

# **Ankündigungseffekt oder Maßnahmewirkung? – Eine Evaluation von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit\***

*Thomas Büttner\*\**

Trainingsmaßnahmen sind ein Instrument der aktiven Arbeitsförderung, mit dem je nach Ausrichtung unterschiedliche Zielsetzungen verfolgt werden. In diesem Beitrag wird eine Variante der Trainingsmaßnahmen, die Maßnahme zur Überprüfung der Verfügbarkeit evaluiert. Diese Maßnahme zielt in erster Linie darauf ab, zu überprüfen, ob Arbeitslose tatsächlich zu einer Arbeitsaufnahme zur Verfügung stehen. Besonderes Augenmerk liegt hier deshalb auf der Fragestellung, wodurch Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit zur Verbesserung von Eingliederungsaussichten beitragen. Unterschieden werden dabei Eingliederungseffekte, die allein aus der Einladung zu einer Maßnahme resultieren, in Abgrenzung zu Effekten, die für die tatsächliche Teilnahme an einer Trainingsmaßnahme geschätzt werden. Zur Untersuchung dieser Fragestellung werden Daten aus einem arbeitsmarktpolitischen Experiment genutzt, die neben Angaben über Maßnahmeteilnahmen auch Informationen zu ausgesprochenen Einladungen liefern.

\* Der Beitrag wurde im März 2008 zur Publikation freigegeben.

\*\* Ich danke Ursula Jaenichen, Susanne Rässler und Gesine Stephan für ihre Unterstützung sowie zwei anonymen Gutachtern für hilfreiche Anmerkungen. Verbleibende Irrtümer gehen zu meinen Lasten.

## **Gliederung**

- 1 Einleitung
- 2 Ausgestaltung von Trainingsmaßnahmen
- 3 Wirkungsmessung und Experiment
- 4 Untersuchungsaufbau
- 5 Mögliche Effekte von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit
- 6 Strukturmerkmale der Teilnehmer- und Vergleichsgruppen
- 7 Ergebnisse
  - 7.1 Teilzeitarbeitslose, nur Frauen
  - 7.2 Vollzeit gewerblich
  - 7.3 Arbeitslose im Alter von 20 bis 27 Jahren
- 8 Fazit

Literatur

## 1 Einleitung

Trainingsmaßnahmen sind ein Instrument der aktiven Arbeitsförderung, das in sehr heterogener Weise zur Verbesserung der Eingliederungschancen von Arbeitsuchenden beiträgt. Es handelt sich dabei um ein Instrument mit relativ kurzer Dauer und daher mit vergleichsweise geringen Kosten. Seit seiner Aufnahme in das Regelinstrumentarium war deshalb ein enormer Anstieg der Förderzahlen zu verzeichnen.<sup>1</sup>

Trotz des großflächigen Einsatzes dieser Maßnahme wird ihr von Seiten der Forschung noch recht geringe Aufmerksamkeit geschenkt. Ein Grund hierfür sind die heterogenen Ziele und Varianten der Maßnahme, die eine Erfolgsbeobachtung erschweren. Hinzu kommt das fundamentale Problem der Evaluationsforschung, die fehlende kontrafaktische Situation. Dennoch befasst sich in jüngerer Zeit eine zunehmende Zahl von Studien mit der Analyse von kürzeren Maßnahmen, wie eben Trainingsmaßnahmen. Hujer et al. (2006) zeigen in einer Studie für Eintritte in Trainingsmaßnahmen im Jahr 2000 in Westdeutschland, dass die Teilnahme an einer Trainingsmaßnahme die Suchzeit für eine neue Beschäftigung von Arbeitsuchenden deutlich reduziert. Demnach führen Trainingsmaßnahmen besonders kurz- und mittelfristig zu einer Verkürzung der Arbeitslosigkeitsdauer. Dieser Effekt nimmt aber sechs Monate nach der Teilnahme stark ab und ist nach 12 Monaten nicht mehr feststellbar. Auch Biewen et al. (2007) weisen darauf hin, dass Maßnahmen mit kurzer Dauer im Vergleich zu deutlich teureren Maßnahmen mit mittlerer und langer Dauer überraschend effektiv sind. Hier werden Personen untersucht, die zwischen Februar 2000 und Januar 2002 arbeitslos geworden sind und an verschiedenen Maßnahmen unterschiedlicher Dauer teilnahmen. Wunsch/Lechner (2007) dagegen stellen fest, dass Maßnahmen generell nicht in der Lage sind, die Chancen von Teilnehmern auf eine reguläre, ungeforderte Beschäftigung zu verbessern. Sie weisen darauf hin, dass Maßnahmeteilnehmer im Vergleich zu Nicht-Teilnehmern desto schlechter dastehen, je länger die Teilnahme an einer Maßnahme dauert. Hier werden Eintritte in verschiedene Maßnahmen der aktiven Arbeitsmarktpolitik in Westdeutschland zwischen Januar 2000 und Dezember 2002 untersucht. Diese Studien wurden auf Basis von Prozessdaten durchgeführt, welche eine Vielzahl von relevanten Informationen enthalten. Leider können diese Daten keine Information über ausgesprochene

Einladungen zu Maßnahmen und den Zeitpunkt dieser Einladungen liefern. Daher ist es auf Basis dieser Daten nicht möglich, die Wirkung der Ankündigung einer Maßnahme zu untersuchen, sondern lediglich die Wirkung der tatsächlichen Teilnahme an der Maßnahme. Deshalb wird hier eine andere Herangehensweise an die Frage, inwieweit Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit zur Verbesserung der Eingliederungsaussichten von Arbeitsuchenden beitragen, gewählt.

Dabei wird das Problem der fehlenden kontrafaktischen Situation mit Hilfe eines Experiments gelöst. Dazu wurden durch eine ausgewählte Agentur für Arbeit aus potenziellen Maßnahmeteilnehmern mit Hilfe einer Zufallsauswahl eine Teilnehmer- und eine Vergleichsgruppe gebildet, wodurch eine Evaluation des Eingliederungserfolgs von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit ermöglicht wurde. Die Durchführung eines Experiments hat den Vorteil, dass nicht nur Teilnehmer untersucht werden können, sondern auch potenzielle Teilnehmer einbezogen werden können, die zwar eine Einladung zu einer Maßnahme erhalten haben, an dieser jedoch aus unterschiedlichen Gründen nicht teilgenommen haben. Dadurch ist es möglich, nicht nur zu untersuchen, welche Wirkung die tatsächliche Teilnahme an der Maßnahme auf die Eingliederungsaussichten der Teilnehmer hat, sondern auch welchen Effekt bereits die Einladung zu einer solchen Maßnahme hat.

Im Folgenden wird zunächst die Ausgestaltung von Trainingsmaßnahmen beschrieben. Anschließend werden Möglichkeiten der Wirkungsmessung mit Hilfe eines Experiments, der Untersuchungsaufbau und mögliche Effekte von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit dargestellt, bevor konkrete Ergebnisse des Experiments präsentiert werden.

## 2 Ausgestaltung von Trainingsmaßnahmen

Trainingsmaßnahmen sollen zur Verbesserung der Eingliederungsaussichten von Arbeitslosen und von Arbeitslosigkeit bedrohten Arbeitsuchenden beitragen. Ihr Ziel ist es, diesem Personenkreis die Möglichkeit zu geben, ihre persönliche Eignung und ihre beruflichen Fertigkeiten zu überprüfen oder alternative Beschäftigungsfelder in Erwägung zu ziehen und zu erproben. Die Teilnehmer an der Trainingsmaßnahme bleiben arbeitslos gemeldet und beziehen weiterhin Lohnersatzleistungen.

<sup>1</sup> Zu ausführlicheren Informationen zu Trainingsmaßnahmen siehe Kurtz (2003).

Tabelle 1

**Eintritte in Trainingsmaßnahmen (TM) und Förderung der beruflichen Weiterbildung (FbW) in den Jahren 2000–2007**

Gesamtdeutschland	2000	2001	2002	2003	2004	2005	2006	2007
FbW Gesamt	522.939	441.907	454.699	254.718	185.041	66.417	144.391	17.516
TM Gesamt	485.339	551.176	864.961	1.064.293	1.188.369	483.598	533.634	477.597
darunter:								
Überprüfung Verfügbarkeit	84.800	106.444	195.929	245.486	270.451	84.142	91.128	88.619
Maßnahmekombination	164.769	169.417	197.828	229.714	212.725	67.198	79.663	66.749
Eignungsfeststellung	102.124	121.178	280.422	316.039	372.217	185.081	205.045	192.932
Vermittlung von Kenntnissen	133.646	154.137	190.782	273.054	332.929	135.082	140.471	113.017
Existenzgründung					47	12.095	17.327	16.280

Quelle: Datawarehouse der Bundesagentur für Arbeit.

Hinweis: Ab 2005 beziehen sich die Zahlen auf Maßnahmen im Rechtskreis des SGB III und sind daher mit den vorherigen Zahlen nicht direkt vergleichbar.

Gefördert werden nach § 49 SGB III Maßnahmen,

- a) in denen die Kenntnisse und Fähigkeiten, das Leistungsvermögen und die beruflichen Entwicklungsmöglichkeiten des Arbeitslosen oder von Arbeitslosigkeit bedrohten Arbeitssuchenden sowie sonstige, für die Eingliederung bedeutsame Umstände ermittelt werden und unter Berücksichtigung der Arbeitsmarktlage festgestellt wird, für welche berufliche Tätigkeit oder Leistung der aktiven Arbeitsförderung er geeignet ist (Eignungsfeststellung),
- b) die die Selbstsuche von Arbeitslosen oder von Arbeitslosigkeit bedrohten Arbeitssuchenden sowie ihre Vermittlung, insbesondere durch Bewerbungstraining und Beratung über Möglichkeiten der Arbeitsplatzsuche, unterstützen oder die Arbeitsbereitschaft und Arbeitsfähigkeit prüfen (Überprüfung der Verfügbarkeit),
- c) die dieser Personengruppe Kenntnisse und Fähigkeiten vermitteln, um eine Vermittlung in Arbeit oder einen erfolgreichen Abschluss einer beruflichen Aus- oder Weiterbildung erheblich zu erleichtern (Kenntnisvermittlung).

