

Nonresponse bei Verbleibsuntersuchungen

Korrekturverfahren zu Antwortausfällen am Beispiel ehemals arbeitsloser Existenzgründer, die mit dem Überbrückungsgeld (§ 57 SGB III) gefördert wurden

Frank Wießner*

Fünf Jahre nach der Existenzgründung wird es zunehmend schwierig, Informationen über den Verbleib der vormals mit Überbrückungsgeld Geförderten zu erhalten. War es in einer früheren Projektetappe – nach durchschnittlich drei Jahren in Selbständigkeit - noch möglich gewesen, bei der Befragung der Gründerinnen und Gründer eine Ausschöpfungsquote von annähernd 90 Prozent der Ausgangsstichprobe zu realisieren, so stellt nun zum Abschluss der Untersuchung nicht nur der „natürliche Schwund“ der Untersuchungseinheiten – etwa durch nicht mehr recherchierbare Adressen – die Analyse vor methodische Probleme. Systematische Verzerrungen entstanden auch durch die Erhebungsorganisation. Denn die abschließenden Telefoninterviews wurden nur mit jenen Gründern geführt, die in der vorletzten Befragungswelle ihre Bereitschaft signalisiert hatten, ein weiteres Mal an der Erhebung teilzunehmen. Daraus resultiert ein Positiv-Bias, der entsprechende Korrekturen erfordert.

Mit Hilfe eines Worst-Case-Szenarios, einer gebundenen Hochrechnung, Datenabgleichen und einem Gewichtungungsverfahren wurde ein „Korridor“ bestimmt, in dem sich die „wahre“ Überlebensquote der ehemaligen Überbrückungsgeld-Empfänger mit großer Wahrscheinlichkeit bewegt. Demnach betreibt gut fünf Jahre nach der Gründung noch immer mindestens jeder zweite der vormals Arbeitslosen sein Unternehmen. Verglichen mit den Befunden anderer Gründerstudien ist für die Gründungsverläufe der Geförderten alles in allem ein ähnlicher Verlauf zu konstatieren wie bei Start-ups, die ohne das Handicap vorangegangener Arbeitslosigkeit getätigt wurden. Trotz der verbleibenden Unschärfen im Datenmaterial und der wachsenden Schwierigkeiten bei der Korrektur dieser Verzerrungen weisen die ehemaligen Überbrückungsgeld-Empfänger noch immer eine respektable Bilanz auf – zumindest auf der Brutto-Ebene. Zu den Netto-Wirkungen der Förderung ist keine Aussage möglich, da eine geeignete Vergleichsgruppe nicht zur Verfügung steht.

Auch der Blick auf die Erfolgsfaktoren zeigt wenig Veränderung im Zeitverlauf. Wie schon rund zwei Jahre zuvor liegen die kritischen Faktoren nach wie vor im Bereich des Humankapitals und des Startkapitals. Eine Einschätzung, inwieweit zusätzliche Förderprogramme auf diesen Gebieten eine sinnvolle Flankierung bieten können, muss nachfolgenden Untersuchungen vorbehalten bleiben. Die vielfältigen Initiativen und Angebote, die in den vergangenen Jahren neu geschaffen wurden, können dazu beitragen, weitere Gründungspotenziale zu erschließen.

Wenn auch der Beitrag des Überbrückungsgeldes zu diesen positiven Entwicklungen ebenso wenig genau beziffert werden kann wie das Ausmaß von Verdrängungs- und Mitnahmeeffekten, die den Nettoerfolg des Förderinstrumentes auf der Arbeitsmarktebene schmälern, so bleibt doch auf jeden Fall festzuhalten, dass das Überbrückungsgeld einen wichtigen Brückenschlag zurück in den Arbeitsmarkt leistet, auch für jene, die in den neugegründeten Unternehmen Beschäftigung finden.

Gliederung

- 1 Begleitforschung zum Überbrückungsgeld
 - 2 Fünf Jahre nach der Gründung
 - 2.1 Telefonische Verbleibsuntersuchung
 - 2.2 Ergebnisse der telefonischen Befragung
 - 2.3 Störquellen und systematische Verzerrungen im Datenmaterial
 - 3 Viele Wege führen zum Ziel
 - 3.1 Worst Case-Szenario zur Bestimmung einer theoretischen Untergrenze
 - 3.1.1 Untersuchungsschritte
 - 3.1.2 Zwischenfazit
 - 3.2 Relativierung der empirischen Befunde durch gebundene Hochrechnung
 - 3.3 Bestimmung einer Obergrenze durch Datenabgleich
 - 3.3.1 Durchgeführte Datenabgleiche
 - 3.3.2 Probleme beim Datenabgleich
 - 3.3.3 Befunde der Datenabgleiche
- 4 Determinanten des Teilnahmeverhaltens – Gewichtung als Lösung?
 - 4.1 Das Problem der Selbstselektivität von Befragungsteilnehmern
 - 4.2 Selbstselektivität in der Untersuchung

* Dr. Frank Wießner ist wiss. Mitarbeiter im IAB. Der Beitrag entstand im Rahmen des IAB-Projekts „Förderung der Aufnahme einer selbständigen Tätigkeit nach § 57 SGB III (vormals § 55a AFG) (Überbrückungsgeld: Strukturen, individuelle Erträge, Arbeitsmarktentwicklungen)“ (Projektnummer 10-471). Er liegt in der alleinigen Verantwortung des Autors. Der Beitrag wurde im Februar 2003 eingereicht und nach der Begutachtung im Mai 2003 zur Veröffentlichung angenommen.

- 4.3 Multivariate Betrachtung möglicher Determinanten selbstselektiven Teilnahmeverhaltens
- 4.4 Bestimmung von Gewichtungsfaktoren
- 5 Determinanten des Erfolges – Neuauflage eines alten Modells
- 6 Fazit
- Literatur

1 Begleitforschung zum Überbrückungsgeld

Seit 1986 werden Arbeitslose, die den Schritt in die berufliche Selbständigkeit wagen wollen, von der Bundesanstalt für Arbeit mit dem Überbrückungsgeld (§ 57 SGB III) gefördert. Über die Jahre hinweg hat sich – von einer durchaus kontroversen Diskussion begleitet – aus den eher bescheidenen Anfängen ein Förderinstrument etabliert, dessen Reichweite und Wirkungsgrad gleichermaßen beachtlich sind. Teilnehmerzahlen in einer Größenordnung von über 123.000 Bewilligungen und ein Jahresbudget von mehr als 1 Mrd. € allein im Jahr 2002 machen das Überbrückungsgeld heute zu einer zentralen Säule des Gründungsgeschehens in der Bundesrepublik. Die jüngsten IAB-Untersuchungen zeigen für den Durchschnitt der vormals arbeitslosen Existenzgründer günstige Entwicklungsverläufe auf. Drei Jahre nach der Gründung waren noch 7 von 10 Überbrückungsgeld-Empfängern selbständig. Jenseits ihrer „Selbstbeschäftigung“ wurden viele der Geförderten darüber hinaus als Arbeitgeber zu wichtigen Multiplikatoren auf dem Arbeitsmarkt: Im Durchschnitt kommt nach drei Jahren auf je einen *ursprünglich* Geförderten ein zusätzlicher Arbeitsplatz.

Die streckenweise rasante Entwicklungsdynamik des Überbrückungsgeldes, die durch eine Reihe von Gesetzesnovellierungen zusätzlichen An Schub erhielt, machte die Gründungsförderung mit Mitteln der Arbeitslosenversicherung schon bald zum Gegenstand der wissenschaftlichen Begleitforschung im IAB (vgl. Kaiser/Otto 1990). Während es bei dieser ersten Überbrückungsgeld-Studie noch möglich war, eine Totalerhebung unter den Förderkohorten der Jahre 1986 – 1988 durchzuführen, konnte die zweite, im Jahre 1995 begonnene IAB-Studie aufgrund der drastisch angestiegenen Zahl der Förderfälle lediglich auf eine Stichprobe rekurren¹. Wie schon in der Vorläuferstudie ging es zunächst darum, systematische Informationen zu den Neugründungen, den Strukturen der Teilnehmer und insbesondere zu deren persönlichem Verbleib zu erschließen. Darüber hinaus sollten weitere Arbeitsmarkteffekte, wie beispielsweise zusätzlich geschaffene Arbeitsplätze, aber auch unbeabsichtigte Wirkungen („Emergenzeffekte“), wie etwa Mitnahme- oder Verdrängungseffekte untersucht werden. Jenseits der grundsätzlichen Beobachtung der Entwicklungsverläufe der neugegründeten Unternehmen galt das weitere Erkenntnisinteresse vor allem der Wirkungsweise der Gründungsförderung und insbesondere möglichen Erfolgsfaktoren bzw. Friktionen im Gründungskontext, um so das Verhältnis zwischen Chancen und Risiken für die jungen Unternehmen weiter optimieren zu können. Entsprechen-

de Befunde wurden zwischenzeitlich vorgestellt (vgl. Wießner 1998, 2000, 2001).

Während der Untersuchungszeitraum für diese Analysen sich erhebungsbedingt auf die ersten drei Jahre nach der Gründung beschränkte, konnte das „Zeitfenster“ der Beobachtung der Gründungsverläufe zwischenzeitlich auf 5 Jahre ausgeweitet werden. Mit der Untersuchung des Verbleibs der vormals Geförderten nach rund 5 Jahren in beruflicher Selbständigkeit soll die IAB-Studie zum Überbrückungsgeld zugleich einen vorläufigen Abschluss finden.

2 Fünf Jahre nach der Gründung

2.1 Telefonische Verbleibsuntersuchung

Nach der vorangegangenen „Aktenrecherche“ im Jahre 1995 und der postalischen bzw. telefonischen Verbleibsuntersuchung im Sommer 1997 wurde die Stichprobe der Existenzgründer, die in den Jahren 1994/95 mit dem Überbrückungsgeld gefördert worden waren (vgl. Fußnote 1 zur Stichprobe), in der Zeit vom 20.09.1999 bis zum 31.12.1999 (Ausnahme: Göppingen; hier bis zum 29.2.2000) von den Forschungssachbearbeitern des IAB abermals telefonisch interviewt. Mit dieser zweiten Telefonbefragung verlängert sich der durchschnittliche Beobachtungszeitraum auf gut 5 Jahre².

Die Ausgangsstichprobe von ursprünglich 4486 Fällen reduzierte sich dabei auf 2299 Probanden. Denn genau so

¹ Grundsätzlich erfordert ein mathematisch korrekter Repräsentationsschluss auf die Grundgesamtheit eine kontrollierte Stichprobenziehung unter Kenntnis von Erhebungseinheit und Grundgesamtheit. Die hier vorgenommene „geschichtete Auswahl“ zählt streng genommen nicht zu den mathematisch abgesicherten Auswahlverfahren, weil sie eben nur kontrolliert, ob die Stichprobe hinsichtlich der mehr oder weniger „willkürlich“ festgelegten Quotierungsmerkmale ein verkleinertes Abbild der Realität darstellt. Einbezogen werden konnten nur wenige Merkmale, die die jeweilige regionale Arbeitsmarktsituation widerspiegeln und mutmaßlich Einfluss auf den Erfolg der Förderung haben. Dabei steuern die verwendeten Quoten in erster Linie die prozentuale Zusammensetzung der Merkmale, ohne jedoch deren Kombination und Beziehungen untereinander wiederzugeben. Die Ergebnisse der Untersuchung können damit lediglich als „quasi-repräsentativ“ (Brinkmann/Otto/Wiedemann 1995: 6) betrachtet werden.

Ausgangsbasis der Untersuchung bildet eine Stichprobe von 15 ausgewählten Arbeitsamtsbezirken (11 im Westen, 4 im Osten Deutschlands), die nach einem kombinierten geschichteten und quotenkontrollierten Verfahren ausgewählt wurden. Berücksichtigt wurden Arbeitslosenquote und Struktur der Arbeitslosigkeit, Beschäftigungsentwicklung sowie Förderanteile. Zusätzlich zu diesem quotenkontrollierten Auswahlverfahren orientierte sich die Stichprobenziehung an einem Zeitrahmen. In den ausgewählten Arbeitsämtern wurden alle im ersten und zweiten Quartal 1994 sowie die im vierten Quartal 1994, die im ersten Quartal 1995 und die in der ersten Hälfte des zweiten Quartals 1995 bewilligten Anträge auf Überbrückungsgeld ausgewertet (Vollerhebung in den ausgewählten Arbeitsämtern). Die im dritten Quartal des Jahres 1994 bewilligten Anträge wurden nicht mit in die Erhebung aufgenommen, weil der damalige § 55a AFG (der Vorläufer des heutigen § 57 SGB III) in diesen Zeitraum durch das Beschäftigungsförderungsgesetz vom 01.08.1994 novelliert wurde und in der Übergangsphase die Behandlung der Förderfälle nach altem und neuem Recht nicht trennscharf abgegrenzt werden kann. (Siehe ausführlich Wießner 2001).

² Das maximale Alter der untersuchten Gründungen beträgt 72 Monate, die letzten in die Untersuchung einbezogenen Gründungen machten sich etwa 55 Monate vor dem Stichtag selbständig.

Übersicht 1: Verteilung der Stichprobe über die ausgewählten Untersuchungsämter

Arbeitsamtsbezirk	Telefoninterviews 1999	Aktenrecherche 1995
Schwerin	198	415
Dessau	158	350
Pirna	280	489
Bremen	100	210
Hannover	265	527
Essen	101	203
Kassel	176	354
Bad Kreuznach	59	107
Mainz	106	181
Göppingen	210	377
Bayreuth	77	131
Hof	86	154
Deggendorf	128	276
Landshut	94	198
Berlin	261	503
Insgesamt	2 299	4 486

viele hatten sich in der schriftlichen Befragung von 1997³ bereit erklärt, an einer weiteren Befragungsrunde teilzunehmen. Von diesen waren 20 Personen nicht eindeutig identifizierbar bzw. konnten in diesen Fällen keine Telefonnummern oder Adressen ermittelt werden. An dieser Stelle entsteht also eine Schiefe in der Datenbasis,

denn wer sich 1997 nicht an der Fragebogen-Aktion beteiligte, aber zumindest ein kurzes Telefon-Interview gab, hatte keine Chance mehr, 1999 in die Auswahl zu gelangen. Übersicht 1 zeigt die Verteilung der reduzierten Stichprobe in den ausgewählten Arbeitsamtsbezirken.

Insgesamt konnten 2055 auswertbare Telefoninterviews geführt werden, was einer Rücklaufquote von gut 90 % entspricht⁴.

