

## **Verknüpfung von Befragungs- und Prozessdaten**

**Selektivität durch fehlende Zustimmung  
der Befragten?**

*Josef Hartmann, Gerhard Krug*

# **Verknüpfung von Befragungs- und Prozessdaten**

## **Selektivität durch fehlende Zustimmung der Befragten?**

*Josef Hartmann (TNS Infratest Sozialforschung), Gerhard Krug (IAB)*

Auch mit seiner neuen Reihe „IAB-Discussion Paper“ will das Forschungsinstitut der Bundesagentur für Arbeit den Dialog mit der externen Wissenschaft intensivieren. Durch die rasche Verbreitung von Forschungsergebnissen über das Internet soll noch vor Drucklegung Kritik angeregt und Qualität gesichert werden.

Also with its new series "IAB Discussion Paper" the research institute of the German Federal Employment Agency wants to intensify dialogue with external science. By the rapid spreading of research results via Internet still before printing criticism shall be stimulated and quality shall be ensured.

## Inhaltsverzeichnis

Abstract .....	4
1 Problemstellung .....	5
2 Optimierung der Datenlage durch die Verknüpfung.....	7
3 Theoretische Überlegungen zum Zustimmungsverhalten .....	12
4 Empirische Ergebnisse.....	18
4.1 Daten und Indikatoren .....	18
4.1.1 Repräsentative Erhebung zur Evaluation des Mainzer Modells ....	18
4.1.2 Indikatoren der Einflussfaktoren der Zustimmungsbereitschaft ..	19
4.2 Erklärung der Zustimmungsbereitschaft.....	21
4.2.1 Unterschiede zwischen Zustimmenden und Nicht-Zustimmenden .....	21
4.2.2 Mehrebenenmodell zur Erklärung der Zustimmungsbereitschaft .	21
4.3 Auswirkung der Selektivität auf Forschungsergebnisse – Zwei Beispiele .....	27
5 Zusammenfassung und Diskussion.....	32
Literatur.....	34
Anhang.....	37

## Abstract

Angesichts sinkender Ausschöpfungsquoten und knapper finanzieller Mittel zur Durchführung von Erhebungen gewinnt die Verknüpfung von Prozess- und Befragungsdaten an Attraktivität. Neben erhebungstechnischen Vorteilen kann die Verknüpfung Schwächen der einzelnen Datenarten ausgleichen und so die Aussagekraft und Validität von Forschungsergebnissen erhöhen. Die Möglichkeit der Zusammenspielung von Prozess- und Befragungsdaten auf individueller Ebene ist allerdings aus datenschutzrechtlichen Gründen von der Zustimmung der betroffenen Personen abhängig. Daher ist es nicht auszuschließen, dass es bei Analysen auf Basis ausdrücklicher Zustimmung der Betroffenen zur selektiven Zusammensetzung der Stichprobe kommt. In Auseinandersetzung mit Theorien des Befragtenverhaltens werden in dieser Arbeit Bestimmungsfaktoren der Zustimmung zur Datenzusammenspielung untersucht und in einer Mehrebenenanalyse anhand von Befragungsdaten auf ihren empirischen Einfluss getestet. Es zeigt sich, dass die Hypothesen zu bestimmten Einflussfaktoren verworfen werden müssen, zu anderen dagegen nicht: Dennoch scheinen alle vermuteten Mechanismen – ein hoher mit der Befragung in Verbindung gebrachter Nutzen, keine erwarteten negativen Konsequenzen, Vertrauen in die Anonymität der Auswertung und der Wunsch nach Schutz der eigenen Privatsphäre – von Bedeutung zu sein.

Zusätzlich wird an zwei beispielhaften Forschungsfragen untersucht, inwieweit sich die Ergebnisse unter Einschluss und unter Ausschluss nicht zustimmender Personen unterscheiden. Für die durchgeführten Analysen zeigen sich keine Unterschiede. Anzumerken ist allerdings, dass diese Ergebnisse nur exemplarischen Charakter haben und nicht ohne weitreichende Annahmen auf andere Fragestellungen übertragen werden können.

**Keywords:** Befragungs- und Prozessdaten

**JEL classification:** C, J, Z

Wir danken Johannes Ludsteck, Roman Lutz, Susanne Rässler, Helmut Rudolph, Mark Trappmann und Katja Wolf für anregende Diskussionen und hilfreiche Kommentare.

## 1 Problemstellung

Angesichts sinkender Ausschöpfungsquoten und knapper finanzieller Mittel zur Durchführung von Erhebungen gewinnt die Verknüpfung von Daten aus verschiedenen Erhebungen an Attraktivität. Zusätzlich zu erhebungstechnischen Vorteilen kann die Verknüpfung Schwächen der einzelnen Datenarten ausgleichen und so die Aussagekraft und Validität von Forschungsergebnissen erhöhen.

Bei der Datenverknüpfung (record linkage) kann zwischen der Ergänzung oder Datenfusion einerseits und exakter Verknüpfung andererseits unterschieden werden. Ergänzung in der rudimentärsten Form findet statt, wenn zwei Datenquellen parallel ausgewertet werden, etwa die Verwaltungsdaten zu sozialversicherungspflichtiger Beschäftigung und das Sozio-oekonomische Panel (SOEP) zur gleichen Fragestellung. Darüber hinaus können Mikrodaten aber auch auf der individuellen Ebene ergänzt werden, indem zu den vorhandenen individuenbezogenen Merkmalen die Angaben von möglichst ähnlichen Individuen hinzugespielt werden. Existieren zu denselben Individuen zwei Datenquellen mit verschiedenen Merkmalen, ist darüber hinausgehend ihre exakte Verknüpfung möglich. Basierend auf geeigneten Identifikatoren wie dem Namen, der Adresse oder der Sozialversicherungsnummer können dabei die Informationen aus den beiden Datenquellen auf der Ebene der Individuen verknüpft werden (z. B. Jenkins et al. 2005). Für die Arbeitsmarktforschung ist dabei neben der Verknüpfung verschiedener prozessproduzierter Daten der Bundesagentur für Arbeit (BA) vor allem die Verknüpfung von Prozess- oder Registerdaten der BA mit Befragungsdaten relevant. Das hieße beispielsweise, zu erhobenen Befragungsdaten die in der Bundesagentur für Arbeit angefallenen Prozessdaten hinzuzuspielen oder umgekehrt die Registerdaten der BA mit Informationen aus Befragungen anzureichern.

Die Wichtigkeit, die der exakten Verknüpfung von Datenquellen in der wissenschaftlichen Diskussion zugeschrieben wird, spiegelt sich auch in den Empfehlungen der vom Bundesministerium für Bildung und Forschung eingesetzten Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik wider: „Die Kommission empfiehlt, die Möglichkeit exakter Datenverknüpfungen – ohne ausdrückliche Zustimmung aller einzelnen Befragten – für rein statistische Zwecke gesetz-

lich zu regeln. Eine exakte Verknüpfung von Erhebungsdaten mit Registerdaten sollte grundsätzlich ebenso beurteilt und geregelt werden“ (Kommission 2001: 24).

Gegenwärtig ist die Zusammenspielung von Prozess- und Befragungsdaten auf individueller Ebene allerdings aus datenschutzrechtlichen Gründen noch von der Zustimmung der betroffenen Personen abhängig. Dabei kann nicht ausgeschlossen werden, dass die Bereitschaft dazu systematisch mit untersuchungsrelevanten Eigenschaften der Personen variiert, wodurch es zu einer Art Verknüpfungs-Bias der Untersuchung kommen kann. Verweigert eine Person ihre Zustimmung, so wirkt sich dies bei multivariaten Analysen, die das gesamte Spektrum der Prozess- und Befragungsdaten nutzen wollen, als Totalausfall dieser Person für die Analyse aus (Unit Nonresponse). Es können also nur Personen einbezogen werden, die keine Datenschutzbedenken haben. Aus der Umfrageforschung ist bekannt, dass die Teilnahmebereitschaft an Befragungen von Merkmalen abhängen kann, die für die Untersuchungsfragen relevant sind, wodurch sich verzerrte Schätzungen (bias) im Hinblick auf die Ergebnisse ergeben können. In ähnlicher Weise ist zu erwarten, dass auch die Bereitschaft, einer Zuspielung von Prozessdaten zuzustimmen, unter den Befragten nicht zufällig verteilt ist.

Daher ist zu fragen, ob eine solche Beschränkung der Analysegesamtheit nicht zu Problemen führt, welche die Vorteile des record linkage aufheben könnten. Vor diesem Hintergrund soll in dieser Arbeit analysiert werden, in welchem Ausmaß und mit welchen Konsequenzen es bei der exakten Verknüpfung auf Basis ausdrücklicher Zustimmung der Betroffenen zur selektiven Zusammensetzung der Stichprobe kommt.

Als Datenquelle dient eine Befragung, die TNS Infratest Sozialforschung im Auftrag des Instituts für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit zur Evaluation der Kombilohnfördermaßnahme „Mainzer Modell“ durchgeführt hat (vgl. Gewiese et al. 2004). Dabei wurden Teilnehmer und Personen einer Kontrollgruppe u. a. gefragt, ob sie der Verknüpfung der Befragungsdaten mit den Prozessdaten der BA zustimmen.

In Abschnitt 2 wird zunächst auf die Vorteile eingegangen, die sich durch die Verknüpfung der Informationen aus den beiden Datenquellen ergeben. Um die Ursache eventueller Zustimmungsverweigerung zu analysieren, werden anschließend theoriegeleitet die Bedingungsfaktoren der Zustimmung identifiziert und Hypothesen über die Richtung ihres Einflusses abgeleitet (Abschnitt 3). Nach einer kurzen Beschreibung des Datensatzes (4.1) und einem deskriptiven Vergleich von Zustimmenden und Nicht-Zustimmenden (4.2) wird aufbauend auf der Auseinandersetzung mit der Theorie ein Modell der Zustimmungswahrscheinlichkeit entwickelt, das den Einfluss der theoretisch identifizierten Faktoren überprüft (4.3). Anschließend wird untersucht, welche Auswirkung die selektive Zustimmungsbereitschaft auf konkrete Forschungsergebnisse hat. Anhand von zwei beispielhaften Forschungsfragen wird analysiert, inwieweit sich die Ergebnisse unter Einschluss und unter Ausschluss nicht zustimmender Personen unterscheiden. Im Schlussabschnitt werden die Ergebnisse zusammengefasst und hinsichtlich ihrer Konsequenzen für die Forschungsarbeit bewertet.

## **2 Optimierung der Datenlage durch die Verknüpfung**

Die Arbeitsmarktforschung ist in der privilegierten Situation, zur Beantwortung ihrer Fragestellungen sowohl Befragungsdaten als auch so genannte prozessproduzierte Daten heranziehen zu können. Prozessproduzierte Daten werden für Verwaltungszwecke erhoben und stehen Wissenschaftlern unter bestimmten Bedingungen aufbereitet zur Verfügung. Beispielsweise finden Prozess- bzw. Registerdaten bei der Evaluation arbeitsmarktpolitischer Maßnahmen Verwendung (z. B. Kaltenborn et al. 2005), aber auch für andere Fragestellungen, wie z. B. zur Mobilität (Winzio 2004), zu geschlechtsspezifischer Entlohnung (Achatz et al. 2005) oder zum Niedriglohnsektor (Eichhorst et al. 2005). Beispiele für solche Datensätze sind die IAB Beschäftigtenstichprobe (IABS-R01, Hamann et al. 2004), die Stichprobe der Integrierten Erwerbsbiografien (IEBS, Hummel et al. 2005) oder Daten, die von der Bundesagentur für Arbeit für die Evaluation arbeitsmarktpolitischer Maßnahmen zur Verfügung gestellt werden.

Diese Daten fallen automatisch im Rahmen der Geschäftstätigkeit an und ihre Erhebung verursacht keine eigenen Kosten.<sup>1</sup> Sie besitzen Vorteile gegenüber Befragungsdaten, sind allerdings auch mit Beschränkungen behaftet, die für Befragungsdaten nicht in dem Maße gelten (Hakim 1983; Schmähl, Fachinger 1994; Wirth, Müller 2004).

Da die Daten im Rahmen eines Verwaltungsprozesses anfallen, sind sie in Bezug auf die Individuen, die mit diesen Daten beschrieben werden, nicht-reaktiv. Probleme, wie beispielsweise die Teilnahmeverweigerung, sozial erwünschtes Antwortverhalten, Antwortverweigerung oder falsche Angaben aufgrund eingeschränkten Erinnerungsvermögens der Befragten treten im Allgemeinen nicht auf.<sup>2</sup> Daher ist davon auszugehen, dass diese Daten im Hinblick auf bestimmte Sachverhalte eine höhere Validität als Befragungsdaten aufweisen. Während beispielsweise bei einem Interview die Erwerbsbiografie retrospektiv erfragt werden muss, können bei den Prozessdaten alle Informationen aus den Geschäftsdaten zur sozialversicherungspflichtigen Beschäftigung verknüpft werden. Auch Effekte sozialer Erwünschtheit spielen hier keine Rolle, z. B. bei Fragen zum aktuellen Bezug von Sozialleistungen oder zum Vorliegen von Sperrzeiten. Prozessdaten sind oft für einen langen Zeitraum verfügbar und Phänomene wie das in Panelstudien beobachtbare Ausfallgeschehen im Zeitverlauf („Panel-Mortalität“), welches zudem oft selektiv ist, also mit inhaltlich interessierenden Merkmalen, z. B. mit bestimmten Entwicklungen im Erwerbsverlauf, zusammenhängt, treten nicht auf.<sup>3</sup>

Ein weiterer Vorteil von Prozessdaten liegt darin, dass sie nicht selten Vollerhebungen darstellen bzw., wenn nötig, in Form von Zufallsstichpro-

---

<sup>1</sup> Das heißt natürlich nicht, dass mit der Aufbereitung solcher prozessproduzierter Daten für die Wissenschaft nicht auch Zeit und Kosten verbunden sind.

