

## Das Wichtigste in Kürze

Dieser Beitrag gibt einen Überblick der neueren wissenschaftlichen Literatur zur Evaluation der Aktiven Arbeitsmarktpolitik (AAMP) in Deutschland. Aufgrund methodischer Fortschritte und der sich verbessernden Verfügbarkeit von Daten wurden in den letzten Jahren von wissenschaftlicher Seite zahlreiche neue Evaluationsstudien vorgelegt, die den großen methodischen Problemen einer adäquaten Evaluation auf Mikro- oder Makroebene in verstärktem Maße Rechnung tragen.

Der Beitrag beginnt mit der Darstellung der ökonomischen Wirkungen, die im Rahmen der AAMP erzielt werden sollen. Im anschließenden methodischen Teil werden die für die existierenden Evaluationsstudien der AAMP in Deutschland relevanten methodischen Aspekte dargestellt. Das Grundproblem der mikroökonomischen Evaluation auf Basis nichtexperimenteller Daten besteht darin, den Vergleichsmaßstab bei Nichtteilnahme zu schätzen, entweder auf der Basis einer vergleichbaren Kontrollgruppe von Nichtteilnehmern oder auf Basis der Situation der Teilnehmer vor der Teilnahme. Entscheidend ist hierbei, eine Vergleichbarkeit hinsichtlich beobachtbarer und nichtbeobachtbarer Charakteristika anzustreben. Evaluationsstudien auf der Makroebene müssen zusätzlich auch die indirekten Effekte der AAMP auf das gesamte Arbeitsergebnis berücksichtigen und die Endogenität der AAMP beachten: Da der Umfang der AAMP je nach Arbeitsmarktlage wechselt, muß eine exogene Variation in der AAMP isoliert werden, um deren Wirkung auf die Arbeitsmarktlage zu bestimmen. Der Hauptteil ist eine Übersicht der einzelnen Evaluationsstudien zur AAMP, die den genannten methodischen Problemen Rechnung tragen. Die Ergebnisse lassen sich wie folgt zusammenfassen:

1. Mikroökonomische Evaluationsstudien finden für den Einsatz der geförderten Weiterbildung in Ostdeutschland zumeist keine signifikanten Wirkungen. In Westdeutschland werden hauptsächlich negative Effekte festgestellt, eine Verbesserung der Arbeitsmarktsituation ergibt sich nur für einzelne Gruppen (z.B. Geringqualifizierte). Für den Einsatz von ABM gibt es bislang nur Ergebnisse für Ostdeutschland, die mehrheitlich signifikant negative Beschäftigungseffekte zeigen.
2. Makroökonomische Analysen finden für ABM zumeist positive Effekte auf die Verminderung struktureller Arbeitslosigkeit, während FuU die regionale Langzeitarbeitslosigkeit senkt.

Gegenüber diesen Ergebnissen sind jedoch gewichtige Vorbehalte vorzubringen: Erstens werden in den einzelnen Arbeiten höchst unterschiedliche Wege zur Lösung des Problems der Selektionsverzerrung beschritten, die zu unterschiedlichen Referenzwerten für die Schätzung des kausalen Effekts führen dürften. Zweitens, und weit entscheidender für die Aussagekraft der Ergebnisse, ist jedoch das zugrundeliegende Datenmaterial, das kleine Ausschnitte aus Umfrageerhebungen nutzt und oft aufgrund der geringen Fallzahlen auch höchst unterschiedliche Maßnahmen zusammenfaßt. Nach Ansicht der Autoren sind die politikrelevanten Folgerungen auf Basis dieser Maßnahmekategorien unklar. Erst durch die Bereitstellung umfangreicher prozeßgenerierter Programmdateien können die vielfältigen Optionen der Arbeitsmarktpolitik angemessen abgebildet werden und robuste Ergebnisse zur Wirkung der einzelnen Maßnahmen erzielt werden. Allerdings ist auch zu konstatieren, daß mögliche positive Effekte der AAMP nicht sehr stark sein können, da sie sonst vermutlich trotz unterschiedlicher Methoden und unzureichender Daten empirisch hätten identifiziert werden können.

# Zur wissenschaftlichen Evaluation der Aktiven Arbeitsmarktpolitik in Deutschland: Ein Überblick\*

Bernd Fitzenberger<sup>1</sup>

Stefan Speckesser<sup>2</sup>

Januar 2000

## Zusammenfassung

Dieser Beitrag gibt einen Überblick der neueren wissenschaftlichen Literatur zur Evaluation der Aktiven Arbeitsmarktpolitik (AAMP) in Deutschland. Das Grundproblem mikroökonomischer Evaluation besteht darin, den Vergleichsmaßstab bei Nichtteilnahme zu schätzen, sei es auf Basis einer vergleichbaren Kontrollgruppe von Nichtteilnehmern oder auf Basis der Situation vor Teilnahme. Entscheidend ist hierbei, den potentiellen Selektionsproblemen Rechnung zu tragen. Evaluationsstudien auf der Makroebene müssen zusätzlich auch die indirekten Effekte der AAMP auf den gesamten Arbeitsmarkt berücksichtigen und der möglichen Endogenität der AAMP Rechnung tragen.

Die meisten Evaluationsstudien für Deutschland zeigen keine signifikant positiven Beschäftigungseffekte der AAMP, sind jedoch mit großer statistischer Unsicherheit behaftet. Wir führen letzteres im wesentlichen darauf zurück, daß wegen zu geringen Datenumfanges meist äußerst heterogene Maßnahmen gepoolt werden und daß über Teilnehmer und Nichtteilnehmer zu wenige Hintergrundinformationen verfügbar sind. Aufgrund dieser Datenrestriktionen verbietet sich daher aus Sicht der Autoren ein vorschnelles Verdikt über die ökonomischen Wirkungen der AAMP, obwohl mögliche positive Effekte nicht sehr stark sein können, da sie sonst vermutlich trotz unterschiedlicher Methoden und unzureichender Daten empirisch hätten identifiziert werden können.

---

\* Wir danken den Teilnehmern der IAB-Tagung "Evaluation of Active Labour Market Policy in Germany" am 28. und 29. Oktober 1999 in der Bundesanstalt für Arbeit, Nürnberg, und Wolfgang Franz sowie zwei anonymen Gutachtern für wertvolle Kommentare. Alle Unzulänglichkeiten dieses Beitrags liegen allein in unserer Verantwortung.

