

BIBB/BauA-
Erwerbstätigenbefragung
2005/2006

Strukturkontrolle, Steuerung und
Gewichtung der Stichprobe

BIBB/BAuA-Erwerbstätigtenbefragung 2005/2006

Strukturkontrolle, Steuerung und Gewichtung der Stichprobe

Vorgelegt von

TNS Infratest Sozialforschung, München

Josef Hartmann

München, Juni 2006
Gew_B_45466

Inhalt	Seite
1. Vorbemerkung	4
2. Grundgesamtheit	5
3. Das Gewichtungsmodell	6
3.1 Gewichtung aufgrund von Stichprobendesign und Ausfallgeschehen	6
3.2 Designgewichtung	7
3.2.1 Auswahlchance des Haushalts	7
3.2.2 Auswahlchance der Person	7
3.3 Korrektur des Ausfallgeschehens über Strukturgewichtung	8
4. Steuerung der Stichprobe	17

1. Vorbemerkung

Das Bundesinstitut für Berufsbildung (BIBB) und die Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) haben im August 2005 TNS Infratest Sozialforschung mit der Durchführung der „BIBB/BAuA-Erwerbstätigenbefragung 2005/2006“ beauftragt. Ziel der Erhebung war es, die Arbeitsbedingungen und –anforderungen, den Erwerb und die Verwertung beruflicher Kenntnisse sowie die diesbezüglichen Entwicklungen und Veränderungen im Zuge technischen und organisatorischen Wandels der Arbeitswelt zu erfassen.

Zu diesem Zweck wurden 20.000 erwerbstätige Personen im Alter ab 15 Jahren telefonisch mit einem computerunterstützten Erhebungsinstrument (CATI) befragt. Vor dem Beginn der Haupterhebung wurden zwei Pretests durchgeführt. Diese Pretests wurden in einem eigenen Bericht dokumentiert.¹ Daneben werden in weiteren Berichten die Feldarbeit und die Vercodung der offenen Angaben zu den Erwerbs- und Ausbildungsberufen sowie zu den Branchen der Erwerbsberufe dargestellt. Der vorliegende Bericht beschreibt die Ergebnisse der Stichprobensteuerung, die Entwicklung des Gewichtungsmodells sowie die Analysen zur Strukturkontrolle.

Das Projekt wurde bei TNS Infratest Sozialforschung von einem Projektteam bearbeitet, dem Agnes Fischer, Miriam Gensicke und Josef Hartmann in der Projektleitung, Günter Steinacker in der Projektkoordination, Werner Abenthum, Manfred Krapf und Ralf Zetzmann in der Fragebogenprogrammierung, Martin Pfister und Susanne Schmitz bei der Gewichtung sowie Sibylla Zynda im Datenmanagement und Eva Mitka für die tabellarische Auswertung angehörten.

¹ TNS Infratest Sozialforschung, 2005: BIBB/BAuA-Erwerbstätigenbefragung 2005/2006 – Pretestbericht. München.

2. Grundgesamtheit

Grundgesamtheit der Untersuchung waren Erwerbstätige ab 15 Jahren mit einer Arbeitszeit von mindestens zehn Stunden pro Woche in Deutschland. Als erwerbstätig wurde definiert, wer zum Zeitpunkt der Befragung eine bezahlte bzw. eine mit einem Einkommen verbundene Tätigkeit ausübt. Diese Definition muss an verschiedenen Stellen präzisiert werden:

- Personen, die ihre Tätigkeit kurzzeitig, d. h. bis zu einer Dauer von maximal drei Monaten, unterbrochen hatten, wurden einbezogen. Hierbei konnte es sich um Mutterschutz oder Elternzeit sowie um Sonderurlaub aus einem anderen Grund handeln.
- Auch mithelfende Familienangehörige wurden befragt. Sie zählen als erwerbstätig, obwohl sie u. a. darüber definiert sind, dass sie ohne feste Vergütung im Betrieb oder Geschäft des anderen Familienmitglieds mitarbeiten.
- Ausländer wurden bei der Untersuchung einbezogen, wenn sie ausreichend deutsch sprachen.²
- Dagegen wurden Personen, die der bezahlten Tätigkeit ehrenamtlich nachgehen, nicht befragt. Hintergrund ist, dass viele ehrenamtliche Tätigkeiten, für die eine Aufwandsentschädigung erstattet wird, keine Erwerbstätigkeit darstellen, da mit ihnen keine Einkünftezielungsabsicht verbunden ist. Und selbst wenn Aufwandsentschädigungen für Verdienstausfall oder für Zeitaufwand gewährt werden und somit eine Einkünftezielungsabsicht vorliegt, unterliegen die Tätigkeiten dann nicht der Einkommensteuer und stellen daher im steuerrechtlichen Sinne keine Erwerbstätigkeit dar, wenn eine Steuerbefreiung wie z. B. nach § 3 Nr. 26 EstG vorliegt.³ Eine Einkünftezielungsabsicht wird zudem nicht bei politischen Mandatsträgern unterstellt, die diese Tätigkeit ehrenamtlich ausüben. Für sie wird davon ausgegangen, dass Aufwandsentschädigungen, die sie erhalten, eine Erstattung für tatsächlich entstandene Aufwendungen darstellen.
- Ausgeschlossen wurden schließlich auch Personen in einem Beschäftigungsverhältnis, das Teil einer Erstausbildung ist.