<sup>2</sup> Das heißt nicht, dass die Prozessdaten fehlerfrei und vollständig sind. Hier ist daran zu denken, dass die Erfassung der Informationen nur im Wege der Erfüllung anderer Aufgaben erfolgt. Die Sachbearbeiter sind keine „Erhebungsexperten“ und haben darüber hinaus durchaus (Ermessens-)Spielräume, was sie erfassen, wie sie es erfassen und teilweise sogar, ob sie überhaupt eine Information erfassen.

<sup>3</sup> Aber auch die Prozessdaten enthalten Informationen zu bestimmten Phasen nicht, beispielsweise wenn es sich um andere Rechtskreise handelt. Ein früheres Beispiel war der Bezug von laufender Hilfe zum Lebensunterhalt, der nicht in den Daten der BA auftauchte. Es ist anzunehmen, dass hier ebenfalls eine selektive Nichterfassung gegeben ist: Für einen bestimmten Zeitpunkt oder Zeitraum können Personen mit bestimmten Merkmalen nicht in den Analysen berücksichtigt werden.



ben ohne Ausfälle aus einer Vollerhebung gezogen werden können. Der Stichprobenumfang kann dann entsprechend groß gewählt werden, so dass auch Detailanalysen mit kleineren Bevölkerungsgruppen durchgeführt werden können, die aufgrund ihrer geringen Inzidenz in bevölkerungsrepräsentativen Befragungen nur in geringer Fallzahl vorhanden sind.

Aber auch Befragungsdaten haben bestimmte Vorteile. Zunächst müssen sich Befragungen nicht auf bestimmte Zielgruppen, beispielsweise Personen, die mit den entsprechenden Behörden, z. B. der Arbeitsverwaltung, in Kontakt kommen, beschränken, sondern sie können prinzipiell alle Zielgruppen einschließen, z. B. auch Nicht-Erwerbspersonen. So liegen bei der BA beispielsweise lückenlose Informationen über die sozialversicherungspflichtigen Beschäftigungszeiten einer Person vor, nicht aber über selbstständige Tätigkeiten, über beamtete Erwerbstätigkeiten und – bis zur Einführung der allgemeinen Meldepflicht 1999 – auch nicht über geringfügige Beschäftigungsverhältnisse.

Weiterhin können in Befragungen zielgerichtet die im Rahmen der Forschungsfrage interessierenden Inhalte erhoben werden. So hat im Falle prozessproduzierter Daten die Forschung keinen oder geringen Einfluss auf die Erhebung der Daten, weder was den Inhalt im Allgemeinen noch was die Messung der einzelnen Merkmale im Besonderen betrifft. Beides folgt keiner wissenschaftlichen Fragestellung bzw. Operationalisierung, sondern einer administrativen Logik.<sup>4</sup> Erfasst werden über bestimmte Zielpersonen genau die Informationen, welche die entsprechende Behörde zur Erfüllung ihres gesetzlichen Auftrages benötigt. Gerade inhaltlich interessante Abweichungen von der Normalität können damit also mit Prozessdaten nicht untersucht werden.<sup>5</sup> Und auch der Kranz der erhobenen und damit verfügbaren Informationen orientiert sich streng an den administrativen Notwendigkeiten, weshalb Wirth und Müller (2004: 101) Registerda-

---

<sup>4</sup> Wenn auch in letzter Zeit zu beobachten ist, dass der Einfluss der Wissenschaft auf die Erhebung forschungsrelevanter Prozessdaten zunimmt. Für diesen Hinweis danken wir Thomas Kruppe.

<sup>5</sup> Ganz zu schweigen von Phänomenen wie freiwilligem Engagement oder Tätigkeiten im Bereich der Schattenwirtschaft.

ten als zum Teil von „buchhalterische(r) Zahlenmentalität und paragrafenorientierte(r) Dürre gekennzeichnet“ beschreiben.

In Befragungen folgen die erhobenen Merkmale im Gegensatz dazu dem Interesse der Forschung. Dies hat den Vorteil, dass neben Merkmalen wie Alter, Beruf oder Bildung auch solche erfasst werden können, die administrativ nicht interessant sind, wie beispielsweise Informationen zum Haushaltskontext (z. B. Höhe des Einkommens eines eventuell vorhandenen Lebenspartners, Vermögen, Hausbesitz, soziale Einbettung etc.), zu Einstellungen (etwa zu Familie und Beruf), zur Zufriedenheit (etwa mit der Tätigkeit, dem Lebensstandard oder dem Lohn) oder zum Arbeitsumfeld (z. B. betriebliche Zusatzleistungen). Die Messung dieser Merkmale folgt zudem einer am Forschungsgegenstand orientierten Operationalisierung, während die Erfassung der Informationen in Prozessdaten meist auf gesetzlichen Definitionen beruht. Dies kann zwar für die Verwaltung selbst zweckmäßig sein, ist aber für spezielle wissenschaftliche Zwecke häufig unbrauchbar oder zumindest zu grob, um daraus analytischen Nutzen zu ziehen.<sup>6</sup>

Schließlich können Befragungen so organisiert werden, dass sie die von den Forschern benötigten Informationen zeitnah erfassen und für Analysen zur Verfügung stellen. Die Aufbereitung von Prozessdaten ist dagegen ein zeitaufwändiges Verfahren, welches dazu führt, dass oft Jahre zwischen der Erfassung und der Bereitstellung für wissenschaftliche Analysen liegen.

Wie aus der kurzen Gegenüberstellung zu ersehen ist, liegen die Vor- und Nachteile beider Datenarten auf unterschiedlichen Gebieten. Es bietet sich daher ihre Kombination an, um die Schwächen der jeweils anderen Datenquelle auszugleichen (z. B. Heckman, LaLonde, Smith 1999: 130). Durch die Verknüpfung von prozessproduzierten Daten mit speziell für die Forschungsfrage erhobenen Befragungsdaten kann ein Maximum an Aussagekraft und Validität der Ergebnisse erreicht werden. Gerade, aber natürlich nicht nur, im Bereich der maßnahmeevaluierenden Arbeitsmarktfor-

---

<sup>6</sup> Daneben liegen Befragungen in der Regel in Händen von „Befragungsexperten“, was die Optimierung der Abläufe zur Datenerhebung garantiert. Dies betrifft sowohl die Entwicklung der Erhebungsinstrumente als auch die tatsächliche Befragung durch einen Interviewer.

schung ist es von hoher Bedeutung, möglichst reichhaltige Informationen zu den Untersuchungseinheiten zu besitzen.<sup>7</sup>

Eine solche direkte Verknüpfung bringt mehrere Vorteile mit sich:

- Zunächst stehen für jedes Individuum mehr Informationen als bei der Verwendung einer einzelnen Datenquelle zur Verfügung. So können beispielsweise die individuen- und arbeitsverhältnisbezogenen Prozessdaten mit Informationen zum Haushaltskontext und zur sozialen Einbettung des Individuums ergänzt werden.
- Daneben liegen die Vor- und Nachteile beider Datenarten auf unterschiedlichen Gebieten, so dass die Merkmale jeweils aus der Datenquelle bezogen werden können, welche die validesten Informationen liefert. Damit ließe sich vor allem eine höhere Genauigkeit, z. B. bei Merkmalen der Ausbildungs- und Erwerbsbiografie, erzielen.
- Da die Befragung zeitnah Daten liefert, die Aufbereitung von Prozessdaten zumindest als „scientific use file“ dagegen teilweise Jahre in Anspruch nimmt, können durch die Verknüpfung von Befragungsdaten mit bereits vorliegenden Prozessdaten frühzeitig Ergebnisse berichtet werden.
- Schließlich müssen bestimmte Informationen im Rahmen der Befragung nicht zusätzlich erhoben werden, da sie in den Prozessdaten bereits vorliegen. Die damit verbundene kürzere Interviewdauer würde die Bereitschaft zu Befragungsteilnahmen erhöhen, das Risiko von Interviewabbrüchen mindern und Befragungskosten verringern.

Allerdings müssen die Befragten aus Gründen des Datenschutzes einer solchen Datenverknüpfung zustimmen. Dabei besteht eventuell das Problem, dass die Bereitschaft dazu systematisch mit untersuchungsrelevanten Eigenschaften der Personen variieren kann, was dann zu einer Art „Ver-

---

<sup>7</sup> Der Grund liegt darin, dass bei der Maßnahmeevaluation kausale Effekte (Holland 1986; Heckman et al. 1999) gemessen werden sollen. Diese sind nur identifizierbar, wenn die gesamte maßnahmeergebnisrelevante Selektivität der Maßnahmeteilnahme mit den verfügbaren Daten abgebildet werden kann. Je mehr sich die Maßnahmeteilnahme an nicht mit Prozessdaten erfassbaren Faktoren (z. B. Suchintensität, Motivation, Haushaltskontext, etc.) orientiert, desto wichtiger werden zusätzliche Informationen aus Befragungen.

knüpfungsbias“ der Untersuchung führt. Da in multivariaten Analysen, die das gesamte Spektrum der Prozess- und Befragungsdaten nutzen wollen, nur Personen berücksichtigt werden können, die der Verknüpfung zustimmen, ist zu fragen, ob eine solche Beschränkung der Analysegesamtheit nicht zu Problemen führt, welche die Vorteile der Datenverknüpfung wieder zunichte machen. Bisher wurde diese Fragestellung weder systematisch verfolgt, noch gab es dazu geeignete Daten.

### **3 Theoretische Überlegungen zum Zustimmungsverhalten**

Im Folgenden sollen zunächst theoriegeleitet Einflussfaktoren der Zustimmung zum Zusammenspielen von Prozess- und Befragungsdaten identifiziert und – darauf aufbauend – Orientierungshypothesen für die anschließenden Analysen formuliert werden. In der Umfrageforschung werden zur Erklärung des Befragtenverhaltens verschiedene Mechanismen diskutiert, die hier speziell zur Erklärung des Zustimmungsverhaltens herangezogen werden.

Zunächst können Überlegungen des Modells der rationalen Wahlhandlung auf das Handeln in der Befragungssituation übertragen werden: Ausgangspunkt dieses Modells sind entscheidungstheoretische Überlegungen, die situationale Unsicherheiten und Interpretationskonflikte über das Modell des subjektiv erwarteten Nutzens berücksichtigen. Die zentrale theoretische Annahme bezieht sich auf das Grundprinzip menschlichen Handelns und besagt, dass Menschen durch ihr Handeln ihren Nutzen maximieren: Sie wählen unter den gegebenen Restriktionen die Handlung, von der sie für sich subjektiv den höchsten Nutzen erwarten.<sup>8</sup> Auch in der Befragungssituation sind von einer Person Entscheidungen zu treffen. So muss sie entscheiden, ob sie überhaupt an einer Befragung teilnimmt, wie sie antwortet – beispielsweise, ob sie bei Faktenfragen die Mühe langen und intensiven Nachdenkens auf sich nimmt, ob sie grundsätzlich wahr-

---

<sup>8</sup> Angemerkt sei, dass hier keine Annahmen über die konkreten kognitiven Prozesse und auch nicht darüber, ob diese bewusst ablaufen, getroffen werden. Damit findet keine Einschränkung auf „Zweckrationalität“ im Sinne von Max Weber statt. „Rationalität“ bedeutet hier nur, dass sich Menschen so verhalten, als ob sie ihre Präferenzen nach einem logischen Muster ordnen und ihr Verhalten nach diesen Präferenzen ausrichten können, dass sie also systematisch auf positive und negative Anreize reagieren können (vgl. Riker, Ordeshook 1973; Wippler 1987).

heitsgemäß antwortet oder lieber so, wie es sozial erwünscht ist – oder eben ob sie erlaubt, dass ihre Befragungsdaten mit Prozessdaten kombiniert werden. Das allgemeine Modell ist also – wie diese kurzen exemplarischen Ausführungen zeigen – prinzipiell auf alle Entscheidungen in der Befragungssituation anwendbar.

Ein Bestimmungsfaktor des Zustimmungsverhaltens liegt also in Kosten-Nutzenerwägungen des Befragten. Die Höhe des Nutzens, den eine befragte Person mit der Zustimmung verbindet, ist von ihrer Erwartung positiver Konsequenzen durch das Zusammenspielen für ihre persönliche Situation abhängig. Es ist allerdings zu vermuten, dass der Nutzen einer Zustimmung für den Befragten eher gering ausfällt. Dagegen ergeben sich durch die Verweigerung für die Befragten soziale Kosten der Rechtfertigung gegenüber dem Interviewer, die sich je nach Unmittelbarkeit des Kontaktes (face-to-face, telefonisch, postalisch) unterscheiden werden. Da aus der Zustimmung wohl kaum Nutzen, aus der Ablehnung aber Kosten entstehen, ist insgesamt von einer eher hohen Zustimmung auszugehen.