³ Der IAB-Untersuchung liegt folgendes Design zugrunde: Zunächst wurde im Sommer 1995 in den ausgewählten Arbeitsamtsbezirken eine Aktenrecherche aller Förderfälle innerhalb des definierten Zeitraumes durchgeführt. Im Anschluss daran wurden die ausgewählten Probanden in kurzen Telefoninterviews um ergänzende Angaben gebeten. In der zweiten Projektphase im Sommer 1997 erfolgte dann bei den Personen der gleichen Stichprobe eine postalische Nachbefragung. Dabei wurden bestimmte zentrale Informationen, wie z. B. der Verbleib in der geförderten Existenz, die Zahl der Mitarbeiter oder der Unternehmensgegenstand erneut abgefragt, um den Einfluss des Zeitverlaufs auf die Entwicklung der Gründungen angemessen berücksichtigen zu können. Hierzu wurde den Untersuchungsteilnehmern zunächst ein Erhebungsbogen mit Anschreiben übersandt. In etwa vierwöchigem Abstand wurden die Nichtantworter noch zweimal mit Erinnerungsschreiben um ihre Teilnahme gebeten. Wer bis zum Stichtag 30. 09.1997 noch nicht geantwortet hatte, wurde in eine telefonische Nachfassaktion (mit einer Kurzfassung des Erhebungsbogens) einbezogen, die bis Mitte November 1997 andauerte.

Zum damaligen Zeitpunkt konnten Gründungsverläufe für einen Zeitraum von durchschnittlich etwa drei Jahren bis maximal 45 Monate betrachtet werden. Die „jüngsten“ Gründungen waren zum Stichtag der Untersuchung des Jahres 1997 etwa 25 Monate alt.

⁴ An dieser Stelle ist das Engagement der beteiligten Sachbearbeiter für Arbeitsmarkt- und Berufsforschung in den Stützpunkt-Arbeitsämtern zu würdigen. Unser Dank gilt hierfür Werner Bossert (Göppingen), Werner Eichinger (Landshut), Ulrich Elsweier (Schwerin), Gerhard Geier (Bamberg), Detlef Güttler (Bonn), Werner Haßenpflug (Kassel), Günter Heuer (Essen), Cristine Hense (Dessau), Manfred Jouaux (Bad Kreuznach), Dr. Hans Kaule (Pirna), Michael Morsch (Elmshorn), Norbert Schön (Berlin III) sowie Otfried Steinmeister (Hannover).

Übersicht 2: Verbleib der Geförderten zur Zeit der ersten Befragung 1997 und der zweiten Befragung 1999/2000

	Befragung 1999/2000	Befragung 1997		
	Nur telefonisch	insgesamt	schriftlich	telefonisch
Selbständig	72,3 %	70,4 %	72,5 %	64,3 %
darunter:				
• geförderte Existenz	67,6 %	66,5 %	69,2 %	58,6 %
• andere selbständige Existenz	4,7 %	3,9 %	3,3 %	5,7 %
nicht mehr selbständig	27,7 %	29,6 %	27,5 %	35,7 %
darunter:				
• sozialversicherungspflichtig beschäftigt	15,3 %	12,0 %	10,8 %	15,7 %
• in anderer Form erwerbstätig	0,5 %	1,0 %	1,1 %	0,6 %
• mithelfende(r) Angehörige(r)	0,3 %	0,3 %	0,4 %	0,1 %
• arbeitslos mit Leistungsanspruch	5,6 %	8,7 %	8,5 %	9,0 %
• arbeitslos ohne Leistungsanspruch	1,6 %	2,0 %	1,3 %	4,2 %
• Fortbildung, Umschulung	1,0 %	1,4 %	1,2 %	1,8 %
• Rente/Vorruhestand	1,6 %	0,8 %	0,7 %	1,0 %
• Mutterschutz/Erziehungsurlaub	0,3 %	0,8 %	0,9 %	0,3 %
• Sonstiges	1,5 %	1,6 %	1,4 %	2,2 %
• keine Angabe	0 %	1,1 %	1,2 %	0,8 %
Zahl der Fälle (=100 %)	2055	3846	2883	963

2.2 Ergebnisse der telefonischen Befragung

Die Telefoninterviews zum Verbleib der Geförderten liefern überaus positive Befunde (siehe Übersicht 2): Von den Befragten waren im Erhebungszeitraum noch annähernd drei Viertel selbständig, davon der Großteil in der geförderten Existenz. Die Tatsache, dass zunehmend die Älteren ihre selbständige Tätigkeit wieder aufgeben, lässt unter gewissen Vorbehalten darauf schließen, dass die berufliche Selbständigkeit in manchen Fällen lediglich als eine vorübergehende Episode auf dem Weg in den Altersruhestand betrachtet wurde. Da jedoch ein guter Teil der älteren Geförderten, die heute nicht mehr selbständig tätig sind, das entsprechende Alter für die Inanspruchnahme von (Vor-)Ruhestandregelungen noch nicht erreicht hatte, stellt sich andererseits die Frage, inwieweit für die Übergangsphase noch andere Sozialversicherungsleistungen in Anspruch genommen werden können.

Mehr als die Hälfte derer, die ihre selbständige Tätigkeit zwischenzeitlich wieder aufgegeben hatten, war bis zur zweiten telefonischen Befragung wieder abhängig beschäftigt. Als arbeitslos waren zum Befragungszeitpunkt knapp 7 % der Interview-Teilnehmer gemeldet. Dies bestätigt abermals die schon in der vorausgegangenen Untersuchungsetappe getroffene Feststellung, dass es sich bei den Existenzgründern insgesamt um eine sehr aktive und agile Gruppe unter den Arbeitslosen handelt.

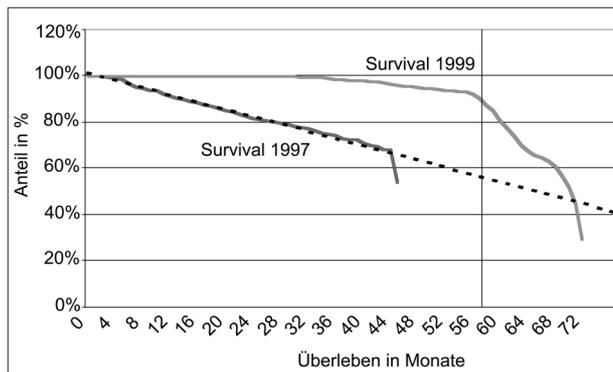
Gerade der ausnehmend günstige Verbleib der ehemaligen Überbrückungsgeld-Empfänger stellt die Untersuchung jedoch vor ein gravierendes Problem. Denn wie Übersicht 2 ebenfalls deutlich macht, sind den Ergebnissen der Telefoninterviews zufolge 5 Jahre nach der Gründung relativ mehr Untersuchungsteilnehmer noch selbständig, als dies bei der Zwischenbefragung nach 3 Jahren der Fall war. Offenbar liegt hier – bedingt durch das Untersuchungsdesign – eine Positivselektion vor, die unplausible und ganz offensichtlich verzerrte Ergebnisse zur Folge hat.

2.3 Störquellen und systematische Verzerrungen im Datenmaterial

Eine naheliegende Erklärung dieses Problems lautet folgendermaßen: Da die Teilnahme an den verschiedenen Etappen der Untersuchung für die Probanden grundsätzlich freiwillig ist, erfolgt sie mithin selbstselektiv. Dies hat systematische Verzerrungen zur Folge, die sich von Befragung zu Befragung kumulieren. So erklärt sich auch die paradoxe Zunahme der Überlebensrate der Neugründungen im Zeitverlauf⁵. Denn die betrachteten Untersuchungseinheiten bilden eine Teilmenge der Ausgangsstichprobe, die – wie gezeigt – von einem Stichtag zum nächsten diskret schrumpft. Die daraus folgenden Verzerrungen gewinnen so im Zeitverlauf proportional an Gewicht und die Informationsbasis wird immer unsicherer.

Abbildung 1 soll diesen Zusammenhang veranschaulichen. Im Diagramm eingezeichnet sind die Survival-Funktionen⁶ für die Untersuchungspopulationen 1997 und 1999. Für die Datenbasis aus 1997 wurde zusätzlich (gestrichelt dargestellt) eine Trendextrapolation eingezeichnet⁷.

Abbildung 1: Survival-Funktionen der Stichproben von 1997 und 1999/2000



Die Graphik veranschaulicht, wie sehr sich die Untersuchungspopulationen unterscheiden. Da es sich bei den Elementen der oben verlaufenden Funktion für das Jahr 1999 um eine Teilmenge der 1997 untersuchten Stichprobe handelt, macht der „Keil“ zwischen den beiden Kurven das Ausmaß der Positivselektion klar. Während die Überlebenskurve für 1997 fast linear abnimmt und nach drei Jahren gut mit der empirisch ermittelten Verbleibsquote von rund 70 % korrespondiert, treten bei der Gruppe, die sich 1999 freiwillig einer weiteren Befragungsrunde unterzog, im selben Zeitraum so gut wie keine Ausfälle auf (die Knicke am Ende der beiden Kurven sind wohl auf Kohorteneffekte zurückzuführen: Da es nur relativ wenige Fälle gibt, die den maximal möglichen Zeitraum bisher durchlaufen haben, fallen die Ausfälle hier relativ stärker ins Gewicht). In der Trendverlängerung weist die Survivalfunktion von 1997 nach 5 Jahren einen Wert von etwa 55 % auf, die 1999-Funktion liegt dagegen um mehr als 25 Prozentpunkte darüber. Dies ist ein weiterer Hinweis auf eine Verzerrung. Offenbar ist die Positivauswahl in der zweiten Befragungswelle so stark, dass die wenigen Ausfälle in den ersten Jahren die im Modell errechnete Survivalkurve soweit nach oben verschieben, dass diese von dem empirisch ermittelten Wert (vgl. Übersicht 2) drastisch abweicht. Damit ist die Basis für weitere Analysen jedoch höchst unsicher.

Interessant ist indes der Schnittpunkt der Überlebenskurve von 1999 und der Trendfunktion von 1997. Demnach wäre 6 Jahre nach der Gründung noch mit einer Überlebensrate von etwa 45 % der ursprünglich Geförderten zu rechnen.

⁵ In Wirklichkeit ist dies natürlich nicht der Fall, auch wenn die Ergebnisse der Telefoninterviews einen solchen Eindruck erwecken.

⁶ Als Survival- oder auch Sterbetafel-Analyse wird ein statistisches Verfahren bezeichnet, das für Untersuchungseinheiten innerhalb eines bestimmten Beobachtungszeitraumes die Wahrscheinlichkeit ermittelt, den Beginn des nächsten Intervalls ($i + 1$) zu erleben unter Berücksichtigung der bedingten Wahrscheinlichkeit, das gegenwärtige Intervall i zu überleben. Dabei wird unterschieden zwischen solchen Untersuchungseinheiten, bei denen das untersuchte Ereignis (hier: Aufgabe der selbständigen Tätigkeit) tatsächlich eingetreten ist und solchen, bei denen das Ereignis im Beobachtungszeitraum (noch) nicht eingetreten ist („Rechtszensurierung“).

⁷ Bei der Survivalkurve von 1997 erfolgt die beste Trendanpassung durch die einfache Linearfunktion $y = -0,0078x + 1,0178$.

Abbildung 2 veranschaulicht den Zusammenhang zwischen Teilnahmeverhalten und Verbleib sowie das systematische Anwachsen der Verzerrungen im Datenmaterial. Aufgrund des Panel-Charakters der Untersuchung können – insbesondere systematische – Ausfälle von Befragungsteilnehmern zu einer Kumulation von Problemen führen. Hartmann/Kohaut (2000: 610) stellen hierzu fest: „Nicht neutrale Ausfälle führen zu Problemen, wenn es sich um *systematische Ausfallprozesse* handelt (...). Derartige systematische Verzerrungen schränken die Möglichkeiten von Schlussfolgerungen von der Stichprobe auf die Grundgesamtheit ein, wobei sich diese Problematik bei Panelerhebungen (...) über die Zeit noch verstärken kann.“

Klar zu erkennen sind zunächst die deutlichen Unterschiede hinsichtlich des Verbleibs zwischen denjenigen, die 1997 den Fragebogen retournierten und denen, die bei der ersten Befragungswelle erst telefonisch erreicht werden konnten. Dies hat im weiteren Untersuchungsverlauf auch Auswirkungen auf die zweite (telefonische) Befragungswelle: Denn *nur* im Fragebogen (schriftliche Befragung 1997) konnten die untersuchten Personen ihre Bereitschaft für die Teilnahme an einer Folgeuntersuchung bekunden (und *nur* solche Probanden wurden tatsächlich ein weiteres Mal befragt). Die daraus folgende Situation ist in Abbildung 2 dargestellt.

Auf die jeweiligen Segmente bezogen sind die Entwicklungsverläufe in sich konsistent und die Befunde durchaus korrekt. Gleichwohl ist damit die Frage nach den Entwicklungsverläufen *aller* ursprünglich Geförderten nicht beantwortet. Mithin fällt die empirisch ermittelte Verbleibsquote aufgrund der immanenten Selektivität offenkundig zu hoch aus. Nachfolgend soll daher mittels eines Worst Case-Szenarios die *maximale* Zahl der Ausfälle im Untersuchungszeitraum hochgerechnet und damit eine theoretische *Untergrenze* bestimmt werden. Zusätzlich

soll durch Abgleiche mit verschiedenen Geschäftsstatistiken der Bundesanstalt für Arbeit die *kleinstmögliche* Zahl der nicht mehr Selbständigen als *Obergrenze* ermittelt werden. Der „wahre“ Wert für den Verbleib in Selbständigkeit ist dann innerhalb dieses Korridors zu verorten.

3 Viele Wege führen zum Ziel

3.1 Worst Case-Szenario zur Bestimmung einer theoretischen Untergrenze

Wie in Abbildung 2 verdeutlicht, sind die Verzerrungen im Datenmaterial nicht zufälliger Natur, sondern systematisch verursacht. Verantwortlich sind zum einen Antwortausfälle und zum anderen die unterschiedlichen Befragungskanäle, auf denen die Filterführung für zentrale Fragen aufbaut. Aus dem Design der Untersuchung resultieren zwei zentrale Einschränkungen, die vor allem für die Selbstselektion der Interviewteilnehmer ursächlich sind:

– In der zweiten Erhebungsrunde 1999/2000 konnte nur telefonisch befragt werden, wer an der *schriftlichen* Befragung im Jahre 1997 teilgenommen hatte *und* bei dieser Gelegenheit seine Einwilligung für eine weitere Befragung gegeben hatte.

– Zweitens standen für die zweite Befragung naturgemäß auch all jene nicht zur Verfügung, die schon in der ersten Befragung nicht erreicht werden konnten. Dabei sind die Gründe für die Nichtteilnahme bzw. Nichterreichbarkeit zunächst unerheblich. Konkret reduziert die Nichterreichbarkeit die Untersuchungsgesamtheit lediglich um 155 „neutrale Ausfälle“, die zum Zeitpunkt der ersten Befragung verstorben, unbekannt verzogen oder aus sonstigen Gründen von vornherein nicht mehr angeschrieben bzw. angerufen werden konnten.