² Nimmt man an, dass die Ausländer, die so in die Befragung einbezogen werden konnten, etwa die Hälfte aller in Deutschland lebenden Ausländer repräsentieren, und geht man beispielsweise von der Annahme aus, dass der Anteil bezüglich eines bestimmten Merkmals in dem nicht repräsentierten ausländischen Bevölkerungsteil 5 Prozentpunkte höher liegt als unter den einbezogenen Bevölkerungsgruppen, dann ergibt sich rechnerisch eine Unterschätzung des gesamtwirtschaftlichen Anteils in der Größenordnung von 0,25 Prozentpunkten.

³ § 3 Nr. 26 EstG definiert den sog. „Übungsleiterfreibetrag“: Er beträgt pro Jahr 1.848 € bzw. 154 € monatlich. Bis zu diesem Betrag unterliegen Aufwandsentschädigungen für Zeitaufwand oder für Verdienstausfall nicht der Einkommensteuer; die entsprechenden ehrenamtlichen Tätigkeiten werden daher nicht als Erwerbstätigkeit eingestuft. Der Übungsleiterfreibetrag gilt allerdings nur für „Einnahmen aus nebenberuflichen Tätigkeiten als Übungsleiter, Erzieher, Betreuer oder vergleichbaren nebenberuflichen Tätigkeiten, aus nebenberuflichen künstlerischen Tätigkeiten oder der nebenberuflichen Pflege alter, kranker oder behinderter Menschen im Dienst oder im Auftrag einer inländischen juristischen Person des öffentlichen Rechts oder einer unter § 5 Abs. 1 Nr. 9 des Körperschaftssteuergesetzes fallenden Einrichtung zur Förderung gemeinnütziger, mildtätiger und kirchlicher Zwecke“ (§ 3 Nr. 26 EstG).

3. Das Gewichtungsmodell

3.1 Gewichtung aufgrund von Stichprobendesign und Ausfallgeschehen

Die Befragung wurde telefonisch durchgeführt. Die Auswahl der zu befragenden Zielperson erfolgt bei einer telefonischen Befragung in zwei Schritten: Im ersten Schritt wird eine Telefonnummer und damit letztendlich ein Haushalt ausgewählt, im zweiten Schritt dann innerhalb des Haushalts die Zielperson.⁴ Aufgrund dieses Vorgehens haben telefonbasierte Befragungen zwei Designkomponenten, die die Proportionalität der Abbildung beeinträchtigen: Die von der Zahl der für Gespräche nutzbaren Anschlüsse abhängige Auswahlwahrscheinlichkeit für Telefonhaushalte und die von der Zahl der zur Zielgesamtheit gehörenden Personen im Haushalt abhängige Auswahlwahrscheinlichkeit für Personen. Darüber hinaus gibt es einen zweiten Ursachenkomplex für Abbildungsprobleme: Einerseits ist bei Zufallsstichproben ein Zufallsfehler unvermeidlich, andererseits sind bei jeder auf Freiwilligkeit basierenden Befragung nonresponse-bedingte Selektivitäten zu konstatieren.

Die Ursachen für Abbildungsprobleme gilt es zu prüfen und im Rahmen des Gewichtungs- und Hochrechnungsmodells zu korrigieren. Das geschieht nach folgenden Grundsätzen:

- Designgewichtung

Die designbedingten Unterschiede in den Auswahlwahrscheinlichkeiten sind theoretisch und praktisch leicht mittels Gewichtung zu korrigieren. Dazu bedarf es lediglich der im Interview erhobenen Zahl der Festnetzanschlüsse, über die der Haushalt erreichbar ist, sowie der Zahl der im Haushalt lebenden Personen der Zielgruppe.

- Ausfallgewichtung

Anders ist es mit der Korrektur der Abbildungsprobleme, die durch selektive Ausfälle entstehen. Denn dazu bedürfte es einer Theorie des Ausfallgeschehens und daraus abgeleiteter Korrekturprozesse. Eine solche Theorie steckt noch in den Anfängen, daraus ableitbare Gewichtungsprozeduren, die den Ausfallprozess nachvollziehen und korrigieren, sind noch in weiter Ferne. Ausfallgewichtungen bei Querschnittsstichproben können sich daher derzeit nur an den Strukturen der realisierten Stichprobe und deren Vergleich mit Strukturen der Grundgesamtheit orientieren und dann die Stichprobenstrukturen denen der Grundgesamtheit mittels Gewichtungsfaktoren anpassen. Voraussetzung ist die Verfügbarkeit von Referenzstatistiken, die (per Konvention) als Abbildung der Grundgesamtheit gelten und die interessierenden Strukturmerkmale in vergleichbarer Weise abbilden. Es empfiehlt sich, solche Strukturgewichtungen so behutsam wie möglich vorzunehmen und dabei Merkmale (und deren Kombinationen) zu verwenden, die möglichst hoch mit den Zielmerkmalen der Untersuchung korreliert sind.⁵

⁴ Vgl. ausführlich zur Stichprobenanlage TNS Infratest Sozialforschung, 2006: BIBB/BAuA-Erwerbstätigtenbefragung 2005/2006 – Feldbericht.

⁵ Vgl. Heyde, Christian von der, 1999: Techniken und Möglichkeiten der Realisierung von Random-Stichproben, Abschnitt Gewichtung. S. 54ff. in: ADM, AG.MA (Hrsg): Stichproben-Verfahren in der Umfrageforschung. Opladen: Leske + Budrich.