Neben den allgemeinen Rechtfertigungskosten entstehen für einige Teilgruppen der Befragten weitere Kosten. Wenn die befragte Person etwa befürchtet, dass ihre Daten nicht vertraulich behandelt werden oder dass sie eventuell sogar Kontrollzwecken dienen, so erhöht dies die Kosten, die für sie mit einer Zustimmung zum Zusammenspielen verbunden sind. Alter, Geschlecht und Bildung können in dieser Hinsicht als Proxy-Indikatoren betrachtet werden: Bedenkt man beispielsweise, dass jüngere Menschen, Männer und Befragte mit höherer Bildung mehr Erfahrungen mit der Nutzung des Internets haben, einem Medium, in dem eine Vielzahl von Daten ausgetauscht werden, so ist zu vermuten, dass diese Gruppen eher einer Verknüpfung der Daten zustimmen. Diese Vermutung wird gestützt durch die Ergebnisse einer Untersuchung von Pascale und Mayer (2004), bei der sich Alter, Geschlecht und Bildung als geeignete Prädiktoren für die Zustimmung zu einer ähnlich sensiblen Frage erwiesen, nämlich zu der Frage, ob für die Zwecke von Folgebefragungen die Informationen aus früheren Befragungen an andere Haushaltsmitglieder weitergegeben werden können, falls die ursprünglich befragte Person nicht angetroffen wird. Es ist zudem zu vermuten, dass die Befürchtung, dass Daten zu Kontrollzwecken verwendet werden, aufgrund der historischen Erfahrungen in Ostdeutschland höher ist.

Zu den Kosten gehört auch, dass die Befragung Zeit kostet und insgesamt eine Unannehmlichkeit darstellt, etwa weil sie als anstrengend empfunden wurde. Als Merkmale, die in diese Richtung wirken, sind neben der Dauer der Befragung und der Arbeitszeit der erwerbstätigen Befragten – die Befragung stellt für sie eine zusätzliche Belastung dar – Indikatoren zu nennen, die auf die Schwierigkeiten verweisen, die mit der Beantwortung der Fragen verbunden sein können: Zu denken ist hier an die Bildung, das Alter und die Nationalität. Personen mit niedriger Bildung oder mit Migrationshintergrund sowie ältere Befragte sollten die Befragung als anstrengender empfinden, damit mit der Befragungssituation eher unzufrieden sein und daher eher ihre Zustimmung zur Datenverknüpfung verweigern. Als weiterer Indikator für die Kosten der Interviewteilnahme kann die Bewertung der Freizeit interpretiert werden, und zwar konkret für die damit verbundenen Opportunitätskosten: Je wichtiger die Freizeit bewertet wird, umso höher sollten diese Kosten sein.

Austauschtheoretische Überlegungen (Blau 1968; Homans 1964) lassen hinsichtlich der Kosten der Zustimmungsverweigerung erwarten, dass Personen, die in irgendeiner Weise vom Staat im Allgemeinen oder vom Arbeitsamt bzw. der Arbeitsagentur im Besonderen Leistungen erhielten, einer Datenverknüpfung eher zustimmen als Personen, für die dies nicht zutrifft. Denn es ist zu vermuten, dass sich Leistungsempfänger den staatlichen Stellen gegenüber, in deren Auftrag die Befragung durchgeführt wird, verpflichtet sehen, die ihnen gewährte Hilfe mit einer Unterstützung der Befragung zu vergelten. Die Tatsache, dass jemand eine staatliche Förderung erhielt, sollte vor dem Hintergrund dieser Argumentation die Wahrscheinlichkeit einer Zustimmung erhöhen. Hier sind einerseits die Geförderten nach dem Mainzer Modell zu nennen, andererseits kann es aber auch unter den Befragten der Kontrollgruppe Personen mit Förderung geben.

Als weiterer Einflussfaktor der Teilnahme an Befragungen – hier allerdings speziell in Betrieben – identifizieren Tomaskovic-Devey, Leiter und Thompson (1995: 80ff.) neben der Fähigkeit („capacity“) und der Motivation, die Fragen zu beantworten („motive to respond“), die Autorität („authority“) der Zielperson. Sie beziehen sich dabei auf „das Ausmaß, in dem der Antwortende die formelle oder informelle Befugnis zur Antwort hat“ (Schnabel 1997: 126). Da mit niedrigerer Befugnis die Kosten der

Teilnahme ansteigen, weil erst die Erlaubnis eingeholt werden muss, und da bei niedriger Befugnis die Wahrscheinlichkeit negativer Sanktionen höher ist, wenn das Interview ohne Erlaubnis gegeben wird, sollte die Teilnahmebereitschaft dieser Personen bei Betriebsbefragungen niedriger sein. Diese Argumentation lässt sich auch auf die Zustimmung zum Zusammenspielen der Daten in der vorliegenden Untersuchung übertragen. Konkret ist zu vermuten, dass die Befugnis, eine eigenständige Entscheidung zu treffen, mit dem Grad der Unabhängigkeit der Zielperson und diese wiederum mit ihrer Stellung im Haushaltskontext zusammenhängt. So ist anzunehmen, dass Personen unabhängiger sind, wenn sie die Hauptverdiener im Haushalt sind. Zudem ist eine Interaktion mit dem Geschlecht der befragten Person zu vermuten: Der formulierte Zusammenhang sollte stärker für weibliche Zielpersonen als für männliche gelten.

Entscheidungen über die Zustimmung werden nicht nur zweckrational getroffen, sondern basieren auch auf den Einstellungen der Befragten gegenüber dem Untersuchungsziel und der auftraggebenden Institution, hier beispielsweise dem Bundesministerium für Arbeit und Sozialordnung.<sup>9</sup> Auch aus dieser Perspektive der relevanten Einstellungen ist zunächst auf eine hohe Zustimmungsbereitschaft zu schließen, da bei der Zielperson bereits die Entscheidung zu Interviewteilnahme positiv ausgefallen war. Dies deutet einerseits auf eine im Allgemeinen positive Einstellung zur Befragung hin, andererseits auf ein gewisses Commitment, welches die Zielperson bereits eingegangen ist. Eine Zustimmung zum Zusammenspielen von Prozess- und Befragungsdaten steht in Einklang mit diesem Commitment und gewährleistet damit eine kognitive Balance, da beide Entscheidungen positiv sind. Eine Verweigerung der Zustimmung dagegen würde ein Ungleichgewicht (kognitive Dissonanz) bedeuten und ist daher weniger wahrscheinlich. Allerdings erkennen die Befragten die konkreten Eigen-

---

<sup>9</sup> Grundsätzlich gibt es verschiedene Ansichten über das Verhältnis von Kosten-Nutzen-Erwägungen und Einstellungen. Einerseits wird argumentiert, dass Einstellungen in das entscheidungstheoretische Modell der Wahlhandlung integrierbar sind, indem sie als Mittel zur Vereinfachung der Zielstruktur aufgefasst werden und somit dann an Bedeutung gewinnen, wenn es sich um „low-cost“-Situationen handelt (z. B. Esser 1986, 1990). Andererseits wird davon ausgegangen, dass Personen auch unabhängig von Kosten-Nutzen-Kalkulationen einstellungsbasiert handeln (Engel et al. 2004: 57).

Noch eine Anmerkung zum Namen des Ministeriums: Es handelt sich hier um die Bezeichnung, die zur Zeit der Befragung gültig war.



schaften der Situation (z. B. der konkrete Inhalt der Befragung) erst im Verlauf der Interaktion mit dem Interviewer, so dass nicht automatisch mit einer Zustimmung zu rechnen ist, sondern diese erfolgt, „dependig on what is made salient and how much the person negatively or positively values the attribute“ (Groves et al. 2004: 177).

Die für die Zustimmung zur Datenverknüpfung relevanten Einstellungen können unterschiedliche Quellen besitzen. Beispielsweise kann die Erwerbsbiografie der Zielpersonen Einfluss auf ihre Einstellung gegenüber der Untersuchung haben: So ist zu vermuten, dass Personen mit Merkmalen, die auf schlechte Arbeitsmarktchancen hindeuten, sich vom Staat und seinen Repräsentanten enttäuscht fühlen und daher weniger bereit sind, die Befragung und die damit verbundenen Ziele zu unterstützen. Zu denken ist hierbei an Langzeitarbeitslose, Geringqualifizierte, allein Erziehende oder an Befragte aus Ostdeutschland.

Umgekehrt sollten Zielpersonen, die positive Erfahrungen mit staatlichen Stellen gemacht haben, in diesem Fall z. B. mit dem Arbeits- oder Sozialamt, auch eine positivere Einstellung zur Verwendung ihrer Daten haben. Positive Erfahrungen können z. B. darin bestehen, dass die angetretene Stelle vom (damaligen) Arbeitsamt vermittelt wurde. Speziell – aber nicht nur – bei zum Zeitpunkt der Befragung Geförderten sollte die Zustimmungsbereitschaft darüber hinaus davon abhängen, wie die Erfahrungen mit der entsprechenden Stelle sind. Dies kann über die subjektive Einschätzung zur Zufriedenheit erfasst werden. Für die Geförderten nach dem Mainzer Modell allgemein könnte schließlich das Ziel von Interesse sein, mit ihrer Teilnahme und mit ihrer Zustimmung zur Verknüpfung der Befragungsdaten mit Prozessinformationen dazu beizutragen, dass die Fördermaßnahme besser an die Bedürfnisse der Betroffenen angepasst wird.

Ob Einstellungen überhaupt aktiviert werden, das heißt handlungsrelevant werden, hängt auch davon ab, inwiefern die durchführende Institution im Bewusstsein des Befragten subjektiv repräsentiert ist (Fetchenhauer 1999) und dieses ist wiederum von der Nähe des Befragten zur Institution abhängig (vgl. auch Engel et al. 2004: 66). Daher wäre eine erhöhte Zustimmungswahrscheinlichkeit dann zu erwarten, wenn die befragte Person sich in staatlicher Anstellung befindet oder sogar beamtet ist. Zudem ist auch aus dieser Perspektive von Personen, deren Beschäftigung staatlich



gefördert wird oder wurde, eine stärkere Bereitschaft zur Zustimmung zu erwarten.

Bei der Aktivierung von Einstellungen spielen auch bestimmte Attribute der Befragungssituation, wie das Thema der Befragung, die (voraussichtliche) Dauer, der Verwendungszweck der Daten, die durchführende Institution oder Merkmale der Interviewer, eine wichtige Rolle. Diese offenbaren sich Befragten oft erst im Verlauf des Interviews. Dies deutet zum einen noch einmal auf die Dauer der Befragung als Einflussfaktor der Zustimmung hin, verweist aber auch auf weitere Merkmale der Interviewersituation. Da in der vorliegenden Untersuchung Telefoninterviews geführt wurden, bei denen ein hohes Maß an Standardisierung der Interviewersituation stattfindet, sollten sich deren Auswirkungen in Grenzen halten. Zu untersuchen ist allerdings der Einfluss von Interviewermerkmalen. Als am Telefon wahrnehmbares Merkmal kann hier das Geschlecht eine Rolle spielen: Es ist beispielsweise zu vermuten, dass insbesondere die Interaktion von junger Interviewerin und älterem männlichen Befragten die Zustimmungswahrscheinlichkeit positiv beeinflusst.

Einstellungen, die unabhängig von den konkreten Inhalten der Untersuchung eine Rolle spielen und für die Bereitschaft zum record linkage relevant werden können, sind die Stärke des Misstrauens bzw. der Wunsch, die eigene Privatsphäre zu schützen (vgl. Singer et al. 1993). Im Normalfall sind in Erhebungen in der Arbeitsmarktforschung keine Skalen zur Messung solcher Eigenschaften vorhanden. Ein geeigneter Proxy-Indikator zur Messung des Misstrauens eines Befragten gegenüber dem Interviewer bzw. der Befragung scheint die Beantwortung der Frage nach dem Einkommen zu sein: So ist für Personen, die bei dieser sensiblen Frage die Angabe verweigert haben, auch eine höhere Ablehnungsquote in Bezug auf das Zusammenspielen von Befragungs- und Prozessdaten zu vermuten. Diese Vermutung wird unterstützt durch die Ergebnisse von Schräpler (2004: 125): In Analysen zur Teilnahme bei Wiederholungsbefragungen im Rahmen des Sozio-oekonomischen Panels (SOEP) fand dieser einen negativen Zusammenhang zwischen der Verweigerung der Frage zum Einkommen und der Wahrscheinlichkeit, an Folgewellen teilzunehmen. Dabei ist davon auszugehen, dass der Grad, bis zu dem solches Misstrauen in Bezug auf die vertrauliche Behandlung der gegebenen Informationen im Interview selbst abgebaut wird, wiederum von Eigenschaften des Inter-

viewers und seiner Erfahrung abhängt (Pickery et al. 2001). Zu denken ist vor allem an ein freundliches und offenes Kontaktverhalten bzw. an die Basis hierfür, und zwar an die verbale Kommunikationsfähigkeit, die durch die Bildung des Interviewers abgebildet werden kann.