Die Förderhöchstdauer variiert je nach Inhalt und Ziel der Maßnahme zwischen zwei und acht Wochen. Es können auch verschiedene Arten von Trainingsmaßnahmen kombiniert werden. Insgesamt dürfen die verschiedenen Maßnahmenteilen jedoch die Dauer von 12 Wochen nicht übersteigen. Trainingsmaßnahmen können sowohl bei einem Bildungsträger als auch in einem Betrieb durchgeführt werden.

Eine Teilnahmeförderung von Arbeitssuchenden kommt in Betracht, wenn die Maßnahme angemessen

und geeignet ist, die Eingliederungsaussichten zu verbessern. Über das Vorliegen der Teilnahmevoraussetzungen entscheidet jeweils die Agentur für Arbeit, in deren Bezirk die Person wohnt. Die Teilnahme an der Trainingsmaßnahme ist verpflichtend, das heißt bei unentschuldigter Nicht-Teilnahme können Sanktionen in Form von Leistungskürzungen verhängt werden. Das Ziel von Trainingsmaßnahmen soll immer die Integration in den ersten Arbeitsmarkt sein. Dazu sieht das Gesetz auch explizit die Möglichkeit vor, Trainingsmaßnahmen zur Überprüfung der Arbeitsbereitschaft und Arbeitsfähigkeit einzusetzen (§ 49 Abs. 2 Nr. 1 SGB III).

Bei Trainingsmaßnahmen handelt es sich um ein Instrument der aktiven Arbeitsförderung mit, im Vergleich zu anderen Maßnahmen, relativ kurzer Förderdauer. Daher ist diese Art von Maßnahme gegenüber anderen qualifizierenden Maßnahmen, wie zum Beispiel der deutlich längeren Förderung der beruflichen Weiterbildung, auch mit niedrigeren Kosten verbunden. Neben diesem finanziellen Aspekt haben in den letzten Jahren auch zahlreiche Gesetzesänderungen, die einem größeren Personenkreis Zugang zu diesem Instrument eröffneten, zu einem deutlichen Anstieg der Förderzahlen im Bereich der Trainingsmaßnahmen geführt (Kurtz 2003). Während die Eintritte in Maßnahmen zur Förderung der beruflichen Weiterbildung (FbW) seit 2000 kontinuierlich abgenommen haben, sind die Förderzahlen im Bereich der Trainingsmaßnahmen (TM) deutlich angestiegen. Tabelle 1 unterstreicht daher die zunehmend wichtigere Rolle, die Trainingsmaßnahmen im Bereich der qualifizierenden Maßnahmen spielen.

Im Jahr 2007 teilten sich die Eintritte in die verschiedenen Formen von Trainingsmaßnahmen im Rechtskreis SGB III folgendermaßen auf: 40,4 Prozent der Eintritte entfielen auf Maßnahmen zur Eignungsfeststellung, auf Kenntnisvermittlung entfielen 23,7 Prozent, auf Maßnahmen zur Überprüfung der Verfügbarkeit 18,6 Prozent sowie auf Maßnahmekombinationen 14 Prozent.<sup>2</sup>

Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit zielen – wie der Name impliziert – in erster Linie darauf ab, zu überprüfen, ob Arbeitslose zu einer Arbeitsaufnahme zur Verfügung stehen und auch bereit sind, eine Arbeit aufzunehmen. Deshalb ist bei dieser Art von Maßnahme nicht zwangsläufig davon auszugehen, dass eine Verbesserung der Eingliederungsaussichten aufgrund der Teilnahme an der Maßnahme (zum Beispiel durch Bewerbungstraining oder Praktikum on-the-job) erreicht wird. Vielmehr ist zu erwarten, dass sich Arbeitslose aus der Arbeitslosigkeit abmelden, die eben nicht für den Arbeitsmarkt zur Verfügung stehen.

Dann ist zwar mit Abmeldungen zu rechnen, welche jedoch nicht unbedingt Eingliederungen in den Arbeitsmarkt darstellen müssen. Um einer Teilnahme aus dem Wege zu gehen, ist auch damit zu rechnen, dass Betroffene ihre Suchaktivitäten verstärken und möglicherweise auch ihren Reservationslohn senken. Als Konsequenz werden sie ihre Arbeitslosigkeit früher beenden als Personen, die nicht eingeladen wurden (Rosholm/Svarer 2004). Diese Abmeldungen sind dann nicht auf das Trainingsprogramm der Maßnahme zurückzuführen, sondern auf den Droheffekt der Einladung zur Maßnahme.

Lediglich die Personen, die für eine Arbeitsaufnahme zur Verfügung stehen, werden auch tatsächlich an der Maßnahme teilnehmen. Anhand dieser Gruppe kann überprüft werden, inwieweit die Teilnahme an der Trainingsmaßnahme selbst die Eingliederungsaussichten verbessert. Hier ist dann mit zwei unterschiedlichen Effekten zu rechnen: dem sogenannten „Lock-in“-Effekt während der Teilnahme an der Maßnahme und der langfristigen Wirkung der Maßnahme nach deren Ende. Da während der Teilnahme die Suchintensität verringert wird, verringern sich meist die Eingliederungschancen während dieser Periode. Die Chancen sollten sich jedoch nach Ende der Maßnahme deutlich verbessern, da sich die Vermittlungsfähigkeit jetzt verbessert hat und die Suchintensität wieder erhöht wird.

Die vorliegende Untersuchung evaluiert mit Hilfe von Daten eines Experiments die Wirkung von Train-

ingsmaßnahmen zur Überprüfung der Verfügbarkeit in Bezug auf eine Wiedereingliederung der betroffenen Personen in den ersten Arbeitsmarkt. Dabei wird überprüft, ob die Einladung zu dieser Maßnahme lediglich zu Abmeldungen aus der Arbeitslosigkeit oder eben auch zu Integrationen in den ersten Arbeitsmarkt führt. Es wird sowohl die Wirkung der Ankündigung (hier also der Einladung zur Maßnahme) auf alle potenziellen Teilnehmer untersucht, wie auch die Wirkung der tatsächlichen Teilnahme an der Maßnahme.

### 3 Wirkungsmessung und Experiment

Zur Untersuchung einer Fragestellung, wie der Wirkungsweise von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit, sind zunächst zwei Wege denkbar. Zum einen die Durchführung eines Experiments, zum anderen der Einsatz nicht-experimenteller Daten. Bei einem Experiment werden Versuchspersonen randomisiert auf die Experimental- und die Kontrollgruppe aufgeteilt und die unabhängige Variable manipuliert, um so den durchschnittlichen Effekt der Teilnahme an der Maßnahme auf die Teilnehmer schätzen zu können. Da die Durchführung von Experimenten im Bereich der aktiven Arbeitsmarktpolitik oftmals nicht möglich ist, wird häufig auf nicht-experimentelle Daten zurückgegriffen. Da hier jedoch eine zufällige Zuweisung und Manipulation der unabhängigen Variablen nicht durchgeführt werden kann, müssen in diesem Fall verschiedene Annahmen getroffen werden, um eine Vergleichbarkeit der beiden Gruppen nachträglich herstellen zu können.

Das fundamentale Problem der Wirkungsmessung liegt dabei bei beiden Herangehensweisen in der Tatsache, dass sich das Erwerbsverhalten von Personen, die an einer Maßnahme teilgenommen haben, nicht ebenso für den hypothetischen Fall einer Nicht-Teilnahme beobachten lässt. Man spricht für die nicht beobachtete Alternative von der sogenannten kontrafaktischen Situation oder nach Rubin (1974, 1977, 1980) von potenziellen Ergebnissen, von denen jeweils nur eines realisiert wird.

T bezeichne zur Veranschaulichung, ob eine Person an der Maßnahme teilnimmt oder nicht, für Teilnehmer gelte  $T=1$ , für Nicht-Teilnehmer  $T=0$ .  $Y$  stellt den Effekt der Maßnahme dar:  $Y_1$  das Ergebnis bei Teilnahme,  $Y_0$  das Ergebnis bei Nicht-Teilnahme. Der durchschnittliche Effekt der Teilnahme an der Maßnahme auf die Teilnehmer („Average Treatment Effect on the Treated“, ATT) ist dann gegeben durch:

$$E(Y_1|T=1) - E(Y_0|T=1). \quad (1)$$

<sup>2</sup> Quelle: Eigene Berechnungen auf Basis Datawarehouse der Bundesagentur für Arbeit.