Abbildung 2: Teilnahmeverhalten und Überlebensquoten in den verschiedenen Phasen der IAB-Untersuchung

Aktenrecherche 1995	Erste Befragung 1997 (schriftlich/telefonisch) Survival gesamt: 70,4 %		Zweite Befragung 1999/2000 (telefonisch)
Untersuchungsgesamtheit 4486 Fälle	155 neutrale Ausfälle 50 Verweigerungen 435 Interview-Ausfälle		
	Telefonische Befragung 963 Fälle (Überlebensquote 64,3 %)		
	Schriftliche Befragung 2883 Fälle (Überlebensquote 72,5 %) *	Bereit zur Teilnahme an der Wiederholungsbefragung: 2299 Fälle (Überlebensquote 77,8 %) *	Telefonische Befragung 2055 Fälle (Überlebensquote 72,3 %) *

* Überlebensquote zum Zeitpunkt der Befragung.

3.1.1 Untersuchungsschritte

Im mathematischen Sinne erfolgt die Bestimmung der Untergrenze für die Verbleibsquote durch eine Art „Abzinsung“. Die Ausgangsbasis bildet dabei die Untersuchungsgesamtheit der Aktenrecherche (4486 Fälle). Im Zeitverlauf verringert sich diese Basis, wobei gewissermaßen unterschiedliche „Zinsfüße“ in Perioden unterschiedlicher Dauer wirksam werden. Auszugehen ist zunächst von folgender Situation:

Schritt 1: Teilnahme an der ersten Erhebungswelle 1997

Abbildung 3a: Gesamtteilnahme an der ersten Erhebung

Untersuchungsgesamtheit	Erste Befragung
Aktenrecherche 4486 Fälle	Neutrale Ausfälle 155 Fälle
	Verbleib unbekannt 485 Fälle
	Teilnahme an der Befragung 3846 Fälle davon 2708 (=70,4 %) noch selbständig 1138 (=29,6 %) nicht mehr selbständig

Neben den bereits angesprochenen 155 neutralen Ausfällen (die Untersuchungsgesamtheit reduziert sich damit auf 4331 Fälle) konnten in der ersten Erhebungswelle weitere 485 Fälle weder postalisch noch telefonisch erreicht werden. Deren Verbleib ist somit unbekannt. Für den ungünstigsten Fall („worst case“) ist anzunehmen, dass diese Personen nicht mehr selbständig sind. Die Zahl der gescheiterten Gründungen würde damit um *maximal* 485 Fälle unterschätzt werden. Die empirisch ermittelte Ausfallrate von 29,6 % wäre damit folgendermaßen zu korrigieren:

$$\frac{1138 + 485}{3846 + 485} = \frac{1623}{4331} = 0,3747 = 37,5\%$$

Das bedeutet, nach drei Jahren sind noch *mindestens* 62,5 % der ursprünglich Geförderten selbständig oder umgekehrt haben *höchstens* 37,5 % ihre berufliche Selbständigkeit wieder aufgegeben.

Schritt 2: Teilnahme an der schriftlichen Befragung (innerhalb der ersten Erhebungswelle)

Aus Schritt 1 ist bekannt, dass die Zahl der gescheiterten Gründungen um *maximal* 485 Fälle unterschätzt wird⁸. Geht man davon aus, dass all diese Personen theoretisch an der postalischen Befragung hätten teilnehmen können, so hätte diese im ungünstigsten Fall folgenden Befund ergeben:

$$\frac{793 + 485}{2883 + 485} = \frac{1278}{3368} = 0,3795 = 37,9\%$$

Abbildung 3b: Teilnahme an der schriftlichen Befragung

Netto-Untersuchungsgesamtheit	Erste Befragung
4331 Fälle (Untersuchungsgesamtheit bereinigt um neutrale Ausfälle)	Verbleib unbekannt 485 Fälle
	Telefoninterviews 963 Fälle davon 619 (=64,3 %) noch selbständig 344 (=35,7 %) nicht mehr selbständig
	Postalische Befragung 2883 Fälle davon 2090 (=72,5 %) noch selbständig 793 (=27,5 %) nicht mehr selbständig

Damit wären also von den Geförderten, die an der *schriftlichen Befragung* hätten teilnehmen können, noch *mindestens* 62,1 % selbständig und *höchstens* 37,9 % dieser Teilgruppe hätten ihre selbständige Existenz bis zum Sommer 1997 wieder aufgegeben.

Zum Vergleich: Empirisch war für diese Gruppe eine Überlebensquote von 72,5 % ermittelt worden.

In Schritt 2 war ermittelt worden, dass es rein theoretisch 2883 + 485 = 3368 Untersuchungspersonen möglich gewesen wäre, an der *schriftlichen* Befragung 1997 teilzunehmen. Es hätten also ebenso viele ihre Teilnahmebereitschaft für eine Wiederholungsbefragung erklären können. Tatsächlich haben dies aber nur 2299 Personen getan. Von denen wiederum waren 510 laut eigener Aussage zum Befragungszeitpunkt nicht mehr selbständig tätig.

Schritt 3: Teilnahme an der zweiten Erhebungswelle 1999/2000

Aus Schritt 2 ergibt sich zudem, dass neben den *tatsächlichen* 2883 Teilnehmern weitere 1069 Untersuchungspersonen *theoretisch* an der schriftlichen Befragung ebenfalls hätten teilnehmen und damit ihre Einwilligung für die zweite Befragung geben können. Da von den *theoretisch* möglichen 3368 Teilnehmern an der ersten *schriftlichen* Befragung 37,9 % bzw. 1278 Personen *theoretisch* aus der Selbständigkeit wieder ausgeschieden sind, wird die Zahl derer, die nicht mehr selbständig waren, aber *theoretisch* dennoch ihre Bereitschaft zur Teilnahme an der zweiten Befragung hätten erklären können, im ungünstigsten Fall um (1278 – 510 =) 768 Fälle unterschätzt.

⁸ In Schritt 2, eigentlich einem Zwischenschritt sollen die Teilnehmer des Telefoninterviews nun unberücksichtigt bleiben, da deren Verbleibe ja bekannt sind. Auch hatten sie ja nicht einmal die theoretische Chance, in die zweite Befragungswelle zu kommen. Die Basis der (theoretisch) maximal möglichen Teilnehmer reduziert sich damit auf 3368 Fälle.

Abbildung 3c: Teilnahme an der zweiten Erhebung

Theoretisch mögliche Teilnahme an der ersten <u>schriftlichen</u> Befragung	Bereitschaft zur Teilnahme an der Wiederholungsbefragung	Tatsächliche Teilnahme an der Wiederholungsbefragung
Theoretisch 3368 Fälle möglich Davon wären gem. Schritt 2 theoretisch 2090 (=62,1 %) noch selbständig 1278 (=37,9 %) nicht mehr selbständig	Bereitschaft theoretisch möglich 1069 Fälle davon theoretisch 301 noch selbständig 768 nicht mehr selbständig	Von weiteren 1069 Fällen wären theoretisch 303 noch selbständig 768 nicht mehr selbständig
	Teilnahmebereitschaft schriftlich erklärt 2299 Fälle davon 1789 (=77,8 %) noch selbständig 510 (=22,2 %) nicht mehr selbständig	244 weitere Probanden hätten real teilnehmen können Teilnahme am Telefoninterview 2055 Fälle davon 1486 (=72,3 %) noch selbständig 569 (=27,7 %) nicht mehr selbständig

Von den insgesamt 2299 Personen, die bei der ersten schriftlichen Befragung im Jahre 1997 ihre Bereitschaft zur Teilnahme an der Wiederholungsbefragung erklärt hatten, haben jedoch nur 2055 Personen *tatsächlich* teilgenommen. Es bleibt also eine Differenz von 244 Untersuchungspersonen, die an der zweiten Befragung nicht mehr teilgenommen haben, obwohl sie ihre Bereitschaft dazu erklärt hatten (Schritt 3). Legt man wieder die ungünstigste Annahme zugrunde, nämlich dass keine dieser Personen mehr selbständig ist, so ergibt sich folgende Korrektur:

$$\frac{569 + 244}{2055 + 244} = \frac{813}{2299} = 0,3536 = 35,4\%$$

Denn von 569 Befragten ist aufgrund des zweiten Telefoninterviews bekannt, dass sie definitiv nicht mehr selbständig waren und von weiteren 244 Personen, die theoretisch an der zweiten Welle hätten teilnehmen können, muss dies im ungünstigsten Fall vermutet werden.

Es ergäbe sich also eine Ausfallrate von 35,4 % bzw. wären noch 64,6 % der ursprünglich Geförderten selbständig tätig.

Nimmt man weiterhin an, dass von den 1069 Fällen, die 1997 theoretisch ebenfalls die Möglichkeit gehabt hätten, ihre Teilnahmebereitschaft zu erklären, Ende 1999 mindestens 768 nicht mehr selbständig waren (wie dies theoretisch aus Schritt 2 folgt), so ergibt sich weiterhin:

$$\frac{569 + 244 + 768}{2055 + 244 + 1069} = \frac{1581}{3368} = 0,4694 = 46,9\%$$

Das bedeutet: Nach gegenwärtigem Kenntnisstand haben – selbst bei Annahme des ungünstigsten Falles – *höchstens* 46,9 % der ursprünglich Geförderten ihre selbständige Tätigkeit wieder aufgegeben. Demnach sind also fünf Jahre nach der Gründung noch *mindestens* 53,1 % der vormaligen Geförderten selbständig tätig.

3.1.2 Zwischenfazit

Die Unterstellung des „worst case“ geht davon aus, dass in den zwei Jahren, die zwischen den beiden Befragungen liegen, die Zahl der Ausfälle unter denen, die *theoretisch* in die zweite Befragungsrunde hätten einbezogen werden können, weiter angestiegen ist. Auch unter der Annahme einer solchen Koinzidenz ungünstigster Umstände folgt, dass 5 Jahre nach der Gründung zumindest noch gut die Hälfte der ehemaligen Überbrückungsgeld-Empfänger in der beruflichen Selbständigkeit verblieben ist.

3.2. Relativierung der empirischen Befunde durch gebundene Hochrechnung

Wenn auch einiges dagegen spricht, den oben diskutierten „worst case“ als gegeben anzunehmen, so unterstreichen die Überlegungen doch, dass die in den Telefoninterviews ermittelte empirische Verbleibsquote mit Sicherheit zu hoch gegriffen ist. Denn soviel ist klar: Bei den Teilnehmern an der Wiederholungsbefragung handelt es sich eindeutig um eine Positivauswahl. Deshalb werden in den nachfolgenden Analyseschritten verschiedene Korrekturmöglichkeiten der so verursachten Verzerrungen diskutiert.

Eine erste Annäherungsmöglichkeit an einen Wert, der näher an der „tatsächlichen“ Verteilung des untersuchten Verbleibsmerkmals innerhalb der *Grundgesamtheit* liegt, bietet die so genannte „gebundene Hochrechnung“. Dies ist ein Schätzverfahren, bei dem durch die Berücksichtigung externer, also nicht aus der Stichprobe stammender Informationen eine Verbesserung der Schätzung erzielt wird. Hierzu wird neben dem eigentlichen Untersuchungsmerkmal (hier: Verbleib in Selbständigkeit zum Erhebungszeitpunkt) ein weiteres, so genanntes *Basismerkmal* erhoben. Dem liegt wiederum die Überlegung zugrunde, dass die Telefoninterviews von 1999/2000 eine Unterstichprobe der Erhebung von 1997 bilden. Auf dieser Grundlage werden nun zwei Annahmen getroffen:

1. Der Auswahlfehler ist sowohl in der Stichprobe von 1997 als auch in der Unterstichprobe von 1999/2000 annähernd gleichverteilt.

2. Das Verhältnis zwischen den Überlebensraten von 1997 und 1999/2000 entspricht in etwa dem Quotienten aus den empirisch ermittelten Selbständigen von 1999/2000 und der damaligen Survivalquote derer, die 1997 ihre Bereitschaft zur Teilnahme an einer Folgeuntersuchung erklärt hatten.

Somit ergibt sich folgender Zusammenhang:

$$\frac{SQ_{99}}{SQ_{97}} \approx \frac{SQ_{99\text{empirisch}}}{SQ_{97\text{Follow-up}}}$$

wobei folgende Bezeichnungen verwendet werden:

SQ_{99} :	Hochgerechnete Überlebensquote 1999/2000
SQ_{97} :	empirisch ermittelte Überlebensquote 1997
$SQ_{99\text{empirisch}}$:	empirisch ermittelte Überlebensquote 1999/2000
$SQ_{97\text{Follow-up}}$:	Überlebensquote derer, die sich 1997 zur Teilnahme an einer weiteren Befragung bereiterklärten

Daraus folgt:

$$SQ_{99} \approx \frac{SQ_{97} \times SQ_{99\text{empirisch}}}{SQ_{97\text{Follow-up}}} = \frac{0,704 \times 0,723}{0,785} = 0,6484 = 64,8 \%$$

Durch die Hochrechnung verringert sich also der empirisch ermittelte Anteil derer, die auch nach fünf Jahren noch selbständig sind, um $7\frac{1}{2}$ Prozentpunkte auf 64,8 %. Inwieweit diese Überlebensquote zutreffender als das Ergebnis der Telefonbefragung ist, soll durch die nachfolgend vorgestellten, ergänzend durchgeführten Datenabgleiche erhellt werden.

3.3 Bestimmung einer Obergrenze durch Datenabgleich

Die Befunde der telefonischen Befragung der Teilpopulation wurden ergänzt und validiert durch einen Abgleich der gesamten Ausgangsstichprobe⁹ (n = 4484) mit verschiedenen Prozessdaten der Arbeitslosenversicherung. Dieser zusätzliche Analyseschritt ist vor allem wegen der

in Abschnitt 2.3 diskutierten Verzerrungen der Befragungsdaten erforderlich. Jenseits des Problems der Selbstselektivität sprechen zudem pragmatische Gründe für den ergänzenden Datenabgleich. Denn zum einen konnte mit den begrenzten verfügbaren Ressourcen keine telefonische Befragung der Gesamtstichprobe durchgeführt werden, zum anderen sind schon aus der vorhergehenden Feldphase des Projekts erhebliche Probleme hinsichtlich der Erreichbarkeit der Probanden bekannt. Namens- und Adressänderungen, wie auch Todesfälle oder schlichtweg das Fehlen eines Eintrages in den amtlichen Telefonverzeichnissen sind hierfür verantwortlich.

3.3.1 Durchgeführte Datenabgleiche

Datenabgleiche¹⁰ wurden konkret durchgeführt

- mit der Leistungsempfängerdatei (LED),
- mit der Beschäftigtendatei (BST)
- und mit der Maßnahmeteilnehmerdatei (ST11).

Damit kann Folgendes festgestellt werden:

1. Der Abgleich mit der LED zeigt, wer zum Untersuchungszeitpunkt Leistungen vom Arbeitsamt bezogen hat. Diese Personengruppe ist also nicht mehr selbständig, sondern arbeitslos.
2. Der Abgleich mit der BST zeigt, wer zum Untersuchungszeitpunkt abhängig beschäftigt und damit ebenfalls nicht mehr selbständig tätig war.
3. Der Abgleich mit der ST11 zeigt, wer zum Untersuchungszeitpunkt an einer Maßnahme des Arbeitsamtes (z. B. ABM, SAM, usw.) teilgenommen hat.