<sup>1</sup> Universität Mannheim, ZEW und IFS

<sup>2</sup> Universität Mannheim

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Aktive Arbeitsmarktpolitik in Deutschland</b>	<b>2</b>
2.1	Reform der Arbeitsmarktpolitik . . . . .	3
2.2	Maßnahmen . . . . .	3
2.3	Umfang der AAMP . . . . .	5
<b>3</b>	<b>Methodische Aspekte des Evaluationsproblems</b>	<b>6</b>
3.1	Erfolgskriterien . . . . .	7
3.2	Das mikroökonomische Evaluationsproblem . . . . .	7
3.3	Makroökonomische Effekte . . . . .	10
<b>4</b>	<b>Empirische Ergebnisse</b>	<b>11</b>
4.1	Mikroökonomische Evaluationsstudien . . . . .	11
4.1.1	Weiterbildung in Ostdeutschland . . . . .	12
4.1.2	Weiterbildung in Westdeutschland . . . . .	14
4.1.3	Arbeitsbeschaffungsmaßnahmen . . . . .	16
4.2	Makroökonomische Evaluationsstudien . . . . .	18
<b>5</b>	<b>Schlußfolgerungen</b>	<b>19</b>

# 1 Einleitung

*Es ist daher dringend erforderlich, die Evaluation arbeitsmarktpolitischer Maßnahmen voranzubringen.*

Sachverständigenrat, SVR (1998, Ziffer 432)

Im Jahr 1998 werden mit dem "Eingliederungstitel" den Arbeitsämtern für Aktive Arbeitsmarktpolitik (AAMP) insgesamt 24,7 Milliarden DM zugewiesen (Bundesanstalt für Arbeit 1999b: 23), die die finanzielle Grundlage der wichtigsten Maßnahmen der Arbeitsmarktpolitik (AAMP) bilden, die Programme der subventionierten Beschäftigung und Bildungsmaßnahmen für verschiedene Zielgruppen umfassen. Zusammen mit weiteren 10,9 Milliarden DM für die Förderung "sonstiger Leistungen der aktiven Arbeitsförderung" werden für die Arbeitsförderung 1998 damit insgesamt 36,6 Milliarden DM aufgewandt (Bundesanstalt für Arbeit 1999b). In die drei wichtigsten Programmgruppen, Arbeitsbeschaffungsmaßnahmen, Weiterbildung und Lohnkostenzuschüsse für regulär Beschäftigte, treten 1998 insgesamt 3,8 vH der gesamten Erwerbspersonen ein, wobei die Arbeitsmarktpolitik vor allem in Ostdeutschland einen großen Umfang hat, wo 1998 13 vH der Erwerbstätigen in eine Maßnahme eintreten (Bundesanstalt für Arbeit 1999a). Der quantitativ bedeutende Umfang der AAMP, insbesondere in Ostdeutschland, bei einem konstant hohen Niveau der Arbeitslosigkeit und zunehmenden Budgetrestriktionen verdeutlicht den Bedarf nach wissenschaftlicher Begleitforschung (vgl. SVR 1998, Ziffer 432), die zuverlässige Verfahren entwickelt, um Maßnahmeneffekte bewerten zu können.

Da das Ziel einer Maßnahme der AAMP ist, Teilnehmer in reguläre Beschäftigung (zurück) zu führen, kann die Frage nach dem kausalen Effekt der Maßnahmen nicht allein durch deren direkte Entlastungseffekte auf die Arbeitslosenstatistik beantwortet werden, sondern erfordert eine ökonomische Analyse des Erfolgs von Maßnahmen nach deren Ende. Aufgrund der besseren Verfügbarkeit von Daten für die Bundesrepublik entstanden in den letzten Jahren eine ganze Anzahl neuer Evaluationsstudien, die die Wirkungen der AAMP auf ökonomische Erfolgsindikatoren prüfen. Dieser Beitrag bietet im folgenden einen Überblick der Evaluationsstudien, die explizit den in Abschnitt 3 skizzierten großen methodischen Problemen der Evaluation der AAMP auf der Mikro- und Makroebene Rechnung tragen. Dies ist im wesentlichen für Weiterbildung und Arbeitsbeschaffungsmaßnahmen der Fall, so daß sich eine Einschränkung auf diese beiden wichtigsten Maßnahmen von selbst ergibt.

Der vorliegende Beitrag gliedert sich in drei Teile: Zunächst rekapituliert der zweite Abschnitt den Katalog der Maßnahmen der AAMP in der Bundesrepublik und verdeutlicht, auf welche Weise in die institutionellen Rahmenbedingungen des Arbeitsmarktes eingegriffen wird, um arbeitsmarktpolitische Ziele zu verwirklichen. Die AAMP bietet dabei zahlreiche unterschiedliche Anreizstrukturen für eine Reintegration in den "ersten Arbeitsmarkt". Der dritte Abschnitt stellt die wesentlichen methodischen Probleme und Vorgehensweisen dar, die in Evaluationsstudien angewandt werden. Dabei ist in mikroökonomischen Evaluationen entscheidend, die Wirkung von Maßnahmen auf die Teilnehmer aufgrund eines angemessenen Vergleichs der Nichtteilnahme zu schätzen, wobei die Situation der Nichtteilnahme jedoch für Teilnehmer nicht beobachtbar ist. In nichtexperimentellen Daten bietet sich dabei die Situation der Nichtteilnehmer ebensowenig an wie die

der Teilnehmer vor einer Maßnahme, denn Personen, die nicht oder noch nicht an einer Maßnahme teilnehmen, unterscheiden sich hinsichtlich beobachtbarer und unbeobachtbarer Charakteristika von den Teilnehmern. Evaluationsstudien auf der Makroebene müssen zusätzlich indirekte Effekte der AAMP auf den gesamten Arbeitsmarkt berücksichtigen. Außerdem ist innerhalb einer Wirkungsanalyse der Endogenität des Ausmaßes der AAMP Rechnung zu tragen.

Der vierte Abschnitt gibt einen Überblick über die wissenschaftlichen Evaluationsstudien zur AAMP in Deutschland auf der mikro- und makroökonomischen Ebene. Über die betrachteten mikroökonomischen Studien hinweg herrscht Uneinigkeit, ob für Selektion allein auf Basis beobachtbarer Variablen hinreichend kontrolliert werden kann und inwieweit Paneldaten helfen, verbleibende Selektionseffekte zu berücksichtigen, da die Maßnahmeteilnahme meist sehr stark mit der Arbeitsmarktgeschichte zuvor zusammenhängt. Entsprechend unterscheiden sich die erzielten Evaluationsergebnisse. Unterschiedliche methodische Vorgehensweisen werden auch in Analysen der makroökonomischen Effekte der AAMP gewählt, die am Ende des Abschnitts dargestellt werden; auch hier erzielen die einzelnen Studien widersprüchliche Ergebnisse. Der letzte Abschnitt zieht Schlußfolgerungen aus den bisherigen Evaluationsergebnissen für Deutschland.

## 2 Aktive Arbeitsmarktpolitik in Deutschland

Aktive Arbeitsmarktpolitik (AAMP) zielt vor allem auf die Problemgruppen auf dem Arbeitsmarkt ab. Nach der Definition der OECD umfaßt AAMP fünf verschiedene Politikfelder: Arbeitsvermittlung und Arbeitsberatung, berufliche Fortbildung für Beschäftigte und Arbeitslose, Maßnahmen für Berufseinsteiger und Jugendliche, subventionierte Beschäftigung und Maßnahmen für die Integration Behinderter in den Arbeitsmarkt (OECD 1993: 39 ff.).

In Deutschland wird der überwiegende Teil der AAMP durch die Bundesanstalt für Arbeit (BA) organisiert. Wie auf Basis des Arbeitsförderungsgesetzes, das bis 1997 die Rechtsgrundlage der AAMP bildete und für die meisten hier zitierten Arbeiten gültig ist, werden auch unter dem neuen rechtlichen Rahmen des Sozialgesetzbuchs III (SGB III) zusätzlich zu den zwei wichtigsten Programmen – Arbeitsbeschaffungsmaßnahmen (ABM) und Förderung der beruflichen Weiterbildung (FbW) – Programme zugunsten Langzeitarbeitsloser, zur beruflichen Rehabilitation und zum Erwerb sprachlicher Fähigkeiten durchgeführt.