Es kann zusammenfassend festgehalten werden, dass auf der Basis verschiedener theoretischer Überlegungen eine Reihe von Merkmalen identifiziert werden konnten, von denen angenommen werden kann, dass sie die Zustimmung zur Verknüpfung von Befragungs- und Prozessdaten beeinflussen. Diese Vermutungen werden im nächsten Abschnitt empirisch überprüft.

## **4 Empirische Ergebnisse**

### **4.1 Daten und Indikatoren**

#### **4.1.1 Repräsentative Erhebung zur Evaluation des Mainzer Modells**

Im Rahmen der Untersuchung zur Wirkung des Mainzer Modells, bei dem zunächst regional begrenzt, später bundesweit ein Kombilohn gewährt wurde, wurden von TNS Infratest Sozialforschung Geförderte und ungefördert erwerbstätige Vergleichspersonen befragt. Die Befragungen sollten möglichst zeitnah zum Förderbeginn bzw. zum Abgang aus Arbeitslosigkeit durchgeführt werden, weil so mit den besten Auskünften zu den besonders interessierenden Verhältnissen unmittelbar vor und nach dem Förderbeginn bzw. dem Abgang aus Arbeitslosigkeit zu rechnen war. Die Stichprobe umfasste Förderzugänge bzw. Abgänge aus Arbeitslosigkeit im Zeitraum von Januar 2001 bis März 2003.

Das Mainzer Modell legt nicht unmittelbar eine bestimmte Vergleichsgruppe nahe: Grundsätzlich stand eine Förderung nach dem Mainzer Modell allen Personen offen, die im Inland eine Beschäftigung aufnehmen dürfen. Es knüpfte nicht an personenbezogenen Merkmalen wie vorheriger Arbeitslosigkeit oder fehlender Qualifikation an. Hier wurde eine Vergleichsgruppe ausgewählt, die sich einerseits an den anvisierten Zielgruppen des Mainzer Modells orientiert und andererseits eine möglichst einfache Stichprobenziehung ermöglicht. Einbezogen wurden Abgängerinnen bzw. Abgänger aus Arbeitslosigkeit, die keine abgeschlossene Berufsausbildung haben oder langzeitarbeitslos (mindestens seit einem Jahr ununterbrochen

arbeitslos gemeldet) waren. Die Stichprobe wurde zeitlich und regional proportional zu den Zugängen in die Förderung gezogen.

Insgesamt wurden 4.523 Personen befragt. Dabei handelt es sich um 3.080 geförderte und 1.443 nicht geförderte Befragte (Vergleichsgruppe). Unter den befragten Geförderten waren 705 Männer und 2.375 Frauen, in der Vergleichsgruppe waren es 695 Männer und 748 Frauen. In Ostdeutschland wurden insgesamt 1.549 Personen befragt, in Westdeutschland waren es 2.974.

Alle Befragten wurden um die Erlaubnis gebeten, ihre Befragungsdaten mit Prozessdaten verknüpfen zu dürfen. Insgesamt stimmten 74,4 % diesem Anliegen zu. Die Frage, mit der die Zustimmung eingeholt wurde, lautete:

*„Zum Schluss haben wir noch eine Bitte an Sie. Für die Ergebnisse dieser Untersuchung wäre es von Vorteil, bei der Auswertung Auszüge aus Ihren Sozialversicherungsdaten einzubeziehen, die beim Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit in Nürnberg vorliegen. Dabei handelt es sich um Informationen zu vorausgegangenen Zeiten der Beschäftigung, der Arbeitslosigkeit und der Beteiligung an Maßnahmen des Arbeitsamtes.*

*Aus Gründen des Datenschutzes kann das nicht ohne Ihr Einverständnis geschehen, um das ich Sie herzlich bitte. Dies ist natürlich ebenso freiwillig, wie das Interview, das Sie freundlicherweise gegeben haben. Alle Vorschriften des Datenschutzes und der anonymen Ergebnisdarstellung gelten selbstverständlich für diese zusätzlichen Daten in gleicher Weise. Wenn Sie einverstanden sind, erhalten Sie von uns noch eine schriftliche Bestätigung, in der das beschrieben ist.*

*Ich darf also notieren: Sind Sie mit der Hinzunahme dieser Sozialversicherungsdaten einverstanden oder nicht einverstanden?“*

#### **4.1.2 Indikatoren der Einflussfaktoren der Zustimmungsbereitschaft**

Auf der Basis der oben dargestellten theoretischen Überlegungen konnten verschiedene Merkmale identifiziert werden, von denen angenommen werden kann, dass sie die Zustimmung zur Verknüpfung von Befragungs-

und Prozessdaten beeinflussen. In Übersicht 1 werden die Merkmalsträger, das Merkmal und die erwartete Wirkungsrichtung auf die Zustimmungswahrscheinlichkeit noch einmal zusammenfassend dargestellt.

### Übersicht 1: Theoretisch identifizierte Einflussfaktoren

Merkmalsträger	Merkmal	Wahrscheinlichkeit der Zustimmung
Interviewer/in	Alter	+
Interviewer/in	Geschlecht	Weiblich: +
Interviewer/in	Bildung	+
Interview	Dauer (Zahl der Variablen)	–
Befragte Person	Alter	–
Befragte Person	Nationalität	Deutsch: +
Befragte Person	Bildung	+
Befragte Person	Arbeitslosigkeit vorher <sup>a)</sup>	–
Befragte Person	Ausbildung	Mit Ausbildung: +
Befragte Person	Erwerbsstatus vorher <sup>a)</sup>	Nicht erwerbstätig: +
Befragte Person	Öffentlicher Dienst	+
Befragte Person	Zufriedenheit mit Kombilohnförderung	+
Befragte Person	Bruttoerwerbseinkommen nachher <sup>a)</sup>	+ Keine Angabe: –
Befragte Person	Sozialhilfebezug vor und/oder nach Referenzzeitpunkt <sup>a)</sup>	Kein Bezug: –
Befragte Person	Geschlecht	Weiblich: –
Befragte Person	Lebensform	Allein: +
Befragte Person	Erwerbstätigkeit nachher <sup>a)</sup>	+
Erwerbstätige Befragte	Arbeitszeit	–
Befragte Person	Erwerbstätigkeit Partner bei Befragung	–
Befragte Person	Vermittlung durch Arbeitsagentur	+
Befragte Person	Wichtigkeit von Freizeit	–
Befragte Person	Region	Ost: –

<sup>a)</sup> Vorher bzw. nachher beziehen sich auf den Referenzzeitpunkt. Mit Referenzzeitpunkt ist der Zeitpunkt des Förderbeginns (Geförderte) bzw. des Abgangs aus Arbeitslosigkeit in reguläre Beschäftigung (Vergleichsgruppe) gemeint.

## **4.2 Erklärung der Zustimmungsbereitschaft**

### **4.2.1 Unterschiede zwischen Zustimmenden und Nicht-Zustimmenden**

Vor der multivariaten Analyse stellen wir im Folgenden kurz bivariat die Zustimmung in Abhängigkeit von den verschiedenen Merkmalen dar (vgl. Anhang) Bezüglich der Interviewermerkmale ist zu erkennen, dass ältere Interviewer, Interviewer mit Fachhochschulreife oder Abitur und Interviewerinnen häufiger die Zustimmung zur Verknüpfung erhalten.

Befragte in mittlerem Alter, deutsche Befragte und Zielpersonen mit höherer Bildung oder mit einer Berufsausbildung stimmen einer Verknüpfung eher zu. Vor dem Referenzzeitpunkt arbeitslose Befragte, besonders Langzeitarbeitslose, erlauben das Zusammenspielen seltener, vorher Erwerbstätige häufiger als die anderen Befragten. Nach dem Referenzzeitpunkt Geförderte, die mit der Förderung zufrieden sind, und Befragte mit niedrigem oder hohem Erwerbseinkommen stimmen eher einem Zusammenspielen zu, Personen, die die Angaben zum Einkommen oder zum Bezug von „laufender Hilfe zum Lebensunterhalt“ (HLU) verweigern, dagegen seltener. Zielpersonen, die vor und/oder nach dem Referenzzeitpunkt HLU erhielten, erlauben überdurchschnittlich oft die Datenverknüpfung. Frauen und Befragte mit Partner geben die Zustimmung seltener, Erwerbstätige, insbesondere im Öffentlichen Dienst, häufiger. Personen mit erwerbstätigem Partner und Westdeutsche erlauben die Datenverknüpfung unterdurchschnittlich oft.

### **4.2.2 Mehrebenenmodell zur Erklärung der Zustimmungsbereitschaft**

Um die empirische Relevanz der theoretisch identifizierten Einflussfaktoren auf die Zustimmungsbereitschaft zur überprüfen, wurde ein Mehr-Ebenen-Probit-Modell mit zufälligen Effekten durchgeführt (z. B. Guo, Zhao 2000). Damit wurde der Tatsache Rechnung getragen, dass neben Merkmalen der Befragten auch auf der höheren Ebene der Interviewer Merkmale auf ihren Einfluss auf die Zustimmung überprüft werden sollen. Das Modell geht von

einer latenten Zustimmungsnéigung  $y_{ij}^*$  beim  $i$ -ten Befragten, der vom  $j$ -ten Interviewer befragt wurde, aus:

$$y_{ij}^* = \beta_0 + x'_{ij}\beta_1 + z'_j\beta_2 + u_j + e_{ij}$$

Die Zustimmungsnéigung ist abhängig von Merkmalen der befragten Individuen  $x_{ij}$  und von Merkmalen der Interviewer  $z_j$ , wobei  $e_{ij}$  ein zufälliger Effekt auf der Befragtenebene und  $u_j$  ein zufälliger Effekt auf der Ebene der Interviewer ist. Beobachtet wird nicht die Zustimmungsnéigung, sondern nur die tatsächliche Zustimmung oder Nichtzustimmung  $y_{ij}$ , die den Wert 1 bei Zustimmung (falls  $y_{ij}^* > 0$ ) und den Wert 0 bei ausbleibender Zustimmung (falls  $y_{ij}^* \leq 0$ ) annimmt. Unter der Annahme, dass  $e_{ij}$  standardnormalverteilt ist, ergibt sich ein Probit-Modell der Zustimmungswahrscheinlichkeit, wobei  $\Phi$  für die Wahrscheinlichkeitsfunktion der Standardnormalverteilung steht.

$$\Pr(y_{ij} = 1 | x, z, u) = \Phi(\beta_0 + x'_{ij}\beta_1 + z'_j\beta_2 + u_j)$$

Die Ergebnisse der Schätzung dieses Modells zeigt Tabelle 1. Dargestellt sind zwei Modelle: eines nur mit Haupteffekten, eines mit zusätzlichen Interaktionseffekten. Die Interaktionseffekte beziehen sich auf Geschlecht und Alter von Interviewerin und Befragtem sowie auf Geschlecht, Partnerschaft und Erwerbstätigkeit der Partnerin oder des Partners, soweit vorhanden, sowie die Erwerbstätigkeit der befragten Person, die Arbeitszeit und Vermittlung der Stelle.

Tabelle 1: Mehr-Ebenen-Probit-Modell mit zufälligen Effekten<sup>10</sup>

	Interaktionseffekte		Ohne		Mit	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
<b>Interviewermerkmale</b>						
<b>Alter</b>						
Bis unter 25 Jahre	-0,058	0,097	-0,056	0,097		
25 bis 34 Jahre	-0,020	0,127	-0,015	0,127		
35 bis 44 Jahre	0,065	0,156	0,077	0,157		
45 bis 54 Jahre	0,421	***	0,157	0,157	0,432	***
55 Jahre oder mehr						
<b>Geschlecht</b>						
Männlich						
Weiblich	-0,018	0,083	-0,027	0,083		
<b>Interaktionseffekt: Interviewer - Befragter</b>						
Andere Konstellationen						
Junge Interviewerin - älterer Befragter			0,157	0,136		
<b>Bildung</b>						
Anderes						
Fachhochschulreife, Abitur	0,157	**	0,080	0,080	0,156	*
<b>Dauer des Interviews</b>						
<b>Zahl der beantworteten Fragen</b>						
	-0,001	0,002	-0,001	0,002		
<b>Befragtenmerkmale</b>						
<b>Alter</b>						
Bis unter 25 Jahre						
25 bis 34 Jahre	0,064	0,075	0,051	0,077		
35 bis 44 Jahre	0,061	0,074	0,046	0,075		
45 bis 54 Jahre	0,015	0,083	-0,021	0,086		
55 Jahre oder mehr	0,006	0,111	-0,042	0,115		
Keine Angabe	-0,468	0,534	-0,483	0,536		
<b>Staatsangehörigkeit</b>						
Nicht deutsch						
Deutsch	0,145	*	0,076	0,076	0,148	*
<b>Bildung</b>						
Kein Abschluss						
Volks-, Hauptschule	0,153	*	0,090	0,091	0,146	0,091
Mittl. Reife, POS 10.Klasse	0,128		0,094	0,094	0,121	0,094
Fachhochschulreife, Abitur	0,057		0,098	0,098	0,050	0,098
<b>Arbeitslosigkeit vor Referenzzeitpunkt laut BA</b>						
Nicht arbeitslos						
Bis unter 12 Monate	-0,150	*	0,083	0,084	-0,146	*
12 Monate oder mehr	-0,222	**	0,094	0,094	-0,221	**
Angaben fehlen	0,067		0,149	0,148	0,062	0,148

<sup>10</sup> Die Referenzkategorien sind jeweils genannt und enthalten keine Einträge.