Hier tritt das fundamentale Evaluationsproblem auf. Da es nicht möglich ist, für eine Person das Ergebnis bei Teilnahme und Nicht-Teilnahme gleichzeitig zu beobachten, ist lediglich die Differenz der beobachteten durchschnittlichen Ergebnisse von Teilnehmern und Nicht-Teilnehmern bekannt,

$$\begin{aligned} & E(Y_1|T=1) - E(Y_0|T=0) \\ &= \underbrace{E(Y_1|T=1) - E(Y_0|T=1)}_{ATT} \\ &+ \underbrace{E(Y_0|T=1) - E(Y_0|T=0)}_{\text{Selektionsverzerrung}}. \end{aligned} \quad (2)$$

Unter anderem aus rechtlichen und organisatorischen Gründen ist es meist nicht möglich, ein Experiment zur Lösung des Evaluationsproblems durchzuführen, weshalb in der Evaluationsforschung häufig nicht-experimentelle Daten verwendet werden. Aus den verfügbaren Daten sind in diesem Fall aber nur die Unterschiede in den beobachteten durchschnittlichen Ergebnissen ermittelbar, um beispielsweise die Wirkung einer Maßnahme zu beurteilen. Dabei ergibt sich eine sogenannte „Selektionsverzerrung“, wenn die Zuordnung zur Maßnahme und das Ergebnis nach Teilnahme nicht unabhängig voneinander sind. So wird sich das Ergebnis bei Nicht-Teilnahme bei Teilnehmern und Nicht-Teilnehmern systematisch unterscheiden, wenn bestimmte Merkmale sowohl Teilnahmewahrscheinlichkeit als auch den erwarteten Fördererfolg beeinflussen (Heckman/Smith 1995). Dieses Problem liegt beispielsweise vor, wenn vor allem Personen mit vergleichsweise guter schulischer Ausbildung eine Weiterbildungsmaßnahme beginnen. Hierfür könnte sowohl die Zuweisung durch die Vermittler als auch die Teilnahmeentscheidung der Teilnehmer verantwortlich sein. Sind die Teilnehmer nun im Durchschnitt besser ausgebildet als der Durchschnitt der Nicht-Teilnehmer, so werden sie bereits aus diesem Grund nach der Maßnahme am Arbeitsmarkt erfolgreicher sein als Nicht-Teilnehmer. Wird dies bei der Evaluation nicht berücksichtigt, würde die Wirkung der Maßnahme überschätzt werden.

Ein Lösungsansatz bei der mikroökonomischen Evaluation ist daher, zu einer Teilnehmergruppe eine geeignete Nicht-Teilnehmergruppe mit vergleichbaren soziodemographischen Merkmalen zu finden, die als Vergleichsgruppe fungiert. Durchgeführt werden kann die Bildung der Vergleichsgruppe mit Matchingverfahren, wie beispielsweise Propensity-Score-Matching. Anschließend kann dann die kausale Wirkung einer Maßnahme gemessen werden (Rubin 1974). Dabei ist darauf zu achten, eine Gruppe zu finden, die der Teilnehmergruppe möglichst ähnlich ist, um die Annahme

sicherzustellen, dass sich die Vergleichsgruppe verhält, wie sich die Teilnehmergruppe im hypothetischen Fall der Nicht-Teilnahme verhalten hätte. Wichtig ist dabei, alle Merkmale in die Berechnung der Propensity Scores einzubeziehen, die sowohl die Teilnahme als auch den Maßnahmeerfolg beeinflussen. Ist nach Bildung der Vergleichsgruppe die Annahme bedingter Unabhängigkeit (Conditional Independence Assumption, CIA) erfüllt,

$$(Y_0, Y_1 \perp T) | X, \quad (3)$$

so entsprechen sich die (potenziellen) Ergebnisse von Teilnehmern und Nicht-Teilnehmern für jeden Teilnahmestatus. Der durchschnittliche Effekt der Teilnahme an der Maßnahme auf die Teilnehmer kann dann durch einen Vergleich der durchschnittlichen Ergebnisse der Teilnehmer- und Vergleichsgruppe geschätzt werden.

Eine andere Möglichkeit zur Lösung des fundamentalen Evaluationsproblems ist die Durchführung eines Experiments. Als Experiment bezeichnet man allgemein die systematische und wiederholbare Generierung von Daten durch kontrollierte und isolierte Variation von Faktoren, die als relevant für die untersuchten Zustände oder Prozesse betrachtet werden (Müntnich/Wießner 2002). Experimente lösen das Evaluationsproblem, indem sie Personen zufällig einer Teilnehmer- und einer Nicht-Teilnehmergruppe zuweisen. Damit ist sichergestellt, dass sich die beiden Gruppen sowohl in beobachtbaren als auch unbeobachtbaren Charakteristika sehr ähnlich sind. Mit anderen Worten: ein Experiment konstruiert die fehlende kontrafaktische Situation (Heckman/Smith 1996). Durch Zufallszuweisung ist das Ergebnis bei Teilnahme und bei Nicht-Teilnahme statistisch unabhängig davon, ob oder wie eine Teilnahmeentscheidung fällt. So wird beispielsweise unterbunden, dass vor allem besser ausgebildete Personen mit höherer Erfolgswahrscheinlichkeit an einer Maßnahme teilnehmen. Bei einem Experiment muss eine geeignete Teilnehmergruppe also nicht erst gefunden werden, diese ist durch die randomisierte Aufteilung bereits vorhanden. Unter bestimmten Bedingungen kann dann durch einen einfachen Vergleich des durchschnittlichen Ergebnisses der Teilnehmer- und Vergleichsgruppe eine unverzerrte Schätzung des Maßnahmeeffektes ermittelt werden (Smith 2000).

Zu diesen Bedingungen gehört, dass bei der Evaluation von Experimenten, wie bei anderen Evaluationsmethoden auch, Ergebnisse nur unter bestimmten Annahmen geschätzt werden können. Diese Annahmen werden von Heckman et al. (1999) beschrieben. Um dies zu zeigen, werden Ergebnisse,

die durch die Zufallszuweisung zustande kommen, mit „\*“ bezeichnet. Deshalb erhalten wir bedingt auf  $X$  für jede Person  $(Y_1^*, Y_0^*, T^*)$ , wenn eine Zufallszuweisung durchgeführt wurde und  $(Y_1, Y_0, T)$  bei der normalen Zuweisung in die Maßnahme durch die Vermittler. Die grundlegende Annahme, um eine Zufallszuweisung zur Lösung des Evaluationsproblems verwenden und dadurch den durchschnittlichen Effekt der Maßnahme auf die Teilnehmer schätzen zu können, ist dann:

$$E(Y_1^* - Y_0^* | X, T^* = 1) = E(Y_1 - Y_0 | X, T = 1). \quad (4)$$

Diese Annahme besagt also, dass die Mittelwerte der Teilnehmer- und Kontrollgruppe, die durch Zufallszuweisung generiert werden, die gewünschten Parameter der Grundgesamtheit liefern. Ist diese Annahme erfüllt, haben experimentelle Daten den Vorteil, dass die geschätzten Maßnahmeeffekte direkt frei von Selektionsverzerrungen sind, von denen Studien auf Basis von nicht-experimentellen Daten betroffen sein können (Burtless 1995).

Bei der Verwendung von experimentellen Daten muss aber dennoch überprüft werden, ob andere mögliche Verzerrungen ausgeschlossen werden können. Nicht möglich sein darf zum Beispiel, dass Teilnehmer der experimentellen Kontrollgruppe während des Experiments an Maßnahmen mit vergleichbarem Inhalt und Zielsetzung teilnehmen. Ist dies der Fall, spricht man von Substitution, die zu einer Verzerrung der Evaluationsergebnisse führt (Heckman/Smith 1996). Ein weiteres Problem, das zu einer Verzerrung der Ergebnisse führt, ist non-compliance (Bijwaard/Ridder 2005). Eine Verzerrung der Ergebnisse entsteht hier, wenn Personen aus der Versuchsgruppe an der zugewiesenen Maßnahme nicht teilnehmen. Trotz gewissenhafter Vorbereitung des Experiments und Auswahl der Personen wird es sich meist nicht vermeiden lassen, dass Personen aus der Teilnehmer- als auch aus der Kontrollgruppe vorzeitig ausscheiden oder die Beteiligung am Experiment verweigern. Hängt die Beteiligung von Faktoren ab, die auch den Maßnahmeerfolg beeinflussen, ist nach Beendigung des Experiments die Vergleichbarkeit der beiden Gruppen in den relevanten Charakteristika nicht mehr gegeben. Ein einfacher Mittelwertvergleich führt dann zu einer verzerrten Schätzung der Maßnahmewirkung. Deshalb muss dann ein Weg gefunden werden, um hier nachzujustieren.

#### 4 Untersuchungsaufbau

Innerhalb eines engen rechtlichen Rahmens ist es der Bundesagentur für Arbeit erlaubt, Maßnahmeteilnahmen zufällig zuzuweisen, wodurch die experi-

mentelle Evaluation von Maßnahmen möglich ist. Im Zeitraum zwischen September und Dezember 2005 wurde durch die Bundesagentur für Arbeit in einer ausgewählten Agentur in Westdeutschland ein Experiment zu Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit durchgeführt.<sup>3</sup> Die Daten dieses Experiments bieten sich zur Untersuchung der Wirkungsweise von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit an, da diese gegenüber nicht-experimentellen Daten einen entscheidenden Vorteil aufweisen: Sie enthalten zusätzlich Angaben zu Einladungen zu der genannten Maßnahme, während in nicht-experimentellen, prozessgenerierten Daten nur Angaben zu tatsächlichen Maßnahmeteilnahmen erfasst werden. Das bedeutet, dass hier nicht nur Personen in die Analyse einbezogen werden können, die tatsächlich an der Maßnahme teilgenommen haben, sondern auch Personen, die zwar zu der Maßnahme eingeladen wurden, aber aus unterschiedlichen Gründen nicht daran teilgenommen haben.