Auf Ebene der Länder und Gemeinden werden weitere Programme implementiert: Die Gemeinden als Träger der Sozialhilfe streben die Reintegration der Sozialhilfeempfänger an und haben dafür im Jahr 1998 200.000 befristete Beschäftigungsverhältnisse mit einem finanziellen Volumen von DM 10 Mrd. geschaffen (Deutscher Städtetag 1999). Schließlich haben die Aktionsprogramme der Bundesregierung (wie etwa JUMP) sowie die Programme der einzelnen Länder und des Europäischen Sozialfonds (ESF) einen zunehmenden Anteil an der Arbeitsmarktpolitik.

## 2.1 Reform der Arbeitsmarktpolitik

Mit dem Inkrafttreten der neuen Rechtsgrundlage SGB III wird bei Maßnahmen der Bundesanstalt für Arbeit eine Ausrichtung der AAMP auf besondere Problemgruppen angestrebt: Während das alte Arbeitsförderungsgesetz (AFG), noch unter den Bedingungen der Vollbeschäftigung geschaffen, auch eine Verbesserung der generellen Funktionsweise des Arbeitsmarkts durch die Arbeitsmarktpolitik vorsah, die insbesondere einen Mangel an qualifizierten Arbeitskräften verhindern sollte (AFG, §§ 1 und 2), ist die Zielrichtung des SGB III, Problemgruppen durch Beratung, berufliche Aus- und Weiterbildung und befristete Schaffung oder Teilfinanzierung von Arbeitsplätzen in reguläre Beschäftigung zu integrieren, so daß insbesondere Leistungen zur Einkommenssicherung bei Arbeitslosigkeit reduziert werden. Die Betonung der Integration in reguläre Beschäftigung und des Versicherungsprinzips zeigen, daß die Arbeitsmarktpolitik einen höheren Grad an Subsidiarität anstrebt (Sell 1998: 545).

Neben der Veränderung der Förderungsgrundsätze werden durch das Inkrafttreten des SGB III vor allem auch Flexibilitätsspielräume für die AAMP auf der regionalen und lokalen Ebene möglich: Die einzelnen Arbeitsämter können im Gegensatz zu vormals üblichen Einzelbudgets flexibel zwischen den einzelnen Maßnahmen wählen, da für die wesentlichen Maßnahmen nur noch ein Budget existiert. Zudem erlaubt das SGB III, eigene Maßnahmen zu entwickeln und zu implementieren (Sell 1998: 541).

Angesichts der Veränderung der finanziellen und institutionellen Grundlagen der AAMP in Deutschland ist davon auszugehen, daß die Arbeitsmarktpolitik in den nächsten Jahren flexibler wird und sich mehr an regionale Arbeitsmarktbedingungen anpaßt als in der Zeit vor Inkrafttreten des SGB III. Damit wird auch im Bereich der Evaluation der AAMP zusätzlich die Einbeziehung der regionalen Ebene bedeutend, um die neu entstehende Vielfältigkeit von Maßnahmen und Maßnahmenpaketen adäquat zu berücksichtigen.

## 2.2 Maßnahmen

Außer der Zunahme regionaler Flexibilität in der Implementation von AAMP sind allerdings die einzelnen Maßnahmen durch die Reform zum SGB III nicht wesentlich verändert worden. Wie unter dem AFG auch, können drei Komplexe von Instrumenten identifiziert werden: Die Integration durch berufliche Weiterbildung bildet dabei den wichtigsten Bereich. Als zweite Säule der AAMP bestehen die Lohnsubventionen aus der Förderung von Maßnahmen zur Arbeitsbeschaffung (ABM), Strukturanpassungsmaßnahmen (SAM) und befristeten Lohnsubventionen bei der Einstellung von Problemgruppen in reguläre Beschäftigung. Schließlich wird durch das Instrumentarium der AAMP auch regionale und berufliche Mobilität unterstützt, vor allem durch die Förderung der Aufnahme einer selbständigen Beschäftigung und bei beruflich bedingten Umzügen.

**Förderung der beruflichen Weiterbildung:** Innerhalb der Förderung der beruflichen Weiterbildung (FbW) herrscht auch unter dem SGB III die Förderung der Fortbildung und die Umschulung (FuU) in einen anderen Beruf vor. Diese Maßnahmen umfassen vor allem die Anpassung der Qualifikation an neue Techniken innerhalb eines erlernten Berufs und die Förderung einer Umschulungsmaßnahme, wenn die Ausübung eines erlernten Berufs aus gesundheitlichen oder Arbeitsmarktgründen nicht mehr möglich ist. Abhängig

vom Umfang der Maßnahme ergeben sich damit unterschiedliche Laufzeiten, in der Regel zwischen drei und acht Monaten für eine Fortbildungsmaßnahme und bis zu 24 Monaten bei einer Umschulung. Neben den Maßnahmekosten können Fahrtkosten und andere Ausgaben erstattet werden. Zusätzlich kann ein Unterhaltsgeld gewährt werden, das zwischen 60 vH und 66 vH des letzten Nettolohns beträgt.

**Subventionierte Beschäftigung:** Das wichtigste Programm im Bereich der subventionierten Beschäftigung sind Arbeitsbeschaffungsmaßnahmen (ABM), die eine befristete Beschäftigung in Projekten ermöglicht, die gesellschaftlich nützliche Arbeiten beinhalten und dem Prinzip der Additionalität folgen, d.h. ohne die Subvention nicht ausgeführt werden könnten. ABM sind kofinanzierte Maßnahmen, bei denen neben der Lohnsubvention durch die BA in Höhe von 30 bis 75 vH der gesamten Vergütung der Maßnahmeträger weitere Kosten tragen muß (z.B. Infrastruktur). Allerdings können in Ausnahmefällen auch weitere Leistungen in Anspruch genommen werden, so daß diese Förderungsrichtlinie nur eine grobe Abgrenzung ist. ABM laufen bis zu 24 Monate und können, wenn eine reguläre Beschäftigung folgt, bis zu 36 Monate andauern. Dabei darf die Vergütung 80 vH eines vergleichbaren nicht subventionierten Arbeitsplatzes nicht übersteigen.

Vor allem in Ostdeutschland spielen neben ABM die Strukturanpassungsmaßnahmen (SAM) eine erhebliche Rolle. Ebenso wie ABM zielen SAM auf eine Integration über den "zweiten Arbeitsmarkt" und die befristete Beschäftigung in einem "additionalen" Beschäftigungsverhältnis. Bei SAM ergeben sich hingegen andere Förderungsgrundsätze: Das Programm steht nicht nur Langzeitarbeitslosen offen, sondern wird oft auch für Beschäftigte implementiert. Unter SAM können nur Projekte gefördert werden, die zur Verbesserung der regionalen Infrastruktur dienen, vor allem in den Bereichen Jugend und Umwelt. Mit Einführung des SGB III wurden SAM in Ostdeutschland allerdings auch Wirtschaftsunternehmen zugänglich. Die Förderung beträgt in der Regel 24 Monate, kann aber auf bis zu 48 Monate verlängert werden, wenn eine Dauerbeschäftigung folgt. Die Förderung entspricht dabei dem Betrag an Arbeitslosengeld oder -hilfe, der bei Fortdauer der Arbeitslosigkeit durchschnittlich gezahlt worden wäre. Ähnlich ABM darf die Vergütung nur maximal 80 vH eines vergleichbaren nicht subventionierten Arbeitsplatzes betragen.