3.2 Designgewichtung

3.2.1 Auswahlchance des Haushalts

An erster Stelle steht die Korrektur der Auswahlwahrscheinlichkeiten der Haushalte. Dazu dient die im Interview erhobene Zahl von Festnetzanschlüssen, über die der jeweilige Haushalt für Gespräche erreichbar ist. Alle Anschlussnummern werden nach dem Gabler-Häder-Verfahren und daher mit identischen Wahrscheinlichkeiten ausgewählt. Die Wahrscheinlichkeit, dass ein Haushalt in die Stichprobenauswahl gelangt, ist daher direkt proportional zur Zahl der Festnetznummern, über die er erreichbar ist.⁶ Ein Gewichtungsfaktor, der reziprok zu dieser Zahl ist, gleicht diese Unterschiede aus. Dieser Korrekturfaktor dient als Eingangsfaktor für die nachfolgenden Gewichtungsschritte.

Die Stichprobe der vorliegenden Untersuchung wird nach dem Prinzip der so genannten Nettosteuerung realisiert. Dabei wird dem Sample-Management-System (SMS) die regionale Zellenverteilung quantitativ vorgegeben, die die realisierte Stichprobe erreichen soll. Ergebnis dieser Steuerung ist eine Nettostichprobe, die nur marginal von der vorgesehenen Regionalverteilung der Privathaushalte abweicht, also praktisch keine oder nur sehr gering von 1 abweichende Gewichtungsfaktoren benötigt.

Diese marginalen Abweichungen resultieren aus der gegebenenfalls unterschiedlichen Zahl der Telefonanschlüsse pro Haushalt, die die mit der Steuerung erreichte Verteilung „verschmutzen“. Diese Abweichungen werden durch die Anpassung an zwei Ränder, die sich aus der Kombination verschiedener regionaler Merkmale ergeben, ausgeglichen (vgl. Tabelle 1, 1. Stufe).

3.2.2 Auswahlchance der Person

Pro Haushalt wird eine Person für die Befragung zufällig ausgewählt. Die Auswahlchance der Person ist umgekehrt proportional zur Zahl der Zielpersonen im Haushalt – bei dieser Untersuchung Personen ab 15 Jahren. Diese Zahl wird im Eingangsteil des Interviews ermittelt und dient jetzt dazu, durch faktorielle Gewichtung jeder Person die gleiche Auswahlwahrscheinlichkeit zu geben (vgl. Tabelle 1, 2. Stufe). Als Produkt aus diesem Faktor und dem Haushaltsgewicht errechnet sich ein Designgewicht, welches auch „Umwandlungsfaktor“ genannt wird, weil es die zunächst haushaltsproportionale Stichprobe in eine personenproportionale Stichprobe umwandelt. Dieses Gewicht bildet den Eingangsfaktor der folgenden Ausfallgewichtung (vgl. Tabelle 1, 3. und 4. Stufe).

⁶ In der entsprechenden Frage werden zudem explizit Telefonnummern, die nur für Fax oder PC benutzt werden, ausgeschlossen.

3.3 Korrektur des Ausfallgeschehens über Strukturgewichtung

Die Berechnung der Gewichtungsfaktoren zur Korrektur des Ausfallgeschehens erfolgt bei TNS Infratest in einem iterativen Randsummenverfahren.⁷ In diese Optimierungsbemühungen fließt auch die Begrenzung der Gewichtungsfaktoren ein, die verhindert, dass einzelnen Fällen in der Auswertung zu großes Gewicht beigegeben wird. Nach Festlegung der Variablenkombinationen wird das Modell als Gewichtung mit mehrdimensionalen Randverteilungen realisiert. D. h. es wird ein iterativer Gewichtungsprozess aufgesetzt, dessen Ergebnis Gewichtungsfaktoren sind, die die realisierte Stichprobe an alle vorgegebenen Sollverteilungen mit vordefinierter Genauigkeit und minimaler Varianz anpassen. Ein integriertes Verfahren ermöglicht es, durch das iterative Vorgehen aktuelle Verschiebungen der Strukturen von in einem vorangegangenen Iterationsschritt berücksichtigten Merkmalen in späteren Iterationsschritten, die sich auf diese Merkmale beziehen, wieder zu korrigieren. Dadurch wird die Anpassung an alle Ränder sukzessive solange verbessert, bis ein vorher definiertes Konvergenzkriterium erreicht ist.

Auswahlgesamtheit der BIBB/BAuA-Erhebung 2005/2006 waren im ersten Schritt alle im Haushalt lebenden Personen ab 15 Jahren. Aus dieser Auswahlgesamtheit wurde durch eine Zufallsauswahl nach dem Schwedenschlüssel (Kish-Methode) die zu befragende Zielperson bestimmt. Anschließend wurden dieser Person mehrere Leitfragen gestellt, anhand derer identifiziert wurde, ob sie erwerbstätig im Sinne der Untersuchung ist.⁸ Nur dann wurde sie befragt. War sie dagegen nicht erwerbstätig, verabschiedete sich der Interviewer mit dem Hinweis darauf, dass sie nicht zur Zielgruppe der Studie gehört. Somit kann die Stichprobe der Bevölkerung ab 15 Jahren in zwei Teilstichproben unterteilt werden: in eine Stichprobe von nach der Definition der Untersuchung erwerbstätigen Personen und in eine Stichprobe von nicht erwerbstätigen Personen.