Für das Experiment wurden von der ausgewählten Agentur für Arbeit Personen bestimmt, die einer von drei für das Experiment vorgesehenen Gruppen angehören und daher für die Beteiligung am Experiment in Frage kommen. Aus diesen wurden dann die am Experiment teilnehmenden Personen zufällig ausgewählt. Die Dauer der Arbeitslosigkeit zu Beginn der Trainingsmaßnahme sollte dabei bis zu fünf Monate betragen. Die Kriterien für die Gruppenbildung wurden so gewählt, dass relevante Gruppen homogener Teilnehmer mit ausreichender Fallzahl gefunden werden konnten. Die potenziellen Teilnehmer der drei Gruppen wurden deshalb anhand der folgenden Kriterien bestimmt:

##### Gruppe 1: Teilzeitarbeitslose, nur Frauen

Diese Gruppe besteht aus Frauen im Alter von 20 bis 55 Jahren, die eine Teilzeitbeschäftigung suchen. Das bedeutet, diese Personen suchen eine Beschäftigung von mindestens 15 Stunden, aber mit weniger als der tariflichen Normalarbeitszeit. Im Dezember 2007 gab es in Gesamtdeutschland 423.460 registrierte teilzeitarbeitsuchende Arbeitslose, darunter 391.615 Frauen.<sup>4</sup>

##### Gruppe 2: Arbeitslose mit gewerblichen Berufen

In dieser Gruppe befinden sich sowohl Männer als auch Frauen im Alter von 20 bis

<sup>3</sup> Nach der Regionaltypisierung von Bezirken der Bundesagentur für Arbeit (Blien/Hirschenauer 2005) handelt es sich um eine Agentur des Vergleichstyps IIIc: Bezirke mit unterdurchschnittlicher Arbeitslosigkeit und wenig Dynamik.

<sup>4</sup> Angaben der Statistik der Bundesagentur für Arbeit.

Tabelle 2

**Teilnehmer am Experiment nach Gruppen und Geschäftstellen**

Gruppe	Geschäftsstelle	Fallzahl	Prozent
1: Teilzeitarbeitslose, nur Frauen	1	43	9,05
	2	44	9,26
	3	45	9,47
2: Arbeitslose mit gewerblichen Berufen	1	58	12,21
	2	58	12,21
	3	52	10,95
3: Arbeitslose im Alter von 20–27 Jahren	1	60	12,63
	2	59	12,42
	3	56	11,79
Total		475	100

55 Jahren, sofern diese eine Beschäftigung in einem gewerblichen Beruf suchen.

**Gruppe 3: Arbeitslose im Alter von 20 bis 27 Jahren**

In der Gruppe 3 befinden sich Arbeitslose im Alter von 20 bis 27 Jahren, dabei spielt es keine Rolle, ob diese bereits eine Ausbildung abgeschlossen haben oder nicht.

Durch die Beteiligung von Personen verschiedener Gruppen am Experiment lässt sich untersuchen, ob eine Überprüfung der Verfügbarkeit auf unterschiedliche Personengruppen unterschiedliche Auswirkungen hat. So wäre durchaus denkbar, dass bestimmte Personengruppen häufiger nicht zur Vermittlung zur Verfügung stehen als andere Personengruppen. Die für das Experiment ausgewählten Personen jeder der drei Gruppen teilen sich wiederum auf drei Geschäftsstellen der Agentur auf. Da es sich um drei Geschäftsstellen einer Agentur für Arbeit handelt, die räumlich direkt aneinander angrenzen, sind diese in Bezug auf die vorherrschenden Arbeitsmarktbedingungen aber durchaus miteinander vergleichbar.

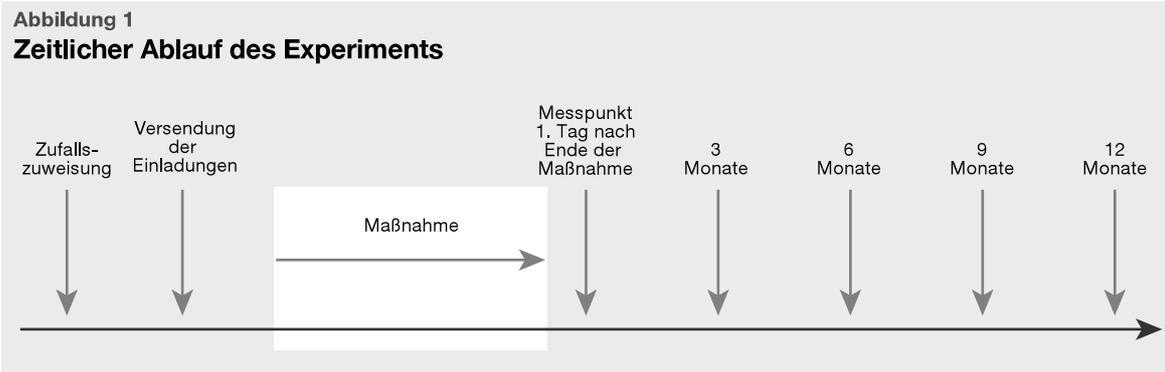
Die so ausgewählten Personen jeder Gruppe wurden anschließend zufällig in den Teil, der zu der Trainingsmaßnahme eingeladen wurde (Teilnahmegruppe), und die jeweilige Vergleichsgruppe (Kontrollgruppe) aufgeteilt.

Die drei Teilgruppen, die so für die tatsächliche Teilnahme vorgesehen waren, wurden zu achtwöchigen Trainingsmaßnahmen eingeladen, die im Zeitraum zwischen September und Dezember 2005 stattfanden. Diese umfassten eine vierwöchige Trainingsmaßnahme zur Eignungsfeststellung sowie ein vier-

wöchiges Anschlusspraktikum. Die Personen der Vergleichsgruppen wurden zu keiner Maßnahme eingeladen, wurden jedoch genauso wie die tatsächlichen Teilnehmer weiter beobachtet.

Während des Beobachtungszeitraums nahmen sowohl die Personen der Teilnahmegruppen als auch die der Kontrollgruppen an keiner weiteren Maßnahme teil, wodurch eine Verzerrung durch Substitution ausgeschlossen werden kann. Zu verschiedenen Messzeitpunkten wurde dann sowohl für die eingeladenen Personen als auch für die Vergleichspersonen durch die Agentur für Arbeit neben weiteren relevanten Merkmalen erfasst, ob und aus welchem Grund Abmeldungen aus der Arbeitslosigkeit vorliegen.

Durch die Zufallsauswahl wurde also zunächst sichergestellt, dass sich Teilnahme- und Kontrollgruppen hinsichtlich relevanter Charakteristika nicht unterscheiden. Leider entstanden nach der Zufallszuweisung Probleme, die eine Nachjustierung notwendig machen, um möglichen Verzerrungen vorzubeugen. Der Hauptgrund für die Nachjustierung ist, dass aufgrund der geringen Fallzahlen die Teilnehmer der einzelnen Geschäftsstellen zusammengefasst werden mussten. Darüber hinaus gab es kleinere Schwierigkeiten bei der tatsächlichen Umsetzung der vorgesehenen Einladungen in den Geschäftsstellen. Es liegen also Fälle von non-compliance vor, die hier sowohl auf die ausgewählten Personen, aber auch auf die zuständigen Vermittler zurückzuführen sind. Darüber hinaus standen zur Auswertung für die Gruppe der Arbeitslosen mit gewerblichen Berufen nur die Daten aus zwei Geschäftsstellen zur Verfügung, weshalb sich in dieser Gruppe die Fallzahlen der Teilnahme- und Kontroll-



gruppe für diesen Personenkreis nach Beendigung der Maßnahme reduziert haben. Ein Teil der zum Experiment zugewiesenen Personen hatte bereits vor dem Versenden der Einladungen die Arbeitslosigkeit beendet und stand somit nicht mehr für die Auswertung zur Verfügung. Da dies jedoch Teilnahme- und Vergleichsgruppen gleichermaßen betrifft, beeinflussen diese Personen nicht das Ergebnis.

Die Nachjustierung erfolgt mit Hilfe von Propensity-Score-Matching (Sacerdote 2004). Die individuellen Wahrscheinlichkeiten (Propensity Scores) werden mit Hilfe von Logit- oder Probit-Modellen geschätzt. Im Rahmen eines Matching-Ansatzes auf der Basis der Propensity Scores können den Personen der Teilnehmergruppe jeweils einzelne Teilnehmer der Vergleichsgruppe zugewiesen werden. Das Ausbalancieren der Verteilungen unter Zuhilfenahme der Propensity Scores entspricht damit einem Ausbalancieren der Verteilungen der einzelnen Personenmerkmale zwischen den beiden Gruppen, um dann wieder über die mittlere Differenz in den individuellen Ergebnisvariablen Unterschiede zwischen den Gruppen schätzen zu können.

Für jede der drei Gruppen wurde daher ein Propensity-Score-Matching durchgeführt, um statistisch eine Situation „sonst gleicher Bedingungen“ wiederherzustellen. Als Matching-Methode wurde ein „Nearest-Neighbour-Matching“ ohne Replacement mit einem Caliper von 0,05 verwendet, also jedem Teilnehmer wurde genau die Vergleichsperson zugeordnet, deren Propensity Score den geringsten Abstand aufweist und eine Distanz von 0,05 nicht überschreitet. Dadurch wurde die Vergleichbarkeit der Teilnehmer- und Kontrollgruppen sichergestellt.

Darüber hinaus ermöglicht ein nachträgliches Matching neben der Untersuchung der Auswirkungen der Androhung einer Trainingsmaßnahme auch die Evaluierung der tatsächlichen Teilnahme an der Maßnahme. Eine Beschreibung der Vorgehensweise

zur Untersuchung der beiden Effekte liefert der folgende Abschnitt.