Die Datenabgleiche fungieren also als eine Art „Negativtest“, der Personen ermittelt, die zu einem bestimmten Stichtag definitiv nicht selbständig tätig waren. Dies rechtfertigt jedoch keineswegs den Umkehrschluss, dass alle diejenigen Personen, die bei den Abgleichen nicht ermittelt wurden, weiterhin beruflich selbständig seien. Da deren Verbleib letztendlich unbekannt ist, fällt die Abschätzung der „Obergrenze“ zwangsläufig zu optimistisch aus.

3.2.2 Probleme beim Datenabgleich

3.2.2.1 Probleme aufgrund von Nichterfassung

Eine gravierende Schwäche der Datenabgleiche besteht darin, nicht alle Personen zu erfassen, die in irgendeiner Weise Arbeitsamtsleistungen beziehen oder Förderungen in Anspruch nehmen. Konkret konnten folgende Gruppierungen nicht identifiziert werden:

- Mithelfende Angehörige oder geringfügig Beschäftigte sind nicht eindeutig identifizierbar, weil diese zum

⁹ Der Datenabgleich bezieht sich nicht auf die Befragungsdaten, sondern auf die Aktenrecherche. Es werden also nur Daten, die *innerhalb* der Arbeitsverwaltung anfallen, abgeglichen.

¹⁰ Mittlerweile wird von der Statistik der Bundesanstalt für Arbeit ein ähnliches Verfahren zur Bestimmung des so genannten „umfassenden Verbleibsnachweises“ durchgeführt.

Untersuchungszeitpunkt in den Geschäftsstatistiken der BA noch nicht entsprechend ausgewiesen wurden;

- Arbeitslose *ohne* Leistungsanspruch oder Arbeitssuchende können nicht eindeutig ermittelt werden;
- Zu den Teilnehmern an Maßnahmen der beruflichen Weiterbildung (FbW) bzw. Fortbildungs- und Umschulungsmaßnahmen (FuU) sind für den Untersuchungszeitpunkt nur aggregierte Daten verfügbar, so dass Einzelfälle ebenfalls nicht identifiziert werden können;
- Zu Personen, die zum Untersuchungszeitpunkt verstorben sind, oder die sich bereits in Rente oder Vorruhestand befinden, die vorübergehend im Mutterschutz sind bzw. Erziehungsurlaub in Anspruch nehmen, sind überhaupt keine Daten verfügbar.
- Leistungsempfänger, die für länger als 6 Wochen erwerbsunfähig erkrankt sind, fallen aus der LED heraus. Sie müssen sich *nach* Genesung erneut arbeitslos melden.

In ihrer Summe haben diese Lücken in den Datenabgleichen zur Folge, dass die „Obergrenze“, d. h. die maximale Zahl derer, die zum Untersuchungszeitpunkt noch selbständig sein könnten, zu hoch geschätzt wird.

Übersicht 3 zeigt die relativen Häufigkeiten der Merkmale, die nicht anhand der Datenabgleiche ermittelt werden können. Im Vergleich zur ersten Erhebung im Sommer 1997 hat sich deren Anteil an der jeweils zugrunde liegenden Untersuchungsgesamtheit bei der zweiten Befragungsrunde leicht verringert. Eine klare Tendenz ist gleichwohl nicht zu erkennen. Da zu den Erwerbsverläufen zwischen den verschiedenen Erhebungszeitpunkten keine Informationen vorliegen, gelte die Annahme, die Chancen der Untersuchungseinheiten, in eine der nach-

Übersicht 3: Verbleib der Geförderten – Ausprägungen, die über Datenabgleich nicht ermittelbar sind

	Befragung 1999/2000 Nur telefonisch	Befragung 1997 insgesamt
• in anderer Form erwerbstätig	0,5 %	1,0 %
• mithelfende(r) Angehörige(r)	0,3 %	0,3 %
• arbeitslos ohne Leistungsanspruch	1,6 %	2,0 %
• Fortbildung, Umschulung	1,0 %	1,4 %
• Rente/Vorruhestand	1,6 %	0,8 %
• Mutterschutz/ Erziehungsurlaub	0,3 %	0,8 %
• Sonstiges	1,5 %	1,6 %
• Keine Angabe	0 %	1,1 %
Beim Datenabgleich unberücksichtigt:	6,8 %	9,0 %
Zahl der Fälle (=100 %)	2055	3846

folgenden Kategorien zu fallen, seien im Zeitverlauf weitgehend konstant. Die Wahrscheinlichkeit des Übergangs in den Altersruhestand nimmt mit der Zeit sogar zu.

Vorsichtig geschätzt werden deshalb die Datenabgleiche die Zahl derer, die ihre selbständige Tätigkeit in der Zwischenzeit wieder aufgegeben haben, um 5 bis 10 % zu niedrig ausweisen.

3.3.2.2 Probleme aufgrund mangelnder zeitlicher Aktualität

Ein weiteres grundlegendes Problem bei der Verarbeitung der Arbeitsamtsdaten besteht in ihrer zeitlichen Aktualität. Denn die in den Ämtern anfallenden Prozessdaten sind der zentralen Verarbeitung in der Hauptstelle der Bundesanstalt für Arbeit nur mit einem time-lag zugänglich.

Für den Abgleich der Überbrückungsgeld-Stichprobe mit den Leistungsempfängern wurde die LED zum Stand 31.12.1999 verwendet, so dass sie recht gut mit den Telefoninterviews konvergiert. Gleiches gilt für die ST11, die den Bestand an Maßnahmeteilnehmern zum Stand Januar 2000 widerspiegelt. Problematischer ist hingegen der Abgleich mit der Beschäftigtenstatistik, da deren Aktualität maßgeblich vom Meldeverhalten der Betriebe abhängt. Für die Ermittlung abhängiger Beschäftigung wurde hier die so genannte Quartalsdatei vom 30.06.1999 verwendet, die aus oben genannten Gründen erst Anfang 2000 in brauchbarem Zustand vorlag¹¹. In dieser Datei sind weitestgehend alle Personen enthalten, die der Bundesanstalt für Arbeit am 30.06.1999 als sozialversicherungspflichtig beschäftigt gemeldet waren. Eine Einschränkung ist darin zu sehen, dass diese Datei noch immer nicht „offiziell“ ist, so dass geringe Unschärfen nach wie vor in Kauf zu nehmen sind.

3.3.2.3 Probleme aufgrund von Lücken und Überlappungen

In der Folge des stichtagsbezogenen Abgleichs und der zeitlichen Differenz zwischen der LED und der Beschäftigtendatei ist es unmöglich, alle Episoden in den Erwerbsbiografien der Untersuchungsteilnehmer lückenlos zu rekonstruieren. So kann beispielsweise ein ehemals mit Überbrückungsgeld geförderter Existenzgründer, der inzwischen seine selbständige Tätigkeit wieder aufgegeben hat, zwischenzeitlich zwar arbeitslos, aber zum

¹¹ Im April 1999 wurde das Meldeverfahren der BST umgestellt. Denn seit diesem Zeitpunkt sind auch die so genannten geringfügig Beschäftigten sozialversicherungspflichtig und müssen entsprechend in der Beschäftigtendatei erfasst werden. Für die vorliegende Fragestellung erweist sich dabei als problematisch, dass die geringfügige Beschäftigung sich zum einen mit der Hauptbeschäftigung überlagern kann (d. h. eine Person kann theoretisch in der BST mehrmals auftauchen) Zum anderen kann ein „630-DM-Job“ auch von Leistungsempfängern ausgeübt werden. Da sich diese Unschärfen im Zeitverlauf kumulieren, verschlechtert sich die Datenqualität, je weiter sich der Untersuchungszeitpunkt vom April 1999 entfernt. Der Stand der BST zum Juni 1999 wird als ein Kompromiss zwischen zeitlicher Nähe zur Befragung einerseits und möglichst geringer Beeinträchtigung aufgrund des neuen Meldeverfahrens andererseits betrachtet.

30.06.1999 (dem Stichtag für die BST) wieder in abhängiger Beschäftigung gewesen sein. Dieser Fall würde dann im Datenabgleich als „sozialversicherungspflichtig beschäftigt“ gezählt werden. Umgekehrt ist es auch denkbar, dass ein ehemals Geförderter, der zum 30.06. 1999 in der BST registriert war, zugleich auch in der LED (Stichtag 31.12.1999) auftaucht, weil das Beschäftigungsverhältnis zwischenzeitlich beendet wurde. Da die LED gegenüber der BST die größere zeitliche Aktualität aufweist, wurden solche Fälle in der BST eliminiert und nur in der LED belassen.

Weiterhin ist es möglich, dass Überlappungen zwischen den Maßnahmeteilnehmern und den Beschäftigten auftreten, da beispielsweise die Teilnahme an Arbeitsbeschaffungs- oder Strukturanpassungsmaßnahmen als abhängige (und sozialversicherungspflichtige) Beschäftigung gilt und in der Statistik auch entsprechend erfasst wird.¹² Aufgrund der zeitlichen Inkongruenz der Datensätze ist es auch möglich, dass beim Datenabgleich Überschneidungen zwischen Leistungsempfängern und Maßnahmeteilnehmern auftreten.¹³ In all diesen Fällen wurde die jeweilige Maßnahme als gegenwärtiger Status betrachtet.

3.3.3 Befunde der Datenabgleiche

Von ursprünglich 4486 Geförderten in der Stichprobe konnten über die Datenabgleiche 1382 Fälle in den Geschäftsstatistiken der Bundesanstalt für Arbeit wiedergefunden werden (siehe Übersicht 4). Von diesen waren 1032 als sozialversicherungspflichtig beschäftigt registriert, 321 bezogen Lohnersatzleistungen vom Arbeitsamt und weitere 29 Personen nahmen an Maßnahmen der aktiven Arbeitsförderung teil.

Übersicht 4: Ergebnisse der Datenabgleiche Anfang 2000

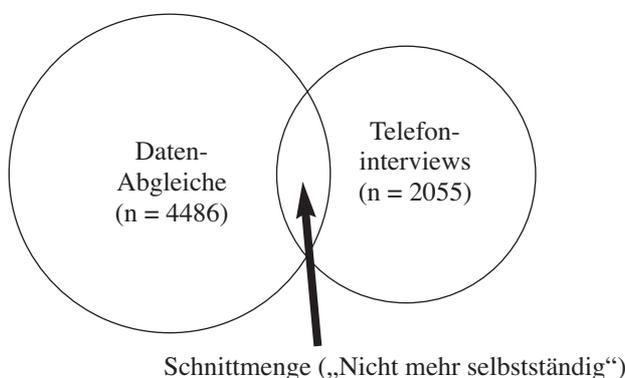
Datenbasis	Zahl der Treffer
Beschäftigtendatei	1032
Leistungsempfängerdatei	321
Maßnahmeteilnehmerdatei	29
Davon:	
ABM	19
SAM	6
LKZ	2
Beschäftigungshilfe für Langzeitarbeitslose	1
Erneutes ÜG	1
Insgesamt:	1382

Demnach haben also bis zum Untersuchungszeitpunkt 30,8 % der ehemaligen Überbrückungsgeld-Empfänger ihre selbständige Tätigkeit wieder aufgegeben. Die gesuchte Obergrenze liegt mithin bei etwa 69 % – wobei dieser Wert aufgrund der Lücken in den Geschäftsstatistiken wohl überhöht ist. In Anbetracht der Abweichungen, die zwangsläufig durch die Verbleibsvarianten entstehen,

die dem Datenabgleich nicht zugänglich sind (vgl. Übersicht 3), dürften 5 Jahre nach der Gründung kaum mehr als zwei Drittel der geförderten Existenzgründer noch selbständig gewesen sein.

Da die Datenabgleiche, wie bereits diskutiert, aufgrund verschiedener Friktionen lückenhaft bzw. unscharf ausfallen, stellt sich zwangsläufig die Frage nach der Validität dieser Resultate. Hierzu soll überprüft werden, inwieweit die Befunde der Datenabgleiche mit den empirischen Ergebnissen der Telefoninterviews deckungsgleich sind. Abbildung 4 zeigt, wie sich die Datenabgleiche und die Telefoninterviews überschneiden müssten. Die Schnittmenge bilden dabei all diejenigen ehemals Geförderten, die zum Untersuchungszeitpunkt nicht mehr selbständig tätig sind.

Abbildung 4: Schnittmenge zwischen Datenabgleichen und Telefoninterviews



Insgesamt konnten 493 Teilnehmer der Telefoninterviews auch bei den Datenabgleichen wieder aufgespürt werden. Da jedoch laut telefonischer Auskunft insgesamt 596 Befragte nicht mehr selbständig tätig waren, ergibt sich hier eine Mindererfassung von gut 13 %. Dies kann unter Umständen darauf hindeuten, dass die bei den Datenabgleichen zu berücksichtigenden Unschärfen sogar noch über den in Abschnitt 3.3.2.1 angenommenen Umfang hinausgehen.

Ogleich für die Schnittmenge eine weitestgehende Übereinstimmung zu erwarten gewesen wäre, kann das Ergebnis insgesamt noch als befriedigend betrachtet werden. Zwar weist der Kontingenzkoeffizient Cramer's V mit einem Wert von 0,64 nur auf einen mittelstarken Zusammenhang hin, doch sind die Abweichungen überwiegend systematisch und einigermaßen plausibel erklärbar. Bei den Datenabgleichen wurden insgesamt 104 Personen (87 sozialversicherungspflichtig Beschäftigte, 5 Leistungsempfänger, 12 Maßnahmeteilnehmer) ermittelt, die

¹² Konkret sind folgende Überlappungen mit der Beschäftigtenstatistik aufgetreten: 15 ABM-Fälle, 5 SAM-Fälle, 2 Fälle von Lohnkostenzuschüssen und eine Beschäftigungshilfe für Langzeitarbeitslose. Ein vormaliger Leistungsempfänger nimmt ebenfalls an einer Strukturanpassungsmaßnahme teil. Neu hinzugekommen sind 4 ABM-Fälle und eine erneute Förderung mit Überbrückungsgeld.

¹³ Beim vorliegenden Datenabgleich war dies genau einmal der Fall.

sich im Telefoninterview noch immer als Selbständige bezeichneten. Damit gab rund jeder fünfte Befragte einen anderen Status an, als er den Geschäftsstatistiken zufolge aufweisen sollte. Wie bereits in Abschnitt 3.3.2.2 angedeutet, ist jedoch ein Teil dieser Unschärfen möglicherweise auf die unterschiedlichen zeitlichen Aktualitäten der für die Abgleiche verwendeten Daten zurückzuführen. Und obgleich die Filterführung des Fragebogens derartige Überschneidungen verhindern bzw. identifizieren sollte, ist es dennoch denkbar, dass zumindest ein Teil dieser Fälle die ehemals geförderte selbständige Tätigkeit noch im Nebenberuf ausübt. Dies könnte Rückwirkungen auf die Selbsteinschätzungen der Befragten haben. Zugleich zeigt das Problem jedoch auch klar die Grenzen des Untersuchungsansatzes wie auch sozialwissenschaftlicher Analysen im Allgemeinen auf. Offenbar sind weder direkte Befragungen noch Geschäftsstatistiken gegen Unschärfen gefeit, besonders dann, wenn letztere im Grunde nicht zu Forschungszwecken geführt werden.

