Boudons hieße das, dass die Handlungsorientierung am Reziprozitätsprinzip nicht nur axiologische, sondern oft auch kognitive Aspekte beinhaltet. Die Fair-wage-effort-Hypothese zieht dies mit in Betracht, indem sie die Größen als abhängig von der subjektiven Wahrnehmung der Beschäftigten beschreibt, repräsentiert mit dem Subskriptum *s*, als subjektiv wahrgenommener Lohn  $I_s$  und als subjektiv wahrgenommener fairer Lohn  $I_s^f$ . Damit ergibt sich die Arbeitsleistung als objektive Konsequenz aus der Relation zweier subjektiver Größen  $e = \min(I_s/I_s^f, 1)$ . In Abgrenzung zu den tatsächlichen Verstößen gegen die Reziprozitätsnorm und im Anschluss an Boudons Typ II der nicht zweckrationalen Überzeugungen werden daher im Folgenden von „kognitiven oder subjektiven Verletzungen der Reziprozitätsnorm“ (KVR) gesprochen. Die in der Reziprozitätsnorm stets inhärente Subjektivität wurde bereits von Gouldner (1960: 172, vgl. auch 2005: 113) thematisiert und zeigt sich auch in den Ergebnissen der Forschung zur Existenz und zum Wirken psychologischer oder auch subjektiver Verträge in Beschäftigungsverhältnissen (Rousseau 1989; Conway/Briner 2005).

Auch der kognitive Aspekt der Reziprozität hängt mit der Unvollständigkeit des Arbeitsvertrages zusammen. Da der Arbeitsvertrag immer nur den allgemeinen Rahmen des Beschäftigungsverhältnisses abstecken kann, füllen die Beschäftigten die „Lücken“ im Vertrag mit der eigenen und subjektiv geprägten Interpretation des Verhaltens der Organisation bzw. ihrer offiziellen Vertreter ihnen gegenüber. Das tun sie, indem sie Annahmen über das Nichtenthaltene machen (vgl. Konow 1996: 22). Dadurch entstehen psychologische Verträge als subjektive Interpretation der reziproken Verpflichtungen, die über die Formalisierung der Erwartungen im Arbeitsvertrag hinausgehen: „When an individual perceives that contributions he or she makes obligates the organization to reciprocity (or vice versa) a psychological contract emerges“ (Rousseau 1989: 124). Psychologische Verträge bestehen also in der subjektiven Interpretation der Beschäftigten darüber, was sie der Organisation schulden und umgekehrt die Organisation ihnen im Rahmen der betrieblichen Austauschbeziehung schuldet. Der Fokus des Begriffes „psychologischer Vertrag“ liegt erstens auf der individuellen Ebene des Beschäftigungsverhältnisses und hier vor allem auf der Seite der Beschäftigten. Nur Individuen können einen psychologischen Vertrag entwickeln, nicht das Unternehmen als ihr Vertragspartner.<sup>18</sup> Die Lokalisierung der psychologischen Verträge auf der individuellen Ebene unterscheidet sie von impliziten Verträgen, die als „betriebliche Praxis“ auf der Ebene der sozialen Beziehung liegen und ein stabiles Muster vorhersagbaren Handelns bilden. Zweitens kennzeichnet einen psychologischen Vertrag, dass dessen Inhalt einseitig ist. „It is the individual's belief in an obligation of reciprocity that constitutes the contract. The belief is unilateral, held by a particular individual, and does not constrain those of any other parties to the relationship“ (Rousseau 1989: 124). Der Inhalt des psychologischen Arbeitsvertrages eines Beschäftigten muss sich daher nicht mit der Einschätzung der Organisationsvertreter im Betrieb oder eines außenstehenden Beobachters decken (Morrison/Robinson 2004: 162). Sie können in diesem Sinne entweder Boudons Typ I oder Typ II der kognitiven Überzeugungen entsprechen und damit entweder intersubjektiv oder nur subjektiv zutreffen. Trotzdem handeln die Beschäftigten unter der Annahme, dass der psychologische Vertrag intersubjektiven Charakter hat und dass „contributions will be reciprocated and that a relationship exists where actions of one party are bound to those of another“ (Rousseau 1989: 128).

Die aufgrund psychologischer Verträge gebildeten Verhaltenserwartungen an den Betrieb können, müssen aber nicht in allen Punkten mit der offiziellen Unternehmenspolitik bzw. deren Umsetzung und subjektiver Interpretation der Vertreter

---

18 Allerdings legt die Forschung zu psychologischen Verträgen nahe, dass auch die Vorgesetzten psychologische Verträge in ihrer Rolle als Vertreter des Unternehmens entwickeln (Lester et al. 2002).

des Betriebes übereinstimmen. Dies führt dazu, dass zwischen Betrieb und Beschäftigten unterschiedliche Einschätzungen darüber bestehen können, wann die Reziprozitätsnorm erfüllt ist. Divergenzen zwischen den Interpretationen der Beschäftigten und der Organisationsvertreter können auf drei Mechanismen zurückgeführt (Morrisson/Robinson 2004: 169 ff.) werden:

Erstens können unterschiedliche Sichtweisen darüber bestehen, welches überhaupt konkret die Verpflichtungen des Gegenübers sind. So kann etwa der Arbeitnehmer davon ausgehen, dass seine geleisteten Überstunden entlohnt werden, während der Betrieb sich lediglich dazu verpflichtet sieht, für einen späteren Ausgleich durch eine Verringerung der Arbeitszeit zu sorgen. Zweitens kann die Wahrnehmung verzerrt sein, ob und in welchem Ausmaß Verpflichtungen erfüllt wurden. Typischerweise wird der Grad der Erfüllung der eigenen Verpflichtungen überbewertet. „People have self-serving views of fairness that can cause them to not see their own role in a conflict situation, hence feeling that they are being unfairly victimized“ (ebd.). Beschäftigte haben demnach oft eine verzerrte Wahrnehmung dessen, ob sie oder der Arbeitgeber die Reziprozitätsnorm erfüllen, und haben ihre eigene Leistung im Vergleich zur Gegenleistung des Betriebes als weit größer in Erinnerung, als dies eigentlich der Fall war (Morrisson/Robinson 1997: 240; vgl. auch Fetchenhauer/Bierhoff 2004: 135). Drittens führen Individuen bei Bewertungen erfüllter Verpflichtungen soziale Vergleiche durch, orientieren sich also daran, was ihnen ähnliche Personen erhalten. Auch dieser Vergleich kann ein hohes Maß an Subjektivität beinhalten und zu subjektiver relativer Deprivation führen (Stouffer et al. 1949; Crosby 1976; vgl. auch Clark/Oswald 1996).

Daher können Verletzungen der Reziprozitätsnorm nicht nur auf absichtliche Handlungen (Ungerechtigkeiten), sondern auf die Divergenz oder Inkongruenz von Erwartungen an den Tauschpartner zurückgehen (Morrisson/Robinson 1997: 231 ff.). Solche allein auf der Divergenz von Erwartungen beruhenden Verletzungen der Reziprozitätsnorm führen jedoch ebenso wie tatsächliche Verletzungen zu entsprechendem Handeln der Beschäftigten.

Aus der empirischen Forschung zu Verletzungen psychologischer Arbeitsverträge sind im Zusammenhang mit der Erklärung von Kombilohnabbrüchen vor allem folgende Ergebnisse relevant. Heriot et al. (1997) untersuchen, welche Aspekte des Beschäftigungsverhältnisses es sind, die sich im psychologischen Vertrag der Beschäftigten niederschlagen, und ob sich diese von den Aspekten unterscheiden, welche die Vertreter des Betriebes als relevant betrachteten. Zunächst ergibt sich in ihrer Untersuchung, dass beide Vertragsparteien gleichermaßen die Einhaltung der Arbeitszeit, die Leistungserbringung und Ehrlichkeit als zentrale Aspekte des Beschäftigungsverhältnisses betrachten. Divergenzen bestehen allerdings dahingehend, dass Betriebsvertreter darüber hinaus mit Humanität und Anerkennung

eher weiche Faktoren betonen, während Beschäftigte die Arbeitsplatzsicherheit, gesicherte Arbeitsbedingungen und einen fairen Lohn als wichtig einschätzen. Die Fairness oder ausgeglichene Reziprozität des Lohnes stellt laut der Untersuchung von Lester et al. (2002; ähnlich auch Porter et al. 1998) auch den Aspekt des psychologischen Arbeitsvertrages dar, bei dem die Divergenzen zwischen Beschäftigten und Betrieb besonders groß sind. In ihrer Untersuchung vergleichen sie die Einschätzung von Beschäftigten und Organisationsvertretern hinsichtlich der Erfüllung wechselseitiger Verpflichtungen in verschiedenen Bereichen des Beschäftigungsverhältnisses. Dabei zeigt sich nicht nur, dass im Bereich der Frage, ob der Lohn fair ist bzw. in angemessener Weise der Leistung des Beschäftigten entspricht, der Unterschied in den Einschätzungen der Vertragspartner besonders stark ist. Darüber hinaus ist der Lohn der einzige Aspekt, bei dem sich im Mittel über die Beschäftigten ein tatsächlicher Bruch des psychologischen Vertrages feststellen lässt (Lester et al. 2002: 49).

Dieser Befund sollte nicht dazu führen, die Frage der Reziprozität als eine Frage der materiellen Aspekte des Beschäftigungsverhältnisses zu betrachten. Vielmehr ist die Lohnhöhe auch deshalb von Divergenzen in den Reziprozitätserwartungen betroffen, weil sie auch Ausdruck der Anerkennung ist, die Beschäftigten entgegengebracht wird, und nicht nur im Rahmen einer finanziellen Aufrechnung von Leistung und Gegenleistung bewertet wird, sondern vor allem auch moralischer und normativer Wertung ausgesetzt ist (Honneth 2003; Voswinkel 2002). Dies erklärt vielleicht auch, warum das Ausmaß der empfundenen Diskrepanz zwischen tatsächlichem und angemessenem Lohn bei Geringverdienern am höchsten ist (Liebig/Schupp 2005: 723).

Allgemein können Brüche psychologischer Verträge auf unterschiedliche Ursachen zurückgeführt werden (Morrison/Robinson 1997: 231; Lester et al. 2002: 47). Es kann sich um absichtliche Vertragsbrüche, um unbeabsichtigte Vertragsbrüche aufgrund externer Hinderungsgründe und um Vertragsbrüche aufgrund der Inkongruenz der Vertragsinhalte zwischen den Vertragsparteien handeln. Die Ergebnisse von Lester et al. (2002) verweisen darauf, dass die Betriebsvertreter arbeitgeberseitige Nichterfüllung von Verpflichtungen tendenziell eher den externen Hinderungsgründen zurechnen oder auf eine Fehlinterpretation der Beschäftigten (Inkongruenz) zurückführen, während Beschäftigte solche Vertragsverletzungen eher als einen vom Betrieb beabsichtigten Vertragsbruch interpretierten.

Was die Konsequenzen von kognitiven Reziprozitätsverletzungen betrifft, so zeigt sich auch bei nur subjektivem Gehalt der psychologischen Verträge eine dem Fall der tatsächlichen Reziprozitätsverletzung (vgl. Abschnitt 3.2.3) entsprechende Reaktion der Beschäftigten, vor allem eine erhöhte Kündigungs- und geringere Leistungsbereitschaft (Herriot et al. 1997: 161; Lester et al. 2002: 40). Im Sinne des

self-serving bias zeigt sich, dass eine kognitive Reziprozitätsverletzung aus Sicht der Beschäftigten ihre Leistung nicht beeinträchtigt, die Vorgesetzten allerdings eine Leistungsverminderung beobachten und auch die Bindung der Beschäftigten an den Betrieb (commitment) zurückgeht. Berichtenswert ist außerdem, dass die Konsequenzen kognitiver Verletzungen der Reziprozität umso stärker ausfallen, je größer die Divergenz zwischen den Auffassungen der Beteiligten ist (Porter et al. 1998).

### 3.3 Negative Wirkung finanzieller Anreize durch subjektiv verletzte Reziprozität

Das Zusammenspiel der normativen und der kognitiven Komponente des Reziprozitätsprinzips führt mitunter dazu, dass ein Eingriff in das Beschäftigungsverhältnis in Form finanzieller Zusatzleistungen durch den Betrieb nicht die intendierten positiven Folgen nach sich zieht. Eines der bekanntesten Beispiele unbeabsichtigt negativer Folgen finanzieller Anreize stammt von Titmuss (1997), der argumentiert, dass die Bezahlung von Blutspendern dazu führen würde, dass nicht mehr, sondern weniger Blut zur Verfügung gestellt wird. Diese These wurde von den Ökonomen Solow (1971) und Arrow (1972) bestritten. Sie gingen davon aus, dass sich das Gesamtangebot an Blutmenge erhöht, da durch eine Bezahlung neben den „altruistischen“ Spendern auch die nicht altruistischen Personen zur Blutspende motiviert werden. Tatsächlich zeigt sich, dass dem nicht so ist, sondern ein negativer Zusammenhang zwischen dem Blutspenden und finanziellen Anreizen besteht (Frey 1997: 82 f.).

Entsprechend berichten Frey und seine Mitautoren, dass von der Leistung abhängige Löhne (Anreizlöhne) nicht immer zu einer Steigerung der Produktivität führen (Frey 1997: 88). Zum Teil kommt es durch sie sogar zu einer Verminderung der Produktivität, also genau zum Gegenteil dessen, was bewirkt werden soll. Frey schreibt dies einem Verdrängungseffekt (crowding-out) zu, bei dem durch einen externen Anreiz die intrinsische Arbeitsmotivation verdrängt wird. Ein solcher Effekt tritt dann ein, wenn die Beschäftigten die externe Belohnung als Kontrollversuch wahrnehmen. Dies untergräbt die Arbeitsmoral von Arbeitnehmern, und daher können sie ihre leistungssteigernde Wirkung nicht entfalten (Frey/Benz 2001: 19). Nehmen Beschäftigte die Belohnung jedoch als Unterstützung wahr, kann sich im Gegenteil die intrinsische Motivation erhöhen (Frey/Jegen 2000: 7). Mit Blick auf Metaanalysen wird dabei festgestellt, dass der Verdrängungseffekt bei monetären stärker als bei nicht monetären Anreizen ausfällt (ebd. 10).

An zwei Beispielen wird der Verdrängungseffekt nachgewiesen. So zeigt sich im Bereich des Ehrenamtes (Frey/Goette 1999), dass zwar die Höhe zusätzlicher finanzieller Belohnung ein Anreiz für Engagement darstellt, die Existenz einer Be-



lohnung sich jedoch negativ auf die Zahl der geleisteten Stunden auswirkt. „Direct incentives may backfire, leading to less volunteering“ (Frey/Jegen 2000: 13). In einem Experiment zu lokal unerwünschten Projekten, bei dem es um die (hypothetische) Zustimmung zum Bau eines Lagers für atomaren Abfall geht, zeigt sich dieser Verdrängungseffekt ebenfalls. Zunächst wurden die Probanden nach ihrer Zustimmung gefragt, ohne dass ein finanzieller Ausgleich angesprochen wurde. In diesem Fall gaben 50,8 % ihre Zustimmung zum Bau des Lagers. Nachdem die Frage allerdings so formuliert wurde, dass für die Gemeinden finanzielle Entschädigungen vorgesehen sind, die einem Bau zustimmen, sinkt die Zustimmung deutlich auf nur noch 24,6 % (Frey/Oberholzer-Gee 1997). „While external interventions, e.g. offering compensation, manages to address concerns regarding the costs of a noxious facility, it reduces the intrinsic motivation to permit the construction of such a facility“ (Frey/Jegen 2000: 15).

Nicht nur im Fall von Belohnungen, auch bei Bestrafungen zeigt sich mitunter eine paradoxe Reaktion der Betroffenen. So zeigt sich, dass die Androhung von Strafzahlungen das zu verhindernde Verhalten mitunter auch verstärkt. Gneezy und Rustichini (2000) berichten, dass in einem Kindergarten, der für Eltern eine Strafzahlung eingeführt hatte, die ihre Kinder zu spät abholten, dadurch nicht etwa die Pünktlichkeit der Eltern erhöht wurde, sondern im Gegenteil der Anteil der unpünktlichen Eltern stieg.

Der Kombilohnparadoxie am nächsten kommt die Analyse von Anreizlöhnen von Fehr und Gächter (2002, 2000: 170 ff.). Sie untersuchen, welche Wirkung Verträge auf die Arbeitsleistung haben, in denen ein Anreizmechanismus gegen Leistungszurückhaltung eingebaut ist. Dazu führen sie ein Experiment durch, das die Bedingungen der Vertragserfüllung in einem Beschäftigungsverhältnis abbilden soll und bei dem drei verschiedene Vertragsausgestaltungen betrachtet werden. Diese sind Verträge ohne Anreizmechanismus (Trust-Treatment, TT), Verträge mit Strafmöglichkeit bei Leistungszurückhaltung (Incentive-Treatment, IT) und Verträge mit Bonuszahlungen bei Leistungserfüllung (Bonus-Treatment, BT).

Das Experiment umfasst drei Phasen, in denen zwei zufällig ausgewählte Personen miteinander interagieren sollen. In der ersten Phase macht der Käufer ein Angebot, das den fixen Preis  $P$  und die erwünschte Qualität  $\hat{Q}$  enthält. In der zweiten Phase entscheidet der Verkäufer, ob er auf das Angebot eingeht. Ist dies der Fall, dann wählt er ein Qualitätsniveau  $Q$  aus, wobei er hierbei nicht an den Wunsch des Käufers gebunden ist, also auch weniger als verlangt leisten kann. Das IT unterscheidet sich vom TT dadurch, dass der Käufer die Möglichkeit hat, dem Verkäufer eine Strafzahlung abzuverlangen, wenn dieser von der abgemachten Qualität  $\hat{Q}$  nach unten abweicht. Das Experiment ist so gestaltet, dass eine Minderleistung mit einer festgelegten Wahrscheinlichkeit  $w$  aufgedeckt wird und mit einer Wahr-

scheinlichkeit 1-w unentdeckt bleibt. Schließlich wird im BT die Situation im IT so variiert, dass grundsätzlich ein Bonus gezahlt wird, der jedoch bei entdeckter Minderleistung einbehalten wird. Damit für die Probanden auch tatsächliche finanzielle Anreize bestanden, wurden nach Beendigung des Experimentes Geldbeträge entsprechend den Ergebnissen des experimentellen Tauschhandels ausgezahlt.

Auch hier zeigen sich paradoxe Ergebnisse des Lohnanreizes. Zum einen zeigt sich, dass sowohl die Qualität als auch die Kooperationsbereitschaft, also Bereitschaft, auf Preisangebote mit einer entsprechenden Gegenleistung zu reagieren, in IT-Anreizverträgen geringer ausfällt als in TT-Verträgen ohne Lohnanreize. Das heißt, obwohl durch die Möglichkeit der Strafzahlung rational handelnde Individuen zu einer Kooperation motiviert wären und insgesamt eine höhere Qualität der Leistungen des Verkäufers zu erwarten wären, tritt genau das Gegenteil ein (Fehr/Gächter 2002: 16). Vergleicht man die Ergebnisse des TT-Vertrages mit denen des BT-Vertrages, dann sieht man, dass auch durch Bonusverträge die Kooperationsbereitschaft nicht steigt, sondern im Gegenteil niedriger ausfällt, wobei dies jedoch nicht so stark ausgeprägt ist wie bei den IT-Anreizverträgen (ebd. 27).

Die Autoren der Analyse ziehen daraus den Schluss, dass die Reaktion der Individuen auf Anreizverträge nicht allein individuell nutzenmaximierend motiviert, sondern an der Reziprozitätsnorm ausgerichtet ist. „Reciprocity means that in response to friendly actions, people are frequently much nicer and much more cooperative than predicted by the self-interest model; conversely, in response to hostile actions they are frequently much more nasty and even brutal (Fehr/Gächter 2000: 159). Nach ihrer Interpretation lassen sich die experimentellen Ergebnisse dadurch erklären, dass die Lohnanreize „may cause a hostile atmosphere of threat and distrust, which reduces any reciprocity-based extra effort“ (Fehr/Gächter 2000: 170). Ähnlich argumentieren auch Tenbrusel und Messick (1999), denen zufolge vertraglich geregelte Anreize einen Signaleffekt entwickeln. Die Möglichkeit der Sanktion oder der Bonuszahlung führt dazu, dass die Beschäftigten das betriebliche Austauschverhältnis eher unter wirtschaftlichen als moralischen Gesichtspunkten betrachten. In diesem Fall findet also ein Verdrängungseffekt der Reziprozität statt (Fehr/Gächter 2000: 171), der erklärt, warum die Leistungsbereitschaft am höchsten ist, wenn im Vertrag keine Anreize festgelegt sind, und am niedrigsten, wenn der finanzielle Anreiz in Form einer Strafzahlung bei nachgewiesener Leistungszurückhaltung erfolgt. Die subjektive Komponente der verletzten Reziprozität zeigt sich vor allem darin, dass bei exakt gleichen finanziellen Konsequenzen als Strafzahlungen<sup>19</sup> formulierte Anreize zu einer stärkeren negativen Reaktion führen als solche Verträge, in de-

---

19 Während Frey die Ergebnisse von Fehr und Gächter als „crowding-out effekt for intrinsic motivation in form of a tendency for reciprocal behavior“ (Frey/Jegen 2000: 11) interpretiert, sehen diese hier eventuell unterschiedliche Mechanismen am Wirken (Fehr/Gächter 2002: 31 f.).

nen der Anreiz als Bonuszahlung formuliert ist. Aus der subjektiven Sicht der „Beschäftigten“ im Experiment stellen Strafandrohungen eine stärkere Verletzung der Reziprozität dar als Bonuszahlungen im Falle der Vertragserfüllung. Ebenfalls auf die Subjektivität der Reziprozitätsverletzung verweist, dass auch Bonuszahlungen negativ interpretiert werden und die Leistungsbereitschaft mindern.

### 3.4 Zusammenfassung: Erklärungsmechanismus und Hypothesen

Wie in den bisherigen Abschnitten dargestellt, sind die Austauschbeziehungen in Beschäftigungsverhältnissen nicht (nur) von einer Orientierung am individuellen Nutzen geprägt, sondern auch von der Wirkung der Reziprozitätsnorm einerseits und andererseits von der Tatsache, dass die Arbeitnehmer die arbeitgeberseitige Einhaltung der Norm auf Basis subjektiver Einschätzungen beurteilen. Daraus lässt sich folgender Erklärungsmechanismus für den fehlenden Erfolg von Kombilohnmaßnahmen bei der dauerhaften Integration in Beschäftigungsverhältnisse ableiten.

Da der Kombilohnzuschuss nicht vom Arbeitgeber gezahlt wird, sondern vom Staat als Verringerung der Beiträge zu den Sozialversicherungen bzw. als Zuschlag zum Kindergeld gewährt wird, entfaltet er nicht die von ihm erwartete Wirkung einer Lohnerhöhung. Im Gegenteil wird die Bezuschussung der Beschäftigung vom Arbeitnehmer als Signal interpretiert, dass der Arbeitgeber nicht bereit ist, einen angemessenen Lohn zu zahlen. Diese Einschätzung ist insofern subjektiv, als diese Sichtweise nicht unbedingt vom Arbeitgeber bzw. seinen Vertretern im Betrieb geteilt werden muss. Zum einen wissen die Betriebe oft nichts von der Bezuschussung der Beschäftigung (Kaltenborn et al. 2005: 35; Gewiese 2004; vgl. auch Whitehouse 1996: 132), zum anderen ist für die Betriebe ausschlaggebend, welchen Preis die Arbeitskraft auf dem Markt erzielt, der Lohn wird also von Angebot und Nachfrage auf dem Markt bestimmt. Die Güter und Dienstleistungen, bei deren Herstellung Geringqualifizierte eingesetzt werden können, sind auf dem Markt nur zu relativ geringen Preisen absetzbar, und daher können Betriebe ihnen nur relativ geringe Löhne zahlen. „Denn nicht Leistungsmaße dienen dem Markt als Norm materieller Verdienste (...), sondern die zahlungskräftige Nachfrage, die Arbeitskräfte, Güter und Dienstleistungen finden“ (Neckel/Dröge 2002: 100; vgl. auch Lengfeld/Krause 2006: 101; Voswinkel 2002: 66). Insofern werden Betriebe gerade auch bei Neueinstellungen davon ausgehen, dass die ausgeglichene Reziprozität von Lohn und Leistung dann besteht, wenn der Lohn den für die Leistung erzielbaren Marktpreis widerspiegelt. Aus Sicht der Betriebe besteht daher ein „ganz normales“ Arbeitsverhältnis bzw. – falls Kenntnis über die Förderung besteht – eines, in dem der Beschäftigte noch zusätzliches Einkommen bezieht. Dagegen werden die Geförderten durch die Bezuschussung darauf hingewiesen,

dass es sich um eine Beschäftigung im Niedriglohnbereich handelt, und es entsteht aufgrund der Empfindung von Unterbezahlung bei den Geförderten, so die erste Hypothese, eine subjektiv empfundene Reziprozitätsverletzung. Obwohl also der Intention nach an der Aufstockung des Einkommens orientiert, signalisiert der Kombilohn dann manchen Beschäftigten eine zu geringe Wertschätzung ihrer Arbeitsleistungen durch den Betrieb. Verstärkt wird das Empfinden einer Reziprozitätsverletzung dadurch, dass die Beschäftigten nach Ablauf der Förderung mit einem Wegfall des Zuschusses und damit mit einem Einkommensverlust rechnen müssen. Schließlich besteht keine Gewissheit, dass der Lohn nach Wegfall der Förderung vom Arbeitgeber erhöht wird. Ist dies der Fall, dann wird sich dies zunächst in einstellungsrelevanten Folgen äußern. Diese können im Anschluss an Jasso (1978) als Gerechtigkeitsurteile beschrieben werden oder, wie hier, als subjektive Wertschätzung des Lohnes (Lohnzufriedenheit  $I_z$ ):

$$I_{Z_{kl}} = \frac{(I + z)_s}{I_s^f} < I_{Z_{kein\ kl}} = \frac{(I)_s}{W_s^f}$$

Trotz Erhöhung des Gesamteinkommens durch den Kombilohn ( $I + z$ ) im Vergleich zu ( $I$ ) hat die Förderung eine negative Auswirkung auf die subjektive Wertschätzung des Lohnes, da sie subjektiv nicht als Lohnerhöhung, sondern Reziprozitätsverletzung wahrgenommen wird.

**Hypothese 1:** *Im Vergleich zu regulärer Beschäftigung führt eine Kombilohnförderung durch die subjektive Verletzung der Reziprozität trotz objektiver Einkommenserhöhung zu geringerer Lohnzufriedenheit.*

Diese Einschätzung der verletzten Reziprozität führt bei Kombilohnbeziehern dazu, dass deren Beschäftigung im Vergleich zu regulären, ungeförderten Beschäftigungsverhältnissen nicht wie erwartet stabiler, sondern im Gegenteil weniger stabil ist. Das Spannungsverhältnis zwischen marktgemäßer und leistungsäquivalenter, reziproker Entlohnung ist keinesfalls auf kombilohngeförderte Beschäftigungsverhältnisse beschränkt. Tendenziell sind alle Beschäftigungsverhältnisse davon betroffen. Neckel und Dröge (2002: 111) vermerken dazu: „Seinen normativen Kern findet das Gerechtigkeitskonzept von Leistung daher im wechselseitigen Verpflichtungsverhältnis, das die Forderung der Äquivalenz von Leistung und Gegenleistung etabliert (...) und eben dieses Äquivalenzverhältnis wird durch reine Ergebnisorientierung des Marktes (...) einer Belastungsprobe unterzogen, deren sinnhafte Auswirkung noch weitgehend unerforscht sind.“ Allerdings tritt dieses Spannungsverhältnis bei kombilohngeförderten Beschäftigungsverhältnissen aufgrund der Signalwirkung

des Zuschusses bewusster zutage. Daher ist dann auch zu erwarten, dass Kombilohnbeschäftigungen häufiger abgebrochen werden als reguläre Beschäftigungsverhältnisse. Zu den üblichen Gründen für eine Beendigung kommen noch kombilohnspezifische Beendigungsgründe hinzu, sowohl was die Arbeitgeber- als auch was die Arbeitnehmerseite betrifft. Diese resultieren aus entsprechenden Anpassungshandlungen, mit denen die Geförderten auf die verletzte Reziprozitätsnorm reagieren. In einer Übertragung der Typologie von Reaktionsweisen auf Ungerechtigkeit im Betrieb (vgl. Abschnitt 3.2.3) auf den Kontext des Kombilohnes (Tabelle 3) wird dies deutlich.

Tabelle 3: Reaktionen auf Reziprozitätsverletzungen im Kontext des Kombilohnes

	proaktiv	reaktiv
konstruktiv	<i>Voice behavior</i> Offene Leistungsreduktion (Risiko: AG-Kündigung)	<i>Loyalty behavior</i> Akzeptanz der Sichtweise des Betriebes
destruktiv	<i>Exit behavior</i> AN-Kündigung bzw. AG-Kündigung provozieren	<i>Withdrawal behavior</i> Verdeckte Leistungszurückhaltung (Risiko: AG-Kündigung)

Die Exit-Option besteht auch im Fall von Kombilohnmaßnahmen darin, das Beschäftigungsverhältnis zu beenden. Dies ist eine sehr starke Reaktion, bei der Beschäftigte eventuell erhebliche Nachteile hinnehmen müssen, z. B. Sperrzeiten beim Bezug von Arbeitslosenunterstützung. Wählt ein Beschäftigter aufgrund der subjektiven Verletzung seiner Reziprozitätserwartung die Exit-Option, muss dies aber nicht gleichbedeutend mit einer Arbeitnehmerkündigung sein. Gleichermaßen kann er, gerade um Sanktionen des Arbeitsamtes zu vermeiden, den Arbeitgeber durch bestimmtes Verhalten zur Kündigung provozieren oder auch mit ihm eine Absprache treffen, sodass der Arbeitgeber ihm eine Kündigung ausspricht (vgl. Bielenski et al. 2002: 443). Andererseits können Beschäftigte durch eine Selbstkündigung auch einer Arbeitgeberkündigung zuvorkommen (Jackofsky 1984: 78).

Neben der Exit-Option ist als Alternative die Voice-Möglichkeit gegeben. Im Kontext des Kombilohnes besteht diese in einer möglichst sichtbaren Leistungsminderung, die vermitteln soll, dass die arbeitgeberseitige Verletzung der ausgeglichenen Reziprozität von Arbeitsleistung und Lohn vom Beschäftigten bemerkt und daher nicht hingenommen wird und eine – aus Arbeitnehmersicht – Wiederherstellung des Ausgleiches erfolgt. Dies kann wiederum vom Arbeitgeber, der bis dahin keinen Grund hatte, von einer nicht ausgeglichenen Reziprozität auszugehen, als eine Verletzung eben dieser in Richtung negative Reziprozität *durch den Beschäftigten* interpretiert werden. In der Folge steigt die Neigung des Arbeitgebers zur Kündigung des Geförderten.

Findet die Leistungszurückhaltung nicht offen, sondern verdeckt statt, dann ist dies als reaktiv-destruktives Verhalten einzuordnen. Solche verdeckte Leistungszurückhaltungen als Reaktion auf Verletzungen der ausgeglichenen Reziprozität können etwa durch Fehlzeiten, Krankheitstage oder sogar Diebstahl erfolgen (Liebig 1997: 218; Rousseau 1989: 125; Lester et al. 2002). Sowohl der offene als auch der verdeckte Ausgleich sind nicht als Exit intendiert, tragen jedoch beide ein Risiko für die Beendigung des Arbeitsverhältnisses in sich, nun aber durch den Arbeitgeber. Auf die offene Leistungszurückhaltung kann der Arbeitgeber ebenso mit einer Kündigung reagieren wie auf die Schädigungsversuche, sollten diese entdeckt werden. Schließlich ist die letzte und als einzige die Beschäftigungsstabilität nicht gefährdende Möglichkeit, mit verletzter Reziprozität umzugehen, diese zu rationalisieren, z. B. indem Gründe gesucht werden, warum es dem Betrieb nicht möglich war, dem Reziprozitätsprinzip nachzukommen. Dies entspräche dann den anderen oben angeführten Interpretationen der Reziprozitätsverletzung, der Interpretation als Inkongruenz der psychologischen Verträge oder der Interpretation als bedingt durch vom Betrieb nicht beeinflussbarer Rahmenbedingungen, z. B. der ungünstigen Wirtschaftslage. Aus diesen Überlegungen ergibt sich die zweite Hypothese:

**Hypothese 2:** *Kombilohngeförderte Beschäftigungsverhältnisse enden früher als reguläre Beschäftigungsverhältnisse.*

Obwohl für das Mainzer Modell bereits festgestellt wurde, dass sich die Förderung zumindest in Westdeutschland negativ auf die Beschäftigungsstabilität auswirkt, lohnt die nochmalige Prüfung, denn die Evaluation wurde lediglich auf Basis von Ereignisanalysen durchgeführt, und die Vorteile des Matchingansatzes (vgl. Abschnitt 4.1.4) wurden nicht genutzt.

Sollten sich beide Hypothesen bestätigen, wäre dies allerdings noch kein Beleg dafür, dass eine geringere Beschäftigungsstabilität in Kombilohnmaßnahmen durch die subjektive Verletzung der Reziprozität hervorgerufen wird. Es wäre also keine Erklärung gegeben. Daher lautet die dritte Hypothese:

**Hypothese 3:** *Die geringere Stabilität kombilohngeförderter Beschäftigungsverhältnisse lässt sich durch die negative Wirkung des Kombilohnes auf die subjektive Wertschätzung des Lohnes erklären.*

Es stellt sich nun die Frage, wie die Hypothesen 1–3 überprüft werden können. Es handelt sich bei allen drei um Kausalhypothesen der besonderen Form, dass die Wirkung einer einzelnen Variable auf eine weitere Variable zu untersuchen ist. Dabei wäre zunächst das einfachste Vorgehen, die Gruppe der Kombilohngeförderten

mit der der regulär Beschäftigten hinsichtlich ihrer Lohnzufriedenheit bzw. ihrer Beschäftigungsstabilität zu vergleichen. Allerdings besteht dabei das Problem, dass sich beide Gruppen nicht nur hinsichtlich des Kombilohnbezuges, sondern auch hinsichtlich weiterer Drittvariablen unterscheiden. Daher ist es möglich, dass sich die Werte der Zielvariablen zwischen den Gruppen auch dann unterscheiden, wenn vom Kombilohnbezug keinerlei Wirkung ausgeht. In die Maßnahme „Kombilohn“ werden ganz bestimmte Personengruppen eintreten, während andere, ebenso spezifische Personengruppen dies nicht tun werden. In einem solchen Fall spricht man von Kompositions- oder Selektionseffekten (Heckman/Smith 1996: 45 f.). Zur Kontrolle der Drittvariablen bietet sich zunächst ein experimentelles Forschungsdesign mit zufälliger Aufteilung in die beiden Gruppen an, das aber mit einigen Nachteilen behaftet ist (vgl. Abschnitt 4.1.1). Kommt dieses nicht infrage, werden gewöhnlich Regressionsanalysen durchgeführt, und die Drittvariablen werden als Kontrollvariablen in die Regression aufgenommen.

Dieses Vorgehen hat allerdings einige wesentliche Nachteile. Zum Ersten treffen Regressionsanalysen stets Annahmen über den funktionalen Zusammenhang zwischen den enthaltenen Variablen, also Ursache, Wirkung und Drittvariablen. Treffen diese Annahmen (z. B. der Linearität) nicht zu, können eventuell vorhandene Effekte nicht erkannt werden. Zum Zweiten benötigen Regressionsanalysen zur Durchführung von Signifikanztests die unverzerrte Schätzung der Standardfehler, zu der wiederum nicht nur die Drittvariablen (also Variablen, die sowohl die Ursache als auch die Wirkung beeinflussen), sondern auch alle weiteren Einflussfaktoren auf die Wirkungsvariable nötig sind (Gujarati 1995: 467 f.). Drittens schließlich werden bei Regressionsanalysen für die Bereiche der Drittvariablen, in denen kein gemeinsamer Stützbereich bei Geförderten und Ungeförderten besteht, Extrapolationen durchgeführt: „(...) when the x-variables [i.e., covariates] show real differences among groups – the case in which adjustment is needed most – covariate adjustment [i.e., regression adjustment] involve a greater or less degree of extrapolation. To illustrate by an extreme case, suppose that we were adjusting for differences in parent's income in a comparison of private and public school children, and that the private-school incomes ranged from \$10,000–\$12,000, while the public-school income ranged from \$4,000–\$6,000. The covariance would adjust results so that they allegedly applied to a mean income of \$8,000 in each group, although neither group has any observation in which incomes are at or near this level“ (Cochran 1957, zit. n. Rubin 2001: 174).

Eine Alternative bietet das Propensity Score Matching, das im Rahmen von Rubins Kausalmodell genau für solche Fälle entwickelt wurde, in denen eine kausale Fragestellung verfolgt wird, ein Experiment aber aus verschiedenen Gründen nicht möglich ist.





## 4 Messung kausaler Effekte in Rubins Kausalmodell

Rubins Kausalmodell (RKM) stellt das aktuell in der Evaluation aktiver Arbeitsmarktpolitik meistverwendete Verfahren der Kausalanalyse dar (Caliendo/Kopeinig im Erscheinen; Fitzenberger/Speckesser 2000; Heckman/Ichimura/Todd 1997, 1998) und findet langsam auch Eingang in die soziologische Forschung (Morgan/Harding 2006; Gangl/DiPrete 2004a, 2004b; DiPrete/Engelhardt 2004; Sobel 1995).

### 4.1 Rubins Kausalmodell

Ursprünglich stammt Rubins Kausalmodell (Rubin 1974, 2005, 2007; Holland 1986) allerdings aus der statistischen Forschung. Das Kausalitätsverständnis des RKM orientiert sich stark am Leitbild des Zufallsexperimentes (Rubin 1974, 2007), was sich unter anderem auch in der verwendeten Terminologie zeigt. So wird die Ursache oft als „Treatment“ bezeichnet, und ihre Wirkung wird stets im Vergleich mit einem Kontrollzustand analysiert, in dem das Treatment nicht verabreicht wurde. Das RKM ist ein sehr allgemeines Konzept und findet in vielen Bereichen der Wissenschaft Verwendung. So wird damit etwa die Wirkung von Medikamenten (Little/Rubin 2000) ebenso analysiert wie die Wirkung von Arbeitsmarktmaßnahmen (Fitzenberger/Speckesser 2000), von Lehrmethoden (Müller 2005) etc.

#### 4.1.1 Das Ideal des Zufallsexperimentes

Bevor das RKM dargestellt wird, soll zunächst einerseits aufgezeigt werden, was Experimente zur Analyse kausaler Zusammenhänge so attraktiv macht und andererseits warum sie gerade im Bereich der Sozialwissenschaften dennoch nicht das „Nonplusultra“ der Kausalanalyse darstellen.

Das Forschungsdesign des Zufalls- oder randomisierten Experimentes ist im Wesentlichen von folgenden Eigenschaften gekennzeichnet (Chao-Ying 2003; Rosenbaum 2002). Das erste zentrale Element ist die Aufteilung der Untersuchungseinheiten nach dem Zufallsprinzip in mindestens zwei Gruppen,<sup>20</sup> die Treatment- und die Kontrollgruppe. Im Anschluss an die Gruppenbildung findet zweitens eine geplante Veränderung in einer der beiden Gruppen statt (per definitionem der Treatmentgruppe), wobei vom Experimentator eine genau festgelegte Variable manipuliert, d. h. in einem gewünschten Sinne verändert wird, wodurch die potenzielle Ursache herbeigeführt wird. Die Gruppe, die diesem Eingriff (Treatment) nicht

<sup>20</sup> Es existiert eine Vielzahl von Varianten, z. B. Einzelgruppendesigns mit Vorher-nachher-Vergleich oder Mehrgruppendesign mit mehreren Treatments (Shadish/Cook/Campbell 2002).

ausgesetzt war, dient als „unbehandelte“ Kontrollgruppe. Schließlich wird drittens nach (bzw. vorzugsweise vor und nach) der Manipulation der Wert einer interessierenden Zielvariable (auch Outcome-, Response- oder Ergebnisvariable) gemessen, und es wird bestimmt, ob sich in Treatmentgruppe und Kontrollgruppe unterschiedliche Werte ergeben. Existiert ein Unterschied, z. B. in Form unterschiedlicher Mittelwerte in beiden Gruppen, dann kann dieser Unterschied als der kausale Effekt der Manipulation interpretiert werden, vorausgesetzt er überschreitet die statistische Grenze der Zufälligkeit.

Bei einem randomisierten Experiment ist damit relativ eindeutig zu belegen, ob ein bestimmter Sachverhalt die Ursache für einen anderen Sachverhalt sein kann. Die argumentative Stärke von Zufallsexperimenten liegt darin begründet, dass die Zuweisung der Untersuchungseinheiten in die Treatment- und Kontrollgruppe ausschließlich auf einem Zufallsmechanismus besteht (wobei die Wahrscheinlichkeit für den Erhalt des Treatments meist auf 0,5 gesetzt wird). Daher wird in Experimenten ausgeschlossen, dass sich Treatment- und Kontrollgruppe in weiteren Eigenschaften unterscheiden, die oben als Selektionsverzerrung oder Kompositionseffekte bezeichnet wurden. Selektionsverzerrungen stellen vor allem dann ein Problem dar, wenn sie dazu führen, dass die Untersuchungseinheiten sich bereits vor der Manipulation in einem Maße unterscheiden, das auch Einfluss auf die Zielvariable hat. Dann kann nicht mehr eindeutig beurteilt werden, ob ein Gruppenunterschied das Ergebnis der durchgeführten Manipulation ist oder (teilweise oder ganz) auf einen A-priori-Unterschied in den Experimentalgruppen zurückzuführen ist.

Eine zufällige Aufteilung der Einheiten auf die Experimentalgruppen verhindert solche Unsicherheiten. Sie führt dazu, dass alle Eigenschaften der Untersuchungseinheiten die gleiche Wahrscheinlichkeitsverteilung aufweisen und die beiden Gruppen daher – zumindest im Erwartungswert der Eigenschaften – sich nicht voneinander unterscheiden, also vollkommen vergleichbar sind. Diese Vergleichbarkeit gilt sowohl hinsichtlich beobachteter Eigenschaften als auch hinsichtlich nicht beobachteter oder auch nicht beobachtbarer Merkmale, zu denen der Experimentator keinen Zugang hat. Da die Vergleichbarkeit nur im Mittel besteht, können zwischen den Gruppen immer noch Unterschiede auftreten, diese beruhen dann aber auf zufälligen Abweichungen und werden das Ergebnis der Zielvariablenmessung nicht systematisch verzerren. Sie können mithilfe statistischer Tests und Konfidenzintervalle berücksichtigt werden.

Diese Eigenschaft des „Ausbalancierens“ der Merkmale der Untersuchungseinheiten macht das Zufallsexperiment zum idealen Instrument der Identifikation kausaler Effekte. Vor allem in den USA werden daher randomisierte Experimente auch zur Messung von Effekten arbeitsmarktpolitischer Maßnahmen eingesetzt. Dabei wird versucht, Fragen nach der Wirkung von Programmen durch *soziale* Experi-

mente zu beantworten. Als soziale Experimente werden solche Vorgehensweisen bezeichnet, in denen die Untersuchungseinheiten Menschen sind und bei denen es sich um Feldexperimente handelt (im Gegensatz zu Laborexperimenten wie dem zum Ultimatumspiel aus Abschnitt 3.2.1). Bei sozialen Experimenten erfolgt zwar die Zuweisung von Personen auf die Analysegruppen im Rahmen eines experimentell kontrollierten Zufallsprozesses. Sowohl die Manipulation als auch die Realisierung der Zielvariable erfolgen allerdings außerhalb der Kontrolle des Experimentators. Die arbeitsmarktpolitische Maßnahme als Manipulation wird meist nicht vom Experimentator selbst durchgeführt, sondern erfolgt durch die Arbeitsmarktakteure (Arbeitsvermittlung, Maßnahmeträger etc.) selbst. Ebenso realisiert sich die Zielvariable nicht in einer kontrollierten Experimentalsituation, sondern sie ergibt sich auf dem Arbeitsmarkt. So wird etwa Personen auf der Basis eines Zufallsprozesses eine bestimmte Maßnahme gewährt oder nicht, und im Anschluss wird beobachtet, wie sich ihre Erwerbsbiografie hinsichtlich der interessierenden Zielvariable (z. B. Ausstieg aus der Arbeitslosigkeit oder Beschäftigungsdauer) entwickelt.

Solche Experimente haben jedoch auch ihre Schwierigkeiten. Eine wichtige Einschränkung erfährt das soziale Experiment dadurch, dass es aus ethischen Gesichtspunkten nicht akzeptabel ist, unter ihrer Arbeitslosigkeit leidenden Personen, die womöglich nach einer Maßnahme wieder Arbeit finden könnten, diese zu verweigern<sup>21</sup> (Müntnich/Wießner 2002). Neben ethischen Bedenken können die Bedenken gegenüber sozialen Experimenten in solche unterteilt werden, die die interne und die externe Validität betreffen, sowie in Bedenken zur Konstruktvalidität (Shadish/Cook/Campbell 2002).

Die Konstruktvalidität bezieht sich auf die Frage, ob die Manipulation in einem Experiment eine korrekte Operationalisierung der interessierenden potenziellen Ursache darstellt. Bei sozialen Experimenten werden die Manipulationen nicht eigens konstruiert, sondern es wird die Wirkung der konkreten in der Praxis durchgeführten Maßnahmen analysiert. Dadurch ist es wohl bedingt, dass die Frage der Konstruktvalidität kaum thematisiert wird. Im Gegenteil liegt der Fokus oft auf der Feststellung eines bzw. keines kausalen Effektes der konkreten Maßnahme, ganz gleich worauf dieser Effekt beruht. Allerdings stellt sich gerade deshalb, weil die konkrete Ausgestaltung der Maßnahme nicht der Kontrolle der Forscher unterliegt, die Frage, was es denn ist, dessen kausaler Effekt gemessen wurde (diese Frage betrifft nicht nur soziale Experimente, sondern auch Beobachtungsstudien, weshalb der Aspekt der Konstruktvalidität auch weiter unten gesondert thematisiert wird, vgl. Abschnitt 5.2.3).

21 Oft wird daher denjenigen, die im Rahmen der zufälligen Aufteilung keine Maßnahme erhalten, die Maßnahmeteilnahme zu einem späteren Zeitpunkt ermöglicht. Dennoch verbleiben sie – zumindest für eine gewisse Zeit – ohne diese eventuell nötige Maßnahme.

Die interne Validität bezieht sich auf die Frage, ob es dem Experiment gelingt, in der Experimentalsituation den kausalen Effekt korrekt zu identifizieren, also ob der gemessene Effekt auch auf die durchgeführte Manipulation zurückgeht. Es ist nicht auszuschließen, dass sich Personen, denen die Teilnahme an einer Maßnahme durch die Zufallsentscheidung verweigert wurde, nach einem Ersatz umsehen und an anderer Stelle an einer ähnlichen oder derselben Maßnahme teilnehmen. Dies kann zu einer Verzerrung bei der Messung führen, die als substitution bias bezeichnet wird (Heckman/Smith 1996: 49 f.).

Eine weitere Beschränkung der internen Validität tritt auf, wenn das Experiment derart in den Arbeitsmarkt eingreift, dass es die Erfolgsbedingungen von Teilnehmern oder der Kontrollgruppe beeinflusst. Teilt man beispielsweise auf einem eher begrenzten lokalen Arbeitsmarkt eine hinreichend große Gruppe Arbeitsloser in Maßnahmeteilnehmer, die eine Weiterbildungsmaßnahme durchlaufen, also dem Arbeitsmarkt nicht zur Verfügung stehen, und eine Kontrollgruppe, die direkt nach Arbeit sucht, und vergleicht man dann den Arbeitsmarkterfolg beider Gruppen, kann dies zu Problemen der internen Validität der Ergebnisse führen. Allein die Tatsache, dass sich das Arbeitsangebot durch die Herausnahme der Teilnehmer aus dem Arbeitsmarkt verringert hat, kann dazu führen, dass nun auf dem lokalen Arbeitsmarkt die Nichtteilnehmer eine bessere Chance auf Arbeit haben als vorher. Ähnlich können sich auch intervenierende Ereignisse auswirken, z. B. wenn während der Laufzeit der Maßnahme aus konjunkturellen Gründen die Nachfrage nach Arbeitskräften steigt und die Nichtteilnehmer davon profitieren.

Auch wenn es gelingt, in der Experimentalsituation den kausalen Effekt zu identifizieren, können die Ergebnisse nicht immer auf Situationen außerhalb des Experimentes übertragen werden (externe Validität). Vor allem spielt die Tatsache, dass die Untersuchungseinheiten rationale (und manchmal auch irrationale) Individuen sind, die auf Veränderungen in ihrer Umwelt reagieren, eine gewichtige Rolle. So reagieren manche Menschen etwa darauf, dass sie Teil eines Zufallsexperimentes sind (bzw. sein sollen). Zum Beispiel werden sich bestimmte Personen, vor allem solche, die sich auf das Risiko eines Zufallsexperimentes nicht einlassen wollen, schon von vornherein nicht bereit erklären, an einem solchen Experiment teilzunehmen. Dies wird als randomization bias bezeichnet (Heckman/Smith 1996: 48 f.), der dazu führt, dass Experimente an einer Population durchgeführt werden, die eventuell nicht der Zielgesamtheit entspricht.

Manski (1995: 53 f.) weist zudem darauf hin, dass es in sozialen Experimenten keine Double-blind-Situation gibt. Das heißt, es ist nicht zu erwarten, dass weder diejenigen, die die Weiterbildung durchführen, noch die Teilnehmer am Experiment im Dunkeln darüber gelassen werden können, wer eine Maßnahme erhält. Dies öffnet Versuchsleiter- und Versuchspersoneneffekten die Türe, etwa indem sich

Nichtteilnehmer besonders bemühen, um zu zeigen, dass sie es auch ohne unterstützende Maßnahme schaffen. Auch der Unterschied von lokaler Erprobung einer Maßnahme im Experiment und einer flächendeckenden Durchführung kann bei Beurteilung ihrer Wirkung von Bedeutung sein. Schließlich kann auch der Kontext der Durchführung eines Experimentes die externe Validität der experimentellen Kausalanalyse entscheidend beeinflussen. In einer Region, in der Arbeitslosigkeit in hohem Maße auf einem Qualifikationsmismatch beruht, wird etwa eine Weiterbildung anders wirken als in einer Region mit hoher konjunktureller Arbeitslosigkeit.

Trotz der hohen Relevanz von Experimenten bei der Ermittlung von Kausaleffekten lohnt es sich also, sich auch mit nicht experimentellen Methoden auseinanderzusetzen. Diese Methoden versuchen unter Bedingungen, die eine Durchführung von Experimenten unmöglich oder zu aufwendig bzw. kostenintensiv machen, kausale Effekte mithilfe von Beobachtungsdaten (Befragungen, Daten der amtlichen Statistik etc.) zu identifizieren.

Rubins Kausalmodell dient nun dazu, in Situationen, in denen aus unterschiedlichen Gründen ein Zufallsexperiment unmöglich, zu aufwendig oder zu kostenintensiv ist, kausale Effekte mithilfe von Beobachtungsdaten zu identifizieren, indem versucht wird, auf statistische Weise eine dem Zufallsexperiment entsprechende Vergleichbarkeit von Treatment- und Kontrollgruppe herzustellen. Solche Analysen werden in Abgrenzung zum Experiment als Beobachtungsstudien (*observational studies*) bezeichnet (Cochran 1965: 234). In der psychologischen Forschungstradition wird auch der Ausdruck *Quasi-Experiment* verwendet (Shadish/Cook/Campbell 2002).

Das RKM besteht im Kern aus zwei miteinander verbundenen Konzepten, dem Ansatz der potenziellen Ergebnisse und dem Zuweisungsmechanismus. Der Ansatz potenzieller Ergebnisse (Abschnitt 4.1.2) formuliert dabei das am Experiment orientierte Kausalitätskonzept, und das Konzept des Zuweisungsmechanismus (Abschnitt 4.1.3) liefert die Begründung für die Möglichkeit der Annäherung an Zufallsexperimente durch Beobachtungsstudien. Das Vorliegen der „strong ignorability“ oder „conditional independence“, wie ein in der Arbeitsmarktforschung gebräuchlicher Ausdruck lautet, ist dabei der zentrale Gradmesser für die interne Validität der gemessenen Ergebnisse. Die Vergleichbarmachung von Treatment- und Kontrollgruppe bei nicht zufälliger Aufteilung erfolgt im RKM dann durch die Methode des Matchings (Abschnitt 4.1.4). Die folgende Darstellung des RKM erfolgt zur Verdeutlichung des Konzeptes bereits mit Bezug auf die hier zu analysierende potenzielle Ursache, der Kombilohnförderung.

#### 4.1.2 Potenzielle Ergebnisse

Das erste für das RKM grundlegende Konzept ist der Ansatz der potenziellen Ergebnisse (Rubin 2005, 2007). Der Ansatz geht davon aus, dass es bei dem betrach-

teten Treatment „Kombilohnförderung“ ( $kl$ ) nur zwei Möglichkeiten gibt, entweder Person  $i$  erhält die Förderung ( $kl_i = 1$ ), oder sie erhält sie nicht ( $kl_i = 0$ ). Das RKM unterscheidet den Wert der Zielvariable, im Folgenden auch Ergebnis- oder Wirkungsvariable (Lohnzufriedenheit bzw. Beschäftigungsstabilität)  $y_i$ , der bei Person  $i$  im Fall des Kombilohnerhalts realisiert wird ( $y_i^1$ ), und den Wert der Zielvariable, der ohne Kombilohn realisiert wird ( $y_i^0$ ). Die Zielvariable kann sowohl kardinal- als auch ordinal- oder nominalskaliert sein. Zum Zweck der Darstellung des Konzeptes wird zunächst von einer kardinalen Variable ausgegangen. Um die Auswirkung für das Individuum  $i$  (oder ein anderes Individuum) zu bestimmen, muss der Wert der Ergebnisvariable bei Ausführen der Intervention „Kombilohn“ mit dem Wert bei ihrer Unterlassung verglichen werden, meist in Form einer Differenz der beiden potenziellen Ergebnisse  $\delta \equiv y_i^1 - y_i^0$ .<sup>22</sup> Diese Differenz gibt den kausalen Effekt der Förderung bei Person  $i$  an, den sogenannten Einheitseffekt (unit effect).

Ein wesentliches Problem bei der Bestimmung des individuellen kausalen Effektes ist, dass sich empirisch nur einer der beiden hypothetischen Ergebnisse beobachten lässt, da ein Individuum zu einem gegebenen Zeitpunkt stets entweder einer Intervention oder Manipulation<sup>23</sup> „Kombilohn“ ausgesetzt ist oder nicht.

$$Y_{beob,i} = \begin{cases} y_i^0, & \text{falls } kl_i = 0 \\ y_i^1, & \text{falls } kl_i = 1 \end{cases} \quad (4.1)$$

Dieser Sachverhalt wird als das Fundamentalproblem der Kausalanalyse bezeichnet (Holland 1986). Die kausale Auswirkung einer Manipulation wird damit nie an einem Individuum beobachtet werden können. Es existiert allerdings eine statistische Lösung für das Problem, indem statt des nicht zu messenden individuellen Effektes der durchschnittliche oder mittlere kausale Effekt (MKE) in einer Gesamtheit der Individuen betrachtet wird.

Unter Einbezug der Eigenschaften  $\mathbf{x}$  der Individuen ergibt sich dann der mittlere kausale Effekt für ein Individuum aus der Teilpopulation mit Charakteristika  $\mathbf{x}$  als Erwartungswert

$$E(\delta_i | \mathbf{x}_i) = MKE(\mathbf{x}_i) = E(y_i^1 | \mathbf{x}_i) - E(y_i^0 | \mathbf{x}_i) \quad (4.2)$$

Der  $MKE(\mathbf{x}_i)$  gibt den zu erwartenden, also durchschnittlichen kausalen Effekt des Kombilohnes bei einer zufällig gezogenen Personen in der Gesamtheit an, die die

22 Möglich, wenn auch kaum gebräuchlich, wäre etwa eine Quotientenbildung:  $\delta'_i \equiv y_i^1 / y_i^0$ .

23 Zum Manipulationsbegriff später mehr (Abschnitt 4.2.1).

konkreten Merkmale  $\mathbf{x}_i$  aufweist, z. B. für eine Person, die männlich, 32 Jahre alt, Facharbeiter etc. ist. Es handelt sich daher um einen bedingten kausalen Effekt. Um den unbedingten kausalen Effekt in der Gesamtheit zu berechnen, wird nochmals der Erwartungswert über alle  $\mathbf{x}$ -bedingten, teilpopulationsspezifischen kausalen Effekte gebildet (Rosenbaum/Rubin 1983).

$$E(\delta_i) = MKE = E_{\mathbf{x}}(MKE(\mathbf{x}_i)) = E_{\mathbf{x}}(E(y_i^1 | \mathbf{x}_i) - E(y_i^0 | \mathbf{x}_i)) \quad (4.3)$$

Da die Auswirkung des Kombilohnes auf die Zielvariable mit den Charakteristika der Individuen variieren kann, können sich die bedingten kausalen Effekte und der unbedingte kausale Effekt unterscheiden (man spricht hier von heterogenen Effekten, siehe etwa Pfeifer 2007). Der mittlere kausale Effekt kann in zwei, jeweils auch getrennt für sich interessante Effekte zerlegt werden, in den mittleren kausalen Effekt der Kombilohnförderung bei den Geförderten ( $MKE_T$ ) und den mittleren kausalen Effekt der Nichtförderung bei den Ungeförderten ( $MKE_{NT}$ ).

Der  $MKE_T$  misst, welche Auswirkung der Kombilohn im Mittel in der *Population der Geförderten* ( $kl_i = 1$ ) hat. Hier werden also nur die Personen betrachtet, welche die Förderung erhalten haben. Für die Teilpopulation mit Charakteristika  $\mathbf{x}$  gilt

$$E(\delta_i | \mathbf{x}_i, kl_i = 1) = MKE_T(\mathbf{x}_i) = E(y_i^1 | \mathbf{x}_i, kl_i = 1) - E(y_i^0 | \mathbf{x}_i, kl_i = 1) \quad (4.4)$$

und entsprechend für die gesamte Population

$$E(\delta_i | kl_i = 1) = MKE_T = E_{\mathbf{x}|kl=1}(MKE_T(\mathbf{x}_i)) = E_{\mathbf{x}|kl=1}(E(y_i^1 | \mathbf{x}_i, kl_i = 1) - E(y_i^0 | \mathbf{x}_i, kl_i = 1)) \quad (4.5)$$

Der  $MKE_T$  gibt an, welche Veränderung von  $y$  sich bei den Geförderten durch die Förderung ergeben hat.  $E(y_i^1 | \mathbf{x}_i, kl_i = 1)$  ist der Wert der Zielvariable der Geförderten im Fall der Förderung und damit ein beobachtbarer oder faktischer Wert. Dagegen ist der Erwartungswert der Zielvariable im Fall der Nichtförderung bei den Geförderten  $E(y_i^0 | \mathbf{x}_i, kl_i = 1)$  nicht beobachtbar und wird daher oft als kontrafaktischer Wert bezeichnet.<sup>24</sup> Er charakterisiert die Situation, die sich ergeben hätte, wenn die Geförderten – bei sonst gleichen Bedingungen – doch keinen Kombilohn erhalten hätten.

Sofern die geförderten Personen sich hinsichtlich ihrer Zusammensetzung von den nicht geförderten Personen zumindest teilweise unterscheiden, messen die

<sup>24</sup> Rubin selbst lehnt die Bezeichnung faktisch und kontrafaktisch ab und bevorzugt den Ausdruck „potential outcomes“.

Größen  $MKE$  und  $MKE_T$  unterschiedliche kausale Parameter, ansonsten liefern sie identische Werte.

Der  $MKE_{NT}$  bildet das Komplement zum  $MKE_T$  für die *Population der Ungeförderten* ( $kl_i = 0$ ). Entsprechend ist er für die Teilpopulation mit Charakteristika  $\mathbf{x}$  gegeben als

$$MKE_{NT}(\mathbf{x}_i) = E(y_i^1 | \mathbf{x}_i, kl_i = 0) - E(y_i^0 | \mathbf{x}_i, kl_i = 0) \quad (4.6)$$

und für die gesamte Population als

$$MKE_{NT} = E_{\mathbf{x}|kl=0}(MKE_{NT}(\mathbf{x}_i)) = E_{\mathbf{x}|kl=0}(E(y^1 | \mathbf{x}_i, kl_i = 0) - E(y^0 | \mathbf{x}_i, kl_i = 0)) \quad (4.7)$$

In der Praxis der Messung kausaler Maßnahmeeffekte bildet meist der  $MKE_T$  die relevanteste Information, da er Aufschluss darüber gibt, ob die Maßnahme speziell bei den Geförderten etwas bewirkt hat.  $MKE_T$  und  $MKE_{NT}$  bilden beide jeweils einen Teilaspekt des gesamten mittleren kausalen Effektes ab. Sie stehen im folgenden Zusammenhang:

$$MKE(\mathbf{x}_i) = \underbrace{\left[ E(y_i^1 | \mathbf{x}_i, kl_i = 1) - E(y_i^0 | \mathbf{x}_i, kl_i = 1) \right]}_{MKE_T(\mathbf{x}_i)} P(kl_i = 1 | \mathbf{x}_i) + \underbrace{\left[ E(y_i^1 | \mathbf{x}_i, kl_i = 0) - E(y_i^0 | \mathbf{x}_i, kl_i = 0) \right]}_{MKE_{NT}(\mathbf{x}_i)} P(kl_i = 0 | \mathbf{x}_i) \quad (4.8)$$

Die Gleichung 4.8 kann allerdings nur dann als kausaler Effekt der Förderung interpretiert werden, wenn einige Bedingungen erfüllt sind. Eine erste Bedingung betrifft die Konstruktvalidität in Form der prinzipiellen Manipulierbarkeit der kausal interessierenden Variable, also hier dem Erhalt einer Kombilohnförderung. Hierzu erfolgen weiter unten noch einige klärende Analysen (Abschnitt 4.2). Darunter ist zu verstehen, dass zu einem bestimmten Zeitpunkt alle Personen – sowohl der Treatment- als auch der Kontrollgruppe – das Treatment hätten erhalten können („potential exposability to treatment“). Bei Maßnahmeteilnahmen, Medikamenteneinnahmen, Teilnahmen an speziellen Therapieformen etc. ist diese Bedingung oft gegeben, bei vielen anderen soziologisch und ökonomisch relevanten Merkmalen wie Geschlecht, Nationalität oder Alter ist dies nicht so. Ebenfalls verletzt ist die Bedingung, wenn es sich bei einer Maßnahme für eine bestimmte Personengruppe verpflichtende Maßnahme handelt, sodass es keine Vergleichsgruppe von Personen gibt, die prinzipiell hätten teilnehmen können, dies aber aus irgendwelchen Grün-



den nicht taten.<sup>25</sup> In einem solchen Fall ist keine Vergleichbarkeit beider Gruppen herstellbar.