## 5 Mögliche Effekte von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit

Bei der Evaluation von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit stellt sich die Frage, wodurch eine mögliche Verbesserung der Eingliederungsaussichten zustande kommt (Black et al. 2003). Mögliche Effekte von Einladungen zu Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit müssen differenziert betrachtet werden. Diese Art von Maßnahme zielt darauf ab, festzustellen, ob Arbeitslose zu einer Arbeitsaufnahme zur Verfügung stehen und dazu auch bereit sind. Zu überprüfen ist also, wie sich die Ankündigung der Maßnahme auswirkt und welchen Effekt anschließend eine tatsächliche Teilnahme an der Maßnahme hat. Ein Treatment-Effekt kann sich dementsprechend allein durch die Tatsache ergeben, dass die Verpflichtung zur Teilnahme an einer Maßnahme besteht, aber auch durch die Teilnahme am Weiterbildungsprogramm der Maßnahme selbst, das ja zur Verbesserung der Eingliederungsaussichten von Arbeitslosen beitragen soll. Bei der Auswertung des Experiments werden deshalb auch diese zwei möglichen Effekte der Trainingsmaßnahmen untersucht.

- Zum einen ist zu überprüfen, welche Auswirkungen sich aufgrund der Einladung der Teilnehmer, also der Ankündigung der Maßnahme ergeben. Da es sich hier um eine Trainingsmaßnahme zur Überprüfung der Verfügbarkeit handelt, liegt der Gedanke nahe, dass einzelne Personen für die tatsächliche Teilnahme an der Maßnahme nicht zur Verfügung stehen, was zu einer höheren Austrittsquote von eingeladenen Personen vor und zu Beginn der Maßnahme gegenüber Vergleichsper-

sonen führt, die sich einer Überprüfung der Verfügbarkeit nicht ausgesetzt sehen. Hier wird das Matching zum Zeitpunkt der Einladungen durchgeführt.

- Darüber hinaus soll auch die Wirkung der tatsächlichen Teilnahme an der Maßnahme geschätzt werden. Hier werden für die Teilnehmergruppe nur Personen einbezogen, die auch wirklich an der Maßnahme teilnehmen, die also die Trainingsmaßnahme zur Eignungsfeststellung sowie das vierwöchige Praktikum durchlaufen. Personen, die sich etwa nach Erhalt der Einladung zur Trainingsmaßnahme aus der Arbeitslosigkeit abmelden oder nicht zur Maßnahme erscheinen, werden nicht mehr in die Analyse einbezogen. Für die Kontrollgruppe können dann auch nur Personen einbezogen werden, die zum Beginn der Maßnahme noch arbeitslos waren und ursprünglich auch als Vergleichspersonen vorgesehen waren, also keine Einladung erhalten haben. Aus der Analyse ausgeschlossen werden Personen, die ihre Arbeitslosigkeit zwischen der Einladung und dem Beginn der Maßnahme beendet haben. Das Matching wird mit Bezug auf den Zeitpunkt des Beginns der Maßnahme durchgeführt. Geschätzt wird hier ausschließlich der Effekt, den die tatsächliche Teilnahme an der Maßnahme hat, ohne Berücksichtigung des Effekts, der sich aus der Überprüfung der Verfügbarkeit, also der Einladung zur Maßnahme ergibt. Dies entspricht einer Evaluation von Trainingsmaßnahmen, wie sie auch in Studien, die nicht auf experimentellen, sondern auf Prozessdaten beruhen, durchgeführt werden kann. Da in diesen Daten normalerweise keine Informationen zu ausgesprochenen Einladungen vorliegen, können in diesen Studien nur die Effekte von tatsächlichen Teilnahmen an einer Maßnahme evaluiert werden. Das wiederum bedeutet, dass nur ein Vergleich zwischen tatsächlichen Maßnahmeteilnehmern und Nicht-Teilnehmern gezogen werden kann. Effekte, die durch die Überprüfung der Verfügbarkeit zustande kommen, müssen hier vernachlässigt werden. Ein Nachteil dieser Herangehensweise ist, dass sich hier in den Kontrollgruppen Personen befinden können, die eigentlich nicht für eine Arbeitsaufnahme zur Verfügung stehen, während aus den Teilnahmegruppen durch die Einladungen bereits alle Personen herausgefiltert wurden, die nicht verfügbar sind. Hier hat die experimentelle Anordnung aber dennoch den Vorteil, dass ausgeschlossen werden kann, dass Arbeitsuchende, die eigentlich für die Teilnahme an der Maßnahme vorgesehen sind, diese aber verweigern oder vorzeitig abbrechen, in den Pool der möglichen Vergleichspersonen

gelangen. Dies kann bei Prozessdaten unter Umständen nicht ausgeschlossen werden, da diese Personen dort nicht identifiziert werden können.

## 6 Strukturmerkmale der Teilnehmer- und Vergleichsgruppen

Die folgenden Tabellen zeigen eine Übersicht über demographische Charakteristika der jeweiligen Teilnahme- und der Kontrollgruppen. Dargestellt sind jeweils die Mittelwerte relevanter Variablen nach dem Nachjustieren für den Zeitpunkt der Einladung zur Trainingsmaßnahme sowie zum Zeitpunkt des Beginns der Maßnahme. Diese Variablen fungieren auch als erklärende Variablen im Probit-Modell zur Schätzung der Propensity Scores. Aufgrund der relativ kleinen Fallzahlen in den jeweiligen Gruppen kann es schwierig sein, eine optimale Übereinstimmung der Mittelwerte der Charakteristika zwischen den jeweiligen Teilnehmer- und Vergleichsgruppen zu erreichen. Es handelt sich hier jedoch um eine Experimentsituation, wobei lediglich ein Nachjustieren notwendig wurde. Deshalb fallen etwaige Abweichungen hier nicht so stark ins Gewicht, weil die Zuweisung mit Hilfe einer Zufallsauswahl durchgeführt wurde.<sup>5</sup> Eine Nachjustierung wurde lediglich notwendig, da einzelne Personen nicht für die Auswertung des Experiments zur Verfügung standen.

Anhand der Tabellen 3 bis 5 zeigt sich auch, in welchen Merkmalen sich die drei Gruppen voneinander unterscheiden. Zum einen sind das, gemäß der Definition der Gruppen, natürlich das Alter und das Geschlecht. Darüber hinaus unterscheiden sich die Gruppen in der Frage, ob die Personen Kinder haben, sowie dem Vorhandensein von gesundheitlichen Einschränkungen. Das aktuelle Qualifikationsniveau wird von zuständigen Vermittlern bei der ersten Gruppe geringer eingeschätzt als bei den beiden anderen Gruppen. Während bei der Gruppe der Frauen der überwiegende Anteil verheiratet ist, ist dies bei der Gruppe der Arbeitslosen im Alter von 20 bis 27 Jahren nur ein kleiner Prozentsatz. Bei den Merkmalen „Dauer der letzten Beschäftigung“ und „Leistungshöhe“ sind ebenfalls starke Unterschiede zwischen den Gruppen zu erkennen. Demnach ist davon auszugehen, dass Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit bei den verschiedenen Gruppen auch unterschiedliche Wirkungen zeigen.

<sup>5</sup> Mit Hilfe von t-tests wurde überprüft, ob sich eine befriedigende Angleichung in der Verteilung der beobachteten Merkmale zwischen Teilnehmer- und Vergleichsgruppe ergeben hat. Die Hypothese der Gleichheit der Mittelwerte der Merkmale kann hier durchgängig bei allen drei Gruppen nicht abgelehnt werden.

Tabelle 3

**Mittelwerte erklärender Variablen, Teilzeitarbeitslose, nur Frauen**

	Einladung (matched)		Teilnahme (matched)	
	Teilnehmer	Vergleich	Teilnehmer	Vergleich
Alter	39.72	36.07	37.14	38.24
männlich	0	0	0	0
Nationalität	0.95	0.93	0.90	0.90
Kinder	0.72	0.74	0.76	0.76
gesundheitliche Einschränkungen	0.07	0.12	0.19	0.19
Dauer Arbeitslosigkeit <sup>a</sup>	2.25	2.09	2.26	2.07
Haupt-/Realschule <sup>b</sup>	0.77	0.74	0.81	0.76
Fachabi/Abitur <sup>c</sup>	0.23	0.25	0.19	0.24
aktuelles Qualifikationsniveau <sup>d</sup>	0.23	0.30	0.29	0.19
Ausbildung	0.23	0.26	0.24	0.19
Fachschule	0.68	0.54	0.57	0.57
Uni/FH	0.09	0.19	0.19	0.24
verheiratet/Partnerschaft	0.81	0.77	0.76	0.76
Alter quadriert	1639.02	1387.65	1454.29	1531.76
Dauer letzte Beschäftigung <sup>e</sup>	39.44	38.61	47.71	68.90
Leistungshöhe <sup>f</sup>	88.93	78.87	97.25	101.88
n <sup>g</sup>	57	57	21	21

*Anmerkungen:*<sup>a</sup> Zur Übersichtlichkeit hier dargestellt in Monaten.<sup>b</sup> Anteil von Haupt-/Realschulabsolventen.<sup>c</sup> Anteil der Abiturienten.<sup>d</sup> Einschätzung des Vermittlers über das aktuelle Qualifikationsniveau.<sup>e</sup> Hier in Monaten.<sup>f</sup> Die Leistungshöhe ging kategorisiert in die Propensity Score Schätzung ein, hier übersichtlicher dargestellt als Mittelwert der tatsächlichen Höhe.<sup>g</sup> Anzahl der Personen nach der Nachjustierung, das heißt ohne Personen für die keine geeignete Vergleichsperson gefunden werden konnte, was zu einer Verringerung gegenüber der ursprünglich für das Experiment ausgewählten Personenzahl führt.