4 Determinanten des Teilnahmeverhaltens – Gewichtung als Lösung?

Nachdem dargelegt wurde, wie die Selbstselektivität der Untersuchungspersonen zu verschiedenen Zeitpunkten an verschiedenen Stellen der Untersuchung zum Tragen kommt¹⁴, war in den vorangegangenen Abschnitten der Versuch unternommen worden, durch die Bestimmung einer theoretischen Ober- und Untergrenze einen „Korridor“ festzulegen, in dem die tatsächliche Verbleibsquote der Geförderten mit großer Wahrscheinlichkeit zu verorten ist. Um diesen Zielbereich noch besser lokalisieren zu können, ist es erforderlich, die *Ursachen* der Selbstselektion zu ergründen. Könnte man die Determinanten des Teilnahmeverhaltens ermitteln, so wäre es in einem zweiten Schritt möglich, die empirisch festgestellte Verteilung des persönlichen Verbleibs der Geförderten so zu gewichten, dass auch die Nichtteilnehmer Berücksichtigung finden.

4.1 Das Problem der Selbstselektivität von Befragungsteilnehmern

Wie bereits in Kapitel 2 geschildert, liegt den Telefoninterviews von 1999/2000 eine Teilstichprobe der ursprünglichen Untersuchungsgesamtheit zugrunde, die – offenbar aufgrund verschiedener Einflüsse – selektiv ist und damit letztendlich verzerrte Befunde zur Folge hat. Die Frage, inwieweit die mit diesem offenkundigen Bias behafteten Ergebnisse noch oder überhaupt für die gesamten Untersuchungseinheiten sprechen können und im Weiteren einen Repräsentativitätsschluss auf die Grundgesamtheit gestatten, stellt zugleich eines der grundlegenden Probleme der Sozialwissenschaften dar. Das so genannte „Repräsentationsproblem“ thematisiert, inwieweit generell „... für ein empirisches Relativ eine zumindest homomorphe Abbildung oder Repräsentation existiert“¹⁵. Entsprechend hängt die Übertragbarkeit von Befunden zu einer Auswahl von Untersuchungseinheiten grundlegend davon ab, inwieweit diese die Grundgesamtheit repräsentieren, d. h. als ein verkleinertes Modell derselben gelten können.

In der Praxis der empirischen Sozialforschung finden sich als Ausprägungen dieser Problemstellung zum einen eher allgemeine messtechnische Tücken wie Reliabilität und Validität und insbesondere die so genannte Reaktivität¹⁶ als unmittelbare Reaktion der Probanden auf das Untersuchungsinstrument. Während hier jedoch die Antwortverzerrungen als Folge inhaltlich oder situativ bedingter Verhaltensweisen auftreten, resultiert aus dem selbstselektiven Teilnahmeverhalten unter Umständen eine überproportionale Beteiligung bestimmter Segmente der Gesamtpopulation. Trotz gegensätzlicher Ursachen sind die Wirkungen ähnlich. Aufgrund der verzerrten Messergebnisse kann die externe Validität der Untersuchung in Frage stehen. Zwar könnten geeignete Stichprobenverfahren oder Gewichtungen die Schief lagen im Datenmaterial zumindest teilweise kompensieren, doch setzt dies zunächst voraus, das Ausmaß der Artefakte zumindest annähernd zu beziffern, denn das entscheidende Problem besteht in der Unterschiedlichkeit der untersuchten und der nicht untersuchten Populationsteile hinsichtlich des Untersuchungsgegenstandes.¹⁷ Da die Teilnahme an der Befragung jedoch freiwillig war¹⁸, bestimmen in erster Linie die Untersuchungsteilnehmer und mehr noch die Nicht-Teilnehmer über die Übertragbarkeit der Untersuchungsergebnisse.¹⁹

Während Antwortausfälle oder Teil-Verweigerungen²⁰ bei einzelnen Fragen des Interviews („Item-Nonresponse“) hier nicht weiter thematisiert werden sollen, hat doch die völlige Teilnahmeverweigerung (Unit-Nonresponse) weitreichende Konsequenzen für die vorliegende Untersuchung. Hermann/Streng (1986: 337) konstatieren hierzu: „*Da davon (von einer rein zufallsbedingten Antwortbereitschaft der Befragten, d. Verf.) aber nicht ohne weiteres ausgegangen werden kann, ist für die Interpretation der Ergebnisse mit weniger als 100 %igem Rücklauf die Kenntnis der Ursachen von Antwortverweigerungen bzw. die Kenntnis von etwaigen Verzerrungsfaktoren wichtig. Und sie ist umso wichtiger, je größer die Antwortverweigerungsquote ausfällt.*“²¹ Vor leichtfertigem Umgang mit

¹⁴ Zur generellen Problematik (exogen verursachter) Bestandsdynamiken in Wiederholungsbefragungen siehe Rendtel/Pannenberg/Daschke (1997).

¹⁵ Heidenreich (1995: 346).

¹⁶ Vgl. hierzu ausführlich z. B. Reinecke (1991).

¹⁷ Esser (1975: 226) folgert hierzu: „Die theoretischen Überlegungen zu den Determinanten von Interaktionsaufnahme und Rollenübernahme in Sekundärkontakten ließen indessen wenig Zweifel daran, daß eine solche Unterschiedlichkeit zu erwarten ist über Prozesse der Selbstselektion von Befragten aus deren Reaktion auf die Interaktionsaufforderung zur Befragung.“

¹⁸ Ausnahmen hierzu bilden amtliche Stichprobenerhebungen wie z. B. der Mikrozensus oder die Arbeitskräftestichprobe der Europäischen Union, bei denen die Auskunftspflicht durch die §§ 7 und 12 Mikrozensusgesetz (MZG) in Verbindung mit § 15 BStatG gesetzlich vorgeschrieben ist.

¹⁹ Das Nonresponse-Problem ist besonders häufig bei schriftlichen Befragungen anzutreffen, wenn der Befragte ohne Konfrontation mit einem Interviewer und daher relativ frei von sozialen Zwängen über die Teilnahme entscheiden kann. Eine international vergleichende Meta-Analyse zu diesem Aspekt wurde von Hox/de Leeuw (1994) vorgelegt.

²⁰ Zu Non-Response oder Response Error siehe z. B. Hartmann/Kohaut (2000).

²¹ Aus diesem Grund versucht man beispielsweise den Fragebogenrücklauf mit verschiedenen Mitteln zu steigern.

Antwortausfällen warnen auch Hox/de Leeuw (1994: 330): „*Research results can be biased if the nonresponse is nonrandom, and if it is in some way correlated with the variables measured in the survey. Since the process leading to nonresponse is usually unknown, it is often optimistically assumed that when the response is high, there is no serious nonresponse bias.*“ Die daraus resultierenden Konsequenzen für die Untersuchung fasst Koch (1998: 66) folgendermaßen zusammen: „*Wenn Ausfälle bei der Realisierung von Zufallsstichproben nicht zufällig erfolgen, gefährden sie die Generalisierbarkeit von Umfrageergebnissen.*“ Einfach formuliert lautet nun die entscheidende Frage: Stellt die befragte Stichprobe noch eine echte Zufallsauswahl der ihr zugrunde liegenden Population dar oder gibt es systematische Ausfälle?

4.2 Selbstselektivität in der Untersuchung

In den vorangegangenen Abschnitten wurde deutlich, dass die Teilnehmer der zweiten Befragungsrunde eine Positivselektion darstellen. Insofern ist die Übertragbarkeit dieser Befunde auf die Grundgesamtheit grundsätzlich fraglich. Voreilig wäre es jedoch, den Vergleich der Telefoninterviews mit der restlichen Untersuchungsgesamtheit auf ein einziges Repräsentativitätsmerkmal, den Verbleib in Selbständigkeit, zu beschränken. Dieses allein lässt noch keine ausreichend gesicherten Rückschlüsse auf die „innere Repräsentativität“ dieser Teilgruppe für die gesamte Untersuchungspopulation zu. Stattdessen können sich die Verzerrungen aufgrund des Nonresponse sogar weiter fortpflanzen. Solange jedoch Auftreten, Ausmaß und Auswirkungen der Selektivität nicht hinreichend genau beziffert sind, bleibt auch der Aussagegehalt der Befragungsdaten unklar. Mithin ist zu klären, ob Teilnehmer und Nichtteilnehmer als – *zumindest einigermaßen* – homogen betrachtet werden und die Telefoninterviews *stellvertretend* für *alle* Untersuchungseinheiten sprechen können. Dann erst kann entschieden werden, ob eine Gewichtung der Daten erforderlich ist, um zu Aussagen über die Untersuchungsgesamtheit zu gelangen²².

4.3 Multivariate Betrachtung möglicher Determinanten selbstselektiven Teilnahmeverhaltens

Zwar gibt es eine ganze Reihe von Forschungsdesigns zur Untersuchung von Nonresponse²³, doch ist keine Methode frei von Problemen. Abgesehen von Designproblemen und praktischen Schwierigkeiten, wie z. B. einer teilweise recht restriktiven Datenschutzgesetzgebung, ist den Ansätzen vor allem ein weitgehend empiristisches und wenig theoriegeleitetes Vorgehen gemeinsam.

Obwohl es auch nach mehr als 50 Jahren Methodenforschung immer noch keine einheitliche Theorie des Interviews gibt²⁴, zeichnet sich zur Erklärung einer Reihe von Spezialproblemen im Rahmen der Interviewforschung (z. B. Intervieweffekte, Effekte sozialer Wünschbarkeit, Zustimmungstendenz, Non-Attitudes, Item-Nonresponse) eine Tendenz zur Verwendung von Rational-Choice-Theorien ab (Schnell 1997: 157). Als Erklärungsmodelle für das Befragtenverhalten sind vor allem handlungstheoretische Konstrukte, die das Verhalten über Merkmale

des Handelnden²⁵ und der Situation²⁶ unter Berücksichtigung der zeitlichen Stabilität der jeweiligen Determinanten bestimmen²⁷, weit verbreitet²⁸. Gleichwohl gestaltet sich die empirische Überprüfung entsprechender Hypothesen außerordentlich schwierig, da es aufgrund der so genannten „kontrafaktischen Situation“ praktisch ausgeschlossen ist, Gründe und Motive für die *Nichtteilnahme* zu eruieren. Das Problem besteht schlichtweg darin, dass für die Nonrespondenten weniger Daten vorliegen als für den Rest der Untersuchungspopulation.

Zwar bleibt damit das Entscheidungskalkül dieser Fälle weiter im Dunkeln, doch weist zumindest der Grundtenor der fragmentarischen Informationen, die die Interviewer in diesem Kontext erschließen konnten, darauf hin, dass generelle Ablehnung nur in seltenen Fällen ursächlich war. Ebenfalls unerforschbar bleiben Faktoren der Populationsselektion, die nicht über das intentionale Verhalten der Befragten wirken. Das wichtigste Beispiel hierzu ist die Nicht-Erreichbarkeit von Befragten.

Zur statistischen Korrektur des Nonresponse ist eine Reihe von Verfahren vorgeschlagen worden, die – je nach Ausgangssituation – ihre Vorzüge, aber auch ihre Tücken haben²⁹. Bei „einfachen“ Gewichtungungsverfahren werden die Korrekturgewichte durch den Vergleich der Verteilung bestimmter Merkmalskombinationen in der Stichprobe mit deren Verteilung in der Grundgesamtheit über einen „Soll-Ist-Vergleich“ ermittelt. Zwar wären im vorliegenden Fall Informationen über die Verteilung soziodemografischer oder regionaler Merkmale, wie sie meist als Gewichtungsvariablen verwendet werden, bekannt, doch haben frühere Untersuchungen (Wießner 2001: 131ff.) auch gezeigt, dass gerade diese Größen nur einen geringen Beitrag zur Erklärung des Gründungserfolges leisten.

Aussichtreicher erscheint hingegen die Verwendung eines Propensity-Modells, mit dessen Hilfe die Teilnahme-wahrscheinlichkeit, d. h. die Wahrscheinlichkeit, Respondent oder Nichtrespondent zu sein, geschätzt werden

²² Für eine theoretische Erklärung des Nonresponse siehe ausführlich Schnell (1997: 133ff.).

²³ Vgl. hierzu beispielsweise die Übersicht bei Schnell (1997: 134ff.).

²⁴ Groves (1987: 156; zitiert nach Schnell 1997: 157) beschreibt „... *survey research as a methodology without a unifying theory.*“

²⁵ Leverkus-Brüning (1966: 114ff.) untergliedert die „Merkmale des Handelnden“ in eine „Horizontkomponente“, die sich aus Wissens- und Bildungsvariablen generiert, sowie eine „Reserviertheitskomponente“, die Dimensionen wie Unschlüssigkeit, Zurückhaltung oder auch bewusste Verheimlichung des Befragten umfasst. Eine weitere personenorientierte Dimension ist das so genannte „Self-Involvement“, d. h. die persönliche Betroffenheit oder das Interesse des Befragten am Untersuchungsgegenstand.

²⁶ In einem weiter gefassten Verständnis ist auch die Aktualität des Untersuchungsgegenstandes zu den zeitlichen und räumlichen Aspekten der Befragungssituation zu zählen.

²⁷ Ein solches handlungstheoretisches Modell benutzen beispielsweise Hermann/Streng (1986: 339). Und Bebbington (1970: 170) konstatierte bereits vor mehr als 30 Jahren: „*Whether or not an individual agrees to participate in a particular survey seems to depend on several conditions – his nature, the subject of the survey, the technique of the interviewer.*“

²⁸ Einem solchen Ansatz folgen beispielsweise auch Hartmann/Kohout (2000: 612ff.) bei ihren Analysen zu Ausfällen im IAB-Betriebspanel.

²⁹ Ein Überblick findet sich bei Schnell (1997: 245ff.).

Exkurs: Modellkonstruktion

Die praktische Realisierung eines Modells, das diesen Anforderungen weitgehend gerecht wird, gestaltet sich jedoch als recht schwierig. Zum einen divergieren die zugrunde liegenden Verhaltensmodelle und handlungstheoretischen Konstrukte in nicht unerheblichem Maße, andererseits unterliegt das verfügbare Variablen-Set verschiedenen Restriktionen. Hosmer/Lemeshow (1989: 82) stellen hierzu fest: „*The goal of any method is to select those variables that result in a „best“ model within the scientific context of the problem.*“ Deshalb ist zur Modellkonstruktion vor allem zweierlei notwendig:

1. Eine Grundvorstellung, nach welchen Kriterien welche Merkmale in das Erklärungsmodell aufgenommen werden sollen, und
2. entsprechende Methoden zur Bestimmung der Modellgüte sowohl hinsichtlich der einzelnen Variable als auch im Hinblick auf das Gesamtmodell.