Eine zentrale Bedingung für die interne Validität der Messung des kausalen Effektes ist die „Stable Unit Treatment Value Assumption“ oder kurz SUTVA (Rubin 1980a). Die Annahme besagt, dass erstens die Tatsache, ob Individuum  $i$  ein Treatment erhält ( $kl_i = 1$ ) oder nicht ( $kl_i = 0$ ), sich nicht auf den Wert  $y_i^0$  oder  $y_i^1$  von Individuum  $j$  auswirkt. So könnte es etwa sein, dass die Handlung „Kauf eines Neuwagens“ bei Person  $a$  sich auf die Zielvariable „allgemeine Lebenszufriedenheit“ bei  $b$  auswirkt, wenn  $i$  von  $j$ 's Autokauf weiß. Zweitens ist die SUTVA auch dann verletzt, wenn unterschiedliche Versionen einer Maßnahme existieren, die jeweils einen anderen Wert der Zielgröße bewirken. Rubin spricht in solchen Fällen von unterschiedlichen Zuordnungsmechanismen (Rubin 1986: 961) oder „hidden versions“ (Rubin 2005: 323) des Treatments. Es ist durchaus möglich, dass unfreiwillige gegenüber freiwilligen Maßnahmeteilnahmen eine solche versteckte Version des Treatments darstellen. Zum Beispiel können freiwillige Teilnahmen an einer Trainingsmaßnahme zum Schreiben von Bewerbungen eine andere Wirkung auf die Arbeitslosigkeitsdauer haben als unfreiwillige Teilnahmen. Mit dieser Voraussetzung wird vermieden, dass es mehr als nur ein potenzielles Ergebnis je Treatmentstatus gibt, etwa

$$svlr_{obs,i} = \begin{cases} y_i^0, & \text{falls } kl_i = 0 \\ y_i^{1,1}, & \text{falls } kl_i = 1, \text{ oder} \\ y_i^{1,2}, & \text{falls } kl_i = 1, \text{ oder} \\ y_i^{1,3}, & \text{falls } kl_i = 1, \dots \end{cases}$$

Beide Möglichkeiten der SUTVA-Verletzungen lassen sich am Beispiel einer Medikamenteneinnahme verdeutlichen. Die Tatsache, dass Person  $i$  ein schmerzlinderndes Medikament eingenommen hat, sollte nicht beeinflussen, ob Person  $j$  eine Schmerzlinderung erfährt oder nicht. Ebenso sollte es keinen Unterschied machen, ob das Medikament als Tablette oder Spritze eingenommen wird, ob man sich das Medikament selbst besorgt oder ob es von einem Arzt verschrieben wird etc.

Die Überprüfung der SUTVA ist sehr schwierig, und so kann ihre Erfüllung meist nur argumentativ begründet werden. Während der erste Aspekt der SUTVA bei der

25 In einem solchen Fall bietet sich ein Regression-Diskontinuity-Modell an (Shadish/Cook/Campbell 2002; Hahn/Todd/Van der Klaauw 2001).

Evaluation aktiver Arbeitsmarktpolitik bisher in Form der Problematisierung von allgemeinen Gleichgewichts-, Peer- (Heckman/Lochner/Taber 1998) oder Verdrängungseffekten (Kangasharju 2005) behandelt wurde, hat der zweite Aspekt bisher offenbar noch keine besondere Beachtung gefunden (vgl. dazu die Analysen in Abschnitt 6.5.2.2). Schließlich ist die letzte und besonders wichtige Bedingung der internen Validität die vollständige Erfassung des sogenannten Zuweisungsmechanismus (assignment mechanism) durch geeignete Variablen, worauf nun im Folgenden ausführlich eingegangen werden soll.

#### 4.1.3 Zuweisungsmechanismus und Conditional Independence Assumption

Wie in Gleichung 4.8 zu sehen ist, werden nicht individuelle Effekte, sondern mittlere kausale Effekte betrachtet. Dabei ist allerdings das Fundamentalproblem der Nichtbeobachtbarkeit beider Zustände noch nicht gelöst worden. Gleichung 4.8 für den MKE enthält zwei dieser nicht beobachtbaren Mittelwerte (und entsprechend die Gleichungen für den  $MKE_T$  und den  $MKE_{NT}$  jeweils einen):

$$MKE(\mathbf{x}_i) = \left[ \underbrace{E(y_i^1 | \mathbf{x}_i, k_i = 1)}_{\text{beobachtbar}} - \underbrace{E(y_i^0 | \mathbf{x}_i, k_i = 1)}_{\text{unbeobachtbar}} \right] \underbrace{P(k_i = 1 | \mathbf{x}_i)}_{\text{beobachtbar}} + \left[ \underbrace{E(y_i^1 | \mathbf{x}_i, k_i = 0)}_{\text{unbeobachtbar}} - \underbrace{E(y_i^0 | \mathbf{x}_i, k_i = 0)}_{\text{beobachtbar}} \right] \underbrace{P(k_i = 0 | \mathbf{x}_i)}_{\text{beobachtbar}} \quad (4.9)$$

Für den Fall der Maßnahmeteilnahme ist nur das potenzielle Ergebnis bei Person  $i$   $y_i^1 | k_i = 1$  beobachtbar (bzw. schätzbar), nicht aber das potenzielle Ergebnis  $y_i^0 | k_i = 1$ , das sich bei Person  $i$  ergeben würde, wenn die Maßnahme unterblieben wäre. Entsprechendes gilt auch umgekehrt für  $y_i^0 | k_i = 0$  und  $y_i^1 | k_i = 0$ . Um das Problem zu lösen, müssen die unbeobachtbaren Größen durch beobachtbare Größen ersetzt werden.

Eine Lösung für dieses Problem stellt das bereits in Abschnitt 4.1.1 dargestellte randomisierte Experiment dar. Zunächst ist  $MKE_{Exp} = E(y_i^1 - y_i^0) = E(y_i^1) - E(y_i^0)$  der mittlere kausale Effekt, wenn die Zuweisung der Personen in die Gruppen der Kombilohnbezieher und der Nichtbezieher im Hinblick auf die Zielvariable zufällig erfolgt. Dies kann der Fall sein, wenn im Rahmen eines Experimentes eine zufällige Aufteilung von Personen in eine Gruppe erfolgt, welche die Maßnahme erhält, und eine weitere, für die dies nicht gilt. Im Experiment gilt durch die zufällige Aufteilung der Gruppen, dass die Individuen keine systematischen Unterschiede aufweisen. Kombilohnersatz und Zielvariable sind dann stochastisch unabhängig. Formal wird

dies im Anschluss an Dawid (1979) in folgender Notation ausgedrückt:  $y_i^0, y_i^1 \perp kl_i$ . Stochastische Unabhängigkeit bedeutet, dass beide Gruppen *hinsichtlich der beobachteten ebenso wie der unbeobachteten* Variablen die gleiche Wahrscheinlichkeitsverteilung aufweisen und damit zumindest im Mittel vollkommen vergleichbar sind, sodass der Erwartungswert der Zielvariable unabhängig vom Maßnahmestatus stets gleich ist, wodurch die einzelnen Werte miteinander austauschbar sind:

$$E(y_i^0 | kl_i = 0) = E(y_i^0 | kl_i = 1) = E(y_i^0) \text{ bzw. } E(y_i^1 | kl_i = 1) = E(y_i^1 | kl_i = 0) = E(y_i^1)$$

In diesem Fall kann der kausale Effekt problemlos aus dem Mittelwertvergleich geschätzt werden. Daher ist auch eine unverzerrte Schätzung über die Differenz der Mittelwerte möglich

$$MK\hat{E}_{Exp} = \hat{E}(y_i^1 - y_i^0) = \hat{E}(y_i^1) - \hat{E}(y_i^0) = \frac{1}{n_{i \in KL1}} \sum_{i=1}^n y_i^1 - \frac{1}{m_{j \in KLO}} \sum_{j=1}^m y_j^0 \quad (4.10)$$

wobei wiederum mit *KL1* der Set der  $i = 1, \dots, n$  Kombilohnbezieher und mit *KLO* der Set der  $j = 1, \dots, m$  ungeförderten Beschäftigten bezeichnet wird. Die stochastische Unabhängigkeit kann im Allgemeinen außerhalb eines Experimentes nur selten angenommen werden. Gerade weil sich Menschen bei der Wahl ihrer Handlungen die möglichen Konsequenzen vor Augen führen, werden Personen, die eine Förderung beantragen bzw. denen eine solche gewährt wird, andere Eigenschaften aufweisen als ungeförderte Beschäftigte. Üben diese auch Einfluss auf die Zielvariable aus, dann besteht Maßnahmeselektivität, und es gilt:

$$E(y_i^0 | kl_i = 0) \neq E(y_i^0 | kl_i = 1) \text{ bzw. } E(y_i^1 | kl_i = 1) \neq E(y_i^1 | kl_i = 0)$$

Dies führt zu Selektionsverzerrungen (selection bias) bei der Messung der kausalen Effekte. Deren Größenordnung hängt davon ab, wie weit der unbeobachtete Erwartungswert von dem beobachteten Erwartungswert entfernt ist, der ihn ersetzt:

$$bias_{MKET} = E(y_i^0 | kl_i = 1) - E(y_i^0 | kl_i = 0)$$

$$bias_{MKENT} = E(y_i^1 | kl_i = 0) - E(y_i^1 | kl_i = 1)$$

$$bias_{MKE} = bias_{MKET} + bias_{MKENT}$$

Bei Selektionsverzerrung ist die statistische Unabhängigkeit von *kl* und *y* verletzt. Die Forderung nach statistischer Unabhängigkeit kann allerdings durch die For-

derung nach *bedingter* statistischer Unabhängigkeit ersetzt werden. Dies wird oft mit dem englischen Ausdruck „Conditional Independence Assumption“ bezeichnet und mit CIA abgekürzt (z. B. Lechner 2002). Im RKM wird diese Forderung auch als „strongly ignorable treatment assignment“ bezeichnet (Rosenbaum/Rubin 1983: 45).

Bei bedingter stochastischer Unabhängigkeit gilt für  $y^1$  (und für  $y^0$  entsprechend):

$$E(y_i^1 | \mathbf{x}_i, k_i = 1) = E(y_i^1 | \mathbf{x}_i, k_i = 0) = E(y_i^1 | \mathbf{x}_i) \quad (4.11)$$

Damit lässt sich eine zu Gleichung 4.8 äquivalente Formulierung des MKE finden, der allein auf beobachtbaren Werten basiert.

$$\begin{aligned} MKE(\mathbf{x}_i) = & \left[ E(y_i^1 | \mathbf{x}_i, k_i = 1) - E(y_i^0 | \mathbf{x}_i, k_i = 0) \right] P(k_i = 1 | \mathbf{x}_i) \\ & + \left[ E(y_i^1 | \mathbf{x}_i, k_i = 1) - E(y_i^0 | \mathbf{x}_i, k_i = 0) \right] P(k_i = 0 | \mathbf{x}_i) \end{aligned} \quad (4.12)$$

Da es für die Validität der Kausalanalyse im RKM von entscheidender Bedeutung ist, dass die  $x$ -Variablen die CIA erfüllen, stellt sich die Frage, welche Eigenschaften diese hierzu aufweisen müssen, damit dies der Fall ist. Das RKM geht davon aus, dass die  $x$ -Variablen den sogenannten Zuweisungsmechanismus abbilden müssen. Als Heuristik dient dabei die Orientierung am Ideal des Zufallsexperimentes. Empirische Fragestellungen sollen demnach als eine Art misslungenes Zufallsexperiment (broken experiment) betrachtet werden (Rubin 2007: 25), wobei dem Zuweisungsmechanismus die Aufgabe zukommt, die wesentlichen Faktoren zu identifizieren, in denen der Gegenstand von der zufälligen Zuweisung in Teilnahme- und Kontrollzustand abweicht.<sup>26</sup> Gelingt dies, so kann davon ausgegangen werden, dass bei „Konstanthalten“ dieser Faktoren die Förderung oder Unterlassung (in Bezug auf das Untersuchungsziel!) nur noch zufällig, also statistisch unabhängig vom Maßnahmeziel erfolgt.

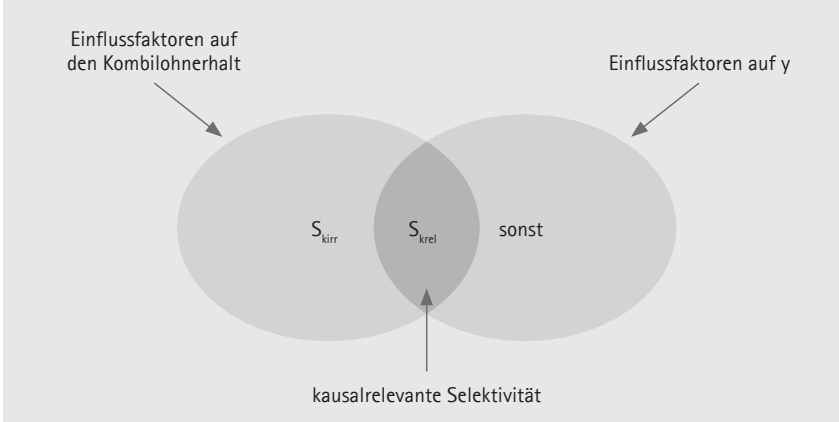
Der Zuweisungsmechanismus liefert diese Wahrscheinlichkeiten für den Treatmenterhalt, indem er den Prozess beschreibt, durch den Individuum  $i$  der Maßnahmenstatus zugewiesen wird und durch den damit gleichzeitig bestimmt wird, welches der potenziellen Maßnahmeergebnisse  $y^0$  oder  $y^1$  realisiert wird (Rubin 1991):

$$P(k_i = 1 | \mathbf{x}_i, y_i^1, y_i^0) = \prod_{j=1}^l f(\mathbf{x}_i)^{k_i} [1 - f(\mathbf{x}_i)]^{1-k_i} \quad \text{mit } l = n + m \quad (4.13)$$

<sup>26</sup> „We view the observed data as having arisen from a hypothetical complex randomized experiment with a lost rule for the propensity scores, whose values we will try to reconstruct“ (Rubin 2007: 25). Die bereits besprochene Randomisierung im Zufallsexperiment gilt damit als idealer, da zufälliger Zuordnungsmechanismus (Rubin 1978).

Die Wahrscheinlichkeit, (nicht) in die Treatmentgruppe zu gelangen, ist also eine Funktion der Eigenschaften  $\mathbf{x}$  der Individuen  $i$ . Die  $\mathbf{x}$  in diesem Zuweisungsmechanismus sind die Einflussfaktoren, die sowohl mit dem Treatment als auch mit der Zielvariablen korreliert sind (Rosenbaum/Rubin 1983). Rubin formuliert sein Konzept des Zuweisungsmechanismus dabei explizit ohne Bezug auf den jeweiligen Gegenstand der Forschung.<sup>27</sup> Der Zuweisungsmechanismus muss so gewählt sein, dass er alle systematischen Einflüsse – ganz gleich, worum es sich dabei handelt – auf die Zuweisung in den Treatmentstatus bzw. den Kontrollstatus enthält.

Abbildung 12: Variablenklassen



Betrachtet man den Sachverhalt der Maßnahmeselektivität als eine Konkretisierung des Zuweisungsmechanismus im Gegenstandsbereich der vorliegenden Untersuchung, dann wird deutlich, dass also nicht alle, sondern nur ein Teil der Einflussfaktoren auf die Kombilohnteilnahme relevant sind (vgl. Abbildung 12). Solche Einflüsse können auch als gemeinsame Ursachen<sup>28</sup> von Teilnahme und Zielvariable bezeichnet werden. Fehlt eine Variable oder mehrere dieser Variablen, so ist eine unverzerrte Schätzung des kausalen Maßnahmeeffektes nicht mehr möglich. Sie bilden daher den kausalrelevanten Aspekt der Selektivität, im Folgenden auch kurz  $S_{krel}$ . Diejenigen Einflussfaktoren, die zwar den Kombilohnerhalt bestimmen, jedoch nicht für die Realisation der Zielvariable relevant sind, sollen als kausalirrelevante Selektionseinflüsse  $S_{kirr}$  bezeichnet werden. In der ökonomischen oder statistischen Terminologie handelt es sich dabei um Instrumente (vgl. Abschnitt 4.3.2.2). Die  $x$ -Variablen müssen daher also „lediglich“ eine geeignete Operationalisierung

27 Dehejia und Wahba gehen dabei so weit zu behaupten, der Zuweisungsmechanismus sei vollkommen frei von Handlungsannahmen, „has no behavioral assumptions attached to it“ (Dehejia/Wahba 2002: 161).

28 Wobei damit natürlich ein anderer Ursachenbegriff impliziert ist als der im RKM, der sich nur auf das Treatment bezieht.

und Messung der  $S_{krel}$  darstellen, sowohl die  $S_{kirr}$  als auch die sonstigen Einflussfaktoren auf die Ergebnisvariable sind zufällig auf die Kombilohngefährdeten und Ungefährdeten verteilt. Wichtig für die Messung dieser Selektionseinflüsse ist, dass diese im „Pre-treatment“-Status, also vor dem der Manipulation, stattfindet. Andernfalls kann die Variable bereits von der Maßnahmeteilnahme verändert worden sein. Wenn eine Beeinflussung durch das Treatment ausgeschlossen ist, können Post-treatment-Variablen jedoch als Ersatz für nicht gemessene Pre-treatment-Variablen verwendet werden (Rosenbaum 1984a).<sup>29</sup>

#### 4.1.4 Matching von Gefährdeten und Ungefährdeten

Der mittlere kausale Effekt ist als Differenz zweier Erwartungswerte definiert. Um den kausalen Effekt in einem konkreten Anwendungsfall zu schätzen, braucht es ein adäquates Schätzverfahren. Die Schätzverfahren im RKM beruhen im Gegensatz zum Experiment auf einer *nachträglichen* Vergleichsgruppenbildung, wodurch die Gruppe der Gefährdeten und Ungefährdeten nicht a priori durch ein Zufallsexperiment, sondern im Nachhinein vergleichbar gemacht werden sollen. Dies geschieht durch das sogenannte Matching (Rosenbaum/Rubin 1983; Rubin/Thomas 1996; Rubin 2001). Dabei wird zu jedem Kombilohnbezieher mindestens ein möglichst ähnlicher Nichtbezieher und umgekehrt zu jedem Nichtbezieher mindestens ein möglichst ähnlicher Kombilohnbezieher gesucht. Man bezeichnet dies gelegentlich auch als die Suche nach „statistischen Zwillingen“ (Bacher 2002). Die Ähnlichkeit der Zwillinge wird durch einen Balancing Score bestimmt. „A balancing score  $b(\mathbf{x})$ , is a function of the observed covariates  $\mathbf{x}$  such that the conditional distribution of  $\mathbf{x}$  given  $b(\mathbf{x})$  is the same for treated ( $kl = 1$ ) and control ( $kl = 0$ ) units“ (Rosenbaum/Rubin 1983: 42, Notation angepasst).

Der differenzierteste Balancing Score, der diese Eigenschaften erfüllt, ist  $\mathbf{x}$  selbst. Gilt die Conditional Strong Ignorability bzw. Conditional Independence Assumption  $(y_i^0, y_i^1) \perp kl_i \mid \mathbf{x}_i$ , d. h., enthält  $\mathbf{x}$  alle kausalrelevanten Selektionseinflüsse und ist  $0 < P(kl_i = 1 \mid \mathbf{x}_i) < 1$  (common support condition), dann führt ein Matching von Teilnehmern zu Nichtteilnehmern (und umgekehrt) mit identischen Ausprägungen der  $\mathbf{x}$ -Variablen zu einer unverzerrten Schätzung des kausalen Effektes. Dabei wäre also zu jedem Gefährdeten eine ungefährdete Person zuzuordnen, die bezüglich der gemeinsamen Einflussfaktoren auf Kombilohnbezug und Zielvariable gleiche Eigenschaften aufweist, z. B. gleichen Lohn, gleiche Arbeitsbedingungen etc. In den meisten Fällen wird es allerdings schwierig sein, für jede Person einen

<sup>29</sup> Ist ein Post-treatment-Merkmal vom Treatment beeinflusst, kann es in bestimmten Fällen ein nicht beobachtetes oder schwierig zu messendes Maßnahmeziel ersetzen (Frangakis/Rubin 2002).

exakten Matchingpartner zu finden, da schon bei wenigen Variablen und Ausprägungen die Zahl der möglichen Kombinationen sehr groß wird. Statt des Matchings auf den Vektor  $\mathbf{x}$  kann allerdings auch auf einen bestimmten Skalar gemacht werden. Zur unverzerrten Schätzung des kausalen Effektes genügt es nämlich, wenn eine Person und ihr Matchingpartner denselben Propensity Score  $P(\mathbf{x}_i)$  aufweisen. Der Propensity Score (Rosenbaum/Rubin 1983; Dehejia/Whaba 2002) ist die bedingte Wahrscheinlichkeit der Maßnahmeteilnahme und besitzt ebenfalls die Eigenschaften eines Balancing Score, weist aber für jede Person nur einen einzigen Wert auf. Damit ist er der grösste Balancing Score.

$$e(\mathbf{x}_i) \equiv P(kl_i = 1 | \mathbf{x}_i) = P(\mathbf{x}_i)$$

Aufgrund der Eigenschaften des Propensity Score als Balancing Score erfüllt dieser auch die Strong Ignorability/Conditional Independence–Annahme, wenn in die Berechnung der Wahrscheinlichkeiten die entsprechenden Selektionseinflüsse eingehen:

$$(y_i^0, y_i^1) \perp\!\!\!\perp kl_i | e(\mathbf{x}_i) \text{ und}$$

$$0 < P(kl_i = 1 | \mathbf{x}_i) < 1$$

Entsprechend lässt sich etwa der  $MKE_T$  analog zur Gleichung 4.6 (für MKM und  $MKM_{NT}$  gilt Entsprechendes) auch mit Bezug auf den Propensity Score (PS) bestimmen:

$$MKE_T = E_{e(\mathbf{x}_i)} \left[ E(y_i^1 | kl_i = 1, e(\mathbf{x}_i)) - E(y_i^0 | kl_i = 1, e(\mathbf{x}_i)) \right] \quad (4.14)$$

Diese Konditionierung auf den Propensity Score wird erreicht, indem jedem Geförderten mit konkretem Propensity Score mindestens ein Ungeförderter mit gleichem Propensity Score zugeordnet wird. Damit ergibt sich als Matchingschätzer für den mittleren kausalen Effekt der Teilnahme bei den Teilnehmern:

$$M\hat{K}E_T = \frac{1}{n_1} \sum_{i \in I_1} \left( (y_i^1) - \sum_{j \in I_0} w(i, j) (y_j^0) \right) = \frac{1}{n_1} \sum_{i \in I_1} (y_i^1) - \frac{1}{n_1} \sum_{i \in I_1} \sum_{j \in I_0} w(i, j) (y_j^0) \quad (4.15)$$

Dabei bezeichnet  $I_1$  den Set von  $i = 1, 2, \dots, n$  Geförderten und  $I_0$  entsprechend den Set von  $j = 1, 2, \dots, m$  ungeförderten, regulär Beschäftigten.

Da der PS eine metrische Größe ist, ist eine exakte Übereinstimmung selten möglich. Dies führt zu einer Vielzahl von Varianten des Propensity Score Matchings,

die sich in dem Algorithmus zur Bestimmung des oder der Matchingpartner(s) unterscheiden. Je nach Algorithmus wird das Gewicht  $w(i, j)$  im Matchingschätzer für den kausalen Effekt anders festlegt.

Die einfachste Möglichkeit, einen geeigneten Matchingpartner zu finden, ist das Nearest Neighbour Matching (NN-Matching) ohne Zurücklegen. Zunächst werden die Geförderten in eine zufällige Reihenfolge gebracht. Dann wird dem ersten Geförderten genau der regulär Beschäftigte zugeordnet, dessen Propensity Score  $P(kl_j = 1 | \mathbf{x}_j) = P(\mathbf{x}_j) = P_j$  am nächsten an dem Propensity Score des Geförderten  $P(kl_i = 1 | \mathbf{x}_i) = P(\mathbf{x}_i) = P_i$  liegt. Im Fall ohne Zurücklegen wird der Ungeförderte aus dem Pool der möglichen Matchingpartner entfernt, und es wird ein NN-Matchingpartner für den nächsten Geförderten gesucht. In dieser Matchingvariante ist dann  $w(i, j) = 1$  für die Person  $j$  mit dem zur Person  $i$  ähnlichsten Propensity Score und  $w(i, j) = 0$  für alle anderen Personen.

Da nicht automatisch der nächste Nachbar auch ein guter Matchingpartner in dem Sinne sein muss, dass er auch einen ähnlichen Propensity Score aufweist, wird oft eine maximale Distanz  $|P_i - P_j|$  festgelegt, bis zu der Personen als Matchingpartner infrage kommen. Diese Distanz wird als Caliper  $c$  bezeichnet und das Vorgehen entsprechend als Caliper-Matching (C-Matching). Es lässt sich keine allgemeingültige Vorschrift zur Wahl des Calipers angeben, Rosenbaum und Rubin (1985; vgl. auch Cochran/Rubin 1973) zeigen allerdings, dass mit  $c = 0,2 \sqrt{(\sigma_{P_i}^2 + \sigma_{P_j}^2)/2}$  eine Bias-Reduktion um 98–99 % erreicht wird.<sup>30</sup> Das NN-Matching kann dabei so durchgeführt werden, dass zu jeder Person jeweils nur eine Kontrollperson hinzugegematcht wird oder dass – sofern im Bereich des Calipers – mehrere Kontrollpersonen zu einer Person gematcht werden. Mit  $\sum_{j \in I_0} w(i, j)$  wird dann eine gewichtete Summe der Zielvariable erzeugt, wobei die Gewichte der jeweiligen Matchingpartner häufig von ihrer PS-Distanz zum Propensity Score des Geförderten abhängig sind.

Eine Erweiterung des C-Matchings ist das Radius-Matching (R-Matching). Hier wird ein bestimmter Mindestabstand  $r$  festgelegt, und es werden alle Kontrollpersonen als Matchingpartner verwendet, bei denen der Betrag der Differenz  $|P_i - P_j| \leq r$  ist. Ähnlich funktioniert das Kernel-Matching (Heckman et al. 1997, 1998). Hier basiert die Gewichtung auf einer – genauer zu spezifizierenden – Kernelfunktion  $K$ :

$$w(i, j) = K\left(\frac{P_j - P_i}{b}\right) / \sum_{k \in I_0} K\left(\frac{P_k - P_i}{b}\right)$$

30 Dabei steht  $\sigma^2$  für die jeweiligen Varianzen der Propensity Scores. Dieses Ergebnis bezieht sich allerdings auf eine spezielle Variante des Matchings.



Dabei ist  $b$  eine Bandbreite, die festlegt, bis zu welchem Abstand eine Gewichtung mit der Kernfunktion stattfindet. Statt des Propensity Scores können auch andere Balancing Scores verwendet werden.<sup>31</sup>

Sind die Wahrscheinlichkeiten für die Maßnahmeteilnahme nicht bekannt – das wird in praktischen Anwendungen meistens der Fall sein –, können sie aus den Daten geschätzt werden. Im Allgemeinen werden die Wahrscheinlichkeiten mit einer logistischen oder Probitregression modelliert, diese liefert:

$$P(kl_i = 1 | \mathbf{x}_i) = \left(1 + \exp(-\mathbf{x}_i' \boldsymbol{\beta})\right)^{-1} \text{ bzw. } P(kl_i = 1 | \mathbf{x}_i) = \Phi(\mathbf{x}_i' \boldsymbol{\beta})$$

Ist in der Stichprobe, aus der die Wahrscheinlichkeiten berechnet werden soll, eine der beiden Gruppen überrepräsentiert (meist wird dies die Gruppe der Geförderten sein), dann können anstelle der PS auch die Odds zum Matching verwendet werden (Smith/Todd 2005). Die Odds sind eine Funktion der Propensity Scores:

$$\text{Odds} = \frac{P(kl_i = 1 | \mathbf{x}_i)}{P(kl_i = 0 | \mathbf{x}_i)}$$

Sie haben einen Wertebereich von 0 bis  $+\infty$ , wobei Werte von 0 bis unter 1 auf eine höhere Wahrscheinlichkeit der Nichtförderung verweisen und Werte von über 1 auf eine höhere Förderwahrscheinlichkeit. Um diese Asymmetrie zu beseitigen, die vor allem beim Caliper-Matching zu Problemen führt (ein Caliper von 0,01 im Bereich zwischen 0 und 1 ist mit einem Caliper von 0,01 im oberen Bereich ab 1 nicht zu vergleichen), werden die Odds logarithmiert (Log Odds). Dann liegt ihr Wertebereich zwischen  $-\infty$  und  $+\infty$  mit 0 bei gleichen Wahrscheinlichkeiten für Förderung und Nichtförderung.

Zentrale Eigenschaft des Matchings ist, dass dieses Verfahren, im Gegensatz etwa zu Regressionen, nur geförderte und ungeförderte Personen verwendet, die einander ähnlich sind. Beim Matchingvorgang werden daher solche Geförderte aus der Analyse ausgeschlossen, für die sich kein Ungeförderter mit ähnlichen Eigenschaften finden lässt. Es muss betont werden, dass, obwohl sich gematchte Individuen in den PS ähneln, zwischen ihnen dennoch Unterschiede in den  $\mathbf{x}$  bestehen bleiben können. Allerdings gleichen sich die Verteilungen der  $\mathbf{x}$  in den beiden Gruppen durch das Matching von Individuen mit ähnlichem PS einander an, im Idealfall entsprechen sie einander vollständig. Dies hat den Vorteil, dass nun nur noch zwei Gruppen mit ähnlicher Zusammensetzung miteinander verglichen werden. Die Konsequenz davon ist allerdings auch, dass sich die Ergebnisse des Vergleichs, also

<sup>31</sup> Eine verbreitete Variante ist das Mahalanobis-Distanz-Matching (Rubin 1980b).

der kausale Effekt, auf die aktuelle Zusammensetzung der Geförderten beziehen und sich eventuell verändern können, falls sich deren Zusammensetzung verändert.

Da die Prüfung der Ausgeglichenheit der beiden Gruppen erfolgen kann, ohne gleichzeitig schon den kausalen Effekt zu berechnen, kann – im Gegensatz etwa zu Regressionsanalysen – die logistische Regression so lange mit Merkmalen und/oder ihren nicht linearen Transformationen, wie etwa Log-Transformation, Quadrierung und Interaktionstermen, angepasst werden, bis der Ausgleich erreicht wurde, ohne dass damit eine – eventuell unbewusste – Suche nach der Spezifikation stattfindet, die einen signifikanten Wert aufweist. Erst nach zufriedenstellendem Ausgleich wird dann der kausale Effekt berechnet (zu den Möglichkeiten zur Prüfung der Ausgleichsqualität vgl. Abschnitt 6.3). Allerdings darf ein erfolgter Ausgleich der Merkmalsverteilungen von Geförderten und Ungeförderten nicht mit einem Beleg für die Gültigkeit der CIA verwechselt werden. Mit dem Matchingverfahren werden jeweils nur die Merkmale ausgeglichen, die vom Forscher spezifiziert werden, ob sie nun die CIA erfüllen oder nicht. Mit dem Ausgleich von Unterschieden ist daher ein Kriterium gegeben, wann aus statistischer Sicht ein Ausgleich der Selektionsverzerrung (hinsichtlich der berücksichtigten Merkmale) erreicht ist, nicht ob tatsächlich alle relevanten Selektionseinflüsse berücksichtigt wurden.

Die Frage, ob ein bestimmter Satz von Variablen die Strong Ignorability bzw. CIA erfüllt, ist keine Frage statistischer Korrelationen. Korrelationen zwischen  $x$  und  $k$ / bzw.  $x$  und  $y$  können aus verschiedenen Gründen bestehen, unter anderem auch dann, wenn  $x$  eine von der Fördergewährung beeinflusste Post-treatment-Variable darstellt. „We must be prepared to ignore irrelevant (...) variables even if they systematically differ in E and C trials, or else we can obtain any estimate desired by eventually finding the ‚right‘ irrelevant (...) variables“ (Rubin 1974: 699).<sup>32</sup> Wie unterscheiden sich aber irrelevante von relevanten Variablen? Dies kann nur eine inhaltliche Auseinandersetzung mit den empirischen Gründen für die Abweichung vom Zufallsprinzip liefern. „Estimating the typical causal effect of treatment versus another is a difficult task unless we understand the actual process well enough to (a) assign most of the variability in  $Y$  to specific causes and (b) ignore associated but causally irrelevant variables“ (Rubin 1974: 700).

Wann die Bedingungen für die CIA gegeben sind, kann daher nur auf Basis der Auseinandersetzung mit dem Gegenstand der Analyse beurteilt werden. Dazu ist es nötig zu untersuchen, was im konkreten Fall die Abweichung vom Zufall erzeugt. Im Bereich arbeitsmarktpolitischer Maßnahmen werden diese Abweichungen als Maßnahmeselektivität bezeichnet. Um diese in Form von Variablen für eine statistische Analyse nutzbar zu machen, muss sich eine soziologische Analyse damit

32 Mit „E and C trials“ sind in diesem Zusammenhang Teilnehmer- und Kontrollgruppe gemeint.

auseinandersetzen, dass im Fall von arbeitsmarktpolitischen Maßnahmen die Teilnahme keine von außen auferlegte Manipulation darstellt, sondern eine soziale Handlung. Soziale Handlungen sind dadurch gekennzeichnet, dass die Handelnden damit einen auf andere Handelnde bezogenen Sinn verbinden. Da an der Maßnahmeteilnahme bzw. -zuweisung stets mindestens zwei, oft sogar drei Personen beteiligt sind (vgl. Abschnitt 5.2.2), kann sogar davon gesprochen werden, dass es dabei um eine soziale Beziehung, also wechselseitig aufeinander bezogenes soziales Handeln, geht (Weber 1976). Die Verwendung des RKM zur Analyse sozialen Handelns ist in der Soziologie jedoch nicht ohne Kritiker. Die Orientierung des RKM gerade am Ideal des Experimentes und das damit verbundene Prinzip der Manipulierbarkeit hat bei einigen Soziologen zu einer kritischen Einschätzung geführt, ob die damit verbundene Vorstellung von Kausalität überhaupt für soziologisch interessante Fragestellungen im Allgemeinen und die Analyse sozialen Handelns im Besonderen relevant ist (Goldthorpe 2001; vgl. auch Heckman 2005a, 2005b sowie Gangl/DiPrete 2004a mit einer Gegenposition).

## 4.2 Das RKM und die kausale Analyse von sozialen Handlungen

Zur soziologischen Kritik hat vor allem Hollands programmatische Formulierung „no causation without manipulation“ (Holland 1986) geführt, die ihm als *conditio sine qua non* für Rubins Kausalmodell gilt. Der Grundgedanke Hollands ist, dass nur solche Faktoren eine Ursache sein können, die sich wie in einem Experiment durch einen Außenstehenden manipulieren lassen.

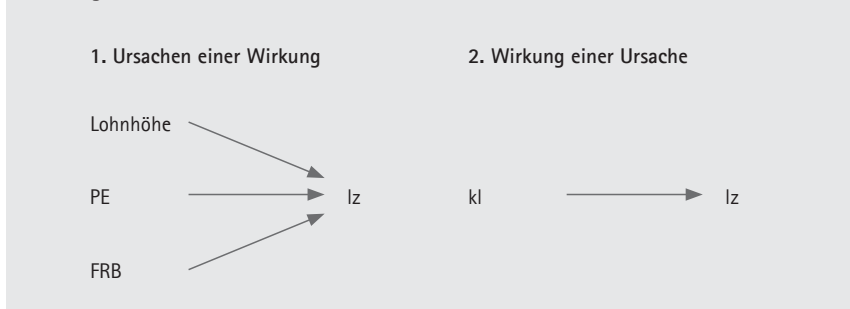
Im Folgenden soll Hollands Manipulationsthese und seine Konsequenzen für die Analyse soziologisch interessanter Gegenstände (wie die Konsequenzen von Maßnahmeteilnahmen) aufgezeigt werden. Danach soll die Kritik von Goldthorpe (2001) an Rubins Kausalmodell dargestellt werden, die diese zum Verwerfen des Konzeptes der Manipulation und damit des RKM führt. Schließlich soll diese Kritik entkräftet werden, indem gezeigt wird, dass Hollands Konzentration auf faktisch manipulierbare Ursachen einem Missverständnis des RKM entspringt sowie dass das Manipulationskonzept, im Sinne Rubins verstanden, als methodologisches Prinzip der Anleitung des Untersuchungsdesigns durchaus mit dem soziologischen Begriff der sozialen Handlung in Einklang steht.

### 4.2.1 „No causation without manipulation“?

Die Frage, wann eine Beziehung zwischen zwei Ereignissen, Zuständen oder sonstigen Phänomenen als kausal gelten kann, beschäftigt Menschen schon immer und lässt sich bisher (und wohl auch in Zukunft) nicht eindeutig beantworten. In den

Sozialwissenschaften lassen sich drei Konzepte unterscheiden, die nicht ohne Bezug auf philosophische Überlegungen entwickelt wurden, jedoch insgesamt eher pragmatischen Charakter haben: Kausalität als robuste Korrelation, Kausalität als Ergebnis einer Manipulation und Kausalität als generativer Mechanismus oder Prozess (Goldthorpe 2001; Cox/Wermuth 2000, 2004). Kausalität als robuster Zusammenhang ist an der Vorstellung orientiert, dass „X is a ‚genuine‘ cause of Y in so far as the dependence of Y on X can be shown to be robust that is to say, cannot be eliminated through one or more other variables being introduced into the analysis and then in some way ‚controlled‘ ” (Goldthorpe 2001: 2). In der Soziologie steht das Konzept des robusten Zusammenhanges in der Tradition der Verhinderung von „Scheinkorrelationen“ durch Drittvariablenkontrolle bzw. dem Ausschalten des Einflusses gemeinsamer Ursachen (vgl. z. B. Lazarsfeld 1976: 6; Simon 1976). In der Ökonomie benutzte bereits Yule um 1900 die Methode der Regression, um statistische Zusammenhänge zwischen zwei Merkmalen so von den Einflüssen weiterer Merkmale zu bereinigen, dass der Zusammenhang als Maß des kausalen Effektes gelten kann (Yule 1899). In Rubins Kausalmodell kommt dagegen das zweite der genannten Kausalkonzepte zum Ausdruck, *Kausalität als Ergebnis einer Intervention oder Manipulation*. Man betrachtet dort im Gegensatz zur gewöhnlichen Regressionsanalyse<sup>33</sup> nur eine einzige Ursache (vgl. Abbildung 13). Es geht also nicht darum, die verschiedenen Ursachen eines bestimmten Phänomens zu entdecken, z. B. verschiedene Ursachen für subjektive Wertschätzung des Einkommens, neben der Lohnhöhe etwa auch Persönlichkeitseigenschaften (PE) der Beschäftigten (vgl. etwa Schumacher/Laubach/Brähler 1995) oder bestimmte sonstige finanzielle Rahmenbedingungen FRB (z. B. Schulden).

Abbildung 13: Causes of effects vs. effects of causes



33 Sie wird hier als „gewöhnliche Regressionsanalyse“ bezeichnet, da die Regressionsanalyse auch zur Messung des kausalen Effektes im Sinne des RKM verwendet werden kann, dann aber einer der Logik des interventionsorientierten Kausalkonzeptes folgt.

Vielmehr konzentriert sich das RKM darauf, welche Wirkung eine genau definierte Manipulation auf eine ebenfalls genau definierte Zielvariable hat, in dieser Arbeit also die Wirkung der Kombientlohnung auf die subjektive Wahrnehmung der Reziprozität von Lohn und Arbeitsleistung. Holland hat dafür den Begriff „effects of causes“ geprägt, der die Messung kausaler Effekte vom üblichen Vorgehen der Suche nach den „causes of effects“ (Holland 1986: 945; kritisch dazu Dawid 1999: 36) unterscheidet.

In Rubins Kausalmodell erfolgt nicht nur eine Beschränkung des Interesses auf genau *eine* Ursache, sondern es kommt auch auf eine möglichst exakte Quantifizierung des Effektes dieser Ursache an,<sup>34</sup> unter anderem auch deshalb, weil zum Maßnahmeerfolg auch gehört, ob der meist hohe Aufwand der Maßnahmedurchführung vom Ausmaß des Erfolges gerechtfertigt wird. Anders ausgedrückt geht es neben der Signifikanz eines Effektes auch um dessen Relevanz.

Die Beschränkung der Kausalanalyse auf die effects of causes bzw. Interventionen oder Manipulationen hängt mit ihrer konzeptionellen Nähe zum experimentellen Vorgehen zusammen: „I take the position that causes are only those things that could, in principle, be treatments in experiments. (...) I believe that the notion of cause that operates in an experiment *and* in an observational study *is the same*. The difference is in the degree of control an *experimenter* has over the phenomenon under investigation compared with that which an *observer* has“ (Holland 1986: 954, Hervorh. im Orig.).

Holland nennt drei Beispiele, um sein Argument zu erläutern:

- (a) Sie war in der Prüfung gut, weil sie eine Frau war.
- (b) Sie war in der Prüfung gut, weil sie dafür gelernt hat.
- (c) Sie war in der Prüfung gut, weil sie von ihrem Lehrer gut vorbereitet wurde.

Fall (c) betrachtet Holland als einen prototypischen Fall für das RKM. Die Vorbereitung durch den Lehrer kann entweder stattfinden oder auch unterbleiben, sodass die für das Manipulationskonzept notwendige potenzielle Aussetzbarkeit (potential exosability) entweder gegenüber dem Experimentalzustand „gute Vorbereitung durch den Lehrer“ oder dem Kontrollzustand „schlechte Vorbereitung durch den Lehrer“ gegeben ist.

34 Dabei liegt der Fokus der Forschungspraxis meist auf der Wirkung einer Maßnahme auf eine genau definierte Zielvariable, sodass eventuelle Nebenwirkungen der Maßnahme nicht berücksichtigt werden. Die untersuchte Wirkung ist das von der Politik intendierte Ziel der Maßnahme (vgl. SGB III), also meist Faktoren wie das Einkommen, der Abgang aus der Arbeitslosigkeit, die Aufnahme einer dauerhaften (stabilen) Beschäftigung oder die Vermeidung von Arbeitslosigkeit. Seltener untersucht, doch trotzdem politisch intendierte Wirkungen, sind etwa die Förderung von Beschäftigungsfähigkeit, das Verhindern sozialer Exklusion, Steigerung der Arbeitszufriedenheit/-produktivität oder die Sanierung des staatlichen Finanzhaushaltes.

In Fall (a) fehlt Holland diese potenzielle Aussetzbarkeit völlig. Die Eigenschaft (attribute) „Frau“ ist bei der betrachteten Person zu jedem Zeitpunkt gegeben, so dass „to change its value is for the unit to change in some way and no longer be the same unit“ (ebd. 955). Das Geschlecht einer Person kann also ebenso wenig wie andere Eigenschaften die Ursache für irgendetwas sein, weil es nicht manipulierbar ist in dem Sinne, dass zu irgendeinem Zeitpunkt die Zuweisung in einen von zwei (oder mehr) Zuständen möglich war.

Einen Grenzfall stellt für Holland der Fall (b) dar, wobei er geneigt zu sein scheint, auch hier eine Nichtmanipulierbarkeit anzunehmen und damit das Wirken von Kausalität zu bestreiten. Probleme bereitet ihm vor allem der freiwillige („voluntary“) Aspekt des Lernens. „It is not clear that we could expose a person to studying or not in any verifiable sense“ (ebd.). Da Maßnahmeteilnahmen zumindest teilweise auch freiwillige Handlungen darstellen, träfe dieses Argument auch deren Analyse im RKM.

#### 4.2.2 Soziologische Kritik am RKM und ihre soziologische Entkräftung

Aufgrund der Tatsache, dass Holland mit seinem Artikel insgesamt eine umfassende und verständliche Darstellung des RKM gelungen ist, fand seine Interpretation des RKM und des Prinzips „no causation without manipulation“ weite Verbreitung. Hollands strikte Verknüpfung des RKM mit dem Konzept der praktischen Manipulierbarkeit führte jedoch zu einer skeptischen Haltung der Soziologie (ausdrücklich für das RKM sprechen sich dagegen Gangl/DiPrete 2004a aus). Eine derartige Einengung dessen, was als Ursache in Betracht kommt, ließe das RKM für die soziologische Forschung tatsächlich als wenig relevant erscheinen. Wenn einerseits soziologisch interessante „Eigenschaften“ wie die Zugehörigkeit zu einer sozialen Gruppe, das Geschlecht, die Bildung des Vaters etc. (vgl. Fall [a] bei Holland) und andererseits (soziale) Handlungen (vgl. Fall [c] bei Holland) von der Kausalanalyse ausgeschlossen werden, scheint das RKM in der Tat für die Soziologie wenig angebracht zu sein.

Der Soziologe John Goldthorpe plädiert daher für die Orientierung soziologischer Kausalanalysen am Konzept der Kausalität als generativer Mechanismus. Dieses Konzept „implies that there is some understanding, albeit provisional, of the process that leads from C to R (i.e. from cause to reaction; G.K.)“ (Cox/Wermuth 2000: 10). Dabei wird es oft argumentativ der statistischen Kausalanalyse als Alternative entgegengestellt (Little 1991: 177 ff.; Hedström/Swedberg 1998: 17). In der statistischen Analyse, so die Kritik, wird suggeriert, dass durch ein geeignetes Korrekturverfahren der wahre kausale Zusammenhang allein auf Basis empirischer Beobachtung von Regelmäßigkeiten und ganz ohne Bezug auf einen diesen Regel-

mäßigkeiten zugrunde liegenden Prozess identifiziert werden kann. Wie genau X seinen Einfluss auf Y ausübt, spielt keine Rolle, sodass es keiner Kenntnis kausaler Gesetzmäßigkeiten bedarf, um Kausalbeziehungen zu entdecken.

Gegen ein solches Vorgehen wird argumentiert, dass statistische Zusammenhänge eben nicht ausreichen, um kausale Effekte zu messen. Es muss eine ausführliche Erklärung des kausalen Zusammenhanges gegeben werden, damit das, was gemessen wurde, mit Recht auch als kausaler Effekt ausgewiesen werden darf. Eine solche Erklärung wird von Prozessen oder Mechanismen geleistet. „In most cases, the explanation takes the form of a more or less complicated set of causal statements. The relations between the elements of the set can be more or less complex; they can be linear, recursive, include feedback loops, and so on. The set is what we usually call a ‚social mechanism‘" (Boudon 1998b: 172; vgl. auch Bunge 1987; Elster 1998: 47 ff.; Hedström/Swedberg 1998: 7 ff.; Salmon 1984: 135 ff.). Ähnlich wird im Bereich der Evaluation arbeitsmarktpolitischer Maßnahmen kritisiert, dass die statistische Messung des kausalen Effektes im RKM keinerlei Aussagen dazu erlaubt, wie eine Maßnahme die Wirkung hervorruft bzw. warum eine Maßnahme die erhoffte Wirkung nicht entfalten konnte. Stellt man in einer Evaluation einen kausalen Effekt fest, so ist daraus nicht ableitbar, dass es gerade die hypothetische Verknüpfung ist, die am Wirken ist, noch kann im Falle des Verwerfens einer Kausalhypothese überprüft werden, woran es lag, dass die Maßnahme keinerlei Wirkung entfaltet hat (Knuth 2004: 188).<sup>35</sup>

Der Grund dafür ist die Orientierung der Evaluationsforschung an der Logik der experimentellen Messung kausaler Effekte, die sie im Bereich empirischer Beobachtungsstudien nachvollzieht. Aus aus Zufallsexperimenten gemessenen kausalen Effekten lässt sich, ohne dass konkret formuliert werden müsste, worin diese Verbindung besteht, darauf schließen, ob zwischen zwei Faktoren eine (kausale) Verbindung existiert (Steel 2004: 62 f.). Gerade darin wird der Vorteil von Experimenten gesehen: Trotz aller Theorie, die vielleicht zur Durchführung eines Experimentes motiviert, sind dessen Ergebnisse unabhängig von theoretischen Vorstellungen: „It is the sense in which experimental outcomes are determined by the workings of the world rather than by theoretical views about the world that provides the possibility of testing theories against the world" (Chalmers 1999: 40).

Insofern ist das Konzept der Kausalität als Manipulation auf die Berücksichtigung eines Mechanismus zwischen Ursache und Wirkung zwar nicht angewiesen, könnte aber davon profitieren. Es wären dann auch Fragestellungen beantwortbar, zu denen eine statistische Evaluation gewöhnlich keine Antwort gibt, etwa

<sup>35</sup> „Aussagen und implizite Annahmen zur Wirkungsweise aktiver Arbeitsförderung beruhen auf Alltagstheorien und Ad-hoc-Hypothesen; es fehlt sowohl an theoretischer Verknüpfung als auch an empirischer Evidenz (...) Bedarf besteht für eine sozialwissenschaftliche Verhaltenstheorie am Arbeitsmarkt" (Knuth 2004: 192, 194).

die nach den Auswirkungen möglicher Modifikationen des Treatments oder nach Sachverhalten, die einen eventuellen kausalen Effekt in der praktischen Durchführung der Maßnahme konterkarieren (vgl. z. B. Deeke/Kruppe 2003: 13). Eine weitere Nutzensmöglichkeit kausaler Mechanismen schlägt Rosenbaum (1984b) vor. Indem ein Mechanismus bestimmte Restriktionen enthält, z. B. wann ein Effekt auftritt und wann nicht, kann er zur Prüfung der Annahmen des RKM in einer konkreten Forschungssituation genutzt werden (vgl. dazu auch Abschnitt 6.5.1.2).

Goldthorpe betrachtet das Mechanismuskonzept allerdings weniger als Ergänzung des Manipulationskonzeptes, sondern als genuin soziologische Alternative, vor allem weil es besser geeignet sei, den Aspekt der kausalen Bedeutung von sozialen Handlungen Rechnung zu tragen. Er schlägt daher vor, das Konzept generativer Prozesse in folgenden Schritten zur Kausalanalyse von Handlungen zu nutzen: Erstens muss das zu erklärende Phänomen festgelegt werden, meist eine soziale Regelmäßigkeit. Zweitens wird eine Hypothese bezüglich des generativen Prozesses formuliert, die in einer Theorie des sozialen Handelns besteht, vorzugsweise in einer Theorie des rationalen Wahlhandelns. Drittens soll diese Hypothese dahin gehend empirisch überprüft werden, ob der hypothetische generative Mechanismus dazu in der Lage ist, die interessierende soziale Regelmäßigkeit hervorzurufen (Goldthorpe 2001: 9 ff.).<sup>36</sup>

Die Vorstellung Goldthorpes und anderer Soziologen, dass das RKM mit der Analyse von Handlungen nicht vereinbar ist, ist allerdings falsch und entspringt der Tatsache, dass die soziologische Kritik ausschließlich auf Hollands Formulierung des RKM Bezug nimmt und Rubins Richtigstellung des Manipulationsprinzips (u. a. in der an Hollands Artikel anschließenden Diskussion!) ignoriert.

Zur Anwendbarkeit des RKM bedarf es nicht unbedingt, dass eine potenzielle Ursache tatsächlich manipulierbar ist, sondern, so wird in Rubins Kommentar zu Hollands Artikel deutlich (Rubin 1986), es geht um eine *hypothetische* Manipulation der kausal interessanten Variable. Wichtig ist für Rubins Kausalmodell nicht, ob eine Manipulation technisch tatsächlich möglich ist, sondern um welche Art von (hypothetischer) Manipulation es sich handelt, deren Effekt gemessen werden soll. Ein (leicht verändertes) Beispiel aus Rubin (1986) soll dies verdeutlichen.

Für Holland hätte die Fragestellung „Hilft Medikament X gegen Kopfschmerzen?“ eindeutig kausalen Gehalt und kann daher im Rahmen des RKM analysiert werden, da die Vergabe eines Medikamentes manipulierbar ist. Dagegen hat die Fragestellung „Erzielen Frauen einen geringeren Lohn, weil sie Frauen sind?“ keinerlei kausale Bedeutung. Eine Frau zu sein ist für ihn eine Eigenschaft des Individuums, welches sich, wenn sich die Eigenschaft ändert, in ein anderes Individuum

<sup>36</sup> Hierin eine Alternative zum RKM zu sehen ist nur bedingt richtig. Wie die hier durchzuführenden Analysen der Wirkung des Kombilohnes auf die Reziprozitätswahrnehmung zeigen, lassen sich die Formulierung eines Wirkungsmechanismus und dessen Test auch mit auf dem Manipulationskonzept basierenden Analysen verbinden.



verwandelt (Holland 1986: 955). Ob das Geschlecht ursächlich für die Lohnhöhe ist, lässt sich also nicht feststellen, da das Geschlecht nicht manipulierbar ist, ohne dass die Analyseinheit in nicht kontrollierbarer Weise verändert wird. Für Rubin ist hingegen die Frage der (hypothetischen) Manipulation vor allem insofern relevant, als sie zu stellen dazu zwingt, die betrachtete Ursache so zu spezifizieren, dass die SUTVA-Bedingung erfüllt ist. Demnach ist die Fragestellung, ob Frauen wegen ihres Geschlechts am Arbeitsmarkt diskriminiert werden, zu vage gestellt und impliziert verschiedene hypothetische Möglichkeiten, das Geschlecht zu verändern (verschiedene versteckte Versionen des Treatments; Rubin 1986: 962; vgl. auch Rubin 2005: 323), und daraus folgend verschiedene potenzielle Ergebnisse im Fall der Treatmentzuweisung (vgl. Abschnitt 6.5.2.2). Rubin plädiert daher dafür, die potenzielle Ursache und damit auch den implizierten Vergleichszustand so genau wie möglich zu definieren, im Fall des Geschlechts etwa als „at conception X to Y chromosome treatment“ oder als „replacing ‚F‘ with an ‚M‘ on a job application form“<sup>37</sup> (Rubin 1986: 962). Damit hat das Konzept der potenziellen Manipulierbarkeit in Rubins Kausalmodell nicht die Funktion einer ontologischen Aussage darüber, welche Phänomene kausale Effekte haben können, die ihm Holland zuschreibt. Eher soll es das Forschungsdesign der Evaluation so anleiten, dass es nicht in Konflikt mit den Grundvoraussetzungen des RKM gerät.

Bei seinem Kommentar zu Hollands strikter Fassung des Manipulationskonzeptes scheint Rubin allerdings vor allem eine experimentelle Umsetzung dieser Fragestellung im Sinn zu haben. In Beobachtungsstudien stellt sich neben dem Ausschluss versteckter Treatmentversionen zusätzlich das Problem, den Zuweisungsmechanismus anzugeben. Hierbei müssen dann diejenigen Faktoren zur Konstruktion des Propensity Score verwendet werden, welche den Prozess der Zuweisung in den einen oder den anderen Status bestimmen. Dies wird bei zeitkonstanten Merkmalen wie dem Geschlecht schwieriger<sup>38</sup> sein als bei zeitvariablen Eigenschaften (Familienstatus, Bildung etc.), da sich der Zuweisungsmechanismus auf die Pre-treatment-Variablen stützt.

Die Frage ist daher also nicht, ob etwas eine Ursache sein kann oder nicht, sondern ob der kausale Effekt mithilfe Rubins Kausalmodell gemessen werden kann oder nicht.<sup>39</sup> Dabei ist die Antwort darauf nicht, wie Holland meint (Holland 1986: 970), von den gegenwärtigen technischen Möglichkeiten abhängig, sondern vielmehr vom gegenwärtigen theoretischen und empirischen Wissen über die Bedingungen, unter denen Menschen so oder anders handeln. Es genügt also

37 In den USA werden Bewerbungen keine Fotos beigefügt.

38 Schwieriger, aber nicht unmöglich, z. B. Messung bei Antritt eines neuen Jobs.

39 Und man muss daher ein anderes, eventuell mit mehr Unsicherheit behaftetes statistisches Modell bemühen. Für den Fall der Wirkung der Geschlechtszugehörigkeit auf die Entlohnung vgl. etwa Achatz/Gartner/Glück 2005.

für die Identifikation eines Zuweisungsmechanismus, dass Wissenschaftler die Einflussfaktoren auf die Zuweisung in den einen oder anderen „Experimentalzustand“ angeben können. Dies wiederum hängt vom Stand der Forschung und den Forschungsmethoden in der jeweiligen sozialwissenschaftlichen Disziplin ab, welcher der Forscher angehört.

Für eine soziologische Analyse, die sich des RKM bedient, gilt es daher, die Handlungslogiken der an der Maßnahmeteilnahme beteiligten Akteure zu erfassen.<sup>40</sup> Den Begriff des Handelns kennzeichnet in der Soziologie, dass die Handelnden damit einen subjektiven Sinn verbinden, soziales Handeln ist es, wenn es dem Sinn nach auf andere bezogen ist, wenn etwa das Handeln der Fallmanager in der Arbeitsagentur an dem Handeln der Arbeitsuchenden orientiert ist und/oder umgekehrt. Dass eine solche Verbindung der Manipulationskonzeptes der Kausalität, wie sie im RKM vorliegt, mit dem Begriff des sozialen Handelns keinen Widerspruch darstellt, zeigt der Begriff der „Wirkhandlung“ bei Alfred Schütz. Dieser bezeichnet neben dem Denken das Wirken als eine Form sozialen Handelns, das sich sowohl einseitig als auch wechselseitig (also mit „Antworthandeln“) vollziehen kann. Schütz verbindet in seinen Analysen das kausale Wirken explizit mit den intentionalen Aspekten des Handelns: „(...) gesellschaftliche Wirkakte und erst recht gesellschaftliche Arbeit (...) werden zu Motiven ‚antwortender‘ Handlungen oder zumindest zu Anlässen von Handlungen, die an den ursprünglichen Handelnden zurückgerichtet sind“ (Schütz/Luckmann 2003: 552; vgl. auch Schütz 1974: 209 ff.). Soziales Handeln ist hier also als Wirkhandlung eine „gezielte Veränderung in der natürlichen und sozialen Umwelt“ (ebd. 567) einerseits und beruht gleichzeitig auf sinnhafter Orientierung.

Einen Schritt weiter geht Robin Collingwood, indem er die menschliche Vorstellung von Kausalität auf ebensolche sozialen Wirkhandlungen zurückführt (Collingwood 1940: 285 ff.). Er unterscheidet drei Typen von Kausalbeziehungen. Der erste Typus besteht in der Kausalität, so wie sie in den theoretischen Naturwissenschaften auftritt. „Here that which is ‚caused‘ is an event or state of things, and its ‚cause‘ is another event or state of things standing to it in a one-one relation of causal priority“ (ebd.). Der zweite Typ beschreibt das kausale Eingreifen des Menschen in die Natur, also den Kausalitätsbegriff in der Technik. Der dritte Typus schließlich bezieht sich auf menschliches Handeln: „Here that which is ‚caused‘ is the free and deliberate act of a conscious and responsible agent, and ‚causing‘ him to do it means affording him a motive for doing it“ (ebd.). Dabei betrachtet Collingwood sowohl die technischen als auch die naturwissenschaftlichen Kausalitätsvorstellungen als von den Erfahrungen im Bereich des Handelns abgeleitet.

<sup>40</sup> Insofern trifft im Bereich der Sozialwissenschaften die Notwendigkeit der Erfassung der Maßnahmeselektivität im RKM mit Max Webers berühmter Definition der Soziologie als eine Wissenschaft, die „soziales Handeln deutend verstehen und dadurch in seinem Ablauf und seinen Wirkungen ursächlich erklären will“ (Weber 1976: 3).

### 4.3 Alternativen zur Messung kausaler Effekte mit dem RKM

Bei der Messung kausaler Effekte im RKM spielt die Annahme der bedingten Unabhängigkeit von Maßnahmeteilnahme und Zielvariable die entscheidende Rolle. Die Annahme ist dann erfüllt, wenn in eine Untersuchung alle gleichzeitigen Einflüsse auf Maßnahmeselektivität und Zielvariable in Form von Kontrollvariablen aufgenommen wurden. Dies setzt dann voraus, dass diese Einflüsse dem Forscher auch bekannt sind. Damit stellt das RKM sehr hohe Ansprüche an den Nachweis eines kausalen Effektes. Die Einflussfaktoren auf die Selektivität müssen nicht nur dem Forscher bekannt, sondern auch in Form entsprechender Variablen in der Datenbasis zur Verfügung stehen. Angesichts dessen stellt sich die Frage, ob damit unerfüllbare Bedingungen gestellt werden. In diesem Abschnitt soll diese Frage auf zweifache Weise beantwortet werden. Zum einen wird gezeigt, dass das Standardverfahren der statistischen Analyse in Soziologie und Ökonomie, die lineare Regression, ebenfalls vom Problem falscher Schlussfolgerungen bei fehlenden Variablen betroffen ist, auch wenn dies in der Forschungspraxis kaum thematisiert wird (Abschnitt 4.3.1). Zum anderen wird gezeigt, dass auch die Verwendung anderer statistischer Verfahren zur Kausalanalyse, die für sich beanspruchen, das Problem fehlender Variablen zu lösen, mit Problemen behaftet ist und damit kein Ersatz, sondern eher eine Ergänzung des RKM darstellen kann (Abschnitt 4.3.2).

#### 4.3.1 Das Problem fehlender Variablen in der OLS- und der ML-Schätzung

Das Problem der Vollständigkeit der theoretischen Modelle in einer empirischen Untersuchung ist nicht auf Kausalanalysen beschränkt. Vielmehr ist es ein generelles Problem der Validität multivariater statistischer Analysen. Gezeigt werden kann dies an den statistisch-algebraischen Grundlagen der multiplen linearen Regression – der Methode der kleinsten Quadrate (KQ) – und der Maximum-Likelihood-(ML-)Schätzung, welche die Basis der meisten in Soziologie und Ökonometrie verwendeten Verfahren darstellen. Allgemein formuliert ist der Zweck einer Regression die Darstellung der Werte einer Variable durch andere Variablen (Verbeek 2004). Dabei wird die zu erklärende Variable auch endogene, abhängige Variable oder Kriterium genannt und mit dem Buchstaben  $y$  bezeichnet. Die erklärenden/exogenen/unabhängigen Variablen bzw. Regressoren werden mit  $x$  bezeichnet und mit  $1, \dots, k$  durchnummeriert:  $x_1, x_2, \dots, x_{k-1}, x_k$ .