Ausgangspunkt für die im Folgenden beschriebene Ausfallgewichtung ist die Gesamtstichprobe der Bevölkerung ab 15 Jahren. Ziel war es, im Rahmen der Ausfallgewichtung die Strukturen dieser Stichprobe in Bezug auf bestimmte Merkmale an die Strukturen in der Grundgesamtheit anzupassen und so eine bevölkerungsrepräsentative Gesamtstichprobe zu erhalten. Erst in einem weiteren Schritt wurden dann die Erwerbstätigen näher betrachtet und in einer Art „Feinjustierung“ wurden bestimmte Verteilungen für sie kontrolliert und angepasst. Zur Ermittlung des endgültigen Gewichtungsmodells wurden drei Varianten (Tabelle 1) mit unterschiedlicher Kombination der Sollverteilungen erprobt, um ein Optimum zwischen Faktorenvarianz einerseits und dem Anpassungserfolg andererseits zu erreichen. Die Referenzstrukturen bei der Entwicklung des Gewichtungsmodells lieferte der Mikrozensus 2004, für die endgültig gewählte Variante wurden zudem Gewichtungsfaktoren auf Basis des Mikrozensus 2005 entwickelt.

⁷ Vgl. z. B. Wauschkuhn, Udo, 1982: Anpassung von Stichproben und n-dimensionalen Tabellen an Randbedingungen. GMD-Bericht Nr. 138. München: Oldenbourg Verlag.

⁸ Ausführlicher vgl. hierzu TNS Infratest Sozialforschung, 2006: BIBB/BAuA-Erwerbstätigenbefragung 2005/2006 – Feldbericht. Vorteil dieses Vorgehens ist eine hohe Genauigkeit bei der Zielpersonenbestimmung und die leichte Handhabbarkeit des Screening-Instruments, denn nur die Zielperson muss die Fragen zur Identifikation des Erwerbsstatus, und zwar nur ihres eigenen, beantworten. Damit ist gewährleistet, dass auch Tätigkeiten am Rand des Arbeitsmarktes identifiziert werden.

Tabelle 1: Ränderkombinationen der Gewichtungsmodelle (Referenzstruktur: Mikrozensus 2004)

	Variante	1	2	3
1. Stufe: Designgewichtung 1 (Personen ab 15 Jahren)				
Auswahlwahrscheinlichkeit des Haushalts: Zahl der Festnetzanschlüsse	X	X	X	
Regierungsbezirk x BIK	X	X	X	
Bundesland x BIK	X	X	X	
2. Stufe: Designgewichtung 2 (Personen ab 15 Jahren)				
Auswahlwahrscheinlichkeit der Person: Zahl der Zielpersonen	X	X	X	
3. Stufe: Ausfallgewichtung 1 (Personen ab 15 Jahren)				
Familienstand x Geschlecht x Alter ^{a)}	X	X	X	
West/Ost x Bildung x Alter	X	X	X	
Bundesland x Geschlecht x Alter	X	X	X	
Bundesland x Geschlecht	X	X	X	
West/Ost x deutsch/nicht deutsch x Geschlecht	X	X	X	
West/Ost x Bildung x Alter x erwerbstätig/nicht erwerbstätig				X
Geschlecht x Alter x erwerbstätig/nicht erwerbstätig				X
Geschlecht x erwerbstätig/nicht erwerbstätig				X
4. Stufe: Ausfallgewichtung 2 (nur Erwerbstätige)				
West/Ost x Bildung x Alter		X		
Geschlecht x Alter			X	
Geschlecht				X
Bevölkerungsstichprobe				
Fallzahl insgesamt	62.253	- ^{b)}	62.253	
Effektivität insgesamt	66,5	-	66,4	

^{a)} Die Merkmalsausprägungen der Variablen sind Tabelle 3 zu entnehmen. Davon weicht nur die Gruppierung der Altersvariable in Kombination mit der Bildungsvariable ab: Hier wurde eine Dichotomisierung in „bis 50 Jahre“ und „ab 51 Jahren“ verwendet. Hintergrund ist, dass im Mikrozensus die Angabe zur Schulbildung freiwillig ist, wenn die befragte Person mindestens 51 Jahre alt ist. Dementsprechend ist für diese Befragten ein deutlich höherer Anteil fehlender Angaben festzustellen.

^{b)} Bei Variante 2 wurden in einem zusätzlichen Schritt nur Erwerbstätige berücksichtigt. Daher gibt es für die anderen Befragten bei dieser Variante keinen Gewichtungsfaktor.

Zwei Indikatoren zur Qualität der Gewichtungsmodelle sind in Tabelle 2 dargestellt. Ihre Bedeutung wird im Folgenden kurz vorgestellt.

- Die Effektivität wird berechnet als das Verhältnis von effektiver Stichprobe n' zur realisierten Stichprobe n , kann daher also bei gegebener Stichprobengröße n zur Berechnung der effektiven Stichprobengröße n' herangezogen werden. Sie nimmt im optimalen Fall den Wert 100 an. Je niedriger ihr Wert ist, umso schlechter ist die Stichprobe zu beurteilen, da man für Folgerungen auf die Grundgesamtheit trotz des tatsächlich realisierten Stichprobenumfangs n nur von der effektiven Stichprobengröße n' ausgehen kann. Dementsprechend größer sind die Konfidenzintervalle anzusetzen.
- Die Standardabweichung der Gewichtungsfaktoren ist ein Maß für ihre Variation. Sie sollte möglichst klein sein.