Zunächst sind also aus den verfügbaren Merkmalen diejenigen auszuwählen, deren Aufnahme in das Erklärungsmodell auch aus theoretischer bzw. induktiver Sicht Sinn macht. Daran anschließend gilt es, das Analysemodell zu verfeinern und zu optimieren, denn eine Modellspezifikation gilt vor allem dann als effizient und korrekt, wenn vor dem Hintergrund der theoretischen Vorüberlegungen weder zu viele (overfitting) noch zu wenige (underfitting) erklärende Variablen enthalten sind. Wird dieses Ziel verfehlt, so können sich verzerrte, inkonsistente oder ineffiziente Schätzer ergeben.

Forschungspraktisch ist diese Anforderung oft schwer zu erfüllen. Erfolgt die Selektion der Regressoren vor allem theoriegeleitet und nach inhaltlichen Aspekten, so besteht die Gefahr, das Modell durch Variablen von fragwürdiger statistischer Signifikanz zu „belasten“ und ein suboptimales Ergebnis zu erzielen. Dem gegenüber generieren streng statistisch begründete Auswahlverfahren oft „biologisch inplausible“ Modelle.

Entsprechend wurde versucht, diese beiden grundsätzlichen „Analyse-Philosophien“ soweit wie möglich zu synthetisieren. A priori ausgeschlossen wurden etwa solche Merkmale, zu denen nur relativ geringe Fallzahlen vorlagen oder deren Ausprägungen eine so schiefe Verteilung aufwiesen, dass sie aufgrund der geringen Varianz für die Regressionsanalyse praktisch unbrauchbar waren.

Daraufhin wurde für die verbliebenen Merkmale eine zunächst univariate Logit-Analyse durchgeführt. Hosmer/Lemeshow (1989) empfehlen, in diesem ersten Durchgang vorerst alle Variablen mit einem p-Wert $< 0,25$ als potenzielle Kandidaten für das multivariate Modell in Betracht zu ziehen. Zwar hat ein solch ungewöhnlich hoher p-Wert (verglichen mit der „traditionellen“ Signifikanzschwelle von $p < 0,05$) zur Folge, dass zunächst auch im streng statistischen Sinne „irrelevante“ Variablen in das Modell aufgenommen werden, doch bleiben auf diese Weise Determinanten, die für eine inhaltliche Interpretation erforderlich sind, vorläufig erhalten. Auf dieser Basis kann die Modellspezifikation so optimiert werden, dass die enthaltenen Merkmale großteils nicht nur signifikante, sondern zugleich auch sinnvoll interpretierbare Erklärungsbeiträge zu leisten vermögen. Das abschließende Modell ist in der Übersicht 5 dargestellt.

kann. Der Gewichtungsfaktor wäre folglich die inverse geschätzte Responsewahrscheinlichkeit³⁰.

Nachfolgend sollen nun mit Hilfe eines logistischen Regressionsmodells mögliche Einflüsse auf das Teilnahmeverhalten der Befragten entdeckt werden (Übersicht 5). Anschließend ist zu diskutieren, wie die aufgezeigten Schief lagen im Datenmaterial durch geeignete Gewichtungsverfahren korrigiert werden können.

Die Relevanz des Unternehmenserfolges, die sich bereits in den deskriptiven Analysen abzeichnete, wird in der multivariaten Betrachtung bestätigt: Der Verbleib in Selbständigkeit beeinflusst die Bereitschaft zur Teilnahme an weiteren Untersuchungen auf höchst signifikante Weise. Der Einfluss des Verbleibsmerkmals ist mit 99,9-prozentiger Sicherheit nicht zufälliger Natur. Wie der logarithmierte Regressionskoeffizient („odds ratios“) zeigt, ist die Wahrscheinlichkeit für die Teilnahme bei einem Selbständigen gegenüber einer Person, die ihre berufliche Selbständigkeit zum Zeitpunkt der schriftlichen Befragung im Sommer 1997 wieder aufgegeben hatte, fast dreimal so hoch.

Vor diesem Hintergrund verwundert es wenig, dass Probanden, die mit ihrer selbständigen Tätigkeit so zufrieden sind, dass sie sich auf jeden Fall wieder selbständig machen würden, ebenfalls eine höchst signifikante Teilnahmebereitschaft aufweisen. Aber auch solche Überbrückungsgeld-Empfänger, die eher zögerlich ein zweites Mal den Schritt in die Selbständigkeit unternehmen würden, sind noch immer signifikant eher bereit, sich an einer weiteren Befragungsrunde zu beteiligen als die Referenzgruppe, die sich auf keinen Fall mehr selbständig machen würde.

Die Hypothese, die Teilnahmebereitschaft an einer weiteren Welle der IAB-Untersuchung hänge nicht zuletzt auch von der Zufriedenheit mit dem Arbeitsamt ab, kann nicht verworfen werden. Zumindest bei der Gruppe derer, die mit dem Service ihres Arbeitsamtes sehr zufrieden waren, ist die Teilnahmebereitschaft signifikant höher als bei der Referenzgruppe und auch bei den anderen Ausprägungen dieses Merkmals weisen die Regressionskoeffizienten zumindest in die erwartete Richtung.

Inhaltlich schwieriger zu deuten, sind hingegen die beiden letzten Merkmale, die in das Modell aufgenommen wurden. Zwar sind sowohl hilfreiche Erfahrungen als auch verwertbare berufliche Kenntnisse aus der vormaligen abhängigen Beschäftigung für die Teilnahmebereitschaft von signifikanter Bedeutung, doch kann zur Wirkungsweise dieser beiden Variablen nicht leicht eine entsprechend plausible Hypothese formuliert werden. Vielleicht repräsentieren gerade diese beiden Einflussgrößen besonders gut die „Horizontkomponente“ von Leverkus-Brüning (1966; vgl. auch Fußnote 25). Möglicherweise

³⁰ Ein Analogon findet sich bei Rendtel (1990), der die Methode der reziproken Auswahlwahrscheinlichkeit vorschlägt als Verfahren für Hochrechnungen von Längs- und Querschnitten in Panelerhebungen und zur Korrektur von Stichprobenfehlern aufgrund demografischer Verluste oder wegen Panelmortalität. Zu Panelausfällen und Panelrepräsentativität siehe auch ausführlich Rendtel (1995).

Übersicht 5: Determinanten der Teilnahme an der telefonischen Wiederholungsbefragung 1999/2000 (logistische Regression)

Ausgewählte Merkmale	Nicht-standar-disierter Regressionskoeffizient	odds ratios (exp(b)) (= logarithm. R.koeffizient)
Gegenwärtiger Verbleib (Referenz: Nicht mehr selbständig) Noch selbständig	1,052***	2,863***
Noch einmal selbständig machen (Referenz: Auf keinen Fall) Auf jeden Fall Wahrscheinlich schon Wahrscheinlich nicht	1,079*** 0,505* 0,517*	2,941*** 1,657* 1,677*
Zufriedenheit mit dem Arbeitsamt (Referenz: Nicht zufrieden) Sehr zufrieden Zufrieden Teilweise zufrieden	0,445* 0,203 -0,076	1,560* 1,225 0,926
Hilfreiche Erfahrungen (Selbsteinschätzung) (Referenz: keine Erfahrung vorhanden) hilfreiche Erfahrung vorhanden	0,470**	1,600**
Verwertbare berufl. Kenntnisse (Referenz: sehr wenig/gar nichts) Sehr viel Ziemlich viel Doch einiges Wenig	0,519** 0,517* 0,337* 0,481*	1,680** 1,677* 1,401* 1,617*
Konstante	-0,635	0,530
Fälle im Modell	2586	
-2LogLikelihood (Step 0)	2333,519	
-2LogLikelihood (Step n)	2113,168	
-2LogLikelihood-Ratio-Test : X ² (df)	220,351 (12)	
R ² (Cox & Snell)	0,082	
R ² (Nagelkerke)	0,137	

Signifikanzniveau: * 95 %, ** 99 %, *** 99,9 %.

Hinweis: Die unabhängigen Variablen wurden bereits im Jahre 1997 mittels postalischer Befragung erhoben.

sehen Personen, die einen großen Teil ihres beruflichen Erfahrungswissens auch in der Selbständigkeit einsetzen können, eher die Chance, aus der Teilnahme an wissenschaftlichen Untersuchungen persönlichen Nutzen ziehen zu können. Indes erscheint in diesem Zusammenhang bemerkenswert, dass eine Reihe weiterer Humankapitalgrößen, die der Untersuchung ebenfalls zugänglich waren, die Voraussetzungen für die Aufnahme in das Modell ebenso wenig erfüllte wie die soziodemografischen Standardgrößen Geschlecht, Alter oder Familienstand.

4.4 Bestimmung von Gewichtungsfaktoren

Die Logit-Analyse zeigt, dass die Bereitschaft zur Teilnahme an der Wiederholungsbefragung ganz offensichtlich von einer Merkmalskombination abhängt, die – zumindest teilweise – erkennbar mit dem Gründungserfolg korreliert. Zugleich muss aber auch auf die relativ geringe Erklärungskraft des vorliegenden Modells hingewiesen werden. Der eher spärliche Anteil erklärter Varianz („Pseudo-R²“) verdeutlicht, dass in der Realität offenbar noch andere und scheinbar auch deutlich stärkere Einflussfaktoren zum Tragen kommen, die jedoch dem Untersuchungsausschnitt nicht zugänglich sind. Mithin hält sich der Erkenntnisgewinn hinsichtlich der Determinanten des Teilnahmeverhaltens in Grenzen und es gilt, was Hermann/Streng (1986: 349) bereits vor mehr als fünfzehn Jahren konstatieren mussten: „Abschließend bleibt festzuhalten, dass uns die Aufhellung des Befragungsdunkelfeldes nur in recht geringem Umfang gelungen ist.“

Gleichwohl soll im „Lichte“ dieser Einschränkungen zumindest ein Versuch unternommen werden, aus den Telefoninterviews Aussagen für die Grundgesamtheit abzuleiten. Um dabei auch die Nichtteilnehmer angemessen zu berücksichtigen, wird ein Gewichtungsfaktor gesucht, der folgende Korrektur realisiert:

$$\text{Verteilung (Telefoninterviews)} \times \text{Gewichtungsfaktor} = \text{Verteilung (insgesamt)}$$

Zur Ermittlung eines solchen Faktors kann wiederum auf das Analysemodell (Übersicht 5) zurückgegriffen werden. Denn bei der logistischen Regression ist die zu erklärende Variable (hier: Bereitschaft zur Teilnahme an weiteren Untersuchungen) entsprechend ihrem Skalenniveau dichotom und nominalskaliert. Der logistische Ansatz geht nicht von einem linearen, sondern einem logarithmischen Zusammenhang aus. Zugleich spezifiziert das Logit-Modell die abhängige Modell-Variable als Wahrscheinlichkeitsgröße („Propensity Score“). Damit besitzt die logistische Regression die Fähigkeit, die Eintrittswahrscheinlichkeit (WS) für ein bestimmtes Ereignis in Abhängigkeit von den erklärenden Variablen direkt abzuschätzen zu können. Dies kann mit folgenden Gleichungen formal dargestellt werden:

$$WS_i = \frac{e^{(B_0+B_1X)}}{1+e^{(B_0+B_1X)}} \quad \text{Gleichung 1.1}$$

oder einfacher

$$WS_i = \frac{1}{1+e^{-(B_0+B_1X)}} \quad \text{Gleichung 1.2}$$

für eine unabhängige Variable,

wobei

WS_i Wahrscheinlichkeit für den Eintritt eines bestimmten Ereignisses i , hier Teilnahme am Telefoninterview

B_0 die geschätzte Regressionskonstante

B_1 die geschätzte Regressionskoeffizient

X die unabhängige Variable

e die Eulersche Zahl als Basis des natürlichen Logarithmus, ($\cong 2,718\dots$).

Für mehr als eine unabhängige Variable gilt dann allgemein:

$$WS = \frac{1}{1+e^{-Z}} \quad \text{Gleichung 1.3}$$

wobei sich für die multivariate Variante der logistischen Regression der Wert von Z in Analogie zur linearen Regression durch die folgende Gleichung verallgemeinern lässt:

$$Z = B_0 + B_1X_1 + B_2X_2 + \dots + B_nX_n \quad \text{Gleichung 1.4}$$

Da die Verteilungen der Regressoren auch für diejenigen Fälle bekannt sind, die nicht an der zweiten Befragungsrunde teilgenommen haben, können die Eintrittswahrscheinlichkeiten zugleich zur Bildung von Gewichtungsfaktoren verwendet werden. Wie gezeigt ergeben sich die Wahrscheinlichkeiten, die Werte zwischen Null und Eins annehmen können, indem die ermittelten Regressionskoeffizienten in die Regressionsgleichung eingesetzt werden. Da in der Ausgangskonstellation hohe Werte dieses Faktors zugleich eine hohe Teilnahmewahrscheinlichkeit indizieren, mit der Gewichtung aber genau die Antipoden der Teilnehmer, nämlich die Nichtteilnehmer, besser berücksichtigt werden sollen, wird als Gewichtungsfaktor der Kehrwert der ermittelten Wahrscheinlichkeiten verwendet.

$$D. h. \text{ Gewicht} = \frac{1}{\text{Teilnahmewahrscheinlichkeit}}$$

Dies führt zu folgenden Verbleibsquoten:

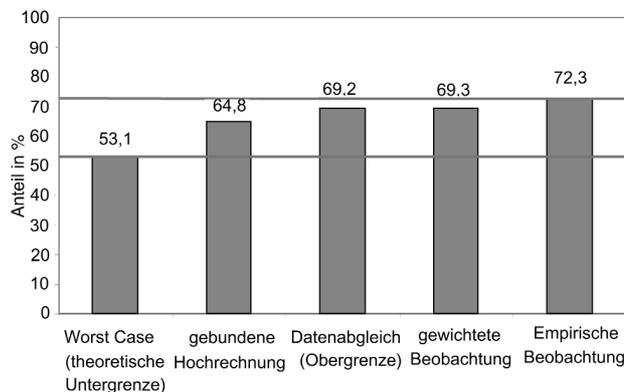
Übersicht 6: Situation der Befragungsteilnehmer 1999/2000 – gewichtet und ungewichtet

Verbleib	Anteil (ungewichtet)	Anteil (gewichtet)
selbständig	72,3 %	69,3 %
darunter:		
• geförderte Existenz	67,6 %	64,5 %
• andere selbständige Existenz	4,7 %	4,7 %
nicht mehr selbständig	27,7 %	30,7 %

Wie Übersicht 6 verdeutlicht, erfüllt die Gewichtung der empirischen Ergebnisse im Grundsatz ihre Aufgabe: Die als zu hoch erkannten Überlebensquoten werden etwas verringert, der Anteil der aus der Selbständigkeit wieder ausgeschiedenen ehemaligen Überbrückungsgeld-Empfänger nimmt hingegen zu³¹. Gleichwohl weisen auch die gewichteten Beobachtungen noch immer Verbleibswerte aus, die leicht über der durch Datenabgleich ermittelten Obergrenze liegen³². Da hier jedoch aufgrund systematischer Lücken eine Untererfassung der nicht mehr selbständig Tätigen in einer Größenordnung von gut 5 bis 10 % vorliegt (vgl. Abschnitt 3.3.2.3), müssen auch die gewichteten Werte noch immer als überhöht betrachtet werden.