Gegeben sei eine Stichprobe von  $i = 1, \dots, n$  Personen, an denen das Merkmal  $y$  gemessen wurde und die  $k$  anderen Merkmale  $x_{i1}, x_{i2}, \dots, x_{ik}$ . Die Variable  $y$  kann etwa – wie in dieser Untersuchung – die subjektiv empfundene Verletzung der Reziprozität sein, die  $x$ -Variablen Merkmale wie der Kombilohnbezug, Persönlich-

keitseigenschaften der Beschäftigten, Persönlichkeitseigenschaften der Vorgesetzten, Arbeitsbedingungen etc. Für jede Person  $i$  stellt die lineare Regression den Wert der  $y$ -Variable als eine Linearkombination der  $x$ -Variablen dar. Im einfachsten Fall existiert nur eine exogene Variable (Einfachregression), in den meisten Fällen werden allerdings mehrere Regressoren verwendet (multiple Regression):

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (4.16)$$

Dabei ist  $\beta_0$  der Niveauparameter (Konstante),  $\beta_1$  bis  $\beta_k$  sind die Regressionskoeffizienten,  $\varepsilon_i$  ein *zufälliger* Störterm. Da die Größe  $\varepsilon_i$  vom Zufall abhängt, können nun auch Personen mit identischen beobachteten Regressoren unterschiedliche  $y$ -Werte aufweisen. Für die Existenz dieser Zufallsgröße werden unterschiedliche Begründungen genannt (von Auer 2003: 35). Erstens hat man es in den Sozialwissenschaften mit Menschen zu tun. Will man das Handeln von Menschen analysieren, muss man auch deren freien Willen Rechnung tragen (vgl. auch Blossfeld/Rohwer 2002: 28), der den Beobachtungen in der Realität – aus Sicht des Forschers – ein Zufallselement verleiht. Die zweite Begründung orientiert sich am Prozess der Datenerhebung und der Messung. Im Störterm zeigen sich unsystematische Erhebungs- oder Messfehler oder andere Abweichungen des beobachteten vom wahren Wert, z. B. wenn ein Merkmal nicht selbst gemessen werden kann, sondern auf eine Proxyvariable<sup>41</sup> ausgewichen werden muss. Auch hier liegt, insofern man nur die *unsystematischen* Abweichungen meint, Zufälligkeit vor. Schließlich verweist eine dritte Begründung für die Störgröße auf die Unvollständigkeit des Modells. Werden bestimmte Merkmale, die  $y_i$  beeinflussen, nicht explizit unter die  $x_{ik}$  aufgenommen, schlagen sie sich in einer Störgröße nieder. Zu diesem Punkt wird später noch einiges auszuführen sein. Schon jetzt muss allerdings betont werden, dass dann die Zufälligkeit der Störgröße nicht mehr gegeben ist. Nichts spricht im Übrigen dagegen, für den Störterm alle drei Begründungen gleichzeitig anzunehmen.

Wichtig ist vor allem, dass es sich bei der Störgröße  $\varepsilon_i$  um einen Einflussfaktor handelt, von dem man einerseits *annimmt*, dass er eine empirische Entsprechung im untersuchten Gegenstandsbereich hat (s. o.: freier Wille, unsystematische Fehler, unerfasste Merkmale). Andererseits kann er in konkreten Beobachtungen jedoch nicht isoliert werden, sondern „verschmilzt“ mit den Einflüssen der anderen Merkmale.

41 Proxyvariablen versuchen, ein Konzept über den Umweg eines ähnlichen Konzeptes zu messen, z. B. wenn statt der tatsächlichen Gesundheit einer Person die Zufriedenheit dieser Person mit ihrer Gesundheit verwendet wird. Die Zufriedenheit ist ein subjektives Merkmal und muss nicht mit der objektiven tatsächlichen Gesundheit übereinstimmen. Der Unterschied zwischen beiden ist dann ein Messfehler. Dieser sollte nicht mit den Regressoren im Modell korrelieren, wenn man an unverzerrten Koeffizienten interessiert ist.

Damit es möglich ist, den Einfluss der Störgröße von den Einflüssen der Regressoren zu trennen, müssen die sog. Gauss-Markov-Annahmen (Verbeek 2004: 16) zutreffen. Sie betreffen die Störgröße und ihre Beziehung zu den Regressoren. Sind diese Bedingungen erfüllt, dann ist  $\hat{\beta}_k$  ein unverzerrter (und effizienter) Schätzer für den wahren Wert. Die Bedingungen sind:

$$1. E(\varepsilon_i) = 0 \quad i = 1, \dots, n$$

Die erste Bedingung ist, dass der Erwartungswert der Störgrößen null ist. Dies muss für die Störgröße einer jeden Beobachtung bzw. Person  $i$  gelten.

$$2. V(\varepsilon_i) = \sigma^2 \quad i = 1, \dots, n$$

In Bedingung zwei wird formuliert, dass die Störgrößen jeder Beobachtung dieselbe Varianz aufweisen (weshalb der Index  $i$  auch wegfällt). Diese Eigenschaft der Störgrößen wird als Homoskedastie bezeichnet und ist nötig, da die Varianz der Störgrößen für die Signifikanztests gebraucht wird. Geschätzt wird die Varianz allerdings mithilfe der Varianz der Residuen  $e_i$ . Wäre die Varianz von  $\varepsilon$  nicht für alle Individuen konstant, könnte man sie nicht aus der Poolung der Residuen aller Individuen berechnen.

$$3. Cov(\varepsilon_i, \varepsilon_j) = 0 \quad i, j = 1, \dots, n, i \neq j$$

Bedingung drei schließt aus, dass die Störgrößen der Einzelbeobachtungen miteinander korreliert sind. Dies ist die Forderung, dass Autokorrelation der Störgrößen ausgeschlossen sein muss. Nimmt man zusätzlich noch an, dass die Störgrößen normalverteilt sind, dann sagen Bedingungen eins bis drei zusammengenommen aus, dass die (wahren) Störgrößen unabhängige Zufallsziehungen aus einer Normalverteilung sind, mit Mittelwert 0 und Varianz  $\sigma^2$ . Die Normalverteilungsbedingung ist jedoch nur für kleine Stichproben zwingend. Bei großen Stichproben gilt der zentrale Grenzwertsatz von Ljapunov.

$$4. (\varepsilon_i) \text{ und } (x_{i1}, \dots, x_{ik}) \text{ sind stochastisch unabhängig.}$$

Die vierte Bedingung ist etwas komplexer als die drei vorherigen. Sie besagt, dass Störgrößen unabhängig von den x-Merkmalen der Regression sind. Es genügt jedoch schon die etwas schwächere Annahme, dass die Regressoren auf der Ebene der Individuen mittlere bedingte Unabhängigkeit (*mean conditional independence*) aufweisen:

$$E(\varepsilon_i | x_{i1}, \dots, x_{ik}) = E(\varepsilon_i) \quad (4.17)$$

Gegeben die x-Regressoren (d. h. in jeder möglichen Kombination von Ausprägungen der x-Regressoren) nimmt  $\varepsilon_i$  im Mittel denselben Wert an. Damit ist es völlig irrelevant, welche Merkmalskombination ein Individuum aufweist, der Mittelwert der Störgröße ist stets der gleiche. Man spricht auch davon, dass die enthaltenen Regressoren keinerlei Informationen bezüglich des Wertes von  $\varepsilon_i$  liefern. Aus der Kombination von Bedingung eins mit Bedingung vier ergibt sich:

$$E(\varepsilon_i | x_{i1}, \dots, x_{ik}) = E(\varepsilon_i) = 0 \quad (4.18)$$

Daraus folgt, dass das stochastische lineare Modell (4.16) im Mittel dem deterministischen Modell ohne Störgröße (4.19) entspricht:

$$E(y_i | x_{i1}, \dots, x_{ik}) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} \quad (4.19)$$

Treffen alle vier Bedingungen zu, dann lassen sich die  $\beta_k$  erwartungstreu, d. h. verzerrungsfrei mit der Methode der kleinsten Quadrate errechnen.

Verletzungen der Bedingungen zwei oder drei sind bei der Durchführung einer multiplen Regression kaum problematisch. Es existieren diverse Möglichkeiten, die Koeffizienten trotzdem unverzerrt zu schätzen.

Das zentrale Problem bei der Verwendung der Regressionsanalyse ist, dass die Bedingungen eins und vier genau dann verletzt werden (können), wenn bedeutungsame Regressoren in der Gleichung nicht berücksichtigt werden. Daraus resultiert dann eine verzerrte Schätzung der Regressionskoeffizienten (der sog. omitted variable bias). Diese Bedingungen, die für die OLS-Regression unverzichtbar sind, entsprechen damit der CIA in Rubins Kausalmodell.

Ähnlich wie im RKM ist die Verfälschung der Regressionskoeffizienten davon abhängig, in welcher Beziehung enthaltene und ausgelassene Variablen zueinander stehen. Um dies zu zeigen, soll angenommen werden, dass ein Merkmal  $y$  mit zwei Regressoren  $x_1$  und  $x_2$  zusammenhängt, die Regression aber nur mit einem Regressor, nämlich  $x_1$ , durchgeführt wird.

Das wahre Modell lautet also  $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$ , und das um das Merkmal  $x_2$  verkürzte Modell, dessen Parameter und Störterm im Folgenden mit einem \* gekennzeichnet werden, lautet  $y_i = \beta_0^* + \beta_1^* x_{i1} + \varepsilon_i^*$ . Dabei besteht dann also der Störterm im verkürzten Modell aus dem eigentlichen Störterm und dem ausgelassenen Merkmal:  $\varepsilon_i^* = \beta_2 x_{i2} + \varepsilon_i$ . Ausgelassene Merkmale „sammeln“ sich bei der Durchführung einer unvollständigen Regression im Störterm und sind somit in der Regression nicht mehr erkennbar.

Führt man dann eine Regression durch, bei der nur eine abhängige Variable aufgenommen wird, der wahre Zusammenhang aber auf zwei abhängigen Variablen beruht, dann sind die in der verkürzten Regression geschätzten Parameter  $\hat{\beta}_0^*$  und  $\hat{\beta}_1^*$  gegenüber den wahren Parametern  $\hat{\beta}_0$  und  $\hat{\beta}_1$  verzerrt. Diese Verzerrung ist abhängig vom Koeffizienten des ausgelassenen Merkmals  $\hat{\beta}_2$  und von dem Ausmaß des linearen Zusammenhanges zwischen der enthaltenen Variable  $x_1$  und der ausgelassenen Variable  $x_2$  – genauer von der Kovarianz zwischen beiden (vgl. etwa von Auer 2003):

$$\hat{\beta}_0^* = \hat{\beta}_0 + \hat{\beta}_2 \left( \bar{x}_2 - \frac{\text{Cov}(x_1, x_2)}{\text{Var}(x_1)} \bar{x}_1 \right) \quad (4.20)$$

$$\hat{\beta}_1^* = \hat{\beta}_1 + \hat{\beta}_2 \frac{\text{Cov}(x_1, x_2)}{\text{Var}(x_1)} \quad (4.21)$$

Diese Darstellung an sich bekannter Eigenschaften der OLS-Regression zeigt, dass jede statistische Analyse mit dem Problem fehlender Variablen zu tun hat. Dies gilt nicht nur, wenn explizit Ursache-Wirkungs-Zusammenhänge untersucht werden sollen, sondern immer dann, wenn die Absicht der jeweiligen Analyse über die Formulierung einer Aussage wie „x ist auf irgendeine Weise mit y korreliert“ hinausgehen soll.

Zu Problemen bei fehlenden Variablen kommt es aber nicht nur bei der OLS-Regression, sondern auch bei Maximum-Likelihood-(ML-)Schätzungen. Dies kann am Beispiel der logistischen Regression dargestellt werden. Ausgangspunkt der ML-Schätzung ist, dass die Koeffizienten so geschätzt werden sollen, dass das Auftreten der vorliegenden Stichprobenwerte maximal wahrscheinlich ist. Diese soll hier kurz skizziert werden.

Die gemeinsame Wahrscheinlichkeitsdichtefunktion einer Stichprobe N unabhängiger und identisch verteilter Beobachtungen – gegeben ein Satz bestimmter Parameter – ist gegeben als Produkt der einzelnen Wahrscheinlichkeitsdichtefunktionen und wird als Likelihood-Funktion bezeichnet:

$$f(y_1, y_2, \dots, y_n | \beta_0, \dots, \beta_k) = \prod_{i=1}^n f(y_i | \beta_0, \dots, \beta_k) = \quad (4.22)$$

$$L(\beta_0, \dots, \beta_k | y_1, y_2, \dots, y_n) =$$

$$L(\beta_0, \dots, \beta_k | \text{Stichprobe})$$

Meist wird die Likelihood-Funktion aber in logarithmierter Form dargestellt, als Loglikelihood-Funktion:

$$\ln L(\beta_0, \dots, \beta_k | \text{Stichprobe}) = \sum_{i=1}^n \ln f(y_i | \beta_0, \dots, \beta_k)$$

Da in der logistischen Regression davon ausgegangen wird, dass die Wahrscheinlichkeit für  $y_i=1$  durch die kumulierte logistische Verteilungsfunktion gegeben ist, ist die spezielle Likelihood-Funktion in diesem Fall (Eckey/Kosfeld/Dreger 2001: 191):

$$L(\beta_0, \dots, \beta_k | \text{Stichprobe}) = \prod_{i=1}^n \left( \frac{1}{1 + \exp(-\mathbf{x}'_i \boldsymbol{\beta})} \right)^{y_i} \left( 1 - \frac{1}{1 + \exp(-\mathbf{x}'_i \boldsymbol{\beta})} \right)^{1-y_i}$$

Die Loglikelihood-Funktion lautet:

$$L(\beta_0, \dots, \beta_k | \text{Stichprobe}) = \sum_{i=1}^n y_i \ln \left( \frac{1}{1 + \exp(-\mathbf{x}'_i \boldsymbol{\beta})} \right) + (1 - y_i) \ln \left( 1 - \frac{1}{1 + \exp(-\mathbf{x}'_i \boldsymbol{\beta})} \right)$$

Ein Maximum im Bezug auf  $(\beta_0, \dots, \beta_k)$  findet sich durch das Nullsetzen der ersten Ableitung. Da eine analytische Bestimmung der  $(\beta_0, \dots, \beta_k)$  nicht möglich ist, wird hierzu ein iteratives Verfahren verwendet (ausführlicher vgl. Eckey/Kosfeld/Dreger 2001: 191).

Auch bei der ML-Schätzung führt das Fehlen von relevanten Variablen zu verzerrten Schätzungen des kausalen Effektes, da dann die Likelihood-Funktion nicht korrekt spezifiziert ist. Im Fall der ML-Schätzung weisen die Forschungsergebnisse von Yatchew und Griliches (1985) darauf hin, dass im Gegensatz zur OLS-Regression „even if the omitted variable is uncorrelated with the included one, the coefficient on the included variable will be inconsistent“ (Greene 2003: 679).

#### 4.3.2 Ansätze für unbeobachtete Maßnahmeselektivität

In der statistischen und ökonometrischen Forschung werden als Alternative zur Erfüllung der CIA Verfahren vorgeschlagen, für die diese Bedingung nicht gilt. Diese Verfahren bezeichnet man als Verfahren für unbeobachtete Heterogenität (Heckman/Hotz 1989). Wäre es dann nicht vernünftiger, das Ziel der CIA aufzugeben und stattdessen zur Analyse des kausalen Effektes der Kombilohnförderung auf diese Verfahren zurückzugreifen? Tatsächlich werden solche Verfahren auch in der Kausalanalyse und zum Teil auch bei der Betrachtung kausaler Effekte arbeitsmarktpolitischer Maßnahmen eingesetzt (Caliendo/Hujer 2006). Bei diesen Verfahren handelt es sich vor allem um Heckmans Switching-Regression (Abschnitt 4.3.2.1) und die Instrumentvariablenschätzung (Abschnitt 4.3.2.2).<sup>42</sup> Diese

<sup>42</sup> Daneben existieren mit dem Timing-of-Events-Ansatz (Abbring/van den Berg 2003), parametrischen (Meyer 1995) und nicht parametrischen Difference-in-Differences-Ansätzen (Smith/Todd 2005) sowie dem Regression-Discontinuity-Ansatz (Hahn et al. 2001) noch weitere Verfahren für unbeobachtete Heterogenität. Da diese jedoch mit den hier verfügbaren Daten (vgl. Abschnitt 6.1) nicht umsetzbar sind, wird auf ihre Darstellung verzichtet.



sollen hier dargestellt werden, und es soll überprüft werden, ob sie besser zur Messung des kausalen Effektes des Kombilohnes geeignet sind als das RKM.

Um dies zu beurteilen, darf der Vergleich des RKM mit den anderen Vorgehensweisen nicht nur auf der Dimension „unbeobachtete Maßnahmeselektivität“ stattfinden, sondern muss auch andere Aspekte mit berücksichtigen. Es zeigt sich, dass die Aufweichung der CIA stets mit dem Treffen zusätzlicher Annahmen verbunden ist, vor allem bezüglich anderer Variablenklassen als der für die Erfüllung der CIA relevanten Klasse  $S_{krel}$  und bezüglich der funktionalen Form des Einflusses der betrachteten Variablen (Abschnitt 4.3.2.3).

#### 4.3.2.1 Die Regression mit Kontrollfunktion als Gegenentwurf zum RKM

Der modelltheoretische Ansatz von James Heckman (2005a, 2005b; Heckman/Navarro-Lozano 2004) versteht sich dezidiert als Gegenentwurf zum RKM. Heckmans Regression mit Kontrollfunktion geht davon aus, dass es in den meisten Fällen unrealistisch sei, von der Gültigkeit der CIA auszugehen. Daher schlägt er vor, unbeobachtete Maßnahmeselektivität zuzulassen und diese indirekt zu messen und in die Schätzung einzubauen. Dies geschieht mithilfe der Formulierung ökonomischer Modelle bzw. Gleichungen und statistischer Annahmen zur funktionalen Form des Zusammenhanges von unbeobachteter Maßnahmeselektivität und unbeobachteten Einflussfaktoren auf die Zielvariable.

Wie in Rubins Kausalmodell ist auch der Ausgangspunkt der Überlegungen in Heckmans modelltheoretischem Ansatz (HMA), dass es für jedes Individuum zwei mögliche Ergebnisse der Zielvariable  $y$  gibt. Das Ergebnis bei Kombilohnerhalt ist  $y_i^1$  und das bei regulärer Beschäftigung  $y_i^0$ . Diese Überlegung wird in Form der Switching-Regression (Quandt 1972) dargestellt, die angibt, welche Version der Zielvariable für das Individuum  $i$  realisiert wird, abhängig davon, ob es eine Förderung erhält  $k_i = 1$  oder nicht  $k_i = 0$ . Dabei wird zunächst davon ausgegangen, dass die Zielvariable metrisch ist.

$$y_i = k_i y_i^1 + (1 - k_i) y_i^0 \quad (4.23)$$

Man spricht bei der Charakterisierung der beiden alternativen Zustände durch zwei Regressionen von alternativen „Regimes“. Die Regimes können durch zwei Gleichungen beschrieben werden, deren funktionale Form beliebig ist, relevant ist allerdings die additive Trennbarkeit, die Tatsache, dass die Störgröße additiv in das Modell eingehen muss. Hier und im Folgenden wird von einer linearen Gleichung ausgegangen. Die einfachste und verbreitetste Version des HMA geht davon aus, dass der Einfluss sowohl der beobachteten als auch der unbeobachteten Heterogenität in beiden Regimes gleich ist (Heckman 1978, 1979). Dies wird als Annahme konstanter Effekte bezeichnet. Die beiden Regimegleichungen lauten:

$$y_i^1 = \alpha^1 + \mathbf{x}_i' \boldsymbol{\beta} + \varepsilon_i \quad (4.24)$$

$$y_i^0 = \alpha^0 + \mathbf{x}_i' \boldsymbol{\beta} + \varepsilon_i \quad (4.25)$$

Der Vektor  $\mathbf{x}_i$  erfasst dabei die beobachtete Heterogenität der (Un-)Geförderten, also alle Merkmale, welche die Ergebnisvariable beeinflussen und in der vorliegenden Datenbasis gemessen wurden. Dagegen bezeichnet  $\varepsilon_i$  die unbeobachtete und/oder die unbeobachtbare Heterogenität, also alle Einflussfaktoren auf die Wirkungsvariable, die nicht gemessen wurden oder die grundsätzlich vom Forscher als nicht messbar betrachtet werden. Neben den Gleichungen für die Regime wird im HMA auch eine sogenannte Selektionsgleichung oder Index-Funktion formuliert, die angibt, welche der beiden Regime für ein bestimmtes Individuum relevant wird:

$$\tilde{kl}_i = \mathbf{z}_i' \boldsymbol{\gamma} + v_i \text{ mit } kl_i = \begin{cases} 0, & \text{falls } \tilde{kl}_i < 0 \\ 1, & \text{falls } \tilde{kl}_i \geq 0 \end{cases} \quad (4.26)$$

Auch hier wird zwischen den beobachteten Variablen  $z$  und der unbeobachteten Heterogenität  $v_i$  unterschieden. Dabei werden sich  $x$  und  $z$  einzelne Elemente teilen, nämlich die Variablen, die gleichzeitig Determinanten der Maßnahmeteilnahme sind und die Zielvariable beeinflussen. Im Falle des PS-Matchings wurden diese Variablen als kausal relevante Selektivität  $S_{krel}$  bezeichnet. Selektionseinflüsse der Kombilohnförderung ohne Einfluss auf die Zielvariable ( $S_{irrel}$ ) sind nur in  $z$ , sonstige Determinanten der Zielvariable ohne Einfluss auf die Kombilohnförderung nur in  $x$  enthalten.

Der Effekt der Teilnahme an der Fördermaßnahme wird im Zusammenhang mit dem HMA oft nicht als kausaler Effekt, sondern als Treatmenteffekt oder auch als Zugewinn (gain) durch die Maßnahme bezeichnet. Formal und inhaltlich gibt es keinen wesentlichen Unterschied zwischen kausalem und Treatmenteffekt. Wie der kausale Effekt im RKM ist Letzterer ebenfalls definiert als  $\delta_i \equiv y_i^1 - y_i^0$ .<sup>43</sup> Wie im RKM werden auch hier nur mittlere kausale Effekte gemessen. Allerdings erlaubt Heckmans Modell – unter bestimmten Annahmen – neben dem mittleren kausalen Effekt grundsätzlich auch die Identifikation weiterer Parameter der Verteilung des kausalen Effektes (siehe Heckman/LaLonde/Smith 1999: 19).

Setzt man die beiden Gleichungen 4.24 und 4.25 in die Switching-Regression (4.23) ein, dann ergibt sich nach dem Ausmultiplizieren und einigen Umformungen die Gleichung 4.27 (vgl. etwa Heckman/LaLonde/Smith 1999: 20 ff.):

<sup>43</sup> Zu den Entsprechungen in beiden Paradigmen siehe Heckman/Navarro-Lozano (2004) aus Sicht des HMA und Holland (1988) aus Sicht des RKM, auch wenn sich die ökonomische Evaluationsforschung bemüht, den Bezug auf den Kausalitätsbegriff zu vermeiden.

$$y_i = \alpha^0 + (\alpha^1 - \alpha^0)kl_i + \mathbf{x}'_i\boldsymbol{\beta} + \varepsilon_i = \alpha^0 + \delta kl_i + \mathbf{x}'_i\boldsymbol{\beta} + \varepsilon_i \quad (4.27)$$

Der mittlere kausale Effekt ist dann  $MKE = E(y_i^1 - y_i^0) = E(\delta_i)$ , wobei es dann keinen Unterschied zwischen  $MKE$  und  $MKE_{NT}$  bzw.  $MKE_T$  gibt.

Gilt für die Störgröße die bedingte Unabhängigkeit vom Teilnahmestatus, also  $E(\varepsilon_i | \mathbf{x}_i, \mathbf{z}_i, kl_i) = E(\varepsilon_i | \mathbf{x}_i, \mathbf{z}_i) = 0$ , dann ist der PS-Matchingschätzer ein unverzerrter Schätzer für den kausalen Effekt  $E(\delta_i)$  (vgl. dazu auch die Anmerkungen zur linearen Regression). Dies ist genau dann der Fall, wenn der Vektor  $\mathbf{x}$  alle  $S_{krel}$  enthält. Dieser Fall wird als Selektion auf Basis der beobachteten Heterogenität (selection on observables) bezeichnet (Heckman/Hotz 1989). In diesem Sinne wird dann auch Rubins Kausalmodell oft als ein Vorgehen für selection on observables bezeichnet, weil es annimmt, dass alle relevanten Determinanten der Maßnahmeteilnahme in  $\mathbf{x}$  enthalten sind. Die Formulierung übersieht allerdings, dass im RKM auch unbeobachtete Heterogenität zugelassen wird, wenn diese dann auch andere Arten von Variablen betrifft als das HMA. Sind alle relevanten Einflussfaktoren auf die Maßnahmeteilnahme Teil der beobachteten Heterogenität, dann kann der kausale Effekt problemlos auch im Rahmen des RKM gemessen werden. Fehlen in  $\mathbf{x}$  jedoch zentrale Determinanten der Maßnahmeteilnahme und haben diese gleichzeitig auch Einfluss auf die Zielvariable  $y$ , dann spricht man von Selektion auf Basis unbeobachteter Heterogenität (selection on unobservables). Dies führt beim PS-Matching zu einer Verletzung der CIA.

In diesem Fall der selection on unobservables enthält der unbeobachtete Teil  $\varepsilon_i$  einen systematischen Teil  $E(\varepsilon_i^S | \mathbf{x}_i, \mathbf{z}_i, kl_i) \neq E(\varepsilon_i^S | \mathbf{x}_i, \mathbf{z}_i)$ , der mit  $kl$  korreliert ist und einen zufälligen Teil  $\varepsilon_i^*$ , für den das nicht gilt und für den ansonsten  $E(\varepsilon_i^* | \mathbf{x}_i, \mathbf{z}_i, kl_i) = E(\varepsilon_i^* | \mathbf{x}_i, \mathbf{z}_i) = 0$  gilt.

Gleichung 4.27 lässt sich damit auch wie folgt ausdrücken:

$$y_i = \alpha^0 + \delta kl_i + \mathbf{x}'_i\boldsymbol{\beta} + \varepsilon_i^S + \underbrace{(\varepsilon_i - \varepsilon_i^S)}_{\varepsilon_i^*} \quad (4.28)$$

Aus dieser Perspektive erscheint das Problem der unbeobachteten Heterogenität als das Problem einer fehlenden Variable in der Regression der Zielvariable (Heckman 1979). Eine Lösung für die stochastische Abhängigkeit von Störgröße und Kombilohnersatz erfordert damit, dass der systematische Teil mit in die Gleichung 4.27 aufgenommen wird. Da es sich aber um unbeobachtete, also mit den vorhandenen Daten nicht erfassbare Heterogenität handelt, wird die fehlende Variable auf dem Umweg über die Indexfunktion (Gleichung 4.26) erfasst (Heckman/MaCurdy 1986; Heckman/Smith 1999: 56 f.):

$$P(kl_i = 1 | \mathbf{z}_i) = P(\tilde{k}l_i > 0) = P(v_i > -\mathbf{z}'_i \boldsymbol{\gamma}) \quad (4.29)$$

$$P(kl_i = 0 | \mathbf{z}_i) = P(\tilde{k}l_i \leq 0) = P(v_i \leq -\mathbf{z}'_i \boldsymbol{\gamma}) \quad (4.30)$$

Ist  $v_i > -\mathbf{z}'_i \boldsymbol{\gamma}$ , dann erhält Person  $i$  einen Kombilohn, ist dagegen  $v_i \leq -\mathbf{z}'_i \boldsymbol{\gamma}$ , dann geht sie eine ungeforderte Beschäftigung ein. Besteht unbeobachtete kausalrelevante Selektivität der Kombilohnförderung, dann kommt es zu einer Korrelation der Störgrößen  $\varepsilon$  und  $v$  in der Gleichung der Zielvariable  $y$  und der Selektionsgleichung der Kombilohnförderung. Unter bestimmten Bedingungen lässt sich daraus eine Korrektur der Selektionsverzerrung konstruieren.

Heckman schlägt vor, eine bivariate Normalverteilung der beiden Störgrößen anzunehmen, also  $(\varepsilon, v) \sim N(0, 0, \sigma_\varepsilon^2, \sigma_v^2, \rho_{\varepsilon v})$ , mit bestimmten Varianzen  $\sigma^2$  und einer Korrelation der Störgrößen  $\rho$ . Daneben muss die Annahme gerechtfertigt sein, dass die Störgrößen  $\varepsilon$  und  $v$  von der beobachteten Heterogenität  $\mathbf{x}$  und  $\mathbf{z}$  unabhängig sind:  $(\varepsilon, v) \perp (\mathbf{x}, \mathbf{z})$ . Diese Annahme bedeutet, dass die unbeobachtete Heterogenität unabhängig von den Regressoren der Regressionsgleichungen 4.26 und 4.28 ist. Schließlich muss in  $\mathbf{z}$  noch mindestens ein Element enthalten sein, das nicht mit  $y$  korreliert (exclusion restriction, vgl. Puhatni 2000).

Gleichung 4.27 lässt sich auch formulieren als:

$$y_i = \alpha^0 + \delta kl_i + \mathbf{x}'_i \boldsymbol{\beta} + kl_i \varepsilon_i + (1 - kl_i) \varepsilon_i. \text{ Dabei gilt:}$$

$$E(y_i^1 | \mathbf{x}_i, kl_i = 1, \mathbf{z}_i) = \alpha^0 + \delta + \mathbf{x}'_i \boldsymbol{\beta} + E(\varepsilon_i | \mathbf{x}_i, kl_i = 1, \mathbf{z}_i)$$

$$\text{und } E(y_i^0 | \mathbf{x}_i, kl_i = 0, \mathbf{z}_i) = \alpha^0 + \mathbf{x}'_i \boldsymbol{\beta} + E(\varepsilon_i | \mathbf{x}_i, kl_i = 0, \mathbf{z}_i)$$

Unter den genannten Annahmen ergibt sich für die Erwartungswerte der Störgrößen:

$$E(\varepsilon_i | \mathbf{x}_i, kl_i = 1, \mathbf{z}_i) = E(\varepsilon_i | v_i > -\mathbf{z}'_i \boldsymbol{\gamma}) = \theta \lambda_i^1 \quad (4.31)$$

$$E(\varepsilon_i | \mathbf{x}_i, kl_i = 0, \mathbf{z}_i) = E(\varepsilon_i | v_i \leq -\mathbf{z}'_i \boldsymbol{\gamma}) = \theta \lambda_i^0 \quad (4.32)$$

mit  $\theta = \rho_{\varepsilon v} \sigma_\varepsilon$ . Dabei ist  $\rho_{\varepsilon v}$  der Korrelationskoeffizient zwischen den Störgrößen und  $\sigma_\varepsilon$  die Standardabweichung der Störgröße aus der Selektionsgleichung. Die  $\lambda_i$  sind wie folgt bestimmt:

$$\lambda_i^1 = \frac{\phi(-\mathbf{z}'_i \boldsymbol{\gamma})}{1 - \Phi(-\mathbf{z}'_i \boldsymbol{\gamma})} = \frac{\phi(\mathbf{z}'_i \boldsymbol{\gamma})}{\Phi(\mathbf{z}'_i \boldsymbol{\gamma})} \quad (4.33)$$

und

$$\lambda_i^0 = \frac{\phi(-\mathbf{z}'_i\boldsymbol{\gamma})}{\Phi(-\mathbf{z}'_i\boldsymbol{\gamma})} = \frac{\phi(\mathbf{z}'_i\boldsymbol{\gamma})}{1-\Phi(\mathbf{z}'_i\boldsymbol{\gamma})} \quad (4.34)$$

Es bezeichnet  $\phi$  die Dichtefunktion der Standardnormalverteilung und  $\Phi$  deren Verteilungsfunktion. Die  $\lambda$  sind damit eine Funktion der Teilnahmewahrscheinlichkeiten, die auch im RKM eine wichtige Rolle spielen. Die  $\mathbf{z}'_i\boldsymbol{\gamma}$  werden wegen der Verteilungsannahme (Normalverteilung) mithilfe einer Probitregression geschätzt:

$$P(kl_i = 1 | \mathbf{z}_i) = \Phi(\mathbf{z}_i\boldsymbol{\beta}).$$

Die Funktionen 4.33 und 4.34 werden als Selektionskorrektur, Kontrollfunktionen, inverse Mills Ratio oder auch generalisierte Residuen bezeichnet. Eine OLS-Regression, welche die Selektionskorrekturen enthält, liefert dann wieder unverzerrte Schätzungen des kausalen Effektes:

$$y = \alpha^0 + \delta kl_i + \mathbf{x}'_i\boldsymbol{\beta} + \theta\lambda_i + \varepsilon^* \text{ mit } \lambda_i = kl_i\lambda_i^1 + (1-kl_i)\lambda_i^0$$

Die Schätzung des kausalen Effektes kann entweder als zweistufige OLS- oder als Maximum-Likelihood-Schätzung erfolgen.

Es lässt sich zeigen, dass die OLS-Regression ein Spezialfall des HMA ist, bei dem die Störgröße mit  $kl$  unkorreliert ist. Dies wird im folgenden Exkurs kurz dargestellt.

---

### Exkurs: die OLS-Regression als Switching-Regression

Unter den Annahmen, dass die Effekte in  $\mathbf{x}$  und  $\varepsilon$  homogen sind und die Störgröße mit  $kl$  unkorreliert ist, entspricht der HMA der gewöhnlichen OLS-Regression. Daraus folgt, dass der kausale Effekt einer Maßnahme unter bestimmten Bedingungen auch mit einer solchen Regression gemessen werden kann, wie sie üblicherweise in Soziologie und Ökonometrie verwendet wird. Zentrale Voraussetzung ist allerdings die Erfüllung der mit der OLS-Regression verbundenen Annahmen, also des Vorliegens linearer Beziehungen, der Korrektheit der Gauss-Markov-(GM-)Annahmen und vor allem der Annahme, die sich aus den GM-Bedingungen eins und vier ergibt, der Vollständigkeit des Modells hinsichtlich aller relevanten Prädiktoren.

In der Notation des HMA mit  $y$  als abhängiger Variable,  $\mathbf{x}$  als Vektor unabhängiger Variablen und  $kl$  als kausal interessierende Variable, kann dies folgendermaßen ausgedrückt werden:

$$y_i = \begin{cases} y_i^1 = \alpha^1 + \mathbf{x}'_i\boldsymbol{\beta}^1 + \varepsilon_i^1, & \text{falls } kl = 1 \\ y_i^0 = \alpha^0 + \mathbf{x}'_i\boldsymbol{\beta}^0 + \varepsilon_i^0, & \text{falls } kl = 0 \end{cases} \quad (4.35)$$

Damit gilt:

$$y_i^0 = \alpha^0 + (\alpha^1 - \alpha^0)kl_i + \mathbf{x}'_i kl_i (\beta^1 - \beta^0) + \mathbf{x}'_i \beta^0 + kl_i (\varepsilon^1 - \varepsilon^0) + \varepsilon^0 \quad (4.36)$$

Diese Gleichung vereinfacht sich zum einfachen linearen Modell  $y_i = \alpha + \delta kl_i + \mathbf{x}'_i \beta + \varepsilon_i$  wenn vor allem zwei zentrale Annahmen gelten. Erstens muss der Effekt des Treatments über alle Personen homogen sein (homogene Effekte). Dies ist dann der Fall, wenn einerseits die Auswirkung der beobachteten Kovariaten  $\mathbf{x}$  in der Teilnahme-situation und in der Nichtteilnahme-situation gleich ist, also  $\beta^1 = \beta^0 = \beta$ . Andererseits muss dies auch für die unbeobachtete Heterogenität in beiden Situationen gelten  $\varepsilon_i^1 = \varepsilon_i^0 = \varepsilon_i$ . Damit ergibt sich aus Gleichung 4.36 folgende Vereinfachung:

$$y_i = \begin{cases} y_i^1 = \alpha^1 + \mathbf{x}'_i \beta + \varepsilon_i, & \text{falls } kl = 1 \\ y_i^0 = \alpha^0 + \mathbf{x}'_i \beta + \varepsilon_i, & \text{falls } kl = 0 \end{cases} \quad (4.37)$$

Damit lässt sich der kausale Effekt der Maßnahmeteilnahme als Differenz der Konstanten in beiden Regimes interpretieren (Heckman/LaLonde/Smith 1999: 22).

Die zweite Annahme bezieht sich ebenfalls auf die unbeobachtete Heterogenität und besagt, dass diese den für die Störgröße formulierten Gauss-Markov-Bedingungen genügt (vgl. Abschnitt 4.3.1).

Damit ist

$$E(y_i^1 | \mathbf{x}_i) = \alpha^1 + \mathbf{x}'_i \beta$$

und

$$E(y_i^0 | \mathbf{x}_i) = \alpha^0 + \mathbf{x}'_i \beta$$

Die Folge beider Annahmen ist, dass der kausale Effekt durch die Differenz der Regressionskonstanten in beiden Regimes

$$\begin{aligned} MKE(\mathbf{x}_i) &= E(y_i^1 - y_i^0 | \mathbf{x}_i) = \alpha^1 + \mathbf{x}'_i \beta - \alpha^0 - \mathbf{x}'_i \beta = \\ \alpha^1 - \alpha^0 &= \delta = MKE = MKE_T \end{aligned} \quad (4.38)$$

gegeben ist und im Rahmen einer OLS-Regression unverzerrt geschätzt werden kann. Wegen der Annahme homogener Effekte besteht dabei kein Unterschied zwischen dem MKE und dem  $MKE_T$ .

Der HMA kann auch dann verwendet werden, wenn nicht nur das Treatment, sondern auch die Zielvariable binär ist (Angrist 1991: 16 f.; Greene 2003: 790; Neal 1997; Evans/Schwab 1995). In diesem Fall werden sowohl das Treatment als auch die Zielvariable als latente Variablen betrachtet.

$$\tilde{y}_i = \begin{cases} 1, & \text{falls } \alpha + \delta kl_i + \mathbf{x}'_i \boldsymbol{\beta} - (\varepsilon_i^* - \varepsilon_i \theta) > 0 \\ 0, & \text{sonst} \end{cases} \quad (4.39)$$

$$\tilde{kl}_i = \begin{cases} 1, & \text{falls } \mathbf{z}'_i \boldsymbol{\gamma} - (v_i - \varepsilon_i) > 0 \\ 0, & \text{sonst} \end{cases} \quad (4.40)$$

Dabei sind  $v_i$ ,  $\varepsilon_i^*$  und  $\varepsilon_i$  trivariat normalverteilte Zufallsvariablen und die beiden kombinierten Störgrößen  $\eta_{i1} = \varepsilon_i^* - \varepsilon_i \theta$  und  $\eta_{i2} = v_i - \varepsilon_i$  ebenfalls normalverteilt mit der Kovarianzmatrix

$$\begin{bmatrix} \sigma_{\varepsilon^*}^2 + \sigma_{\varepsilon}^2 \theta^2 & \sigma_{\varepsilon}^2 \theta \\ \sigma_{\varepsilon}^2 \theta & \sigma_v^2 + \sigma_{\varepsilon}^2 \end{bmatrix}$$

Die Berechnung des kausalen Effektes kann mithilfe eines bivariaten Probit-Modells erfolgen. Der MKE ist dann gegeben durch

$$MKE = E_x \left[ \Phi \left( (\alpha + \delta kl_i + \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i \theta) / \sigma_{\varepsilon^*} \right) - \Phi \left( (\alpha + \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i \theta) / \sigma_{\varepsilon^*} \right) \right]$$

bzw. äquivalent (Angrist 1991)

$$MKE = E_x \left[ \Phi \left( (\alpha + \delta kl_i + \mathbf{x}'_i \boldsymbol{\beta}) / (\sigma_{\varepsilon^*}^2 + \sigma_{\varepsilon}^2 \theta^2) \right) - \Phi \left( (\alpha + \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i \theta) / (\sigma_{\varepsilon^*}^2 + \sigma_{\varepsilon}^2 \theta^2) \right) \right]$$

Die ML-Schätzung eines bivariaten Probit-Modells identifiziert die standardisierten Koeffizienten  $\beta / (\sigma_{\varepsilon^*}^2 + \sigma_{\varepsilon}^2 \theta^2)$  und kann zur Berechnung des kausalen Effektes verwendet werden (Angrist 1991).

Erweiterungen der Regression mit Kontrollfunktion existieren außerdem in Hinsicht auf die gemeinsame Verteilung der unbeobachteten Heterogenität in der Selektions- und der Responsegleichung. Hier wurden z. B. Modelle mit der Annahme einer gemeinsamen Student-t-Verteilung statt der Normalverteilung entwickelt (Heckman/Tobias/Vytlacil 2003) oder semiparametrische Modelle (Frölich 2002).

#### 4.3.2.2 Instrumentvariablen- oder „Two Stage Least Square“-Schätzung

Eine weitere Möglichkeit der indirekten Erfüllung der CIA ist die Instrumentvariablen-schätzung (IV-Schätzung), auch „Two Stage Least Square (2SLS)“-Verfahren genannt. Dieses Verfahren besitzt einige Ähnlichkeiten zur Regression mit Kontrollfunktion, folgt aber einer anderen statistischen Grundlage (Wooldridge 2002: 83 ff., 621 ff.). Für die folgende Darstellung wird zunächst davon ausgegangen, dass sowohl die Variable  $y$  als auch die Variable  $kl$  metrisches Skalenniveau aufweisen. Grundsätzlich macht es in der Instrumentvariablen-schätzung keinen Unterschied, ob die zu instrumentierende Variable metrisch oder binär ist (Wooldridge 2002: 84). Daher wird die Vereinfachung später wieder aufgehoben, zumal im Anwendungsfall beide Variablen binär sind.

Ausgangspunkt der Instrumentvariablen-schätzung ist ein lineares Modell der abhängigen Variable  $y$ :

$$y_i = \delta kl_i + \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i \quad (4.41)$$

Dabei gelten für den Variablenvektor  $\mathbf{x}'_i$  die Standardannahmen der OLS-Regression, also vor allem deren Unabhängigkeit von der Störgröße  $\varepsilon_i$ . Wie beim Kontrollfunktionsansatz wird davon ausgegangen, dass eine selection on unobservables besteht und daher die CIA nicht erfüllt ist. Die Störgröße und die Variable  $kl$  sind weder unbedingte noch bedingte voneinander unabhängig:  $E(\varepsilon_i | \mathbf{x}_i, kl_i) \neq E(\varepsilon_i | \mathbf{x}_i)$ .  $kl_i$  wird daher als endogen bezeichnet. In einem solchen Fall würde jede OLS- oder ML-Schätzung zu verzerrten Schätzungen des kausalen Effektes führen. Um den kausalen Effekt trotzdem messen zu können, benötigt man mindestens eine Instrumentvariable, hier mit dem Vektor  $\mathbf{z}'_i$  bezeichnet. Instrumentvariablen zeichnen aus, dass sie Teil der oben (vgl. Abbildung 12) als kausalirrelevant bezeichneten Selektivität sind. Für Instrumentvariablen oder kurz Instrumente gilt daher einerseits  $E(y_i | kl_i, \mathbf{x}_i, \mathbf{z}_i) = E(y_i | kl_i, \mathbf{x}_i)$ . Andererseits muss das Instrument gleichzeitig Einfluss auf die Selektion in den Kombilohn haben, sodass gilt:  $E(kl_i | \mathbf{z}_i) \neq E(kl_i)$ .

Die reduzierte Form (reduced form)-Gleichung bezeichnet die Darstellung der endogenen Variable als Linearkombination der  $x$ -Variablen und der Instrumente:

$$kl_i = \mathbf{z}'_i \boldsymbol{\gamma} + \mathbf{x}'_i \boldsymbol{\psi} + v_i \quad (4.42)$$

Im Gegensatz zur Selektionsgleichung beansprucht die reduzierte Form keine inhaltliche Interpretierbarkeit. Das zeigt sich auch darin, dass die Gleichung mit dem Vektor  $\mathbf{x}'_i$  auch die Einflussfaktoren auf die Zielvariable enthält, die nicht mit der Maßnahmeselektivität korreliert sind. Das „Two Stage Least Square“-Verfahren besteht nun darin,



1. eine OLS-Schätzung der reduzierten Form durchzuführen und die vorhergesagten Werte  $k_i^l$  zu berechnen,
2. eine OLS-Regression basierend auf der Gleichung 4.41 durchzuführen, bei der  $k_i^l$  durch  $k_i^l$  ersetzt wird.

Grundsätzlich genügt bereits ein Instrument zur Durchführung der 2SLS-Schätzung. Die IV-Schätzung führt zwar nicht zu einer unverzerrten, allerdings aber zu einer konsistenten Schätzung des kausalen Effektes. Dies soll anhand einer Einfachregression gezeigt werden (von Auer 2003; vgl. dazu die OLS-Regression in Abschnitt 4.3.1).

Gegeben ist einerseits ein lineares Modell der Zielvariable  $y$  und eine lineare Projektion der Variable  $x$ :

$$y_i = \alpha + \beta x_i + \varepsilon_i \tag{4.43}$$

$$x_i = \omega + \gamma z_i + v_i \tag{4.44}$$

Unter Verwendung der IV-Schätzung ist die Schätzformel:

$$\hat{\beta}^{IV} = \frac{\hat{Cov}(z, y)}{\hat{Cov}(z, x)} = \frac{\sum (z_i - \bar{z})(y_i - \bar{y})}{\sum (z_i - \bar{z})(x_i - \bar{x})} = \beta + \frac{\sum (z_i - \bar{z})\varepsilon_i}{\hat{Cov}(z, x)} \tag{4.45}$$

Zwar liefert dieser Schätzer Werte, die vom wahren Wert des Parameters  $\beta$  um den Term  $\frac{\sum (z_i - \bar{z})\varepsilon_i}{\hat{Cov}(z, x)}$  abweichen, doch handelt es sich bei der IV-Schätzung um eine konsistente Schätzung, da mit Stichprobenumfang  $n \rightarrow \infty$  die Wahrscheinlichkeit für Abweichungen von  $\beta$  gegen null geht:  $\lim_{n \rightarrow \infty} Pr(|\hat{\beta}^{IV} - \beta| > c) = 0$ .

Für Fragestellungen, in denen die Treatmentvariable bzw. die Zielvariable und die Treatmentvariable binär sind, wird häufig statt der OLS-Regression ein lineares Wahrscheinlichkeitsmodell verwendet. Dieses entspricht dem zweistufigen Verfahren oben, wobei bei der linearen OLS-Regression anstatt einer metrischen eine binäre Zielvariable verwendet wird (Angrist 1991; Evans/Schwab 1995).<sup>44</sup>

Ein dreistufiges Verfahren, das speziell für die Analyse binärer Treatmentvariablen geeignet ist, beschreibt Wooldridge (2002: 623):

44 „Linear IV estimation of average treatment effects in nonlinear models can often be justified“ (Angrist 1991: 24).

1. wird ein Modell für binäre Antwortkategorien (Logit- oder Probit-Modell) der reduzierten Form durchgeführt, und die Wahrscheinlichkeiten  $\hat{P}(kl_i = 1 | \mathbf{x}_i, \mathbf{z}_i)$  werden geschätzt;
2. werden diese dann statt der z-Variablen als Instrumente in einer linearen OLS-Regression der reduzierten Form verwendet, und es werden die vorhergesagten Werte  $k_i^{\hat{}}$  berechnet;
3. wird  $kl_i$  in einer auf Gleichung 4.41 basierenden linearen OLS-Regression (einer kardinalen oder binären Zielvariable) durch  $k_i^{\hat{}}$  ersetzt.

#### 4.3.2.3 Der Preis unbeobachteter Maßnahmeselektivität

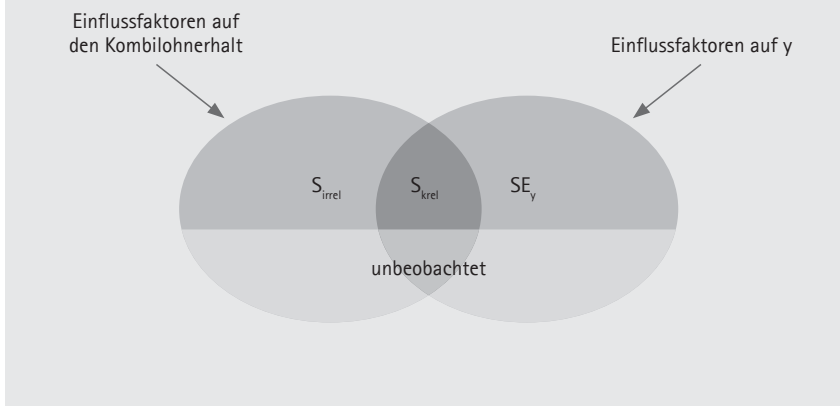
Die in diesem Abschnitt 4.3.2 vorgestellten statistischen und ökonometrischen Verfahren verstehen sich als Lösungen des Problems unbeobachteter kausalrelevanter Selektionseinflüsse und kommen daher grundsätzlich als Alternativen zum RKM infrage. Sowohl in methodischen als auch in anwendungsbezogenen Arbeiten wird dabei oft der Eindruck erweckt, dass bei Unsicherheit bezüglich der CIA der Einsatz dieser Verfahren automatisch empfehlenswert ist. Ebenso scheinen diese Verfahren den Anwender davon zu entlasten, sich mit dem Gegenstand der Forschung und den Selektionseinflüssen möglichst ausführlich zu befassen. Schließlich wird das, was man nicht in der Analyse berücksichtigt hat, automatisch mit der Kontrollfunktion indirekt berücksichtigt oder durch die Verwendung der Instrumentvariablen irrelevant.

In diesem Abschnitt soll gezeigt werden, dass der Schein trügt. Damit Verfahren für unbeobachtete Heterogenität mit weniger Information über die kausalrelevante Maßnahmeselektivität auskommen können, benötigen sie im Gegenzug an anderer Stelle mehr Informationen. Diese Informationen betreffen erstens funktionale Zusammenhänge der Variablen untereinander, welche im PS-Matching nicht nötig sind. Zweitens werden teils funktionale, teils auch andersartige Annahmen über die ausgelassenen Variablen nötig. Zum dritten betreffen die Informationen weitere, im Propensity Score Matching ebenfalls nicht nötige Variablen sowie deren Beziehung zu den unbeobachteten Variablen.

Bei einer Analyse der Wirkung des Kombilohnes auf die Verletzung subjektiver Reziprozitätserwartungen lassen sich die potenziell nötigen Informationen in drei Klassen aufteilen (vgl. Abbildung 14): kausalrelevante und kausalirrelevante Selektionseinflüsse  $S_{krel}$  bzw.  $S_{irrel}$  sowie sonstige Einflussfaktoren auf die subjektive Verletzung der Reziprozitätsnorm  $SE_y$ . Grundsätzlich können nicht nur die kausalrelevanten Selektionseinflüsse unbeobachtet sein, sondern auch die anderen beiden Informationsklassen. Für das PS-Matching wird nur das Fehlen kausalrelevanter Selektionseinflüsse zum Problem (vgl. Abschnitt 4.1.4), also alle die Einflussfaktoren auf den Erhalt der Kombilohnförderung, die gleichzeitig auch Einflussfaktoren

auf  $y$  sind. In der Abbildung 14 bilden sie die Schnittfläche zwischen den beiden Kreisen. Solche Selektionseinflüsse, die keinerlei Einfluss auf die Wirkungsvariable haben, können für die Analyse als zufällig und daher im Mittel gleich auf die Geförderten und Ungeförderten verteilt betrachtet werden. Ihre Berücksichtigung ist daher nicht nötig. Entsprechendes gilt für die Einflussfaktoren auf die Ergebnisvariable, die nichts mit der Kombilohnförderung zu tun haben.

Abbildung 14: Informationsklassen



Die als Alternativen zum RKM vorgestellten Verfahren beanspruchen für sich, einen Teil der  $S_{krel}$  unbeobachtet zu lassen und damit das Problem der CIA zu lösen. Allerdings zeigt eine nähere Betrachtung, dass dies nur um den Preis zusätzlicher Informationen möglich ist, die über das im PS-Matching nötige Wissen weit hinausgehen. Dies soll für die Verfahren der Reihe nach gezeigt werden.

### Lineare Regression mit Kontrollfunktion

In der Regression mit Kontrollfunktion sind zwei Gleichungen nötig. Die erste muss das Zustandekommen der Werte der Zielvariable erklären (Gleichung 4.46). Dazu sind neben der Kombilohnbeteiligung einerseits die  $S_{krel}$  nötig und andererseits die  $SE_y$ , die im Vektor  $\mathbf{x}'_i$  zusammengefasst sind. Neben einer Störgröße  $\varepsilon_i$ , die die Gauss-Markov-Bedingungen erfüllt, enthält die Gleichung noch unbeobachtete Variablen und deren Einfluss  $u_i$ , wobei es sich sowohl um eine einzelne als auch um mehrere Variablen handeln kann.

$$y = \alpha + \delta kl_i + \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i^S + \varepsilon_i^* \quad (4.46)$$

Zweitens muss der Erhalt einer Kombilohnförderung erklärt werden:

$$kl_i = \omega + \mathbf{z}'_i \boldsymbol{\gamma} + \nu_i \quad (4.47)$$

Dazu müssen sowohl die  $S_{krel}$  als auch die  $S_{irrel}$  im Vektor  $\mathbf{z}'_i$  der Gleichung 4.47 enthalten sein. Daneben können auch hier eine Variable oder mehrere Variablen  $\nu_i$  unbeobachtet bleiben.

Aus den Annahmen des Modells (vgl. Abschnitt 4.3.2.1) lässt sich ableiten, dass die Störgröße  $\varepsilon_i^S$  in Gleichung 4.46 – wegen des nötigen Einflusses auf die Maßnahmeteilnahme und der nötigen Unabhängigkeit von den übrigen enthaltenen Variablen – nur aus ausgelassenen  $S_{krel}$  bestehen darf. Im Erklärungsmodell der Wirkung müssen somit alle  $SE_y$  enthalten sein. Dies setzt natürlich voraus, dass diese alle dem Forscher bekannt sind und in den Daten gemessen wurden. Die Variablen in der Störgröße  $\nu_i$  können dagegen sowohl aus ausgelassenen  $S_{krel}$  als auch ausgelassenen  $S_{irrel}$  bestehen.

Aus den statistischen Annahmen des Modells folgt weiterhin, dass sowohl  $\varepsilon_i^S$  als auch  $\nu$  unabhängig von den in den beiden Regressionen enthaltenen Variablen sein müssen. Kennzeichnet man enthaltene Variablen mit einem Asterisk, dann gilt:  $\varepsilon_i^S, \nu \perp S_{krel}^*, S_{irrel}^*, SE_y$ , wobei für die sonstigen Einflussfaktoren auf  $y$  immer noch gilt, dass alle in der Gleichung 4.46 enthalten sein müssen. Spätestens an dieser Stelle wird klar, dass die Regression mit Kontrollfunktion ein intensives Wissen über die ausgelassene kausalrelevante Selektivität benötigt, nicht nur, was die Tatsache betrifft, welche Variablen ausgelassen wurden, sondern auch, in welcher Beziehung sie zu anderen ausgelassenen und auch in welcher Beziehung sie zu den in den Gleichungen 4.46 und 4.47 enthaltenen Variablen stehen.

Verstärkend kommt hinzu, dass die Regression mit Kontrollfunktion nur unverzerrte Schätzungen des kausalen Effektes liefert, wenn die in Gleichung 4.46 und 4.47 ausgelassenen Variablen(bündel) einer bivariaten Normalverteilung folgen:

$$(\varepsilon_i^S, \nu) \sim N(0, 0, \sigma_u^2, \sigma_v^2, \rho_{uv})$$

Es zeigt sich damit, dass ein Auslassen kausalrelevanter Selektivität in der Regression mit Kontrollfunktion nur unter der Einführung zahlreicher Annahmen möglich ist. Diese Annahmen sind zum Teil ebenso wenig überprüfbar wie die CIA, betreffen jedoch wesentlich mehr Aspekte des Gegenstandsbereiches.

#### *Instrumentvariablenregression*

Ähnlich wie in der Regression mit Kontrollfunktion setzt auch die Instrumentvariablen-schätzung unüberprüfbares Wissen voraus. Dieses bezieht sich hauptsächlich

auf den Zusammenhang zwischen dem Instrument und eventuellen ausgelassenen Variablen. Gegeben sei (vgl. Wooldridge 2002: 101 f.) eine einfache Regression mit  $y = \alpha_0 + \delta k_i + u_i$ , wobei  $u_i$  eine unbeobachtete Variable ist. Mit  $z_i$  als Instrument gilt  $p \lim \hat{\delta} = \delta + \frac{\text{Cov}(z, u)}{\text{Cov}(z, k)}$ . Daher muss die Kovarianz zwischen dem Instrument und einem unbeobachteten kausalrelevanten Selektionseinfluss  $\text{Cov}(z, u) = 0$  bekannt sein bzw. angenommen werden, damit eine konsistente Schätzung des kausalen Effektes erfolgt. Darüber hinaus ergeben sich für die IV-Schätzung Probleme bei geringer Fallzahl, da die Argumentation bezüglich der Konsistenz des Schätzers auf den Wahrscheinlichkeitslimes (Stichprobenumfang  $n \rightarrow \infty$ ) Bezug nimmt.

In verschiedenen Untersuchungen zeigte sich zudem, dass die Schätzung eines kausalen Effektes mit Instrumentvariablen, die nur schwach mit der Treatmentvariable korreliert sind, problematisch ist. Dies gilt auch bei Fallzahlen, die relativ hoch sind, z. B. bis zu 500.000 (Wooldridge 2002: 104).

#### 4.4 Zusammenfassung

In diesem Kapitel wurde Rubins Kausalmodell als konzeptioneller Rahmen für die Kausalanalyse auch im Bereich soziologisch interessanter Fragestellungen vorgestellt. Es wurde darauf hingewiesen, dass die nicht parametrische Umsetzung des RKM durch das Propensity Score Matching auf bestimmte Annahmen beruht, vor allem der Stable Unit Treatment Value Assumption und der Conditional Independence Assumption. Letztere geht davon aus, dass in der Kausalanalyse alle relevanten Selektionsvariablen berücksichtigt wurden. Da diese Annahme oft als unüberprüfbar kritisiert wird, wurden alternative Verfahren vorgestellt, welche diese Annahme nicht treffen. Es hat sich jedoch gezeigt, dass auch die alternativen Verfahren „Regression mit Kontrollfunktion“ und „Instrumentvariablenschätzung“ bestimmte Annahmen treffen, deren Zutreffen ebenfalls nicht zu überprüfen ist. Daher wird zur empirischen Analyse der kausalen Zusammenhänge zwischen Kombilohn, Lohnzufriedenheit und Beschäftigungsstabilität folgendes Vorgehen gewählt. Zunächst wird in Abschnitt 5 eine möglichst umfassende Untersuchung der Selektivität in die Kombilohnförderung angestrebt, um die CIA zu rechtfertigen. In Abschnitt 6 wird dann die Kausalanalyse mit dem Propensity Score Matching durchgeführt, wobei die parallele Durchführung einer Regression mit Kontrollfunktion und einer Instrumentvariablenschätzung Aufschluss über die Sensitivität der Ergebnisse in Bezug auf die Verfahrenswahl geben wird (Abschnitt 6.5.1.3). Zusätzlich wird noch eine Reihe weiterer Tests durchgeführt, welche die CIA und andere relevante Verfahrensannahmen zwar nicht beweisen können, für ihre Gültigkeit jedoch gute Argumente liefern können (Abschnitt 6.5).



## 5 Selektivität des Kombilohnbezuges

Zentrale Voraussetzung zur Messung kausaler Effekte ist die Erfassung der kausalrelevanten Selektivität. Durch ihre Aufnahme als Variablen zur Konstruktion des Propensity Score wird die Conditional-Independence-Annahme erfüllt ist. Dies setzt dann voraus, dass diese Einflüsse dem Forscher auch bekannt sind. Die zentrale Herausforderung der Kausalanalyse ist, dass die zur Erfüllung der CIA nötige Vollständigkeit der  $S_{krel}$  nicht zu beweisen ist.

Wenn aber der faktische Erfolg bei der Erfüllung der CIA nicht zu überprüfen ist, welcher Maßstab sollte dann an eine Analyse gelegt werden? Auf diese Frage gibt es in der Forschungspraxis keine eindeutige Antwort. In dieser Arbeit wird folgender Standpunkt vertreten: Da die faktische Erfüllung der CIA nicht zu beweisen ist, kommt es in zentraler Weise darauf an, den Prozess der Suche nach den kausalrelevanten Selektivitäten so zu gestalten, dass keine Aspekte der Maßnahmeselektivität von vornherein systematisch ausgeblendet werden. Daher wird zunächst das noch junge Forschungsgebiet nach möglichen Strategien bei der Identifikation der kausalrelevanten Maßnahmeselektivität untersucht und auf ihren Nutzen für die vorliegende Arbeit geprüft (Abschnitt 5.1). Daraus wird die Konsequenz gezogen, dass die Identifikation der Maßnahmeselektivität sich möglichst genau mit den Handlungsbedingungen der an der konkreten Maßnahme beteiligten Akteure auseinandersetzen sollte. Diese Konsequenz wird im Abschnitt 5.2 für die Analyse der Maßnahmeselektivität im Mainzer Modell umgesetzt.

### 5.1 Mögliche Strategien zur Identifikation von Maßnahmeselektivitäten

In der Forschungspraxis sind bei der Suche nach der Maßnahmeselektivität sehr verschiedene Vorgehensweisen anzutreffen. Eine Literaturrecherche zeigt, dass sich diese vier unterschiedlichen Strategien zuordnen lassen, wobei einzelne Forschungsprojekte zum Teil auch mehrere dieser Strategien gleichzeitig nutzen. Die Identifikation relevanter Selektionseinflüsse kann erstens auf Basis formal-mathematischer allgemeiner ökonomischer Modelle erfolgen (Abschnitt 5.1.1). Zweitens basiert die Suche nach den CIA-Variablen auf allgemeinen Überlegungen zu den Bedingungen von Maßnahmeselektivität, die man als Ansätze zu einer allgemeinen, von einer konkreten Maßnahme unabhängigen Theorie der Maßnahmeselektivität betrachten kann (Abschnitt 5.1.2). Eine dritte Strategie ist die Durchführung qualitativ-empirischer Studien zur konkreten Maßnahme (Abschnitt 5.1.3), und die vierte Strategie besteht schließlich in der Orientierung an bereits durchgeführten Forschungsarbeiten zu ähnlichen Maßnahmen (Abschnitt 5.1.4).