Tabelle 4

**Mittelwerte erklärender Variablen, Vollzeit gewerblich**

	Einladung (matched)		Teilnahme (matched)	
	Teilnehmer	Vergleich	Teilnehmer	Vergleich
Alter	40.16	41.60	40.79	40.33
männlich	0.80	0.88	0.83	0.79
Nationalität	0.82	0.78	0.83	0.79
Kinder	0.24	0.25	0.21	0.33
gesundheitliche Einschränkungen	0.22	0.12	0.17	0.17
Dauer Alo	2.33	2.17	2.44	2.29
Haupt-/Realschule	0.80	0.80	0.79	0.75
Fachabi/Abitur	0.06	0.08	0.04	0.04
aktuelles Qualifikationsniveau	0.43	0.53	0.46	0.58
Ausbildung	0.39	0.49	0.46	0.54
Fachschule	0.60	0.51	0.54	0.46
Uni/FH	0	0	0	0
verheiratet/Partnerschaft	0.57	0.59	0.46	0.54
Alter quadriert	1716.08	1816.47	1745.46	1700.25
Dauer letzte Beschäftigung	62.32	57.45	57.88	64.90
Leistungshöhe	155.25	136.11	131.89	112.03
n	51	51	24	24

Tabelle 5

**Mittelwerte erklärender Variablen, Arbeitslose im Alter von 20–27 Jahren**

	Einladung (matched)		Teilnahme (matched)	
	Teilnehmer	Vergleich	Teilnehmer	Vergleich
Alter	21.05	20.80	21.44	20.97
männlich	0.67	0.68	0.63	0.63
Nationalität	0.94	0.95	1	1
Kinder	0.04	0.02	0	0.03
gesundheitliche Einschränkungen	0.07	0.02	0.03	0.03
Dauer Alo	1.96	2.08	2.11	2.13
Haupt-/Realschule	0.83	0.88	0.91	0.88
Fachabi/Abitur	0.14	0.12	0.09	0.13
aktuelles Qualifikationsniveau	0.32	0.23	0.06	0.09
Ausbildung	0.31	0.23	0.06	0.09
Fachschule	0.65	0.72	0.94	0.88
Uni/FH	0.04	0.05	0	0.03
verheiratet/Partnerschaft	0.06	0.04	0	0.06
Alter quadriert	446.70	435.94	462.81	442.53
Dauer letzte Beschäftigung	15.70	19.53	21.33	20.79
Leistungshöhe	59.72	60.33	72.39	83.51
n	81	81	32	32

Festzustellen ist außerdem, dass der Anteil der eingeladenen Personen, der tatsächlich an der Maßnahme teilnimmt, in allen drei Gruppen eher gering ausfällt. Interessant ist die Tatsache, dass bei der Gruppe der Teilzeitarbeitslosen und der Gruppe der Arbeitslosen im Alter von 20 bis 27 Jahren die durchschnittliche Leistungshöhe der Teilnehmer an der Maßnahme deutlich über der durchschnittlichen Leistungshöhe aller eingeladenen Personen liegt. Dies deutet darauf hin, dass Personen dieser Gruppen eher tatsächlich an der Maßnahme teilnehmen, wenn sie durch eine Nichtteilnahme auf höhere Transferleistungen verzichten müssen.

## 7 Ergebnisse

Zunächst dargestellt werden in diesem Kapitel die Eingliederungseffekte aufgrund der Einladung zur Trainingsmaßnahme. Es wird also untersucht, welche Auswirkungen eine Einladung zu einer Trainingsmaßnahme gegenüber der kontrafaktischen Situation der Nichteinladung hat. Unterschieden werden dabei die beiden Zielvariablen Abgang aus Arbeitslosigkeit und Abgang in Erwerbstätigkeit, die zweite Zielvariable entspricht hier auch gleichzeitig dem Abgang in ungeforderte Beschäftigung, da die ausgewählten Personen zunächst in keine weiteren Maßnahmen vermittelt wurden. Der Schätzer für den Effekt der Einladung wird berechnet als Differenz der Anteile von Personen aus Teilnehmer- und Vergleichsgruppe, die zum jeweiligen Messzeitpunkt bereits aus der Arbeitslosigkeit abgegangen sind beziehungsweise eine Erwerbstätigkeit aufgenommen haben. Er wird jeweils zu Beginn der Maßnahme, am ersten Tag nach Ende der Maßnahme sowie nach 3, 6, 9, und 12 Monaten gemessen (durchschnittlicher Effekt der Einladung zur Maßnahme auf die eingeladenen Personen). Wünschenswert wäre hier eine Darstellung der Effekte in kürzeren Intervallen, was aber auf Basis der vorhandenen experimentellen Daten leider nicht möglich ist. Ausgewiesen werden auch Zeitpunkte nach Ende der Maßnahme. Selbst wenn sich hier eventuell bereits Effekte der tatsächlichen Maßnahmeteilnahme zeigen, ist es dennoch interessant zu untersuchen, welche Auswirkungen Einladungen zu Trainingsmaßnahmen mittel- und langfristig zeigen. So wäre es ja denkbar, dass sich ein Teil der Personen, die sich aufgrund der Teilnahme aus der Arbeitslosigkeit abgemeldet haben, nach wenigen Wochen wieder arbeitsuchend meldet und die Wirkung der Überprüfung der Verfügbarkeit deshalb nur von relativ kurzer Dauer ist.

Analog dazu werden die durchschnittlichen Eingliederungseffekte der tatsächlichen Teilnahme an der

Maßnahme geschätzt. Untersucht wird hier der Effekt, den eine tatsächliche Teilnahme an der Trainingsmaßnahme gegenüber der Nicht-Teilnahme hat. Effekte definieren wir hier als positiv, wenn eine höhere Abgangsrate der Teilnehmergruppe gegenüber der Kontrollgruppe zu verzeichnen ist. Vergleichspersonen sind in diesem zweiten Fall Personen, die nicht zur Maßnahme eingeladen wurden und zu Beginn der Maßnahme noch arbeitslos gemeldet sind. Beträgt der geschätzte Effekt zu einem Messzeitpunkt beispielsweise 0,2, so ist zu diesem Zeitpunkt die Wahrscheinlichkeit, die Arbeitslosigkeit beendet zu haben bzw. eine Erwerbstätigkeit aufgenommen zu haben, in der Gruppe der Maßnahmenteilnehmer 20 Prozentpunkte höher als in der Vergleichsgruppe. Die Varianz der Eingliederungseffekte wird nach folgender Formel (Lechner 2001) geschätzt:

$$\text{Var}(\hat{\theta}) = \frac{1}{N_1} \text{Var}(Y_1) + \frac{\sum_{i=1}^{N_0} (w_i)^2}{(N_1)^2} \text{Var}(Y_0). \quad (5)$$

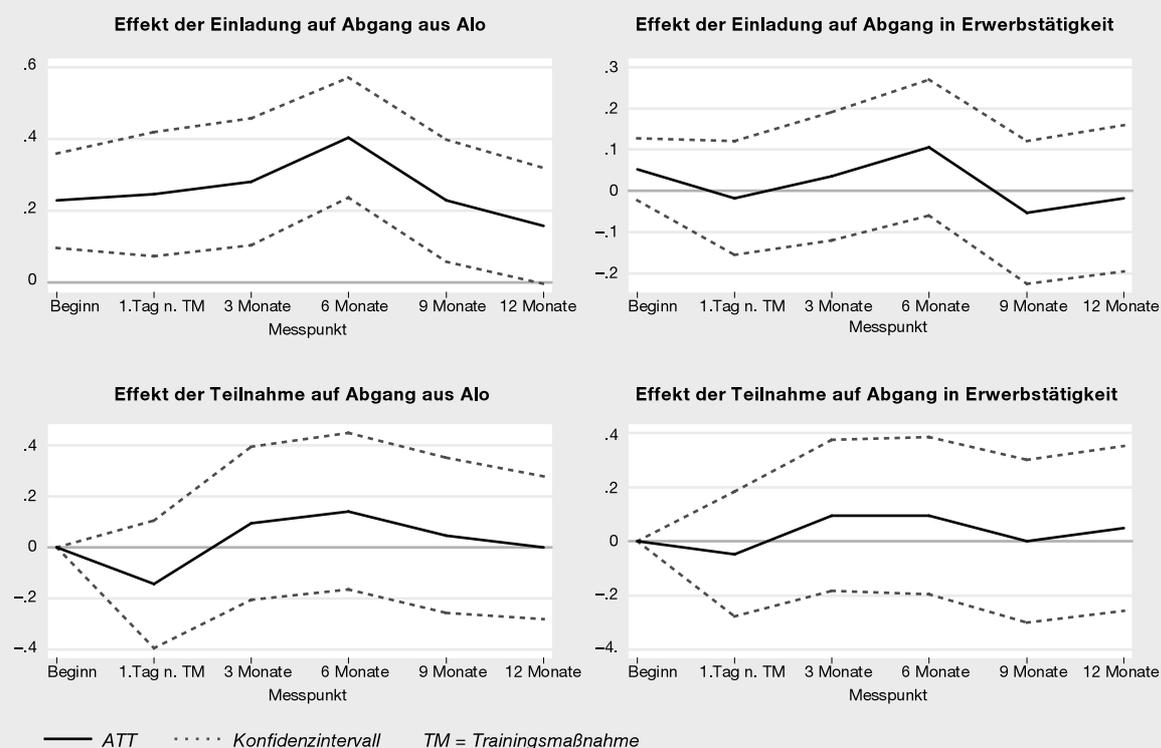
Dabei ist  $\hat{\theta}$  die geschätzte Maßnahmewirkung,  $N_1$  ist die Anzahl der Versuchspersonen, hier also im ersten Fall die Zahl der eingeladenen Personen und im zweiten Fall die Zahl der tatsächlichen Teilnehmer,  $N_0$  ist dann die Anzahl der jeweiligen Vergleichspersonen, sowie  $w_i$  das Gewicht mit dem Person  $i$  aus der Kontrollgruppe in die Schätzung der Maßnahmewirkung eingeht.