Abbildung 5 zeigt die auf verschiedenem Wege ermittelten Verbleibsquoten der ehemals mit Überbrückungsgeld geförderten Gründerinnen und Gründer. Zwar hat der Korridor, in dem der „wahre“ Verbleibswert zu verorten ist, noch immer eine Bandbreite von annähernd 20 Prozentpunkten, doch sind sowohl der untere Extremwert („worst case“) als auch die Obergrenze (empirische Beobachtung) als eher unrealistisch einzustufen. So kann als weitgehend gesichert angenommen werden, dass fünf Jahre nach der Gründung wohl mehr als die Hälfte, jedoch sicher auch nicht mehr als zwei Drittel der ehemaligen Überbrückungsgeld-Empfänger noch selbständig tätig sind.

Abbildung 5: Auf verschiedenen Wegen ermittelte Verbleibsquoten ehemaliger Überbrückungsgeld-Empfänger etwa 5 Jahre nach der Gründung



5 Determinanten des Erfolges – Neuaufgabe eines alten Modells

Abschließend sollen nun die Determinanten des Erfolges mit Hilfe eines logistischen Regressionsmodells ermittelt

³¹ Um die grundsätzliche Brauchbarkeit der ermittelten Gewichtungsfaktoren bewerten zu können, müssten sich die Sensitivitätsanalysen auf mehrere Merkmale bzw. Merkmalskombinationen erstrecken und insbesondere auch die Beziehungen der Variablen untereinander berücksichtigen. Da es bei der vorliegenden Arbeit jedoch ausschließlich um den Verbleib der Geförderten geht, wird dieses Problem nicht weiter thematisiert.

³² Die Tatsache, dass die Gewichtungsfaktoren zwar grundsätzlich in die richtige Richtung, jedoch in unzureichendem Ausmaß weisen, erklärt sich möglicherweise dadurch, dass die regressionsanalytisch ermittelten Determinanten des Teilnahmeverhaltens zwar größtenteils die geforderte statistische Signifikanz aufweisen, insgesamt jedoch nur eine geringe Varianzaufklärung leisten.

werden. Die abhängige Variable ist der Verbleib der Untersuchungspersonen. Für das Logit-Modell wurde sie dichotomisiert, d. h. sie weist nur noch die beiden Ausprägungen „noch selbständig tätig“ und „nicht mehr selbständig tätig“ auf. Das zu untersuchende Ereignis wurde als „noch selbständig tätig“ definiert.

Ein solches Modell war bereits bei den früheren Untersuchungen mit der vorliegenden Stichprobe verwendet worden³³. Es stellt sich nun die Frage, ob die damals ermittelten Erfolgsfaktoren auch heute noch ihre Gültigkeit besitzen oder ob auf längere Sicht andere Einflussgrößen zum Tragen kommen. Zum Vergleich wurde das „alte“ Modell repliziert. Da ein Großteil der Teilnehmer der telefonischen Befragung ihre persönliche Einwilligung zur Verknüpfung der vorliegenden Datensätze gab, steht für die Analyse ein überaus umfangreiches Merkmalspektrum bereit. Zu bedenken ist jedoch, dass auch dieses Data Set mit einem Positiv-Bias behaftet ist, der für die multivariate Analyse mit den oben dargestellten Verfahren nicht sinnvoll korrigiert werden kann. Für eine einfache Gegenüberstellung von Tendenzen mag dies jedoch trotzdem ausreichen.

In Übersicht 7 werden die Regressionskoeffizienten für die Samples von 1997 und 1999/2000 einander gegenübergestellt. Da der unstandardisierte Regressionskoeffizient in der logistischen Regression – anders als beim linearen Regressionsmodell – nicht direkt interpretierbar ist, werden zur besseren Verständlichkeit der Befunde zusätzlich logarithmierte Koeffizienten, die so genannten „Odds Ratios“, ausgewiesen. Dieser Wert drückt aus, um welchen Faktor sich bei Vorliegen einer bestimmten Merkmalsausprägung die Wahrscheinlichkeit für den Eintritt des untersuchten Ereignisses gegenüber der Referenzkategorie ändert.

Gegenüber dem Ursprungsmodell hat sich die Erklärungskraft kaum verändert. Die konservativste der drei ausgewiesenen Kennziffern, das so genannte „Pseudo-R²“ liegt nun sogar etwas niedriger. Demnach werden durch das neue Modell 9,1 % der Varianz der abhängigen Variablen erklärt – gegenüber 9,9 % mit den Daten von 1997. Andere, an das R² der linearen Regression angelehnte Indizes fallen minimal günstiger aus, doch auch so kann die Anpassungsgüte des Modells nicht befriedigen. Zwar wird im Klassifizierungstableau für rund drei Viertel der Fälle das richtige Ergebnis prognostiziert, doch zeigt sich hier eine deutliche Schiefelage: Während das Modell für 1999/2000 fast 96 % der in Selbständigkeit Verbliebenen richtig einschätzt, kommt es bei denen, die ihre berufliche Selbständigkeit in der Zwischenzeit wieder aufgegeben haben, nur in jedem sechsten Fall zum richtigen Ergebnis. Diese Schwäche des Modells ist einerseits natürlich durch die schiefe Verteilung des Regressanden bedingt. Zum anderen zeigt dies aber auch, dass der Erfolg einer Unternehmensgründung zum großen Teil von Einflussgrößen abhängt, die der Untersuchung nach wie vor nicht zugänglich sind.

Die Gegenüberstellung der Modelle mit den Daten von 1997 und 1999/2000 zeigt weiterhin, dass die Befunde der früheren Untersuchung im Großen und Ganzen nach wie vor Gültigkeit haben. So gab es – mit Ausnahme der

Merkmale „Geschlecht“ und „Familienstand“, bei denen sich das Signifikanzniveau deutlich veränderte – nur marginale Änderungen bei einzelnen Ausprägungen. Auch ist die Wirkungsrichtung (d. h. das Vorzeichen) der untersuchten Einflussfaktoren weitgehend gleich geblieben und die Stärke der Einflüsse weist kaum sprunghafte Veränderungen auf.

Frauen haben immer noch schlechtere Erfolgsaussichten als Männer und im neuen Modell ist dieser Zusammenhang sogar hochsignifikant, d. h. mit einer Sicherheit von 99 % nicht zufälliger Natur. Die geschlechtsspezifischen Unterschiede dürften sich aber vor allem durch die unterschiedlichen Rahmenbedingungen für Männer und Frauen und die unterschiedlichen Gründungskonzepte erklären³⁴.

Die Erfolgsaussichten *ostdeutscher Gründer* sind immer noch günstiger als die der westdeutschen. Zudem hat sich das Signifikanzniveau der Herkunftsregion zum Untersuchungszeitpunkt Ende 1999 sogar noch erhöht. Im Lichte der gegenwärtigen ökonomischen Situation der neuen Bundesländer erscheint dieser Befund zunächst als sehr erstaunlich. Eine Erklärung könnte darin bestehen, dass die Untersuchungspersonen zum Zeitpunkt der Gründung 1994/95 zumindest teilweise noch an dem Wiedervereinigungsboom partizipieren konnten. Die damaligen Märkte im Osten waren sicher noch längst nicht gesättigt, so dass eine Etablierung in der Folgezeit gut vorstellbar ist.

Der *Familienstand* hat im Zeitverlauf der Untersuchung nicht nur seine höchst signifikante Bedeutung eingebüßt, sondern zugleich auch noch seine Wirkungsrichtung verändert. Eine plausible Erklärung für diese Entwicklung ist nicht leicht zu finden. Möglicherweise kommt an dieser Stelle der oben bereits angesprochene Bias im Datenmaterial zum Tragen.

Schon auf Grundlage der Erhebung von 1997 war das *Alter der Gründerpersonen* – abgesehen von der letzten Alterskategorie – nicht signifikant wichtig für den Erfolg in der Selbständigkeit. Mit den Daten von 1999/2000 ist nun auch dieser letzte Rest von nicht-zufälligem Einfluss verschwunden.

Erfolgsrelevant sind hingegen *hilfreiche Erfahrungen* aus früheren beruflichen Tätigkeiten. Interessanterweise zeigt sich auch bei der Neuauflage des Analysemodells die Einschätzung der verwertbaren Erfahrungen durch Außenstehende als hochsignifikant, während die Selbsteinschätzung der Befragten zu dieser Größe statistisch unbedeutend bleibt. Die hohe Bedeutung der Fremdeinschätzungen deutet darauf hin, dass externe Fachleute durchaus in der Lage sein können, die Erfolgsaussichten Gründungswilliger zutreffend zu taxieren³⁵. Ein weiterer

³³ Vgl. Wießner (2000), Wießner (2001)

³⁴ Vgl. z. B. Welter (2002), Lauxen-Ulbrich/Leicht (2002).

³⁵ Die Erfahrungen der Gründer wurden von den Forschungssachbearbeitern des IAB im Rahmen einer Aktenrecherche im Arbeitsamt bewertet. Hierzu ist kritisch anzumerken, dass den IAB-Mitarbeitern zum damaligen Zeitpunkt für einen Teil der Fälle bereits erste Verbleibsinformationen vorlagen, so dass es sich bei den Einschätzungen möglicherweise um Tautologien oder self-fulfilling prophecies handeln kann. Gleichwohl ist die Validität von Expertenurteilen im Gründungskontext nicht zu unterschätzen.

Übersicht 7: Logistisches Regressionsmodell zur Erklärung des Gründungserfolges

Ausgewählte Merkmale	Nichtstandardisierter Regressionskoeffizient		odds ratios (exp(b)) (= logarithm. R.koeffizient)	
	1997	1999/2000	1997	1999/2000
<i>Geschlecht (Referenz: Männer)</i>				
Frauen	-0,173	-0,398**	0,841	0,672**
<i>Alte / Neue Bundesländer (Referenz: West)</i>				
Ost	0,257*	0,390**	1,293*	1,477**
<i>Familienstand (Referenz: alleinstehend/sonstiges)</i>				
Verheiratet/zusammenlebend	0,421***	-0,036	1,523***	0,964
<i>Alter (Referenz: bis unter 30 Jahre)</i>				
30 bis unter 35 Jahre	0,027	0,180	1,027	1,198
35 bis unter 40 Jahre	0,065	0,003	1,067	1,003
40 bis unter 45 Jahre	0,064	0,102	1,066	1,107
45 bis unter 50 Jahre	-0,243	0,181	0,784	1,198
50 bis unter 55 Jahre	0,010	-0,190	1,010	0,827
55 Jahre und älter	-0,617*	-0,914	0,539*	0,401
<i>Hilfreiche Erfahrungen (Selbsteinschätzung) (Referenz: keine)</i>				
vorhanden	0,121	0,040	1,129	1,041
<i>Hilfreiche Erfahrungen (Fremdeinschätzung) (Referenz: keine)</i>				
vorhanden	0,547***	0,497***	1,729***	1,644***
<i>Verwertbare berufl. Kenntnisse (Referenz: sehr viel)</i>				
Ziemlich viel	-0,136	-0,120	0,873	0,887
Doch einiges	-0,276	-0,466*	0,759	0,627*
Wenig	0,051	-0,189	1,053	0,827
Sehr wenig/gar nichts	0,167	-0,231	1,182	0,794
<i>Gründungsberuf (Berufsbereiche) (Referenz: kein Dienstleistungsberuf)</i>				
Dienstleistungsberuf	-0,318	0,148	0,727	1,160
<i>Gründungsbranche (Referenz: nicht tertiärer Sektor)</i>				
Handel	-0,441	-0,785**	0,644	0,456**
Verkehr und Nachrichten	-0,195	-0,381	0,823	0,683
Kredit und Versicherung	-0,579*	-1,076***	0,561*	0,341***
Sonstige Dienstleistungen	-0,350	-0,616*	0,705	0,540*
<i>Branchenmobilität (Referenz: branchentreu)</i>				
Branchenwechsler	0,029	0,048	1,029	1,049
<i>Förderrecht (Referenz: „altes“ Recht)</i>				
„neues“ Förderrecht	-0,075	-0,360	0,928	0,698
<i>Gesamtbetrag Überbrückungsgeld (Referenz: bis zu 3.000 DM)</i>				
3.001 bis 6.000 DM	0,174	0,147	1,190	1,158
6.001 bis 9.000 DM	0,177	0,270	1,194	1,310
9.001 DM und mehr	0,365	0,249	1,441	1,283
<i>Höhe des Startkapitals (Referenz: kein Startkapital)</i>				
Unter 5.000 DM	-0,286	-0,064	0,751	0,938
5.000 bis unter 10.000 DM	0,426**	0,416	1,531**	1,515
10.000 bis unter 25.000 DM	0,457***	0,528**	1,579***	1,696**
25.000 bis unter 50.000 DM	0,817***	0,756***	2,264***	2,130***
50.000 bis unter 100.000 DM	1,156***	1,074***	3,178***	2,926***
100.000 DM und mehr	1,455***	1,674***	4,285***	5,333***
Konstante	0,447	0,872		
Fälle im Modell	2213	1714		
-2LogLikelihood (Step 0)	2497,001	2013,620		
-2LogLikelihood (Step n)	2272,892	1830,802		
-2LogLikelihood-Ratio-Test : X ² (df)	224,209 (31)	182,818 (31)		
„Pseudo-R ² “	≈ 9,9 %	≈ 9,1 %		
R ² (Cox & Snell)	0,096	0,101		
R ² (Nagelkerke)	0,142	0,146		

Signifikanzniveau: * 95 %, **99 %, *** 99,9 %.

Baustein aus dem Bereich des Humankapitals, die „wertbaren beruflichen Kenntnisse“, weist nur für eine einzige Ausprägung die geforderte statistische Signifikanz auf, so dass hier allenfalls in der Tendenz eine Bestätigung der tragenden Rolle von Know-How für den Gründungserfolg konstatiert werden kann.

Die Befunde aus früheren Arbeiten legten nahe, im multivariaten Modell auch den Einfluss eines Branchenwechsels auf den Gründungserfolg einer kritischen Würdigung zu unterziehen oder umgekehrt zu hinterfragen, inwieweit die durch einen Wechsel des Wirtschaftszweiges bedingten Risiken durch die bereits diskutierten beruflichen Erfahrungen, Kenntnisse oder andere Humankapitalgrößen (teil-)kompensiert werden können. In der Gesamtbetrachtung gehen von einem *Branchenwechsel* beim Schritt in die Selbständigkeit allerdings keine signifikanten Einflüsse auf den Unternehmenserfolg aus.