### 5.1.1 Formal-mathematische Modelle

Die Strategie zur Bestimmung der kausalrelevanten Selektivität, die vor allem von James Heckman vertreten wird, ist das Aufstellen eines formalen Modells zur Erklärung der Maßnahmeteilnahme, meist unter expliziter Bezugnahme auf das neoklassische Handlungsmodell der Nutzenmaximierung und ökonomische Theorien wie die Suchtheorie (Mortensen 1970; McCall 1970) oder die Humankapitaltheorie (Mincer 1974). Es wird eine Entscheidungssituation unterstellt, in der der Maßnahmeteilnehmer die Alternative wählt, von der er sich den höheren Nutzen erwartet. Da sich für Kombilohnmaßnahmen kein Beispiel aus der Literatur finden lässt, wird dieses Vorgehen anhand der Teilnahme an einer Trainingsmaßnahme ( $tm$ ) exemplifiziert (Heckman/LaLonde/Smith 1999: 50 ff.). Zur Vereinfachung der Notation wird auf den Personenindex  $i$  verzichtet. Es sei die Fördermaßnahme eine Trainingsmaßnahme und das Maßnahmeziel das nach ihrem Abschluss erzielte Einkommen. Die Erwerbsbiografie beginnt zur Erwerbsperiode 1 ( $t = 1$ ), und die Weiterbildung ist nur in einer Periode, der Periode  $t = k$ , möglich. Das Einkommen ohne Weiterbildung ist  $y_j^0$  mit  $j = 1, \dots, k$ . Nach dem möglichen Weiterbildungszeitpunkt  $k$  sind die potenziellen Ergebnisse  $(y_j^0, y_j^1)$  mit  $j = k + 1, \dots, T$ , wobei  $T$  die letzte Periode im Erwerbsleben ist. Welches Einkommen zur Periode  $t$  beobachtet wird, ist von der Teilnahme oder Nichtteilnahme an der Trainingsmaßnahme<sup>45</sup> abhängig:

$$y_t = tm y_t^1 + (1 - tm) y_t^0, \quad tm = \begin{cases} 1 \\ 0 \end{cases}$$

Handlungstheoretische Basis für die Bestimmung der Elemente der Indexfunktion bzw. Selektionsgleichung ist die Modellierung der Teilnahme als Entscheidungssituation und die Annahme, dass Individuen ihre Entscheidungen so treffen, dass sie ihren Nutzen maximieren, wobei der Nutzen hier im Einkommen besteht (Heckman/LaLonde/Smith 1999: 50). Entsprechend wird auch die Teilnahmeentscheidung dem betroffenen Individuum zugeschrieben, andere Einflüsse (z. B. Arbeitsverwaltung oder Familie) werden explizit ausgeschlossen. Unter der Annahme perfekter Kapitalmärkte und der Fähigkeit der Individuen zur korrekten Vorhersage ihrer zukünftigen Einkommen ergibt sich, dass eine Person teilnimmt ( $tm = 1$ ), wenn die erwartete Differenz der Einkommen ohne und mit Trainingsmaßnahme abzüglich der Kosten  $c$  für die Maßnahme selbst (Opportunitätskosten, eventuelle Gebühren etc.) positiv ist. Formal lässt sich dies so ausdrücken:

45 Trainingsmaßnahmen sind relativ kurze Maßnahmen, in denen Arbeitslosen oder Arbeitssuchenden bestimmte Kenntnisse vermittelt werden sollen.



$$E \left[ \sum_{j=1}^{T-k} \frac{y_{k+j}^1}{(1+r)^j} - c - \sum_{j=1}^{T-k} \frac{y_{k+j}^0}{(1+r)^j} \middle| I_k \right] \geq 0$$

Dabei ist  $r$  die Zinsrate und  $I_k$  sind die Informationen zum Zeitpunkt  $k$ , auf deren Basis die Individuen für sich den Erwartungswert berechnen. Die  $I_k$  können sich von Person zu Person unterscheiden.<sup>46</sup>

Aus dieser Entscheidungsregel werden von Heckman/LaLonde/Smith (1999) folgende Aussagen abgeleitet:

- Je älter die Personen und je höher die Abschreibungsrate, desto geringer die Teilnahmewahrscheinlichkeit.
- Das Einkommen vor Periode  $k$  ist unwichtig für die Teilnahmeentscheidung, es sei denn als Teil von  $I_k$ , wo es dazu dient, zukünftige Einkommen zu prognostizieren.
- Nur die Kosten  $c$  und die entgangenen Zinserträge aus dem Einkommen bestimmen die Teilnahme. Je geringer das Einkommen vor Periode  $k$  und je geringer die Teilnahmekosten, desto wahrscheinlicher die Teilnahme.
- Jede Verletzung der Unabhängigkeit von Teilnahmeentscheidung und realisiertem Maßnahmeergebnis basiert ausschließlich auf dieser Entscheidungsregel. „It is the relationship between expected outcomes at the time decisions are made and the realized outcomes that generate the structure of the bias“ (ebd. 51).

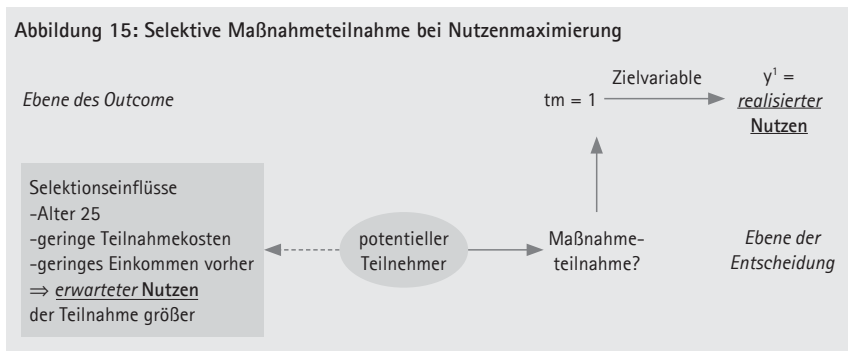
Trifft man zusätzlich noch die Annahme konstanter kausaler Effekte  $y_k^1 - y_k^0 = \delta_t, t > k$ , dann ergibt sich als vereinfachte Entscheidungsregel:

$$tm = \begin{cases} 1, & \text{wenn } E \left[ \sum_{j=1}^{T-k} \frac{\delta_{k+j}}{(1+r)^j} - c - y_k^0 \middle| I_k \right] \geq 0 \\ 0 & \text{sonst} \end{cases}$$

Individuen orientieren sich bei der Entscheidung für oder gegen eine Maßnahmeteilnahme an einem Vergleich der Kosten der Maßnahmeteilnahme (inklusive der Opportunitätskosten während der Maßnahmeperiode  $y_k^0$ ) mit dem aufsummierten erwarteten zukünftigen Einkommen  $E \left[ \sum_{j=1}^{T-k} \frac{\delta_{k+j}}{(1+r)^j} \middle| I_k \right]$ . Letztere sind für Per-

46 Dadurch erhält das Modell grundsätzlich einen Platzhalter für unterschiedliche handlungstheoretische Herangehensweisen, diese werden allerdings nicht genutzt.

sonen mit gleichem Alter und gleicher Abschreibungsrate identisch. Zukünftiges Einkommen ohne Weiterbildung hat keinen Einfluss, außer über  $y_k^0$ . Aus dieser Formulierungsregel folgt dann, dass die Selektivität der Maßnahmeteilnahme „comes through the dependence with  $y_k^0$  and any dependence on costs  $c$ “ (ebd. 52). Aufgabe des evaluierenden Forschers muss es nun sein,  $\delta$ ,  $c$  bzw.  $y_k^0$  direkt oder indirekt, also über Proxyvariablen oder mithilfe eines Verfahrens für unbeobachtete Heterogenität, zu messen und in die Indexfunktion zu integrieren, um damit die Selektivität des Maßnahmezugangs zu berücksichtigen.



In der Ökonomie sind derartige Modelle oft anzutreffen. Die dabei typische Formalisierung und Vereinfachung dient der Konzentration einer Erklärung auf die wesentlichen Elemente eines Phänomens. Im Anschluss an Milton Friedman wird in der Ökonomie die These vertreten, dass es bei Erklärungen nicht um den Realitätsgehalt der dem Modell zugrunde liegenden Annahmen wie der vollständigen Information, der rationalen Wahl zwischen Alternativen etc. ankomme. Wichtig sei vielmehr, ob die aus einem Erklärungsmodell abgeleitete Prognose zutrifft (Friedman 1953; Brüderl 2004: 177). Aus soziologischer Sicht lässt sich hieran die Kritik formulieren, dass die Ökonomie „größeren Wert auf die Axiomatisierbarkeit und Formalisierbarkeit dieser Annahmen gelegt [hat] als auf ihre inhaltliche Angemessenheit“ (Albert 1999: 223). Eine solche Strategie läuft aus zwei Gründen Gefahr, bei der Messung des kausalen Effektes ins Leere zu laufen. Zum einen kommt es vor allem auf die Erfassung der empirisch wirksamen Abläufe an, um Verzerrungen aus der statistischen Messung zu eliminieren. Wichtiger als eine formal stimmige Erklärung ist daher die empirisch stimmige Deskription des Maßnahmezugangs. Der Begriff Deskription darf hier allerdings nicht mit einer theorieleeren Beobachtung verwechselt werden. Zum anderen ist das Ziel der Erfassung der Maßnahmeselektivität nicht die Prognose der Maßnahmeteilnahme, für die sowohl die kausalrelevanten als auch irrelevanten Selektionseinflüsse nötig wären, sondern ausschließlich der Ausgleich der kausalrelevanten Unterschiede zwischen Geförderten und Ungeförderten.

Lindenberg, selbst Vertreter der Rational-Choice-Theorie, hebt die Konsequenzen solcher Vereinfachungen treffend hervor: „Simplification has a powerful influence on the description of reality that is to be explained rather than just an influence on the explanation of this reality“ (Lindenberg 1998: 60). Das zu Erklärende verschwindet hinter dem, was die Theorie als zu Erklärendes formalisiert.

Leicht lassen sich auch in Heckmans formaler Darstellung der Teilnahmeentscheidung solche Vereinfachungen finden, nicht zuletzt auch deshalb, weil sie bewusst getroffen werden, um die Erstellung eines konsistenten, mathematisch formulierbaren und statistisch schätzbaren Modells zu ermöglichen. Heckmans Modell beginnt mit folgender Annahme: „To simplify the discussion (...) assume that persons have only one period in their lives (...) where they have the chance to take job training“ (Heckman/LaLonde/Smith 1999: 50). Weiter heißt es: „(The model) ignores family and bureaucratic influences on participation decisions“ (ebd.) und „Assume, that credit markets are perfect so that the agents can lend and borrow freely at interest rate  $r$ “ (ebd.). Damit das theoretische Modell als Probit- oder Logitregression geschätzt werden kann, wird zusätzlich angenommen, dass erstens der kausale Effekt für alle Individuen und alle Erwerbsperioden  $t$  – gegeben bestimmte Kovariatenkonstellationen – gleich ist (konstanter kausaler Effekt), dass zweitens die Zahl der Erwerbsperioden gegen unendlich geht und drittens „agents are assumed to be farsighted, and possess the ability to make accurate forecasts“ (ebd. 52) etc.

Indem der Formalisierung dienliche simplifizierende Annahmen getroffen werden, erscheint die Frage der Teilnahme oder Nichtteilnahme dann erstens allein als eine Frage des (finanziellen) Nutzens der einzelnen Alternativen und ist zweitens in einem einfachen statistischen Modell abbildbar. Da die Definition des Explanandums als erwartete Nutzendifferenz die Suche nach den Kovariaten anleitet, bleibt dabei völlig unklar, welche Konsequenzen diese simplifizierenden Annahmen auf die Identifikation der Selektionseinflüsse haben. Was, wenn die Individuen nicht „farsighted“ sind oder wenn nicht nur die Individuen selbst, sondern auch ihre Familienangehörigen mitentscheiden. Welchen Einfluss hat der Fallmanager in der Arbeitsagentur, der in vielen Fällen als „gate keeper“ wirkt? „Thus, if I simplify by disregarding important tendencies, I also disregard the problems people will deal with in their dealings“ (Lindenberg 1998: 63). Die Folge davon ist, dass bei der Identifikation der Selektionseinflüsse ein Teil der tatsächlich auftretenden Selektionsverzerrungen ausgeblendet wird.

In der Forschungspraxis sind solche formalen Modelle der Teilnahmeentscheidung eher selten anzutreffen. Stattdessen werden die Determinanten der Teilnahmeentscheidung meist verbal formuliert. Auch dieses Vorgehen orientiert sich jedoch letztlich stark an formalisierten ökonomischen Theorien, z. B. der Humanka-

pitaltheorie (Reinowski/Schultz 2006: 12) oder einer Vielzahl anderer Theorien (z. B. Jirjahn/Pfeiffer/Tsertsvadze 2006, vgl. unten).

### 5.1.2 Ansätze einer allgemeinen Theorie der Maßnahmeselektivität

Neben der Orientierung an ökonomischen Theorien besteht eine andere Möglichkeit der Bestimmung von Selektionseinflüssen darin, die nötigen Kontrollvariablen aus einer allgemeinen Theorie der Maßnahmeselektivität abzuleiten. Allgemein ist eine solche Theorie in dem Sinne, dass sie losgelöst von den Besonderheiten der einzelnen Maßnahmen der Arbeitsmarktpolitik die Handlungsmuster von potenziellen Maßnahmeteilnehmern und Beschäftigten der Arbeitsagenturen abbildet, die bei der Selektion in die Maßnahme eine Rolle spielen. Solche Theorien zeichnen – wie es Lechner, Miquel und Wunsch (2005a: 24) formulieren – den „average' selection process“ nach. Bisher existiert eine solche allgemeine Theorie (noch) nicht, in der Literatur lassen sich aber einige Ansätze dazu finden, die hier dargestellt werden sollen.

So betrachten Caliendo, Hujer und Thomsen (2004, 2005a, 2005b, 2005c) Maßnahmeselektivität als eine Teilnahmeentscheidung, die sich an vier Merkmalsgruppen orientiert. Die erste Merkmalsgruppe bilden „soziodemografische Hintergrundvariablen“ der Individuen wie Geschlecht, Alter oder Familienstand. Zur zweiten Merkmalsgruppe, die mit „Qualifikation“ überschrieben ist, zählen etwa die vorhandene Berufserfahrung oder die Stellung im Beruf. Die dritte Merkmalsgruppe „Erwerbskarriere“ deckt die Arbeitsmarkthistorie des Arbeitsuchenden ab, z. B. die Dauer bisheriger Arbeitslosigkeit oder frühere Teilnahmen an Maßnahmen. Schließlich umfasst die vierte und letzte Merkmalsgruppe den „regionalen Kontext“ (Caliendo/Hujer/Thomsen 2004: 14 ff.; 2005a: 16 f.; 2005b: 8 f.; 2005c: 11 ff.). Ähnlich gehen auch Fitzenberger, Osikominu und Völter (2006) vor. Sie unterscheiden „personal characteristics“, „information about the last employer“, z. B. Branche oder Betriebsgröße, „information on earnings“, „employment and program participation history“ und „regional labor market conditions“ (ebd. 27 f.).

Lechner und Wunsch (2006a: 11, 2006b: 20) sowie Lechner, Miquel und Wunsch (2005a: 23 f., 2005b: 25) verweisen zum einen darauf, dass neben dem Arbeitsuchenden auch der jeweilige Fallmanager zur Maßnahmeselektivität beiträgt. Sie betrachten daher die Zuweisung (oder Fremdselektion) einerseits und die Selbstselektion andererseits als zentrale Aspekte des Selektionsprozesses. In ähnlicher Weise basiert die Modellierung des Entscheidungsprozesses bei Biewen et al. (2006) darauf, dass die Maßnahmeselektivität vor allem auf „motivation, plans and labor market prospects of the unemployed and the way they are perceived by the caseworker in the labor office“ (Biewen et al. 2006: 20) zurückgeht. Anzumerken ist dabei, dass trotz Berücksichtigung eines zweiten Akteurs der Fo-

kus auf den Eigenschaften des Arbeitssuchenden liegt, wenn auch nun aus Sicht der Fallmanager.

In einer Untersuchung der Schweizer Arbeitsverwaltung wird hinsichtlich der Frage „What is the Value Added by the Caseworker“ von Lechner und Smith (2003) festgestellt: „First, swiss caseworkers are making use of the flexibility available to them (...). Second, the caseworkers do not allocate persons at random with respect to their observed characteristics. (...) Third, the caseworkers' allocation shows evidence of systematic, reasonable patterns. It makes sense to assign foreigners to language courses and the most qualified among the unemployed to computer courses“ (Lechner/Smith 2003: 7). Sowohl Lechner et al. (2005a, b) als auch Dyke et al. (2005: 9) verweisen zusätzlich auf die Wirksamkeit von formalen Regelungen des Maßnahmezuganges („eligibility“, „program rules“) als zentrale Determinanten der Maßnahmeteilnahme.

In einer Analyse der Auswirkungen der Hartz-Reformen auf die Arbeitsvermittlung stellen Behrend et al. (2006) fest, dass Fallmanager bei ihrer Einschätzung, welche Arbeitssuchenden welche Leistungen bzw. Maßnahmen erhalten sollten, auf eine „Alltagspädagogik“ zurückgreifen. Diese kann unterschiedliche Ausformungen haben, von der Motivation und Aktivierung der Eigenbemühungen der Arbeitslosen („Kampf gegen Resignation“) über Strategien der Herstellung persönlicher Beziehungen („Kooperation“) bis zur Vermittlung der Einsicht, dass sich die Arbeitslosen bestimmter gesetzlicher Vorgaben nicht entziehen können („Einsicht in das Notwendige“).

Wie aus der Zusammenstellung deutlich wird, ist es bis zu einer allgemeinen Theorie der Maßnahmeselektivität noch ein längerer Weg, bisher sind hierzu nur Ansätze vorhanden. Es besteht also noch Bedarf an weiterer Forschungsarbeit. Allerdings ist zu erwarten, dass eine solche allgemeine Theorie nur einen Rahmen bilden kann, da konkrete Maßnahmen stets Besonderheiten aufweisen und so nicht nur zu maßnahmeartspezifischen (z. B. Weiterbildungsmaßnahmen oder Kombilohnmodelle), sondern auch zu unmittelbar maßnahmespezifischen Besonderheiten führen. Die Ausgestaltung jeder einzelnen Maßnahme interagiert auf spezifische Weise mit den Handlungsmustern der beteiligten Akteure, also der Arbeitssuchenden, Fallmanager in den Arbeitsagenturen und (bei Kombilohnförderungen, vgl. unten) einstellenden Betriebe. Aus diesem Grund ist neben der Orientierung an ökonomischen Theorien und an einer allgemeinen Theorie der Maßnahmeteilnahme eine dritte Möglichkeit die empirische Analyse der betrachteten Maßnahme in Hinblick auf die konkreten Handlungsmuster der beteiligten Akteure.

### 5.1.3 Empirisch-qualitative Studien

Die Durchführung qualitativer Studien zur Aufdeckung der Selektionsprozesse ist vergleichsweise selten und wird meist im Rahmen der Implementationsforschung

durchgeführt. Typischerweise bestehen Maßnahmeevaluationen aus vier Teilbereichen (z. B. Kruppe 2006). Neben der Messung des kausalen Maßnahmeeffektes auf individueller Ebene, wie sie auch in dieser Arbeit vorgenommen wird, beinhalten sie eine Betrachtung von Makroeffekten der Maßnahme, eine Kosten-Nutzenanalyse und eine Implementationsanalyse. Letztere wird meist auf Basis nicht oder teilstandardisierter Erhebungen (z. B. Experteninterviews, Gruppendiskussionen etc.) durchgeführt und befasst sich neben der Analyse der Schwierigkeiten bei der Umsetzung politischer Vorgaben vor Ort auch mit den Handlungslogiken der Akteure bei der Gewährung oder Nichtgewährung von Maßnahmen bzw. der Teilnahmeentscheidung der Arbeitsuchenden oder Entscheidung gegen eine Teilnahme.

Die Vorteile qualitativer Analysen der Maßnahmeselektivität liegen vor allem darin, dass sie im Gegensatz zu allgemeinen theoretischen Analysen der Maßnahmeselektivität auf die Umsetzungsbedingungen der konkreten Maßnahme eingehen. Dabei können qualitative Analysen auch Sachverhalte aufdecken, die nur für diese eine Maßnahme und die konkreten Rahmenbedingungen der Implementation dieser konkreten Maßnahme gelten. Dies wird im Bereich der Messung kausaler Maßnahmeeffekte allerdings oft nur eingeschränkt genutzt. Zum einen werden qualitative Studien meist nur in Großprojekten zur Evaluation aktiver Arbeitsmarktpolitik realisiert. Zum anderen werden auch dann, wenn nicht standardisierte Untersuchungen durchgeführt werden, deren Ergebnisse selten systematisch in die statistische Analyse aufgenommen. Dafür gibt es wohl mehrere Gründe.

Ein Grund ist der „time lag“ der Untersuchungen. Da qualitative Analysen zeitintensiv sind, liegen die Ergebnisse der Implementationsforschung meist erst dann vor, wenn schon ein Großteil der standardisierten Erhebungen zur Messung des kausalen Effektes abgeschlossen ist. So können die Ergebnisse der Implementationsforschung nur bedingt in die Kausalanalyse einfließen, das heißt, nur insofern bereits bei der Konzeption des standardisierten Erhebungsinstrumentes daran gedacht wurde.<sup>47</sup> Ein anderer Grund ist, dass qualitative Untersuchungen und statistische Analysen auf unterschiedlichen Forschungstraditionen beruhen. Erstere werden oft von Politologen oder Soziologen durchgeführt und Letztere oft von Ökonomen. Hinzu kommt, dass der inhaltliche Fokus der Implementationsforschung, welche die qualitativen Untersuchungen durchführt, oft nicht nur oder nicht einmal hauptsächlich auf der Maßnahmeselektivität liegt, sondern auf den Problemen der Akteure in den Arbeitsagenturen bei der Umsetzung der Maßnahmen.

Ein weiterer wichtiger Grund ist sicherlich, dass in der Arbeitsmarktforschung zur Messung kausaler Maßnahmeeffekte vor allem die im Rahmen der Verwal-

---

<sup>47</sup> Dies widerspricht natürlich gerade einem zentralen Mehrwert qualitativer Analysen, nämlich dem Aufdecken bisher nicht bekannter Sachverhalte.

tungstätigkeit der Bundesagentur für Arbeit anfallenden Prozessdaten verwendet werden. Dabei handelt es sich um Geschäftsdaten der BA, die vom Institut für Arbeitsmarkt- und Berufsforschung (IAB) gesammelt und für wissenschaftliche Analysen aufbereitet werden.

Da im Rahmen der Evaluation des Mainzer Modells eine Befragung von Teilnehmern und Nichtteilnehmern mit umfangreichem Fragenspektrum durchgeführt wurde, gilt diese Einschränkung für die hier durchgeführte Analyse nicht. Statt der losen Verknüpfung (die oft lediglich ein Nebeneinander darstellt) zwischen der qualitativen Untersuchung der Maßnahmeselektivitäten und der statistischen Analyse des kausalen Maßnahmeeffektes wird hier eine direktere Orientierung der Variablenwahl für die statistische Analyse an den Ergebnissen einer im Rahmen der Implementationsanalyse zum Mainzer Modell durchgeführten qualitativen Untersuchung angestrebt.

#### 5.1.4 Untersuchungen zu ähnlichen Maßnahmen

Eine letzte Möglichkeit, sich den Selektivitäten einer Maßnahme zu nähern, ist die Orientierung an Untersuchungen anderer, ähnlicher Maßnahmen (z. B. Mueser/Troske/Gorislavsky 2003: 16; Reinowski/Schultz/Wiemers 2003: 19). Dieses Vorgehen ist insofern sinnvoll, als man annehmen kann, dass die beteiligten Akteure bei ähnlichen Maßnahmen mit ähnlichen Rahmenbedingungen konfrontiert sind, ähnliche Vor- und Nachteile in einer Maßnahmeteilnahme sehen etc.

In den vergangenen Jahren wurden neben dem Mainzer Modell auch eine Reihe anderer Kombilohnmodelle erprobt, die als Lieferanten für Informationen zu den Selektionsprozessen infrage kommen.<sup>48</sup> Allerdings sollte der Nutzen solcher Untersuchungen ähnlicher Maßnahmen nicht überschätzt werden. Da die Maßnahmen stets in nationale Institutionen eingebunden sind, ist auch die Bezugnahme auf internationale Forschung oft nur bedingt hilfreich. So reagieren Kombilohnmodelle vor allem auf die in Deutschland als problematisch betrachtete hohe Transferentzugsrate (Walwei 2002) und müssen in das deutsche Arrangement der Arbeitsmarktinstitutionen, vor allem der Lohnfindung und der Arbeitslosenversicherung, eingepasst werden. Auch sind sie von bestimmten kulturellen Vorstellungen über die Rolle des Staates in Bezug auf den Arbeitsmarkt (Esping-Andersen 1991) geprägt.

Auch die vollständige Vergleichbarkeit des Mainzer Modells (MZM) mit anderen Kombilohnmodellen in Deutschland ist nicht unbedingt gesichert und abhängig von einer Reihe von Entscheidungen, die für jede Kombilohnförderung zu treffen waren (vgl. Abschnitt 2.2). Damit kann eine Orientierung an ähnlichen Maßnahmen

<sup>48</sup> Wissenschaftliche Untersuchungen zu den kausalen Effekten dieser Kombilohnmodelle sind jedoch selten.

nur Anhaltspunkte zur Maßnahmeselektivität bringen. Eine aktuelle Untersuchung stellt die des Hamburger Modells (HM) dar (Jirjahn/Pfeiffer/Tsertsvadze 2006). Sie kann als Beispiel für den Nutzen der Bezugnahme auf andere Maßnahmen dienen. In der Untersuchung werden eine Reihe von Einflussfaktoren auf die Selektivität (vgl. Tabelle 4) des HM aufgeführt, und ihre Operationalisierung wird aufgezeigt.

Tabelle 4: Gründe für die Maßnahmeselektivität im Hamburger Modell und deren Operationalisierung

Grund für Selektivität des Maßnahmezuganges	Operationalisierung
Qualifikation	Schulabschluss
überproportionale Verantwortung von Frauen für die Haushaltsproduktion	Geschlecht
starke Verantwortung für die Kinder	alleinerziehend ja/nein
obsoleter Qualifikation, geringe Leistungsfähigkeit oder Diskriminierung älterer Arbeitssuchender; kurzer Beschäftigungshorizont wegen bevorstehenden Renteneintritts	Alter
Stigmatisierung, Humankapitalentwertung, Entmutigung	Arbeitslosigkeitsdauer; wiederholte Arbeitslosigkeit in Biografie ja/nein
vorherige Teilnahme an Trainingsmaßnahme	vorherige Teilnahme an Trainingsmaßnahme
Produktivitätsnachteile durch gesundheitliche Einschränkungen	gesundheitliche Einschränkungen
Quelle: Jirjahn/Pfeiffer/Tsertsvadze 2006: 17 ff.	

Dass solche Untersuchungen anderer Maßnahmen nur bedingt geeignet sind, um Aufschluss über die Maßnahmeselektivität im MZM zu erhalten, zeigt sich unter anderem darin, dass das HM zum Beispiel keine reine Kombilohnförderung darstellt, sondern diese mit einem Lohnkostenzuschuss an den Arbeitgeber verbindet. Darin unterscheidet es sich vom MZM, bei dem der überwiegende Anteil der Förderfälle nur die Kombientlohnung beinhaltet.<sup>49</sup>

## 5.2 Identifikation der Maßnahmeselektivität im Mainzer Modell

Die Auseinandersetzung mit den bestehenden Strategien zur Suche nach den Selektionseinflüssen zeigt, dass es bisher keinen Königsweg gibt. Im Folgenden wird

<sup>49</sup> Die Übertragbarkeit der Ergebnisse aus dieser speziellen Untersuchung wird überdies noch durch die Wahl der Vergleichsgruppe erschwert. In der Analyse des kausalen Effektes des HM wurde als Vergleichsgruppe die Gruppe der Arbeitslosen gewählt, die keine Beschäftigung aufgenommen haben. Dagegen basiert die Datenerhebung für das Mainzer Modell auf einem Vergleich mit Personen, die zwar eine Beschäftigung aufgenommen haben, die allerdings nicht mit einem finanziellen Zuschuss des Staates gefördert wurden.



zum einen zur Identifikation der Selektionseinflüsse unmittelbar auf die betrachtete Maßnahme Bezug genommen, wobei vor allem die qualitative Untersuchung der Maßnahmeimplementation als Lieferant der nötigen Informationen verwendet wird. Zunächst werden allerdings noch einige offenen Fragen diskutiert.

### 5.2.1 Vorbemerkung

Die Gründe, warum Arbeitsuchende eine Kombilohnförderung erhalten oder nicht, können vielfältiger Art sein, und es wäre eine unlösbare Aufgabe, sie alle zu identifizieren. Die CIA bezieht sich allerdings nur auf Teilaspekte des Selektionsprozesses. In einer Untersuchung des kausalen Effektes einer Maßnahme müssen „nur“ die Aspekte des Selektionsprozesses in der Konstruktion des Propensity Scores einfließen, die gleichzeitig den Zugang in die Maßnahme und die betrachtete Zielvariable der Untersuchung beeinflussen. Ihr Fehlen führt zu einer verfälschten Aussage zur Wirkung der Maßnahme. Alle anderen Aspekte der Selektivität, die nicht mit der Zielvariable in Verbindung stehen, können vernachlässigt werden.<sup>50</sup>

Dies hat zunächst zur Folge, dass das Ausmaß des Selektivitätsproblems immer auch von der gewählten Zielvariable abhängt. Die Zielvariable ist meist eine Operationalisierung des Maßnahmeerfolges und damit von der konkreten Zielstellung der Forschungsarbeit abhängig. Je Forschungsinteresse kommen andere Zielvariablen in Betracht. In der vorliegenden Arbeit wird nicht nur eine, sondern zwei Zielvariablen betrachtet, die Lohnzufriedenheit einerseits und die Beschäftigungsstabilität andererseits.

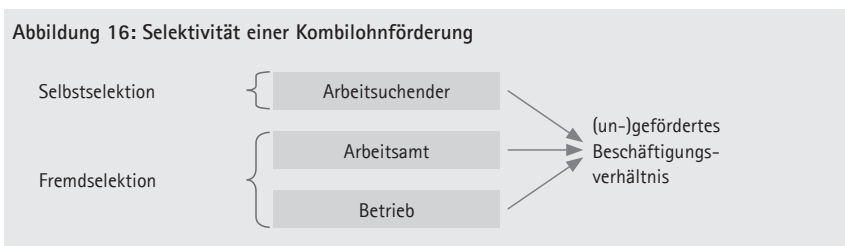
Obwohl grundsätzlich im RKM nur kausalrelevante Selektionseinflüsse zur Identifikation des kausalen Effektes benötigt werden, ist es doch zweckmäßig, zunächst ohne diese Beschränkung möglichst alle Selektionseinflüsse zu erfassen. Dies liegt zum einen generell darin begründet, dass die Beschränkung auf kausalrelevante Selektivitäten immer auch Wissen über die Zielvariable der Untersuchung voraussetzt. Damit würde eine Beschränkung des Informationsbedarfes an einer Stelle mit einer Erweiterung an anderer Stelle erkaufte. Typischerweise stellen Zielvariablen der Maßnahme das im Vergleich zur Maßnahmeteilnahme komplexere Phänomen dar, sodass es strategisch sinnvoller ist, sich mit der Maßnahmeselektivität so umfassend wie möglich auseinanderzusetzen. Zum anderen bietet das Vorgehen im speziellen Fall zweier Zielvariablen den Vorteil, dass damit die conditional independence auch für beide Zielvariablen erfüllbar ist (dies ist eine Grundvoraussetzung für die Prüfung der dritten Hypothese, vgl. Abschnitt 6.4).

<sup>50</sup> Das gilt zumindest im RKM, sofern es durch Propensity Score Matching umgesetzt wird. Für andere statistische Verfahren sind die hier als kausalirrelevant bezeichneten Aspekte der Maßnahmeselektivität von zentraler Bedeutung, nämlich als Instrumente (vgl. Abschnitt 4.3.2).

### 5.2.2 An der Maßnahmeselektivität beteiligte Akteure

Die Selektivität einer Kombilohnmaßnahme kann zunächst danach unterschieden werden, ob sie vom Arbeitsuchenden selbst ausgeht oder von anderen Akteuren, die ebenfalls Einfluss auf die Maßnahmegewährung haben (vgl. Abbildung 16). Geht die Selektivität vom Arbeitsuchenden selbst aus, spricht man von *Selbstsektion*. Man kann sich als Arbeitsuchender mehr oder weniger aktiv um eine Maßnahme bemühen und diese beim Arbeitsamt beantragen. Personen, die dies aktiv tun, sind etwa besonders informiert über Fördermöglichkeiten oder haben gewisse negative oder positive Erfahrungen auf dem Arbeitsmarkt bzw. mit der Arbeitsverwaltung, welche die Bereitschaft zur Aufnahme der Maßnahme beeinflussen. Es ist auch zu erwarten, dass sich diese Personen in irgendeiner Weise einen Vorteil von der Förderung versprechen, worin dieser auch bestehen mag.

Ob ein Arbeitsuchender eine Kombilohnförderung erhält, hängt aber nicht nur von ihm selbst ab, sondern auch von anderen Personen, Institutionen und Kontextfaktoren (*Fremdsektion*). Es liegt im Ermessen der zuständigen Stelle der Agentur für Arbeit bzw. des Sozialamtes, welchen Personen die Förderung aktiv angeboten wird. So kann der Fallmanager (Vermittler) vor Ort daran interessiert sein, in seinen Augen besonders bedürftigen Arbeitsuchenden eine Stelle mit Förderung zu vermitteln, oder sich auf solche Kunden konzentrieren, von denen er annimmt, dass sie sonst nicht bereit wären, eine niedrig entlohnte Stelle aufzunehmen. Während die meisten arbeitsmarktpolitischen Maßnahmen in der Phase der Arbeitslosigkeit durchgeführt werden, wie z. B. Trainingsmaßnahmen oder Umschulungen, ist es gerade Bedingung des Erhaltes einer Kombilohnförderung, dass eine Beschäftigungsaufnahme erfolgt. Damit hängt die Maßnahmeselektivität von Kombilohnförderungen zusätzlich noch davon ab, welche Betriebe welchen Arbeitsuchenden welche Art von förderfähigen (d. h. niedrig entlohnenden) Arbeitsplätzen anbieten.



Jeder dieser drei Akteure, Arbeitsuchender, Arbeitsamt und Betrieb, trägt zur Selektivität von kombilohngeförderten Beschäftigungen bei (vgl. Abbildung 16). Daher muss eine Untersuchung der Zugangsselektivität einer Kombilohnmaßnahme im Gegensatz zu anderen Maßnahmen nicht nur die Handlungslogiken zweier, sondern

dreier unterschiedlicher Klassen von Akteuren betrachten (allerdings wäre zu überlegen, welche Rolle bei anderen Maßnahmen z. B. die Maßnahmeträger spielen).

### 5.2.3 „Kombilohnförderung“ als hypothetische Manipulation im RKM

Eine zentrale und nur auf den ersten Blick triviale Bedingung der Validität von Forschungsergebnissen ist, dass die Ergebnisse der Forschungsbemühungen, egal ob im Rahmen eines Experimentes oder einer Beobachtungsstudie, sich auch tatsächlich auf den Sachverhalt beziehen, den man untersuchen wollte. Dieser Aspekt wird als „Konstruktvalidität“ (Shadish/Cook/Campbell 2002: 64 ff.) bezeichnet und spielt in Rubins Kausalmodell unter dem Aspekt der genauen Festlegung der hypothetischen Manipulation (vgl. Rubin 1986 und die Ausführungen in Abschnitt 4.2 dieser Arbeit) eine wichtige Rolle. In diesem Zusammenhang stellt sich die Frage, was das Prototypische (Shadish/Cook/Campbell 2002: 67) am Konstrukt „Kombilohnförderung“ ist, dessen Auswirkung gemessen werden soll, bzw. in Rubins Terminologie, was genau die hypothetische Manipulation darstellt. Dies ist insofern wichtig, als Mängel in der Konstruktvalidität in der nachfolgenden empirischen Analyse dazu führen, dass ein eventuell gemessener Effekt (oder fehlender Effekt) fälschlich den subjektiven Verletzungen der Reziprozitätsnorm zugerechnet wird, während aufgrund der Operationalisierung eigentlich etwas anderes gemessen wurde.

Im Gegensatz zu Experimenten existiert in Beobachtungsstudien kein „natürlicher“ Kontrollzustand. Dieser wird erst vom Forscher explizit oder (meist) implizit durch die Wahl der Vergleichsgruppe einerseits und die Wahl der Merkmale für die Berechnung des Propensity Score andererseits konstruiert. Gerade bei Kombilohnmaßnahmen ist dies eine besonders komplexe Aufgabe. Betrachtet man die wenigen Analysen der Wirkung von Kombilohnförderungen oder Lohnkostenzuschüssen,<sup>51</sup> dann wird deutlich, dass diese Problematik bisher wenig berücksichtigt wurde. Oft werden Vergleichsgruppe und Kontrollvariablen aus bestimmten pragmatischen, forschungspraktischen Gründen der Datenverfügbarkeit gewählt. Dadurch bekommt das Konstrukt Kombilohn allerdings je nach Untersuchung unterschiedliche Bedeutung. Bevor dargelegt wird, wie in der vorliegenden Analyse Vergleichszustand und Kontrollvariablenwahl auf das zu untersuchende hypothetische Manipulationskonzept abgestimmt ist, soll die Schwierigkeit der Bestimmung der hypothetischen Manipulation anhand der bisherigen Praxis verdeutlicht werden (vgl. Abbildung 17).

51 Im Hinblick auf die Vergleichsgruppenwahl sind Kombilöhne und Lohnkostenzuschüsse mit ähnlichen Problemen konfrontiert.

Im Rahmen des Matchings wird für jeden Kombilohnbezieher ( $kl_i^1$  in Abbildung 17) eine vor dem Maßnahmeeintritt möglichst ähnliche Vergleichsperson ohne Kombilohn gesucht ( $kl_i^0$ ). Im Folgenden wird zur Vereinfachung der Darstellung davon ausgegangen, dass alle Kombilohnbezieher aus der Arbeitslosigkeit kommen. Der Geförderte  $kl_i^1$  hat hier etwa die Eigenschaft, dass er bis zum Förderbeginn arbeitslos war und danach eine Beschäftigung als Hilfsarbeiter aufgenommen hat. Diese wurde mit einem Lohn  $L$  von 1.500 Euro vergütet und mit einem Zuschuss von 200 Euro gefördert. Die Zielvariable sei hier der Beschäftigungsstatus (arbeitslos vs. erwerbstätig), dessen Betrachtung zeitdynamisch erfolgt, indem er  $t$  Monate ( $t = 0, 1, \dots, T$ ) nach Beginn der Förderung als Zielvariable gemessen wird. Damit kann festgestellt werden, ob z. B. 6 Monate nach Aufnahme der geförderten Beschäftigung der Anteil der erwerbstätigen Geförderten höher ist als der der Ungeförderten.

Wer als ungefördert infrage kommt, wird durch das Vergleichskonzept bestimmt. Grundsätzlich gibt es für den Vergleichszustand zwei Möglichkeiten. Entweder betrachtet man Arbeitslose (Vergleichskonzept A) oder Erwerbstätige in einer Beschäftigung ohne Kombilohnförderung<sup>52</sup> (Vergleichskonzept B). Wie im Folgenden dargelegt wird, führen die verschiedenen Vergleichskonzepte allerdings automatisch zu unterschiedlichen Kombilohnkonstrukten, weil sie das beeinflussen, was das RKM als hypothetische Manipulation misst. Dies beeinflusst wiederum, ob und welcher Effekt des Kombilohnes gemessen wird.

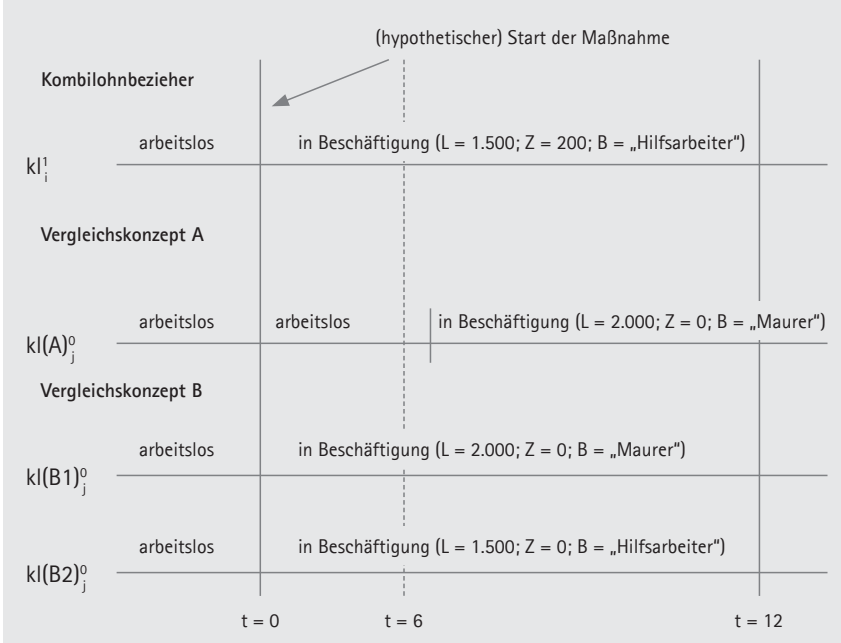
Vergleichskonzept A besteht darin, dass als Vergleichszustand eines Kombilohngeförderten der Verbleib in der Arbeitslosigkeit gewählt wird (z. B. Jirjahn/Pfeiffer/Tsertsvadze 2006; ähnlich auch bei Lohnkostenzuschüssen Jaenichen 2002a; Heinemann/Jaenichen/Stephan 2006). Dabei besteht zunächst das Problem, dass bei Arbeitslosen kein Zeitpunkt des Maßnahmestarts existiert. Daher muss für sie ein hypothetischer Startzeitpunkt  $t = 0$  bestimmt werden, ab dem dann der Verbleib in oder Austritt aus der Arbeitslosigkeit  $t$  Monate nach Förderbeginn bzw. hypothetischen Förderbeginn gemessen wird.

In Abbildung 17 wird dem Geförderten (z. B. im Rahmen des PS-Matchings) im Konzept A beispielhaft eine Vergleichsperson  $kl(A)^0$  zugeordnet, die folgende Eigenschaften hat: Sie hat vor dem (hypothetischen) Maßnahmebeginn im Idealfall sehr ähnliche<sup>53</sup> Eigenschaften, mit der Ausnahme, dass keine Arbeitsaufnahme zum Zeitpunkt  $t = 0$  erfolgt.

52 Bei Arbeitslosen lässt sich dann noch unterscheiden, ob man die Zeit in geförderter Beschäftigung bereits als Erfolg bewertet oder der Erfolg erst dann eintritt, wenn die Förderung ausgelaufen ist. Bei Beschäftigten als Kontrollgruppe kann nach Personen in vollkommen ungeförderten und Personen in „kombilohnungeförderten“ Beschäftigungen, d. h. eventuell mit anderer Förderung versehenen Beschäftigungen, unterschieden werden.

53 Eigentlich ist dies nur beim exakten Matching der Fall. Gewöhnlich sind nach dem Matching die Verteilungen der Merkmale bei Geförderten und Ungeförderten ähnlich, nicht unbedingt die Matchingpartner selbst, die ähneln sich oft nur hinsichtlich des Propensity Scores.

Abbildung 17: Vergleichskonzepte bei der Evaluation von Kombilohnmaßnahmen



Dieses Vorgehen ist insofern sinnvoll, als das Ziel des Kombilohnes unter anderem die Aktivierung von Arbeitslosen zur Aufnahme einer Beschäftigung ist (Anreizmechanismus). Das Problem ist nun allerdings, dass mit der Wahl von Arbeitslosen als Vergleichspersonen der Erfolg dieses Ziels bereits unterstellt wird. Die Konstruktion des Vergleichszustandes impliziert, dass eine ungeförderte Person zum Zeitpunkt  $t = 0$  die Möglichkeit der Beschäftigungsaufnahme hatte und sich aus finanziellen Gründen dagegen entschieden hat.

Damit ist in die eigentlich zu messende Wirkung des Kombilohnes zum Teil bereits im Untersuchungsdesign eingebaut. Dies wird besonders deutlich, wenn man beachtet, dass zum Zeitpunkt  $t = 1$  fast alle Kombilohnbezieher beschäftigt sind, während fast alle Vergleichspersonen im Konzept A arbeitslos sind. Dieser „Startvorteil“ (Jaenichen 2002a: 345) verschwindet erst im Laufe der weiteren Beschäftigungsbiografie, wenn überhaupt, und es ist nicht auszuschließen, dass so unabhängig vom tatsächlichen Maßnahmeerfolg zumindest kurz- und mittelfristig automatisch zu hohe „kausale“ Effekte gemessen werden. Darüber hinaus besteht, wie oben dargestellt, ein wesentlicher Aspekt der Maßnahmeselektivität in der Fremdselektion durch die Betriebe. Dieser Aspekt kann mit einer solchen Vergleichsgruppenwahl nicht berücksichtigt werden (vgl. das Variablenspektrum der Untersuchung in Tabelle 4), da für Arbeitslose natürlich keine Betriebsinformationen vorliegen. Da sich die Betriebseigenschaften (wohl) nicht zufällig auf die

kombilohngeförderten Aussteiger aus der Arbeitslosigkeit und die (noch) arbeitslosen Ungeförderten verteilen (vgl. Abschnitt 5.2.4.3 unten), wird als Manipulation „Kombilohn“ hauptsächlich die Sortierung in bestimmte Beschäftigungsverhältnisse gemessen, z. B. in Abbildung 17 in eine Beschäftigung nicht als Hilfsarbeiter, sondern als Maurer (etwa weil mehr Suchzeit nach einer solchen Tätigkeit genutzt wurde) mit einem Lohn  $L$  von 2.000 Euro und damit 300 Euro mehr als der Geförderte. Dies ist insofern plausibel, als die Hoffnung besteht, dass durch Kombilöhne Beschäftigungen entstehen/aufgenommen werden, die ohne Förderung nicht entstanden/aufgenommen worden wären. Allerdings ist unklar, wie sich dies auf die Zielgröße Beschäftigungsstatus auswirkt bzw. welcher Teil eines eventuellen Effektes auf den Kombilohn und welcher auf die Sortierung in entsprechende Betriebe zurückgeht. Da qualifizierte Beschäftigungen in der Regel stabiler sind (Erlinghagen/Knuth 2004: 29), könnte der Startvorteil langfristig in einen Nachteil umschlagen, obwohl der Kombilohn erfolgreich zur Aufnahme einer Beschäftigung im Niedriglohnssektor motiviert hatte.

Wäre es dann vielleicht besser, eine ungeförderte Beschäftigung als Vergleichszustand zu verwenden (z. B. Hartmann/Krug 2004; für Lohnkostenzuschüsse siehe Cockx/Van der Linden/Karaa 1998, die darin ein Problem des endogenen Samplings sehen und dies mit einer Selektionskorrektur lösen), wie es im Vergleichskonzept B erfolgt? Nicht unbedingt, denn auch hier werden bestimmte Sachverhalte unterstellt und fließen in den kausalen Effekt ein. So impliziert die Betrachtung von beschäftigten Vergleichspersonen, dass die Kombilohnbezieher auch ohne den Zuschuss eine Beschäftigung aufgenommen hätten, eben eine ungeförderte.<sup>54</sup> Damit wird nun umgekehrt zum Konzept A vorausgesetzt, dass der Anreizmechanismus keinen Erfolg hatte. Je nach Wahl der Beobachtungseinheit (bzw. dadurch bedingt: Kontrollvariablen) kann hier wiederum zwischen zwei Vergleichskonzepten der „Kombilohnförderung“ unterschieden werden. In Konzept B1 betrachtet man das Individuum als Untersuchungseinheit und verwendet zur Berechnung des Propensity Score nur dessen Eigenschaften (inklusive Eigenschaften der für sie zuständigen Arbeitsämter). In einem solchen Fall ist ein Kombilohn dann automatisch definiert als ein Mechanismus zur Sortierung in andere, eventuell niedriger entlohnte Beschäftigung – nun aber unter der Annahme, dass die Geförderten ohne Förderung nicht arbeitslos geblieben wären, sondern eine andere Beschäftigung aufgenommen hätten. Hier ist es durchaus möglich, dass trotz eines Erfolges des Kombilohnes in der Dimension „Anreizwirkung“ die Messung eines (scheinbareren) Misserfolges in der Dimension „dauerhafte Reintegration“ erfolgt.

---

54 Das ist für das Mainzer Modell zumindest bei einer Vielzahl der Fälle plausibel (vgl. Hollederer/Rudolph 2001), doch wohl auch nicht uneingeschränkt der Fall.

In Konzept B2 nimmt man hingegen das Beschäftigungsverhältnis als Beobachtungseinheit an, und es werden daher auch die Eigenschaften der einstellenden Betriebe kontrolliert. In diesem Fall besteht die Manipulation „Kombilohn“ ausschließlich in der Einkommenserhöhung. Geförderte ( $kl_i^1$ ) und Ungeförderte ( $kl(B)_i^0$ ) unterscheiden sich lediglich dadurch, dass Erstere den Zuschuss erhalten und Letztere nicht. Daher wird hier als Manipulation „Kombilohn“ der zusätzliche Lohn durch Kombilohnförderung betrachtet.

Keine der Vorgehensweisen kann als eindeutig korrekt oder eindeutig falsch bezeichnet werden. Bei allen Vergleichskonzepten ist es jedoch so, dass durch die jeweilige Wahl des Vergleichszustandes eine spezifische Manipulation „Kombilohn“ konstruiert wird, ohne dass dies offensichtlich wäre. Daher ist bei der Wahl des Vergleichszustandes bzw. der Wahl der Beobachtungseinheit bei Kombilohnfördermaßnahmen besondere Vorsicht geboten.

In der vorliegenden Arbeit wird das Vergleichskonzept B2 gewählt, da zum einen die interessierende Einheit das Beschäftigungsverhältnis darstellt und die Manipulation in dem zusätzlichen, aus einem anderen als der Beschäftigung selbst bezogenen Einkommen besteht. Der in einer Kombilohnmaßnahme bezogene Lohn wird im Gegensatz zum Lohn aus regulärer Beschäftigung um den Zuschuss erhöht. Andere Unterschiede, auch andere Wirkungen des Kombilohnes, etwa die Sortierwirkung in bestimmte Beschäftigungsverhältnisse oder die Einkommenserhöhung, sollen dabei konstant gehalten werden. Um die interessierende Manipulation „herzustellen“, müssen alle Selektivitäten ausgeglichen werden, die aufgrund von Selbst- und Fremdselektion auch durch die Betriebe entstanden sind, sowie zusätzlich noch die Höhe des Arbeitgeberanteils am Lohn, sodass der einzige Unterschied von Geförderten und Ungeförderten im Bezug des Zuschusses liegt. Dies ist mit den verwendeten Daten (vgl. unten) insofern möglich, als hier der *Bruttolohn* abgefragt wurde, während die Förderung in der Verringerung der Sozialversicherungsbeiträge für den Arbeitnehmer und gegebenenfalls einem Aufschlag zum Kindergeld besteht, also in einer Erhöhung des *Nettolohnes* resultiert (vgl. Abbildung 8 in Abschnitt 2.3).

#### 5.2.4 Selektivität der Kombilohnmaßnahme „Mainzer Modell“

Die Selektivität einer Maßnahme lässt sich nicht angemessen als Summe der einzelnen durch die verschiedenen beteiligten individuellen und kollektiven Akteure erfassen. Durch die gleichzeitige Beteiligung dreier Akteure entstehen komplexe Wechselwirkungen, die zwar auf den Intentionen der Einzelnen beruhen, jedoch nicht auf diese reduzierbar sind. So können bestimmte Arbeitsuchende bestrebt sein, bestimmte Ziele zu erreichen, zu denen auch die Maximierung des Einkommens zählt, sie können dies aber nur unter Mitwirkung der Fallmanager und Be-

triebe. Diese verfolgen wiederum eigene Ziele und sind in komplexe organisations-spezifische Handlungszusammenhänge eingebunden. Hinzu kommt, dass gerade bei Modellprojekten wie dem Mainzer Modell oder Reformen, die eine neue Maßnahme einführen, über diese Wechselwirkungen noch keine Informationen vorliegen. Daher ist das Vorliegen nicht standardisierter Analysen besonders wichtig, um die Maßnahmeselektivität möglichst vollständig zu erfassen.

Im Folgenden soll die Maßnahmeselektivität der Kombilohnmaßnahme „Mainzer Modell“ analysiert werden. Diese Analyse basiert zum einen auf den Ergebnissen der vom Institut Arbeit und Technik (IAT) durchgeführten qualitativen Studie zur Implementation des Mainzer Modells (Kalina/Vanselow/Weinkopf 2004: 29 ff.) und wird zum anderen durch gegenstandsrelevante Theorien aus der Arbeitsmarktforschung ergänzt.

### 5.2.4.1 Selbstselektion der Teilnehmer

Ein wichtiger, allerdings nur selten beachteter Aspekt der Maßnahmeselektivität ist, dass der intensive Kontakt zur Arbeitsverwaltung bzw. dessen Fehlen die Wahrscheinlichkeit der Förderung maßgeblich beeinflussen kann. Es sind die Arbeitsämter – im Mainzer Modell unter Mithilfe der Sozialämter –, die „Förderanträge ausgeben, ausgefüllte Anträge entgegennehmen, prüfen und ggf. nach einer Bewilligung zahlbar machen“ (Kalina/Vanselow/Weinkopf 2004: 42).<sup>55</sup> Es ist daher zu erwarten, dass Personen, die bei ihrer Arbeitsplatzsuche verstärkt andere Wege nutzen, unter den Kombilohngeförderten unterrepräsentiert sind. Solche anderen Suchwege sind etwa informelle Suchwege (vgl. Noll/Weick 2002) wie soziale Netzwerke (Granovetter 1973), Initiativbewerbung, eigene Stellenanzeige etc.

Wenn eine Förderung einem Arbeitsuchenden nicht aktiv angeboten wird, dann kann er diese trotzdem erhalten, wenn er sich selbst darum bemüht. Dies setzt allerdings voraus, dass er von der Möglichkeit zur Beantragung eines Zuschusses weiß. Die Implementationsforschung zum Mainzer Modell verweist darauf, dass das wenige Informationsmaterial, wenn überhaupt, dann nur in deutscher Sprache zur Verfügung stand (Kalina/Vanselow/Weinkopf 2004: 49). Daher ist zu erwarten, dass sich unter den Geförderten im Vergleich zu Ungeförderten weniger Personen mit sogenanntem Migrationshintergrund befinden. Da zum Teil Weiterbildungsträger genutzt wurden (ebd. 53), um über die Möglichkeit einer Kombilohnförderung im Mainzer Modell zu informieren, könnte auch eine Selektivität der Teilnahme dahingehend bestehen, dass verstärkt Abgänger aus der Weiterbildung gefördert wurden.

---

55 Grundsätzlich war es auch möglich, den Förderantrag bis zu sechs Wochen nach Aufnahme einer Beschäftigung zu stellen, doch konzentrierten sich die Mitarbeiter im Arbeitsamt ihrem Beratungsverständnis und der beschäftigungsschaffenden Intention des MZM folgend darauf, Arbeitsuchende vor der Arbeitsaufnahme über die Fördermöglichkeit zu informieren.



Bei Personen, die von einer Fördermöglichkeit wissen, kann man im Sinne des unterstellten Anreizmechanismus zunächst von einer Positivselektion auszugehen. Es werden sich verstärkt solche Arbeitsuchenden für eine geförderte Beschäftigung entscheiden, die sich einen höheren Lohn als ohne Förderung erwarten. Dies wird gerade bei solchen Arbeitsuchenden der Fall sein, die aufgrund geringer Qualifikation und bisher negativer Erfahrungen auf dem Arbeitsmarkt nicht mit der Aufnahme höher entlohnter Beschäftigungsverhältnisse rechnen können. Dabei spielt auch der Haushaltskontext eine Rolle. Die Förderung wird dabei umso attraktiver sein, je geringer das Haushaltseinkommen insgesamt ausfällt.

#### 5.2.4.2 *Fremdselektion durch die Arbeitsämter*

Die Handlungslogik der Arbeitsämter bei der Selektion von Arbeitsuchenden in geförderte oder ungeförderte Beschäftigung weist zwei wesentliche Elemente auf. Erstens müssen sich die Vermittler an die für die speziellen Maßnahmen geltenden Förderrichtlinien halten. Zweitens besitzen sie gleichzeitig einen breiten Ermessensspielraum, den sie im Rahmen ihrer Berufsauffassung und in interpretativer Auseinandersetzung mit dem Förderinstrument und den Rahmenbedingungen ausfüllen.

Demnach wird die Selektivität des Maßnahmezuganges im Mainzer Modell also maßgeblich von den Förderrichtlinien beeinflusst. Diese schließen zwar keine Personengruppe von der Förderung mit einem Kombilohnzuschuss aus, nennen jedoch auch klare Zielgruppen der Maßnahmen. Zu diesen gehören vor allem Langzeitarbeitslose und Geringqualifizierte, weshalb sich auch die im Rahmen des CAST-Projektes durchgeführte Befragung auf diese Personengruppe beschränkt. Der Erhalt einer Förderung war weniger von den individuellen Eigenschaften, sondern vielmehr von Eigenschaften des Haushaltskontextes und der aufgenommenen Beschäftigung abhängig. Nur wenn das Einkommen aus der Beschäftigung eine bestimmte Höchstgrenze nicht überschritt, konnte eine Förderung gewährt werden. Die Fördergrenzen waren dabei vom jeweiligen Haushaltskontext abhängig, vor allem von der Zahl der Kinder unter 18 Jahren im Haushalt und davon, ob im Haushalt ein erwerbstätiger Lebenspartner vorhanden war bzw. wie hoch dessen Erwerbseinkommen war. Schließlich waren diese Förderbedingungen so gestaltet, dass sie gerade zur Aufnahme von Teilzeitarbeit motivierten. Zentraler Bestandteil des Mainzer Modells war die Unterstützung der Kooperation zwischen Arbeitsämtern und Trägern der Sozialhilfe. Obwohl diese im Einzelnen nicht immer das gewünschte Ausmaß annahm, ist trotzdem zu erwarten, dass unter den kombilohngeförderten Beschäftigten verstärkt ehemalige Sozialhilfeempfänger zu finden sind. Und dies umso mehr, als in den Förderrichtlinien die Überwindung der „Sozialhilfefalle“ als eines der Hauptziele des MZM festgeschrieben war.

Die qualitativen Analysen des Implementationsprozesses verweisen darauf, dass der Anreizmechanismus des Kombilohnzuschusses nicht eine so herausragende Rolle gespielt hat, wie man es sich bei der Konzeption des Mainzer Modells erwartet hat (Kalina/Vanselow/Weinkopf 2004: 45 f.). Zunächst war man davon ausgegangen, dass sowohl die Arbeitgeber aus Eigeninteresse verstärkt förderfähige offene Stellen im Niedriglohnbereich melden als auch Arbeitsuchende aus dem Eigeninteresse des durch den Zuschuss gewährten Vorteils die Förderung verstärkt nachfragen. Allerdings stellte sich im Laufe der Maßnahme Folgendes heraus: „Anstelle einer weitgehend spontanen Inanspruchnahme durch Arbeitskräfte musste das Programm durch aktive Beratung und vermittlerische Aktivitäten begleitet werden, was wiederum besondere Anforderungen an das Personal stellte. Ähnliches gilt auch für die Meldung offener, niedrig entlohnter Stellen durch die Arbeitgeber“ (ebd. 46). Daher tritt im Mainzer Modell vor allem der Aspekt der Fremdselektion durch die Arbeitsämter in den Vordergrund, wobei diese stark von den konkreten Rahmenbedingungen der Maßnahmeumsetzung beeinflusst wurde.<sup>56</sup>

Eines der wesentlichen Merkmale der Umsetzung des Mainzer Modells war die komplexe Ausgestaltung der Förderbedingungen (Kalina/Vanselow/Weinkopf 2004: 47, 62; Cords 2003), vor allem was die Bestimmung der Förderfähigkeit eines Beschäftigungsverhältnisses und der Höhe der Zuschüsse betraf. Diese hing von einer Reihe persönlicher Merkmale und Merkmale der finanziellen und sonstigen Lebenssituation ab und war erst dann exakt zu bestimmen, wenn das Einkommen aus der – mitunter nur angestrebten – Beschäftigung bekannt war. Dies bedeutete für das Personal der Arbeitsämter neben dem Arbeitsaufwand bei der Suche nach infrage kommenden Arbeitsuchenden und der Ansprache von Arbeitgebern einen hohen Aufwand bei der Identifikation der Förderfähigkeit und der Zuschusshöhe. Oft stellte sich erst relativ spät heraus, ob eine konkrete Beschäftigung förderfähig war oder nicht. Trotz des hohen Mehraufwandes wurden den Arbeitsämtern jedoch keine zusätzlichen Ressourcen zur Verfügung gestellt. Unter Umständen führt die Erwartung eines breiten Einsatzes des Förderinstrumentes, ohne für ausreichende Ressourcen zu sorgen, dazu, dass die Vermittler sich nur eingeschränkt um das Matching von Bewerbern und Stellen kümmern können. Dieser Aspekt ist umso wichtiger, als darin eine alternative Erklärung für die hohe Abbruchquote bestehen könnte.