	Interaktionseffekte		Ohne		Mit	
	Coef.	S.E.	Coef.	S.E.		
<b>Befragtenmerkmale</b>						
<b>Ausbildung</b>						
Keine						
Ausbildung/Studium	-0,010	0,054	-0,015	0,054		
<b>Erwerbsstatus vor Referenzzeitpunkt</b>						
Erwerbstätig						
Anderes	0,044	0,080	0,044	0,080		
<b>Merkmale des Status nach dem Referenzzeitpunkt</b>						
<b>Öffentlicher Dienst</b>						
Nein						
Ja	-0,039	0,094	-0,038	0,094		
<b>Zufriedenheit mit Höhe der Kombilohnförderung</b>						
Nicht Kombilohngefördert						
Weniger zufrieden, keine Angabe	-0,082	0,075	-0,081	0,075		
Sehr zufrieden, zufrieden	0,191	***	0,187	***	0,065	
<b>Bruttoerwerbseinkommen im Monat nach dem Referenzzeitpunkt</b>						
Nicht vorhanden						
Bis 511 EURO	0,014	0,085	0,018	0,085		
512 - 1023 EURO	0,038	0,070	0,039	0,070		
1024 EURO oder mehr	0,185	**	0,177	**	0,080	
Angabe fehlt	-0,565	***	-0,565	***	0,085	
<b>Merkmale des Haushaltskontextes</b>						
<b>Sozialhilfe vor und nach Referenzzeitpunkt</b>						
Weder vorher, noch nachher						
Vorher und nachher	0,294	***	0,285	***	0,090	
Nur nachher	0,156	0,194	0,161	0,194		
Nur vorher	0,127	*	0,123	0,076		
Mindestens einmal keine Angabe	-0,477	***	-0,474	***	0,164	
<b>Geschlecht</b>						
Männlich						
Weiblich	-0,207	***	0,055			
<b>Lebensform</b>						
Ohne Partner						
Mit Partner	-0,017	0,059				
<b>Erwerbstätigkeit befragte Person</b>						
Nein						
Teilzeit, Vermittlung durch Arbeitsagentur	0,132	0,114	0,129	0,114		
Teilzeit, Vermittlung nicht durch Arbeitsagentur	0,065	0,091	0,064	0,091		
Vollzeit, Vermittlung durch Arbeitsagentur	0,016	0,112	0,012	0,112		
Vollzeit, Vermittlung nicht durch Arbeitsagentur	0,014	0,090	0,012	0,090		
<b>Erwerbstätigkeit Partner</b>						
Nein						
Ja	0,036	0,064				
<b>Geschlecht und Lebensform</b>						
Frau mit erwerbstätigem Partner						
Mann mit erwerbstätiger Partnerin			0,233	**	0,106	
Frau mit nicht erwerbstätigem Partner			-0,036		0,081	
Mann mit nicht erwerbstätiger Partnerin			0,182	**	0,088	



	Interaktionseffekte		Ohne		Mit	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Frau ohne Partner					0,003	0,067
Mann ohne Partnerin					0,138	0,092
<b>Wichtigkeit von Freizeit</b>						
nicht wichtig						
wichtig	-0,024	0,052			-0,024	0,052
<b>Region</b>						
West						
Ost	0,158	***	0,053		0,164	***
<b>Konstante</b>	0,416	*	0,230		0,251	0,245
<b>Var at level 1</b>	0,173		0,004		0,173	0,004
<b>level 2</b>	0,092		0,025		0,094	0,025
<b>Fallzahl</b>	4.523				4.523	
<b>Gruppen</b>	142				142	
<b>-2LL Nullmodell</b>	-2.624,0				-2.624,0	
<b>-2LL volles Modell</b>	-2.505,2				-2.504,1	
<b>Pseudo R<sup>2</sup></b>	0,05				0,05	

Im Hinblick auf Alter und Geschlecht des Interviewers war argumentiert worden, dass die Zielperson ältere und weibliche Interviewerinnen eher als vertrauenswürdig einschätzt und daher eher die Zustimmung zum Zusammenspielen von Befragungs- und Prozessdaten gibt. Für das Alter sprechen Ergebnisse für die Hypothese, für das Geschlecht ist kein Effekt feststellbar. Ebenfalls nicht bestätigt wird die Vermutung, dass ältere Befragte bei einer jungen Interviewerin eher die Datenverknüpfung erlauben. Der Interaktionseffekt zwischen Alter und Geschlecht von Interviewerin und Befragtem ist nicht signifikant. Mit höherer Bildung der Interviewer geht dagegen eine erhöhte Zustimmung der Befragten einher. Dies war mit dem Argument, dass mit der Bildung die Kommunikationsstärke steigt, erwartet worden.

Entgegen der Erwartungen zeigt sich kein Einfluss der Dauer des gerade gegebenen Interviews sowie des Alters und der Bildung der befragten Person. Diese drei Merkmale waren als Indikatoren für die Belastung durch das Interview eingeführt worden. Bei den beiden letzten Merkmalen war zudem argumentiert worden, dass sie als Proxy-Indikatoren für das Vertrauen in die Gewährleistung der Anonymität der Daten interpretiert werden können. Beide Argumentationsstränge waren auch für die Staatsangehörigkeit angeführt worden. Diesbezüglich bestätigen sich die Vermutungen: Deutsche Zielpersonen stimmen der Verknüpfung eher zu als nicht deutsche.

Für Zielpersonen ohne Berufsausbildung und solche, die vor dem Referenzzeitpunkt arbeitslos waren, war erwartet worden, dass sie die Datenverknüpfung eher nicht erlauben, weil sie vom Arbeitsmarkt und den arbeitsmarktpolitischen Institutionen enttäuscht sind. Die Hypothese wird für Arbeitslose bestätigt, für Personen ohne Berufsausbildung allerdings nicht. Ebenso wenig hängt die Tatsache, dass die Zielpersonen vor dem Referenzzeitpunkt einer Erwerbstätigkeit nachgingen, mit der Zustimmung zusammen.<sup>11</sup>

In Bezug auf die Situation nach dem Referenzzeitpunkt, also dem Zeitpunkt der Aufnahme einer geförderten Beschäftigung bzw. dem Abgang aus Arbeitslosigkeit, zeigt sich erwartungsgemäß, dass Personen, die mit der Förderung zufrieden sind, und solche, die ein hohes Erwerbseinkommen erzielen, der Datenverknüpfung eher zustimmen. Es bestätigt sich auch die Vermutung, dass der Wunsch nach Schutz der Privatsphäre sich darin äußert, dass bestimmte Fragen nicht beantwortet werden und aus dem gleichen Grund auch eher die Erlaubnis zur Datenverknüpfung nicht gegeben wird: Personen, die die Angabe zum Einkommen im Monat nach dem Referenzzeitpunkt verweigern, lassen mit geringerer Wahrscheinlichkeit das Zusammenspielen der Daten zu.

Betrachtet man den Haushaltskontext, so sprechen die Ergebnisse für die Vermutung, dass der Bezug von staatlichen Unterstützungsleistungen die Bereitschaft zum Erlauben der Datenverknüpfung erhöht. Zudem stimmen weibliche Zielpersonen tatsächlich seltener als Männer dem Zusammenspielen zu. Dabei sind keine Unterschiede zwischen Frauen mit erwerbstätigem Partner, mit nicht erwerbstätigem Partner und ohne Partner festzustellen. Unter den Männern zeigt sich die höchste Zustimmungsbereitschaft bei denen mit erwerbstätiger Partnerin. Die eigene Erwerbstätigkeit hat dagegen in keinem Modell einen Einfluss auf die Zustimmungswahrscheinlichkeit. Schließlich ist noch festzustellen, dass entgegen der Erwartung in den Regionen Ostdeutschlands die Erlaubnis zur Datenverknüpfung signifikant höher ist.

---

<sup>11</sup> Es sei darauf hingewiesen, dass es beim Arbeitsamt bzw. mittlerweile der Arbeitsagentur arbeitslos Gemeldeten innerhalb gewisser Grenzen erlaubt ist, eine Erwerbstätigkeit auszuüben.

Die Tatsache, dass die befragte Person erwerbstätig ist, zeigt keinen Effekt.<sup>12</sup> Auch Interaktionseffekte der Erwerbstätigkeit mit der Arbeitszeit und der Vermittlung der Stelle durch das Arbeitsamt sind nicht feststellbar. Schließlich zeigt sich auch für die Bewertung der Freizeit in dem Modell kein Einfluss.

Zusammenfassend kann festgehalten werden: Die Resultate der multivariaten Analysen zeigen, dass die Hypothesen teilweise verworfen werden, teilweise aber auch nicht: Ein hoher mit der Befragung in Verbindung gebrachter Nutzen, keine erwarteten negativen Konsequenzen, Vertrauen in die Anonymität der Auswertung und der Wunsch nach Schutz der eigenen Privatsphäre – alle vermuteten Mechanismen scheinen zu wirken. Insgesamt erklären die in den Modellen enthaltenen Merkmale aber nur wenig von der Zustimmungsbereitschaft. Dies kann als eine eher geringe Selektivität der Zustimmungsbereitschaft interpretiert werden, wobei anzumerken ist, dass meist nur Proxy-Variablen für die theoretisch relevanten Konzepte verwendet werden konnten. Insofern können erst weitere Analysen, die auf bessere Indikatoren zugreifen können, die Frage des Ausmaßes der Selektivität klären. Bis zum Vorliegen solcher Analysen sind die folgenden Vergleiche zwischen Modellen mit und ohne Zustimmung als eine weitere Annäherung zu verstehen.

### **4.3 Auswirkung der Selektivität auf Forschungsergebnisse – Zwei Beispiele**

Um näher zu untersuchen, ob die festgestellten Selektivitäten das Ergebnis konkreter Forschungsarbeiten beeinflussen, wenn nur Beobachtungen von Personen ohne Datenschutzbedenken vorliegen, wird im Folgenden eine Regression der Arbeitszeit durchgeführt. Diese wird auf Basis zweier unterschiedlicher Analysegesamtheiten berechnet. Modell 1 wird mit allen verfügbaren Fällen in der Stichprobe durchgeführt, Modell 2 dagegen nur mit solchen Fällen, die der Datenzusammenspielung zugestimmt haben. Zeigen sich in den Modellen unterschiedliche Regressionskoeffizienten, kann davon ausgegangen werden, dass die Selektivität hinreichend groß

---

<sup>12</sup> Das Modell mit dem entsprechenden Haupteffekt wurde aus Platzgründen hier nicht dargestellt.

ist, um solche Unterschiede zu erzeugen. Ergeben sich keine Unterschiede, ist die Selektivität der Zustimmungsbereitschaft eher gering.

Allerdings lassen sich die so gewonnenen Ergebnisse nicht auf andere Analysen übertragen, da sich bei dem Vorgehen stets nur die Selektivität auswirkt, die auch mit der Arbeitszeit bzw. den Regressoren korreliert ist. Um zumindest ein breiteres Spektrum an inhaltlichen Fragestellungen abzudecken, wurde neben der linearen Regression der Arbeitszeit auch eine logistische Regression zur Zufriedenheit mit der Tätigkeit berechnet, eine Analyse also, in der es weniger um „harte“ Informationen als vielmehr um Einstellungsfragen geht.

Im Folgenden werden die beiden Fragestellungen jeweils einmal mit allen verfügbaren Fällen bearbeitet und einmal so, als hätte die Datenverknüpfung stattgefunden, also ohne die Nichtzustimmer.<sup>13</sup> Es werden die resultierenden Modelle dahingehend betrachtet, ob sie zu unterschiedlichen Schlussfolgerungen bei der Identifikation statistisch signifikanter Einflussfaktoren auf die Arbeitszeit bzw. die Zufriedenheit mit der Tätigkeit führen. Da die Modelle 1 und 2 auf sich überschneidenden Stichproben beruhen, werden sie als *seemingly unrelated regressions* (siehe Greene 2003: 340 ff.) durchgeführt und anschließend werden mittels eines Wald-Tests jeweils für beide Fragestellungen die Koeffizienten des Modells 1 mit denen des Modells 2 verglichen.

Die Tabelle 2 und Tabelle 3 weisen die Ergebnisse der beiden Analysen aus, jeweils differenziert nach Modell 1 und Modell 2. Betrachtet man die Regressoren im Einzelnen, dann zeigt sich in der Analyse zur Arbeitszeit bei drei Regressoren eine unterschiedliche Schlussfolgerung zur Signifikanz (bei einem Signifikanzniveau 0,05). Beim Branchendummy „Callcenter“ wird der p-Wert von 0,05 im zweiten Modell überschritten und ebenso wird der Koeffizienten der „Kinderzahl“ nun im Gegensatz zum Referenzmodell 1 nicht mehr als signifikant ausgewiesen.