### 7.1 Teilzeitarbeitslose, nur Frauen

Bei der Gruppe der Frauen in Teilzeitarbeitslosigkeit zeigt sich ein positiver Eingliederungseffekt aufgrund der Einladung. In Bezug auf Abgänge aus Arbeitslosigkeit ist dieser Effekt zu allen Messpunkten signifikant positiv. Bei der Messung der Wirkung auf Integrationen in den ersten Arbeitsmarkt hingegen zeigt sich ein geringerer, nicht signifikanter Effekt. Diese Ergebnisse deuten also darauf hin, dass eine Überprüfung der Verfügbarkeit bei diesem Personenkreis zu einer erhöhten Abgangswahrscheinlichkeit aus Arbeitslosigkeit gegenüber der Kontrollgruppe führt, die dann auch über einen längeren Zeitraum hinweg konstant positiv bleibt, jedoch nur zu einer unwesentlich höheren Wahrscheinlichkeit, eine Beschäftigung aufzunehmen.

Bei der Evaluierung der tatsächlichen Teilnahme an der Trainingsmaßnahme zeigt sich hier jedoch weder bei den Abgängen aus Arbeitslosigkeit, noch bei Abgängen in Erwerbstätigkeit ein deutlicher Effekt. Zudem zeigt sich, dass ein positiver Effekt sich erst

**Abbildung 2**  
**Geschätzte Maßnahmewirkung, Teilzeitarbeitslose, nur Frauen\***



\* Anzumerken ist hier, dass der Effekt der tatsächlichen Teilnahme zu Beginn der Maßnahme aufgrund des erneuten Matching hier null sein muss.

nach Ende des „Lock-in“-Effekts einstellt. Diese Befunde legen den Schluss nahe, dass in dieser Gruppe Trainingsmaßnahmen tatsächlich zu Abmeldungen aus Arbeitslosigkeit führen, die jedoch, wie erwartet, auf die Überprüfung der Verfügbarkeit durch die Ankündigung einer Maßnahme zurückzuführen sind. Durch die tatsächliche Teilnahme an der Trainingsmaßnahme wird kaum eine Verbesserung der Eingliederungsaussichten erreicht.

### 7.2 Vollzeit gewerblich

Auch in der Gruppe der Vollzeitbeschäftigten mit gewerblichen Berufen zeigt sich ein positiver Effekt aufgrund der Einladung, der auch über den Beobachtungszeitraum hinweg erhalten bleibt. Der Effekt ist allerdings weder in Bezug auf Abmeldungen aus Arbeitslosigkeit noch in Bezug auf Eingliederungen in Erwerbstätigkeit signifikant. Dagegen wird bei dieser Gruppe ein signifikant negativer Eingliederungseffekt aufgrund der tatsächlichen Teilnahme an der Maßnahme zu den Messpunkten am Ende und drei Monate nach Ende der Maß-

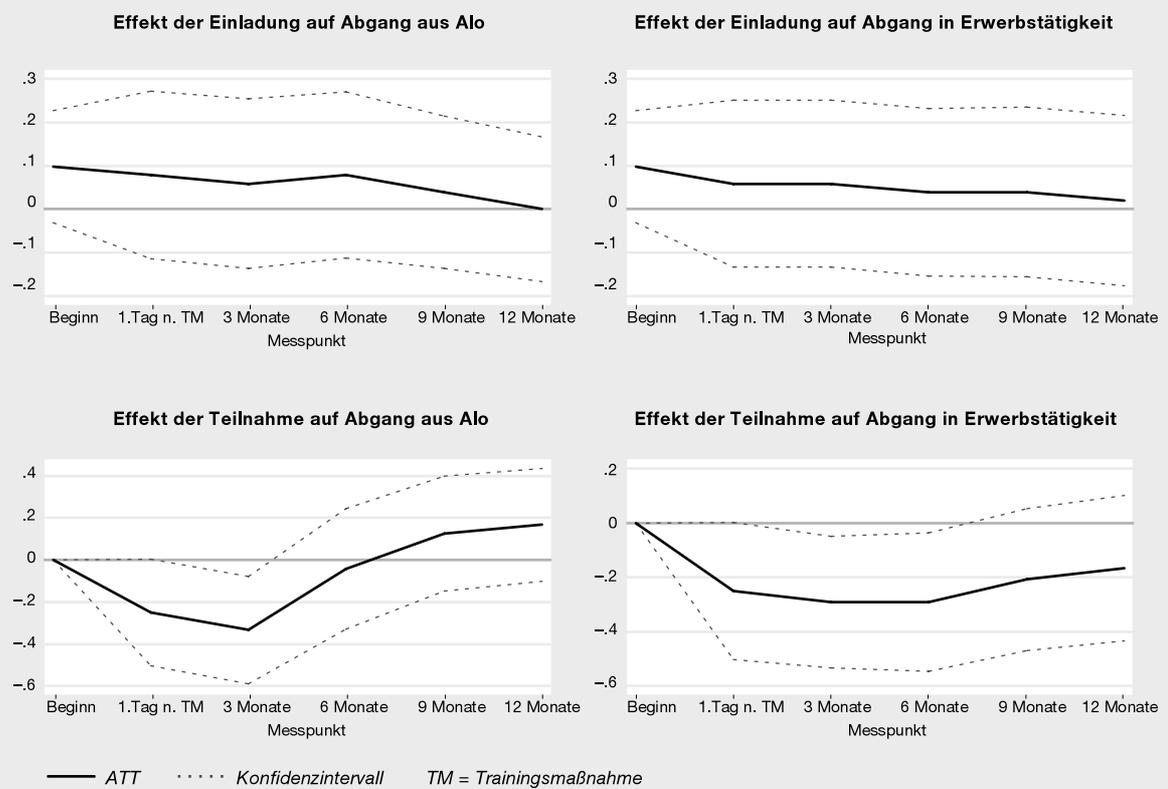
nahme verzeichnet. Zum Messpunkt nach neun Monaten zeigt sich dann ein positiver, aber insignifikanter Effekt in Bezug auf Abgänge aus Arbeitslosigkeit.

Die tatsächliche Teilnahme an der Maßnahme hat also nach einer gewissen Anlaufzeit einen positiven Effekt in Bezug auf Abmeldungen aus der Arbeitslosigkeit, nicht jedoch in Bezug auf Integrationen in den Arbeitsmarkt. Bei diesem Personenkreis ist folglich kein nennenswerter Effekt allein durch die Einladung zu einer Trainingsmaßnahme zu beobachten, dafür ist davon auszugehen, dass sich die tatsächliche Teilnahme an der Maßnahme mittelfristig positiv auf die Eingliederungsaussichten auswirkt.

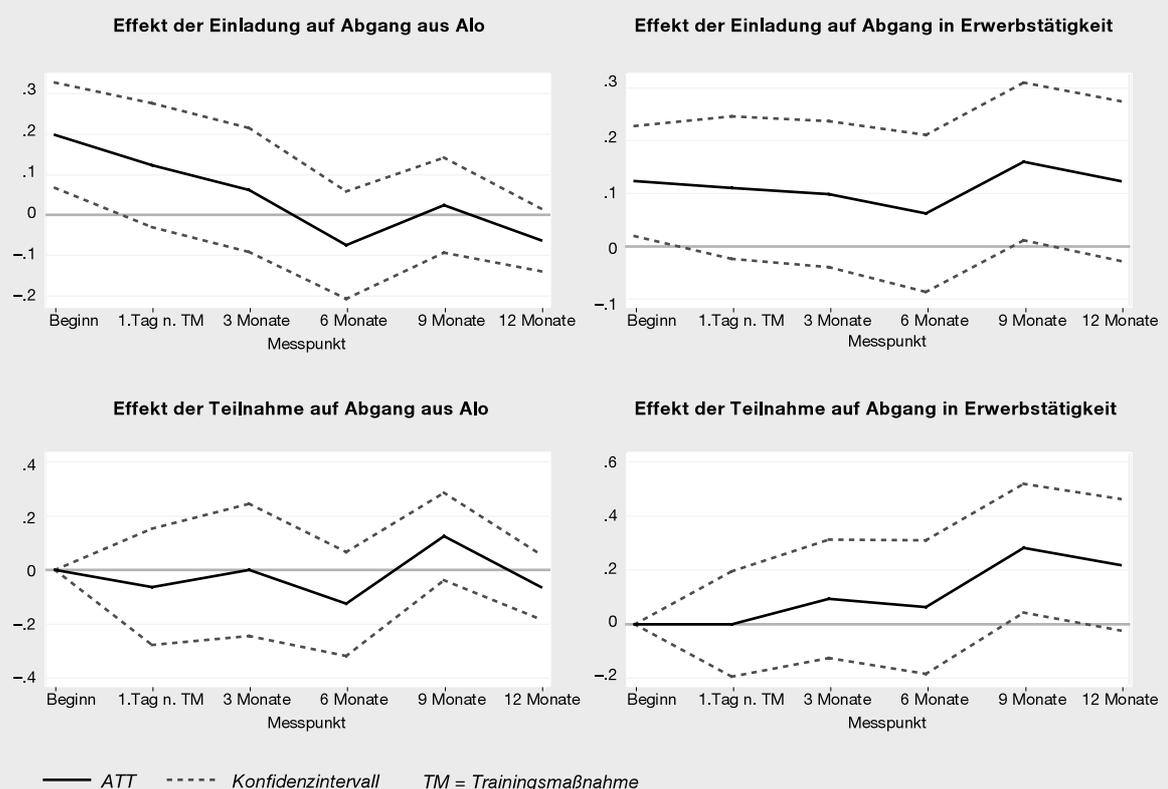
### 7.3 Arbeitslose im Alter von 20 bis 27 Jahren

Etwas differenzierter stellt sich das Bild bei der Gruppe der jüngeren Arbeitslosen dar. Auch hier

**Abbildung 3**  
**Geschätzte Maßnahmewirkung, Vollzeit gewerblich**



**Abbildung 4**  
**Geschätzte Maßnahmewirkung, Arbeitslose im Alter von 20–27 Jahren**



hat die Einladung zur Trainingsmaßnahme zunächst einen deutlichen, signifikant positiven Effekt auf die Abgangswahrscheinlichkeit aus Arbeitslosigkeit. Dieser nimmt jedoch zwischen den Messpunkten nach drei und nach sechs Monaten deutlich ab und wird insignifikant. Bei der Betrachtung der Zielvariable Abgänge in Erwerbstätigkeit bleibt der Effekt positiv, ist jedoch ebenfalls nicht über den gesamten Beobachtungszeitraum hinweg signifikant. Bei der Evaluierung der tatsächlichen Teilnahme an der Maßnahme zeigt sich in Bezug auf Abgänge in Erwerbstätigkeit nach neun Monaten ein signifikant positiver Effekt. Hier kann also davon ausgegangen werden, dass sich hier eine Teilnahme an der Maßnahme nach längerer Anlaufzeit positiv auf die Abgangsrate in Erwerbstätigkeit auswirkt.