Eine wichtige Rolle beim Zugang zum MZM spielte auch die damalige Trennung von Arbeitslosenunterstützung und Sozialhilfe. Während registrierte Arbeitslose den Förderantrag direkt beim Arbeitsamt stellen konnten, mussten Sozialhilfeempfänger

---

56 Dass eine solche „Überzeugungsarbeit“ kein Spezifikum des Mainzer Modells ist, zeigen die Analysen von Behrend et al. (2006: 3).

dazu mehrere Stellen anlaufen (Kalina/Vanselow/Weinkopf 2004: 49).<sup>57</sup> Geförderte Personen können sich daher zum einen von Ungeförderten hinsichtlich des Leistungsbezuges (Arbeitslosengeld bzw. -hilfe vom Arbeitsamt bzw. Sozialhilfe vom Sozialamt) unterscheiden und zum anderen, damit zusammenhängend, hinsichtlich bisheriger Berufserfahrungen (Bezugsvoraussetzung für Arbeitslosenunterstützung ist das Vorliegen bestimmter Beschäftigungszeiten) und ihres Qualifikationsniveaus.

Generell wird der Arbeitsmarkthistorie und der Qualifikation potenzieller Maßnahmeteilnehmer bei der Selektivität von Maßnahmezugängen eine hohe Bedeutung beigemessen. Die Vermittler werden sich bei der Suche nach besonders geeigneten Kandidaten an deren bisherigen Erfolgen oder Misserfolgen bei der Arbeitsuche orientieren. Dies gilt für das Mainzer Modell in besonderem Maße, da für dessen Umsetzung keine zusätzlichen Ressourcen zur Verfügung standen. „Hinsichtlich des Einsatzes personeller Kapazitäten stand das Mainzer Modell dadurch in Konkurrenz zu anderen Aufgaben, Instrumenten und Förderprogrammen. In einer solchen Situation werden Prioritäten u. a. auch durch die individuelle Einschätzung der Leitung und der einzelnen Mitarbeiter/innen im Hinblick darauf bestimmt, wo Aufwand und Ertrag in einem besonders günstigen Verhältnis stehen und wo die erzielbaren Wirkungen (ggf. mutmaßlich) am größten sind“ (Kalina/Vanselow/Weinkopf 2004: 61). Zudem besteht die Hoffnung, dass sich schwer oder kaum messbare motivationale Eigenschaften der Individuen über die Arbeitsmarkthistorie zumindest indirekt erfassen lassen (Heckman/Ichimura/Smith/Todd 1998; Winterhager/Heinze/Spermann 2006: 510). So sollen bei der Auswahl der Kontrollvariablen Faktoren wie Berufserfahrung, Erfahrung mit Arbeitslosigkeit oder frühere Maßnahmeteilnahme berücksichtigt werden sowie ob die Arbeitslosigkeit vor der Maßnahmeteilnahme auf eine arbeitnehmer- oder arbeitgeberseitige Kündigung zurückgeht.

Die Arbeitsämter müssen ihre Vermittlungstätigkeit auf die spezifischen Bedingungen der regionalen Arbeitsmarktsituationen abstimmen. Je nach Höhe der Arbeitslosigkeit oder Zahl der offenen Stellen wird eine Kombilohnmaßnahme für die Vermittler ein mehr oder weniger geeignetes Mittel zur Förderung von Beschäftigungsaufnahmen darstellen und daher mehr oder weniger offensiv angeboten. Die Disparität bezüglich der regionalen Arbeitsmarktsituation kann durch eine Typisierung der Arbeitsamtsbezirke (Blien et al. 2004) abgebildet werden. Diese unterscheidet mithilfe einer Clusteranalyse auf Basis verschiedener Indikatoren zur Beschaffenheit der Rahmenbedingungen<sup>58</sup> auf dem Arbeitsmarkt fünf Typen: Be-

57 Nach der bundesweiten Ausdehnung fanden in Sozialämtern im Vergleich zu den Arbeitsämtern dann weit weniger Ansprachen von Arbeitslosen statt, obwohl sich das MZM ursprünglich vor allem auch an die Personen in der sogenannten Sozialhilfefälle richtete.

58 Diese sind die Unterbeschäftigungsquote, die Bevölkerungsdichte, die Quote der offenen Stellen, die Einstellungsquote, die Quote der Sozialhilfeempfänger, der Tertiarisierungsgrad und die Saisonspanne (Blien et al. 2004: 174).

zirke in Ostdeutschland mit dominierendem Arbeitsplatzdefizit (Typ 1); großstädtisch geprägte Bezirke vorwiegend in Westdeutschland mit hoher Arbeitslosigkeit (Typ 2); mittelständische und ländliche Gebiete in Westdeutschland mit durchschnittlicher Arbeitslosigkeit (Typ 3); Zentren in Westdeutschland mit günstiger Arbeitsmarktlage und hoher Dynamik (Typ 4); Bezirke in Westdeutschland mit guter Arbeitsmarktlage und hoher Dynamik (Typ 5).

#### 5.2.4.3 *Fremdselektion durch die Betriebe*

Grundsätzlich wurde bei der Entwicklung des Förderinstrumentes „Mainzer Modell“ davon ausgegangen, dass eine finanzielle Motivation der Arbeitsuchenden zur Aufnahme niedrig entlohnter Beschäftigung über den Umweg sinkender Lohnkosten zur Entstehung neuer Stellen führen werde, da Betriebe nun Stellen im Niedriglohnbereich besetzen könnten, für die bisher keine Bewerber zu finden waren. Allerdings zeigte sich, dass „die meisten Arbeitgeber keinen eigenen Vorteil in dem Programm erkennen konnten. Das etablierte Instrument der Lohnkostenzuschüsse an den Arbeitgeber wurde von diesen als viel attraktiver eingeschätzt“ (Kalina/Vanselow/Weinkopf 2004: 51). Daher fällt im Fall des Mainzer Modells die gezielte Selektivität der Förderung durch die Betriebe gering aus. Dies zeigt sich auch darin, dass viele der einstellenden Betriebe nicht wussten, dass ein Beschäftigter im Rahmen des Mainzer Modells gefördert wurde (Gewiese 2004; vgl. auch Cords 2003: 114). Wie die Befragung ergab, war von den Betrieben mit mindestens einem MZM-Geförderten (je nach Befragungszeitpunkt) überhaupt nur 62 bzw. 55 % das Mainzer Modell bekannt, und nur 14 bzw. 12 % der Betriebe wussten, dass sie einen MZM-Geförderten beschäftigen (Gewiese 2004: 275 f.). Daraus folgt nicht, dass Betriebe bei der Selektivität des Maßnahmezuganges keinerlei Rolle spielten (wenn sie auch vergleichsweise gering sein dürfte). So stellten bestimmte Branchen verstärkt geförderte Beschäftigte ein (Krug/Rudolph/Holleder 2004; Kalina/Vanselow/Weinkopf 2004: 56), vor allem solche, die zu den typischen Niedriglohnbereichen gehören. Insbesondere waren Geförderte besonders häufig im Reinigungsgewerbe, im Hotel- und Gaststättengewerbe, im Handel und in der Leiharbeitsbranche zu finden.

Aus den Ergebnissen der Implementationsforschung (Kalina/Vanselow/Weinkopf 2004: 52) lässt sich überdies die Vermutung ableiten, dass die Förderung besonders selten an Beschäftigte in Kleinstbetrieben vermittelt wurde, sodass die geförderten Beschäftigungsverhältnisse in diesem Bereich selektiv sein können, d. h., dass unter den Geförderten besonders häufig Beschäftigte in mittleren und großen Betrieben sind und weniger Beschäftigte in kleinen Betrieben.

## 6 Der kausale Effekt des Kombilohnes auf Lohnzufriedenheit und Beschäftigungsstabilität

In diesem Kapitel werden die in Abschnitt 3.4 entwickelten Hypothesen einer empirischen Überprüfung unterzogen. Dies wird auf Basis von Befragungsdaten geschehen, die eingangs des Kapitels vorgestellt werden. Nachdem auf die Operationalisierung der Selektionseinflüsse aus Abschnitt 5.2 und der betrachteten Zielvariablen eingegangen wird, erfolgt das Propensity Score Matching, und die Hypothesen werden getestet. Schließlich wird die Validität der Hypothesentests noch überprüft, indem einige Plausibilitätstests der zentralen Verfahrensannahmen im RKM durchgeführt werden. Zusätzlich wird noch eine zu den vorgestellten Hypothesen alternative Erklärung des Abbruchverhaltens getestet und verworfen. Schließlich erfolgt eine Diskussion, inwieweit sich die aus dem Hypothesentest gewonnenen Erkenntnisse verallgemeinern lassen.

### 6.1 Verwendete Datenbasis

Die vorliegende Untersuchung greift auf eine bereits bestehende Datenbasis zurück. Diese Situation ist typisch für die Mehrzahl der Arbeiten zur Messung kausaler Effekte arbeitsmarktpolitischer Maßnahmen. Untypisch ist allerdings, dass es sich bei den Daten um eine eigens durchgeführte Befragung handelt, während Maßnahmen der AAMP meist auf Basis der Prozessdaten der Bundesagentur für Arbeit untersucht werden. Der Grund ist wohl, dass die Prozessdaten der BA Massendaten liefern und so erstens auch bei sehr speziellen Teilpopulationen wie den Maßnahmeteilnehmern eine hohe Fallzahl zu realisieren ist und zweitens eine Vielzahl potenzieller Vergleichspersonen zur Verfügung stehen, wodurch die Suche nach möglichst ähnlichen Matchingpartnern erleichtert wird. Allerdings besteht dabei immer ein Trade-off zwischen der Erzielung ausbalancierter Teilstichproben von Geförderten und Ungeförderten und der CIA. Die Prozessdaten der BA enthalten nur sehr spezifische Informationen. Erhoben wird, was administrativ relevant ist, sodass wissenschaftlich interessante Informationen oft fehlen. Vor allem fehlen für Beschäftigte Informationen zum Haushaltskontext einerseits und subjektive Einschätzungen der Lohn- und Arbeitssituation und Einstellungen andererseits (vgl. zu den Unterschieden von Prozess- und Befragungsdaten allgemein Hartmann/Krug 2009). Solche Merkmale können in Prozessdaten nur indirekt erfasst werden, etwa über die Arbeitsmarkt- und Maßnahmeteilnahmebiografie (Stephan/Rässler/Schewe 2006: 459; Winterhager/Heinze/Spermann 2006: 509 ff.). Dadurch können Messfehler entstehen, und es bleibt zudem die Frage, ob so tatsächlich die interessierenden Merkmale der Individuen erfasst wurden.

Daher und weil in der vorliegenden Untersuchung nicht nur die faktische Lohnhöhe, sondern gerade deren subjektive Einschätzung relevant ist, wird hier auf Befragungsdaten zurückgegriffen<sup>59</sup> und das Problem kleiner Fallzahlen zugunsten der höheren Plausibilität der CIA in Kauf genommen. Verwendet werden Daten, die zur Evaluation des Mainzer Modells im Rahmen des vom Bundesministerium für Wirtschaft und Arbeit finanzierten und 2004 abgeschlossenen Projektes „Begleitforschung zum Sonderprogramm zur Erprobung von Modellansätzen zur Förderung der Beschäftigung von Geringqualifizierten und Langzeitarbeitslosen (Mainzer Modell, Modell der Saar-Gemeinschaftsinitiative)“ erhoben wurden. Das Projekt wurde von einem Forschungsverbund „Mainzer Modell“ unter der Koordination von Bruno Kaltenborn durchgeführt. Bei der Befragung handelt es sich um eine in fünf Wellen unterteilte, repräsentative Befragung, die von TNS Infratest Sozialforschung im Auftrag des Instituts für Arbeitsmarkt- und Berufsforschung (IAB) durchgeführt wurde. Befragt wurden sowohl Teilnehmer des Mainzer Modells als auch Vergleichspersonen, die im gleichen Zeitraum eine Beschäftigung aufgenommen haben. Die Stichprobenziehung erfolgte zeitlich und regional proportional zu den Zugängen in die Förderung. In den ursprünglichen Förderregionen wurden Zugänge von Januar 2001 bis August 2002 und nach der bundesweiten Ausdehnung Zugänge von September 2002 bis März 2003 in die Stichprobe einbezogen. Damit wurde in beiden Fällen die „Anlaufphase“ wegen möglicher Sondereffekte ausgespart. Durchgeführt wurde die Befragung als CATI (Computer Assisted Telephone Interview). Da das Mainzer Modell keine eindeutig definierte Zielgruppe besaß, allerdings der Intention nach auf Arbeitslose und/oder Geringqualifizierte ausgerichtet war, wurde die Vergleichsgruppe auf Personen beschränkt, die aus der Arbeitslosigkeit abgingen und dabei entweder langzeitarbeitslos (mindestens ein Jahr ohne Unterbrechung arbeitslos gemeldet) oder ohne abgeschlossene Berufsausbildung waren. Teilnehmer am Mainzer Modell, die diese Kriterien nicht erfüllten, wurden allerdings trotzdem befragt.

Um möglichst valide Angaben auf Retrospektivfragen zur Situation vor der geförderten oder ungeförderten Beschäftigung zu erhalten, wurde die Befragung in zeitlicher Nähe zum Beschäftigungs- bzw. Förderbeginn durchgeführt. Ursprünglich war noch eine Zweitbefragung mit längerem Abstand geplant, die jedoch wegen der Einstellung des Mainzer Modells zur Kostenreduktion nur noch mit den Teilnehmern durchgeführt wurde. Die hier verwendete Erhebung ist die Erstbefragung, die nicht unmittelbar für eine Analyse der Beschäftigungsstabilität, die möglichst lange Erhebungszeiträume benötigt, konzipiert worden war. Dies zeigt sich auch darin,

---

59 Neben der Befragung werden für einzelne Berechnungen (siehe Anhang) auch vom IAB aufbereitete Prozessdaten der BA herangezogen (Hummel et al. 2005).

dass in jeder Welle zuerst die Teilnehmer befragt wurden und abhängig von der Zahl der realisierten Interviews dann ein bestimmtes Quantum an Vergleichspersonen. Dieser Umstand führt u. a. dazu, dass die Chance für ein Ende der (un-)geförderten Beschäftigung im Beobachtungszeitraum, sei es durch den Arbeitgeber oder den Arbeitnehmer, für die Vergleichspersonen durch den längeren Beobachtungszeitraum systematisch erhöht ist. Aufgrund der systematischen Zensurierung kann ein reiner Mittelwertvergleich der Beschäftigungsdauern auf Basis dieser Daten und im Rahmen des Propensity Score Matching zu verzerrten Ergebnissen führen. Eine Lösung bietet die Kombination beider Verfahren als spezielle Variante des regression adjusted matching (Rubin 1973).<sup>60</sup>

Realisiert wurden schließlich 3.080 Interviews mit geförderten und 1.443 Interviews mit Personen der Vergleichsgruppe. Es handelt sich damit um eine disproportionale Stichprobe, bei der die Kombilohngeförderten übergewichtet wurden. Unter den befragten Geförderten waren 1.176 Personen, die den Kriterien „Arbeitsaufnahme aus der Arbeitslosigkeit bei gleichzeitiger Langzeitarbeitslosigkeit“ und/oder „fehlende Berufsausbildung“ entsprachen. Da in die Vergleichsgruppe auch Personen gelangten, die keine Erwerbstätigkeit aufgenommen hatten, etwa weil sie in schulische oder berufliche Aus- oder Weiterbildung übergegangen waren, blieben schließlich 671 Vergleichspersonen.<sup>61</sup> In der folgenden Analyse werden zusätzlich noch die Alleinerziehenden ausgeschlossen. Der Grund hierfür ist, dass deren Häufung in der Stichprobe auf eine Besonderheit im Mainzer Modell zurückgeht, nämlich der Orientierung des Zuschusses nicht am Stundenlohn, sondern am Monatslohn, wodurch in der Kombination mit Teilzeitarbeit Beschäftigungen förderfähig wurden, die – unter Betrachtung des Stundenlohnes – nicht zum Niedriglohnsegment zählen. Diese Sondersituation kann erheblichen Einfluss auf ihre Reaktion im Sinne der Lohnzufriedenheit auf den Kombilohn bewirken, wodurch sich ein SUTVA-Problem ergibt. Wird ein Nichtniedriglohn-Beschäftigungsverhältnis mit einem Zuschuss versehen, so ist eventuell die Neigung, dem Betrieb eine Reziprozitätsverletzung zuzurechnen, geringer oder sogar überhaupt nicht vorhanden. Mit getrennten Analysen für Alleinerziehende wäre dieses zwar zu lösen gewesen, doch reicht dazu die Fallzahl in der Kontrollgruppe nicht aus. Damit verbleiben 1.199 Befragte, davon 564 Kombilohngeförderte und 635 Vergleichspersonen.

Die Stichprobe der Vergleichspersonen war so gewählt, dass sie eine Zufallsstichprobe aus der Gesamtheit der Personen darstellt, die eine Beschäftigung aus der Arbeitslosigkeit aufnahmen und die entsprechenden bereits genannten Ei-

60 Bei der ursprünglichen MaBnahmeevaluation wurde versucht, dieses Problem mithilfe von Ereignisanalysen ohne vorheriges Matching zu lösen. Bei diesen besteht allerdings keine kausale Interpretierbarkeit im Rahmen des RKM.

61 Eine ausführlichere Beschreibung der Stichprobe findet sich im Methodenbericht von TNS Infratest Sozialforschung (Hartmann 2004).

genschaften besaßen. Eine gezogene Person muss also nicht eine Beschäftigung aufgenommen haben, die im Mainzer Modell förderfähig war. Entscheidend ist, dass sie vor der Maßnahmeteilnahme die Möglichkeit gehabt hätte, einen solchen Arbeitsplatz aufzunehmen. Dies war für jede(n) Neubeschäftigte(n) zum Zeitpunkt der Durchführung des Mainzer Modells prinzipiell der Fall.

Die durchgeführte Befragung enthält eine Vielzahl von Fragen zu unterschiedlichen Themen, von der Erwerbsbiografie über den Haushaltskontext bis hin zu Einstellungsfragen. Sie enthält damit ein breites Spektrum, um möglichst viele Aspekte des selektiven Maßnahmezugangs zu berücksichtigen.

## 6.2 Messung von Selektionseinflüssen und Zielvariable

In diesem Abschnitt sollen die aufgefundenen Selektionseinflüsse sowie die Zielvariablen mit den Informationen aus der Befragung zum Mainzer Modell operationalisiert werden. Die auf den Arbeitsuchenden und potenziellen Maßnahmeteilnehmer zurückgehenden Einflussfaktoren werden mit den Variablen „Wie wurde die Stelle gefunden“ (*über die Arbeitsvermittlung, Bekannte/Freunde, durch eigene Initiative oder Sonstiges*), Suchintensität (*Zahl der genutzten Suchwege*), Migrationshintergrund (*ja/nein*), Teilnahme an einer Weiterbildungsmaßnahme vor Beschäftigungsaufnahme (*nein/ja*), der Berufsausbildung (*keine, Facharbeiter/Lehre, Berufs-/Fachschule/Meister/Techniker, [Fach-]Hochschule und Sonstiges*) und der kumulierten Arbeitslosigkeitsdauer in der Erwerbsbiografie abgedeckt. Der Frage, ob die aufgenommene Tätigkeit eine Übergangslösung darstellt, sowie dem Alter und dem Geschlecht wird ebenfalls potenzielle Relevanz für oder gegen die Teilnahmeentscheidung beigemessen.

Die Förderrichtlinien werden mit folgenden Variablen abgebildet: Haushaltskontext<sup>62</sup> (*alleinlebend ohne Kind, erwerbsloser Partner/ohne Kind; erwerbstätiger Partner/ohne Kind; erwerbsloser Partner/mit Kind; erwerbstätiger Partner/mit Kind*) und Arbeitszeit (*in Stunden*). Die sonstigen Aspekte der Fremdselektion durch die Arbeitsämter sollen durch die Variablen Matchingqualität (*keine Qualifikation nötig, Passung auf Stelle gut, Passung auf Stelle mittel oder schlecht*), Leistungsbezug (*keine Leistung, Arbeitslosengeld, Arbeitslosenhilfe, Sozialhilfe*), Teilnahme an einer Arbeitsbeschaffungsmaßnahme vor der aktuellen Beschäftigung, Teilnahme an einer anderen Maßnahme, Berufserfahrung (*Jahre in Beschäftigung kumuliert*), Art der Kündigung der letzten Beschäftigung (*Eigenkündigung, Arbeitgeberkündigung, Sonstiges*) sowie die regionale Arbeitsmarkttypologie von Blien et al. (2004) erfasst werden.

---

62 Die Alleinerziehenden wurden aus dem Sample entfernt, s. o.



Die Teilnahmeselektivität, die auf die einstellenden Betriebe zurückgeht, soll mithilfe einer Variable für die Branchenzugehörigkeit sowie der Frage erfasst werden, ob das Beschäftigungsverhältnis befristet ist. Da die Betriebsgröße in der verwendeten Befragung nicht erhoben wurde, kann dieser Aspekt auch nicht als Variable in die Berechnung des PS aufgenommen werden und verbleibt so als unbeobachtete Heterogenität. Dies könnte zur verfälschten Messung des kausalen Effektes führen, wenn die Betriebsgröße auch die Lohnzufriedenheit bzw. die Beschäftigungsdauer beeinflusst. Da dies ohne Daten nicht zu beurteilen ist, wird weiter unten (Abschnitt 6.5.1.1) ein Sensitivitätstest (Rosenbaum-Bounds) durchgeführt, der angibt, wie groß der unbeobachtete Einfluss der Betriebsgröße sein müsste, damit die Ergebnisse der Messung des kausalen Effektes einer Kombilohnförderung in Zweifel gezogen werden müssen.

Für die hier durchzuführende Hypothesenprüfung ist die Kontrolle der Lohnhöhe unentbehrlich. Die Tatsache, dass der Kombilohn gerade für Niedriglohnbeschäftigte gezahlt wird, führt dazu, dass die Teilnahme auch selektiv bezüglich der Höhe des Individuallohnes ist. Dafür kommt grundsätzlich der Stundenlohn oder der Monatslohn infrage. Soll geprüft werden, ob der Kombilohnbezug als Zuschlag zum Marktlohn einen kausalen Effekt auf die Lohnzufriedenheit/Beschäftigungsdauer aufweist, muss also der Lohnunterschied bezüglich des reinen Arbeitgeberlohnes (ohne Zuschuss) zwischen Geförderten und Ungeförderten ausgeglichen werden. Dies erfolgt durch die Verwendung des Bruttolohnes, wodurch sich der Nettolohn der Geförderten gerade um die Höhe des Zuschusses von dem regulär Beschäftigter unterscheidet. Dies hat zur Folge, dass sich dann, wenn durch das Matching ein Ausgleich des Stundenlohnes in den beiden Teilgruppen erfolgt ist, der verfügbare Lohn der Geförderten stets höher ist als der der Ungeförderten. Da die Förderung nur erteilt wurde, wenn das Haushaltseinkommen unterhalb einer bestimmten Bedürftigkeitsschwelle lag, ist auch die Höhe des Haushaltseinkommens ein wichtiges Element der Zuweisung einer Förderung. Es ist vorstellbar, dass die subjektive Verletzung der Reziprozität auch davon abhängt, ob und wie viel Einkommen ein eventueller Lebenspartner zum Haushaltseinkommen beiträgt. Neben der faktischen Lohnhöhe ist es zur Stützung der Kausalitätsbehauptung auch nötig, dass sich Geförderte und Ungeförderte nicht schon von vornherein in ihrer Bewertung des finanziellen Aspektes eines Beschäftigungsverhältnisses unterscheiden, also darin, welchen Stellenwert die Einkommenshöhe bei den Beschäftigten einnimmt. Daher muss auch dieser Sachverhalt berücksichtigt werden.

Um das finanzielle Anspruchsniveau der Beschäftigten zu erfassen, werden verschiedene Fragen zur Wichtigkeit bestimmter Lebensbereiche aufgenommen (neben der *Wichtigkeit des Einkommens* auch die nach *Arbeit, Beruf, Freizeit und Familie*). Zudem wird erfasst, ob der Beschäftigte bereits ein Stellenangebot des

Arbeitsamt wegen zu niedrigen Einkommens abgelehnt hat. Die faktische Lohnsituation wird mithilfe der Variablen „Stundenlohn“ und „Haushaltseinkommen“ abgedeckt, und es wird erfasst, ob zusätzliches Einkommen aus einer Nebenbeschäftigung besteht. Es erfolgt eine Einteilung des Stundenlohnes in DM-Klassen (umgerechnet in Euro), da oft fehlende Angaben zum Lohn existieren. Der Anteil von Personen, die keine oder nur unplausible Angaben zum Stundenlohn lieferten, ist in beiden Gruppen gleich. Der Stundenlohn wird aus den Variablen „Arbeitszeit“ und „monatliches Bruttoeinkommen“ gebildet, das Nettoeinkommen wurde nicht abgefragt. Der individuelle Monatslohn wird hier nicht berücksichtigt. Zum einen ergibt er sich aus der Kombination von Arbeitszeit und Stundenlohn, und zum anderen dient er weiter unten zum Test der CIA (vgl. Abschnitt 6.5.1.2).

Die Befragung zum Mainzer Modell enthält keine expliziten Informationen zum Themenbereich Reziprozität, allerdings aber eine Frage nach der Zufriedenheit mit dem Lohn: „Wie zufrieden waren sie mit dem Verdienst? Waren Sie ...“ Die Variable Verdienstzufriedenheit besitzt ursprünglich vier Ausprägungen, „sehr zufrieden“, „zufrieden“, „nicht zufrieden“ und „überhaupt nicht zufrieden“. Die Angaben wurden dichotomisiert mit „nicht zufrieden“ bzw. „überhaupt nicht zufrieden“ einerseits und „sehr zufrieden“ und „zufrieden“ andererseits.

Für die nachfolgenden Analysen wird im Anschluss an die Überlegungen in Kapitel 3 davon ausgegangen werden, dass bei Konstanthalten des Lohnes und anderer Faktoren durch das Matchingverfahren die Antworten auf diese Frage nach der Lohnzufriedenheit Ausdruck der Interpretation des ausgeglichenen Verhältnisses von Lohn und Arbeitsleistung ist, so wie sie der Befragte/Beschäftigte vornimmt. Ein Unterschied der Verdienstzufriedenheit von Geförderten und Ungeförderten wird damit zurückgeführt auf eine durch den Kombilohn erzeugte subjektive Verletzung der Reziprozität.

### 6.3 Propensity Score Matching von Kombilohngeförderten und Ungeförderten

Zentrales Instrument zur Schätzung des kausalen Effektes ist die Bestimmung der Propensity Scores. Im Folgenden wird diese Bestimmung durchgeführt, wobei dies im Gegensatz zu Regressionsanalysen grundsätzlich ohne Kenntnis der Auswirkung verschiedener Spezifikationen der Kontrollvariablen auf das Ergebnis der Analyse erfolgen kann (Rubin 2007: 25).<sup>63</sup> Im Folgenden wird also, noch bevor es an die Berechnung des kausalen Effektes geht, die Güte des Ausgleiches der Selektions-

63 So erfolgen Urteile über die Qualität des Schätzmodelles unabhängig davon, ob sie aus Sicht des Forschers zu einem wünschenswerten (das heißt meist: signifikanten) Ergebnis führt.

Einflüsse zwischen den Gruppen durch das Matching überprüft. Das Matching wurde mit dem Stata Supplement „psmatch2“ (Leuven/Sianesi 2003) in der Version von 2006 durchgeführt. Bei unbekanntem Propensity Scores müssen diese aus den Daten geschätzt werden. Dies kann auf verschiedene Weise erfolgen, etwa mithilfe einer logistischen Regression:

$$P(\mathbf{x}_i) = E(kl = 1 | \mathbf{x}_i) = \left(1 + \exp(-\mathbf{x}_i' \beta)\right)^{-1} \quad (6.1)$$

Dabei ist  $kl = 1$ , wenn ein Beschäftigter eine Kombilohnförderung erhält, und  $kl = 0$ , wenn es sich um eine reguläre Beschäftigung handelt. Der Vektor  $\mathbf{x}_i'$  muss die Aspekte des Selektionsprozesses abdecken, die sowohl die Maßnahmeteilnahme als auch die jeweilige Zielvariable beeinflussen. Die Analyse des Selektionsprozesses erfolgte in Kapitel 5. Die Ergebnisse der logistischen Regression sind in Tabelle 5, Modell 1 abzulesen. Da die Schätzung nur dazu dient, die Propensity Scores zu berechnen, brauchen die Ergebnisse hier nicht im Einzelnen interpretiert werden. Hervorzuheben ist die Erreichung eines mit 0,34 recht hohen Pseudo-R<sup>2</sup>-Wertes. Ein Likelihood-Ratio-Test (Greene 2003: 678) verwirft die Hypothese, dass keine der Variablen einen von null verschiedenen Einfluss hat.

Während sich bei einigen Variablen ihr Einfluss auf den Erhalt der Kombilohnförderung bestätigt, fällt dieser bei anderen insignifikant aus. Zur Frage, ob in solchen Fällen die nicht signifikanten Einflussfaktoren aus dem Modell für den PS entfernt werden sollen, gibt es unterschiedliche Antworten. Einerseits verschärft die Aufnahme von nur schwach mit Treatment- und/oder Zielvariable korrelierten Einflussfaktoren das Common-Support-Problem und erhöht die Varianz der Schätzer (Bryson/Dorsett/Purdon 2002; Augurzky/Schmidt 2001; Black/Smith 2004; Imbens 2004). Andererseits sollten bei der Berechnung der PS möglichst die gesamten kausalrelevanten Selektionsprozesse von den Variablen  $\mathbf{x}_i'$  abgedeckt werden, weshalb ein Auslassen von Variablen eventuell zu einer Verletzung der CIA führt. Daher schlagen Rubin und Thomas (1996: 253 f.) vor, nur dann Variablen nicht zur Schätzung der Propensity Scores zu verwenden, wenn eindeutig ist, dass sie keinen Einfluss auf die Zielvariable haben.

Allerdings hat gerade in kleinen Stichproben – wie sie hier mit ca. 1.000 Fällen vorliegt – die Aufnahme von Variablen, die nur schwach mit Fördererhalt *und* der Zielvariablen korreliert sind, kaum Einfluss auf die Behebung der Selektivität, führt aber stattdessen zu Problemen der Präzision bei der Schätzung des kausalen Effektes (Imbens 2004: 23).

Tabelle 5: Logistische Regressionen der Förderwahrscheinlichkeit vor und nach dem Matching

Wahrscheinlichkeit des Erhaltes einer Kombilohnförderung	Modell 1: alle Variablen (vor dem Matching)		Modell 2: reduzierte Variablenzahl (vor dem Matching)		Modell 3: reduzierte Variablenzahl (nach dem Matching)	
	Odds Ratio	p	Odds Ratio	p	Odds Ratio	p
Haushaltskonstellation						
(R: Alleinlebend ohne Kind)						
Erwerbsloser Partner/ohne Kind	2,615	0,00	2,418	0,00	0,611	0,20
Erwerbstätiger Partner/ohne Kind	1,107	0,76	1,196	0,56	0,704	0,40
Erwerbsloser Partner/mit Kind	6,716	0,00	5,843	0,00	0,726	0,35
Erwerbstätiger Partner/mit Kind	2,340	0,00	2,047	0,01	0,457	0,06
Arbeitszeit	0,954	0,00	0,958	0,00	0,999	0,90
Passung Bewerber/Stelle						
(R: Keine Qualifikation erforderlich)						
Passung gut	1,095	0,70				
Passung mittel/schlecht	1,185	0,61				
Wie wurde die Stelle gefunden						
(R: Arbeitsvermittlung)						
Bekannte/Freunde	0,884	0,63	0,853	0,51	0,688	0,24
Eigene Initiative	1,033	0,90	1,052	0,83	0,591	0,10
Sonstiges	0,601	0,05	0,612	0,05	0,954	0,89
Vorher Weiterbildung ja	2,943	0,03	2,765	0,04	1,197	0,78
Lohnersatzleistungen vom						
Arbeitsamt (R: Arbeitslosengeld)						
Arbeitslosenhilfe	2,694	0,00	2,673	0,00	0,802	0,46
Keine Leistungen vom Arbeitsamt	0,861	0,57	0,803	0,38	1,094	0,79
Sozialhilfebezug	2,885	0,00	2,870	0,00	0,777	0,53
Berufsausbildung (R: Keine)						
Facharbeiter/Lehre	0,940	0,79				
(Berufs-)Fachschule/Meister, Techniker	1,106	0,74				
(Fach-)Hochschule	0,890	0,78				
Sonst., k.A.	0,563	0,25				
Nationalität deutsch	0,688	0,19				
Geschlecht weiblich	0,893	0,60				
Alter	0,997	0,85				
Beschäftigung befristet	2,485	0,00	2,546	0,00	1,050	0,83
Branche (R: Leih-/Zeitarbeitsfirma)						
Reinigungsgewerbe	0,903	0,79	0,834	0,61	0,931	0,88
Hotel- und Gaststättengewerbe	1,100	0,80	0,899	0,77	0,844	0,73
Callcenter	4,604	0,04	4,038	0,05	1,083	0,94
Energie und Wasser; Bergbau, verarbeitendes Gewerbe	0,187	0,00	0,192	0,00	1,023	0,96
Handel	0,813	0,50	0,825	0,52	0,776	0,52
Verkehr/Nachrichtenübermittlung, Banken, Versicherungen	0,527	0,25	0,511	0,22	1,171	0,82
Öffentlicher Dienst/ Sozialversicherung	0,184	0,00	0,200	0,00	1,114	0,84
Bereich andere Dienstleistungen, k.A.	0,613	0,07	0,618	0,07	0,995	0,99
Beruf wichtig	1,022	0,91				
Familie wichtig	0,841	0,43				
Freizeit wichtig	1,197	0,39				

Wahrscheinlichkeit des Erhaltes einer Kombilohnförderung	Modell 1: alle Variablen (vor dem Matching)		Modell 2: reduzierte Variablenzahl (vor dem Matching)		Modell 3: reduzierte Variablenzahl (nach dem Matching)	
	Odds Ratio	p	Odds Ratio	p	Odds Ratio	p
Einkommen wichtig	1,048	0,81	1,073	0,68	1,165	0,52
Arbeit wichtig	1,015	0,94				
Suchintensität (Zahl der genutzten Suchwege)	1,194	0,00	1,169	0,00	0,997	0,97
Bereits ein Stellenangebot vom AA wegen zu niedrigem Einkommen abgelehnt?	1,526	0,16	1,682	0,07	1,035	0,93
Ist Beschäftigung Übergangslösung?	1,123	0,51				
Berufserfahrung in Jahren	1,017	0,28				
Bisherige Arbeitslosigkeitsdauer insgesamt (R: < 6 Monate)						
6-12 Monate	0,418	0,02	0,420	0,01	1,529	0,37
12-24 Monate	1,123	0,71	1,172	0,61	1,619	0,26
> 24 Monate/k.A.	1,006	0,98	1,070	0,81	1,388	0,39
Von ABM in Beschäftigung	1,349	0,37				
Von anderer Maßnahme in Beschäftigung	1,085	0,83				
Art der Kündigung der letzten Beschäftigung (R: Sonstiges)						
Eigenkündigung	1,562	0,19	1,608	0,15	0,991	0,98
Arbeitgeberkündigung	1,585	0,03	1,735	0,01	0,873	0,60
Arbeitsamttypen (R: Typ 1)						
Typ 2	1,055	0,87	0,944	0,85	1,197	0,68
Typ 3	0,570	0,03	0,560	0,01	0,962	0,90
Typ 4	0,762	0,70	0,612	0,48	0,387	0,32
Typ 5	0,423	0,00	0,370	0,00	0,993	0,98
Einkommen aus Nebentätigkeit	0,442	0,07	0,478	0,09	1,308	0,65
Haushaltseinkommen	0,999	0,00	0,999	0,00	1,000	0,20
Stundenlohn (R: bis 4,60 Euro)						
4,61 bis 6,90 Euro	0,952	0,87	1,072	0,81	0,761	0,45
6,91 bis 9,20 Euro	0,687	0,24	0,773	0,40	0,970	0,94
9,21 Euro oder mehr	0,594	0,14	0,694	0,28	0,865	0,74
Keine (plausible) Angabe	0,658	0,19	0,740	0,32	0,794	0,56
Fälle (davon Kombilohn)	974 (500)		983 (503)		372 (186)	
Pseudo-R <sup>2</sup>	0,337		0,328		0,026	
Log likelihood	-447,400		-457,874		-251,119	
ch <sup>2</sup>	454,757		446,440		13,46	
LR-Test (prob > chi2)	0,000		0,000		1,000	
aic	1.008,800		995,749		582,238	
bic	1.287,040		1191,373		738,993	
Anmerkung: Die Odds Ratio errechnet sich als $e^{\beta}$ .						

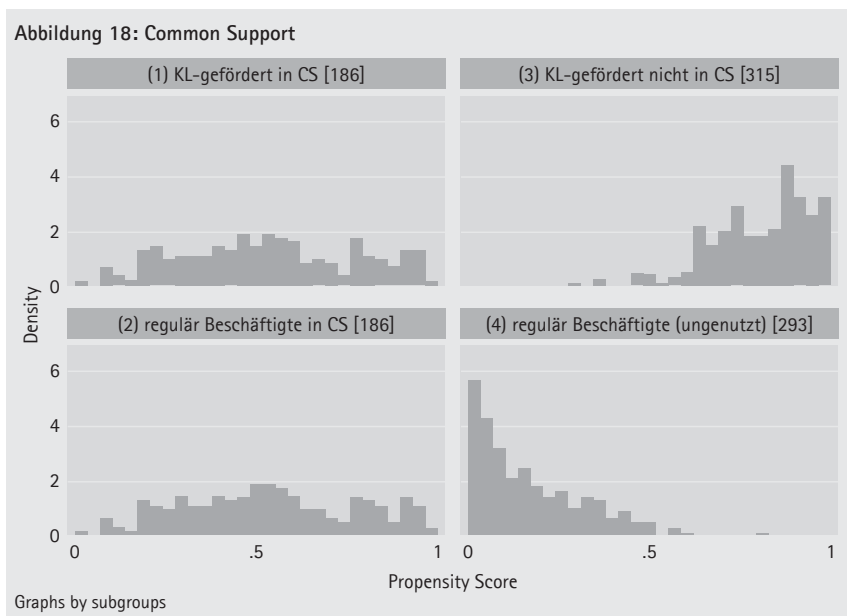
Daher soll das Modell 1 um insignifikante Variablen verkürzt werden,<sup>64</sup> wobei mit der Wichtigkeit des Einkommens und dem Stundenlohn auch insignifikante, aus theoretischen Gesichtspunkten aber relevante Variablen behalten wurden (Modell 2). Ein

64 Es hat sich gezeigt, dass die Ergebnisse der Schätzung des kausalen Effektes sehr stabil in Bezug auf das Auslassen der schwachen Selektionseinflüsse sind. Dies ist an Tabelle A2 im Anhang abzulesen, wo der kausale Effekt mit dem vollen Variablenpektrum berechnet wurde.

Likelihood-Ratio-Test<sup>65</sup> zeigt, dass die Hypothese, von den ausgelassenen Variablen weise keine einen von null verschiedenen Koeffizienten auf, nicht abgelehnt werden kann ( $p = 0,906$ ). Der Vergleich der Werte der Informationskriterien AIC (Akiakes Informationskriterium) und BIC (Bayesianisches Informationskriterium) zeigt, dass mit dem Weglassen der Variablen die Güte des Modells sogar steigt.

Wichtig für die Aussagekraft des später zu berechnenden kausalen Effektes ist der Common Support. Dieser bezeichnet den gemeinsamen Stützbereich der Verteilung der PS. Für die Geförderten lassen sich nur innerhalb dieses Common Support Matchingpartner finden.

In Abbildung 18 werden die geschätzten PS der Geförderten (1) und der Ungeförderten (2) im Common Support gegenübergestellt. Dabei zeigt sich, dass es keine Lücken im Common Support gibt, abgesehen von einer dünnen Besetzung im Bereich sehr niedriger Teilnahmewahrscheinlichkeiten. Kombilohngeförderte, für die keine Matchingpartner im Rahmen des angegebenen Calipers gefunden wurde (3), weisen hauptsächlich eine sehr hohe Teilnahmewahrscheinlichkeit auf, während umgekehrt die regulär Beschäftigten, welche nicht als Matchingpartner für Geförderte verwendet wurden, eher niedrige Teilnahmewahrscheinlichkeiten besitzen (4). Insgesamt erstreckt sich der Common Support also fast über die gesamte Spanne der Propensity Scores, auch wenn für eine Reihe von Geförderten (im Rahmen des allerdings sehr engen Calipers, siehe unten) keine Matchingpartner gefunden wurden.



65 Dabei wurde aus technischen Gründen die Fallzahl auf die 974 Fälle des Modells 1 beschränkt.

Das Matching sollte, wenn es erfolgreich ist, zu einem Ausgleich der Verteilungen der in der logistischen Regression verwendeten Variablen in den beiden Gruppen (Geförderte und Ungeförderte) führen. Man spricht dann von ausbalancierten Stichproben oder Teilgruppen. Ob in einer konkreten Analyse mithilfe der PS eine Vergleichbarkeit der Teilgruppen erreicht wurde, kann mithilfe verschiedener Techniken überprüft werden, auch wenn es kein eindeutiges Kriterium gibt.

Zunächst kann die Verteilung der Propensity Scores betrachtet werden (Tabelle 6). Es zeigt sich, dass die Verteilungen vor dem Matching in beiden Gruppen sehr unterschiedlich waren. Im Mittel weisen Ungeförderte weit geringere Wahrscheinlichkeit für eine Kombilohnförderung auf als die Geförderten, sowohl was den Median als auch was den Mittelwert betrifft. Auch im Vergleich der Quantilswerte zeigt sich diese Tendenz. Der Eindruck ungleicher Verteilungen lässt sich mithilfe eines statistischen Tests prüfen. Ein Kolmogorov-Smirnov-Anpassungstest (K-S-Test) unter Verwendung des Bootstrap-Verfahrens zur Schätzung der Standardfehler testet die Hypothese der Gleichheit beider Verteilungen und kommt im vorliegenden Fall zum Verwerfen der Hypothese. Demgegenüber erweisen sich die Verteilungen nach dem Matchen als äußerst ähnlich, sowohl in den Mittelwerten als auch der Standardabweichung. Bei den gematchten Daten lässt sich die Hypothese der Gleichheit der Verteilungen nun auch nicht mehr verwerfen.

Tabelle 6: Verteilung der PS vor und nach dem Matchen

Percentiles	Vor dem Matching		Nach dem Matching	
	Kombilohn	Regulärer Lohn	Kombilohn	Regulärer Lohn
1 %	0,111	0,005	0,073	0,073
5 %	0,240	0,016	0,174	0,174
10 %	0,367	0,028	0,212	0,211
25 %	0,572	0,086	0,357	0,357
Median	0,750	0,246	0,522	0,523
75 %	0,884	0,475	0,745	0,740
90 %	0,947	0,735	0,883	0,882
95 %	0,971	0,848	0,927	0,922
99 %	0,989	0,949	0,963	0,960
N	503	480	186	186
Arithm. Mittelwert	0,704	0,310	0,535	0,533
Standardabweichung	0,224	0,259	0,239	0,238
K-S-Test auf Gleichheit der Verteilungen, Bootstrap mit Stata	t = 25,36		t = 0,16	

Eine weitere und intuitiv plausible Möglichkeit des Tests der Matchingqualität besteht darin, die logistische Regression der Maßnahmeteilnahme nochmals durchzuführen, allerdings nur mit den Personen, für die ein Matchingpartner gefunden wurde. Die Ergebnisse sind in Modell 3 der Tabelle 5 den Ergebnissen der ursprünglichen logistischen Regression gegenübergestellt. Dabei zeigt sich, dass sich nun ein LR-Test auf Gleichheit aller Koeffizienten mit null nicht verwerfen lässt. Die Wirkung des Matchings zeigt sich schließlich auch darin, dass nun das Pseudo-R<sup>2</sup> nicht mehr 0,34, sondern lediglich 0,03 beträgt und auch die Werte der Informationskriterien stark zurückgegangen sind.

Eine andere Möglichkeit ist der t-Test auf Gleichheit der Mittelwerte der verwendeten Variablen (Rosenbaum/Rubin 1985). Wird etwa das Haushaltseinkommen zur Konstruktion des PS verwendet, dann dürfen sich Geförderte und Ungeförderte in dieser Hinsicht nicht mehr signifikant unterscheiden. Von Vorteil und typisch für Analysen, die sich des Matchingverfahrens bedienen, ist eine im Vergleich zur Gruppe der Geförderten sehr große Gruppe der Ungeförderten. Dies wird z. B. mit der Verwendung prozessproduzierter Massendaten der BA erreicht. In einer solchen Situation lassen sich oft nahezu vollständig ausbalancierte Stichproben erreichen. Bei relativ wenig Ungeförderten (oder, wie es hier der Fall ist, sogar weniger ungeförderte als geförderte Beschäftigte) besteht die Gefahr, dass keine Ausbalancierung zu erreichen ist.<sup>66</sup> Tabelle 7 zeigt, dass hier trotz relativ weniger Personen in der Gruppe der Ungeförderten eine befriedigende Ausbalancierung der Eigenschaften beider Gruppen erreicht wurde. Im Vergleich der Mittelwerte ist zu sehen, dass nach dem Matching zwischen den Geförderten und Ungeförderten nun keine signifikanten Unterschiede mehr bestehen.

Eine weitere Möglichkeit ist die Berechnung des standardisierten Bias (sbias). Er gibt für jedes Merkmal x das Ausmaß der Unterschiede zwischen den Gruppen an (Rosenbaum/Rubin 1985):

$$sbias = \frac{\bar{x}_{kl=0} - \bar{x}_{kl=1}}{\sqrt{(s_{kl=0}^2 + s_{kl=1}^2)} \cdot 0,5} \cdot 100 \quad (6.2)$$

Dabei ist  $\bar{x}_{kl=0} - \bar{x}_{kl=1}$  die Differenz zwischen den Mittelwerten des Merkmals x unter den Nichtgeförderten und den Geförderten, und  $s^2$  bezeichnet die jeweiligen Varianzen innerhalb der Gruppen. Der sbias wird sowohl vor dem Matching als auch danach berechnet.

<sup>66</sup> In einem solchen Fall kann eine nachträgliche Regressionsanpassung, das heißt eine Regression nur mit den auch im Matching verwendeten Personen, durchgeführt werden.



Tabelle 7: Mittelwertvergleiche und sbias vor und nach dem Matchen

		Mittelwert		sbias	bias red.	t-test	
		Kombilohn	Regulärer Lohn			t	P > t
Haushaltskonstellation (Alleinlebend ohne Kind)	vor	0,202	0,290	-20,70		-3,24	0,00
	nach	0,280	0,237	10,00	51,50	0,95	0,35
Erwerbsloser Partner/ ohne Kind	vor	0,130	0,113	5,20		0,81	0,42
	nach	0,129	0,140	-3,30	36,80	-0,30	0,76
Erwerbstätiger Partner/ ohne Kind	vor	0,080	0,180	-30,00		-4,71	0,00
	nach	0,118	0,118	0,00	100,00	0,00	1,00
Erwerbsloser Partner/ mit Kind	vor	0,401	0,192	47,00		7,33	0,00
	nach	0,312	0,285	6,00	87,10	0,57	0,57
Erwerbstätiger Partner/ mit Kind	vor	0,188	0,225	-9,30		-1,46	0,14
	nach	0,161	0,220	-14,60	-56,30	-1,45	0,15
Arbeitszeit	vor	31,5	35,5	-33,10		-5,19	0,00
	nach	33,8	33,7	1,20	96,40	0,11	0,91
Wie wurde die Stelle gefunden (Arbeitsvermittlung)	vor	0,281	0,194	20,60		3,22	0,00
	nach	0,280	0,220	14,00	32,20	1,32	0,19
Bekannte/Freunde	vor	0,248	0,278	-6,90		-1,07	0,28
	nach	0,253	0,280	-6,10	10,90	-0,59	0,56
Eigene Initiative	vor	0,289	0,248	9,20		1,45	0,15
	nach	0,247	0,301	-12,10	-31,20	-1,16	0,25
Sonstiges	vor	0,182	0,280	-23,40		-3,67	0,00
	nach	0,220	0,199	5,10	78,10	0,51	0,61
Vorher Weiterbildung ja	vor	0,032	0,027	2,80		0,44	0,66
	nach	0,038	0,027	6,30	-124,20	0,59	0,56
Lohnersatzleistungen vom Arbeitsamt (Arbeitslosengeld)	vor	0,214	0,374	-35,70		-5,59	0,00
	nach	0,269	0,263	1,20	96,60	0,12	0,91
Arbeitslosenhilfe	vor	0,591	0,301	61,00		9,54	0,00
	nach	0,484	0,489	-1,10	98,10	-0,10	0,92
Keine Leistungen vom Arbeitsamt	vor	0,196	0,326	-29,90		-4,69	0,00
	nach	0,247	0,247	0,00	100,00	0,00	1,00
Sozialhilfebezug	vor	0,180	0,063	36,40		5,67	0,00
	nach	0,102	0,113	-3,30	90,80	-0,33	0,74
Beschäftigung befristet	vor	0,477	0,305	35,80		5,60	0,00
	nach	0,409	0,371	7,80	78,20	0,74	0,46
Branche Leih-/Zeitarbeitsfirma	vor	0,242	0,111	34,80		5,43	0,00
	nach	0,194	0,188	1,40	95,90	0,13	0,90
Reinigungsgewerbe	vor	0,100	0,058	15,30		2,39	0,02
	nach	0,075	0,075	0,00	100,00	0,00	1,00
Hotel- und Gaststättengewerbe	vor	0,084	0,058	9,90		1,54	0,12
	nach	0,065	0,070	-2,10	78,80	-0,21	0,84
Callcenter	vor	0,046	0,008	23,30		3,61	0,00
	nach	0,011	0,011	0,00	100,00	0,00	1,00
Energie und Wasser; Bergbau, verarbeitendes Gewerbe	vor	0,074	0,278	-55,50		-8,74	0,00
	nach	0,124	0,118	1,50	97,40	0,16	0,87
Handel	vor	0,166	0,121	12,70		1,99	0,05
	nach	0,124	0,151	-7,70	39,70	-0,75	0,45
Verkehr/Nachrichtenüberm., Banken, Versicherungen	vor	0,020	0,027	-4,70		-0,74	0,46
	nach	0,032	0,032	0,00	100,00	0,00	1,00
Öffentlicher Dienst/ Sozialversicherung	vor	0,034	0,077	-19,00		-2,98	0,00
	nach	0,065	0,054	4,70	75,20	0,44	0,66

## Der kausale Effekt des Kombilohnes auf Lohnzufriedenheit und Beschäftigungsstabilität

		Mittelwert		sbias	bias red.	t-test	
		Kombilohn	Regulärer Lohn			t	P > t
Bereich andere Dienstleistungen, k.A.	vor	0,236	0,261	-5,90		-0,92	0,36
	nach	0,312	0,301	2,50	57,70	0,22	0,82
Einkommen wichtig	vor	0,509	0,445	12,90		2,02	0,04
	nach	0,462	0,457	1,10	91,60	0,10	0,92
Suchintensität (Zahl der genutzten Suchwege)	vor	6,164	5,269	45,20		7,10	0,00
	nach	6,054	6,070	-0,80	98,20	-0,10	0,92
Bereits ein Stellenangebot vom AA wegen zu niedrigem Einkommen abgelehnt?	vor	0,108	0,084	8,30		1,29	0,20
	nach	0,091	0,091	0,00	100,00	0,00	1,00
Bisherige Arbeitslosigkeitsdauer insgesamt < 6 Monate	vor	16,122	15,706	4,10		0,64	0,52
	nach	16,848	15,467	13,60	-231,60	1,30	0,20
6–12 Monate	vor	0,062	0,177	-36,20		-5,68	0,00
	nach	0,113	0,113	0,00	100,00	0,00	1,00
12–24 Monate	vor	0,188	0,203	-3,80		-0,59	0,56
	nach	0,199	0,167	8,10	-116,80	0,80	0,42
> 24 Monate/k.A.	vor	0,659	0,449	43,10		6,75	0,00
	nach	0,570	0,565	1,10	97,40	0,10	0,92
Art der Kündigung der letzten Beschäftigung (Sonstiges)	vor	0,617	0,628	-2,40		-0,37	0,71
	nach	0,629	0,624	1,10	53,80	0,11	0,92
Eigenkündigung	vor	0,068	0,077	-3,60		-0,57	0,57
	nach	0,070	0,070	0,00	100,00	0,00	1,00
Arbeitgeberkündigung	vor	0,315	0,294	4,60		0,71	0,48
	nach	0,301	0,306	-1,20	74,40	-0,11	0,91
Arbeitsamttypen Typ 1	vor	0,445	0,246	42,70		6,67	0,00
	nach	0,376	0,355	4,60	89,20	0,43	0,67
Typ 2	vor	0,114	0,086	9,40		1,47	0,14
	nach	0,113	0,097	5,40	42,80	0,51	0,61
Typ 3	vor	0,263	0,349	-18,50		-2,90	0,00
	nach	0,317	0,328	-2,30	87,40	-0,22	0,83
Typ 4	vor	0,018	0,015	2,60		0,41	0,68
	nach	0,011	0,022	-8,50	-220,90	-0,82	0,41
Typ 5	vor	0,160	0,305	-34,90		-5,47	0,00
	nach	0,183	0,199	-3,90	88,90	-0,39	0,69
Einkommen aus Nebentätigkeit	vor	0,034	0,052	-9,00		-1,41	0,16
	nach	0,043	0,038	2,60	70,60	0,26	0,79
Haushaltseinkommen	vor	1423	1811	-49,40		-7,77	0,00
	nach	1498	1479	2,40	95,20	0,27	0,79
Stundenlohn bis 4,60 Euro	vor	0,164	0,098	19,50		3,05	0,00
	nach	0,151	0,140	3,20	83,60	0,29	0,77
4,61 bis 6,90 Euro	vor	0,291	0,217	17,10		2,67	0,01
	nach	0,226	0,269	-9,90	42,10	-0,96	0,34
6,91 bis 9,20 Euro	vor	0,206	0,228	-5,30		-0,83	0,40
	nach	0,242	0,204	9,10	-71,30	0,87	0,39
9,21 Euro oder mehr	vor	0,126	0,240	-29,90		-4,69	0,00
	nach	0,161	0,161	0,00	100,00	0,00	1,00
Keine (plausible) Angabe	vor	0,214	0,217	-0,90		-0,13	0,89
	nach	0,220	0,226	-1,30	-51,60	-0,12	0,90
Mittel über alle Variablen	vor			21,10			
	nach			4,10			

Ist der Wert des standardisierten Bias vor dem Matching groß, sollte er nach dem Matching geringer ausfallen und möglichst gegen null gehen. Tabelle 7 zeigt, dass der Bias in fast allen Fällen deutlich zurückgegangen ist. Im Mittel über alle Variablen betrug er vor dem Matching 21,1 %, danach lediglich 4,1 %. In einzelnen Fällen nahm der Bias allerdings leicht zu, ohne jedoch zu signifikanten Ergebnissen des Mittelwertvergleiches zu führen. Daher kann von einer hinreichend guten Balancierungsqualität des Matchingvorganges ausgegangen werden.

Insgesamt lässt sich damit auf Basis der durchgeführten Tests feststellen, dass das durchgeführte Propensity Score Matching zu einem Ausgleich der Verteilungen aller beobachteten Selektionseinflüsse geführt hat. Unter der Bedingung, dass keine unbeobachteten kausalrelevanten Selektionseinflüsse existieren, spricht dies für die interne Validität der nun im Anschluss zu messenden kausalen Effekte.

## 6.4 Prüfung der Hypothesen

*Hypothese 1: Im Vergleich zu regulärer Beschäftigung führt eine Kombilohnförderung durch die subjektive Verletzung der Reziprozität trotz objektiver Einkommenserhöhung zu geringerer Lohnzufriedenheit.*

Auf der Basis der gematchten Stichprobe soll nun zunächst die Hypothese 1 überprüft werden. Es wurde ein „Ein-nächster-Nachbar“-Matching ohne Zurücklegen durchgeführt, auch weil dies zur Bestimmung der Standardfehler nötig ist, wenn das Matching später mit einer Regressionsanpassung ergänzt wird. Für die Wahl des Calipers geben Cochran und Rubin (1973) den Rat, sich an  $c = 0,2\sqrt{(\hat{\sigma}_{kl=0}^2 + \hat{\sigma}_{kl=1}^2)}/2$  zu orientieren. Im vorliegenden Fall wäre dies der Wert von ca. 0,048. Da eine Differenz von 5 Prozentpunkten im Propensity Score allerdings sehr groß erscheint, wird hier der ungleich engere und daher anspruchsvollere Caliper von  $c = 0,005$  gewählt. Damit dürfen sich Geförderte und zugematchte Ungeförderte nur um einen halben Prozentpunkt bei der Teilnahmewahrscheinlichkeit unterscheiden. Dies verringert zwar die Fallzahl, führt aber zu einer stärkeren Reduktion der Selektionsverzerrung.

Von den 501 Kombilohngeförderten wurden für 186 Personen adäquate Matchingpartner gefunden. Das Ergebnis des Matchings in Bezug auf den kausalen Effekt des Kombilohnes auf die Lohnzufriedenheit zeigt Tabelle 8. Vor dem Matching gaben 36 % der Kombilohngeförderten und 59 % der Ungeförderten an, dass sie mit ihrem Lohn zufrieden sind. Der von Selektionseffekten unbereinigte Unterschied beträgt damit –23 Prozentpunkte. Die Veränderung der Anteile nach dem Matchen zeigt, dass alles in allem eine Negativselektion in die Kombilohnförderung stattfand. Es wurden verstärkt solche Personen bzw. Beschäftigungsverhältnisse gefördert, die auch ohne Kombilohnförderung unzufriedener mit dem Lohn gewesen

wären. Diese Selektivität ist für einen Teil der unterschiedlichen Lohnzufriedenheit verantwortlich, kann sie allerdings nicht vollständig erklären, da auch nach dem PS-Matching ein Unterschied zwischen den Anteilen in beiden Gruppen besteht. Dabei ergibt sich eine Schätzung des mittleren kausalen Effektes der Kombilohnförderung auf die Geförderten von rund  $-0,14$  ( $M\hat{K}E_T = 0,41 - 0,55 = -0,14$ ). Damit verringert sich der Anteil der mit ihrem Lohn zufriedenen Personen durch den Kombilohnbezug um 14 Prozentpunkte gegenüber regulären Beschäftigungsverhältnissen.

Tabelle 8: Der kausale Effekt des Kombilohnes auf die Lohnzufriedenheit

	Zufrieden mit dem Lohn (Anteil)		Effekt	Standardfehler	t-Wert
	Kombi-lohn	Regulärer Lohn			
Vor dem Matching	0,36	0,59	-0,23	0,03	-7,38
Nach dem Matching	0,41	0,55	-0,14	0,05	-2,82
Zahl der Matches	186				

Mit einem t-Wert von  $-2,82$  ist der kausale Effekt des Kombilohnes auf die Lohnzufriedenheit signifikant. Damit liefert das Propensity Score Matching ein – im Rahmen statistischer Unschärfen – deutliches Ergebnis, welches die Hypothese 1 stützt.

Caliendo und Kopeinig (im Erscheinen) schlagen allerdings vor, die Sensitivität der Ergebnisse des Matchings noch im Hinblick auf den verwendeten Matchingalgorithmus zu prüfen. Dies soll im Folgenden durchgeführt werden. Wie bereits in Abschnitt 4.1.4 aufgeführt, erfolgt die Schätzung des mittleren kausalen Effektes des Kombilohnes bei den Geförderten über folgende Mittelwertdifferenz, wobei je nach Wahl des Algorithmus die Gewichte  $w(i, j)$  unterschiedlich festgelegt werden:

$$M\hat{K}E_T = \frac{1}{n_1} \sum_{i \in I_1} (Iz_i^1) - \frac{1}{n_1} \sum_{i \in I_1} \sum_{j \in I_0} w(i, j) (Iz_j^0) \tag{6.3}$$

Insgesamt zeigt die Tabelle 9, dass sich bei allen durchgespielten Varianten des Matchingalgorithmus ein signifikant positiver kausaler Effekt des Kombilohnes ergibt. Dies spricht für die Validität des in Tabelle 8 dargestellten Ergebnisses. Im Folgenden sollen die Ergebnisse kurz besprochen werden.

Den einfachsten Fall stellt das Ein-nächster-Nachbar-Verfahren auf Basis des Propensity Score ohne Zurücklegen dar. Dabei vereinfacht sich die Gewichtung zu  $w(i, j) = 0$  für Ungeförderte, die nicht als Matchingpartner ausgewählt werden, und  $w(i, j) = 1$  für diejenigen, die ausgewählt werden. Für jeden Geför-

derden wird ein ungeförderter Matchingpartner gesucht, und zwar der, der dem geförderten im Hinblick auf die Differenz der geschätzten PS am ähnlichsten ist ( $\min_j |P_i - P_j|$ ). Dabei kann jeder Ungeförderte nur einmal berücksichtigt werden (ohne Zurücklegen). Wie in Tabelle 9 abzulesen ist, führt dieses Vorgehen mit einem geschätzten kausalen Effekt von  $-23$  Prozentpunkten zum selben Ergebnis wie ein simpler Vergleich beider Teilgruppen ohne Matching. Der Grund ist, dass auch solche Ungeförderten als Matchingpartner verwendet werden, die zwar die nächsten Nachbarn sind, dennoch aber einen sehr unterschiedlichen Propensity Score aufweisen. Generell empfiehlt es sich daher stets, nur solche Ungeförderte als Matchingpartner zuzulassen, bei denen die Differenz der Propensity Scores  $|\hat{P}_i - \hat{P}_j| < c$ , dem Caliper, ist. Führt man einen solchen Caliper ein, etwa von der Größenordnung  $c = 0,01$ , dann ist zu sehen, dass für 196 Geförderte ein Matchingpartner gefunden wird, der die Abstandsbedingung erfüllt. Der kausale Effekt ist mit  $-10$  Prozentpunkten gegenüber dem oben in Tabelle 8 ausgewiesenen Effekt von  $-14$  Prozentpunkten etwas geringer, aber immer noch signifikant von null verschieden.