Neben der Förderung des zweiten Arbeitsmarkts unterstützt die BA auch die Aufnahme einer regulären Beschäftigung für bestimmte Problemgruppen am Arbeitsmarkt über gezielte **Lohnsubventionen**. Dabei gibt die BA Zuschüsse in Höhe von bis zu 50 vH der Gesamtvergütung. Nur graduelle Unterschiede zu dieser Förderung bietet ein Programm, das die Einstellung von Arbeitslosen in neu gegründete Unternehmen fördert. Mit gleichen finanziellen Leistungen durch die BA hat dieses Programm eine Laufzeit von höchstens zwölf Monaten. Als drittes Programm im Bereich der Lohnsubventionen für die Zielgruppe der Langzeitarbeitslosen schließlich wurde das Bundeshilfeprogramm eingerichtet, das einen Zuschuß von 50 bis 75 vH der Gesamtvergütung für bis zu zwölf Monate vorsieht, wenn eine Dauerbeschäftigung folgt.

Mit der Einführung des SGB III wurden zusätzlich Eingliederungsverträge eingeführt. In diesem Programm können Arbeitgeber, die Langzeitarbeitslose eingestellt haben, eine hundertprozentige Lohnsubvention für die Zeit erhalten, in denen die Einarbeitung auf den Arbeitsplatz erfolgt. Dieses Programm ist auf eine Dauer von sechs Monaten begrenzt.

Die Förderung der **Mobilität** mit dem Ziel der Aufnahme einer neuen regulären Beschäftigung bildet einen weiteren Bereich der AAMP in Deutschland. Hierbei sind vor allem zwei Programme zu nennen: (1) Mobilitätshilfen sind darauf ausgerichtet, daß die Aufnahme einer Beschäftigung in einer anderen Region erleichtert wird. Dabei erhalten Anspruchsberechtigte (im wesentlichen ehemalige Arbeitslose) wahlweise a) ein zinsloses Darlehen (80 vH des ersten Gehalts), b) einen Zuschuß zu einer notwendigen Erstausrüstung für die Aufnahmen einer Beschäftigung, c) eine Reisekostenbeihilfe für die Dauer von zwei Jahren zwischen Wohnbezirk und Arbeitsstätte oder d) eine einmalige Beihilfe zum Umzug in einen neuen Bezirk. (2) Überbrückungsgeld ist eine Hilfe für die Aufnahme einer selbständigen Beschäftigung durch ehemalige Arbeitslose. Um die erste Phase der Selbständigkeit finanziell abzusichern, können Arbeitslose für einen Zeitraum von sechs Monaten ihre Bezüge an Einkommensersatzleistungen ausweiten und damit eine Förderung ihrer Existenzgründung erhalten.

## 2.3 Umfang der AAMP

Den Umfang der Teilnahme an Maßnahmen der AAMP verdeutlicht Tabelle 1, die die Eintritte in Maßnahmen als Anteil an allen zivilen Erwerbspersonen darstellt. Für den Zeitraum 1993 bis 1998 zeigen sich hierbei auf einem in Gesamtdeutschland um 3 vH liegenden Niveau leichte Veränderungen in der Teilnahme an einzelnen Maßnahmen und deutliche regionale Unterschiede: Zunächst wird der höhere Anteil der Arbeitsmarktpolitik in Ostdeutschland deutlich, wo zwischen 7 vH der Erwerbspersonen im Jahre 1997 und 13 vH im Jahre 1998 von AAMP betroffen sind, während dieser Anteil in Westdeutschland stabil bei etwa 1,5 vH liegt. In zwei Jahren, 1994 und 1998, zeigt sich eine deutliche Ausweitung der AAMP in Ostdeutschland, während in Westdeutschland keine derartigen Schwankungen festzustellen sind. Diese stilisierten Fakten sind hinlänglich bekannt, so daß hier ergänzend nur einige weitere Tendenzen aufgezeigt werden sollen:

1. Traditionelle ABM und Strukturanpassungsmaßnahmen (SAM) spielen in Ostdeutschland nach wie vor eine große Rolle und betreffen 1998 immer noch 4,4 vH der gesamten Erwerbspersonen.
2. Gleichzeitig ist ein starker Anstieg der Teilnahme in Programmen zu beobachten, die reguläre Beschäftigung in Unternehmen subventionieren. Seitdem ab dem Jahre 1997 auch Wirtschaftsunternehmen in Ostdeutschland SAM in Anspruch nehmen können, ist dieses Instrument mit 3,4 vH der Erwerbspersonen nahezu ebenso bedeutend wie ABM und SAM.
3. Fortbildungsprogramme weisen seit 1994 eine rückläufige Teilnehmerentwicklung auf. Traten in Ostdeutschland 1994 noch 3,9 vH der Erwerbspersonen in eine Fortbildungs- oder Umschulungsmaßnahme ein, so ergibt sich 1998 nur noch ein Anteil von 2,4 vH.

In Westdeutschland ergeben sich auf einem insgesamt weit niedrigeren Niveau ähnliche Tendenzen. Auch hier zeigt sich tendenziell eine zunehmende Teilnahme in Programmen der subventionierten Beschäftigung, insbesondere der Lohnkostenzuschüsse für regulär Beschäftigte.



In einer regional differenzierenden Analyse der Teilnehmerentwicklung der AAMP zeigen Schmid et al. (2000a: 10ff.), daß der Umfang der einzelnen Programme über die Regionen in West- und Ostdeutschland erheblich variiert. Die oben dargestellten aggregierten Daten zur Maßnahmenteilnahme vermitteln mithin nur grobe Anhaltspunkte für die individuelle Wahrscheinlichkeit, ein bestimmtes Programm zu absolvieren. Insbesondere zeigt sich, daß in den einzelnen Regionen oft ein Förderungsschwerpunkt zugunsten von ABM oder FuU gelegt wird (vergl. ebd.). Regionale Variation der Maßnahmen und der extensivere Einsatz der Lohnkostenzuschüsse deuten in Richtung einer zunehmenden Heterogenisierung der Arbeitsmarktpolitik: Die Arbeitsämter können innerhalb des durch die Rechtsetzung des SGB III vorgegeben Portfolios flexibler wählen. Bei den Lohnkostenzuschüssen, die im Verhältnis zu den anderen Maßnahmen an Bedeutung gewonnen haben, haben die Arbeitsämter dabei verschiedene Optionen.

Für die Evaluation der AAMP kann angesichts des institutionellen Designs der Maßnahmen und der Teilnehmerentwicklung als Ergebnis festgehalten werden: Eine umfassende Wirkungsanalyse muß die drei prinzipiellen Programmgruppen abdecken, neben FuU und ABM/SAM also insbesondere auch den zunehmenden Einsatz von Lohnkostenzuschüssen. Da durch die Programmviefalt den Arbeitsämtern unterschiedliche Optionen gegeben werden, Problemgruppen in Beschäftigung zu bringen, kann eine simultane Evaluation heterogener Maßnahmen erforderlich sein. Wegen der regionalen Variation der Arbeitsmarktpolitik ist auch ein Unterschied zwischen Regionen hinsichtlich der Ausrichtung auf einzelne Maßnahmen zu konstatieren, welcher einen Teil der Maßnahmenheterogenität darstellt.