Tabelle 2: Qualitätsindikatoren der Gewichtungsmodelle für die Erwerbstägenstichprobe
(Referenzstruktur: Mikrozensus 2004)

	Variante	1	2	3
Erwerbstägenstichprobe				
Fallzahl ungewichtet		20.000	20.000	20.000
Fallzahl gewichtet nach Anpassung		20.000	20.000	20.000
<i>Effektivität</i>				
		64,0	67,1	67,4
<i>Standardabweichung</i>				
		0,751	0,701	0,695

Vergleicht man diese Indikatoren über die Modelle hinweg, so deuten sie alle auf die Überlegenheit von Gewichtungsmodell in der Variante 3 hin. Es weist die höchste Effektivität und die geringste Streuung der Gewichtungsfaktoren auf. Zudem ist die „Strukturinkongruenz“ (vgl. Tabelle 3) am niedrigsten. Die Strukturinkongruenz wird hier gemessen als die Summe der absoluten Abweichungen zwischen Soll- und Stichprobenstruktur nach entsprechender Gewichtung für die in Tabelle 3 dargestellten Merkmale. Anzumerken ist, dass die Abweichung noch geringer ausfallen würde, wenn man jeweils nur die gültigen Angaben heranziehen und darauf prozentuieren würde. Dies liegt an dem Vorgehen bei der Entwicklung der Gewichtungsfaktoren, welches darauf ausgelegt ist, die Verteilungen bezüglich der gültigen Angaben anzupassen. Dies sei am Beispiel der Bildung verdeutlicht: Bezogen auf die gültigen Angaben beträgt der Anteil der Personen mit maximal Hauptschulabschluss im Mikrozensus 35,2 %, in der nach Variante 3 gewichteten Stichprobe 35,1 %. Die aus der Tabelle ablesbare Abweichung von 1,5 Prozentpunkten verringert sich bei dieser Betrachtung also auf 0,1 Prozentpunkte.

Tabelle 3: Vergleich zwischen Stichproben- und Referenzstruktur (Mikrozensus 2004)

	MZ 2004	Variante 1	Variante 2	Variante 3
Geschlecht				
Männlich	56,4	59,0	56,4	56,4
Weiblich	43,6	41,0	43,6	43,6
<i>Summe</i>	100	100	100	100
Familienstand				
Ledig	28,7	28,4	29,0	28,9
Verheiratet	61,9	61,8	61,2	61,3
Verwitwet	1,5	1,5	1,5	1,5
Geschieden	7,9	8,1	8,1	8,2
Angabe fehlt		0,1	0,1	0,1
<i>Summe</i>	100	100	100	100
Deutsch - Nicht-Deutsch				
Nicht-deutsch	8,2	8,2	8,1	8,2
Deutsch	91,8	91,8	91,9	91,8
<i>Summe</i>	100	100	100	100
Alter				
15 bis 20	1,4	1,0	1,4	1,4
21 bis 30	16,3	15,4	16,3	16,3
31 bis 40	29,0	29,9	28,9	28,9
41 bis 50	29,9	30,6	29,8	29,8
51 bis 64	22,4	21,7	22,4	22,4
65 und älter	1,0	1,1	1,0	1,0
Angabe fehlt		0,2	0,2	0,2
<i>Summe</i>	100	100	100	100
Stellung im Beruf				
Arbeiter	29,9	31,6	30,5	30,5
Angestellte	50,9	50,6	51,3	51,3
Beamte	6,7	6,9	7,2	7,2
Selbst./Freiber./freier Mit.	11,5	9,5	9,4	9,5
Mith. Familienangeh.	1,1	1,2	1,2	1,2
Arbeiter/Angestellter		0,1	0,1	0,1
Angabe fehlt		0,1	0,1	0,1
<i>Summe</i>	100	100	100	100

Fortsetzung Tabelle 3

	MZ 2004	Variante 1	Variante 2	Variante 3
Höchster Schulabschluss				
Maximal Volks-/Hauptschule	33,4	35,6	34,8	34,9
Mittlere Reife/POS	33,1	35,5	34,9	34,9
Fachhochschulreife, Abitur	28,2	28,4	29,7	29,7
Angabe fehlt	5,1	0,5	0,5	0,5
System Missing	0,1			
<i>Summe</i>	<i>100</i>	<i>100</i>	<i>100</i>	<i>100</i>
Bundesland, Wohnort				
Schleswig-Holstein	3,4	3,2	3,3	3,3
Hamburg	2,2	2,0	2,0	2,0
Niedersachsen	9,4	9,2	9,2	9,2
Bremen	0,7	0,8	0,8	0,8
Nordrhein-Westfalen	20,8	21,0	21,0	21,1
Hessen	7,7	7,4	7,4	7,4
Rheinland-Pfalz	4,9	5,0	5,0	5,0
Baden-Württemberg	13,7	13,7	13,8	13,7
Bayern	16,3	16,2	16,3	16,2
Saarland	1,2	1,2	1,2	1,2
Berlin	4,0	3,8	3,8	3,8
Brandenburg	3,1	3,2	3,1	3,1
Mecklenburg-Vorpommern	1,9	2,2	2,2	2,1
Sachsen	5,0	5,2	5,0	5,0
Sachsen-Anhalt	2,9	3,1	3,1	3,1
Thüringen	2,9	2,9	2,9	2,9
<i>Summe</i>	<i>100</i>	<i>100</i>	<i>100</i>	<i>100</i>