Neben der Relevanz beruflicher und sektoraler Mobilität stellt sich weiterhin die Frage nach bestimmten „Erfolgsberufen“ oder besonders vielversprechenden *Wirtschaftszweigen*. Da auch im Rahmen der Überbrückungsgeld-Förderung das Gründungsgeschehen dominiert wird von Serviceberufen und Gründungen im Dienstleistungssektor, wurden alle Nicht-Dienstleistungsberufe zu einer gemeinsamen Kategorie zusammengefasst. Bei den Wirtschaftszweigen wurden primärer und sekundärer Sektor ebenfalls zu einer gemeinsamen Kategorie aggregiert, während die verschiedenen Arten von Dienstleistungen ungruppiert blieben.

Für einen Hinweis auf Vor- oder auch Nachteile der selbständigen Ausübung eines *Dienstleistungsberufes* fehlt im neuen Logit-Modell die statistische Signifikanz. Bei der Betrachtung der *Gründungsbranchen* zeigen sich nur geringe Unterschiede zwischen den beiden Modellen. Wie zu erwarten war, sind die Erfolgsaussichten im Dienstleistungsgewerbe deutlich schlechter als in den Referenzbranchen. Abgesehen vom Bereich „Verkehr und Nachrichten“ ist für alle Ausprägungen des tertiären Sektors eine Signifikanz auf mindestens 95 %-Niveau nachgewiesen. So haben beispielsweise Gründungen im Wirtschaftszweig „Kredit und Versicherung“ *ceteris paribus* nur etwa ein Drittel der Erfolgswahrscheinlichkeit von Start-ups in der Urproduktion oder im verarbeitenden Gewerbe. Zwar bieten die Dienstleistungsberufe ebenso wie der Sektors durch die niedrigen Einstiegsbarrieren vordergründig ein attraktives Umfeld. Doch gewinnt gerade dadurch der Wettbewerb eine Intensität, die offenbar weder durch Nischenstrategien noch durch berufliche oder sektorale Mobilität vollständig kompensiert werden kann.

Der *leistungsrechtliche Hintergrund* der Geförderten wird in den Modellen repräsentiert durch die Fördervarianten³⁶ „altes Recht“ versus „neues Recht“ sowie die Höhe des insgesamt bezogenen Überbrückungsgeldes. Da das Überbrückungsgeld zumindest bei den so genannten „Neufällen“ in seiner Höhe dem Arbeitslosengeld entspricht, kann dieses Merkmal unter Vorbehalt als Proxy-Variable für das vormalige Erwerbseinkommen aus abhängiger Beschäftigung betrachtet werden. Für beide X-

Variablen weisen die Befunde beider Modelle in die selbe Richtung, jedoch ohne statistische Signifikanz³⁷.

Bekanntermaßen ist das Überbrückungsgeld zur Sicherung des Lebensunterhaltes und nicht zur Finanzierung des Gründungsvorhabens vorgesehen. Deshalb sollte es insbesondere im Zusammenhang mit der *Kapitalausstattung* der Gründer betrachtet werden. In beiden Modellen erweist sich die Höhe des verfügbaren Startkapitals als eine wichtige Erklärungsvariable im Hinblick auf den Verbleib der Geförderten in Selbständigkeit. Zwischen der Höhe der Startinvestitionen und dem Erfolg zeigt sich ein ziemlich klarer proportionaler Zusammenhang. Wie auch schon im alten Modell bleibt lediglich die erste Ausprägung „bis unter 5.000 DM Startkapital“ insignifikant. Dies ist möglicherweise auf missverständliche Angaben der Befragten zurückzuführen. Denn im betriebswirtschaftlichen Sinne ist die hier als Referenzkategorie verwendete Ausprägung „kein Startkapital“ unsinnig – Gründungen ohne Kapital gibt es nicht. Anzunehmen ist, dass bei diesen Antworten die in die Unternehmung eingebrachten Sachmittel nicht berücksichtigt wurden. Dies würde zugleich erklären, warum Kleinstgründungen mit einer Kapitalausstattung von weniger als 5.000 DM eine geringere Erfolgswahrscheinlichkeit aufweisen als die Referenzkategorie. Wer sich mit einem Startkapital von mindestens 25.000 DM selbständig machte, hat auch bei Verwendung des Datenmaterials von 1999/2000 mehr als doppelt so hohe Erfolgsaussichten wie ein Gründer ohne Kapital. Bei einem Startkapital von 100.000 DM sind die Chancen mittlerweile gut fünfmal so hoch und damit noch günstiger als in der ersten Analyse. Die herausragende Bedeutung der Höhe der verfügbaren Finanzmittel entspricht durchaus der allgemeinen Lebenserfahrung, insofern sind gerade diese Befunde wenig überraschend. Zu bedenken ist allerdings, dass mit einer Startinvestition von 100.000 DM nach aller Wahrscheinlichkeit ganz andere Gründungen in anderen Branchen realisiert werden, als bei Unternehmungen, die anfangs mit 10.000 DM auskamen. Bei der Diskussion über die Bedeutung des Startkapitals sollte man deshalb das Gebot der Kommensurabilität nicht aus den Augen verlieren: Nur gleiche oder zumindest ähnliche Unternehmen können sinnvoll miteinander verglichen werden.

³⁶ Als Folge einer Gesetzesänderung zum 01.08.1994 wurde das Überbrückungsgeld wesentlich komfortabler ausgestattet. Während die alte Regelung Überbrückungsgeld für bis zu 26 Wochen bis maximal der Höhe der Lohnersatzleistung vorsah, wurde dann ÜG für grundsätzlich 6 Monate in Höhe des Arbeitslosengeldes bzw. der Arbeitslosenhilfe gewährt. Die durchschnittliche Förderdauer hat sich dadurch für die „Neufälle“ mehr als verdoppelt, die durchschnittliche pro-Kopf-Förderung erhöhte sich sogar um den Faktor 2,5. Da der Untersuchung eine Stichprobe mit Geförderten zugrunde liegt, die sich zwischen Anfang 1994 und etwa Mitte 1995 mit Hilfe des Arbeitsamtes selbständig machten, kann hier unterschieden werden zwischen den „Altfällen“, die nach dem alten Recht gefördert wurden und den „Neufällen“, die in Genuss der neuen, großzügigeren Ausgestaltung des Förderprogramms kamen.

³⁷ Möglicherweise sind bestimmte Selektionsmechanismen dafür verantwortlich, dass die „Neufälle“ trotz der höheren Förderung über einen längeren Zeitraum im Vergleich zu den „Altfällen“ weniger erfolgreich abschneiden. Dabei ist jedoch bei diesen Fällen umgekehrt mit höherer Mitnahme zu rechnen. Siehe ausführlich Wießner (2001: 173ff.).

6 Fazit

Mit der vorliegenden Arbeit sind die Grenzen des Untersuchungsdesigns erreicht. Die zunehmenden Unschärfen im Datenmaterial und die wachsenden Schwierigkeiten bei der Korrektur dieser Verzerrungen könnten allenfalls durch deutlich aufwendigere und ressourcenintensivere Untersuchungsmethoden gemildert werden. Gleichwohl ist auch auf dieser Datenbasis klar erkennbar, dass die ehemaligen Überbrückungsgeld-Empfänger noch immer eine respektable Bilanz aufweisen – zumindest auf der Brutto-Ebene. Da eine geeignete Vergleichsgruppe nicht zur Verfügung steht, ist zu den Netto-Wirkungen der Förderung nach wie vor keine Aussage möglich. Doch spricht die Tatsache, dass gut fünf Jahre nach der Gründung noch immer mindestens jeder zweite der vormals Arbeitslosen sein Unternehmen betreibt, jedenfalls für die Qualität der geförderten Gründungen. Auch unter denen, die ihre berufliche Selbständigkeit in der Zwischenzeit wieder aufgegeben haben, finden sich nur wenige wieder beim Arbeitsamt ein. Verglichen mit den Befunden anderer Gründerstudien ist für die Gründungsverläufe der Geförderten alles in allem ein ähnlicher Verlauf zu konstatieren wie bei Start-ups, die ohne das Handicap vorangegangener Arbeitslosigkeit getätigt wurden.

Auch der Blick auf die Erfolgsfaktoren zeigt wenig Veränderung. Die kritischen Faktoren liegen nach wie vor im Bereich der Ausstattung mit Humankapital und Startkapital. Eine Einschätzung, inwieweit ergänzende Förderprogramme auf diesen Gebieten eine sinnvolle Flankierung bieten können, muss künftigen Untersuchungen vorbehalten bleiben. Die vielfältigen Initiativen und Angebote, die in den vergangenen Jahren neu geschaffen wurden, können dazu beitragen, weitere Gründungspotenziale zu erschließen. Es wäre jedoch zu kurz gegriffen, sich allein auf eine bunte und üppig ausgestattete Förderlandschaft verlassen zu wollen. Denn wie die Analysen zeigen, hängen Erfolg und Nachhaltigkeit der geförderten Gründungen zuallererst vom Wollen und Können der Gründerpersonen ab.

Wenn auch der Beitrag des Überbrückungsgeldes zu diesen positiven Entwicklungen ebenso wenig genau beziffert werden kann wie das Ausmaß von Verdrängungs- und Mitnahmeeffekten, die den Nettoerfolg des Förderinstrumentes auf der Arbeitsmarktebene schmälern, so bleibt doch auf jeden Fall festzuhalten, dass das Überbrückungsgeld einen wichtigen Brückenschlag zurück in den Arbeitsmarkt leistet. Und daran können auch jene teilhaben, die in den neugegründeten Unternehmen Beschäftigung finden.

Literaturverzeichnis

- Bebbington, A.C. (1970): The Effect of Non-Response in the Sample Survey with an Example. In: *Human Relations*, Vol. 23, No. 3, 169-180.
- Brinkmann, Christian/Otto, Manfred/Wiedemann, Eberhard (1995): Existenzgründungen mit Hilfe der BA. IAB-Werkstattbericht Nr. 10/1995 vom 22.11.1995.

- Esser, Hartmut (1986a): Können Befragte lügen? In: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 38, 314-336.
- Esser, Hartmut (1986b): Über die Teilnahme an Befragungen. In: *ZUMA-Nachrichten* 18, 38-47.
- Esser, Hartmut (1975): Soziale Regelmäßigkeiten des Befragtenverhaltens. Meisenheim an der Glan: Verlag Anton Hain.
- Groves, Robert M. (1987): Research on Survey Data Quality. In: *Public Opinion Quarterly* 4, 156-172.
- Hartmann, Josef/Kohaut, Susanne (2000): Analysen zu Ausfällen (Unit-Nonresponse) im IAB-Betriebspanel. In: *MittAB* 4, 609-618.
- Heidenreich, Klaus (1995): Grundbegriffe der Meß- und Testtheorie. In: Roth, Erwin (Hrsg.): *Sozialwissenschaftliche Methoden. Lehr- und Handbuch für Forschung und Praxis*. München, Wien: Oldenbourg Verlag. 4. Aufl.
- Hermann, Dieter/Streng, Franz (1986): Das Dunkelfeld der Befragung. Unit-nonresponse und item-nonresponse bei einer schriftlichen Befragung von Richtern und Staatsanwälten. In: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 2, 337-351.
- Hosmer, David W./Lemeshow, Stanley (1989): *Applied Logistic Regression*. New York: Wiley.
- Hox, Joop J./de Leeuw, Edith (1994): A Comparison of Nonresponse in Mail, Telephone and Face-to-Face Surveys. Applying multilevel Modelling to Meta-Analysis. In: *Quality & Quantity* 28, 329-344.
- Kaiser, Manfred/Otto, Manfred (1990): Übergang von Arbeitslosigkeit in berufliche Selbständigkeit. Erste Ergebnisse aus IAB-Verlaufserhebungen bei Überbrückungsgeld-Empfängern nach §55a AFG der Jahre 1986-1988. In: *MittAB* 2, 284-299.
- Koch, Achim (1998): Wenn ‚mehr‘ nicht gleichbedeutend mit ‚besser‘ ist: Ausschöpfungsquoten und Stichprobenverzerrungen in allgemeinen Bevölkerungsumfragen. In: *ZUMA-Nachrichten* 42, 22. Jg., Mai 1998, 66-90.
- Kühnel, Steffen/Jagodzinski, Wolfgang/Terwey, Michael (1989): Teilnehmen oder boykottieren: Ein Anwendungsbeispiel der binären logistischen Regression mit SPSSx⁴. In: *ZA-Informationen* 25, 11/89, 44-75.
- Lauxen-Ulbrich, Maria/Leicht, René (2002): *Gründerinnen und selbständige Frauen in Deutschland: Entwicklung, Potenziale und Profile*. Institut für Mittelstandsforschung Mannheim: Strukturbericht – Kurzinformativ Nr. 7/2002.
- Leverkus-Brüning, Iris (1966): Die Meinungslosen. Die Bedeutung der Restkategorie in der empirischen Sozialforschung. Berlin: Duncker & Humblot.
- Reinecke, Jost (1991): *Interviewer- und Befragtenverhalten. Theoretische Ansätze und methodische Konzepte*. Opladen: Westdeutscher Verlag.
- Rendtel, Ulrich/Pannenberg, Markus/Daschke, Stefan (1997): Die Gewichtung der Zuwanderer-Stichprobe des Sozio-oekonomischen Panels (SOEP). In: *DIW-Vierteljahrshefte* 2, 271-286.
- Rendtel, Ulrich (1995): *Lebenslagen im Wandel: Panelausfälle und Panelrepräsentativität*. Frankfurt/New York: Campus Verlag.
- Rendtel, Ulrich (1990): *Hochrechnung und Stichprobenfehler in Panelerhebungen*. Frankfurt/Main.

- Schnell, Rainer (1997): Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen. Opladen: Leske + Budrich.
- Urban, Dieter (1993): Logit-Analyse. Statistische Verfahren zur Analyse von Modellen mit qualitativen Response-Variablen. Stuttgart/Jena/New York: Gustav Fischer Verlag.
- Welter, Friederike (2002): Gründerinnen in Deutschland: Potenziale und das institutionelle Umfeld. Essen: RWI.
- Wießner, Frank (2001): Arbeitslose werden Unternehmer. Eine Evaluation von Existenzgründungen vormals Arbeitsloser mit dem Überbrückungsgeld nach § 57 SGB III (vormals § 55a AFG). Beiträge zur Arbeitsmarkt- und Berufsforschung (BeitrAB) 241. Nürnberg.
- Wießner, Frank (2000): Erfolgsfaktoren von Existenzgründungen aus der Arbeitslosigkeit. Eine multivariate Betrachtung vormals arbeitsloser Existenzgründer, die vom Arbeitsamt mit dem Überbrückungsgeld (§ 57 SGB III) gefördert wurden. In: MittAB 3, 518-532.
- Wießner, Frank (1998): Das Überbrückungsgeld als Instrument der Arbeitsmarktpolitik - eine Zwischenbilanz. In: MittAB 1, 123-142.