(Tabelle 1 etwa hier)

### 3 Methodische Aspekte des Evaluationsproblems

Es existiert bereits eine umfangreiche Literatur über methodische Aspekte der Wirkungsanalyse von AAMP, vgl. Heckman et al. (1999) als aktuellen Überblick sowie Calmfors (1994) im Hinblick auf makroökonomische Evaluationen. Dieser Abschnitt kann keinen umfassenden Überblick der Methoden geben. Stattdessen werden die Aspekte skizziert, die bei der Diskussion der empirischen Evaluationsforschung zur AAMP in Deutschland notwendig sind.

Eine vollständige Evaluation einer spezifischen Maßnahme der AAMP umfaßt vier Teile: (1) Die Definition eines Erfolgskriteriums. (2) Führt die Maßnahme zu einer kausalen Verbesserung für die an der Maßnahme teilnehmenden Individuen? (3) Rechtfertigt der individuelle Erfolg die direkten Kosten der Maßnahme? (4) Ist die Maßnahme auf der makroökonomischen Ebene erfolgreich? Die Punkte (2) und (3) definieren das mikroökonomische und der Punkt (4) das makroökonomische Evaluationsproblem. Da der unter Punkt (3) benannte Kosten-Nutzen-Aspekt in Evaluationsstudien zur AAMP in Deutschland kaum vorkommt, beschränkt sich unser methodischer Überblick auf (2) und (4).

### 3.1 Erfolgskriterien

Wir beschränken uns auf ökonomische Ergebnisvariablen wie (Wieder-)Beschäftigungschancen, Stabilität der Beschäftigung oder Verdienste, die auch als Ziele der AAMP in Deutschland genannt werden (vgl. Abschnitt 2), ohne daß wir die Bedeutung nichtökonomischer Ziele im Rahmen des politischen Prozesses in Abrede stellen wollen.<sup>1</sup> Die ökonomischen Ergebnisvariablen können auf individueller und auf aggregierter Ebene untersucht werden.

### 3.2 Das mikroökonomische Evaluationsproblem

Hierbei geht es darum, ob die Programmteilnahme einen Erfolg für die Teilnehmer hinsichtlich der relevanten Ergebnisvariable aufweist und ob dieser Erfolg die direkten Kosten des Programms rechtfertigt.<sup>2</sup> Obwohl die Berücksichtigung der Kosten für eine umfassende Bewertung notwendig ist, wurden diese jedoch bisher in der neueren deutschen Literatur nicht berücksichtigt. Dies hängt sicherlich auch damit zusammen, daß eine intensive Debatte darüber geführt wird, ob selbst bei Vernachlässigung der Kosten überhaupt ein positiver Effekt der AAMP gefunden werden kann. Wir beschränken uns daher auf letztere Fragestellung. Weiter behandeln wir den Fall einer homogenen arbeitsmarktpolitischen Maßnahme, deren Wirkung sich jedoch für die (potentiellen) Teilnehmer unterscheiden kann.<sup>3</sup>

Mikroökonomische Evaluationen bauen auf dem Ansatz der potentiellen Ergebnisse auf, d.h. der kausale Effekt einer Maßnahme für Individuum  $i$  ist die Differenz zwischen dem Wert der Ergebnisvariable  $YT_i$  bei Teilnahme und dem Wert  $YC_i$ , der sich für dasselbe Individuum bei Nichtteilnahme ergeben würde. Das Evaluationsproblem besteht darin, daß  $YT_i$  und  $YC_i$  nicht gleichzeitig für dasselbe Individuum beobachtet werden können. Um den Effekt der Maßnahmenteilnahme für die Teilnehmer abzuschätzen, muß das durchschnittliche Ergebnis bei Nichtteilnahme  $E\{YC|q = 1\}$  für die Teilnehmergruppe ( $q = 1$ ) auf Basis der Ergebnisvariablen bei Nichtteilnahme ( $q = 0$ ) geschätzt werden. Dies kann auf der Situation vor Teilnahme für die Teilnehmer ("Vorher-Nachher-Vergleich") oder auf der Situation für Nichtteilnehmer ("Kontrollgruppe") beruhen. Die Definition eines solchen Vergleichsmaßstabes beruht immer auf nicht testbaren, identifizierenden Annahmen.

Formal ergibt sich der durchschnittliche Maßnahmeeffekt für die Teilnehmergruppe ( $q = 1$ ) ("Treatment-on-the-Treated") als

---

<sup>1</sup>Berthold und Fehn (1997) argumentieren, daß die AAMP in Deutschland hauptsächlich dazu dient, die öffentliche Besorgnis hinsichtlich der Höhe der Arbeitslosigkeit zu beruhigen, ohne die wirklichen Ursachen zu bekämpfen.

<sup>2</sup>Einen Überblick der methodischen Aspekte des mikroökonomischen Evaluationsproblems im Hinblick auf die deutsche Literatur geben Fitzenberger und Prey (1998) und Lechner (1998b).

<sup>3</sup>Einige der im nächsten Abschnitt vorgestellten Evaluationsstudien berücksichtigen die Heterogenität der Maßnahmen, indem sie verschiedene Maßnahmenteilnehmer separat behandeln. Lechner (1999b) behandelt explizit den Fall heterogener Maßnahmen unter der im folgenden (Gleichung 2) dargestellten Annahme der bedingten Unabhängigkeit. Der von ihm vorgeschlagene Evaluationsansatz wurde jedoch bisher noch nicht auf die AAMP in Deutschland angewendet.

$$(1) \quad E\{YT|q = 1\} - E\{YC|q = 1\}$$

und das Evaluationsproblem erfordert eine Schätzung von  $E\{YC|q = 1\}$ , welches sich typischerweise nicht adäquat durch das durchschnittliche beobachtbare Ergebnis  $YC$  für die Nichtteilnehmer ( $q = 0$ ) schätzen läßt: Diese Selektionsverzerrung kann auf Unterschieden in beobachtbaren und unbeobachtbaren Charakteristika zwischen Teilnehmern und Nichtteilnehmern beruhen. Darüberhinaus weist die neuere methodische Literatur auf die Verzerrung durch fehlspezifizierte ökonomische Modelle der Selektionskorrektur hin (z.B. im Zusammenhang mit "Ashenfelters Tal", siehe unten) .

**Selektion hinsichtlich beobachtbarer Variablen:** Für die beobachtbaren Charakteristika  $x$  eliminiert die nicht testbare Annahme der bedingten Unabhängigkeit (ABU)

$$(2) \quad E\{YC|q = 1, x\} = E\{YC|q = 0, x\} \quad ,$$

die Selektionsverzerrung (im Mittelwert) bei gleichem  $x$ , obwohl Teilnehmer ( $q = 1$ ) und Nichtteilnehmer ( $q = 0$ ) a priori nicht vergleichbar sind. ABU ist die Basis des populären Matching-Ansatzes (Zuordnungsansatz), der den Maßnahmeerfolg als die durchschnittliche Differenz zwischen Teilnehmern und Nichtteilnehmern mit möglichst gleichen beobachtbaren Charakteristika (gematchte Stichprobe) schätzt.