In dieser Gruppe kann zusammenfassend also ein kurzfristiger, positiver Effekt auf die Abgangswahrscheinlichkeit aus Arbeitslosigkeit aufgrund der Einladung zur Maßnahme sowie ein langfristiger, positiver Effekt auf die Eingliederungswahrscheinlichkeit in Erwerbstätigkeit durch die Teilnahme an der Trainingsmaßnahme festgestellt werden.

## 8 Fazit

Bei der Evaluation von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit stellt sich die Frage, ob und wenn ja wodurch dieses Instrument der aktiven Arbeitsmarktpolitik zu einer höheren Abgangsrate von Arbeitssuchenden führt. Führt die tatsächliche Teilnahme an einer Maßnahme zu einer Verbesserung der Eingliederungschancen oder kommen Abgänge aus Arbeitslosigkeit vor allem aufgrund der Ankündigung einer Maßnahme, also der Überprüfung der Verfügbarkeit, zustande? Diese beiden Effekte können auf Basis von Daten eines arbeitsmarktpolitischen Experiments separat betrachtet werden, da diese neben Informationen zu Maßnahmeteilnehmern auch Informationen zu ausgesprochenen Einladungen enthalten. So kann hier untersucht werden, ob Trainingsmaßnahmen in der Lage sind, zu überprüfen, ob Arbeitssuchende für den Arbeitsmarkt zur Verfügung stehen.

Die Evaluation dieses Experiments zeigt zunächst einen positiven Abgangs- beziehungsweise Eingliederungseffekt, der sich durch die Einladung zu einer Maßnahme ergibt. Dieser Effekt beruht in erster Linie auf Abgängen aus Arbeitslosigkeit, die bereits vor Beginn der Maßnahme stattfinden. Mit Hilfe des Propensity-Matchings kann auch der Effekt der tatsächlichen Teilnahme an der Maßnahme evaluiert werden. Eine positive Maßnahmewirkung aufgrund

der tatsächlichen Teilnahme kann jedoch nur in geringem Ausmaß festgestellt werden. Eine höhere Abgangsrate der eingeladenen Personen gegenüber der Vergleichsgruppe aufgrund der Überprüfung der Verfügbarkeit ist hier vor allem bei der Gruppe der teilzeitarbeitslosen Frauen und bei der Gruppe der Arbeitslosen im Alter von 20 bis 27 Jahren in den ersten Monaten festzustellen. Bei der Gruppe der Arbeitslosen mit gewerblichen Berufen ist dieser Effekt dagegen nicht so deutlich ausgeprägt, dafür ist hier nach einer gewissen Anlaufzeit ein leicht positiver Effekt durch die Teilnahme an der Maßnahme zu erkennen. Ein signifikant positiver Effekt der tatsächlichen Teilnahme ist jedoch nur bei der Gruppe der jüngeren Arbeitslosen nach Ablauf von neun Monaten nach Ende der Maßnahme festzustellen.

Die Ergebnisse dieser Evaluation von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit mit Hilfe von Daten eines Experiments deuten daher darauf hin, dass diese sehr wohl geeignet sind, zu überprüfen, ob arbeitslos gemeldete Personen tatsächlich zur Teilnahme an einer Maßnahme und damit eben auch zur Vermittlung zur Verfügung stehen. Dies kann anhand der höheren Wahrscheinlichkeiten der Personen der Teilnehmergruppe, nach Erhalt der Einladung zur Maßnahme die Arbeitslosigkeit zu beenden, festgestellt werden. Allerdings muss in diesem Zusammenhang auch darauf hingewiesen werden, dass die Abmeldungen aufgrund der Einladung überwiegend nicht zu Integrationen in den ersten Arbeitsmarkt führen. Soll das Ziel aber auch eine Verbesserung der Eingliederungschancen durch die Maßnahme selbst sein, kann hier nur mäßiger Erfolg festgestellt werden: Die Evaluation der tatsächlichen Teilnahme an der Maßnahme zeigt kurzfristig kaum eine Auswirkung auf die Abgangs- und Eingliederungsrate. Die Ergebnisse zeigen jedoch, dass sich eine Teilnahme an der Maßnahme langfristig positiv auf die Eingliederungsaussichten auswirken kann. Um dies abschließend beurteilen zu können, wären jedoch weitere Beobachtungspunkte über den Zeitraum von zwölf Monaten hinaus nötig.

Schließlich muss aber auch, gerade vor dem Hintergrund der relativ geringen Fallzahlen, darauf hingewiesen werden, dass sich diese Schlussfolgerungen lediglich auf das hier untersuchte Experiment in einem bestimmten regionalen Arbeitsmarkt (in Westdeutschland) beziehen und daher nicht verallgemeinert werden sollten.

## Literatur

- Biewen, M./Fitzenberger, B./Osikominu, A./Waller, M.* (2007): Which program for whom? Evidence on the comparative effectiveness of public sponsored training programs in Germany, ZEW Discussion Paper 07-042.
- Bijwaard, G./Ridder, G.* (2005): Correcting for selective compliance in re-employment bonus experiment, *Journal of Econometrics*, 125, S. 77–111.
- Black, D./Smith, J./Berger, M./Noel, B.* (2003): Is the Threat of Reemployment Services More Effective than the Services Themselves? Evidence from Random Assignment in the UI System, *The American Economic Review*, Vol. 93, No. 4, 1313–1327.
- Blien, U./Hirschenauer, F.* (2005): Vergleichstypen 2005 – Neufassung der Regionaltypisierung für Vergleiche zwischen Agenturbezirken, IAB-Forschungsbericht 24/2005, Nürnberg.
- Burtless, G.* (1995): The Case for Randomized Field Trials in Economic and Policy Research, *The Journal of Economic Perspectives*, Vol. 9, No. 2, 63–84.
- Heckman, J./Lalonde, R./Smith, J.* (1999): The Economic and Econometrics of Active Labour Market Programs, *Handbook of Labor Economics*, Volume 3, 1865–2095.
- Heckman, J./Smith, J.* (1996): Experimental and nonexperimental evaluation. In: Schmidt, G./O'Reilly, J./Schoorman, K. (Eds.): *International Handbook of Labour Market Policy and Evaluation* (Cheltenham, Elgar).
- Heckman, J./Smith, J.* (1995): Assessing the Case for Social Experiments, *Journal of Economic Perspectives*, Volume 9 (2), 85–110.
- Hujer, R./Thomsen, S./Zeiss, C.* (2006): The Effects of Short-Term Training Measures on the Individual Unemployment Duration in West Germany, ZEW Discussion Paper 06–65.
- Kurtz, B.* (2003): Trainingsmaßnahmen – was verbirgt sich dahinter?, IAB-Werkstattbericht 08/2003, Nürnberg.
- Lechner, M.* (2001): Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption. In: Lechner, M./Pfeiffer, F. (ed): *Econometric Evaluation of Labour Market Policies*, Heidelberg, 43–58.
- Müntnich, M./Wießner, F.* (2002): Soziale Experimente und Modellversuche: Ein Beitrag zur Evaluation von Neuanlässen in der Arbeitsmarktpolitik. In: IAB-Kompandium Arbeitsmarkt- und Berufsforschung, Beiträge zur Arbeitsmarkt- und Berufsforschung, BeitrAB 250, 415–427.
- Rosholm, M./Svarer, M.* (2004): Estimating the Threat Effect of Active Labour Market Programmes, IZA Discussion Paper No. 1300.
- Rubin, D. B.* (1980): Discussion of Randomization Analysis of Experimental Data in the Fisher Randomization Test by Basu, *Journal of the American Statistical Association*, 75, 591–593.
- Rubin, D. B.* (1977): Assignment to Treatment Group on the Basis of a Covariate, *Journal of Educational Statistics*, 2, 1–26, Printer's correction note 3, 384.
- Rubin, D. B.* (1974): Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies, *Journal of Educational Psychology*, 66, 688–701.
- Sacerdote, B.* (2004): Fixing broken experiments using the propensity score. In: *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, 61–71.
- Smith, J.* (2000): Evaluation aktiver Arbeitsmarktpolitik: Erfahrungen aus Nordamerika, *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, Jg. 33, H. 3, 345–356.
- Statistik der Bundesagentur für Arbeit*, *Arbeitsmarkt in Zahlen, Aktuelle Daten* (Dezember 2007).
- Wunsch, C./Lechner, M.* (2007): What Did All the Money Do? On the General Ineffectiveness of Recent West German Labour Market Programmes, IZA Discussion Paper 2800.