Fortsetzung Tabelle 3

	MZ 2004	Variante 1	Variante 2	Variante 3
Bundesland, Betriebsort				
Schleswig-Holstein	3,0	3,0	3,1	3,1
Hamburg	2,8	2,7	2,7	2,7
Niedersachsen	8,7	8,7	8,7	8,7
Bremen	1,0	1,0	1,0	1,0
Nordrhein-Westfalen	20,5	21,1	21,1	21,2
Hessen	7,8	7,6	7,6	7,6
Rheinland-Pfalz	4,4	4,6	4,6	4,6
Baden-Württemberg	13,6	13,8	13,8	13,8
Bayern	15,9	16,3	16,4	16,4
Saarland	1,2	1,2	1,2	1,2
Berlin	4,2	4,1	4,0	4,1
Brandenburg	2,6	2,7	2,7	2,7
Mecklenburg-Vorpommern	1,7	2,0	2,0	2,0
Sachsen	4,8	5,2	5,0	5,0
Sachsen-Anhalt	2,6	2,8	2,7	2,7
Thüringen	2,6	2,6	2,5	2,5
Ausland	0,5	0,6	0,6	0,6
Angabe fehlt	2,1	0,1	0,1	0,1
<i>Summe</i>	<i>100</i>	<i>100</i>	<i>100</i>	<i>100</i>
<i>Strukturinkongruenz</i>		24,1	15,2	14,9
<i>Durchschnittliche Strukturinkongruenz^{a)}</i>		0,44	0,28	0,27

^{a)} Bezogen auf alle inhaltlichen Merkmalsausprägungen.

In Tabelle 3 werden auch zwei Merkmale ausgewiesen, die bei der Entwicklung der Gewichtungsmodelle nicht berücksichtigt wurden, und zwar die Stellung im Beruf und der Ort, an dem der Betrieb liegt. Für beide Merkmale zeigen sich nur geringe Abweichungen von der Referenzstruktur. Dies verweist auf die Repräsentativität der Stichprobe für die Grundgesamtheit und damit auf die Verallgemeinerbarkeit der Analyseergebnisse.

Wie bereits angesprochen, wurden für die in den bisherigen Tabellen dokumentierte Entwicklung der Gewichtungsfaktoren die Referenzstrukturen aus dem Mikrozensus 2004 gezählt, da zunächst die Daten für 2005 noch nicht zur Verfügung standen. Neben dem auf dem Mikrozensus 2004 basierenden Gewichtungsfaktor der Variante 3 wurde der Stichprobe auch der Gewichtungsfaktor nach Variante 3 zugespielt, der auf Referenzstrukturen aus Zählungen des Mikrozensus 2005 beruht. Die diesbezüglichen Strukturvergleiche und Qualitätsindikatoren zeigt Tabelle 4. Auch hier ist eine sehr gute Anpassung der Stichprobenstrukturen an die Referenzstrukturen festzustellen.

Tabelle 4: Vergleich zwischen Stichprobenstruktur und Referenzstruktur im Mikrozensus 2005 sowie Qualitätsindikatoren für das Gewichtungsmodell der Variante 3

	MZ 2005	Variante 3
Geschlecht		
Männlich	56,3	56,3
Weiblich	43,7	43,7
<i>Summe</i>	100	100
Familienstand		
Ledig	30,0	30,1
Verheiratet	60,4	59,9
Verwitwet	1,5	1,5
Geschieden	8,1	8,5
Angabe fehlt		0,1
<i>Summe</i>	100	100
Deutsch - Nicht-Deutsch		
Nicht-deutsch	8,4	8,8
Deutsch	91,6	91,2
<i>Summe</i>	100	100
Alter		
15 bis 20	1,6	1,6
21 bis 30	16,8	16,8
31 bis 40	27,8	27,7
41 bis 50	30,4	30,3
51 bis 64	22,3	22,2
65 und älter	1,2	1,2
Angabe fehlt		0,2
<i>Summe</i>	100	100
Stellung im Beruf		
Arbeiter	28,0	29,8
Angestellte	52,6	51,5
Beamte	6,5	7,3
Selbst./Freiber./freier Mit.	11,8	9,7
Mith. Familienangeh.	1,1	1,3
Arbeiter/Angestellter		0,2
Angabe fehlt		0,2
<i>Summe</i>	100	100

Fortsetzung Tabelle 4

	MZ 2005	Variante 3
Höchster Schulabschluss		
Maximal Volks-/Hauptschule	33,9	33,5 ^{a)}
Mittlere Reife/POS	34,9	35,4
Fachhochschulreife, Abitur	30,4	30,8
Angabe fehlt	0,6	0,4
System Missing	0,2	
<i>Summe</i>	<i>100</i>	<i>100</i>
Bundesland, Wohnort		
Schleswig-Holstein	3,4	3,3
Hamburg	2,2	2,0
Niedersachsen	9,3	9,1
Bremen	0,7	0,8
Nordrhein-Westfalen	20,9	21,0
Hessen	7,5	7,5
Rheinland-Pfalz	5,0	5,0
Baden-Württemberg	13,7	13,8
Bayern	16,2	16,1
Saarland	1,2	1,2
Berlin	4,0	3,9
Brandenburg	3,1	3,1
Mecklenburg-Vorpommern	2,0	2,1
Sachsen	5,0	5,2
Sachsen-Anhalt	2,8	3,0
Thüringen	2,8	2,9
<i>Summe</i>	<i>100</i>	<i>100</i>

^{a)} Als Sollwert für den Anteil in der niedrigsten Ausprägung der Bildung wurde hier ein etwas niedrigerer Wert eingesetzt, um zu erreichen, dass der Anteil der Arbeiter ebenfalls etwas niedriger ausfällt.