Bei mehreren  $x$ -Variablen besteht das Problem, daß für einen Teilnehmer nicht unbedingt ein Nichtteilnehmer mit denselben Eigenschaften gefunden werden kann. In diesem Fall wird für jeden Teilnehmer auf möglichst ähnliche Nichtteilnehmer zurückgegriffen. Falls der Maßnahmeerfolg von den  $x$ -Variablen abhängt, unterscheidet sich der durchschnittliche Maßnahmeerfolg für die Teilnehmer von dem für die gesamte Grundgesamtheit. Zudem hängt dann der geschätzte durchschnittliche Maßnahmeerfolg für die Teilnehmer davon ab, inwieweit für alle vergleichbare Nichtteilnehmer gefunden werden können. Im folgenden sollen zwei Verfahren beschrieben werden, die in den Studien für Deutschland verwendet werden:

1. Matching auf Basis der Teilnahmeneigung ("Propensity Score Matching", PSM, – oft ergänzt um weitere  $x$ -Variablen): Versucht man Teilnehmer und Nichtteilnehmer nichtparametrisch zu matchen, dann ist es bei nicht exakt übereinstimmenden Charakteristika erforderlich, die Ähnlichkeit zwischen Teilnehmer und Nichtteilnehmern zu messen. Grundsätzlich ist dann zu unterscheiden, ob immer der beste Match verwendet wird ("Nearest-Neighbor-Matching") oder nur solche Fälle ("Caliper-Matching"), bei denen der beste Match eine hinreichende Ähnlichkeit aufweist. Der Matching-Ansatz auf Basis der Teilnahmeneigung basiert auf einem Ergebnis von Rosenbaum and Rubin (1983), das besagt, daß bei Gültigkeit der ABU für die  $x$ -Variablen die ABU auch für die Wahrscheinlichkeit der Maßnahmenteilnahme  $P(x)$  ("Propensity Score") gilt, d.h. wenn in Gleichung (2)  $x$  durch  $P(x)$  ersetzt wird. Dieses Ergebnis erlaubt es, auf Basis der eindimensionalen Teilnahmeneigung die Ähnlichkeit von Teilnehmern und Nichtteilnehmern zu bestimmen. Durch diese Dimensionsreduktion vereinfacht sich das Matching. Allerdings muß nun  $P(x)$  geschätzt werden. Meist wird in der neueren deutschen Literatur der hybride Ansatz von Lechner (1998a) verwendet, bei dem das Matching mit Hilfe einer Gewichtungsfunktion

gleichzeitig  $P(x)$  und ausgewählte (oft zeitlich variierende)  $x$ -Variablen verwendet. Die deutsche Literatur hat es bisher versäumt, bei der Schätzung der Varianz des Maßnahmeeffektes die Tatsache zu berücksichtigen, daß  $P(x)$  selbst eine geschätzte Größe ist.

2. Parametrische Regression (PR): Traditionelle ökonometrische Regressionsansätze, die bei einer Vielzahl von Variablen im Vergleich zur Stichprobengröße große praktische Vorteile aufweisen, verwenden eine parametrische Spezifikation, um den Einfluß von beobachtbaren Charakteristika auf die Ergebnisvariable zu modellieren. Implizit wird unterstellt, daß das nichtbeobachtbare Ergebnis auf der geschätzten Regressionsfunktion liegt. Diese Annahme wurde in der neueren Literatur stark dahingehend kritisiert, daß die funktionale Form meist zu restriktiv sei und daß auf Basis nicht glaubwürdiger identifizierender Annahmen die geschätzte Funktion in Bereiche extrapoliert werde, für die keine Beobachtungen vorliegen.

**Selektion hinsichtlich unbeobachtbarer Variablen:** Ein Teil der Literatur argumentiert, daß in vielen Fällen die ABU keine angemessene identifizierende Annahme darstellt, vgl. Heckman et al. (1999). Unter plausiblen Annahmen ist es möglich, daß unbeobachtbare Charakteristika (z.B. die individuelle Einschätzung des potentiellen Maßnahmeerfolges oder das Angebot einer Maßnahmenteilnahme in naher Zukunft) gleichzeitig einen Einfluß auf die Teilnahmeentscheidung und die Ergebnisvariable ausüben. In diesem Fall läßt sich kein angemessener Vergleichsmaßstab auf Basis der beobachtbaren Charakteristika konstruieren. Um diesem Problem Rechnung zu tragen, werden in der Literatur zwei Ansätze verfolgt:

1. Ökonometrische Selektionsmodelle: Ebenso wie PR bei Selektionsverzerrung durch beobachtbare Variablen, wird die Anwendung klassischer ökonometrischer Selektionsmodelle (Schätzung parametrischer Teilnahme- und Ergebnisgleichungen bei gleichzeitiger Berücksichtigung der Korrelation der Fehlerterme) sehr stark in der Literatur kritisiert, vgl. Heckman et al. (1999, Abschnitt 7.4.2). Der Hauptvorwurf besteht darin, daß meist auf Basis fehlspezifizierter und restriktiver funktionaler Formen eine Vielzahl von Schätzergebnissen möglich sind.
2. Differenz-von-Differenzen (DvD) Schätzer: Der DvD-Schätzer benötigt Paneldaten und basiert auf der Annahme zeitinvarianter, linearer Selektionseffekte. Dieser Ansatz erweitert den einfachen Vorher-Nachher-Vergleich dahingehend, daß die zeitliche Veränderung der Ergebnisvariable für den Teilnehmer der entsprechenden Veränderung für die Nichtteilnehmer im gleichen Zeitraum gegenüberzustellen ist. Dahinter steht die Beobachtung, daß sich die Arbeitsmarktsituation im Zeitablauf auch aufgrund anderer Einflußfaktoren als der Maßnahme verändern kann.

Eine Erweiterung des einfachen DvD-Schätzers stellt der sogenannte konditionale DvD-Schätzer dar, vgl. Heckman et al. (1999), bei dem zunächst hinsichtlich ihrer beobachtbaren Charakteristika vergleichbare Individuen  $i$  and  $j$  gematcht werden und dann für diesen Match die DvD in den Ergebnisvariablen den Maßnahmeerfolg schätzt. In einem Regressionsansatz kann dieser konditionale DvD-Schätzer implementiert werden, indem Dummyvariablen für eine Maßnahmenteilnahme in der Vergangenheit ("Postprogram") und der Zukunft ("Preprogram") als Regresso-

ren verwendet werden. Der konditionale DvD-Schätzer ergibt sich als die Differenz zwischen dem Koeffizienten der Postprogram-Dummy und der Preprogram-Dummy.

**Preprogram Test (PPT) und Ashenfelters Tal:** Preprogramm-Dummyvariablen werden auch als Spezifikationstests für Regressionsansätze und nichtparametrische Matching-Ansätze verwendet, da vor Teilnahme Teilnehmer und Nichtteilnehmer vergleichbar sein sollten, vgl. Heckman und Hotz (1989). Die Idee des PPT besteht darin, daß sich nach Kontrolle für alle Ursachen der Selektionsverzerrung Teilnehmer und Nichtteilnehmer vor der Teilnahme in der Ergebnisvariable YC nicht signifikant unterscheiden sollen. Ist dies nicht der Fall, dann ist es naheliegend, daß die Evaluationsmethode nicht hinreichend für Selektionsverzerrung kontrolliert. Umgekehrt ist es naheliegend, daß ein signifikanter Preprogrammeffekt – wie oben schon skizziert – im Rahmen eines DvD-Ansatzes als Schätzung der verbleibenden zeitinvarianten Selektionsverzerrung verwendet werden kann.