Da Ungeförderte nicht nur für einen Geförderten, sondern für mehrere als Matchingpartner verwendet werden können, kann auch eine Auswahl mit Zurücklegen erfolgen. Nimmt man etwa an, dass zwei Kombilohngeförderte den selben Propensity Score aufweisen, und es liegt ein Nichtgeförderter vor, der auch diesen PS aufweist, alle anderen weisen einen mehr oder weniger davon verschiedenen Wert auf, dann führt das Verfahren ohne Zurücklegen dazu, dass für einen der Geförderten der „perfekte“ Matchingpartner gefunden wird, für den anderen aber eine Person mit unterschiedlichen PS hinzugematcht wird. Das Verfahren mit Zurücklegen würde den „perfekten“ Matchingpartner hingegen für beide Geförderte verwenden. Es zeigt sich, dass auch hier ein negativer kausaler Effekt gemessen wird, unabhängig davon, ob ein eher weiter Caliper von  $0,01$  oder ein enger Caliper von  $0,001$  gewählt wird. Generell fällt er weit stärker aus als im Fall ohne Zurücklegen.

Auch bei der Verwendung der Log Odds als Basis des Matchings statt der Propensity Scores kommt es zu Ergebnissen, die keine substanziellen Abweichungen von der Schätzung in Tabelle 8 aufweisen.

Bei mehr als einem (in Tabelle 9 beispielhaft drei) nächsten Nachbarn ist ebenfalls sowohl für das Propensity Matching als auch für das Log Odds Matching ein negativer und signifikanter kausaler Effekt zu beobachten. Das ändert sich auch dann nicht, wenn statt des Ein-nächster-Nachbar-Verfahrens ein Radius-Matching erfolgt, bei dem alle im Bereich des Calipers liegenden Matchingpartner verwendet werden.

Tabelle 9: Sensitivität des geschätzten kausalen Effektes gegenüber Varianten des Matchingalgorithmus

	Zufrieden mit dem Lohn (Anteil)		Kausaler Effekt	Standardfehler	t-Wert	Zahl der Matches
	Kombilohn	Regulärer Lohn				
Ohne Matching	0,36	0,59	-0,23	0,03	-7,38	-
<b>Ein nächster Nachbar; Propensity Score (ohne Zurücklegen)</b>						
Ohne Caliper	0,36	0,59	-0,23	0,03	-7,26	476
Caliper = 0,01	0,44	0,54	-0,10	0,05	-2,03	196
Caliper = 0,005	0,41	0,55	-0,14	0,05	-2,82	186
Caliper = 0,001	0,39	0,56	-0,16	0,06	-2,51	117
<b>Ein nächster Nachbar; Propensity Score (mit Zurücklegen)</b>						
Caliper = 0,01	0,36	0,63	-0,27	0,06	-4,18	471
Caliper = 0,005	0,38	0,61	-0,23	0,06	-3,75	411
Caliper = 0,001	0,38	0,57	-0,19	0,06	-2,93	195
<b>Ein nächster Nachbar; Log Odds (ohne Zurücklegen)</b>						
Caliper = 0,05	0,40	0,54	-0,14	0,05	-2,85	179
Caliper = 0,01	0,44	0,57	-0,13	0,06	-2,33	149
Caliper = 0,005	0,38	0,56	-0,19	0,07	-2,76	108
<b>Ein nächster Nachbar; Log Odds (mit Zurücklegen)</b>						
Caliper = 0,05	0,36	0,62	-0,26	0,06	-4,09	439
Caliper = 0,01	0,40	0,58	-0,18	0,06	-3,08	247
Caliper = 0,005	0,40	0,58	-0,19	0,07	-2,75	166
<b>Maximal drei nächste Nachbarn; Propensity Score (mit Zurücklegen)</b>						
Caliper = 0,01	0,36	0,59	-0,23	0,06	-3,96	471
Caliper = 0,005	0,38	0,60	-0,22	0,06	-3,89	411
Caliper = 0,001	0,38	0,58	-0,19	0,06	-3,06	195
<b>Maximal drei nächste Nachbarn; Log Odds (mit Zurücklegen)</b>						
Caliper = 0,05	0,36	0,60	-0,24	0,06	-4,09	439
Caliper = 0,01	0,40	0,59	-0,19	0,06	-3,36	247
Caliper = 0,005	0,40	0,57	-0,18	0,07	-2,69	166
<b>Radius-Matching; Propensity Score (mit Zurücklegen)</b>						
Caliper = 0,005	0,38	0,60	-0,22	0,06	-3,94	411
<b>Kernel-Matching; Propensity Score (mit Zurücklegen)</b>						
Bandwidth = 0,006	0,38	0,60	-0,22	0,06	-3,98	411
Rundungsbedingte Ungenauigkeiten						

Damit lässt sich feststellen, dass über die gesamte Bandbreite der möglichen Matchingalgorithmen stets ein negativer kausaler Effekt des Kombilohnes gemessen wird, wobei dieser meist sogar größer ausfällt, als der in der gewählten Variante des Ein-nächster-Nachbar-Matchings ohne Zurücklegen. Diese hat jedoch den Vorteil, dass sie durch die nur einmalige Verwendung der Kontrollpersonen bei der für die Prüfung von Hypothese 2 und 3 nötigen Regressionsanpassung eine Berechnung der Standardfehler mithilfe der üblichen Vorgehensweise ermöglichen. Bei mehrmaliger Verwendung einzelner Beobachtungen müssten auch die Standardfehler entsprechend angepasst werden.

**Hypothese 2:** *Kombilohngeförderte Beschäftigungsverhältnisse enden früher als reguläre Beschäftigungsverhältnisse.*

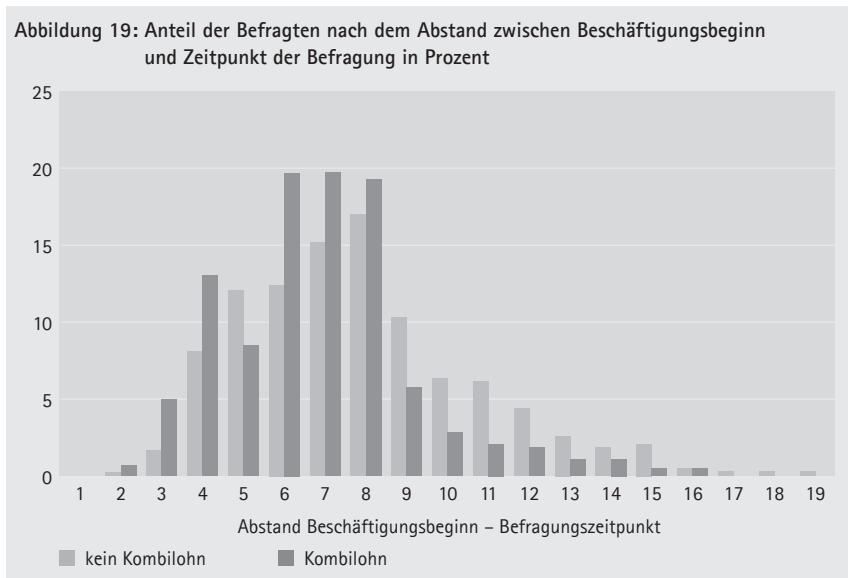
Auch zum Test der Hypothese 2 bietet sich zunächst eine Berechnung des Matchingschätzers an. Dabei kommen zwei Operationalisierungen der Beschäftigungsstabilität infrage: die Beschäftigungsdauer und der Beschäftigungsstatus (in Beschäftigung ja/nein) zum Zeitpunkt der Befragung. Für beide Zielvariablen sind die Ergebnisse des Matchings in Tabelle 10 ausgewiesen.

Tabelle 10: Propensity Score Matching: Effekte des Kombilohnes auf die Beschäftigungsdauer und den Beschäftigungsstatus

	Kombilohn	Regulärer Lohn	Effekt	Standardfehler	t-Wert
Zeit in Beschäftigung (Monate)					
Vor dem Matching	5,428	6,474	-1,046	0,197	-5,30
Nach dem Matching	5,242	6,067	-0,825	0,300	-2,75
Zahl der Matches	194				
Beschäftigungsstatus: Beschäftigt					
Vor dem Matching	0,330	0,338	-0,008	0,030	-0,27
Nach dem Matching	0,340	0,365	-0,025	0,048	-0,53
Zahl der Matches	197				

Es ergeben sich hierbei zwei unterschiedliche Effekte. Der kausale Effekt des Kombilohnes auf den Beschäftigungsstatus ist zwar negativ, jedoch mit 2,5 Prozentpunkten sehr gering und zudem nicht signifikant. Im Gegensatz dazu zeigt sich für die Beschäftigungsdauer ein signifikant negativer Effekt des Kombilohnes, der der Hypothese 2 entspricht. Allerdings enthält die Variable „Beschäftigungsdauer“ sowohl zensierte als auch nicht zensierte Werte, also solche, bei denen zum Befragungszeitpunkt die Beschäftigung noch andauert, und solche, bei denen diese bereits geendet hat. Aufgrund des Erhebungsdesigns existiert

für Kombilohnempfänger allerdings tendenziell eine frühere Zensierung als für Ungeförderte, wodurch sich auch tendenziell die gemessene Zeit in Beschäftigung verringert. Daher kann der ausgewiesene Effekt auch darauf und nicht (oder nicht nur) auf die Wirkung des Kombilohnes zurückgehen. Abbildung 19 zeigt die Verteilung der relativen Häufigkeiten. Zu sehen ist, dass die Verteilung bei den Ungeförderten deutlich nach rechts verschoben ist. Insofern ergeben sich Zweifel, ob der in Tabelle 10 ausgewiesene Effekt tatsächlich der kausale Effekt des Kombilohnes ist.



Eine Möglichkeit, mit solchen Zensierungen umzugehen, ist die Ereignisanalyse. Diese basiert auf vier miteinander verbundenen Konzepten, der Dichtefunktion einer Zufallsvariable, ihrer Verteilungsfunktion, der sogenannten Survivorfunktion und der Hazardrate (Wooldridge 2002: 685 ff.; Blossfeld/Rohwer 2002: 31 ff.).  $T$  sei eine solche Zufallsvariable. Sie gibt die Zeit an, zu der ein bestimmtes Ereignis eintritt.  $T$  folgt einer vorerst noch un spezifizierten Dichtefunktion

$$f(t) = \frac{dF}{dt}(t) \tag{6.4}$$

bzw. einer kumulierten Verteilungsfunktion

$$F(t) = P(T \leq t) \text{ für } t \geq 0 \tag{6.5}$$

Die Dichtefunktion bestimmt die Wahrscheinlichkeit, dass das relevante Ereignis eintritt, während die Verteilungsfunktion die Wahrscheinlichkeit angibt, dass es



bis zum Zeitpunkt  $t$  eingetreten ist. Im Gegensatz dazu ist die Survivor- oder auch Überlebensfunktion definiert als die Wahrscheinlichkeit, dass ein Ereignis erst nach einem bestimmten Zeitpunkt realisiert wird:

$$S(t) \equiv P(T > t) \quad (6.6)$$

Sie ist damit die Gegenwahrscheinlichkeit zu Gleichung 6.5:  $S(t) = 1 - F(t)$ .

Das letzte zentrale Konzept der Ereignisanalyse ist die zeitabhängige Übergangs- oder Hazardrate  $\lambda(t)$  von einem Zustand in den anderen. Sie wird aus der Wahrscheinlichkeit  $P(t \leq T < t' | T \geq t)$  abgeleitet, die angibt, mit welcher Wahrscheinlichkeit ein Ereignis im Zeitintervall  $[t, t']$  eintritt, wenn es bisher, d. h. seit dem Zeitpunkt  $t_0$ , dem Beginn des Beobachtungszeitraumes bzw. dem Einsetzen der Maßnahme, noch nicht eingetreten war. Da die Zeit ein stetiges Merkmal ist, muss die Hazardrate  $\lambda(t)$  dann als der Grenzwert definiert werdenden (Blossfeld/Rohwer 2002: 32).

$$\lambda(t) = \lim_{t' \rightarrow t} \frac{P(t \leq T < t' | T \geq t)}{t' - t} \quad (6.7)$$

Die verschiedenen Konzepte lassen sich wie folgt zueinander in Beziehung setzen:

$$\lambda(t) = \lim_{t' \rightarrow t} \frac{P(t \leq T < t' | T \geq t)}{t' - t} = \lim_{t' \rightarrow t} \frac{F(t' - t) - F(t)}{(t' - t)(1 - F(t))} = \frac{f(t)}{S(t)} \quad (6.8)$$

Soll nun der Effekt verschiedener Merkmale auf die Übergangswahrscheinlichkeit geschätzt werden, muss ein Verlauf der Hazardrate spezifiziert werden. In den meisten Fällen werden die Wahrscheinlichkeiten  $P(T = t_i | T \geq t_i)$ , welche die Hazardrate bilden, allerdings unbekannt sein. Daher ist die Formulierung einer Hypothese zum Verlauf des Übergangsrisikos einer der zentralen Arbeitsschritte in der Ereignisanalyse. Ein semiparametrisches Verfahren, das ohne eine solche Festlegung auskommt, stellt dagegen die sogenannte Cox-Regression dar (Cox 1972), welche die Hazardrate wie folgt spezifiziert, wobei  $\alpha(t)$  das Basisrisiko (base line hazard) darstellt:

$$\lambda_i(t, x) = \alpha(t) \exp(x_i' \beta) \quad (6.9)$$

Einzige Einschränkung der Hazardrate im Cox-Modell ist die Annahme, dass die Kovariaten proportional auf sie einwirken (proportional hazard assumption).

In einigen Forschungszusammenhängen hat man es allerdings nicht mit einer stetig, sondern diskret gemessenen Zeit zu tun, sei es weil Messungen auf Basis von diskreten Zeiteinheiten wie Jahren, Monaten oder Wochen erfolgen oder weil

bestimmte Ereignisse empirisch meist zu bestimmten Zeitpunkten eintreten. In Modellen für stetige Zeit würden die dadurch entstehenden Bindungen zu Schwierigkeiten führen. Es existiert allerdings auch eine Variante der Cox-Regression, die solche Bindungen verarbeiten kann. Diese wird im Folgenden verwendet.

Damit würde jedoch auf die Vorteile des Matchingverfahrens verzichtet, vor allem was dessen Beschränkung auf die miteinander vergleichbaren Fälle betrifft. Eine Möglichkeit, diesen Vorteil dennoch zu nutzen, besteht darin, die Cox-Regression ausschließlich mit den Fällen durchzuführen, welche auch für das Matching verwendet wurden (vgl. etwa Hujer/Wellner 2000). Ein solches Vorgehen kann als Variante der Regressionsanpassung (regression adjusted matching, vgl. Rubin 1973) betrachtet werden, die den zusätzlichen Vorteil der Berücksichtigung der zeitlichen Dimension der Zielvariable besitzt.

Tabelle 11: Gesamt- und Nettoeffekt des Kombilohnes auf die Beschäftigungsdauer – Cox-Regressionen auf Basis gematchter Samples

Cox-Regressionen	Model 1: alle		Model 2: zufrieden		Model 3: unzufrieden	
	Risk Ratio	p	Risk Ratio	p	Risk Ratio	p
Kombilohnbezug	1,015	0,94	0,702	0,29	0,929	0,82
Haushaltskonstellation (R: Alleinlebend ohne Kind)						
Erwerbsloser Partner/ohne Kind	0,684	0,27	0,919	0,89	0,932	0,91
Erwerbstätiger Partner/ohne Kind	0,775	0,50	0,890	0,86	1,549	0,52
Erwerbsloser Partner/mit Kind	0,769	0,35	1,902	0,25	0,701	0,50
Erwerbstätiger Partner/mit Kind	0,804	0,52	1,151	0,82	1,284	0,68
Arbeitszeit	1,014	0,08	1,003	0,84	1,051	0,00
Wie wurde die Stelle gefunden (R: Arbeitsvermittlung)						
Bekannte/Freunde	0,669	0,17	0,533	0,20	1,041	0,93
Eigene Initiative	1,035	0,90	1,191	0,71	1,538	0,34
Sonstiges	0,749	0,30	0,723	0,52	0,577	0,35
Vorher Weiterbildung ja	0,375	0,34	0,775	0,83		
Lohnersatzleistungen vom Arbeitsamt (R: Arbeitslosengeld)						
Arbeitslosenhilfe	0,947	0,83	0,544	0,21	0,949	0,90
Keine Leistungen vom Arbeitsamt	0,980	0,94	0,588	0,36	1,101	0,83
Sozialhilfebezug	1,137	0,68	2,048	0,24	0,863	0,77
Beschäftigung befristet	1,317	0,17	1,481	0,28	0,733	0,38
Branche (R: Leih-/Zeitarbeitsfirma)						
Reinigungsgewerbe	0,397	0,02	0,706	0,62	0,557	0,34
Hotel- und Gaststättengewerbe	0,617	0,16	2,781	0,09	0,437	0,17

Cox-Regressions	Model 1: alle		Model 2: zufrieden		Model 3: unzufrieden	
	Risk Ratio	p	Risk Ratio	p	Risk Ratio	p
Callcenter	0,395	0,38	1,279	0,81	0,347	0,41
Energie und Wasser; Bergbau, verarbeitendes Gewerbe	0,452	0,02	0,199	0,02	0,695	0,51
Handel	0,304	0,00	0,152	0,01	0,293	0,06
Verkehr/Nachrichtenübermittlung, Banken, Versicherungen	0,328	0,09	0,142	0,05		
Öffentlicher Dienst/ Sozialversicherung	0,405	0,07			0,289	0,14
Bereich andere Dienstleistungen, k.A.	0,442	0,01	0,386	0,06	0,293	0,02
Einkommen wichtig	1,090	0,67	1,460	0,30	0,716	0,33
Suchintensität (Zahl der genutzten Suchwege)	1,022	0,71	1,018	0,86	0,940	0,54
Bereits ein Stellenangebot vom AA wegen zu niedrigem Einkommen abgelehnt?	0,645	0,24	0,771	0,63	0,520	0,35
Bisherige Arbeitslosigkeitsdauer insgesamt (R: < 6 Monate)						
6–12 Monate	1,252	0,61	4,631	0,07	1,384	0,63
12–24 Monate	1,484	0,29	4,348	0,05	2,046	0,21
> 24 Monate/k.A.	1,766	0,08	4,341	0,04	1,663	0,33
Art der Kündigung der letzten Beschäftigung (R: Sonstiges)						
Eigenkündigung	1,141	0,74	2,369	0,18	1,052	0,95
Arbeitgeberkündigung	1,084	0,72	1,558	0,24	0,658	0,28
Arbeitsamttypen (R: Typ 1)						
Typ 2	0,779	0,53	0,261	0,08	0,507	0,34
Typ 3	1,154	0,59	1,236	0,63	1,132	0,82
Typ 4	0,586	0,50			0,695	0,75
Typ 5	1,353	0,31	0,854	0,77	2,239	0,15
Einkommen aus Nebentätigkeit	2,499	0,02	0,877	0,86	6,332	0,02
Haushaltseinkommen	1,000	0,59	0,999	0,04	0,999	0,10
Stundenlohn (R: bis 4,60 Euro)						
4,61 bis 6,90 Euro	0,742	0,30	0,284	0,02	1,071	0,89
6,91 bis 9,20 Euro	0,563	0,11	0,583	0,37	0,930	0,91
9,21 Euro oder mehr	0,847	0,66	0,296	0,10	3,840	0,04
Keine (plausible) Angabe	0,966	0,92	0,352	0,09	1,323	0,64
Fälle	363		176		177	
Log likelihood	-633,360		-228,014		-251,711	

Anmerkung: Werte der Risk Ratio unter 1 verweisen auf einen positiven Einfluss auf die Beschäftigungsstabilität, Werte über 1 auf einen negativen Einfluss. Fehlende Koeffizienten verweisen darauf, dass hier die Fallzahl im gematchten Sampel zu gering war. Die entsprechenden Fälle wurden für die Schätzung aus der Stichprobe entfernt.

Modell 1 in Tabelle 11 zeigt, dass in einem Matching mit anschließender ereignisanalytischer Regressionsanpassung der Effekt des Kombilohnes auf die Beschäftigungsstabilität mit einer Steigerung der Hazardrate um 1,5 % im Vergleich zu Ungeförderten zwar leicht negativ ausfällt, jedoch nicht signifikant ist. Damit muss die Hypothese 2, kombilohngeförderte Beschäftigungsverhältnisse endeten früher als reguläre Beschäftigungsverhältnisse, abgelehnt werden.<sup>67</sup> Allerdings, es zeigt sich vom höheren Einkommen im Kombilohnfall auch kein positiver Effekt auf den Verbleib in Beschäftigung, wie er angesichts der höheren Opportunitätskosten in Kombilohnbeschäftigungen zu erwarten gewesen wäre. Insofern ist es möglich, dass der Effekt der subjektiven Reziprozitätsverletzung durch den Kombilohnzuschuss zwar nicht ausreicht, um einen negativen Effekt auf die Beschäftigungsdauer hervorzurufen, sondern lediglich einen positiven Effekt des höheren Einkommens unterdrückt. Um dies zu untersuchen, wird im Folgenden eine im Hinblick auf die eben erzielten Ergebnisse angepasste Hypothese 3 getestet.

**Hypothese 3 (angepasst):** *Die fehlende Wirkung des Kombilohnes auf die Stabilität des Beschäftigungsverhältnisses lässt sich durch dessen negative Wirkung auf die subjektive Wertschätzung des Lohnes erklären.*

Hypothese 3 kann analog zur klassischen Drittvariablenkontrolle (Lazarsfeld 1976; Nowak 1960) untersucht werden. Der (fehlende) Zusammenhang zwischen zwei Variablen kann demnach als erklärt gelten, wenn er durch die Kontrolle einer Drittvariable verschwindet (auftaucht). Hier wäre also zu untersuchen, ob der fehlende Zusammenhang zwischen Kombilohn und Beschäftigungsstabilität durch die Kontrolle der Lohnzufriedenheit zutage tritt. Der einfachste Fall einer solchen Kontrolle der Lohnzufriedenheit ist die Aufspaltung der Stichprobe in lohnzufriedene und lohnunzufriedene Beschäftigte, wobei die Kausalanalyse in jeder Unterstichprobe erneut durchgeführt wird. Allerdings stellt dieses Vorgehen im Fall des Matchings bzw. im Rahmen von Rubins Kausalmodell spezifische Voraussetzungen, vor allem die bedingte Unabhängigkeit nicht nur von Kombilohn und Lohnzufriedenheit, sondern auch von Kombilohn und Beschäftigungsstabilität (Rosenbaum 1984a). Ist diese Voraussetzung gegeben, kann die regressionsangepasste Matchinganalyse für die durch die Lohnzufriedenheit gebildeten Teilgruppen getrennt durchgeführt werden (Tabelle 11, Modell 2 und 3). Aufgrund der ohnehin geringen Fallzahl wurde

---

67 Dieses Ergebnis widerspricht den Ergebnissen der Evaluationsforschung zum Mainzer Modell, wo zumindest für Westdeutschland ein negativer Effekt des Kombilohnes auf die Beschäftigungsstabilität gemessen wurde (Hartmann/Krug 2004). Der Unterschied lässt sich vermutlich dadurch erklären, dass das hier durchgeführte Matching nur vergleichbare Fälle verwendet.

mit  $c = 0,01$  ein größerer Caliper gewählt, der allerdings noch immer weit unter der oben genannten Schwelle von 0,048 liegt.

Betrachtet man die Risk Ratios des Kombilohnbezuges in Modell 2 und 3, dann zeigt sich die in Hypothese 3 (angepasst) erwartete Richtung. Eliminiert man den Einfluss des Kombilohnes auf die Lohnzufriedenheit, dann zeigt sich ein positiver Effekt des in geförderten Beschäftigten höheren Einkommens auf die Beschäftigungsstabilität. In der Gruppe der Lohnzufriedenen verringert sich das Risiko des Beschäftigungsendes durch den Kombilohn um 30 % und in der Gruppe der Lohnunzufriedenen immerhin noch um 7 %. In diesem Sinne erklärt der negative Nebeneffekt des Kombilohnes auf die subjektive Wahrnehmung der Reziprozität von Lohn und Arbeitsleistung tatsächlich den ausbleibenden positiven Effekt des Kombilohnes auf die dauerhafte Integration in Beschäftigung. Allerdings weisen die Ergebnisse zwar in die durch Hypothese 3 (angepasst) implizierte Richtung, sind jedoch nicht statistisch signifikant. Daher kann die Frage, ob die negative Auswirkung des Kombilohnes auf die subjektive Wertschätzung des Lohnes zu Konsequenzen im Verhalten der Geförderten führt, welche die Integrationswirkung der Maßnahme konterkarieren, nicht eindeutig beantwortet werden. Die Daten lassen damit auch eine alternative Interpretation zu, nämlich dass der Kombilohn zwar die Lohnzufriedenheit trotz der Einkommenserhöhung negativ beeinflusst, dass dies jedoch keinerlei Einfluss auf die Beschäftigungsstabilität in geförderten Beschäftigungsverhältnissen hat.

## 6.5 Prüfung zentraler Verfahrensannahmen

In den meisten Fällen, in denen das RKM oder seine ökonomische Entsprechung im HMA verwendet wird, wird auch auf die Einschränkung hingewiesen, dass die Messung des kausalen Effektes nur dann Gültigkeit besitzt, wenn die zentralen Verfahrensannahmen CIA und SUTVA erfüllt sind. Bisweilen gewinnt man als Leser solcher Arbeiten den Eindruck, dass dabei ein zentrales Problem, nämlich ob der ausgewiesene Effekt tatsächlich ein kausaler Effekt ist oder nur ein bedingter Mittelwertunterschied ohne kausale Interpretierbarkeit, etwas zu schnell (oft nur in einem Satz) abgehandelt wird. Allerdings existiert eine Reihe von Möglichkeiten, diese Annahmen zumindest auf grobe Widersprüche zu den verfügbaren Daten zu prüfen. Damit lassen sich beide Annahmen zwar nicht beweisen, sie können in manchen Fällen aber widerlegt werden und helfen in jedem Fall dem Leser, der (nicht nur wenn es sich um eine Arbeit im Rahmen der Politikberatung handelt) oft mit den statistischen Details des RKM nicht umfassend vertraut ist, die Validität der Ergebnisse besser einzuschätzen. Im Folgenden sollen einige der in der Fachliteratur vorgeschlagenen Tests der Annahmen im RKM durchgeführt werden, so-

fern es die Datenlage erlaubt. Dabei erfolgt größtenteils eine Konzentration auf die Hypothese 1, da es sich dabei um eine reine Matchinganalyse handelt, während bei der Prüfung der Hypothesen 2 und 3 eine Regressionsanpassung erfolgte. Dadurch sind Tests wie z. B. die Rosenbaum-Bounds nicht mehr aussagekräftig.

### 6.5.1 Die bedingte Unabhängigkeit des Kombilohnzuganges (CIA)

Grundsätzlich ist die Annahme der bedingten Unabhängigkeit nicht zu beweisen und muss sich stets auf eine umfassende Analyse des Gegenstands der Forschung stützen. Je gründlicher diese Analysen sind, desto plausibler ist die CIA. Ein Aspekt dieser Analysen wurde bereits mit der Untersuchung der Maßnahmeselektivität behandelt. Die Analysen der Maßnahmeselektivität (Kapitel 5) verweisen z. B. darauf, dass die Betriebsgröße solch ein Einflussfaktor sein könnte. Im Folgenden wird versucht, mithilfe einiger selten genutzter Tests die Annahme der bedingten Unabhängigkeit weiter zu stützen. Dazu gibt es mehrere Möglichkeiten, von denen hier zwei genutzt werden sollen. Die erste Möglichkeit ist weniger ein Test der CIA als vielmehr eine Sensitivitätsanalyse der Kausalaussage gegenüber dem Auslassen von Variablen. Die zweite Möglichkeit nutzt die Tatsache, dass aus den theoretischen Überlegungen zum Wirkmechanismus Aussagen abgeleitet werden können, die zur Überprüfung der CIA herangezogen werden können.

#### 6.5.1.1 Rosenbaum-Bounds

Die Gültigkeit der CIA und damit auch die Gültigkeit der Ergebnisse zur Schätzung des kausalen Effektes ist infrage gestellt, wenn eine kausalrelevante Variable in der logistischen Regression nicht enthalten war. Trotz aller Gründlichkeit der Analyse lassen sich nie alle Zweifel darüber ausräumen, ob nicht doch ein kausalrelevanter Einflussfaktor übersehen wurde. Weiter oben wurde bereits ausgeführt, dass die Betriebsgröße nach den Erkenntnissen der Implementationsforschung eventuell Einfluss auf die Maßnahmeselektivität im Mainzer Modell hatte. Dies wird dann zum Problem, wenn es in großen (kleinen) Betrieben häufiger (seltener) zu subjektiv empfundenen Verletzungen der Reziprozität kommt. Auch wenn aus der Theorie kein solcher Einfluss der Betriebsgröße bekannt ist, soll im Folgenden untersucht werden, welche Auswirkungen das Fehlen der Betriebsgröße oder einer anderen einflussreichen Variablen bei der Konstruktion des PS auf die Messung des kausalen Effektes haben würde. Könnte das Fehlen den beobachteten Effekt erklären und ihn damit seiner kausalen Interpretation berauben? Um dies zu überprüfen, wird ein von Rosenbaum (2002) entwickelter und von Becker und Caliendo (2006) in Stata implementierter Test durchgeführt. Der Test besteht darin, unter der Annahme eines „worst case“ (Gangl/DiPrete 2004b: 291), dass also die Betriebsgröße sowohl

die Maßnahmeselektivität als auch die Zielvariable beeinflusst, anzugeben, wie groß dieser Einfluss mindestens sein muss, damit der in Abschnitt 6.4 geschätzte kausale Effekt von -14 Prozentpunkten nicht mehr von null signifikant verschieden ist. Dabei handelt es sich um eine Sensitivitätsanalyse der Kausalanalyse, die diese weder bestätigen noch widerlegen kann, sondern lediglich die Robustheit des Kausalschlusses in Bezug auf die Größe des Einflusses eines ausgelassenen, aber relevanten Einflussfaktors angibt.

Die Grundüberlegung, auf der dieser Test basiert, ist folgende: Folgt man den Überlegungen der Implementationsforschung, dann ist die Wahrscheinlichkeit des Erhalts einer Kombilohnförderung für das Individuum  $i$  nicht wie oben unterstellt

$$P(x_i) = E(kl_i = 1 | \mathbf{x}_i) = \left(1 + \exp(-\mathbf{x}'_i \boldsymbol{\beta})\right)^{-1} \quad (6.10)$$

sondern zusätzlich besteht noch der Einfluss der Variable Betriebsgröße  $bg$ , sodass

$$P(x_i) = E(kl_i = 1 | \mathbf{x}_i) = \left(1 + \exp(-(\mathbf{x}'_i \boldsymbol{\beta} + \gamma bg_i))\right)^{-1} \quad (6.11)$$

mit  $\gamma \neq 0$ , wenn dieser Einfluss der Betriebsgröße tatsächlich besteht, und  $\gamma = 0$ , wenn die Betriebsgröße keinen Einfluss auf den Kombilohnerhalt hatte.

Die Chancen (Odds) dafür, dass Individuum  $i$  eine Kombilohnförderung erhält, sind gegeben durch das Verhältnis der Wahrscheinlichkeit, einen Kombilohn zu erhalten, und der Gegenwahrscheinlichkeit, keinen Kombilohn zu erhalten  $P(\mathbf{x}_i) / (1 - P(\mathbf{x}_i))$ . Entsprechendes gilt für Person  $j$ , die als Matchingpartner für  $i$  aufgrund gleicher Kovariaten  $\mathbf{x}$  ausgesucht wurde:<sup>68</sup>  $P(\mathbf{x}_j) / (1 - P(\mathbf{x}_j))$ . Die Odds Ratio ist dann (Becker/Caliendo 2006):

$$\frac{P(\mathbf{x}_i) / (1 - P(\mathbf{x}_i))}{P(\mathbf{x}_j) / (1 - P(\mathbf{x}_j))} = \frac{P(\mathbf{x}_i) (1 - P(\mathbf{x}_j))}{P(\mathbf{x}_j) (1 - P(\mathbf{x}_i))} = \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + \gamma bg_i)}{\exp(\mathbf{x}'_j \boldsymbol{\beta} + \gamma bg_j)} = \exp(\gamma (bg_i - bg_j)) \quad (6.12)$$

Person  $j$  wird dann ein adäquater Matchingpartner für  $i$  sein, wenn nicht nur eine Gleichheit bezüglich der Kovariaten  $\mathbf{x}$  besteht, sondern entweder wenn die Betriebsgröße bei beiden Beschäftigten „zufällig“ gleich ausfällt ( $bg_i - bg_j = 0$ ) oder wenn  $\gamma = 0$  ist, das heißt, die Betriebsgröße für die Kombilohnwahrscheinlichkeit irrelevant ist.

<sup>68</sup> Der Einfachheit halber wird von exaktem Matching ausgegangen.

Geht man von der Betriebsgröße als einer Dummyvariable mit den Werten 0 und 1 aus, dann lässt sich angeben, dass die Odds Ratio von Person  $i$  und  $j$  zwischen folgenden Werten (Bounds) liegen muss (Rosenbaum 2002: 108):

$$\frac{1}{e^\gamma} \leq \exp(\gamma(bg_i - bg_j)) \leq e^\gamma \quad (6.13)$$

Dies bedeutet, dass die beiden Odds der Individuen  $i$  und  $j$  gleich sind und damit die Odds Ratio 1, wenn  $\gamma = 0$  bzw.  $e^\gamma = 1$  ist. Ist etwa  $e^\gamma = 2$ , dann bedeutet dies, dass Personen, die auf Basis eines Propensity Score Matchings mit den Variablen  $\mathbf{x}$  als ähnlich erschienen und als zwei Personen mit gleicher Kombilohnwahrscheinlichkeit ausgewählt wurden, in Wirklichkeit bezüglich der Betriebsgröße unterschiedlich waren und deswegen Person  $i$  eigentlich eine doppelt so hohe Teilnahmewahrscheinlichkeit besaß wie Person  $j$ . Da in einem solchen Fall Personen mit unterschiedlichen Teilnahmewahrscheinlichkeiten aneinander gematcht wurden, ist ein gemessener Effekt des Treatments kein kausaler Effekt, sondern geht auf einen unbeobachteten Selektionseinfluss zurück. Auf Basis der Teststatistik  $Q_{MH}$  (Mantel/Haenszel 1959; Becker/Caliendo 2007) lässt sich nun testen, wie groß  $e^\gamma$  mindestens sein muss, damit der Effekt auf die Lohnzufriedenheit (Abschnitt 6.4, Tabelle 8) nicht mehr dem Kombilohn zugerechnet werden kann, sondern der unbeobachteten Variable Betriebsgröße. Dabei ist  $Q_{MH}^+$  die Teststatistik bei Überschätzung des kausalen Effektes und  $Q_{MH}^-$  bei Unterschätzung und  $p_{MH}^+$  bzw.  $p_{MH}^-$  die entsprechenden p-Werte für den kausalen Effekt.

Die Berechnung der Werte der Teststatistiken und der zugehörigen p-Werte für verschiedene Odds Ratios erfolgt mit dem Stata Supplement „mhbounds“ (Becker/Caliendo 2006). Die Ergebnisse sind in Tabelle 12 angegeben. Im Fall des positiven kausalen Effektes auf die Lohnzufriedenheit ist vor allem die Möglichkeit der Überschätzung des wahren kausalen Effektes relevant (Spalten  $Q\_mh+$  bzw.  $p\_mh+$ ). Besteht entweder kein Einfluss der Betriebsgröße auf die Förderwahrscheinlichkeit oder besteht dieser, wirkt sich aber nicht auf die Lohnzufriedenheit aus, so ist auch hier der kausale Effekt mit einem p-Wert von 0,005 signifikant von null verschieden. Geht man mit der Implementationsforschung davon aus, dass die Betriebsgröße eventuell bei der Maßnahmeselektivität eine Rolle gespielt hat (Abschnitt 5.2.4.3), und geht man weiter davon aus, dass bei einem Betriebsgrößendummy 1 für große und 0 für kleine Betriebe steht, dann lassen sich die Werte der Tabelle 12 wie folgt interpretieren.

Sofern sich die Chance für eine Förderung gegenüber der Chance einer Nichtförderung durch die Betriebsgröße um 25 % (eine Odds Ratio von 1,25) erhöht, bleibt der kausale Effekt auf dem Niveau von 0,05 signifikant, bei einer Erhöhung um 50 % bleibt er auf einem Niveau von 0,1 signifikant. Aufgrund der geringen



Fallzahl (186 gematchte Paare) erscheint es gerechtfertigt, erst dann den obigen Effekt infrage zu stellen, wenn die Betriebsgröße zu einer Veränderung der Odds der Geförderten im Vergleich zu den ungeförderten um mehr als das 1,40-Fache führt. Damit zeigt sich das Ergebnis der Messung des kausalen Effektes relativ robust gegenüber ausgelassenen Variablen.

Tabelle 12: Rosenbaum-Bounds

$e^{\gamma}$	Q_mh+	Q_mh-	p_mh+	p_mh-
1	2,584	2,584	0,005	0,005
1,05	2,769	2,411	0,003	0,008
1,1	2,940	2,241	0,002	0,013
1,15	3,104	2,078	0,001	0,019
1,2	3,261	1,923	0,001	0,027
1,25	3,413	1,775	0,000	0,038
1,3	3,558	1,632	0,000	0,051
1,35	3,699	1,495	0,000	0,067
1,4	3,835	1,363	0,000	0,086
1,45	3,966	1,236	0,000	0,108
1,5	4,093	1,113	0,000	0,133
.				
.				
.				
3	6,771	1,110	0,000	0,134
3,1	6,902	1,228	0,000	0,110
3,2	7,029	1,343	0,000	0,090
3,3	7,153	1,454	0,000	0,073
3,4	7,274	1,562	0,000	0,059
3,5	7,391	1,667	0,000	0,048

Q\_mh+: Mantel-Haenszel Statistik unter der Annahme der Überschätzung des kausalen Effektes  
Q\_mh-: Mantel-Haenszel Statistik unter der Annahme der Unterschätzung des kausalen Effektes  
p\_mh+: Signifikanzniveau unter der Annahme der Überschätzung des kausalen Effektes  
p\_mh-: Signifikanzniveau unter der Annahme der Unterschätzung des kausalen Effektes

Zur Interpretation der Ergebnisse ist wichtig, dass es sich dabei um eine Analyse unter der strengen Annahme handelt, dass die Betriebsgröße für jedes gematchte Paar in der Analyse auch zu einer höheren Wahrscheinlichkeit der Reziprozitätsverletzung führt (vgl. Gangl/DiPrete 2004b). Positiv ausgedrückt ist der gemessene Effekt also selbst dann noch als ein negativer kausaler Effekt zu interpretieren, wenn zwar der Effekt der Betriebsgröße so stark ausfällt, dass für jedes Paar von Geförderten und hinzugematchten Ungeförderten gilt, dass die Geförderten eher eine subjektive Verletzung der Lohnreziprozität erfahren und gleichzeitig die Be-

triebsgröße die Odds Ratio der Förderung um mehr als das 1,40-Fache erhöht. Ein signifikant positiver Effekt des Lohnzuschusses im Kombilohn ergibt sich im Übrigen erst ab der Odds Ratio von 3,1 für einen ausgelassenen Selektionseinfluss (Tabelle 12).

Die Rosenbaum-Bounds liefern zwar zusätzliche Informationen, können aber die Unsicherheit nie beseitigen, solange die Größe des Einflusses der Betriebsgröße nicht bekannt ist. Zwar sind in der verwendeten Befragung keine Informationen zur Betriebsgröße vorhanden, doch existiert mit der IEB (Integrierte Erwerbsbiografien; vgl. Hummel et al. 2005) eine Datenquelle, in der die Betriebsinformation vorhanden ist<sup>69</sup> und aus der somit die Odds Ratio der Betriebsgröße<sup>70</sup> auf die Maßnahmeteilnahme empirisch geschätzt werden kann. Für die Wahrscheinlichkeit des Erhalts einer Kombilohnförderung im Mainzer Modell ergibt sich für die Betriebsgröße unter Kontrolle der sonstigen in der IEB verfügbaren Selektionseinflüsse eine Odds Ratio von 1,06 mit einem Konfidenzintervall auf dem 95 %-Niveau 0,96 bis 1,17. Damit stellt das Fehlen der Betriebsgröße wohl kein gravierendes Problem für die Validität der Kausalanalyse dar.

#### 6.5.1.2 Nutzung alternativer Zielvariablen

Neben einer Sensitivitätsanalyse schlägt Rosenbaum (1984b) auch vor, theoretisches Wissen über die Hervorbringung des kausalen Effektes durch das Treatment (kausaler Mechanismus) dazu zu nutzen, Tests der CIA zu konstruieren. Ausgang der Überlegungen ist, dass ohne Erfüllung der CIA im Propensity Score Matching kein Ausgleich der Maßnahmeselektivität erfolgt, sodass eine verzerrte Messung des kausalen Effektes durch Vergleich der unbedingten Erwartungswerte erfolgt:

$$MKE \neq E(y_i^1) - E(y_i^0) \quad (6.14)$$

Für die Zielvariable einer Untersuchung ist der wahre kausale Effekt nicht bekannt, sondern soll aus den Daten geschätzt werden. Daher kann auch nicht beurteilt werden, ob die CIA gültig ist. Wäre der kausale Effekt bereits bekannt, dann könnte anhand der Übereinstimmung des bekannten mit dem im PS-Matching gemessenen Effekt die CIA als gesichert gelten. Dies ist etwa der Fall, wenn neben der Beobachtungsstudie zusätzlich noch ein soziales Experiment durchgeführt wurde (z. B. Dehejia/Wahba 2002).<sup>71</sup> Dann kann der Effekt aus dem Zufallsexperiment als

69 Allerdings fehlen in dieser Datenquelle dann andere wesentliche Variablen, so z. B. auch die Lohnzufriedenheit.

70 Aus der Variable „Anzahl der Beschäftigten“ wird ein Dummy gebildet, der für Kleinbetriebe den Wert 1 annimmt. In Anlehnung an die Definition der EU (ABl. der EU L124/36) werden als Kleinbetriebe solche mit unter 10 Beschäftigten betrachtet.

71 Auch hier kann allerdings noch Unsicherheit bezüglich der internen Validität des Experimentes bestehen, vgl. Abschnitt 4.1.1.

Maßstab für den Effekt aus der Beobachtungsstudie verwendet werden. Gewöhnlich wird eine Beobachtungsstudie aber dann durchgeführt, wenn ein randomisiertes Experiment nicht möglich oder zu fehleranfällig ist. Rosenbaum (1984b) schlägt verschiedene Möglichkeiten vor, die Gültigkeit der CIA auf indirekte Weise zu überprüfen, durch die Verwendung alternativer Zielvariablen, alternative Untersuchungspopulationen oder Varianten des Treatments. Einige der von Rosenbaum vorgeschlagenen Tests werden hier im Folgenden durchgeführt. Diese Tests sind darauf ausgelegt, die CIA zu falsifizieren, und daher ist klar, dass ein Bestehen solcher Tests nur Argumente für die Plausibilität der CIA liefern kann, sie dadurch aber nicht bestätigt wird.

Ein Plausibilitätstest der CIA ist etwa möglich, wenn die Wirkungen eines Treatments auf eine bestimmte, von der Zielvariable verschiedene Variable unter Erfüllung der CIA durch den Variablenvektor  $\mathbf{x}$  bekannt oder relativ gesichert ist (im Folgenden als  $\delta'$  bezeichnet). Für Ursachen gilt generell, dass bei Gültigkeit der CIA (und bei Ausschluss rückwirkender Kausalität) für alle Messungen der Zielvariable vor Einsetzen der Ursache, also zum Zeitpunkt  $t < 0$ , das Ergebnis  $\delta' = 0$  zu erwarten ist (Imbens 2004: 22). Ebenso sollte dies für andere, zeitlich vor dem Treatmentinsatz gemessene Variablen gelten, da das Treatment (und unter Ausschluss rückwirkender Kausalität bzw. von Antizipationseffekten) nur Faktoren beeinflussen kann, die nach seinem Einsetzen gemessen werden. In diesem Sinne wäre die Messung des kausalen Effektes des Kombilohnes auf die Lohnzufriedenheit vor Kombilohnbezug ein geeigneter Test. Eine solche Messung muss sich dann natürlich auf die letzte Beschäftigung beziehen. Leider ist diese Information in den Daten nicht vorhanden.<sup>72</sup>

Abhängig vom konkreten Forschungsgegenstand und dem theoretischen Mechanismus, welcher der Wirkung des Treatments auf die Zielvariable zugrunde liegt, lassen sich eventuell auch für die Zeitpunkte  $t > 0$  bestimmte alternative Zielvariablen angeben, für die aus der Theorie ein bestimmter Wert von  $\delta'$  folgt. So könnte die Wirkung des Treatments auf bestimmte Zielvariablen etwa größer sein als auf andere, oder es sollte für andere Zielvariablen keine Wirkung nachzuweisen sein. Eine speziell im vorliegenden Anwendungsfall bestehende Möglichkeit ist es, den kausalen Effekt auf eine nicht bei der Konstruktion der Propensity Scores berücksichtigte Variable zu berechnen, von der aber trotzdem ein Effekt von  $\delta' = 0$  zu erwarten ist. Da sowohl der Stundenlohn als auch die Arbeitszeit berücksichtigt wurden, wäre zu erwarten, dass sich auch bei der nicht berücksichtigten Variable *Monatseinkommen*, wenn die Ausbalancierung der Stichproben von Geförderten und Ungeförderten geglückt ist, kein signifikanter Unterschied ergibt.

72 Daran scheitert auch die Anwendung des Difference-in-Differences-Matchings.

Tabelle 13: „Kausaler Effekt“ des Kombilohnes auf das individuelle Bruttoeinkommen

	Individuelles Bruttoeinkommen		Effekt	Standardfehler	t-Wert
	Kombilohn	Regulärer Lohn			
<i>Ein nächster Nachbar; Propensity Score (ohne Zurücklegen); Caliper 0,005</i>					
<i>Vor dem Matching</i>	885,07	1.249,15	-364,08	42,15	-8,64
<i>Nach dem Matching</i>	1.025,10	1.008,63	16,47	62,67	0,26
<i>Zahl der Matches</i>	136				
Anmerkung: Fallzahlverlust wegen fehlender Einkommensangaben.					

Das Ergebnis der Berechnungen zeigt Tabelle 13. Der Mittelwertunterschied vor dem Matching zeigt, dass Geförderte sich erwartungsgemäß in eher geringer entlohnten Tätigkeiten befinden als Ungeförderte. Im Mittel zeigt sich ein signifikanter Unterschied von ca. 364 Euro. Nach dem Matching besteht mit einer Differenz im mittleren Einkommen von 16,5 Euro bei einem t-Wert von 0,26 kaum mehr ein Unterschied zwischen beiden Teilgruppen. Dies spricht zumindest nicht gegen die Gültigkeit der CIA.

Besonders geeignet für den Test der CIA sind auch alternative Zielvariablen, die der jeweils untersuchten Zielvariable möglichst ähnlich sind, für die aber bei Gültigkeit der CIA kein kausaler Effekt des Kombilohnes zu erwarten ist. Da die betrachtete Ursachensvariable Kombilohn in einem Lohnzuschuss besteht und eine der zentralen Zielvariablen die Lohnzufriedenheit, bieten sich andere auf die Beschäftigung bezogene Bereichszufriedenheiten an. Gelingt es dem Matching, alle relevanten Unterschiede zwischen Kombilohngeförderten und regulär Beschäftigten, insbesondere die Arbeitsbedingungen, auszugleichen, dann sollte sich weder ein Einfluss der Förderung auf die Zufriedenheit mit der Arbeit insgesamt noch mit der Tätigkeit finden lassen. Wenn sich der einzige Unterschied auf den Lohn bezieht, so wäre es ein Hinweis auf eine verletzte CIA, wenn durch den Kombilohn nicht nur die Lohnzufriedenheit, sondern auch die Zufriedenheit mit anderen Bereichen der Beschäftigung beeinflusst würde, etwa weil in die Kombilohnförderung tendenziell eher unzufriedene Personen selektiert wurden. Positive und negative Aspekte einer Beschäftigung (z. B. sympathische Kollegen und Vorgesetzte oder ein interessanter Aufgabenbereich) sollten bei Gültigkeit der CIA jedoch zufällig auf Geförderte und Ungeförderte verteilt sein, sodass nach dem Matchen kein Unterschied zu beobachten sein sollte. In der Befragung zum Mainzer Modell wurde neben der Lohnzufriedenheit noch die Zufriedenheit mit der Art der Tätigkeit und die Zufriedenheit mit der Arbeitsstelle insgesamt erhoben. Im Fall der Zufriedenheit mit der Tätigkeit (Tabelle 14) zeigt sich ein negativer Effekt vor dem Matching von ca. 5 %, der allerdings nicht signifikant ist. Dieser Effekt bleibt auch nach dem Matching insignifikant.

Tabelle 14: Kausaler Effekt des Kombilohnes auf die Zufriedenheit mit der Tätigkeit

	Zufrieden mit der Tätigkeit (Anteil)		Effekt	Standard- fehler	t-Wert
	Kombilohn	Regulärer Lohn			
<i>Ein nächster Nachbar; Propensity Score (ohne Zurücklegen); Caliper 0,005</i>					
<i>Vor dem Matching</i>	0,773	0,821	-0,047	0,026	-1,85
<i>Nach dem Matching</i>	0,770	0,818	-0,048	0,042	-1,15
<i>Zahl der Matches</i>	187				

Ähnlich verhält es sich auch mit dem kausalen Effekt des Kombilohnbezuges auf die Zufriedenheit mit der Arbeitsstelle insgesamt (Tabelle 15). Auch hier lässt sich vor und vor allem auch nach dem Matchen kein Unterschied zwischen Geförderten und Ungeförderten nachweisen. Damit stützen die Ergebnisse die Plausibilität der CIA.

Tabelle 15: Kausaler Effekt des Kombilohnes auf die Zufriedenheit mit der Arbeitsstelle insgesamt

	Zufrieden insgesamt (Anteil)		Effekt	Standard- fehler	t-Wert
	Kombilohn	Regulärer Lohn			
<i>Ein nächster Nachbar; Propensity Score (ohne Zurücklegen); Caliper 0,005</i>					
<i>Prima facie</i>	0,789	0,804	-0,015	0,026	-0,58
<i>Nach Matching</i>	0,786	0,802	-0,016	0,042	-0,38
<i>Zahl der Matches</i>	187				

Eine weitere Möglichkeit zur Plausibilitätsprüfung der CIA ist nach Rosenbaum (1984b) die Verwendung einer anderen Population, für die ebenfalls der kausale Effekt bekannt oder in ausreichendem Maße aufgrund der theoretischen Überlegungen gesichert ist. So kann etwa der Effekt für Population 2 bekanntermaßen größer sein als der für die betrachtete Population 1 ( $\delta_2 > \delta_1$ ), oder es kann für eine Population bekannt sein, dass das Treatment abgesehen von Selektionseffekten keinerlei Wirkung auf die Zielvariable hat ( $\delta_2 = 0$ ).

Tabelle 16: Kausaler Effekt des Kombilohnes bei Austausch der Treatmentpopulation

	Zufrieden mit dem Lohn (Anteil)		Effekt	Standard- fehler	t-Wert
	Kombilohn	Regulärer Lohn			
<i>Ein nächster Nachbar; Propensity Score (ohne Zurücklegen); Caliper 0,005</i>					
<i>Prima facie</i>	0,396	0,589	-0,193	0,030	-6,41
<i>Nach Matching</i>	0,383	0,572	-0,189	0,044	-4,29
<i>Zahl der Matches</i>	248				

In den vorliegenden Daten lässt sich keine alternative Population identifizieren, für die ein kausaler Effekt bekannt wäre. Es kann sich allerdings trotzdem lohnen, die betrachteten Populationen zu variieren, auch wenn die Auswirkung unbekannt ist. So kann statt der Kombilohngefährdeten, die wie die Vergleichsgruppe vor Beschäftigungsaufnahme langzeitarbeitslos und/oder geringqualifiziert waren, die bisher ausgeschlossene Teilpopulation der weder Langzeitarbeitslosen noch Geringqualifizierten als Treatmentgruppe betrachtet werden (Tabelle 16).

Auch hier zeigt sich ein kausaler Effekt des Kombilohnes, der mit 18,9 Prozentpunkten etwas größer ist als in der Population der Langzeitarbeitslosen/Geringqualifizierten. Schließlich lautet ein weiterer Vorschlag Rosenbaums zur Überprüfung der CIA zu versuchen, aus dem kausalen Mechanismus Informationen darüber abzuleiten, ob ähnliche Maßnahmen eine unterschiedliche Wirkung auf die Lohnzufriedenheit haben. In den vorliegenden Daten sind jedoch solche ähnliche Maßnahmen nicht erfasst.

#### 6.5.1.3 Verfahren mit Kontrolle unbeobachteter Selektivität

Bisher wurde im Rahmen des Matchingverfahrens davon ausgegangen, dass keine unbeobachteten kausalrelevanten Selektionseinflüsse existieren, bzw. es wurde die Sensitivität der Ergebnisse in Bezug auf das Auslassen eines kausalrelevanten Selektionseinflusses untersucht. In Abschnitt 4.3.2 wurden Verfahren vorgestellt, deren Ergebnisse nicht von der vollständigen Erfassung der  $S_{krel}$  abhängen. Wie allerdings ebenfalls ausgeführt wurde, treffen solche Verfahren dann andere Annahmen, die mindestens ebenso problematisch sind wie die CIA. Aufgrund der Komplementarität der Annahmen sind die Verfahren allerdings gut dazu geeignet, um die Sensitivität der Ergebnisse in Bezug auf die Annahme der Gültigkeit der CIA zu prüfen. So kann in der Regression mit Kontrollfunktion – unter der Annahme der Gültigkeit aller Verfahrensannahmen – das Vorliegen unbeobachteter Heterogenität getestet werden, indem die Korrelation  $\rho$  zwischen den Störgrößen der Response- und der Selektionsgleichung bestimmt wird. In der Instrumentvariablenschätzung kann – wiederum unter der Annahme der Gültigkeit aller Verfahrensannahmen – getestet werden, ob eine von der Störgröße der Selektionsgleichung konditional unabhängige Variable ähnliche Ergebnisse liefert wie das Matching, das davon ausgeht, dass auch das Treatment eine solche konditional unabhängige Variable darstellt. Ähnlich wie in den vorangegangenen Abschnitten gilt, dass erstens eine Übereinstimmung der Ergebnisse verschiedener Verfahren deren Validität nur stützen und nicht beweisen kann. Zweitens kann aus einer fehlenden Übereinstimmung nicht darauf geschlossen werden, welches der verwendeten Verfahren gültige Ergebnisse liefert.

Sowohl für die Regression mit Kontrollfunktion als auch für die Instrumentvariablenschätzung wird eine Variable benötigt, welche zwar die Selektion in die

Kombilohnmaßnahme, jedoch nicht die Zielvariable beeinflusst bzw. nicht mit in der Regression ausgelassenen Selektionseinflüssen korreliert ist. Eine solche Variable stellt im Zusammenhang mit der Kombilohnmaßnahme „Mainzer Modell“ die Arbeitszeit dar. Die Förderbedingungen im MZM waren so gestaltet, dass die Aufnahme einer kombilohngeförderten Beschäftigung gerade für Teilzeitbeschäftigte unabhängig von ihrem Stundenlohn besonders attraktiv war.<sup>73</sup> Gleichzeitig lässt sich kein Grund dafür finden, warum die Arbeitszeit, wiederum unabhängig vom Stundenlohn, die Wahrnehmung der Reziprozität von Lohn und Arbeitsleistung beeinflusst. Daher wird im Folgenden auch die Arbeitszeit als Instrumentvariable verwendet, sowohl in der Regression mit Kontrollfunktion als auch in der eigentlichen Instrumentvariablen schätzung.

Tabelle 17: Regression mit Kontrollfunktion

	Responsegleichung		Selektionsgleichung	
	Koeffizient	p	Koeffizient	p
Kombilohnbezug	-0,827	0,04		
Haushaltskonstellation (R: Alleinlebend ohne Kind)				
Erwerbsloser Partner/ohne Kind	-0,137	0,40	0,513	0,00
Erwerbstätiger Partner/ohne Kind	0,027	0,86	0,134	0,44
Erwerbsloser Partner/mit Kind	-0,006	0,97	1,011	0,00
Erwerbstätiger Partner/mit Kind	0,077	0,60	0,412	0,01
Wie wurde die Stelle gefunden (R: Arbeitsvermittlung)				
Bekannte/Freunde	0,029	0,81	-0,087	0,54
Eigene Initiative	-0,178	0,14	0,048	0,73
Sonstiges	-0,084	0,51	-0,302	0,04
Vorher Weiterbildung ja	0,006	0,98	0,528	0,07
Lohnersatzleistungen vom Arbeitsamt (R: Arbeitslosengeld)				
Arbeitslosenhilfe	0,112	0,40	0,558	0,00
Keine Leistungen vom Arbeitsamt	-0,010	0,94	-0,146	0,31
Sozialhilfebezug	0,010	0,95	0,594	0,00
Beschäftigung befristet	0,188	0,06	0,535	0,00
Branche (R: Leih-/Zeitarbeitsfirma)				
Reinigungsgewerbe	0,092	0,63	-0,091	0,66
Hotel- und Gaststättengewerbe	0,002	0,99	-0,017	0,93
Callcenter	0,518	0,07	0,784	0,04
Energie und Wasser; Bergbau, verarbeitendes Gewerbe	0,193	0,33	-0,954	0,00

73 Dies führte dadurch, dass die Förderberechtigung vom Einkommen abhing, unter anderem dazu, dass im MZM auch solche Personen einen Zuschuss erhielten, die zwar einen hohen Stundenlohn erzielten, durch eine geringe Arbeitszeit jedoch unterhalb der Förderhöchstgrenze des Einkommens blieben.

	Responsegleichung		Selektionsgleichung	
	Koeffizient	p	Koeffizient	p
Handel	0,359	0,02	-0,065	0,70
Verkehr/Nachrichtenübermittlung, Banken, Versicherungen	0,754	0,01	-0,375	0,23
Öffentlicher Dienst/Sozialversicherung	0,417	0,08	-0,947	0,00
Bereich andere Dienstleistungen, k.A.	0,283	0,05	-0,277	0,08
Einkommen wichtig	-0,147	0,09	0,027	0,79
Suchintensität (Zahl der genutzten Suchwege)	0,007	0,76	0,087	0,00
Bereits ein Stellenangebot vom AA wegen zu niedrigem Einkommen abgelehnt?	-0,001	1,00	0,287	0,08
Bisherige Arbeitslosigkeitsdauer insgesamt (R: < 6 Monate)				
6-12 Monate	0,157	0,39	-0,458	0,03
12-24 Monate	-0,282	0,08	0,047	0,79
> 24 Monate/k.A.	-0,144	0,32	0,051	0,75
Art der Kündigung der letzten Beschäftigung (R: Sonstiges)				
Eigenkündigung	0,123	0,46	0,282	0,12
Arbeitgeberkündigung	-0,061	0,55	0,276	0,02
Arbeitsamtypen (R: Typ 1)				
Typ 2	-0,121	0,45	-0,018	0,92
Typ 3	-0,123	0,32	-0,326	0,02
Typ 4	-0,245	0,49	-0,218	0,62
Typ 5	-0,214	0,14	-0,561	0,00
Einkommen aus Nebentätigkeit	0,131	0,57	-0,464	0,06
Haushaltseinkommen	0,000	0,18	0,000	0,00
Stundenlohn (R: bis 4,60 Euro)				
4,61 bis 6,90 Euro	0,204	0,15	0,069	0,69
6,91 bis 9,20 Euro	0,186	0,22	-0,118	0,50
9,21 Euro oder mehr	0,660	0,00	-0,213	0,28
Keine (plausible) Angabe	0,213	0,15	-0,154	0,39
Konstante	-0,112	0,74	0,444	0,22
Arbeitszeit (Instrument)			-0,024	0,00
$\rho$ (rho)	0,286	0,267		
Mittlerer kausaler Effekt	-0,301			
Anmerkung: Test der Signifikanz von $\rho$ auf Basis des inversen hyperbolischen Tangens; Test der Signifikanz des mittleren kausalen Effektes nicht ohne Weiteres möglich, zur Schätzung vgl. Abschnitt 4.3.2.1.				

Die Regression mit Kontrollfunktion (Tabelle 17) wurde als eine Maximum-Likelihood-Schätzung durchgeführt, wobei Response- und Selektionsgleichung als Seemingly Unrelated Regression (Weesie 1999; Greene 2003: 340 ff.) durchgeführt wurden. Der Wert des Parameters  $\rho = 0,29$  bezeichnet das Ausmaß und die Richtung der Korrelation zwischen den Störgrößen der beiden Gleichungen. Das positive Vorzeichen verweist dabei darauf, dass in der Selektionsgleichung ausgelassene Fakto-



ren (wie z. B. die Betriebsgröße oder auch andere, bisher vollkommen unbekannte Einflüsse), welche die Wahrscheinlichkeit für den Kombilohnersatz erhöhen, auch die Lohnzufriedenheit positiv beeinflussen (Positivselektion). Ein Signifikanztest dieses Parameters fungiert – im Rahmen der Modellannahmen – als Test der CI-Annahme. Es zeigt sich, dass die Annahme, dass die Störgrößen beider Gleichungen unkorreliert sind, nicht verworfen werden kann. Damit spricht auch die Regression mit Kontrollfunktion nicht gegen das in Abschnitt 6.4 erzielte Ergebnis zur Wirkung des Kombilohnes auf die Lohnzufriedenheit, wenn auch die Schätzung des kausalen Effektes mit –30 Prozentpunkten deutlich höher ausfällt. Der Unterschied geht vermutlich auf die Tatsache zurück, dass das Matching nur im Common Support erfolgt, während eine Regression alle, auch die weniger vergleichbaren Fälle verwendet.

Die Koeffizienten der Selektionsgleichung in Tabelle 17 entsprechen in Richtung und Signifikanz denen der logistischen Regression vor dem Matching in Tabelle 5.