Fortsetzung Tabelle 4

	MZ 2005	Variante 3
Bundesland, Betriebsort		
Schleswig-Holstein	3,0	3,1
Hamburg	2,7	2,7
Niedersachsen	8,9	8,7
Bremen	1,1	1,0
Nordrhein-Westfalen	20,6	21,1
Hessen	7,5	7,6
Rheinland-Pfalz	4,8	4,6
Baden-Württemberg	13,4	13,9
Bayern	16,6	16,3
Saarland	1,3	1,2
Berlin	4,2	4,1
Brandenburg	2,7	2,7
Mecklenburg-Vorpommern	2,4	1,9
Sachsen	4,7	5,1
Sachsen-Anhalt	2,6	2,6
Thüringen	2,6	2,5
Ausland	0,6	0,5
Angabe fehlt	0,3	0,1
<i>Summe</i>	<i>100</i>	<i>100</i>
<i>Standardabweichung</i>		<i>0,747</i>
<i>Strukturinkongruenz</i>		<i>14,7</i>
<i>Durchschnittliche Strukturinkongruenz^{a)}</i>		<i>0,27</i>

^{a)} Bezogen auf alle inhaltlichen Merkmalsausprägungen.

4. Steuerung der Stichprobe

Werden bei einer Stichprobe Abbildungsprobleme in Bezug auf die interessierende Grundgesamtheit festgestellt, ist es üblich, Abweichungen der Verteilungen in der Stichprobe von Referenzverteilungen durch eine faktorielle Gewichtung zu korrigieren: Die in der Stichprobe realisierten Strukturen werden so an die Referenzstrukturen angepasst. Dabei werden den Einheiten in der Stichprobe über die Gewichtungsfaktoren Bedeutungsgewichte zugeordnet, wodurch ihre Bedeutung im Hinblick auf die Analyseergebnisse verändert und somit quasi eine Schichtung à posteriori durchgeführt wird.

Um eine Veränderung des Bedeutungsgewichts zu erreichen, kann allerdings auch über eine Schichtung à priori Einfluss auf die Zusammensetzung der Stichprobe genommen werden. Wenn man die entsprechenden Strukturen in der Stichprobe zeitnah beobachtet, besteht die Möglichkeit, entsprechende Abweichungen über eine auf der Schichtung aufsetzende Steuerung zu verringern. Ein solches Vorgehen reduziert den Anpassungsbedarf im Rahmen einer nachträglichen Gewichtung.

So war bei der vorliegenden Untersuchung zu erwarten, dass Zielpersonen aus unteren Bildungsschichten eher unterproportional teilnehmen. Ein niedriges Bildungsniveau geht meist einher mit einer geringeren kommunikativen Kompetenz. Dies ist möglicherweise der erklärende Mechanismus, der hinter der geringeren Neigung steht, an telefonischen Befragungen teilzunehmen. Bei der Befragung zeigte sich dann erwartungsgemäß, dass über eine telefonische Ansprache der Zugang zu den unteren Bildungsschichten schwieriger ist als über eine mündlich-persönliche. Da von einem Zusammenhang zwischen Bildungsniveau und Stellung im Beruf auszugehen ist, war zu vermuten, dass die Schwierigkeit, untere Bildungsschichten zu erreichen, einhergeht mit einer Unterrepräsentation von Arbeitern in der Stichprobe. Auch diese Vermutung wurde bestätigt.

Daher wurde entschieden, bereits im Vorhinein mit gezielten Maßnahmen gegenzusteuern. Ziel war es, über ein Screening-Verfahren die Wahrscheinlichkeit zu erhöhen, dass Arbeiter in die Stichprobe gelangen. Für die Wahl des Merkmals „Stellung im Beruf“ und der Merkmalsausprägung „Arbeiter“ sprachen eine Reihe von Gründen

- Zunächst zeigen bereits die Erfahrungen mit der BIBB/IAB-Erhebung 1998/99, dass die Arbeiter unterrepräsentiert sind.⁹

⁹ In der von Infratest erhobenen ungewichteten Teilstichprobe lag die Abweichung bei 3 Prozentpunkten, in der ungewichteten Teilstichprobe von infas betrug sie sogar 7,6 Prozentpunkte (vgl. Josef Hartmann, Harald Bielenski, Bernhard von Rosenbladt, 1999: Erwerb und Verwertung beruflicher Qualifikationen von Erwerbstätigen. BIBB/IAB-Erhebung 1998/99. Gewichtung und Strukturkontrolle der Stichprobe. Bonn: Bundesinstitut für Berufsbildung). Es ist allerdings darauf hinzuweisen, dass es sich damals um eine CAPI-Stichprobe handelte. Unsere Erfahrungen in jüngerer Zeit mit telefonischen Erhebungen ließen erwarten, dass die Abweichungen in einer CATI-Stichprobe eher noch größer ausfallen.