---

<sup>13</sup> Dazu muss auf die Zuspiegelung der Prozessdaten verzichtet werden, da das Ersetzen von Variablen aus der Befragung durch solche aus den Prozessdaten bereits durch deren Unterschiedlichkeit (vergleiche Abschnitt 2) eventuell zu Abweichungen beider Modelle führen könnte.

Tabelle 2: Lineare Regression der Arbeitszeit

<i>Analyse 1: Arbeitszeit</i>	Modell 1: vollständige Stichprobe		Modell 2: eingeschränkte Stichprobe	
	Coef.	p	Coef.	p
<b>Alter</b>	0.053	0.74	-0.007	0.97
<b>Alter (quadr.)</b>	-0.001	0.66	0.000	1.00
<b>Stellung im Beruf</b>				
Arbeiter				
Angestellter	-1.809	*** 0.00	-2.039	*** 0.00
sonst.	-1.204	0.37	-1.918	0.20
<b>Branche</b>				
Leih-/Zeitarbeitsfirma				
Reinigungsgewerbe	-6,597	*** 0,00	-5,900	*** 0,00
Hotel- und Gaststättengewerbe	-0,687	0,41	-0,749	0,43
Callcenter	-2,315	** 0,01	-1,570	0,13
Energie- und Wasser; Bergbau, Verarb. Gewerbe	0,335	0,58	0,184	0,78
Handel	-2,558	*** 0,00	-2,766	*** 0,00
Verkehr/Nachrichtenüberm., Banken, Versicherungen	0,344	0,80	0,705	0,63
Öffentlicher Dienst/Sozialversicherung	-4,395	*** 0,00	-4,554	*** 0,00
Bereich andere Dienstleistungen, k.A.	-1,792	*** 0,00	-1,681	*** 0,01
<b>Geschlecht, Lebensform und Partnererwerbstätigkeit</b>				
Frau, allein lebend				
Mann, allein lebend	4,383	*** 0,00	4,139	*** 0,00
Frau, allein erziehend	-1,325	* 0,05	-1,738	** 0,03
Mann, allein erziehend	7,221	*** 0,00	7,284	*** 0,00
Frau, Partner nicht erwerbstätig	-0,516	0,51	-0,604	0,51
Frau, Partner erwerbstätig	-0,350	0,65	-0,933	0,30
Mann, Partner nicht erwerbstätig	10,143	*** 0,00	10,045	*** 0,00
Mann, Partner erwerbstätig	10,329	*** 0,00	10,201	*** 0,00
<b>Zahl der Kinder (unter 18 Jahren im Haushalt)</b>	-0,469	** 0,03	-0,357	0,13
<b>Nebeneinkommen vorhanden</b>	-3,212	*** 0,00	-3,373	*** 0,00
<b>Erfolg im Beruf sehr wichtig</b>	-0,356	0,37	-0,357	0,42
<b>Einkommen sehr wichtig</b>	0,460	0,20	0,309	0,45
<b>Arbeit sehr wichtig</b>	1,012	*** 0,00	1,000	** 0,01
<b>Freizeit sehr wichtig</b>	0,013	0,97	0,617	0,17
<b>Familie sehr wichtig</b>	-0,742	* 0,10	-1,090	** 0,03
<b>Region: Ost</b>	5,343	*** 0,00	5,241	*** 0,00
<b>Konstante</b>	29,143	*** 0,00	30,345	*** 0,00
<b>N</b>	3408		2638	
<b>R<sup>2</sup></b>	0,3056		0,3211	
Wald-Test auf Gleichheit aller Koeffizienten in Modell 1 und 2			chi <sup>2</sup> ( 27) = 30,58 Prob > chi <sup>2</sup> = 0,2886	
Legende: *p<0,1; **p<0,05; ***p<0,01				

Auch in der zweiten Analyse zur Zufriedenheit (Tabelle 3) treten einige Unterschiede zwischen Modell 1 und 2 auf. So ist der Einfluss des Stundenlohns auf die Zufriedenheit mit der Tätigkeit nur in Modell 1 signifikant ebenso wie der Einfluss einiger Branchen und der Einfluss des vorherigen Bezuges von Arbeitslosengeld im Vergleich zu Arbeitslosenhilfe. Umgekehrt werden nun in Modell 2 vorher insignifikante Effekte als signifikant ausgewiesen. Dies sind der Effekt eines Volks- oder Hauptschulabschlusses im Vergleich zu keinem Schulabschluss und der Einfluss der Befristung des Arbeitsvertrages auf die Zufriedenheit.

**Tabelle 3: Logistische Regression der Zufriedenheit mit der Tätigkeit**

Analyse 2: Zufriedenheit mit der Tätigkeit	Modell 1: vollständige Stichprobe		Modell 2: eingeschränkte Stichprobe	
	Odds Ratio	p	Odds Ratio	p
<b>Stundenlohn</b>	1,039 *	0,08	1,040	0,12
<b>Stundenlohn (quadr.)</b>	1,000	0,12	1,000	0,16
<b>Förderung Mainzer Modell</b>				
Geförderte (qualifiziert und nicht langzeitarbeitslos)				
Geförderte (gering qualifiziert u/o langzeitarbeitslos)	0,759 **	0,04	0,724 **	0,04
Vergleichsgruppe (gering qualifizierte u/o langzeitarbeitslos)	0,892	0,54	0,854	0,47
<b>Zufriedenheit mit Verdienst</b>	3,935 ***	0,00	4,513 ***	0,00
<b>Alter</b>				
Bis unter 25 Jahre				
25 bis 34 Jahre	0,854	0,47	0,997	0,99
35 bis 44 Jahre	0,989	0,97	1,018	0,95
45 bis 54 Jahre	0,826	0,54	0,763	0,46
55 Jahre oder mehr	3,080 *	0,05	3,686 *	0,06
<b>Bildung</b>				
kein Abschluss				
Volks-, Hauptschule	0,706	0,17	0,545 **	0,05
Mittl. Reife, POS 10.Klasse	0,538 **	0,02	0,423 ***	0,01
Fachhochschulreife, Abitur	0,408 ***	0,00	0,370 ***	0,00
<b>Staatsangehörigkeit deutsch</b>	1,651 **	0,01	2,013 ***	0,00
<b>Erwerbsstatus vor Referenzzeitpunkt</b>				
Erwerbstätig				
Anderes	0,956	0,81	0,956	0,83
<b>Stellung im Beruf</b>				
Arbeiter				
Angestellter	1,018	0,89	1,040	0,80
sonst.	0,657	0,20	0,563	0,12
<b>Befristung</b>				
Nicht befristet / trifft nicht zu				
befristet	1,188	0,13	1,260 *	0,08

<b>Ausbildung</b>				
Keine				
Ausbildung/Studium	1,051	0,74	1,000	1,00
<b>Berufserfahrung</b>	0,987	0,21	0,995	0,69
<b>Branche</b>				
Leih-/Zeitarbeitsfirma				
Reinigungsgewerbe	1,032	0,90	0,903	0,71
Hotel- und Gaststättengewerbe	1,253	0,33	1,123	0,66
Callcenter	1,253	0,48	1,309	0,49
Energie- und Wasser; Bergbau, Verarb. Gewerbe	1,340	0,16	1,183	0,48
Handel	1,415 *	0,08	1,279	0,28
Verkehr/Nachrichtenüberm., Ban- ken, Versicherungen	2,249 **	0,04	2,403 *	0,07
Öffentlicher Dienst/ Sozialversi- cherung	1,873 **	0,04	1,172	0,63
Andere Dienstleistungen, k.A.	1,685 ***	0,00	1,382	0,11
<b>Geschlecht</b>	1,405 **	0,03	1,678 ***	0,00
<b>Lebensform und Partnererwerbstätigkeit</b>				
allein lebend				
allein erziehend	1,170	0,45	1,098	0,71
Partner, nicht erwerbstätig	1,085	0,68	1,018	0,94
Partner, erwerbstätig	1,299	0,20	1,289	0,29
<b>Zahl der Kinder (unter 18 Jah- ren im Haushalt)</b>	0,985	0,80	0,997	0,97
<b>Lohnersatzleistungen</b>				
Arbeitslosenhilfe				
Arbeitslosengeld	1,380 **	0,05	1,342	0,12
keine Leistungen	0,920	0,61	0,969	0,87
Sozialhilfe (auch ergänzend)	0,967	0,84	0,929	0,70
<b>Nebeneinkommen vorhanden</b>	1,275	0,48	0,974	0,94
<b>Arbeit sehr wichtig</b>	1,062	0,61	1,003	0,98
<b>Einkommen sehr wichtig</b>	1,044	0,72	1,127	0,38
<b>Beschäftigung ist Übergangslö- sung</b>	0,293 ***	0,00	0,269 ***	0,00
<b>Region Ost</b>	1,325 *	0,05	1,355 *	0,07
<b>Konstante</b>	2,681 *	0,03	2,519 *	0,08
<b>N</b>	2781		2164	
<b>Pseudo R<sup>2</sup></b>	0,1732		0,1930	
<b>Log likelihood</b>	-1073,0		-818,9	
Wald-Test auf Gleichheit aller Koeffizienten in Modell 1 und 2		chi <sup>2</sup> ( 40) = 37,42 Prob > chi <sup>2</sup> = 0,5870		
Legende: *p<0,1;**p<0,05;***p<0,01				

Wie der Vergleich der Modelle 1 und 2 zeigt, gäbe es bei der Betrachtung der Modelle ohne Zustimmung zum Teil abweichende Interpretationen der Einflussfaktoren auf die Arbeitszeit und die Zufriedenheit mit der Tätigkeit. Ob sich die Modelle 1 und 2 tatsächlich, das heißt statistisch signifikant unterscheiden, oder ob Unterschiede in Größenordnung und Signifikanz

der Koeffizienten eher auf den Fallzahlverlust bzw. zufällige Abweichungen zurückgehen, zeigt ein Test auf Gleichheit der Koeffizienten. In beiden Analysen führt der Wald-Test der Hypothese der Gleichheit aller Koeffizienten des Modells 2 mit denen in Modell 1 nicht zum Verwerfen dieser Hypothese. Daher gibt es keinen Grund anzunehmen, dass sich in den beiden beispielhaften Analysen durch die Beschränkung auf die Befragten ohne Datenschutzbedenken systematische Veränderungen der Schlussfolgerungen ergeben würden. Wie zu sehen war, bedeutet dies nicht, dass im Einzelfall ein Forscher nicht zu unterschiedlichen Schlussfolgerungen kommen würde. Die Unterschiede entstehen aber – zumindest in den hier betrachteten Anwendungsbeispielen – durch die Orientierung an einem relativ willkürlichen Schwellenwert des p-Wertes. So weichen denn auch die p-Werte zwischen den in Modell 1 signifikanten und in Modell 2 insignifikanten Merkmalen (bzw. umgekehrt) kaum voneinander ab.

## 5 Zusammenfassung und Diskussion

Das Zusammenspielen von Prozess- und Befragungsdaten kann eine Reihe von Vorteilen für die Forschung haben. Diese Vorteile können allerdings nicht zum Tragen kommen, wenn es aufgrund selektiver Zustimmung zum Zusammenspielen zu verzerrten Stichproben kommt. Eine Beschränkung der Analysegesamtheit auf die Fälle, die einer Verknüpfung zustimmen, ließe dann andere Ergebnisse erwarten als eine Einbeziehung aller Fälle in die Analysen. Der vorliegende Beitrag untersucht dieses Problem unter Verwendung der Daten einer Untersuchung zur Wirkung des Mainzer Modells.

Es wurde zunächst analysiert, wovon die Zustimmungsbereitschaft abhängt. Dabei zeigte sich, dass insbesondere Merkmale, die auf positive Erfahrungen mit bzw. Einstellungen gegenüber staatlichen Institutionen und der Arbeitsmarktpolitik hindeuten, positiv mit dem Zustimmungsverhalten zusammenhängen. So ist die Einwilligung beispielsweise positiv korreliert mit dem Bezug staatlicher Transferleistungen oder mit der Zufriedenheit mit der staatlich finanzierten Förderung der eigenen Erwerbstätigkeit, negativ dagegen mit der Erfahrung von Arbeitslosigkeit. Andererseits ist für eine Merkmalsausprägung, die auf fehlendes Vertrauen hinweist, und zwar die Verweigerung der Einkommensangabe, ein negativer Zusammenhang mit der Zustimmungsbereitschaft festzustellen. Wei-



terhin weisen gut verdienende, deutsche und männliche Zielpersonen sowie solche, die einen Hauptschulabschluss haben bzw. die in Ostdeutschland leben, eine erhöhte Wahrscheinlichkeit der Zustimmung auf. Insofern, als eigentlich nur in Bezug auf die Staatsangehörigkeit ein Effekt festzustellen ist, der auch für das Ausfallgeschehen gilt, weisen die Analysen darauf hin, dass bezüglich der Zustimmung zur Datenverknüpfung andere Merkmale eine Rolle spielen als hinsichtlich der Teilnahmeentscheidung an Befragungen. Ausfallbedingte Selektivitäten werden damit nicht verstärkt. Anzumerken ist dabei allerdings, dass die theoretisch identifizierten Einflussfaktoren in dem Modell vor allem über Proxy-Indikatoren kontrolliert wurden. Daher kann nicht ausgeschlossen werden, dass theoretisch bestimmte Mechanismen auch deshalb nicht bestätigt werden konnten, weil keine geeigneten Indikatoren vorlagen.