Tabelle 18: Instrumentvariablen-schätzung

	Modell 1: Probit Kombilohnbezug		Modell 2: Lineares Wahrscheinlich- keitsmodell Kombilohnbezug		Modell 3: Lineares Wahrscheinlich- keitsmodell Lohnzufriedenheit	
	Koeff.	p	Koeff.	p	Koeff.	p
Arbeitszeit (Instrument)	-0,024	0,000				
Propensity Score (Instrument)			1,145	0,00		
<b>Instrument (mittlerer kausaler Effekt)</b>					<b>-0,283</b>	<b>0,04</b>
Haushaltskonstellation (R: Alleinlebend ohne Kind)						
Erwerbsloser Partner/ohne Kind	0,526	0,002	-0,024	0,63	-0,050	0,39
Erwerbstätiger Partner/ohne Kind	0,109	0,523	-0,004	0,92	0,007	0,90
Erwerbsloser Partner/mit Kind	0,995	0,000	-0,031	0,50	-0,006	0,92
Erwerbstätiger Partner/mit Kind	0,424	0,005	-0,020	0,63	0,027	0,63
Wie wurde die Stelle gefunden (R: Arbeitsvermittlung)						
Bekannte/Freunde	-0,107	0,453	0,009	0,82	0,010	0,82
Eigene Initiative	0,032	0,819	-0,002	0,97	-0,060	0,17
Sonstiges	-0,296	0,047	0,015	0,71	-0,026	0,58
Vorher Weiterbildung ja	0,547	0,059	-0,009	0,91	0,004	0,97
Lohnersatzleistungen vom Arbeitsamt (R: Arbeitslosengeld)						
Arbeitslosenhilfe	0,569	0,000	-0,024	0,57	0,042	0,38
Keine Leistungen vom Arbeitsamt	-0,123	0,394	0,001	0,98	0,002	0,96
Sozialhilfebezug	0,592	0,001	-0,020	0,69	0,004	0,95
Beschäftigung befristet	0,525	0,000	-0,016	0,59	0,064	0,08

## Der kausale Effekt des Kombilohnes auf Lohnzufriedenheit und Beschäftigungsstabilität

	Modell 1: Probit Kombilohnbezug		Modell 2: Lineares Wahrscheinlich- keitsmodell Kombilohnbezug		Modell 3: Lineares Wahrscheinlich- keitsmodell Lohnzufriedenheit	
	Koeff.	p	Koeff.	p	Koeff.	p
Branche (R: Leih-/Zeitarbeitsfirma)						
Reinigungsgewerbe	-0,077	0,713	-0,012	0,84	0,033	0,62
Hotel- und Gaststättengewerbe	-0,033	0,871	-0,008	0,89	-0,004	0,95
Callcenter	0,808	0,037	-0,037	0,59	0,185	0,09
Energie und Wasser; Bergbau, verarbeitendes Gewerbe	-0,940	0,000	0,036	0,54	0,077	0,26
Handel	-0,064	0,701	-0,013	0,78	0,129	0,02
Verkehr/Nachrichtenübermittlung, Banken, Versicherungen	-0,363	0,253	0,014	0,89	0,275	0,01
Öffentlicher Dienst/Sozialversicherung	-0,924	0,000	0,031	0,66	0,158	0,06
Bereich andere Dienstleistungen, k.A.	-0,268	0,091	0,008	0,86	0,102	0,05
Einkommen wichtig	0,042	0,664	-0,002	0,94	-0,051	0,11
Suchintensität (Zahl der genutzten Suchwege)	0,088	0,000	-0,002	0,71	0,002	0,77
Bereits ein Stellenangebot vom AA wegen zu niedrigem Einkommen abgelehnt?	0,296	0,070	-0,012	0,78	-0,001	0,98
Bisherige Arbeitslosigkeitsdauer insgesamt (R: < 6 Monate)						
6-12 Monate	-0,451	0,028	0,006	0,91	0,061	0,34
12-24 Monate	0,100	0,564	-0,004	0,92	-0,098	0,08
> 24 Monate/k.A.	0,058	0,716	-0,008	0,85	-0,055	0,30
Art der Kündigung der letzten Beschäftigung (R: Sonstiges)						
Eigenkündigung	0,273	0,135	-0,007	0,88	0,041	0,48
Arbeitgeberkündigung	0,296	0,011	-0,004	0,90	-0,020	0,59
Arbeitsamttypen (R: Typ 1)						
Typ 2	-0,028	0,877	0,000	0,99	-0,044	0,45
Typ 3	-0,316	0,019	0,006	0,89	-0,041	0,36
Typ 4	-0,296	0,481	0,016	0,89	-0,089	0,47
Typ 5	-0,558	0,000	0,017	0,71	-0,070	0,19
Einkommen aus Nebentätigkeit	-0,441	0,065	0,015	0,81	0,053	0,53
Haushaltseinkommen	0,000	0,000	0,000	0,39	0,000	0,14
Stundenlohn (R: bis 4,60 Euro)						
4,61 bis 6,90 Euro	0,061	0,722	-0,009	0,84	0,072	0,16
6,91 bis 9,20 Euro	-0,124	0,480	-0,005	0,92	0,065	0,24
9,21 Euro oder mehr	-0,206	0,293	0,001	0,98	0,239	0,00
Keine (plausible) Angabe	-0,158	0,380	-0,002	0,96	0,075	0,16
Konstante	0,404	0,269	-0,051	0,62	0,444	0,00
Fälle	983		983		980	
(Pseudo-)R <sup>2</sup>	0,325		0,395		0,1286	
Log pseudolikelihood	-459,74995					
ch <sup>2</sup> /F	278,82		30,08		4,84	
LR-Test (prob > chi2)/F-Test	0,000		0,000		0,000	
Anmerkung: Robuste Standardfehler; zur Schätzung vgl. Abschnitt 4.3.2.2.						

Auch im Rahmen der dreistufigen Instrumentvariablenschätzung (Tabelle 18) wird der kausale Effekt des Kombilohnes auf die Lohnzufriedenheit bestätigt. Wie von Angrist (1991) für den Fall binärer Zielvariablen festgestellt, kommt die IV-Schätzung des kausalen Effektes auch hier mit 0,28 (Modell 3) zu einem sehr ähnlichen Ergebnis wie die Regression mit Kontrollfunktion, das sich als signifikant ausweisen lässt. Da die Modelle 1 und 2 lediglich zur Berechnung der Instrumente dienen, wird auf ihre Interpretation verzichtet.

### 6.5.2 Die Stable Unit Treatment Value Assumption (SUTVA)

Wie die Conditional Independence Assumption ist auch die Stable Treatment Value Assumption nicht zu beweisen, sondern kann nur im Hinblick auf ihre Plausibilität im betrachteten Untersuchungsfeld geprüft werden. Damit die SUTVA erfüllt ist, müssen zwei Bedingungen erfüllt sein (Rubin 2005: 323; 1991: 1219). Zum einen ist dies die Bedingung, dass zwischen den betrachteten Individuen keine Wechselwirkungen bestehen. Der Wert der Zielvariable bei Individuum  $i$  im Treatment- oder Kontrollzustand darf nicht davon beeinflusst werden, ob die Person  $j$  ein Treatment erhält oder nicht. So darf etwa die Wirkung eines computerunterstützten Kurses (Treatment) bei Schüler  $i$  nicht davon abhängig sein, ob sein Freund, Schüler  $j$ , ebenfalls am Kurs teilnimmt oder nicht (Rosenbaum/Rubin 1984: 26). In der Evaluation aktiver Arbeitsmarktpolitik treten solche Wechselwirkungen in Form allgemeiner Gleichgewichtseffekte oder von Verdrängungseffekten auf (Jaenichen 2002b). Zum anderen beinhaltet die SUTVA die Forderung, dass keine versteckten Versionen des Treatments existieren (Rubin 2005: 323). Ein Beispiel für versteckte Versionen wurde bereits weiter unten angeführt (Abschnitt 4.2.1). Dort wurde darauf hingewiesen, dass die Frage „Erzielen Frauen einen geringeren Lohn, weil sie Frauen sind?“ im RKM nur dann eindeutig zu beantworten ist, wenn klar ist, auf welche Version des Treatments, welche hypothetische Manipulation sich die Aussage bezieht. In Abschnitt 5.2.3 wurde dann thematisiert, welche hypothetische Manipulation der vorliegenden Untersuchung zugrunde liegt.

Das Vorliegen von Wechselwirkungen oder unterschiedlichen Versionen des Treatments ist ein Problem, das stets im Kontext der gestellten Forschungsfrage thematisiert werden muss. Hier soll die SUTVA in Bezug auf den Kombilohn und die Lohnzufriedenheit bzw. den Kombilohn und die Beschäftigungsstabilität thematisiert werden.

#### 6.5.2.1 Wechselwirkungen zwischen Geförderten und Ungeförderten

Zunächst sollen Wechselwirkungen untersucht werden. Betrachtet man etwa die Wirkung des Kombilohnes auf die Arbeitslosigkeitsdauer (Aktivierung), dann kön-

nen Wechselwirkungen in Form von Verdrängungseffekten auftreten. So wäre es möglich, dass Arbeitslose, die eine Beschäftigung suchen, bei konstanter Stellenzahl von solchen verdrängt werden, die eine Beschäftigung suchen und vom Arbeitsamt die Zusage eines Kombilohnes erhalten haben. Sei  $ald_i^1$  die Arbeitslosigkeitsdauer im Falle der Kombilohnförderung und  $ald_i^0$  die im Fall ohne Kombilohn, dann träte bei Verdrängungseffekten der Fall ein, dass die Tatsache, dass Person  $j$  einen Kombilohn erhält ( $kl_j = 1$ ), den Wert der Zielvariable von Person  $i$  im Fall des Nichterhaltes eines Kombilohnes  $ald_i^0$  verändert. Kombilohngeförderte können aufgrund des Zuschusses niedrigere Lohnforderungen stellen und trotzdem noch einen höheren Lohn erzielen als Arbeitssuchende ohne Zuschuss. Daher kann ein Arbeitgeber statt Person  $i$  die Person  $j$  einstellen, weshalb  $i$  nun länger in Arbeitslosigkeit verbleibt. Die Tatsache, dass Person  $i$  einen Kombilohn erhält, verlängert hier also die Arbeitslosigkeitsdauer von Person  $j$ .

Es ist nun zu fragen, ob sich im Fall der Wirkung des Kombilohnes auf die Lohnzufriedenheit  $lz$  bzw. die Beschäftigungsstabilität  $bs$  ähnliche Effekte ergeben können, welche die kausale Interpretation des gemessenen Effektes fraglich erscheinen lassen. Für die Messung des  $MKE_T$  ist es nötig auszuschließen, dass das Ausmaß der Werte  $lz^0$  und  $bs^0$ , der Lohnzufriedenheit bzw. Beschäftigungsstabilität der Ungeförderten im ungeförderten Zustand, von der Tatsache abhängt, ob Person  $j$  einen Kombilohn erhält oder nicht.<sup>74</sup> Nur dann sind die Werte angemessene Repräsentationen des kontrafaktischen Zustands der Geförderten ohne Förderung. Eine solche Wechselwirkung ist etwa dann möglich, wenn die Lohnzufriedenheit der ungeförderten Person  $j$  dadurch positiv beeinflusst wird, dass sie die Lohnzahlung seines eigenen Arbeitgebers nun positiver als in Abwesenheit von geförderten Personen  $i$  bewertet, weil sie – seiner subjektiven Einschätzung nach – hoch genug ist, dass im Vergleich zu geförderten Personen keine sozialstaatliche Unterstützung nötig ist. Existiert eine solche Wechselwirkung, so würde der Effekt von  $-0,14$  Prozentpunkten nicht nur den Effekt des Kombilohnes auf die Geförderten, sondern auch die Wirkung des Kombilohnes auf die Ungeförderten einschließen. Im Extremfall hat der Kombilohn keinen Einfluss auf die Geförderten, sondern lediglich auf die Ungeförderten, weil deren Wahrscheinlichkeit für Lohnzufriedenheit steigt. Ähnlich kann es auch bei der Beschäftigungsstabilität sein, dass der Kombilohn diese tatsächlich erhöht, dass aber die Ungeförderten das Vorhandensein von Geförderten als Signal für die Schwierigkeiten betrachten, als Arbeitsloser eine adäquat entlohnte Anschlussbeschäftigung zu finden. Daher könnten sie bestrebt sein, verstärkt an ihrer bestehenden Beschäftigung

74 Grund ist die Bedingung, dass beim  $MKE_T$  die Forderung gilt:  $E(y^i | kl = 1, \mathbf{x}) = E(y^i | kl = 0, \mathbf{x})$ . Für den  $MKE$  wäre zusätzlich noch die Beeinflussung von  $y^i$  durch die Kombilohnbeteiligung anderer Personen auszuschließen.

festzuhalten. Die Folge wäre der beobachtete Nulleffekt des Kombilohnes auf die Beschäftigungsstabilität.

Ob eine Beeinflussung der Zielvariablen durch die Kombilohnbeteiligung der Geförderten stattfinden kann, hängt zunächst davon ab, inwiefern die Tatsache der Förderung durch die Ungeförderten überhaupt beobachtbar ist. Grundsätzlich muss für Außenstehende nicht bekannt sein, ob eine Person eine staatliche Förderung erhält. Am ehesten wird dies sein, wenn geförderte und ungeforderte Beschäftigungsaufnahmen in einem Betrieb erfolgen. Aufgrund der geringen Förderzahlen im Mainzer Modell ist dies aber unwahrscheinlich. Mit steigender Förderzahl in der Umgebung eines Ungeförderten kann sich allerdings die Sichtbarkeit des Mainzer Modells als finanzielle Förderung erhöhen, sodass dennoch eine Beeinflussung stattfinden kann. Auch hier sprechen die geringen Förderzahlen eher gegen eine solche Annahme.

Tabelle 19: Einfluss der Zahl der Geförderten auf Lohnzufriedenheit und Beschäftigungsstabilität von Ungeförderten

	Logistische Regression der Lohnzufriedenheit		Cox-Regression des Beschäftigungsendes	
	Odds Ratio	p	Risk Ratio	p
Zahl der Förderfälle im jeweiligen Arbeitsamtbezirk	1,000	0,85	1,000	0,40
Haushaltskonstellation (R: Alleinlebend ohne Kind)				
Erwerbsloser Partner/ohne Kind	0,736	0,39	1,389	0,26
Erwerbstätiger Partner/ohne Kind	1,420	0,28	0,875	0,64
Erwerbsloser Partner/mit Kind	0,958	0,89	0,982	0,95
Erwerbstätiger Partner/mit Kind	1,308	0,38	0,919	0,74
Arbeitszeit	0,992	0,40	1,012	0,12
Wie wurde die Stelle gefunden (R: Arbeitsvermittlung)				
Bekannte/Freunde	1,100	0,76	0,596	0,04
Eigene Initiative	0,629	0,15	0,874	0,58
Sonstiges	0,819	0,52	0,456	0,00
Vorher Weiterbildung ja	0,938	0,92	1,953	0,14
Lohnersatzleistungen vom Arbeitsamt (R: Arbeitslosengeld)				
Arbeitslosenhilfe	1,524	0,13	1,084	0,73
Keine Leistungen vom Arbeitsamt	0,808	0,43	0,897	0,65
Sozialhilfebezug	0,634	0,29	1,059	0,88
Beschäftigung befristet	0,934	0,76	1,210	0,30
Branche (R: Leih-/Zeitarbeitsfirma)				
Reinigungsgewerbe	0,703	0,50	0,611	0,23
Hotel- und Gaststättengewerbe	0,718	0,53	0,663	0,27
Callcenter	3,845	0,28	0,732	0,69
Energie und Wasser; Bergbau, verarbeitendes Gewerbe	1,669	0,16	0,497	0,01
Handel	1,794	0,17	0,387	0,01

	Logistische Regression der Lohnzufriedenheit		Cox-Regression des Beschäftigungsendes	
	Odds Ratio	p	Risk Ratio	p
Verkehr/Nachrichtenübermittlung, Banken, Versicherungen	4,830	0,04	0,343	0,06
Öffentlicher Dienst/Sozialversicherung	2,619	0,05	0,221	0,00
Bereich andere Dienstleistungen, k.A.	1,786	0,11	0,402	0,00
Einkommen wichtig	0,741	0,15	0,687	0,04
Suchintensität (Zahl der genutzten Suchwege)	0,967	0,48	0,977	0,57
Bereits ein Stellenangebot vom AA wegen zu niedrigem Einkommen abgelehnt?	0,840	0,64	0,518	0,08
Bisherige Arbeitslosigkeitsdauer insgesamt (R: < 6 Monate)				
6–12 Monate	1,413	0,34	1,330	0,36
12–24 Monate	0,516	0,06	1,596	0,12
> 24 Monate/k.A.	0,665	0,19	1,405	0,22
Art der Kündigung der letzten Beschäftigung (R: Sonstiges)				
Eigenkündigung	1,526	0,31	1,487	0,24
Arbeitgeberkündigung	0,966	0,89	0,957	0,83
Arbeitsamttypen (R: Typ 1)				
Typ 2	0,896	0,80	0,777	0,51
Typ 3	0,778	0,41	1,181	0,52
Typ 4	0,735	0,73	0,365	0,34
Typ 5	0,837	0,64	1,642	0,12
Einkommen aus Nebentätigkeit	0,746	0,54	1,123	0,75
Haushaltseinkommen	1,000	0,09	1,000	0,63
Stundenlohn (R: bis 4,60 Euro)				
4,61 bis 6,90 Euro	1,066	0,87	1,212	0,57
6,91 bis 9,20 Euro	0,957	0,91	1,091	0,81
9,21 Euro oder mehr	1,815	0,15	1,215	0,60
Keine (plausible) Angabe	1,351	0,45	1,548	0,21
Fälle	479,000		468,000	
Pseudo-R <sup>2</sup>	0,089			
Log likelihood	-295,368		-841,787	
chi <sup>2</sup>	58,136		63,313	
aic	672,735		1763,575	
bic	843,775		1929,514	
Anmerkung: Die „Zahl der Förderfälle im jeweiligen Arbeitsamtbezirk“ stellen eine eigene Berechnung mithilfe von Prozessdaten zum Mainzer Modell dar, der CAST-Förderdatei (vgl. Kaltenborn et al. 2005: 13).				

Im Anschluss an das Vorgehen von Kangasharju (2005) kann dieser Aspekt der SUTVA auch statistisch durch die Aufnahme einer Kontrollvariable in eine Regression überprüft werden, die die Sichtbarkeit des MZM für den Ungeförderten abbildet. Als Kontrollvariable wird hier die Zahl der Förderfälle im Arbeitsamtbezirk

verwendet. Zeigt diese in einer logistischen Regression der Lohnzufriedenheit bzw. in einer Cox-Regression der Beschäftigungsdauer bei den Ungeförderten einen signifikanten Effekt, wäre dies als ein Hinweis auf die Verletzung der SUTVA zu deuten. Das Ergebnis dieses Tests ist in Tabelle 19 abzulesen. Dort erweist sich die Zahl der MZM-Förderfälle weder als signifikant für die Zufriedenheit der Ungeförderten noch für deren Beschäftigungsstabilität (daran ändert auch eine Aufnahme der quadrierten Zahl der Förderfälle nichts). Demnach spricht dieser Test nicht gegen die Gültigkeit der SUTVA.

#### 6.5.2.2 *Versteckte Zuweisungsmechanismen*

Die Vermischung verschiedener Zuweisungsmechanismen ist ein zweiter Grund für die Verletzung der SUTVA. Eine Frage, die sich bei Maßnahmeteilnahmen im Allgemeinen und bei Kombilohnmaßnahmen im Besonderen stellt, jedoch bisher im Zusammenhang mit der SUTVA selten thematisiert wurde, ist die Frage, ob unterschiedliche Versionen des Treatments „Kombilohn“ mit unterschiedlichen potenziellen Ergebnissen existieren. Im Fall der Kombilohnförderung „Mainzer Modell“ waren grundsätzlich zwei Möglichkeiten der Förderung gegeben, die eventuell zu unterschiedlichen Einschätzungen der Reziprozität im Beschäftigungsverhältnis bzw. daraus folgend der Lohnzufriedenheit führen können. Zum einen konnte der Zuschuss vor Aufnahme der Beschäftigung beantragt worden sein. Das heißt, dem Geförderten war von vornherein bekannt, dass die zukünftige Beschäftigung der Förderbarkeit mit einem Lohnzuschuss unterlag und in diesem Sinne niedrig entlohnt wird. Zum anderen konnte der Zuschuss auch im Nachhinein (im Rahmen einer bestimmten Frist) für eine bereits aufgenommene Beschäftigung beantragt werden. In diesem Fall erfuhr der Geförderte erst nach Aufnahme der Beschäftigung, dass es sich um eine Niedriglohnstelle handelte. Beide Varianten können eventuell zu unterschiedlichen potenziellen Ergebnissen hinsichtlich der Lohnzufriedenheit führen, das heißt unterschiedliche Wirkungen entfalten. So könnte die Aufnahme einer Beschäftigung unter Wissen der Bezuschussung bei den Geförderten nicht zur Einschätzung der Verletzung der Reziprozitätsnorm führen, da sie die Stelle eventuell ohne die Förderung nicht erhalten hätten. Erfährt ein Beschäftigter dagegen erst nach der Aufnahme der Beschäftigung davon, so entsteht womöglich gerade deshalb die negative Signalwirkung des Kombilohnes und nicht durch die Tatsache der Förderung an sich.

Da in der Befragung zum Mainzer Modell u. a. die Frage gestellt wird, ob der Zuschuss erst nach der Einstellungszusage des Arbeitgebers beantragt wurde („Antragstellung nach Beschäftigungsaufnahme“, AnB) oder ob erst die Förderung beantragt wurde und dann die Beschäftigung aufgenommen wurde („Antragstellung vor Beschäftigungsaufnahme“, AvB), lässt sich diese Information zum Test

der SUTVA verwenden. Dazu wird die Variable  $kl$  nicht mehr binär, sondern mit drei Ausprägungen betrachtet, mit 0 für die regulär Beschäftigten, 1 für einen AvB-Kombilohn und 2 für einen AnB-Kombilohn:

$$lz_{obs,i} = \begin{cases} lz_i^0, & \text{falls } kl_i = 0 \\ lz_i^{1f}, & \text{falls } kl_i = 1 \\ lz_i^{1u}, & \text{falls } kl_i = 2 \end{cases}$$

In diesem Fall kann das Matching als „multiple treatment matching“ durchgeführt werden (Imai/van Dyke 2004; Joffe/Rosenbaum 1999). Dabei werden die nötigen Propensity Scores entweder mit einem multinomialen Logit- oder Probit-Modell geschätzt, oder sie können auch – wie hier – auf Basis einzelner logistischer Regressionen bestimmt werden (Lechner 2001).

Tabelle 20: Logistische Regressionen zur Bestimmung der Propensity Scores im Multiple-Treatment-Fall

Binäre logistische Regressionen	Modell 1: AvB-Kombilohn vs. kein Kombilohn		Modell 2: AnB-Kombilohn vs. kein Kombilohn		Modell 3: AnB- vs. AvB- Kombilohn	
	Odds Ratio	p	Odds Ratio	p	Odds Ratio	p
Haushaltskonstellation (R: Alleinlebend ohne Kind)						
Erwerbsloser Partner/ohne Kind	2,228	0,05	2,662	0,00	0,958	0,91
Erwerbstätiger Partner/ohne Kind	1,174	0,72	1,433	0,34	0,762	0,58
Erwerbsloser Partner/mit Kind	7,411	0,00	5,418	0,00	1,379	0,32
Erwerbstätiger Partner/mit Kind	1,974	0,07	2,493	0,01	0,832	0,64
Arbeitszeit	0,955	0,00	0,960	0,00	1,006	0,52
Wie wurde die Stelle gefunden (R: Arbeitsvermittlung)						
Bekannte/Freunde	0,850	0,62	0,865	0,62	0,854	0,59
Eigene Initiative	1,023	0,94	1,191	0,53	0,933	0,80
Sonstiges	0,594	0,13	0,735	0,30	0,992	0,98
Vorher Weiterbildung ja	2,973	0,13	2,207	0,20	0,939	0,92
Lohnersatzleistungen vom Arbeitsamt (R: Arbeitslosengeld)						
Arbeitslosenhilfe	3,167	0,00	2,791	0,00	0,945	0,85
Keine Leistungen vom Arbeitsamt	0,696	0,32	0,960	0,89	0,857	0,68
Sozialhilfebezug	1,514	0,30	2,578	0,01	0,827	0,56
Beschäftigung befristet	2,701	0,00	1,971	0,00	1,156	0,49
Branche (R: Leih-/Zeitarbeitsfirma)						
Reinigungsgewerbe	0,791	0,65	0,785	0,56	0,934	0,87
Hotel- und Gaststättengewerbe	1,583	0,33	0,813	0,62	1,447	0,37
Callcenter	4,985	0,08	2,795	0,21	1,917	0,24



Binäre logistische Regressionen	Modell 1: AvB-Kombilohn vs. kein Kombilohn		Modell 2: AnB-Kombilohn vs. kein Kombilohn		Modell 3: AnB- vs. AvB- Kombilohn	
	Odds Ratio	p	Odds Ratio	p	Odds Ratio	p
Energie und Wasser; Bergbau, verarbeitendes Gewerbe	0,238	0,00	0,156	0,00	1,143	0,76
Handel	1,348	0,44	0,703	0,31	1,672	0,12
Verkehr/Nachrichtenübermittlung, Banken, Versicherungen	0,794	0,75	0,415	0,17	1,122	0,88
Öffentlicher Dienst/Sozialversicherung	0,269	0,02	0,190	0,00	1,478	0,52
Bereich andere Dienstleistungen, k.A.	0,660	0,24	0,498	0,02	1,363	0,32
Einkommen wichtig	1,492	0,09	0,933	0,74	1,164	0,47
Suchintensität (Zahl der genutzten Suchwege)	1,223	0,01	1,134	0,03	1,022	0,75
Bereits ein Stellenangebot vom AA wegen zu niedrigem Einkommen abgelehnt?	1,808	0,12	1,623	0,16	1,559	0,17
Bisherige Arbeitslosigkeitsdauer insgesamt (R: < 6 Monate)						
6–12 Monate	0,260	0,01	0,490	0,11	0,777	0,66
12–24 Monate	0,715	0,44	1,317	0,47	0,638	0,31
> 24 Monate/k.A.	0,760	0,50	1,373	0,36	0,711	0,40
Art der Kündigung der letzten Beschäftigung (R: Sonstiges)						
Eigenkündigung	1,349	0,53	1,831	0,12	0,830	0,67
Arbeitgeberkündigung	1,378	0,23	1,686	0,03	0,927	0,75
Arbeitsamttypen (R: Typ 1)						
Typ 2	1,418	0,40	0,768	0,48	1,668	0,16
Typ 3	0,657	0,18	0,498	0,01	0,966	0,91
Typ 4	1,392	0,73	0,428	0,30	0,780	0,76
Typ 5	0,320	0,00	0,397	0,00	0,634	0,21
Einkommen aus Nebentätigkeit	0,545	0,27	0,446	0,14	1,644	0,37
Haushaltseinkommen	1,000	0,22	0,999	0,00	1,001	0,01
Stundenlohn (R: bis 4,60 Euro)						
4,61 bis 6,90 Euro	0,987	0,97	1,214	0,56	1,261	0,47
6,91 bis 9,20 Euro	0,660	0,33	1,162	0,67	0,820	0,59
9,21 Euro oder mehr	0,426	0,07	1,026	0,95	0,919	0,85
Keine (plausible) Angabe	0,853	0,69	0,609	0,16	1,606	0,18
Fälle	668		746		454	
Pseudo-R <sup>2</sup>	0,346		0,321		0,057	
Log likelihood	-259,47		-330,22		-290,29	
ch <sup>2</sup>	275,05		311,48		35,33	
LR-Test (prob > chi2)	0,00		0,00		0,64	
aic	598,94		740,44		660,58	
bic	779,11		925,03		825,30	
Wald-Test der Gleichheit aller Koeffizienten in beiden Modellen auf Basis einer Seemingly- Unrelated-Schätzung	chi2 (39)	42,06				
	Prob > chi2	0,34				

Bei einem Kontrollzustand und zwei Treatments gibt es drei Vergleiche. Die zugehörigen Berechnungen der Propensity Scores erfolgen mit drei logistischen Regressionen, deren Ergebnisse in Tabelle 20 aufgeführt sind. Der Frage einer Verletzung der SUTVA lässt sich zunächst im Vergleich der beiden Teilgruppen von Geförderten nachgehen. Betrachtet man die Modelle 1 und 2 der Tabelle, dann erweisen sich einige Einflussfaktoren in Modell 1 als signifikant und im Modell 2 als insignifikant bzw. umgekehrt. Dies könnte darauf verweisen, dass tatsächlich Unterschiede zwischen den beiden Gruppen existieren und sich dies in unterschiedlichen Zuweisungsmechanismen niederschlägt. Allerdings zeigt ein verallgemeinerter Wald-Test auf Basis einer Seemingly Unrelated Regression<sup>75</sup> (Weesie 1999; Greene 2003: 340 ff.), dass die Koeffizienten der Variablen, die den Zuweisungsmechanismus abbilden, sich beim AvB- und beim AnB-Kombilohn statistisch nicht unterscheiden.

Im Modell 3 der Tabelle 20 werden die Determinanten der Maßnahmeteilnahme nochmals dahin gehend betrachtet, ob sie auf die AnB-Kombilohnförderung einen anderen Einfluss hat als auf die AvB-Förderung. Ein Unterschied würde wiederum darauf verweisen, dass hinter den beiden Arten von Geförderten doch unterschiedliche Zuweisungsmechanismen stehen. Der LR-Test des Modells 3 zeigt jedoch, dass die Hypothese der Gleichheit aller Koeffizienten mit null nicht verworfen werden kann ( $p = 0,57$ ), das Ergebnis eines einheitlichen Zuweisungsmechanismus wird also nochmals bestätigt.

Allerdings, auch wenn dies der Fall ist, könnte es trotzdem noch sein, dass der Wert der Zielvariable Lohnzufriedenheit in beiden Varianten unterschiedlich ausfällt. Dies kann überprüft werden, wenn beide Gruppen als Personen mit zwei verschiedenen Treatments betrachtet werden und die Ergebnisse miteinander verglichen werden.

Tabelle 21: Der kausale Effekt des AnB-Kombilohnes

	Zufrieden mit dem Lohn (Anteil)		Effekt	Standardfehler	t-Wert
	AnB-Kombilohn	Regulärer Lohn			
Vor dem Matching	0,368	0,589	-0,220	0,037	-5,89
<b>Ein nächster Nachbar; Propensity Score (ohne Zurücklegen); Caliper 0,005</b>					
Nach dem Matching	0,394	0,512	-0,118	0,062	-1,90
Zahl der Matches	127				
<b>Radius-Matching; Log Odds (mit Zurücklegen); Caliper 0,01</b>					
Nach dem Matching	0,376	0,512	-0,136	0,059	-2,32
Zahl der Matches	243				

75 Der Test muss auf Basis einer Seemingly-Unrelated-Schätzung durchgeführt werden, weil beide Male dieselben Personen die Vergleichsgruppe der Ungeförderten bilden (vgl. Weesie 1999).

In Tabelle 21 wird der kausale Effekt der AvB-Variante auf die Lohnzufriedenheit geschätzt. Dieser beträgt im Fall des Ein-nächster-Nachbar-Matchings ca. -12 Prozentpunkte und kann bei einem t-Wert von -1,90 als signifikant betrachtet werden. Das Radius-Matching, das mehr Fälle und damit mehr Informationen verwendet, zeigt dies noch deutlicher. Hier ist der gemessene kausale Effekt ca. -14 Prozentpunkte und bei einem t-Wert von 2,32 ebenfalls signifikant. Die Frage ist nun, ob auch dann noch ein Effekt des Kombilohnes nachzuweisen ist, wenn dieser erst nach Aufnahme der Beschäftigung beantragt wird.

Tabelle 22: Der kausale Effekt des AvB-Kombilohnes

	Zufrieden mit dem Lohn (Anteil)		Effekt	Standard- fehler	t-Wert
	AvB- Kombilohn	Regulärer Lohn			
Vor dem Matching	0,362	0,589	-0,227	0,042	-5,39
<b>Ein nächster Nachbar; Propensity Score (ohne Zurücklegen); Caliper 0,005</b>					
Nach dem Matching	0,389	0,526	-0,137	0,072	-1,90
Zahl der Matches	95				
<b>Radius-Matching; Log Odds (mit Zurücklegen); Caliper 0,01</b>					
Nach dem Matching	0,343	0,546	-0,203	0,065	-3,11
Zahl der Matches	163				

Wie Tabelle 22 zeigt, ist dies entgegen der oben geäußerten Vermutung der Fall, der Effekt des AvB-Kombilohnes ist ebenfalls in beiden ausgewiesenen Matchingvarianten signifikant negativ und beträgt im SNNM-Fall ca. -14 Prozentpunkte, beim Radius-Matching sogar -20 Prozentpunkte. Damit scheint hier zumindest in Bezug auf das Vorher oder Nachher der Förderung kein Problem der unterschiedlichen Zuweisungsmechanismen und der SUTVA zu bestehen.

Tabelle 23: Antragstellung vor vs. nach Beschäftigungsaufnahme

	Zufrieden mit dem Lohn (Anteil)		Effekt	Standard- fehler	t-Wert
	AvB- Kombilohn	AnB- Kombilohn			
Vor dem Matching	0,362	0,368	-0,006	0,046	-0,15
<b>Ein nächster Nachbar; Propensity Score (ohne Zurücklegen); Caliper 0,005</b>					
Nach dem Matching	0,339	0,411	-0,072	0,065	-1,10
Zahl der Matches	95				
<b>Radius-Matching; Log Odds (mit Zurücklegen); Caliper 0,01</b>					
Nach dem Matching	0,342	0,384	-0,042	0,052	-0,81
Zahl der Matches	181				

Das zeigt auch die Tabelle 23, in der die beiden Gefördertengruppen miteinander verglichen werden. Es zeigt sich kein signifikant von null verschiedener Effekt auf die Lohnzufriedenheit.

## 6.6 Fehlender job-worker-match als alternative Erklärung?

Der Theorie des Matchings (Jovanovic 1979) zufolge ist der job-worker-match, die Übereinstimmung der Anforderungen des Betriebes mit den tatsächlichen Qualifikationen bzw. der tatsächlichen Grenzproduktivität des Beschäftigten, eine entscheidende Determinante der Beschäftigungsdauer. Die Theorie geht davon aus, dass verschiedene Arbeitskräfte auf unterschiedlichen Arbeitsplätzen eine unterschiedliche Produktivität aufweisen und der Lohn in individuellen Verhandlungen festgelegt wird. Für den einstellenden Betrieb ist die reale Produktivität eines Stellenbewerbers vor der Stellenbesetzung nicht zu erkennen, sodass sich die Entlohnung an der Grenzproduktivität (gegeben die vorhandenen Informationen) orientiert. Die tatsächliche Produktivität stellt sich aber relativ bald nach Aufnahme der Beschäftigung heraus. Besitzen Arbeitskräfte in Relation zu ihren Arbeitsaufgaben eine hohe Produktivität, dann wird ihr Lohn mit der Dauer der Beschäftigung steigen, und sie verbleiben daher im Betrieb. Beschäftigte mit niedriger Produktivität neigen dagegen dazu, den Betrieb zu verlassen und nach einer Beschäftigung zu suchen, in der ihre Produktivität und dadurch auch ihr Lohn hoch ist.

Tabelle 24: Einfluss des Kombilohnes auf den qualifikatorischen Mismatch

	Kein qualifikatorischer Mismatch		Effekt	Standardfehler	t-Wert
	Kombilohn	Regulärer Lohn			
Vor dem Matching	0,074	0,072	0,002	0,023	0,10
<b>Ein nächster Nachbar; Propensity Score (ohne Zurücklegen); Caliper 0,005</b>					
Nach dem Matching	0,074	0,051	0,023	0,026	0,88
Zahl der Matches	174				
<b>Radius-Matching; Log Odds (mit Zurücklegen); Caliper 0,01</b>					
Nach dem Matching	0,076	0,070	0,006	0,026	0,22
Zahl der Matches	224				

Wenn aufgrund der Fremdselektion durch die Arbeitsämter unter den Teilnehmern qualifikatorische Mismatches häufiger vorkommen als unter Nichtteilnehmern, kann auch dies zu erhöhter Lohnunzufriedenheit und darauf folgendem Maßnahmeabbruch führen. Daher stellt die job-worker-match-Theorie entweder eine weitere, eventuell sogar alternative Erklärung für die hohen Quoten der Beschäfti-

gungsabbrüche in Kombilohnmaßnahmen dar. Dies allerdings nur, wenn sich nachweisen lässt, dass die Wirkung des Kombilohnes auf den Beschäftigungsabbruch über den Mechanismus des qualifikatorischen Mismatches verläuft. Analog zur Überprüfung des Reziprozitätsmechanismus zeigt Tabelle 24, dass der Kombilohn keinen Einfluss auf den job-worker-match besitzt.<sup>76</sup>

## 6.7 Reichweite der Ergebnisse

Zur abschließenden Einordnung der voranstehenden Analysen ist noch deren Reichweite zu diskutieren. Es stellt sich die Frage, inwiefern das Ergebnis auf andere Kombilohnmaßnahmen als das Mainzer Modell verallgemeinerbar ist. Dabei sind drei Aspekte zu berücksichtigen. Zunächst war die Analyse auf Personen beschränkt, die entweder langzeitarbeitslos waren oder keine formale Berufsqualifikation aufwiesen. Dies hatte Gründe im Erhebungsdesign der Befragung, die zu Beginn der vorliegenden Forschungsarbeit bereits abgeschlossen war und auf die daher kein Einfluss genommen werden konnte. Daher ist nicht eindeutig, ob der gefundene Effekt eine allgemeine Reaktion auf Kombilöhne darstellt oder vom Qualifikationsniveau abhängt. Die Analyse in Abschnitt 6.5.1.2, Tabelle 16, spricht jedoch dafür, dass die Ergebnisse über den Kreis der geringqualifizierten bzw. ehemals langzeitarbeitslosen Beschäftigten hinaus verallgemeinerbar sein dürften, vorausgesetzt sie sind im Niedriglohnbereich beschäftigt.

Anders dürfte dies bei den Alleinerziehenden der Fall sein, obwohl aufgrund zu wenig Kontrollpersonen in der Vergleichsgruppe keine gesicherten Aussagen möglich sind (vgl. Abschnitt 6.1). Hier ist denkbar, dass der Kombilohnbezug nicht als Reziprozitätsverletzung empfunden wird, da bei Alleinerziehenden die Vereinbarkeit von Erwerbstätigkeit und Kinderbetreuung im Vordergrund stand. Überdies wurden – konstruktionsbedingt im Mainzer Modell – bei Alleinerziehenden oft Beschäftigungen gefördert, die lediglich wegen ihres Teilzeitcharakters förderfähig waren, dem Stundenlohn nach jedoch kaum als Niedriglohnbeschäftigung betrachtet werden können. Es lässt sich weder aus den theoretischen noch empirischen Analysen ableiten, ob eine subjektive Reziprozitätsverletzung im Rahmen des Kombilohnes auch außerhalb von Niedriglohnbeschäftigung auftritt. Allerdings hat dies insofern wenig praktische Relevanz, als Kombilöhne hauptsächlich auf die Förderung von Niedriglohnbeschäftigung abzielen.

Schließlich werden zum Teil auch arbeitsmarktpolitische Maßnahmen als Kombilöhne bezeichnet, die sich von der hier analysierten Variante erheblich unter-

<sup>76</sup> Der qualifikatorische Mismatch wurde durch die Selbsteinschätzung der Beschäftigten erfasst, inwiefern ihre Stelle besondere Qualifikationen erfordert und ob sie diese erfüllen.

scheiden, etwa weil sie den Kombilohn mit einem Lohnkostenzuschuss verbinden (Hamburger Modell), weil der Zuschuss zeitlich unbefristet gewährt wird (Mini- und Midi-Jobs) oder die geförderte Tätigkeit nicht im ersten Arbeitsmarkt lokalisiert ist (Ein-Euro-Jobs). All diese Varianten weisen Besonderheiten auf, deren Wirkung auf die Reziprozitätswahrnehmung der Beschäftigten einer gesonderten Analyse bedürfte, für die jedoch (noch) keine Daten vorliegen. Daher ist auch unklar, inwiefern diese Varianten Eigenschaften aufweisen, die der subjektiven Empfindung von Reziprozitätsverletzungen entgegenwirken oder sie nochmals verstärken.

## 7 Zusammenfassung der Ergebnisse und Ausblick

Zum Abschluss dieser Arbeit sollen die Ergebnisse zusammengefasst, die Grenzen der durchgeführten Analyse aufgezeigt und mit einem Ausblick verknüpft werden. Ausgangspunkt der Arbeit war die Beobachtung aus der Begleitforschung zu Kombilohnmaßnahmen, dass diese von einer außergewöhnlich hohen Abbruchquote gekennzeichnet sind. In Auseinandersetzung mit dem Begriff der Reziprozität als Charakteristikum betrieblicher Sozialbeziehungen wurden Hypothesen zur Erklärung des Abbruchverhaltens in Kombilohnmaßnahmen formuliert und auf der Basis von Befragungsdaten zu kombilohngeförderten und regulär Beschäftigten einer empirischen Prüfung unterzogen. Zunächst wurde ein Propensity Score Matching von Geförderten und Ungeförderten durchgeführt, und es wurden mehreren Tests zur Sicherstellung des Ausgleiches von Unterschieden im Bereich kausalrelevanter Selektivität durchgeführt. Trotz einer relativ kleinen Stichprobe erwies sie sich als hinreichend für ein erfolgreiches Balancing der gematchten Teilgruppen. Auf der Grundlage des gematchten Datensatzes wurden dann die Hypothesen 1 bis 3 geprüft. Hypothese 1, die besagt, dass der Kombilohn eine negative Wirkung auf die Lohnzufriedenheit aufweist, konnte dabei durch die Ergebnisse des Tests eindeutig gestützt werden. Hypothese 2 zur negativen Wirkung des Zuschusses auf die Beschäftigungsstabilität wurde jedoch von den Daten nicht bestätigt. Zwar betont die Begleitforschung die hohen Abbruchquoten von Kombilohnfördermaßnahmen, doch scheinen diese nicht das auch in regulären Beschäftigungsverhältnissen übliche Maß zu übersteigen. Da jedoch auch kein positiver Effekt zu verzeichnen war, wurde die Fragestellung der Hypothese 3 entsprechend angepasst. Es wurde untersucht, ob die Wirkung des Kombilohnes auf die Lohnzufriedenheit zwar nicht groß genug ist, um eine positive Wirkung auf die Beschäftigungswirkung ins negative umzukehren, er jedoch Nebenwirkungen erzeugt, die eine positive Wirkung unterdrücken. Die Antwort auf diese Frage fiel jedoch nicht eindeutig aus. Zwar erwies sich der Einfluss des Kombilohnes nach dem „Herausrechnen“ des Einflusses auf die Wertschätzung des Lohnes als für das Beschäftigungsverhältnis stabilisierend, der entsprechende Wert in der Cox-Regression war jedoch nicht signifikant.

Die Ergebnisse der Hypothesenprüfung wurden anschließend – im Rahmen der statistischen und durch die verfügbaren Daten (z. B. als Variablen enthaltene Informationen, Fallzahl) gegebenen Möglichkeiten – umfangreichen Tests unterzogen, um ihre Validität sicherzustellen. Diese Tests stützten ausnahmslos die Gültigkeit der Annahmen, auf denen die durchgeführten statistischen Analysen beruhen, d. h., weder die Gültigkeit der Conditional Independence Assumption noch der Stable Unit Treatment Value Assumption wurde durch die Tests in Zweifel gezogen.

So zeigte der Kombilohn zwar einen negativen Effekt auf die Lohnzufriedenheit, jedoch nicht auf andere auf die Beschäftigung bezogene Bereichszufriedenheiten. Damit lässt sich argumentieren, dass die Gruppe der Geförderten nicht aus per se unzufriedenen Personen besteht, sondern dass die Unzufriedenheit auf den Kombilohn zurückzuführen ist. Ansonsten wären auch andere Lebensbereiche von der Unzufriedenheit betroffen. Dass die CIA wohl gültig ist, zeigen auch die Ergebnisse der Verfahren zur indirekten Berücksichtigung ausgelassener Selektionseinflüsse. So wies die Regression mit Kontrollfunktion keinen signifikanten Einfluss unbeobachteter kausalrelevanter Variablen nach und kam, wie auch die Instrumentvariablenschätzung, zu einem negativen kausalen Effekt des Kombilohnes auf die Lohnzufriedenheit. Schließlich wurde noch gezeigt, dass selbst dann, wenn doch unbeobachtete kausalrelevante Heterogenität bestünde, diese einen sehr starken Einfluss ausüben müsste, um den negativen Effekt des Kombilohnes auf die Lohnzufriedenheit als Scheinkorrelation zu erklären. Auch Tests der SUTVA stützen die Validität der Hypothesentests. So kann kein Einfluss der Geförderten auf die Zielvariablen der Ungeforderten nachgewiesen werden, und auch im Bereich unterschiedlicher Varianten des Treatments scheinen keine Probleme mit den Modellannahmen zu bestehen. Schließlich wurde noch eine alternative Erklärung des fehlenden Einflusses des Kombilohnes auf die Integration in Beschäftigung widerlegt, die auf einen Effekt des Kombilohnes auf die Übereinstimmung von Qualifikationen der Beschäftigten und Anforderungen der Tätigkeit Bezug nimmt.

Betrachtet man die Ergebnisse dieser Arbeit in einem breiteren Rahmen, so liefert die durchgeführte Analyse insgesamt einige Anhaltspunkte, dass im Bereich der Konzeption wie nachheriger Evaluation der Aspekt der Subjektivität menschlicher Deutungsprozesse zu wenig berücksichtigt wird. So ließ sich in dieser Arbeit nachweisen, dass objektive Maßnahmen der aktiven Arbeitsmarktpolitik von deren Adressaten durchaus anders interpretiert werden können, als dies von politischen Entscheidungsträgern intendiert ist. Der Kombilohn „Mainzer Modell“ erhöhte zwar objektiv das Einkommen der Geförderten, beeinflusst jedoch zudem die subjektive Wertschätzung des Arbeitsentgeltes negativ. Die für die Arbeitsmarktpolitik zentrale Frage allerdings, inwiefern eine solche subjektive Interpretation objektiver Maßnahmen auch den Erfolg der Arbeitsmarktpolitik beeinflusst, blieb in dieser Arbeit letztlich jedoch ohne zufriedenstellende Antwort. Die Ergebnisse lassen weder den eindeutigen Schluss zu, dass die Beeinflussung der subjektiven Wertschätzung des Arbeitgeberlohnes zum Ausbleiben des Integrationserfolges beim Kombilohn beigetragen hat. Noch ist der Schluss zwingend, dass der subjektive Effekt zu keinerlei Konsequenzen im für den Maßnahmeerfolg relevanten Handeln der Beschäftigten führt. Für die Handlungsrelevanz spricht, dass sich die Koeffizienten in die von Hypothese 3 formulierte Richtung verändern. Rechnet man den subjektiven Effekt des



Kombilohnes auf die Wertschätzung des Lohnes heraus, verweisen die Daten auf eine stabilere Integration der Geförderten in Beschäftigung. Dagegen spricht, dass diese Koeffizienten im statistischen Sinne nicht signifikant von null verschieden sind und somit nicht auszuschließen ist, dass die Übereinstimmung der Ergebnisse mit der formulierten Hypothese 3 zufällig erfolgte. Antworten auf die Frage der Erfolgsrelevanz subjektiver Interpretationen objektiver Maßnahmen der Arbeitsmarktpolitik sind daher auf zukünftige Analysen angewiesen.

Ebenfalls ein lohnendes Ziel zukünftiger Forschung dürfte die intensivere Auseinandersetzung der Evaluation aktiver Arbeitsmarktpolitik mit dem Begriff der Reziprozität sein. Die Argumentation der vorliegenden Arbeit bezüglich der Wirkung der Reziprozitätsnorm im Fall des Kombilohnes beruht einerseits auf dem argumentativen Verweis auf Ergebnisse der Experimentellen Ökonomie und empirischen Analysen aus der Soziologie und Sozialpsychologie, in denen sehr ähnliche Auswirkungen einer Verletzung der Reziprozitätsnorm nachgewiesen wurden. Andererseits wurden für die beobachtete paradoxe Wirkung des Kombilohnes auf die Lohnzufriedenheit alternative Erklärungen ausgeschlossen, so vor allem Unterschiede im Hinblick auf die Zufriedenheit mit der Tätigkeit oder mit der Arbeitsstelle insgesamt oder die Wirkung unbeobachteter Heterogenität bzw. unbeobachteter Drittvariablen. Für eine explizite empirische Auseinandersetzung mit den Deutungsprozessen der (kombilohngeförderten) Beschäftigten in Bezug auf deren Reziprozitätswahrnehmung fehlten der vorliegenden Analyse jedoch entsprechende Daten. Besonders geeignet wären hier wohl qualitative Daten gewesen, deren Erhebung jedoch wegen der Einstellung des Mainzer Modells im Jahr 2003 nicht mehr möglich ist. Hier könnten zukünftige Forschungsprojekte zu ähnlichen oder anderen Maßnahmen einen Beitrag liefern, diese Lücke zu schließen.



## Abbildungsverzeichnis

Abbildung 1: Verlauf der Arbeitslosenquote in Deutschland .....	15
Abbildung 2: Qualifikationsspezifische Arbeitslosenquoten .....	16
Abbildung 3: Anteil der Langzeitarbeitslosen an allen Arbeitslosen im Zeitverlauf.....	16
Abbildung 4: Erklärung der Höhe der Arbeitslosigkeit im europäischen Arbeitsmarktmodell.....	18
Abbildung 5: Erklärung des hysteretischen Verlaufs der Arbeitslosigkeit.....	19
Abbildung 6: Lohnsetzung als Ansatzpunkt der Arbeitsmarktpolitik .....	21
Abbildung 7: Senkung des Anspruchslohnes durch Kombilohn .....	25
Abbildung 8: Arbeitnehmerbeitrag zur Sozialversicherung für Allein- stehende: mit und ohne Förderung im Mainzer Modell.....	30
Abbildung 9: Abbruchquoten in verschiedenen Kombilohnmaßnahmen.....	36
Abbildung 10: Beschäftigungsstatus nach Abbruch der geförderten Beschäftigung im Mainzer Modell .....	37
Abbildung 11: Das Kontinuum der Reziprozitätsbeziehung .....	52
Abbildung 12: Variablenklassen.....	83
Abbildung 13: Causes of effects vs. effects of causes.....	90
Abbildung 14: Informationsklassen.....	113
Abbildung 15: Selektive Maßnahmeteilnahme bei Nutzenmaximierung.....	120
Abbildung 16: Selektivität einer Kombilohnförderung.....	128
Abbildung 17: Vergleichskonzepte bei der Evaluation von Kombilohn- maßnahmen .....	131
Abbildung 18: Common Support.....	148
Abbildung 19: Anteil der Befragten nach dem Abstand zwischen Beschäf- tigungsbeginn und Zeitpunkt der Befragung in Prozent.....	158



## Tabellenverzeichnis

Tabelle 1:	Gegenüberstellung: ökonomische und soziologische RC-Theorie...	39
Tabelle 2:	Mögliche Reaktionen auf Reziprozitätsverletzungen in Beschäftigungsverhältnissen.....	58
Tabelle 3:	Reaktionen auf Reziprozitätsverletzungen im Kontext des Kombilohnes .....	67
Tabelle 4:	Gründe für die Maßnahmeselektivität im Hamburger Modell und deren Operationalisierung.....	126
Tabelle 5:	Logistische Regressionen der Förderwahrscheinlichkeit vor und nach dem Matching .....	146
Tabelle 6:	Verteilung der PS vor und nach dem Matchen.....	149
Tabelle 7:	Mittelwertvergleiche und sbias vor und nach dem Matchen .....	151
Tabelle 8:	Der kausale Effekt des Kombilohnes auf die Lohnzufriedenheit..	154
Tabelle 9:	Sensitivität des geschätzten kausalen Effektes gegenüber Varianten des Matchingalgorithmus .....	156
Tabelle 10:	Propensity Score Matching: Effekte des Kombilohnes auf die Beschäftigungsdauer und den Beschäftigungsstatus.....	157
Tabelle 11:	Gesamt- und Nettoeffekt des Kombilohnes auf die Beschäftigungsdauer – Cox-Regressionen auf Basis gematchter Samples .....	160
Tabelle 12:	Rosenbaum-Bounds.....	167
Tabelle 13:	„Kausaler Effekt“ des Kombilohnes auf das individuelle Bruttoeinkommen.....	170
Tabelle 14:	Kausaler Effekt des Kombilohnes auf die Zufriedenheit mit der Tätigkeit.....	171
Tabelle 15:	Kausaler Effekt des Kombilohnes auf die Zufriedenheit mit der Arbeitsstelle insgesamt.....	171
Tabelle 16:	Kausaler Effekt des Kombilohnes bei Austausch der Treatmentpopulation .....	171
Tabelle 17:	Regression mit Kontrollfunktion .....	173
Tabelle 18:	Instrumentvariablenschätzung.....	175
Tabelle 19:	Einfluss der Zahl der Geförderten auf Lohnzufriedenheit und Beschäftigungsstabilität von Ungeförderten.....	179
Tabelle 20:	Logistische Regressionen zur Bestimmung der Propensity Scores im Multiple-Treatment-Fall .....	182
Tabelle 21:	Der kausale Effekt des AnB-Kombilohnes .....	184
Tabelle 22:	Der kausale Effekt des AvB-Kombilohnes.....	185
Tabelle 23:	Antragstellung vor vs. nach Beschäftigungsaufnahme.....	185
Tabelle 24:	Einfluss des Kombilohnes auf den qualifikatorischen Mismatch.	186



## Literatur

- Abbring, Jaap H.; van den Berg, Gerard J. (2003): The non-parametric identification of treatment effects in duration models. *Econometrica* 71, 1491–518.
- Achatz, Juliane; Gartner, Hermann; Glück, Timea (2005): Bonus oder Bias? Mechanismen geschlechtsspezifischer Entlohnung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 57 (3), 466–493.
- Adams, J. S. (1963): Toward an Understanding of Inequity. *Journal of Abnormal and Social Psychology*, 67, 422–436.
- Adloff, Frank (2005): Die Reziprozität der Gesellschaft – Zum Paradigma der Gabe in der Moderne. Corsten, Michael; Rosa, Hartmut; Schrader, Ralph (Hrsg.): *Die Gerechtigkeit der Gesellschaft*. VS: Wiesbaden, 25–51.
- Adloff, Frank; Mau, Steffen (2005): Zur Theorie der Gabe und Reziprozität. Adloff, Frank; Mau, Steffen (Hrsg.): *Vom Geben und Nehmen. Zur Soziologie der Reziprozität*. Frankfurt am Main: Campus, 9–57.
- Akerlof, George A. (1982): Labor Contracts as Partial Gift Exchange. *Quarterly Journal of Economics* 97, 543–569.
- Akerlof, George A.; Yellen, Janet L. (1990): The Fair Wage-Effort Hypothesis and Unemployment. *Quarterly Journal of Economics*, 105, 255–283.
- Albert, Hans (1999): Die Soziologie und das Problem der Einheit der Wissenschaften. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 51, 215–231.
- Amtsblatt der Europäischen Union L 124 vom 6. Mai 2003, 36–41.
- Angrist, Joshua D. (1991): Instrumental Variables Estimation of Average Treatment Effects in Econometrics and Epidemiology. NBER Technical Working Papers 0115, National Bureau of Economic Research, Inc.
- Arrow, Kenneth J. (1972): Gifts and Exchanges. *Philosophy and Public Affairs* 1, 343–362.
- Augurzky, Boris; Schmidt, Christoph (2001): The Propensity Score: A Means to an End. IZA Discussion Paper No. 271.
- Bacher, Johann (2002): Statistisches Matching: Anwendungsmöglichkeiten, Verfahren und ihre praktische Umsetzung in SPSS. *ZA-Informationen*, 51, 38–66.
- Bäcker, Gerhard (2002): Armut trotz Sozialhilfe? Zum Verhältnis von Einkommensarmut und Hilfe zum Lebensunterhalt. Sell, Stefan (Hrsg.): *Armut als Herausforderung. Bestandsaufnahme und Perspektiven der Armutsforschung und Armutsberichterstattung*. Berlin: Duncker & Humbolt.
- Becker, Sascha O.; Caliendo, Marco (2006): mhbounds – Sensitivity Analysis for Average Treatment Effects. Software, <http://www.caliendo.de/Papers/mhbounds/mhbounds.zip>.

- Becker, Sascha O.; Caliendo, Marco (2007): Sensitivity Analysis for Average Treatment Effects. *The Stata Journal* 7 (1), 71–83.
- Behrend, Olaf; Ludwig-Mayerhofer, Wolfgang; Sondermann, Ariadne; Hirsland, Andreas (2006): Reform der Arbeitsverwaltung: Im Schatten der Aufmerksamkeit – die Arbeitsvermittler. IAB-Kurzbericht 21/2006.
- Berliner Zeitung vom 3.1.2006: Ziele und Modelle.
- Bethwaite, Judy; Tompkinson, Paul (1996): The ultimatum game and non-selfish utility functions. *The Journal of Economic Psychology* 17, 259–271.
- Bielenski, Harald; Hartmann, Josef; Mauer, Andreas; Seifert, Hartmut (2002): Übergänge zu einem neuen Arbeitsplatz. *WSI Mitteilungen* 8/2002.
- Biewen, Martin; Fitzenberger, Bernd; Osikominu, Aderonke; Waller, Marie (2006): Employment Effects of Short and Medium Term Further Training Programs in Germany in the Early 2000s.
- Black, Dan; Smith, Jeff (2004): How robust is the evidence on the effects of the college quality? Evidence from matching. *Journal of Econometrics*, 121 (1), 99–124.
- Blau, Dawid M. (1991): Search for Nonwage Job Characteristics: A Test of the Reservation Wage Hypothesis. *Journal of Labor Economics* Vol. 9 (2), 186–205.
- Blau, Peter M. (2005): Sozialer Austausch. Adloff, Frank; Mau, Steffen (Hrsg.): Vom Geben und Nehmen. Zur Soziologie der Reziprozität, Frankfurt am Main: Campus, 125–138.
- Blien, Uwe; Hirschenauer, Franziska; Arendt, Manfred; Braun, Hans Jürgen; Gunst, Dieter-Michael; Kilcioglu, Sibel; Kleinschmidt, Helmut; Musati, Martina; Roß, Hermann; Vollkommer, Dieter; Wein, Jochen (2004): Typisierung von Bezirken der Agenturen für Arbeit. *Zeitschrift für ArbeitsmarktForschung*, 37 (2), 146–175.
- Blossfeld, Hans Peter; Rohwer, Götz (2002): Techniques in Event History Modeling. *New Approaches to Causal Analysis*. Second Edition. London.
- Blundell, Richard (2000): Work incentives and „in-work“ benefit Reforms: A Review. *Oxford Review of Economic Policy*, 16 (1), 27–44.
- Bode, Ingo; Brose, Hanns-Georg (1999): Die neuen Grenzen organisierter Reziprozität. *Berliner Journal für Soziologie* 2, 179–196.
- Bonin, Holger; Kempe, Wolfram; Schneider Hilmar (2002): Kombilohn oder Workfare? Zur Wirksamkeit zweier arbeitsmarktpolitischer Strategien. IZA Diskussion Paper No. 587.
- Bonß, Wolfgang; Ludwig-Mayerhofer, Wolfgang (2000): Arbeitsmarkt. Allmendiger, J.; Ludwig-Mayerhofer, W. (Hrsg.): *Soziologie des Sozialstaats. Grundsätzliche Grundlagen, historische Zusammenhänge und aktuelle Entwicklungstendenzen*. Weinheim und München.



- Bosch, Gerhard; Kalina, Thorsten (2005): Entwicklung und Struktur der Niedriglohnbeschäftigung in Deutschland. Institut Arbeit und Technik im Wissenschaftszentrum Nordrhein-Westfalen. Jahrbuch 2005. O.O., 29–46.
- Boudon, Raymond (1996): The „Cognitivist Model“. A Generalized „Rational-Choice Model“. *Rationality and Society* 8 (2), 123–150.
- Boudon, Raymond (1997): The present relevance of Weber's Wertrationalität (value rationality). Koslowski, Peter (Hrsg.): *Methodology of the social sciences, ethics, and economics in the newer historical school: from Max Weber and Rickert to Sombart and Rothacker*. Berlin/New York: Springer, 4–29.
- Boudon, Raymond (1998a): Limitations of Rational Choice Theory. *American Journal of Sociology* 104 (3), 817–828.
- Boudon, Raymond (1998b): Social mechanisms without black boxes. Hedström, Peter; Swedberg, Richard (Hrsg.): *Social mechanisms: an analytical approach to social theory*. Cambridge: Cambridge University Press, 172–203.
- Bowles, Samuel; Gintis, Herbert (2000): Reciprocity, Self-Interest and the Welfare State. *Nordic Journal of Political Economy*, 26 (4), 33–53.
- Brinkmann, Christian; Hartmann, Josef; Holleder, Alfons; Kalina, Thorsten; Kaltenborn, Bruno; Rudolph, Helmut; Vanselow, Achim; Weinkopf, Claudia; Wiedemann, Eberhard (2004): Drei Jahre Mainzer Modell. Eine Zwischenbilanz. 3. Zwischenbericht. Berlin: BMWA-Dokumentation Nr. 528.
- Brüderl, Josef (2004): Die Überprüfung von Rational-Choice-Modellen mit Umfragedaten. Dieckmann, Andreas; Voss, Thomas (Hrsg.): *Rational-Choice-Theorie in den Sozialwissenschaften. Anwendungen und Probleme*. München: R. Oldenbourg Verlag, 163–180.
- Bryson, Alex; Dorsett, Richard; Purdon, Susan (2002): The Use of Propensity Score Matching in the Evaluation of Labour Market Policies. Working Paper No. 4, Department for Work and Pensions.
- Bunge, Mario (1987): *Kausalität, Geschichte und Probleme*.
- Cahuc, Pierre; Zylberberg, André (2004): *Labor Economics*. Cambridge, Mass.: MIT Press.
- Caliendo, Marco; Hujer, Reinhard (2006): The Microeconometric Estimation of Treatment Effects – An Overview. *Allgemeines Statistisches Archiv*, 90 (1), 197–212.
- Caliendo, Marco; Hujer, Reinhard; Thomsen, Stephan L. (2004): Evaluation der Eingliederungseffekte von Arbeitsbeschaffungsmaßnahmen in reguläre Beschäftigung für Teilnehmer in Deutschland. ZEW Discussion Paper 04/46.
- Caliendo, Marco; Hujer, Reinhard; Thomsen, Stephan L. (2005a): Individual Employment Effects of Job Creation Schemes in Germany with Respect to Sectoral Heterogeneity. IAB-Discussion Paper 13/2005.