- Die Stellung im Beruf ist inhaltlich relevant, weil zu vermuten ist, dass die anderen im Rahmen der geplanten Untersuchung interessierenden Merkmale damit zusammenhängen.
- Die berufliche Stellung ist vergleichsweise einfach zu erheben und der Zielperson ist leicht vermittelbar, warum dieses Merkmal im Zusammenhang mit ihrer Auswahl bedeutsam ist.
- Da zudem, wie oben bereits angesprochen, ein Zusammenhang zwischen der beruflichen Stellung und der Bildung zu erwarten war – was sich auch tatsächlich bestätigte –, sollte die entsprechende Steuerung zu einer Reduzierung des Bildungsbias in der Stichprobe führen.

Zunächst wurde die Erhebung von Oktober 2005 bis Anfang Januar 2006 ohne Steuerung durchgeführt. Regelmäßige Strukturkontrollen zeigten allerdings die erwarteten Abweichungen der Verteilung der Stellung im Beruf. Daher wurde ab dem 12. Januar 2006 steuernd eingegriffen. Dieses zeitlich versetzte Vorgehen hatte den Vorteil, dass mehr Sicherheit in Bezug auf den Umfang der erforderlichen Steuerung bestand. Durch die Steuerung der Stichprobe nach der Stellung im Beruf konnte die entsprechende Verteilung angenähert werden. Tabelle 5 zeigt die mit dem Umwandlungsfaktor designgewichtetete Verteilung vor und nach der Stichprobensteuerung sowie insgesamt.¹⁰ Es ist erkennbar, dass die Maßnahme den Anteil der Arbeiter in der Gesamtstichprobe um ca. sechs Prozentpunkte erhöht hat. Dies entsprach in etwa einer vorher vorgenommenen Abschätzung, eine weitere Anpassung sollte nicht stattfinden.

Tabelle 5: Stellung im Beruf vor und nach der Steuerung (designgewichtet mit Umwandlungsfaktor)

Berufliche Stellung	Vor Steuerung	Nach Steuerung	Insgesamt
Fallzahl	13.190	6.810	20.000
Arbeiter	19,7	37,6	25,8
Nicht-Arbeiter	80,3	62,4	74,2
<i>Summe</i>	<i>100</i>	<i>100</i>	<i>100</i>

¹⁰ Im Rahmen der Designgewichtung (vgl. Abschnitt 3.2) wird die Auswahlwahrscheinlichkeit des Haushalts und der Person innerhalb des Haushalts berücksichtigt. Sie ist u. a. erforderlich, um von der Haushaltsebene auf die Personenebene zu gelangen.

Die Steuerung nach der Stellung im Beruf hatte auch Einfluss auf den Anteil der niedrig Gebildeten und der Frauen in der Stichprobe. Die entsprechenden Auswirkungen sind in Tabelle 6 dargestellt. So wurde der Anteil der niedrig Gebildeten erhöht und der der Frauen reduziert. Dies war gleichbedeutend mit einer Annäherung der Verteilung in der Stichprobe an die Struktur in der Grundgesamtheit.

Tabelle 6: Bildung und Geschlecht vor und nach der Steuerung (designgewichtet)

	Vor Steuerung	Nach Steuerung	Insgesamt
Fallzahl	13.190	6.810	20.000
Bildung			
Bis Hauptschule (inkl. fehlende Angabe)	24,2	30,6	26,4
Realschule	37,5	38,1	37,7
FHR/Abitur	38,3	31,3	35,9
<i>Summe</i>	100	100	100
Geschlecht			
Männlich	49,0	51,8	50,0
Weiblich	51,0	48,2	50,0
<i>Summe</i>	100	100	100

Abschließend ist noch auf Wirkung der Steuerung der Stichprobe auf die Effektivität einzugehen. Ausgangspunkt ist folgende Überlegung: Durch die Steuerung wurde in der Stichprobe bei Gewichtung nach Variante 1 (Bevölkerungsgewichtung, vgl. Tabelle 3) ein Anteil von 31,6 % Arbeitern erreicht. Es stellt sich nun die Frage, welche Auswirkungen es auf die Effektivität der Stichprobe hat, wenn man diesen Anteil ohne Steuerung über ein entsprechendes Gewichtungsmodell erreichen will.

Um dies abschätzen zu können, wurde für alle im Rahmen der Untersuchung interessierenden Erwerbstätigen (s. Kapitel 2), d. h. konkret die 20.000, die befragt wurden, plus die 8.714 Erwerbstätigen, die nicht Arbeiter sind und daher nicht befragt wurden, ein Gewichtungsmodell entwickelt. Es entspricht Variante 1 und passt zudem noch den Arbeiteranteil auf 31,6 % an. Für das so entwickelte Modell ergibt sich eine Effektivität von 60. Sie liegt damit 4 Punkte

unter dem nach Steuerung mit Variante 1 erzielten Wert. Der Abstand wäre noch größer, wenn statt der 28.714 Personen nur 20.000 Personen in das Modell eingegangen wären: Sie wäre dann sogar 6 Punkte niedriger. Schließlich ist als weiterer Vorteil der Steuerung zu werten, dass dadurch die ungewichtete Fallzahl der Arbeiter in der Stichprobe erhöht wurde, was die Möglichkeiten zu differenzierten Auswertungen für diese Gruppe verbessert.