Ergänzend wurden die Ergebnisse zweier beispielhaft durchgeführter Analysen verglichen, die einmal mit allen Personen und einmal ohne diejenigen durchgeführt wurden, die ihre Zustimmung zur Verknüpfung verweigert haben. Es hat sich gezeigt, dass die Nichtberücksichtigung von Personen mit Datenschutzbedenken zumindest in den beiden ausgewählten Analysen keinen systematischen Einfluss auf die Analyseergebnisse hat. Obwohl sich die Ergebnisse nicht unbedingt dahingehend verallgemeinern lassen, dass das Zusammenspielen von Prozess- und Befragungsdaten nie zu selektiven Ausfällen führen wird, sprechen sie doch dafür, dass sich das Problem zumindest in Grenzen hält. Weitere Untersuchungen, eventuell mit eigens zur Untersuchung der Zustimmungselektivität konzipierten Befragungen, wären nötig, um zu Aussagen mit höherem Allgemeingrad zu gelangen.

## Literatur

- Achatz, Juliane; Gartner, Hermann; Glück, Timea (2005): Bonus oder Bias? Mechanismen geschlechtsspezifischer Entlohnung. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie, Jg. 57, H. 3. S. 466-493.
- Blau, Peter M. (1968): Social Exchange. In: David L. Sills (Hg.): International Encyclopedia of the Social Sciences, Vol. 7. New York London: S. 452-457.
- Eichhorst, Werner; Gartner, Hermann; Krug, Gerhard; Rhein, Thomas; Wiedemann, Eberhard (2005): Niedriglohnbeschäftigung in Deutschland und im internationalen Vergleich. In: J. Allmendinger, Jutta; Eichhorst, Werner; Walwei, Ulrich (Hrsg.): IAB Handbuch Arbeitsmarkt : Analysen, Daten, Fakten, Frankfurt am Main u. a. : Campus Verlag S. 107-142.
- Engel, Uwe; Plötschke, Manuela; Schnabel, Christiane; Simonson, Julia (2004): Nonresponse und Stichprobenqualität. Ausschöpfung in Umfragen der Markt- und Sozialforschung. Bremen.
- Esser, Hartmut (1986): Über die Teilnahme an Befragungen. ZUMA-Nachrichten 18: S. 38-47.
- Esser, Hartmut (1990): „Habits“, „Frames“ und „Rational Choice“. Die Reichweite der Theorie der rationalen Wahl (am Beispiel der Erklärung des Befragtenverhaltens). Zeitschrift für Soziologie 19: S. 231-247.
- Fetchenhauer, Detlef (1999): Möglichkeiten und Grenzen von Rational Choice – Erklärungen für betrügerisches Verhalten am Beispiel des Versicherungsbetruges. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie, Jg. 51, H. 2. S. 283-312.
- Gewiese, Tilo, Hartmann, Josef, Krug, Gerhard, Rudolph, Helmut (2004): Das Mainzer Modell aus der Sicht der Arbeitnehmer und Betriebe. Befunde aus der Begleitforschung. Download unter: <http://doku.iab.de/externe/2004/k040823w09.pdf>
- Greene, William H. (2003): Econometric Analysis. Upper Saddle River, New Jersey: Pearsons Education, Inc.
- Groves, Robert M.; Floyd J. Fowler, Jr. et al. (2004): Survey Methodology. Hoboken, New Jersey: John Wiley & Sons.
- Guo, Guang; Zhao, Hongxin (2000): Multilevel Modelling for Binary Data. In: Annual Review of Sociology 26: S. 441-462.
- Hakim, Cathrine (1983): Research based on administrative records. In: The Sociological Review, New Series, Vol. 31, S. 489-519.
- Hamann, Silke; Krug, Gerhard; Köhler, Markus; Ludwig-Mayerhofer, Wolfgang; Hackett, Anne (2004): Die IAB-Regionalstichprobe 1975-2001: IABS-RO1. In: ZA-Information, Nr. 55. S. 34-59.

- Heckman, James J.; LaLonde, Robert J.; Smith Jeffrey A. (1999): The Economics and Econometrics of Active Labor Market Programs. In: Ashenfelter, Card (Hrsg.): Handbook of Labor Economics.
- Heckman, James J. (1979): Sample Selection as Specification Error. In: *Econometrica* 47: S. 153-162.
- Holland, Paul W. (1986): Statistics and Causal Inference. In: *Journal of the American Statistical Association* Vol. 81, Nr. 396, S. 945-960.
- Homans, George Casper (1964): Bringing Man Back In: *American Sociological Review*, Bd. 29, S. 809-818.
- Hummel, Elisabeth; Jacobebbinghaus, Peter; Kohlmann, Annette; Oertel, Martina; Wübbecke, Christina; Ziegerer, Manfred (2005): Stichprobe der Integrierten Erwerbsbiografien: IEBS 1.0, Handbuch-Version 1.0.0. FDZ Datenreport Nr. 06/2005.
- Jenkins, Stephen P.; Lynn, Peter; Jäckle, Annette; Sala, Emanuela (2005): Linking household survey and administrative record data: what should the matching variables be? *DIW Discussion Papers* 489.
- Kaltenborn, Bruno; Krug, Gerhard; Rudolph, Helmut; Weinkopf, Claudia; Wiedemann, Eberhard (2005): Evaluierung der arbeitsmarktpolitischen Sonderprogramme CAST und Mainzer Modell. Bundesministerium für Wirtschaft und Arbeit. Forschungsbericht Nr. 552, Berlin.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (2001): Wege zu einer besseren informationellen Infrastruktur. Gutachten der vom Bundesministerium für Bildung und Forschung eingesetzten Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. Baden-Baden: Nomos.
- Pascale, Joanne; Thomas S. Mayer (2004): Exploring Confidentiality Issues Related to Dependent Interviewing: Preliminary Findings. In: *Journal of Official Statistics* 20: S. 357-377.
- Pickery, Jan; Losenveldt, Geert; Carton, Ann (2001): The Effects of Interviewer and Respondent Characteristics on Response Behavior in Panel Surveys. In: *Sociological Methods & Research* 29: S. 509-523.
- Riker, William H.; Ordeshook, Peter C. (1973): *An Introduction to Positive Political Theory*. Englewood Cliffs, N. J.: Prentice-Hall.
- Schmähl, Winfried; Fachinger, Uwe (1994): Prozessproduzierte Daten als Grundlage für sozial- und verteilungspolitische Analyse – Erfahrungen mit Daten der Rentenversicherungsträger für Längsschnittanalysen. In: Deutsche Forschungsgemeinschaft (Hrsg.): *Mikroanalytische Grundlagen der Gesellschaftspolitik*. Band 2, S. 179-200.

- Schnabel, Annette (1997): Teilnahmeverhalten bei Unternehmensbefragungen. *Arbeit*, Jg. 6, Heft 2, S. 154-172.
- Schnell, Rainer (1997): Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Umfang. Opladen: Leske + Budrich.
- Schräpler, Jörg-Peter (2004): Respondent Behavior in Panel Studies. A Case Study for Income Nonresponse by Means of the German Socio-Economic Panel (SOEP). In: *Sociological Methods & Research* 33: S. 118-156.
- Singer, Eleanor; Mathiowetz, Nancy; Couper, Mick (1993): The Impact of Privacy and Confidentiality Concerns on Survey Participation: The Case of the 1990 U.S. Census. In: *Public Opinion Quarterly* 57: S. 465-482.
- Tomaskovic-Devey, Donald; Leiter, Jeffrey; Thompson, Shealy (1995): Item Nonresponse in Organizational Surveys. In: *Sociological Methodology* 25, S. 77-110.
- Windzio, Michael (2004): Kann der regionale Kontext zur „Arbeitslosenfalle“ werden? In: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56: S. 257-278
- Wippler, Reinhard (1987): Kulturelle Ressourcen, gesellschaftlicher Erfolg und Lebensqualität. In: Bernhard Giesen, Hans Haferkamp (Hrsg.): *Soziologie der sozialen Ungleichheit*. Opladen: Westdeutscher Verlag. S. 221–254.
- Wirth, Heike; Müller, Walter (2004): Mikrodaten der amtlichen Statistik als eine Datengrundlage der empirischen Sozialforschung. In: Andreas Diekmann (Hrsg.): *Methoden der Sozialforschung (Sonderheft 44 der Kölner Zeitschrift für Soziologie und Sozialpsychologie)*. Wiesbaden: VS Verlag für Sozialwissenschaften. S. 93-127

## Anhang

### Zustimmung in Abhängigkeit von verschiedenen Merkmalen

	Mit Zusammenspielen einverstanden	Anteil
<b>Interviewer:</b>		
Geschlecht	Frau	74,6%
	Mann	73,8%
Alter:	Bis unter 25 Jahre	73,1%
	25 bis 34 Jahre	72,0%
	35 bis 44 Jahre	72,8%
	45 bis 54 Jahre	75,6%
	55 Jahre oder mehr	83,7%
	Sonst	73,4%
Bildung	FHR, Abitur	75,3%
<b>Befragte:</b>		
Alter	Bis unter 25 Jahre	72,9%
	25 bis 34 Jahre	76,1%
	35 bis 44 Jahre	75,4%
	45 bis 54 Jahre	72,8%
	55 Jahre oder mehr	67,9%
	Angabe fehlt	50,0%
Deutsche Staatsangehörigkeit	Nicht deutsch	69,6%
	Deutsch	74,8%
Bildung	Kein Hauptschulabschluss	70,8%
	Volks-/Hauptschule	75,2%
	Mittlere Reife/POS10.Klasse	75,0%
	Abi/FHS	72,0%
	Keine/sonstige Ausbildung, keine Angabe	
Berufsausbildung		73,9%
Arbeitslosigkeit vor Referenzzeitpunkt	Ausbildung/Studium	74,6%
	Nicht arbeitslos	80,3%
	Bis unter 12 Monate	73,9%
	12 Monate oder mehr	71,4%
Erwerbsstatus vor Referenzzeitpunkt	Angaben fehlen	82,0%
	Erwerbstätig	75,8%
	Anderes	74,2%
Zufriedenheit mit Förderhöhe	Nicht kombilohngefördert	69,9%
	Weniger zufrieden, keine Angabe	69,3%
	Sehr/zufrieden	78,8%
Bruttoerwerbseinkommen Befr. nach Referenzzeitpunkt	Nicht vorhanden	73,3%
	1 - 325 EURO	78,5%
	326 - 511 EURO	74,3%
	512 - 1023 EURO	76,8%
	1024 EURO oder mehr	79,2%
	Angabe fehlt	51,9%

	<b>Mit Zusammenspielen einverstanden</b>	<b>Anteil</b>
HLU vor/nach Referenzzeitpunkt	Weder-noch	73,1%
	Vor und nach	82,7%
	Nur nach	78,6%
	Nur vor	79,8%
	Mindestens einmal keine Angabe	36,2%
Geschlecht	Frau	73,7%
	Mann	75,7%
Lebensform	Ohne Partner	76,3%
	Mit Partner	73,6%
Erwerbstätigkeit befragte Person	Nein	69,4%
	Ja	75,4%
Öffentlicher Dienst	Nein	74,3%
	Ja	75,8%
Erwerbstätigkeit Partner/in	Nein	74,6%
	Ja	73,3%
Region	West	73,8%
	Ost	75,4%