- Caliendo, Marco; Hujer, Reinhard; Thomsen, Stephan L. (2005b): Identifying Effect Heterogeneity to Improve the Efficiency of Job Creation Schemes in Germany. IAB-Discussion Paper 8/2005.
- Caliendo, Marco; Hujer, Reinhard; Thomsen, Stephan L. (2005c): The Employment Effects of Job Creation Schemes in Germany: A Microeconometric Evaluation. IZA Discussion Papers 1512.
- Caliendo, Marco; Kopeinig, Sabine (im Erscheinen): Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*.
- CDU, CSU und SPD (2005): Gemeinsam für Deutschland – mit Mut und Menschlichkeit. Koalitionsvertrag zwischen CDU, CSU und SPD auf Bundesebene für die 16. Legislaturperiode, 11. November.
- Chalmers, Alan F. (1999): *What ist his thing called science?* Indianapolis/Cambridge: Hackett Publishing Company, Inc.
- Chao-Ying, Joanne Peng (2003): Experiment. Lewis-Beck, Michael S.; Bryman, Alan; Futing Liao, Tim (Hrsg.): *The SAGE Encyclopedia of Social Science Research Methods*. Volume 1. Thousand Oaks et al.: SAGE Publications.
- Christensen, Björn (2003): Anspruchslohn und Arbeitslosigkeit in Deutschland. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 36 (4), 573–598.
- Clark, Andrew E.; Oswald, Andrew J. (1996): Satisfaction and Comparison Income. *Journal of Public Economics* 61, 359–381.
- Cochran, William G. (1957): Analysis of covariances: its nature and uses. *Biometrics* 13, 261–281.
- Cochran, William G. (1965): The Planning of Observational Studies of Human Populations. *Journal of the Royal Statistical Society, Series A* 182, 234–255.
- Cochran, William G.; Rubin, Donald B. (1973): Controlling Bias in Observational Studies: A Review. *Sankhya, Series A*, 35, 417–446.
- Cockx, Bart; Van der Linden, Bruno; Karaa, Adel (1998): Active labour market policies and job tenure. *Oxford Economic Papers* 50, 685–708.
- Collingwood, Robin G. (1940): *An Essay on Metaphysics*. Oxford: Clarendon Press.
- Conway, Neil; Briner Rob B. (2005): *Understanding Psychological Contracts at Work: A Critical Evaluation of Theory and Research*. New York: Oxford University Press.
- Cords, Torsten (2003): „Mainzer Modell“ – praktische Erfahrung eines Arbeitsamtes. Jahn, Elke; Wiedemann, Eberhard (Hrsg.): *Beschäftigungsförderung im Niedriglohnsektor*. Nürnberg, 113–118.
- Cox, David (1972): Regression models and life tabels. *Journal of the Royal Statistical Society Series B*, 34, 187–220.
- Cox, D. R.; Wermuth, Nanny (2000): Some statistical aspects of causality. ZUMA-Arbeitsbericht 00/07.

- Cox, D. R.; Wermuth, Nanny (2004): Causality: a statistical view. *International Statistical*.
- Crosby, Faye (1976): A model of egoistical relative deprivation. *Psychological Review* 83 (2), 85–113.
- Czommer, Lars; Kalina, Thorsten; Vanselow, Achim (2003): Beschäftigungsförderung im Niedriglohnbereich: keine Massenwirkung, aber eine individuelle Hilfe zu neuer Beschäftigung. Institut Arbeit und Technik: Jahrbuch 2002/2003. Gelsenkirchen, 37–46.
- D'Ambrosio, Conchita; Frick, Joachim R. (2004): Subjektive Well-Being and Relative Deprivation: an Empirical Link. *DIW Discussion Paper* 449.
- Dann, Sabine; Kirchmann, Andrea; Spermann, Alexander; Volkert, Jürgen (2002): Einstiegsgeld in Baden-Württemberg. Schlussbericht. Sozialministerium Baden-Württemberg, Stuttgart.
- Dawid, A. P. (1999): Who needs counterfactuals? Gammernan, A. (Hrsg.): *Causal Models and Intelligent Data Management*. Berlin: Springer Verlag, 33–50.
- Dawid, A. P. (1979): Conditional Independence in Statistical Theory. With Discussion. *Journal of the Royal Statistical Society, Series B*, 41 (1), 1–31.
- Deeke, Axel; Kruppe, Thomas (2003): Beschäftigungsfähigkeit als Evaluationsmaßstab? Inhaltliche und methodische Aspekte der Wirkungsanalyse beruflicher Weiterbildung im Rahmen des ESF-BA-Programms. IAB-Werkstattbericht Nr. 01/2003. Nürnberg.
- Dehejia, Rajeev H.; Wahba, Sadek (2002): Propensity Score-Matching Methods for Nonexperimental Causal Studies. *The Review of Economics and Statistics* 84 (1), 151–161.
- Dieckhoff, Martina (2007): Does it work? The Effect of Continuing Training on Labour Market Outcomes: A Comparative Study of Germany, Denmark and the UK. *European Sociological Review* 23 (3), 295–308.
- Dieckhoff, Martina (2007): Does it Work? The Effect of Continuing Training on Labour Market Outcomes: A Comparative Study of Germany, Denmark, and the United Kingdom. *European Sociological Review* 23, 295–308.
- Dietz, Martin; Koch, Susanne; Walwei, Ulrich (2006): Kombilohn: Ein Ansatz mit Haken und Ösen. IAB-Kurzbericht Nr. 3/2006, Nürnberg.
- DiPrete, Thomas A.; Engelhardt, Henriette (2004): Estimating Causal Effects With Matching Methods in the Presence and Absence of Bias Cancellation. *Sociological Methods & Research* 32, 501–528.
- Dyke, Andrew; Heinrich, Carolyn J.; Mueser, Peter R.; Troske, Kenneth R. (2005): The Effects of Welfare-to-Work Program Activities on Labor Market Outcomes. IZA Discussion Paper No. 1520.

- Eckel, Catherine C.; Grossman Philip J. (1996): Altruismus in Anonymus Dictator Games. *Games and Economic Behavior* 16, 181–191.
- Eckey, Hans-Friedrich; Kosfeld, Reinhold; Dreger, Christian (2001): *Ökonometrie, Grundlagen, Methoden, Beispiele*. Wiesbaden: Gabler-Verlag.
- Eichhorst, Werner (2006): Kombilöhne und Mindestlöhne als Instrumente der Beschäftigungspolitik – Erfahrungen und Handlungsoptionen. IZA Discussion Paper No. 2120.
- Eichhorst, Werner; Gartner, Hermann; Krug, Gerhard; Rhein, Thomas; Wiedemann, Eberhard (2005): Niedriglohnbeschäftigung in Deutschland und im internationalen Vergleich. Allmendinger, Jutta; Eichhorst, Werner; Walwei, Ulrich (Hrsg.): *IAB Handbuch Arbeitsmarkt: Analysen, Daten, Fakten*. Frankfurt am Main u. a.: Campus Verlag, 107–142.
- Elkeles, Thomas; Seifert, Wolfgang (1992): Arbeitslosigkeit und Gesundheit. Langzeitanalysen mit dem Sozio-Ökonomischen Panel. *Soziale Welt* 43, 278–300.
- Elster, Jon (1998): A Plea for Mechanisms. Hedström, Peter; Swedberg, Richard (Hrsg.): *Social mechanisms: an analytical approach to social theory*. Cambridge: Cambridge University Press, 45–73.
- Erlinghagen, Marcel; Knuth, Mathias (2004): Beschäftigungsstabilität in der Wissensgesellschaft. Struck, Olaf; Köhler, Christoph (Hrsg.): *Beschäftigungsstabilität im Wandel?* München und Mering: Rainer Hampp Verlag.
- Esping-Andersen, Gosta (1991): *The three worlds of welfare capitalism*. Cambridge: Polity Press.
- Evans William N.; Schwab, Robert M. (1995): Finishing High School and Starting College: Do Catholic Schools Make a Difference? *The Quarterly Journal of Economics*, 110 (4), 941–974.
- Falk, Armin (2003): Homo Oeconomicus versus Homo Reciprocans: Ansätze für ein neues wirtschaftspolitisches Leitbild? *Perspektiven der Wirtschaftspolitik* 4 (1), 141–172.
- Falk, Armin; Fischbacher, Urs (2000): A Theory of Reciprocity. Institute for Empirical Research in Economics. University of Zurich Working Paper Series No. 6.
- Falk, Armin; Fischbacher, Urs (2006): A Theory of Reciprocity. *Games and Economic Behavior* 54 (2), 293–315.
- Fehr, Ernst; Gächter, Simon (2000): Fairness and retaliation: The Economics of Reciprocity. *The Journal of Economic Perspectives* 14 (3), 159–181.
- Fehr, Ernst; Gächter, Simon (2002): Do incentive contracts undermine voluntary cooperation? Institute of Empirical Research in Economics University of Zurich Working Paper Series No. 34.
- Fehr, Ernst; Schmidt, Klaus M. (1999): A theory of fairness, competition, and Cooperation. *The Quarterly Journal of Economics* 114 (3), 817–868.

- Fetchenhauer, Detlef; Bierhoff, Hans-Werner (2004): Altruismus aus evolutions-theoretischer Perspektive. *Zeitschrift für Sozialpsychologie* 35 (3), 131–141.
- Firth, Raymond (1959): *Economics of the New Zealand Maori*. Wellington, New Zealand.
- Fitzenberger, Bernd; Osikominu, Aderonke; Völter, Robert (2006): Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany. IAB-Discussion Paper 17/2006.
- Fitzenberger, Bernd; Speckesser, Stefan (2000): Zur wissenschaftlichen Evaluation der Aktiven Arbeitsmarktpolitik in Deutschland: Ein Überblick. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 33 (3), 357–370.
- Frangakis, Constantine E.; Rubin, Donald B. (2002): Principal Stratification in Causal Inference. *Biometrics* 58, 21–29.
- Frankfurter Allgemeine Zeitung vom 3.1.2006: Der deutsche Kombilohn existiert bereits.
- Frankfurter Rundschau vom 2.1.2006: Anreize zum Ausbau des Niedriglohnssektors sind fragwürdig.
- Franz, Wolfgang (1999): *Arbeitsmarktökonomik*. Berlin u. a.: Springer.
- Franz, Wolfgang (2005): Will the (German) NAIRU Please Stand Up? *German Economic Review* 6 (2), 131–153.
- Frey, Bruno S. (1997): *Wie ökonomische Anreize die (Arbeits-)Moral verdrängen*. München: Vahlen.
- Frey, Bruno S.; Benz, Mathias (2001): *Ökonomie und Psychologie: eine Übersicht*. Institute for Empirical Research in Economics University of Zurich Working Paper No. 92.
- Frey, Bruno S.; Jegen, Reto (2000): Motivation crowding theory: A survey of empirical evidence. CESifo Working Paper Series No. 245.
- Frey, Bruno S.; Oberholzer-Gee, Felix (1997): The cost of price incentives: An empirical analysis of motivation. *The American Economic Review*, 87 (4), 746–755.
- Frey, Bruno S.; Goette, Lorenz (1999): Does Pay Motivate Volunteers? IEW Working Paper Series No. 007.
- Friedman, Milton (1953): *The Methodology of Positive Economics*. Friedman, Milton (Hrsg.): *Essays in Positive Economics*. Chicago: Chicago University Press, 3–43.
- Frölich, Markus (2002): *Semiparametric Estimation of Selectivity Models*. Hauptpage, New York: Nova Science Publishers.
- Gangl, Markus; DiPrete, Thomas A. (2004a): Kausalanalyse durch Matchingverfahren. Diekmann, Andreas (Hrsg.): *Methoden der Sozialforschung*. Sonderheft der Kölner Zeitschrift für Soziologie und Sozialpsychologie 44.

- Gangl, Markus; DiPrete, Thomas A. (2004b): Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimates and instrumental variables with imperfect instruments. *Sociological Methodology* 34, 271–310.
- Gerhardt, Michael; Meyer Larsen, Ingolf (2005): Das „Hamburger Modell“ zur Beschäftigungsförderung. Auswertungsbericht. Freie Hansestadt Hamburg/Bundesanstalt für Arbeit.
- Gerhardt, Michael; Prang, Oliver (2003): Das „Hamburger Modell“ zur Beschäftigungsförderung. Erste Untersuchungsergebnisse. Freie Hansestadt Hamburg/Bundesanstalt für Arbeit.
- Gerster, Florian; Deubel, Ingolf (2001): Das Mainzer Modell für Beschäftigung und Familienförderung. Gerster, Florian (Hrsg.): Arbeit muss sich lohnen. Baden-Baden: Nomos, 59–77.
- Gewiese, Tilo (2004): Das Mainzer Modell aus Perspektive der Betriebe – Die Betriebsbefragung im Rahmen der Evaluation zum Mainzer Modell. Gewiese, Tilo; Hartmann, Josef; Krug, Gerhard; Rudolph, Helmut: Das Mainzer Modell aus Sicht der Arbeitnehmer und Betriebe. Befunde aus der Begleitforschung. Berlin, 262–320. Abzurufen unter: <http://doku.iab.de/externe/2004/k040823w09.pdf>.
- Gneezy, Uri; Rustichini, Aldo (2000): Pay Enough Or Don't Pay At All. *The Quarterly Journal of Economics* 15 (3), 791–810.
- Goldthorpe, John H. (1998): Rational Action Theory for Sociology. *The British Journal of Sociology* 49 (2), 167–192.
- Goldthorpe, John H. (2001): Causation, Statistics, and Sociology. *European Sociological Review* 17, 1–20.
- Gouldner, Alvin W. (1960): The norm of reciprocity: A preliminary statement. *American Sociological Review* Jg. 25, Nr. 2, 161–178.
- Gouldner, Alvin W. (2005): Etwas gegen Nichts. Reziprozität und Asymmetrie. Adloff, Frank; Mau, Steffen (Hrsg.): Vom Geben und Nehmen. Zur Soziologie der Reziprozität. Frankfurt am Main: Campus, 109–124.
- Granovetter, Mark (1973): The Strength of Weak Ties. *American Journal of Sociology*, Vol. 78 (6), 1360–1380.
- Greene, William H. (2003): *Econometric Analysis*. Prentice Hall.
- Greve, Jens (2003): Handlungserklärung und die zwei Rationalitäten. Neuere Ansätze zur Integration von Wert- und Zweckrationalität in ein Handlungsmodell. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 55 (4), 621–653.
- Güth, Werner; Schmittberger, Rolf; Schwarze, Bernd (1982): An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization* 3 (4), 367–388.

- Gujarati, Damodar N. (1995): *Basic Econometrics*. New York u. a.: McGraw-Hill Inc.
- Hahn, Jinyong; Todd, Petra; Van der Klaauw, Wilbert (2001): Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica* 69 (1), 201–209.
- Hart, Oliver; Moore, John (1999): Foundations of Incomplete Contracts. *The Review of Economic Studies*, 66 (1), 115–138.
- Hartmann, Josef (2004): Repräsentative Erhebung zur Evaluation des Mainzer Modells. Gewiese, Tilo; Hartmann, Josef; Krug, Gerhard; Rudolph, Helmut: Das Mainzer Modell aus Sicht der Arbeitnehmer und Betriebe. Befunde aus der Begleitforschung, 49–139. Abzurufen unter: <http://doku.iab.de/externe/2004/k040823w09.pdf>.
- Hartmann, Josef; Krug, Gerhard (2004): Analysen zur Beschäftigungsstabilität und zur weiteren erwerbsbiographischen Entwicklung der nach dem Mainzer Modell Geförderten. Gewiese, Tilo; Hartmann, Josef; Krug, Gerhard; Rudolph, Helmut: Das Mainzer Modell aus Sicht der Arbeitnehmer und Betriebe. Befunde aus der Begleitforschung, 175–261. Abzurufen unter: <http://doku.iab.de/externe/2004/k040823w09.pdf>.
- Hartmann, Josef; Krug, Gerhard (2009): Verknüpfung von personenbezogenen Prozess- und Befragungsdaten – Selektivität durch fehlende Zustimmung der Befragten? *Zeitschrift für ArbeitsmarktForschung* 42 (2), 121–139.
- Heckman, James J. (1978): Dummy Endogenous Variables in a Simultaneous Equation System. *Econometrica* 46 (4), 931–959.
- Heckman, James J. (1979): Sample Selection Bias as a Specification Error. *Econometrica* 47 (1), 153–162.
- Heckman, James J. (2005a): The Scientific Model of Causality. *Sociological Methodology* 35 (1), 1–97.
- Heckman, James J. (2005b): Rejoinder: Response to Sobel. *Sociological Methodology* 35 (1), 1–97.
- Heckman, James J.; Hotz Joseph V. (1989): Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training. *Journal of the American Statistical Association* 84 (408), 862–874.
- Heckman, James J.; Ichimura, Hidehiko; Smith, Jeff; Todd, Petra (1998): Characterizing Selection Bias Using Experimental Data. *Econometrica*, 66 (5), 1017–1098.
- Heckman, James J.; Ichimura, Hidehiko; Todd, Petra E. (1997): Matching as an Econometric Evaluation Estimator. *Review of Economic Studies* 64, 605–654.
- Heckman, James J.; Ichimura, Hidehiko; Todd, Petra E. (1998): Matching as an Econometric Evaluation Estimator. *Review of Economic Studies* 65, 261–294.

- Heckman, James J.; LaLonde, R. J.; Smith, Jeffrey (1999): The Economics and Econometrics of Active Labor Market Programs. Ashenfelter, O.; Card, David (Hrsg.): Handbook of Labor Economics. Vol. 3a. Amsterdam: North Holland, 1865–2097.
- Heckman, James J.; Lochner, Lance; Taber, Christopher (1998): General Equilibrium Treatment Effects: A Study of Tax and Tuition Policy. NBER Working Paper No. 6426.
- Heckman, James J.; MaCurdy, Thomas E. (1986): Labor Econometrics. Griliches, Zvi: Handbook of Econometrics. Handbooks in Economics series, book 2 Amsterdam; Oxford and Tokyo: North-Holland, 1917–1977.
- Heckman, James J.; Navarro-Lozano, Salvador (2004): Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models. Review of Economics and Statistics 86 (1), 30–57.
- Heckman, James J.; Smith, Jeffrey (1996): Experimental and Non-Experimental Evaluation. Günther Schmid, et al. (Hrsg.): International Handbook of Labour Market Policy and Evaluation. Cheltenham/Brookfield: Edward Elgar Publishing, 37–88.
- Heckman, James; Tobias, Justin L.; Vytlacil, Edward (2003): Simple Estimators for treatment parameters in a latent-variable framework. Review of Economics and Statistics 85 (3), 748–755.
- Hedström, Peter; Swedberg, Richard (1998): Social mechanisms: An introductory essay. Hedström, Peter; Swedberg, Richard (Hrsg.): Social Mechanisms: An Analytical Approach to Social Theory. Cambridge: Cambridge University Press, 1–31.
- Heinemann, Sarah; Jaenichen, Ursula; Stephan, Gesine (2006): Eingliederungszuschüsse: Förderumfang, Strukturen, Effektivität. Bundesarbeitsblatt 3-2006.
- Herriott, Peter; Manning, W. E. G.; Kidd, Jennifer M. (1997): The content of the psychological contract. British Journal of Management 8, 151–162.
- Hieming, Bettina; Jaehrling, Karen; Kalina, Thorsten; Vanselow, Achim; Weinkopf, Claudia (2005): Stellenbesetzung im Bereich „einfacher“ Dienstleistungen: Abschlussbericht einer Studie im Auftrag des Bundesministeriums für Wirtschaft und Arbeit. Berlin: Bundesministerium für Wirtschaft und Arbeit. BMWA-Dokumentation Nr. 550.
- Hinz, Thomas; Wagner, Simone (2006): Gib und Nimm. Lokale Austauschnetzwerke zwischen sozialer Bewegung und Marktergänzung. Soziale Welt 57, 65–81.
- Hirschman, Albert O. (1970): Exit, Voice, and Loyalty. Responses to Decline in Firms, Organizations, and States. Cambridge: Harvard UP.
- Hohendanner, Christian (2007): Verdrängen Ein-Euro-Jobs sozialversicherungspflichtige Beschäftigung in den Betrieben? IAB-Discussion Paper 08/2007.



- Holland, Paul W. (1986): Statistics and Causal Inference. *Journal of the American Statistical Association* 81 (396), 945–960.
- Holland, Paul W. (1988): Causal inference, path analysis and recursive structural equation models. Clogg, Clifford (Hrsg.): *Sociological Methodology 1988*. Washington, D.C.: ASA, 449–484.
- Hollederer, Alfons; Rudolph, Helmut (2001): Bisher geringe Nachfrage im Niedriglohnsektor. Ein Jahr Erfahrungen mit dem Sonderprogramm CAST. *IAB-Materialien Nr. 3*, 14–15.
- Hollederer, Alfons; Rudolph, Helmut (2002): Arbeitsanreize und Niedriglöhne: Konzeptionen und erste Erfahrungen des Mainzer Modells und des SGI-Modells. Dann, Sabine; Kirchmann, Andrea; Spermann, Alexander; Volkert Jürgen (Hrsg.): *Kombi-Einkommen – Ein Weg aus der Sozialhilfe?* Baden-Baden: Nomos, 53–66.
- Hollstein, Betina (2005): Reziprozität in familialen Generationenbeziehungen. Adloff, Frank; Mau, Steffen (Hrsg.): *Vom Geben und Nehmen. Zur Soziologie der Reziprozität*. Frankfurt am Main: Campus, 187–210.
- Homans, George C. (1961): *Social Behaviour: Its Elementary Forms*. New York: Harcourt Brace Javanovich.
- Honneth, Axel (2003): *Kampf um Anerkennung*. Frankfurt am Main: Suhrkamp.
- Hujer, Reinhard; Caliendo, Marco (2003): Lohnsubventionen in Deutschland: Wie sieht eine optimale Evaluierungsstrategie aus? *Vierteljahreshefte zur Wirtschaftsforschung* 72, 109–123.
- Hujer, Reinhard; Wellner, Marc (2000): Berufliche Weiterbildung und individuelle Arbeitslosigkeitsdauer in West- und Ostdeutschland: Eine mikroökonomische Analyse. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 33 (3), 405–420.
- Hummel, Elisabeth; Jacobebbinghaus, Peter; Kohlmann, Annette; Oertel, Martina; Wübbecke, Christina; Ziegerer, Manfred (2005): Stichprobe der Integrierten Erwerbsbiografien – IEBS 1.0, Handbuch-Version 1.0.0. *FDZ Datenreport Nr. 06/2005*. Nürnberg.
- Imai, Kosuke; van Dyke, David A. (2004): Causal Inference with General Treatment Regimes: Generalizing the Propensity Score. *Journal of the American Statistical Association* 99 (467), 854–866.
- Imbens, Guido W. (2004): Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *The Review of Economics and Statistics*, Vol. 86 (1), 4–29.
- Jackofsky, Ellen F. (1984): Turnover and Job Performance: An Integrated Model. *Academy of Management Review* 9 (1), 74–83.

- Jaenichen, Ursula (2002a): Lohnkostenzuschüsse und individuelle Arbeitslosigkeit. Analysen auf der Grundlage kombinierter Erhebungs- und Prozessdaten unter Anwendung von Propensity Score Matching. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 35 (3), 327–351.
- Jaenichen, Ursula (2002b): Mikroevaluationen: Bildung von Vergleichsgruppen zur Schätzung individueller Förderwirkungen. Kleinhenz, Gerhard (Hrsg.): *IAB-Kompodium Arbeitsmarkt- und Berufsforschung. Beiträge zur Arbeitsmarkt- und Berufsforschung, BeitrAB 250*, 387–397.
- Jahoda, Marie (1981): Work, employment, and unemployment – values, theories, and approaches in social research. *American Psychologist*, 36 (2), 184–191.
- Jasso, Guillermina (1978): On the Justice of Earnings. A new Specification of the Justice Evaluation Function. *American Journal of Sociology* 83, 1398–1419.
- Jasso, Guillermina (1980): A new Theory of Distributive Justice. *American Sociological Review* 45, 3–32.
- Jirjahn, Uwe; Pfeiffer, Christian; Tsertsvadze, Georgi (2006): Mikroökonomische Beschäftigungseffekte des Hamburger Modells zur Beschäftigungsförderung. *IAB-Discussion Paper 25/2006*.
- Joffe, M. M.; Rosenbaum Paul R. (1999): Propensity Scores. *American Journal of Epidemiology* 150, 327–333.
- Jovanovic, Boyan (1979): Job matching and the theory of turnover. *Journal of Political Economy* 87 (5), 972–990.
- Kalina, Thorsten; Vanselow, Achim; Weinkopf, Claudia (2004): Umsetzung. Brinkmann, Christian; Hartmann, Josef; Holleder, Alfons; Kalina, Thorsten; Kaltenborn, Bruno; Rudolph, Helmut; Vanselow, Achim; Weinkopf, Claudia; Wiedemann, Eberhard: *Drei Jahre Mainzer Modell. Eine Zwischenbilanz. 3. Zwischenbericht*. Berlin: *BMWA-Dokumentation 528*.
- Kaltenborn, Bruno (2001): Kombilöhne in Deutschland. *IAB-Werkstattbericht Nr. 14*.
- Kaltenborn, Bruno; Krug, Gerhard; Rudolph, Helmut; Weinkopf, Claudia; Wiedemann, Eberhard (2005): Evaluierung der arbeitsmarktpolitischen Sonderprogramme CAST und Mainzer Modell. *Bundesministerium für Wirtschaft und Arbeit. Forschungsbericht Nr. 552*, Berlin.
- Kaltenborn, Bruno; Wielage, Nina (2005): Kombilöhne: Erfahrungen und Ausblick. *Blickpunkt Arbeit und Wirtschaft* 4/2005.
- Kangasharju, Aki (2005): Do Wage subsidies increase employment in subsidised Firms? *VATT Discussion Papers 378*.
- Kleinbeck, Uwe (1996): *Arbeitsmotivation: Entstehung, Wirkung und Förderung*. Weinheim und München: Juventa.

- Knabe, Andreas; Schöb, Ronnie; Weimann, Joachim (2006): Marginal Employment Subsidization: A New Concept and a Reappraisal. *Kyklos* 59 (4), 557–577.
- Knuth, Mathias (2004): Fragen und Herausforderungen an die sozialwissenschaftliche Arbeitsmarktforschung – Der Stachel der mikroökonomischen Evaluationsstudien (Podiumsbeitrag). Gensior, Sabine; Mendijs, Hans Gerhard; Seifert, Hartmut: 25 Jahre SAMF. Perspektiven Sozialwissenschaftlicher Arbeitsmarktforschung.
- Konow, James (1996): A positive theory of economic fairness. *Journal of Economic Behavior and Organization* 31 (1), 13–35.
- Kotthoff, Hermann (2000): Anerkennung und sozialer Tausch. Die soziale Konstruktion von Betriebsbürgerschaft. Holtgrewe, Ursula; Voswinkel, Stephan; Wagner, Gabriele (Hrsg.): Anerkennung und Arbeit. Konstanz: Universitätsverlag Konstanz, 27–36.
- Krug, Gerhard; Rudolph, Helmut (2004): Multivariate Analysen zur Zielerreichung des Mainzer Modells. Forschungsverbund Evaluierung Mainzer Modell (Hrsg.): Das Mainzer Modell aus Sicht der Arbeitnehmer und Betriebe. Befunde aus der Begleitforschung, 140–174. Abzurufen unter: <http://doku.iab.de/externe/2004/k040823w09.pdf>.
- Krug, Gerhard; Rudolph, Helmut; Holleder, Alfons (2004): Förderstrukturen im Mainzer Modell. Forschungsverbund Evaluierung Mainzer Modell (Hrsg.): Das Mainzer Modell aus Sicht der Arbeitnehmer und Betriebe. Befunde aus der Begleitforschung, 17–48. Abzurufen unter: <http://doku.iab.de/externe/2004/k040823w09.pdf>.
- Kruppe, Thomas (Hrsg.) (2006): Private Vermittlung als Unterstützung – eine Evaluation von Vermittlungsgutscheinen und Beauftragungen Dritter. Beiträge zur Arbeitsmarkt- und Berufsforschung/BeitrAB 301, Nürnberg.
- Layard, Richard (1997): Preventing long-term unemployment: an economic analysis. Snower, Dennis J.; de la Dehesa, Guillermo (Hrsg.): Unemployment policy. Government options for the labour market. Cambridge: Cambridge University Press, 333–349.
- Layard, Richard; Nickel, Stephen; Jackman, Richard (1995): *The Unemployment Crisis*. Oxford University Press: Oxford.
- Lazarsfeld, Paul F. (1976): Die Interpretation statistischer Beziehungen als Forschungsoperation. Hummel, Hans J.; Ziegler, Rolf (Hrsg.): Korrelation und Kausalität. Band 1. Stuttgart: Ferdinand Enke Verlag, 1–15.
- Lechner, Michael (2001): Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In Lechner, Michael; Pfeiffer, F. (Hrsg.): *Econometric Evaluation of Labour Market Policies*. Physica-Verlag: Heidelberg, 1–18.

- Lechner, Michael (2002): Mikroökonometrische Evaluation arbeitsmarktpolitischer Massnahmen. University of St. Gallen Department of Economics working paper series 2002–20, Department of Economics, University of St. Gallen.
- Lechner, Michael; Miquel Ruth; Wunsch, Conny (2005a): Long-Run Effects of Public Sector Sponsored Training in West Germany. IAB-Discussion Paper 3/2005.
- Lechner, Michael; Miquel Ruth; Wunsch, Conny (2005b): The Curse and Blessing of Training the Unemployed in a Changing Economy. The Case of East Germany After Unification. IAB-Discussion Paper 14/2005.
- Lechner, Michael; Smith, Jeffrey (2003): What is the Value Added by Caseworkers? University of St. Gallen Department of Economics working paper series 2003–05, Department of Economics, University of St. Gallen.
- Lechner, Michael; Wunsch, Conny (2006a): Are Training Programs more effective when unemployment is high? IZA discussion paper No. 2355.
- Lechner, Michael; Wunsch, Conny (2006b): Active Labour Market Policy in East Germany. Waiting for the economy to take off. IAB-Discussion Paper 20/2006.
- Lengfeld, Holger; Krause; Alexandra (2006): Wann gilt der Arbeitsmarkt als sozial gerecht? Der Einfluss des Unternehmenskontexts auf die Akzeptanz ertragsabhängiger Entlohnung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 58 (1), 98–116.
- Lessenich, Stephan; Mau, Steffen (2005): Reziprozität und Wohlfahrtsstaat. Adloff, Frank; Mau, Steffen (Hrsg.): Vom Geben und Nehmen. Zur Soziologie der Reziprozität. Frankfurt am Main: Campus, 257–276.
- Lester, Scott W.; Turnley, William H.; Bloodgood, James M.; Bolino, Mark C. (2002): Nor seeing eye to eye: differences in supervisor and subordinate perceptions of and attributions for psychological contract breach. *Journal of Organizational Behavior* 23 (1), 39–56.
- Leuven, E.; Sianesi, Barbara (2003): PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. Software, <http://ideas.repec.org/c/boc/bocode/s432001.html>.
- Liebig, Stefan (1997): Soziale Gerechtigkeitsforschung und Gerechtigkeit in Unternehmen. München und Mering: Reiner Hampp Verlag.
- Liebig, Stefan; Schupp, Jürgen (2005): Empfinden die Erwerbstätigen in Deutschland ihre Einkommen als gerecht? DIW Wochenbericht 48/2005.
- Lindenberg, Siegwart (1998): The influence of simplification on explananda: phenomenon-centered versus choice-centered theories in the social sciences. Blossfeld, Hans-Peter; Prein, Gerald (Hrsg.): Rational Choice Theory and Large-Scale Data Analysis. Boulder, San Francisco, Oxford: Westview Press, 54–69.

- Little, Daniel (1991): *Varieties of Social Explanation: An Introduction to the Philosophy of Social Science*. Summertown, Oxford: Westview Press.
- Little, Roderick J.; Rubin, Donald R. (2000): *Causal Effects in Clinical Epidemiological Studies Via Potential Outcomes: Concepts and Analytical Approaches*. *Annual Review of Public Health* 21, 121–145.
- Malinowski, Bronislaw (2001): *Argonauten des westlichen Pazifik*. Klotz: Eschborn.
- Manski, Charles F. (1995): *Identification Problems in the Social Sciences*. Harvard.
- Mantel, N.; Haenszel, W. (1959): *Statistical aspects of the analysis of data from retrospective studies*. *Journal of the National Cancer Institute* 22, 719–748.
- Mauss, Marcel (1990): *Die Gabe. Form und Funktion des Austausches in archaischen Gesellschaften*. Frankfurt am Main: Suhrkamp.
- McCall, John J. (1970): *Economics of Information and Job Search*. *Quarterly Journal of Economics* 84 (1), 113–126.
- Meyer, Bruce D. (1995): *Natural and Quasi-Experimental Experiments in Economics*. *Journal of Business and Economic Statistics* 13 (2), 151–161.
- Mills, C. Wright (1951): *White Collars: The American Middle Classes*. New York: Oxford University Press.
- Mincer, Jacob (1974): *Schooling, experience and earnings*. New York: Columbia University Press.
- Molm, Linda D.; Peterson, Gretchen; Takahashi, Nobuyuki (2001): *The Value of Exchange*. *Social Forces*, 80 (1), 159–184.
- Morgan, Stephen L.; Harding, David J. (2006): *Matching estimators of causal effects: Prospects and Pitfalls in theory and practice*. *Sociological Methods & Research*, 35 (1), 3–60.
- Morgenroth, Christine (1990): *Sprachloser Widerstand. Zur Sozialpathologie der Lebenswelt von Arbeitslosen*. Frankfurt/M.
- Morgenroth, Christine (2003): *Arbeitsidentität und Arbeitslosigkeit – ein depressiver Zirkel*. *Aus Politik und Zeitgeschichte* B6–7.
- Morrison, Elisabeth Wolfe; Robinson, Sandra L. (1997): *When employees feel betrayed: a model of how psychological contract violations develops*. *Academy of Management Review*, 22 (1), 226–256.
- Morrison, Elisabeth Wolfe; Robinson, Sandra L. (2004): *The employment relationship from two sides: Incongruence in employees' and employers' perception of obligations*. Coyle-Shapiro, Jaqueline A.-M.; Shore, Lynn M.; Taylor, M. Susan; Tetrick, Lois E. (Hrsg.): *The employment relationship. Examining psychological and contextual perspectives*. Oxford: Oxford University Press, 161–180.

- Mortensen, Dale T. (1970): A Theory of Wage and Employment Dynamics. Phelps, E. S. et al. (Hrsg.): *Macroeconomic Foundations of Employment and Inflation Theory*. New York: W. W. Norton, 124–166.
- Mueser, Peter; Troske, Kenneth R.; Gorislavsky, Alexey (2003): Using State Administrative Data to Measure Program Performance. IZA Discussion Papers No. 786.
- Müller, Oleg (2005): *Evaluation multimedialer Teachware mit Kontrollgruppe*. Aachen: Shaker Verlag.
- Müntnich, Michael; Wießner, Frank (2002): Soziale Experimente und Modellversuche: Ein Beitrag zur Evaluation von Neuansätzen in der Arbeitsmarktpolitik. Kleinhenz, Gerhard (Hrsg.): *IAB-Kompodium Arbeitsmarkt- und Berufsforschung*. Beiträge zur Arbeitsmarkt- und Berufsforschung, Bd. 250, Nürnberg, 415–427.
- Neal, Derek (1997): The Effects of Catholic Secondary Schooling on Educational Achievement. *Journal of Labor Economics*, 15 (1), Part 1, 98–123.
- Neckel, Sighard; Dröge, Kai (2002): Die Verdienste und ihr Preis: Leistung in der Marktgesellschaft. Honneth, Axel (Hrsg.): *Befreiung aus der Mündigkeit. Paradoxien des gegenwärtigen Kapitalismus*. Frankfurt/New York: Campus, 93–116.
- Nelson, Margaret N. (2000): Single Mothers and Social Support: The Commitment to, and Retreat from, Reciprocity. *Qualitative Sociology* 23 (3), 291–317.
- Noll, Heinz-Herbert; Weick, Stefan (2002): Informelle Kontakte für Zugang zu Jobs wichtiger als Arbeitsvermittlung. *Informationsdienst Soziale Indikatoren* 28, 6–10.
- Nowak, Stefan (1960): Some Problems of Causal Interpretation of Statistical Relationships. *Philosophy of Science*, 27, 23–38.
- Nürnberger Zeitung vom 3.1.2006: Rettet der Kombilohn den Arbeitsmarkt? Einfaches Modell – schwierige Bewertung (Isabel Strohschein).
- Ockenfels, Axel (1999): *Fairness, Reziprozität und Eigennutz: Ökonomische Theorie und experimentelle Evidenz*. Tübingen: Mohr Siebeck.
- OECD (1993): *Employment Outlook*. Paris.
- Opp, Karl-Dieter (2004): Die Theorie rationalen Handelns im Vergleich mit alternativen Theorien. Gabriel, Manfred (Hrsg.): *Paradigmen akteurszentrierter Soziologie*. Wiesbaden: VS Verlag für Sozialwissenschaften, 43–68.
- Pfeifer, Christian (2007): Homogene und heterogene Teilnahmeeffekte des Hamburger Kombilohnmodells. Ein Verfahrensvergleich von Propensity Score Matching und OLS-Regression. IAB-Discussion Paper 22/2007.
- Polanyi, Karl (2002): *The Great Transformation*. Frankfurt/M.: Suhrkamp.
- Porter, Lyman W.; Pearce, Jone L.; Tripoli, Angela M.; Lewis, Tristi M. (1998): Differential perceptions of employers' inducements: implications for psychological contracts. *Journal of Organizational Behavior* 19, 769–782.

- Puhatni, Patrick A. (2000): The Heckman correction for sample selection and its critique. *Journal of economic surveys* 14 (1), 53–69.
- Quandt, Richard E. (1972): A New Approach to Estimating Switching Regressions. *Journal of the American Statistical Association* 67 (338), 306–310.
- Ramb, Bernd-Thomas (1993): Die allgemeine Logik menschlichen Handelns. Ramb, Bernd-Thomas; Tietzel, Manfred (Hrsg.): *Ökonomische Verhaltenstheorie*. München: Vahlen.
- Reinberg, Alexander; Hummel, Markus (2007): Qualifikationsspezifische Arbeitslosigkeit im Jahr 2005 und die Einführung der Hartz-IV-Reform. Empirische Befunde und methodische Probleme. IAB-Forschungsbericht 09/2007.
- Reinowski, Eva; Schultz, Birgit (2006): Microeconomic Evaluation of Selected ESF-funded ALMP-Programmes. IWH-Diskussionspapiere No. 17.
- Reinowski, Eva; Schultz, Birgit; Wiemers, Jürgen (2003): Evaluation von Maßnahmen der aktiven Arbeitsmarktpolitik mit Hilfe eines iterativen Matching-Algorithmus. Eine Fallstudie über langzeitarbeitslose Maßnahmeteilnehmer in Sachsen. IWH Diskussionspapiere Nr. 173.
- Rhein, Thomas; Gartner, Hermann; Krug, Gerhard (2005): Niedriglohnsektor: Aufstiegschancen für Geringverdiener verschlechtert. IAB-Kurzbericht 03/2005.
- Robinson, Sandra L. (1995): Violation of psychological contracts: impact on employee attitudes. Tetrick, Lois E.; Barling J. (Hrsg.): *Changing employment relations. Behavioral and social perspectives*. Washington, DC: American Psychological Association, 91–108.
- Robinson, Peter (2000): Active labour-market policies: a case of evidence-based policy-making? *Oxford Review of Economic Policy* 16 (1), 13–26.
- Rosenbaum, Paul R. (1984a): The Consequences of Adjustment for a Concomitant Variable That Has Been Affected by the Treatment. *Journal of the Royal Statistical Society, Series A*, 147, 656–666.
- Rosenbaum, Paul R. (1984b): From Association to Causation in Observational Studies: The Role of Tests of Strongly Ignorable Treatment Assignment. *Journal of the American Statistical Association*, 79 (385), 41–48.
- Rosenbaum, Paul, R. (2002): *Observational Studies*. New York u. a.: Springer.
- Rosenbaum, Paul R.; Rubin, Donald B. (1983): The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 76, 41–55.
- Rosenbaum, Paul R.; Rubin, Donald B. (1984): Estimating the Effects Caused by Treatments. Comment to J. W. Pratt and Robert Schlaifer. *Journal of the American Statistical Association* 79 (385), 27–28.
- Rosenbaum, Paul R.; Rubin, Donald B. (1985): Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score. *The American Statistician* 39 (1), 33–38.

- Rousseau, Denise M. (1989): Psychological and implied contracts. *Employee Responsibilities and Rights Journal*, 2 (2), 121–139.
- Rubin, Donald B. (1973): The Use of Matched Sampling and Regression Adjustment to Remove Bias in Observational Studies. *Biometrics*, 29 (1), 185–203.
- Rubin, Donald B. (1974): Estimating causal effects of treatment in randomized and nonrandomized studies. *Journal of Educational Studies* 66, 688–701.
- Rubin, Donald B. (1978): Bayesian Inference for Causal Effects: The Role of Randomization. *Annals of Statistics* 6, 34–58.
- Rubin, Donald B. (1980a): Discussion of paper by D. Basu. *Journal of the American Statistical Association* 75, 591–593.
- Rubin, Donald B. (1980b): Bias Reduction Using Mahalanobis-Metric Matching. *Biometrics* 36, 293–298.
- Rubin, Donald B. (1986): Which Ifs Have Causal Answers. Comment to Holland 1986. *Journal of the American Statistical Association* 81 (396), 961–962.
- Rubin, Donald B. (1991): Practical Implications of Modes of Statistical Inferences for Causal Effects and the Critical Role of the Assignment Mechanism. *Biometrics* 47, 1213–1234.
- Rubin, Donald B. (2001): Using Propensity Scores to Help Design Observational Studies: Application to Tobacco Litigation. *Health Services & Outcomes Research Methodology* 2, 169–188.
- Rubin, Donald B. (2005): Causal inference using potential outcomes: Design, Modeling, Decisions. *Journal of the American Statistical Association* Vol. 100, Nr. 469, 322–331.
- Rubin, Donald B. (2006): *Matched Sampling for Causal Effects*. Cambridge University Press.
- Rubin, Donald B. (2007): The design versus the analysis of observational studies for causal effects: Parallels with the design of randomized trials. *Statistics in Medicine* 26, 20–36.
- Rubin, Donald B.; Thomas, Neil (1996): Matching Using Estimated Propensity Scores: Relating Theory to Praxis. *Biometrics* 52, 249–264.
- Sachverständigenrat (2005/06): Die Chance nutzen – Reformen mutig voranbringen. Jahresgutachten.
- Sahlins, Marshall D. (1999): Zur Soziologie des Primitiven Tauschs. *Berliner Journal für Soziologie*, 9 (2), 149–178.
- Salmon, Wesley (1984): *Scientific Explanation and the Causal Structure of the World*. Princeton.
- Scharpf, Fritz (1994): Die negative Einkommenssteuer – Ein beschäftigungspolitisches Instrument? *Wirtschaftsdienst* 1994/3, 111–114.



- Schmid, Günther (2003): Moderne Dienstleistungen am Arbeitsmarkt: Strategie und Vorschläge der Hartz-Kommission. Aus Politik und Zeitgeschichte B 6-7, 3-6.
- Schmidt, Christoph M. (2000): Arbeitsmarktpolitische Maßnahmen und ihre Evaluierung: eine Bestandsaufnahme. IZA Discussion Paper No. 207, Bonn.
- Schneider, Hilmar (2002): Kombi-Einkommen: Arbeitsangebotseffekte mit und ohne Zielgruppenbegrenzung. Dann, Sabine; Kirchmann, Andrea; Spermann, Alexander; Volkert, Jürgen (Hrsg.): Kombi-Einkommen – Ein Weg aus der Sozialhilfe? Baden-Baden: Nomos, 33-52.
- Schöb, Ronnie; Weimann, Joachim (2006): Kombilohn und Mindestlohn: Das kleine Steuerüberwälzungseinmaleins. Wirtschaftsdienst 2, 102-104.
- Schumacher, Jörg; Laubach, Wilfried; Brähler, Elmar (1995): Wie zufrieden sind wir mit unserem Leben? Soziodemographische und psychologische Prädiktoren der allgemeinen und bereichsspezifischen Lebenszufriedenheit. Zeitschrift für Medizinische Psychologie 4, 17-26.
- Schütz, Alfred (1974): Der sinnhafte Aufbau der Welt. Eine Einleitung in die verstehende Soziologie. Frankfurt/M.: Suhrkamp.
- Schütz, Alfred; Luckmann, Thomas (2003): Strukturen der Lebenswelt. Konstanz: UVK.
- Sesselmeier, Werner (1999): Mikroökonomische Theorien der Arbeitslosigkeit. Ökonomie und Gesellschaft: Jahrbuch 15. Unternehmungsverhalten und Arbeitslosigkeit.
- Sesselmeier, Werner; Blauermel, Gregor (1997): Arbeitsmarkttheorien, 2. Aufl.
- Shadish, William R.; Cook, Thomas D.; Campbell, Donald T. (2002): Experimental & Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin.
- Shapiro, Carl; Stiglitz Joseph E. (1984): Equilibrium Unemployment as a Worker Discipline Device. American Economic Review 74 (3), 433-444.
- Sheppard B. H.; Lewicki, R. J.; Minton, J. W. (1992): Organizational justice: The search for fairness in the workplace. New York, NY: Macmillan, Inc.
- Simon, Herbert A. (1976): Scheinkorrelation: ihre kausale Interpretation. Hummel, Hans J.; Ziegler, Rolf (Hrsg.): Korrelation und Kausalität. Band 1. Stuttgart: Ferdinand Enke Verlag, 55-67.
- Sinn, Hans-Werner; Hozner, Christian; Meister, Wolfgang; Ochel, Wolfgang; Werdinger, Martin (2006): Aktivierende Sozialhilfe 2006 – das Kombilohn-Modell des ifo Instituts. ifo-Schnelldienst 2/2006, 49-51.
- Smith, Jeff A.; Todd, Petra E. (2005): Does matching overcome LaLonde's critique of nonexperimental estimators? Journal of Econometrics Vol. 125 (1-2), 305-353.
- Snower, Denis J. (1997): Evaluating unemployment policies: what do the underlying theories tell us? Snower, Dennis J.; de la Dehesa, Guillermo (Hrsg.): Unemployment policy. Government options for the labour market. Cambridge, 15-53.

- Sobel, Joel (2005): Independent preferences and reciprocity. *Journal of Economic Literature* XLIII, 2, 392–436.
- Sobel, Michael E. (1995): *Causal Inference in the Social and Behavioral Sciences*. Armingier, Gerhard; Clifford C. Clogg und Michael E. Sobel (Hrsg.): *Handbook of Statistical Modelling for the Social and Behavioral Sciences*. New York/London: Plenum, 1–38.
- Solow, Robert M. (1971): Blood and Thunder. *Yale Law Journal* 80, 170–183.
- Spermann, Alexander (2002): Das Mainzer Kombilohn-Modell zur Verbesserung von Arbeitsanreizen – ein Placebo für gering Qualifizierte? *Zeitschrift für Wirtschaftspolitik* 51 (2), 89–198.
- Steel, Daniel (2004): Social Mechanisms and Causal Inference. *Philosophy of the Social Sciences*, 34 (1), 55–78.
- Stephan, Gesine (1996): Grundlagen des europäischen Arbeitsmarktmodells. *Wirtschaftswissenschaftliches Studium* 25, 238–242.
- Stephan, Gesine; Rässler, Susanne; Schewe, Torben (2006): Das TrEffeR-Projekt der Bundesagentur für Arbeit. Die Wirkung von Maßnahmen aktiver Arbeitsmarktpolitik. *Zeitschrift für ArbeitsmarktForschung* 39 (3/4), 447–465.
- Stouffer, S. A.; Suchman, E. A.; DeVinney, L. C.; Star, S. A.; Williams, R. M. (1949): *The american soldier: Adjustment during army life* (Vol. 1). Princeton, NJ: Princeton University Press.
- Tenbrusel, Ann E.; Messick, David M. (1999): Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly*, 44 (4), 684–707.
- Titmuss, Richard M. (1997): *The Gift Relationship. From Human Blood to Social Policy*. Original edition with new chapters. Herausgegeben von Oakeley, Ann; Ashton, John. New York: New York Press.
- Verbeek, Marno (2004): *A Guide to Modern Econometrics*. Chichester: John Wiley & Sons.
- Vierling, Michael (2002): Arbeitsanreize für gering Qualifizierte: Anforderungen an das Transfersystem. *Zeitschrift für Wirtschaftspolitik* 51 (3), 376–395.
- von Auer, Ludwig (2003): *Ökonometrie. Eine Einführung*. Berlin: Springer.
- Voswinkel, Stephan (2002): Bewunderung ohne Würdigung? Paradoxien der Anerkennung doppelt subjektiver Arbeit. Honneth, Axel (Hrsg.): *Befreiung aus der Mündigkeit. Paradoxien des gegenwärtigen Kapitalismus*. Frankfurt/New York: Campus, 65–92.
- Voswinkel, Stephan (2005): Reziprozität und Anerkennung in Arbeitsbeziehungen. Adloff, Frank; Mau, Steffen (Hrsg.): *Vom Geben und Nehmen. Zur Soziologie der Reziprozität*. Frankfurt am Main: Campus.
- Walwei, Ulrich (2002): Kombilohn: Ein Weg zu mehr Beschäftigung? *Wirtschaftsdienst* 82 (2), 82–91.

- Warr, Peter (1987): *Work, Unemployment and Mental Health*. Clarendon Press: Oxford.
- Weber, Max (1922/1976): *Wirtschaft und Gesellschaft*. Tübingen.
- Weesie, Jeroen (1999): sg121: Seemingly unrelated estimation and the cluster-adjusted sandwich estimator. *Stata Technical Bulletin* 52, 34–47.
- Werner, Heinz (1999): Niedriglohnssektor in den USA: Der „Earned Income Tax Credit“ soll Armut bei Arbeit lindern. *IAB-Kurzbericht* 12/1999.
- Whitehouse, Edward (1996): Designing and Implementing In-Work Benefits. *The Economic Journal* 106, 130–141.
- Winterhager, Henrik; Heinze, Anja; Spermann, Alexander (2006): Deregulating Job Placement in Europe: A Microeconomic Evaluation of an Innovative Voucher Scheme in Germany. *Labour Economics* 13, 505–517.
- Wooldridge, Jeffrey (2002): *Econometric Analysis of Cross Section and Panel Data*. B&T.
- Wunder, Christoph (2005): Arbeitslosigkeit und Alterssicherung. Der Einfluss früherer Arbeitslosigkeit auf die Höhe der gesetzlichen Altersrente. *Zeitschrift für ArbeitsmarktForschung*, 8 (4), 493–509.
- Yatchew, Adonis; Griliches, Zvi (1985): Specification Error in Probit Models. *The Review of Economics and Statistics*, 67 (1), 134–39.
- Yule, G. Udny (1899): An investigation into the causes of changes in Pauperism in England, chiefly during the last two intercensal decades. *Journal of the Royal Statistical Society* 62, 249–295.
- Ziegler, Frank (2002): Misserfolg des Mainzer Modells. *Zeitschrift für Sozialhilfe und Sozialgesetzbuch (ZfSH/SGB)* 9/2002, 527–532.



# Anhang

## Notation

$$y_i = \alpha_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \dots + \beta_k x_{ik} + \varepsilon_i \quad i = 1, \dots, n$$

Kleinbuchstaben unfett: Variablen

$$y_i = \alpha_0 + \mathbf{x}_i' \boldsymbol{\beta} + \varepsilon_i \quad i = 1, \dots, n$$

Kleinbuchstaben fett: Vektoren

$E(y_i | x_i)$ : Erwartungswert der Variable  $y$  bei Person  $i$  an der Stelle  $x_i$  der Variable  $x$

$E(y_i | \mathbf{x}_i)$ : Erwartungswert der Variable  $y$  bei Person  $i$  an den Stellen  $\mathbf{x}_i$  des Vektors  $\mathbf{x}_i$

## Wichtige Abkürzungen

- MKE: Mittlerer kausaler Effekt
- T: Treatmentgruppe/Teilnehmer/Geförderte
- NT: Non-Treatmentgruppe/Nichtteilnehmer/Ungeförderte
- Iz: Lohnzufriedenheit
- CIA: Conditional Independence Assumption
- R: Referenzkategorie
- SUTVA: Stable Unit Treatment Value Assumption

## Tabellenanhang

Tabelle A1: Einfluss der Betriebsgröße auf die Kombilohnförderung unter Kontrolle weiterer Selektionseinflüsse

	Odds Ratio	p	95 %-Konfidenzintervall	
Kleinstbetriebe	1,061	0,257	0,958	1,175
Geschlecht weiblich	1,644	0,000	1,493	1,811
Haushaltskontext (R: Alleinlebend)				
Zusammen mit mind. einer Person	1,678	0,000	1,415	1,989
Verheiratet	1,293	0,000	1,129	1,481
Angabe fehlt	0,495	0,515	0,059	4,116
Zahl der Kinder (R: keine)				
Ein Kind	1,949	0,000	1,711	2,220
Zwei Kinder	2,377	0,000	2,074	2,724
Drei Kinder	3,015	0,000	2,479	3,666
Vier und mehr Kinder	5,703	0,000	4,224	7,700
Alter	0,998	0,422	0,993	1,003
Arbeitszeit wie gewünscht	0,490	0,000	0,442	0,543
Beruf wie gewünscht	0,687	0,000	0,624	0,757
Wunsch nach Befristung	0,186	0,101	0,025	1,390
Zahl der Stellenangebote (R: keine)				
1–5 Angebote	1,169	0,065	0,990	1,380
6–10 Angebote	1,409	0,000	1,181	1,680
11–20 Angebote	1,804	0,000	1,504	2,164
Mehr als 20 Angebote	2,015	0,000	1,635	2,484
Kumulierte Tage in Arbeitslosigkeit in Erwerbsbiografie	1,000	0,001	1,000	1,001
Ausbildung (R: Volks-/Hauptschule/ Mittlere Reife ohne Berufsausbildung)				
Volks-/Hauptschule/Mittlere Reife mit Berufsausbildung	1,181	0,009	1,043	1,339
Abitur ohne Berufsausbildung	0,822	0,644	0,359	1,885
Abitur mit Berufsausbildung	0,791	0,183	0,560	1,117
Abschluss FHS	0,406	0,003	0,224	0,736
Hochschul-/Uniabschluss	0,497	0,014	0,284	0,870
Unbekannt	1,248	0,000	1,128	1,380
Nicht deutsche Staatsbürgerschaft	1,079	0,416	0,898	1,297
Qualifikation (R: Volks-/Hauptschule/ Mittlere Reife ohne Berufsausbildung)				
Volks-/Hauptschule/Mittlere Reife mit Berufsausbildung	0,743	0,623	0,228	2,428
Abitur ohne Berufsausbildung	0,370	0,113	0,108	1,264
Abitur mit Berufsausbildung	0,792	0,697	0,246	2,551
Fachhochschule	1,225	0,722	0,400	3,756
Universität	1,455	0,511	0,475	4,459

	Odds Ratio	p	95 %-Konfidenzintervall	
Wirtschaftszweige nach NACE 2003 (R: Wirtschaftszweig 0)				
Wirtschaftszweig 1	2,248	0,000	1,580	3,199
Wirtschaftszweig 2	1,324	0,092	0,955	1,835
Wirtschaftszweig 3	1,560	0,016	1,085	2,244
Wirtschaftszweig 4	0,730	0,074	0,517	1,031
Wirtschaftszweig 5	2,556	0,000	1,911	3,418
Wirtschaftszweig 6	2,554	0,000	1,867	3,493
Wirtschaftszweig 7	3,348	0,000	2,520	4,449
Wirtschaftszweig 8	1,488	0,011	1,096	2,020
Wirtschaftszweig 9	2,637	0,000	1,913	3,634
Kumulierte Tage in Beschäftigung in den letzten 7 Jahren	1,000	0,000	1,000	1,000
Kumulierte Tage in Maßnahmen in den letzten 7 Jahren	1,000	0,144	1,000	1,000
Art des Endes der letzten Beschäftigung (R: Arbeitgeber)				
Arbeitnehmer	0,956	0,702	0,758	1,205
Aufhebungsvertrag	1,115	0,544	0,784	1,586
Befristung	0,895	0,078	0,790	1,012
Ungekündigt	1,189	0,500	0,719	1,967
Sonstiges	1,563	0,031	1,043	2,341
Unbekannt	1,070	0,187	0,968	1,183
Lohn in letzter Beschäftigung	0,996	0,000	0,994	0,998
Gesundheitliche Einschränkungen mit Auswirkung auf die Vermittlung (R: nein)	1,221	0,012	1,045	1,426
Berufserfahrung laut Arbeitsvermittler (R: ja)	1,080	0,214	0,957	1,219
Regionaler Arbeitsmarktkontext* (R: Typ 1a)				
Typ 1b	1,039	0,858	0,682	1,584
Typ 1c	1,006	0,977	0,653	1,550
Typ 2a	0,948	0,799	0,630	1,427
Typ 2b	0,968	0,887	0,615	1,523
Typ 3a	1,004	0,988	0,631	1,596
Typ 3b	1,334	0,269	0,801	2,220
Typ 3c	0,847	0,432	0,559	1,282
Typ 4	0,936	0,748	0,625	1,401
Typ 5a	0,863	0,494	0,567	1,315
Typ 5b	0,858	0,473	0,564	1,304
Typ 5c	1,091	0,689	0,711	1,675

\* Zur inhaltlichen Abgrenzung der Typen siehe Blien et al. 2004.

Tabelle A2: Kausaler Effekt bei Verwendung des vollen Variablenspektrums

	Effekte	Standardfehler	t-Wert	Zahl der Matches
Vor dem Matching	-0,228	0,031	-7,31	-
<b>Ein nächster Nachbar; Propensity Score (ohne Zurücklegen)</b>				
Ohne Caliper	-0,226	0,032	-7,15	473
<b>Ein nächster Nachbar; Propensity Score (ohne Zurücklegen)</b>				
Caliper = 0,01	-0,101	0,051	-1,96	188
Caliper = 0,05	-0,106	0,051	-2,07	188
<b>Ein nächster Nachbar; Propensity Score (mit Zurücklegen)</b>				
Caliper = 0,05	-0,218	0,060	-3,61	482
Caliper = 0,005	-0,174	0,059	-2,96	438
<b>Ein nächster Nachbar; Log Odds (ohne Zurücklegen)</b>				
Caliper = 0,01	-0,119	0,056	-2,13	160
Caliper = 0,05	-0,128	0,051	-2,49	188
<b>Ein nächster Nachbar; Log Odds (mit Zurücklegen)</b>				
Caliper = 0,005	-0,151	0,067	-2,27	152
Caliper = 0,01	-0,139	0,058	-2,38	231
Caliper = 0,05	-0,215	0,059	-3,66	442
<b>Maximal zwei nächste Nachbarn; Log Odds (mit Zurücklegen)</b>				
Caliper = 0,005	-0,151	0,065	-2,33	152
Caliper = 0,01	-0,139	0,055	-2,54	231
Caliper = 0,05	-0,187	0,053	-3,52	442
<b>Maximal drei nächste Nachbarn; Log Odds (mit Zurücklegen)</b>				
Caliper = 0,005	-0,151	0,065	-2,33	152
Caliper = 0,01	-0,144	0,054	-2,65	231
Caliper = 0,05	-0,186	0,052	-3,56	442
<b>Maximal sechs Nachbarn; Log Odds (mit Zurücklegen)</b>				
Caliper = 0,01	-0,146	0,054	-2,7	231
<b>Radius-Matching; Log Odds (mit Zurücklegen)</b>				
Caliper = 0,01	-0,175	0,052	-3,35	482
<b>Kernel-Matching; Log Odds (mit Zurücklegen)</b>				
Bandwidth = 0,01	-0,190	0,053	-3,6	482
Rundungsbedingte Ungenauigkeiten				



## Kurzfassung

Die vorliegende Forschungsarbeit setzt sich mit der Frage nach der Stabilität finanziell geförderter Beschäftigungsverhältnisse (Kombilöhne) auf drei Ebenen auseinander.

Inhaltlicher Schwerpunkt ist zunächst die Evaluation der arbeitsmarktpolitischen Maßnahme *Kombilohn*. Ausgangspunkt ist die bisher wenig thematisierte empirische Beobachtung, dass finanziell bezuschusste Beschäftigungsverhältnisse oft nur von kurzer Dauer sind. Da mehr Einkommen scheinbar zu geringerer Neigung zum Verbleib in Beschäftigung führt, wird dies im Anschluss an Boudon als eine Rationalitäts-Paradoxie aufgefasst. Es wird untersucht, welche Faktoren einer dauerhaften Integration in Beschäftigungsverhältnisse durch finanzielle Kombilohnzuschüsse entgegenstehen könnten.

Dies führt auf der Ebene der Handlungstheorie dazu, das Konzept der Rationalität als finanzielle Nutzenmaximierung zu hinterfragen, wie es in Bezug auf die erhoffte Wirkung arbeitsmarktpolitischer Maßnahmen wie dem Kombilohn oft unterstellt wird. Als Alternative wird der Begriff der Reziprozität von Arbeitslohn und Arbeitsleistung als handlungsleitende Norm aus der aktuellen soziologischen, ökonomischen und sozialpsychologischen Forschung herausgearbeitet.

In methodologischer Hinsicht diskutiert die Arbeit, inwiefern das in der Arbeitsmarktforschung verbreitete Ideal des Zufallsexperimentes und das daran orientierte Kausalmodell von Donald B. Rubin sowie seine Umsetzung im Propensity Score Matching für die Analyse von sozialen Handlungen überhaupt angemessen ist. Es wird herausgearbeitet, dass Rubins Kausalmodell entgegen kritischer Stimmen aus der Soziologie einen geeigneten Analyserahmen für die soziologische Forschung liefert.

Aus den handlungstheoretischen Überlegungen wird die Hypothese abgeleitet, dass die subjektiv-normative Orientierung der Geförderten am Reziprozitätsprinzip für das frühzeitige Ende ihrer Beschäftigungen verantwortlich ist. Auf Basis einer Befragung aus den Jahren 2002/2003 zur Kombilohnmaßnahme „Mainzer Modell“ und unter Verwendung des Propensity Score Matchings wird diese Hypothese überprüft und im Wesentlichen bestätigt. Zusätzlich werden umfassende Sensitivitätsanalysen und Robustheitschecks durchgeführt, welche die Validität der erlangten Ergebnisse stützen.

Aus der Analyse wird der Schluss gezogen, dass sowohl die Konzeption als auch die nachträgliche Evaluation arbeitsmarktpolitischer Maßnahmen davon profitieren könnte, neben dem Motiv der Nutzenmaximierung auch subjektive und normative Gründe des Handelns zu berücksichtigen.