Bei der Interpretation des Preprogramm-Effektes besteht jedoch das Problem einer nicht adäquaten Zuordnung, vgl. Heckman et al. (1999, Abschnitt 8.4), da häufig zu beobachten ist, daß sich die Arbeitsmarktsituation von Teilnehmern kurz vor der Maßnahme verschlechtert. Dieses als Ashenfelters Tal (“Ashenfelter’s Dip”) bezeichnete Phänomen wurde in der Literatur zuerst bei den Verdiensten festgestellt, kann aber auch bei anderen Ergebnisvariablen auftreten. Da diese Verschlechterung der Arbeitsmarktsituation typischerweise mit der zukünftigen Maßnahmenteilnahme (z.B. aufgrund reduzierter Suchaktivitäten) im Zusammenhang steht, sind Teilnehmer und Nichtteilnehmer schon kurz vor der Maßnahme nicht mehr vergleichbar (Spezifikationstest) bzw. kann ein bestehender zeitinvarianter Selektionseffekt (DvD-Ansatz) nicht auf Basis des Zeitraums kurz vor der Maßnahme bestimmt werden. Angesichts dieses Problems schlagen Fitzenberger und Prey (1998, 2000) vor, einen langfristigen Preprogramm-Effekt zu schätzen, der auf einem hinreichend lange vor dem Maßnahmebeginn liegenden Zeitraum basiert.

### 3.3 Makroökonomische Effekte

Eine makroökonomische Evaluation muß über die im vorherigen Abschnitt behandelte *ceteris paribus* Analyse hinausgehen, da die AAMP die Gleichgewichtspreise und das individuelle Verhalten in der gesamten Volkswirtschaft verändern kann. Solche indirekten Effekte sind angesichts des hohen Umfangs der AAMP in Deutschland (insbesondere in Ostdeutschland) sehr wahrscheinlich. Calmfors (1994) unterscheidet Verdrängungseffekte (Teilnehmer besetzen die Arbeitsplätze, die sonst von Nichtteilnehmern besetzt worden wären), Mitnahmeeffekte (Förderung von Maßnahmen, die auch ohne Förderung durchgeführt worden wären), Substitutionseffekte (Änderung der Lohnrelationen reduziert Arbeitsnachfrage nach anderen Typen von Nichtteilnehmern) und Steuereffekte (Finanzierung der AAMP).

Um solche indirekten Effekte schätzen zu können, muß für die mögliche Endogenität der AAMP im Hinblick auf die aggregierte Arbeitsmarktsituation kontrolliert werden. Beispielsweise lassen sich Verdrängungseffekte nicht einfach dadurch bestimmen, daß man die Beschäftigung von Nichtteilnehmern auf die Zahl der Teilnehmer an arbeitsmarktpolitischen Maßnahmen regressiert, da letztere vermutlich selbst negativ von der aggregierten

Arbeitsmarktsituation abhängt. Hierdurch würden die Verdrängungseffekte überschätzt, und es sind Instrumentvariablen notwendig, um für die Endogenität der AAMP zu kontrollieren.

Die makroökonomischen Evaluationsstudien für Deutschland behandeln im wesentlichen zwei Fragestellungen: (1) Reduziert die AAMP die Lohnzurückhaltung der Gewerkschaften, da die Folgen von Lohnerhöhungen in Form erhöhter Arbeitslosigkeit von der Allgemeinheit getragen werden? Hier wird meist das Arbeitsergebnis als Schnittpunkt der Lohnsetzungsfunktion und der Arbeitsnachfrage interpretiert, vgl. Calmfors and Skedinger (1995). (2) Reduziert die AAMP den Mismatch zwischen offenen Stellen und Arbeitslosen, vgl. Bellmann und Jackman (1996)? Beide Fragestellungen erlauben es, die Wirkung der AAMP auf die strukturelle Arbeitslosigkeit von kurzfristigen Schocks zu trennen. Gleichwohl ist auch hier für die Endogenität der AAMP zu kontrollieren.

## 4 Empirische Ergebnisse

Dieser Abschnitt beschreibt die Ergebnisse der uns bekannten wissenschaftlichen Evaluationsstudien, die den im vorherigen Abschnitt skizzierten methodischen Problemen Rechnung tragen. Angesichts der inzwischen sehr umfangreichen Literatur konnten rein deskriptive Arbeiten – so informativ sie im Einzelfall auch sein mögen – nicht berücksichtigt werden. Die hier skizzierten Studien verwenden sehr unterschiedliche methodische Ansätze. Daher ist nicht zu erwarten, daß die Studien zu quantitativ gleichen Ergebnissen kommen, es sei denn die in Abschnitt 3 skizzierten methodischen Probleme wären nicht relevant und es käme dann auch nicht darauf an, wie ihnen Rechnung getragen wird, vgl. Heckman und Hotz (1989). Die gelegentlich vorgetragene Kritik an der mikro- und makroökonomischen Evaluationsforschung, daß keine einheitlichen Ergebnisse erzielt werden, greift dann zu kurz, wenn sich die Unterschiede in den Ergebnissen durch unterschiedliche identifizierende Annahmen hinsichtlich des Selektionsmechanismus bzw. der Endogenität der AAMP erklären lassen.

### 4.1 Mikroökonomische Evaluationsstudien

Die in diesem Überblick zitierten empirischen Evaluationen der Arbeitsmarktpolitik in der Bundesrepublik basieren fast ausschließlich auf Daten aus Haushalts- und Personenbefragungen, vor allem dem Sozioökonomischen Panel (GSOEP) und dem Arbeitsmarktmonitor für Ostdeutschland (AMM Ost). Beide Paneldatensätze stellen Informationen zur Bevölkerung im erwerbsfähigen Alter zur Verfügung, differieren jedoch erheblich hinsichtlich der Anzahl der repräsentativ Befragten und des Zeitraumes, für den Informationen gesammelt werden: Im GSOEP sind Informationen zusammengefaßt, die jährlich von ca. 12.000 Personen in Ost- und Westdeutschland erhoben werden. Während für Westdeutschland Informationen für den Zeitraum 1984 bis 1999 verfügbar sind, können für die Neuen Länder Erwerbstätigkeit und Arbeitslosigkeit erst für die Zeit nach 1990 verfolgt werden. Der AMM Ost ist ein Datensatz, der für den Zeitraum 1990 bis 1994 Informationen zur erwerbsfähigen Bevölkerung in den Neuen Bundesländern enthält. Mit einer Stichprobe von anfangs knapp 11.000 Personen hat der AMM Ost eine deutlich größere Personenzahl als

das GSOEP Ost (ca. 4.000 Personen), die allerdings im Längsschnitt durch eine erhebliche Panelmortalität reduziert wird: Für den gesamten Zeitraum stehen lediglich Angaben für etwa 4.000 Personen zur Verfügung.