## Recently published

No.	Author(s)	Title	Date
<a href="#">1/2004</a>	Bauer, T. K. Bender, S. Bonin, H.	Dismissal protection and worker flows in small establishments <a href="#">published in: <i>Economica</i></a>	7/04
<a href="#">2/2004</a>	Achatz, J. Gartner, H. Glück, T.	Bonus oder Bias? : Mechanismen geschlechtsspezifischer Entlohnung <a href="#">published in: <i>Kölner Zeitschrift für Soziologie und Sozialpsychologie</i> 57 (2005), S. 466-493 (revised)</a>	7/04
<a href="#">3/2004</a>	Andrews, M. Schank, T. Upward, R.	Practical estimation methods for linked employer-employee data	8/04
<a href="#">4/2004</a>	Brixy, U. Kohaut, S. Schnabel, C.	Do newly founded firms pay lower wages? : first evidence from Germany <a href="#">published in: <i>Small Business Economics</i>, (2007)</a>	9/04
<a href="#">5/2004</a>	Kölling, A. Rässler, S.	Editing and multiply imputing German establishment panel data to estimate stochastic production frontier models <a href="#">published in: <i>Zeitschrift für ArbeitsmarktForschung</i> 37 (2004), S. 306-318</a>	10/04
<a href="#">6/2004</a>	Stephan, G. Gerlach, K.	Collective contracts, wages and wage dispersion in a multi-level model <a href="#">published as: <i>Wage settlements and wage setting : results from a multi-level model</i>. In: <i>Applied Economics</i>, Vol. 37, No. 20 (2005), S. 2297-2306</a>	10/04
<a href="#">7/2004</a>	Gartner, H. Stephan, G.	How collective contracts and works councils reduce the gender wage gap	12/04
<a href="#">1/2005</a>	Blien, U. Suedekum, J.	Local economic structure and industry development in Germany, 1993-2001	1/05
<a href="#">2/2005</a>	Brixy, U. Kohaut, S. Schnabel, C.	How fast do newly founded firms mature? : empirical analyses on job quality in start-ups <a href="#">published in: <i>Michael Fritsch, Jürgen Schmude (Ed.): Entrepreneurship in the region</i>, New York et al., 2006, S. 95-112</a>	1/05
<a href="#">3/2005</a>	Lechner, M. Miquel, R. Wunsch, C.	Long-run effects of public sector sponsored training in West Germany	1/05
<a href="#">4/2005</a>	Hinz, T. Gartner, H.	Lohnunterschiede zwischen Frauen und Männern in Branchen, Berufen und Betrieben <a href="#">published in: <i>Zeitschrift für Soziologie</i> 34 (2005), S. 22-39, as: <i>Geschlechtsspezifische Lohnunterschiede in Branchen, Berufen und Betrieben</i></a>	2/05
<a href="#">5/2005</a>	Gartner, H. Rässler, S.	Analyzing the changing gender wage gap based on multiply imputed right censored wages	2/05
<a href="#">6/2005</a>	Alda, H. Bender, S. Gartner, H.	The linked employer-employee dataset of the IAB (LIAB) <a href="#">published as: <i>The linked employer-employee dataset created from the IAB establishment panel and the process-produced data of the IAB (LIAB)</i>. In: <i>Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften</i> 125 (2005), S. 327-336 (shortened)</a>	3/05
<a href="#">7/2005</a>	Haas, A. Rothe, T.	Labour market dynamics from a regional perspective : the multi-account system	4/05

<a href="#">8/2005</a>	Caliendo, M. Hujer, R. Thomsen, S. L.	Identifying effect heterogeneity to improve the efficiency of job creation schemes in Germany	4/05
<a href="#">9/2005</a>	Gerlach, K. Stephan, G.	Wage distributions by wage-setting regime <a href="#">published as: Bargaining regimes and wage dispersion. In: Jahrbücher für Nationalökonomie und Statistik, Bd. 226, H. 6 (2006)</a>	4/05
<a href="#">10/2005</a>	Gerlach, K. Stephan, G.	Individual tenure and collective contracts	4/05
<a href="#">11/2005</a>	Blien, U. Hirschenauer, F.	Formula allocation : the regional allocation of budgetary funds for measures of active labour market policy in Germany <a href="#">published in: Economics Bulletin, Vol. 18, no. 7 (2006)</a>	4/05
<a href="#">12/2005</a>	Alda, H. Allaart, P. Bellmann, L.	Churning and institutions : Dutch and German establishments compared with micro-level data	5/05
<a href="#">13/2005</a>	Caliendo, M. Hujer, R. Thomsen, S. L.	Individual employment effects of job creation schemes in Germany with respect to sectoral heterogeneity	5/05
<a href="#">14/2005</a>	Lechner, M. Miquel, R. Wunsch, C.	The curse and blessing of training the unemployed in a changing economy : the case of East Germany after unification	6/05
<a href="#">15/2005</a>	Jensen, U. Rässler, S.	Where have all the data gone? : stochastic production frontiers with multiply imputed German establishment data <a href="#">published in: Zeitschrift für ArbeitsmarktForschung, Jg. 39, H. 2, 2006, S. 277-295</a>	7/05
<a href="#">16/2005</a>	Schnabel, C. Zagelmeyer, S. Kohaut, S.	Collective bargaining structure and its determinants : an empirical analysis with British and German establishment data <a href="#">published in: European Journal of Industrial Relations, Vol. 12, No. 2, S. 165-188</a>	8/05
<a href="#">17/2005</a>	Koch, S. Stephan, G. Walwei, U.	Workfare: Möglichkeiten und Grenzen <a href="#">published in: Zeitschrift für ArbeitsmarktForschung 38 (2005), S. 419-440</a>	8/05
<a href="#">18/2005</a>	Alda, H. Bellmann, L. Gartner, H.	Wage structure and labour mobility in the West German private sector 1993-2000	8/05
<a href="#">19/2005</a>	Eichhorst, W. Konle-Seidl, R.	The interaction of labor market regulation and labor market policies in welfare state reform	9/05
<a href="#">20/2005</a>	Gerlach, K. Stephan, G.	Tarifverträge und betriebliche Entlohnungsstrukturen <a href="#">published in: C. Clemens, M. Heinemann &amp; S. Soretz (Hg.): Auf allen Märkten zu Hause, Marburg 2006, S. 123-143</a>	11/05
<a href="#">21/2005</a>	Fitzenberger, B. Speckesser, S.	Employment effects of the provision of specific professional skills and techniques in Germany	11/05
<a href="#">22/2005</a>	Ludsteck, J. Jacobebbinghaus, P.	Strike activity and centralisation in wage setting	12/05
<a href="#">1/2006</a>	Gerlach, K. Levine, D. Stephan, G. Struck, O.	The acceptability of layoffs and pay cuts : comparing North America with Germany	1/06
<a href="#">2/2006</a>	Ludsteck, J.	Employment effects of centralization in wage setting in a median voter model	2/06
<a href="#">3/2006</a>	Gaggermeier, C.	Pension and children : Pareto improvement with heterogeneous preferences	2/06
<a href="#">4/2006</a>	Binder, J. Schwengler, B.	Korrekturverfahren zur Berechnung der Einkommen über der Beitragsbemessungsgrenze	3/06



<a href="#">5/2006</a>	Brixy, U. Grotz, R.	Regional patterns and determinants of new firm formation and survival in western Germany	4/06
<a href="#">6/2006</a>	Blien, U. Sanner, H.	Structural change and regional employment dynamics	4/06
<a href="#">7/2006</a>	Stephan, G. Rässler, S. Schewe, T.	Wirkungsanalyse in der Bundesagentur für Arbeit : Konzeption, Datenbasis und ausgewählte Befunde <a href="#">published as: Das TrEffeR-Projekt der Bundesagentur für Arbeit : die Wirkung von Maßnahmen aktiver Arbeitsmarktpolitik. In: Zeitschrift für ArbeitsmarktForschung, Jg. 39, H. 3/4 (2006)</a>	4/06
<a href="#">8/2006</a>	Gash, V. Mertens, A. Romeu Gordo, L.	Are fixed-term jobs bad for your health? : a comparison of West-Germany and Spain <a href="#">published in: European Societies, 2007</a>	5/06
<a href="#">9/2006</a>	Romeu Gordo, L.	Compression of morbidity and the labor supply of older people	5/06
<a href="#">10/2006</a>	Jahn, E. J. Wagner, T.	Base period, qualifying period and the equilibrium rate of unemployment	6/06
<a href="#">11/2006</a>	Jensen, U. Gartner, H. Rässler, S.	Measuring overeducation with earnings frontiers and multiply imputed censored income data	6/06
<a href="#">12/2006</a>	Meyer, B. Lutz, C. Schnur, P. Zika, G.	National economic policy simulations with global interdependencies : a sensitivity analysis for Germany	7/06
<a href="#">13/2006</a>	Beblo, M. Bender, S. Wolf, E.	The wage effects of entering motherhood : a within-firm matching approach	8/06
<a href="#">14/2006</a>	Niebuhr, A.	Migration and innovation : does cultural diversity matter for regional R&D activity?	8/06
<a href="#">15/2006</a>	Kiesl, H. Rässler, S.	How valid can data fusion be? <a href="#">published in: Journal of Official Statistics, (2006)</a>	8/06
<a href="#">16/2006</a>	Hujer, R. Zeiss, C.	The effects of job creation schemes on the unemployment duration in East Germany	8/06
<a href="#">17/2006</a>	Fitzenberger, B. Osikominu, A. Völter, R.	Get training or wait? : long-run employment effects of training programs for the unemployed in West Germany	9/06
<a href="#">18/2006</a>	Antoni, M. Jahn, E. J.	Do changes in regulation affect employment duration in temporary work agencies?	9/06
<a href="#">19/2006</a>	Fuchs, J. Söhnlein, D.	Effekte alternativer Annahmen auf die prognostizierte Erwerbsbevölkerung	10/06
<a href="#">20/2006</a>	Lechner, M. Wunsch, C.	Active labour market policy in East Germany : waiting for the economy to take off	11/06
<a href="#">21/2006</a>	Kruppe, T.	Die Förderung beruflicher Weiterbildung : eine mikroökonomische Evaluation der Ergänzung durch das ESF-BA-Programm	11/06
<a href="#">22/2006</a>	Feil, M. Klinger, S. Zika, G.	Sozialabgaben und Beschäftigung : Simulationen mit drei makroökonomischen Modellen	11/06
<a href="#">23/2006</a>	Blien, U. Phan, t. H. V.	A pilot study on the Vietnamese labour market and its social and economic context	11/06
<a href="#">24/2006</a>	Lutz, R.	Was spricht eigentlich gegen eine private Arbeitslosenversicherung?	11/06

<a href="#">25/2006</a>	Jirjahn, U. Pfeifer, C. Tsertsvadze, G.	Mikroökonomische Beschäftigungseffekte des Hamburger Modells zur Beschäftigungsförderung	11/06
<a href="#">26/2006</a>	Rudolph, H.	Indikator gesteuerte Verteilung von Eingliederungsmitteln im SGB II : Erfolgs- und Effizienzkriterien als Leistungsanreiz?	12/06
<a href="#">27/2006</a>	Wolff, J.	How does experience and job mobility determine wage gain in a transition and a non-transition economy? : the case of east and west Germany	12/06
<a href="#">28/2006</a>	Blien, U. Kirchhof, K. Ludewig, O.	Agglomeration effects on labour demand	12/06
<a href="#">29/2006</a>	Blien, U. Hirschenauer, F. Phan, t. H. V.	Model-based classification of regional labour markets : for purposes of labour market policy	12/06
<a href="#">30/2006</a>	Krug, G.	Kombilohn und Reziprozität in Beschäftigungsverhältnissen : eine Analyse im Rahmen des Matching-Ansatzes	12/06
<a href="#">1/2007</a>	Moritz, M. Gröger, M.	The German-Czech border region after the fall of the Iron Curtain: Effects on the labour market : an empirical study using the IAB Employment Sample (IABS)	1/07
<a href="#">2/2007</a>	Hampel, K. Kunz, M. Schanne, N. Wapler, R. Weyh, A.	Regional employment forecasts with spatial interdependencies	1/07
<a href="#">3/2007</a>	Eckey, H.- F. Schwengler, B. Türck, M.	Vergleich von deutschen Arbeitsmarktregionen	1/07
<a href="#">4/2007</a>	Kristen, C. Granato, N.	The educational attainment of the second generation in Germany : social origins and ethnic inequality	1/07
<a href="#">5/2007</a>	Jacob, M. Kleinert, C.	Does unemployment help or hinder becoming independent? : the role of employment status for leaving the parental home	1/07
<a href="#">6/2007</a>	Konle-Seidl, R. Eichhorst, W. Grienberger-Zingerle, M.	Activation policies in Germany : from status protection to basic income support	1/07
<a href="#">7/2007</a>	Lechner, M. Wunsch, C.	Are training programs more effective when unemployment is high?	2/07
<a href="#">8/2007</a>	Hohendanner, C.	Verdrängen Ein-Euro-Jobs sozialversicherungspflichtige Beschäftigung in den Betrieben?	2/07
<a href="#">9/2007</a>	Seibert, H.	Frühe Flexibilisierung? : regionale Mobilität nach der Lehrausbildung in Deutschland zwischen 1977 und 2004	2/07
<a href="#">10/2007</a>	Bernhard, S. Kurz, K.	Familie und Arbeitsmarkt : eine Längsschnittstudie zum Einfluss beruflicher Unsicherheiten auf die Familienerweiterung	2/07
<a href="#">11/2007</a>	Drechsler, J. Dundler, A. Bender, S. Rässler, S. Zwick, T.	A new approach for disclosure control in the IAB Establishment Panel multiple imputation for a better data access	2/07
<a href="#">12/2007</a>	Fuchs, J. Söhnlein, D.	Bedeutung von Demografie und Erwerbsverhalten für das Erwerbspersonenpotenzial in Ost- und Westdeutschland	3/07

## Impressum

**IAB Discussion Paper**  
**No. 13 / 2007**

### Herausgeber

Institut für Arbeitsmarkt- und Berufsforschung  
der Bundesagentur für Arbeit  
Weddigenstr. 20-22  
D-90478 Nürnberg

### Redaktion

Regina Stoll, Jutta Palm-Nowak

### Technische Herstellung

Jutta Sebald, Claudia Halbmeier

### Rechte

Nachdruck – auch auszugsweise – nur mit  
Genehmigung des IAB gestattet

### Bezugsmöglichkeit

Volltext-Download dieses Discussion Paper  
unter:

<http://doku.iab.de/discussionpapers/2007/dp1307.pdf>

### IAB im Internet

<http://www.iab.de>

### Rückfragen zum Inhalt an

Gerhard Krug, Tel. 0911/179-3387,  
oder E-Mail: [gerhard.krug@iab.de](mailto:gerhard.krug@iab.de)