Informationen zur Teilnahme in beruflicher Fortbildung sind in beiden Datensätzen enthalten, Angaben zur Teilnahme an ABM und SAM dagegen nur im AMM, so daß für diesen Bereich bislang nur Evaluationen für Ostdeutschland vorliegen. Die Datensätze unterscheiden sich hinsichtlich wichtiger sozioökonomischer Strukturmerkmale, die im AMM Ost weniger genau abgefragt werden. So ist im GSOEP eine Rekonstruktion der Erwerbsbiographie für den gesamten Zeitraum der Erhebung enthalten, die sich sehr gut für Abgangsratenmodelle eignet (das sogenannte "Kalendarium"), während monatsgenaue Angaben zum Erwerbsstatus aus dem AMM Ost nur mit großem Aufwand konstruiert werden können (und vermutlich mit höherem Meßfehler). Die im folgenden aufgeführten empirischen Analysen zur Wirksamkeit der AAMP in Ostdeutschland zeigen diesen Unterschied deutlich. Außer der beiden hier genannten werden nur wenige weitere Datensätze genutzt: Bender und Klose (1999) basieren ihre Auswertung auf einer 1 vH-Stichprobe der amtlichen Beschäftigtenstatistik (IAB-S), während Eichler und Lechner (1999) die Daten des Arbeitsmarktmonitors Sachsen-Anhalt (AMM SA) nutzen.

Wir beschränken uns im folgenden auf die Darstellung von Arbeiten, die die Wirksamkeit von ABM und FuU analysieren, da diese den Großteil der bisherigen empirischen Analysen ausmachen.<sup>4</sup>

#### 4.1.1 Weiterbildung in Ostdeutschland

**Datensätze und Evaluationsperioden:** Die im folgenden dargestellten Evaluationen der beruflichen Weiterbildung basieren entweder auf dem GSOEP Ost (Lechner 1998a, 1999a, Pannenberg 1995a, 1995b, Staat 1997, Hujer und Wellner 1999) oder auf dem AMM Ost (Hübler 1994, 1997, Fitzenberger und Prey, 1998, 2000, Prey 1999, Kraus, Puhani und Steiner 1999). In Analysen auf Basis beider Datensätze wird staatlich unterstützte Weiterbildung zumeist über den Bezug von Unterhaltsgeld während einer Maßnahme identifiziert, was eine leichte Einschränkung der Teilnehmergruppe mit sich bringt, da Unterhaltsgeld nicht von allen Teilnehmern bezogen wird. Im folgenden werden aber auch Analysen zitiert, die diese Abgrenzung nicht vornehmen (Hübler 1994, Pannenberg 1995a, Lechner 1999a). In der Studie von Hübler (1994) wird Vollzeitweiterbildung evaluiert, unabhängig davon, ob diese privat oder öffentlich finanziert wird. Pannenberg (1995a) und Lechner (1999a) betrachten die gesamte berufliche Weiterbildung außerhalb von Betrieben. Im Fall von Ostdeutschland kann aufgrund der hohen Reichweite der AAMP im Transformationsprozeß davon ausgegangen werden, daß der überwiegende Teil dieser Maßnahmen durch die AAMP gefördert wurde.

Die Evaluationen der Weiterbildung betreffen zudem auch unterschiedliche Zeiträume: In den Studien von Hübler (1994) und Pannenberg (1995a) werden Maßnahmen für den Zeitraum 1990–92 evaluiert, bei den meisten anderen Autoren für die Zeit 1990–94. Wegen

---

<sup>4</sup>Weitere empirische Arbeiten betreffen die Wirksamkeit des Bezugs von Überbrückungsgeld (Pfeiffer, Reize 2000, Pfeiffer 1999) und regional begrenzte Programme (z.B. die Einführung befristeter Leiharbeit in Rheinland-Pfalz zum Zweck der Reintegration, vgl. Lechner et. al 2000). Ökonometrische Evaluationen anderer, oben dargestellter Programme sind uns nicht bekannt.

der relativ geringen Stichprobengrößen beider Datensätze finden sich für den gesamten Zeitraum 1990–94 nur selten mehr als 300 Teilnehmer an Weiterbildung in den einzelnen Analysen (vor allem bei Analysen mit dem AMM Ost), und die Anzahl der Teilnehmer liegt im GSOEP in der Regel bei ca. 100 Personen.

**Methodisches Vorgehen und Erfolgskriterien:** In den meisten Arbeiten werden komplementäre Ansätze verwendet und verschiedene Spezifikationen für unterschiedliche Ergebnisvariablen implementiert. In unserem Überblick werden die empirischen Analysen hinsichtlich der Auswirkungen von AAMP auf die Beschäftigung als Ergebnisvariable dargestellt, was allerdings in vielen Arbeiten nur einen Teilaspekt darstellt.

In Evaluationsstudien, in denen die Auswirkungen der AAMP auf Löhne eingeschätzt werden, werden zumeist lineare Panelschätzungen angewandt (Hübler 1994, Pannenberg 1995a). Die Ergebnisse hinsichtlich der Beschäftigung werden entweder als Probit Schätzungen der Beschäftigung zu einem Zeitpunkt nach der Maßnahme (Hübler 1997, 1998, Fitzenberger und Prey 1998, 2000, Prey 1999) oder in einem diskreten Abgangsratemodell für die Abgänge aus Arbeitslosigkeit bzw. Programmteilnahme in Beschäftigung spezifiziert (Pannenberg 1995a, 1995b, Hujer und Wellner 1999, Kraus, Puhani und Steiner 1999).

Auch hinsichtlich der Lösung des Evaluationsproblems werden in den zitierten Arbeiten unterschiedliche Wege beschritten: Eine Kontrolle der Selektionsverzerrung aufgrund beobachtbarer Charakteristika in den Schätzungen finden sich in den Arbeiten von Pannenberg (1995a) und Kraus, Puhani und Steiner (1999), während Staat (1997) in einem zweistufigen Schätzverfahren die Teilnahmedummy durch die Teilnahmeneigung in seiner Ergebnisschätzung ersetzt. Fitzenberger und Prey (1998, 2000) und Prey (1999) kontrollieren für Selektion hinsichtlich beobachtbarer Charakteristika durch simultane Probitschätzungen für Teilnahme und Erfolgskriterium und fügen zusätzlich Dummy-Variablen für die Beschäftigung vor und nach der Maßnahme in die Schätzgleichungen ein, so daß ein DvD-Schätzer implementiert wird, der für verbliebene Selektionsverzerrung hinsichtlich unbeobachtbarer Charakteristika kontrolliert.

In vier Studien werden gematchte Stichproben verwendet, indem entweder eine Einschränkung der Kontrollgruppe nach einem Ähnlichkeitsmaß erfolgt, bis eine ausreichende Ähnlichkeit von Teilnahme und Kontrollgruppe vorliegt und der PPT keine signifikanten Unterschiede mehr zwischen Teilnehmern und Nichtteilnehmern ausweist (Hübler 1997, 1998), oder auf Basis der geschätzten Teilnahmeneigung und weiterer zeitvariabler Informationen einem Teilnehmer genau ein entsprechender Nichtteilnehmer zugeordnet wird (Hujer und Wellner 1999, Fitzenberger und Prey 1998, Lechner 1999a). Während bei Hübler (1998) und Hujer und Wellner (1999) auf Basis der gematchten Stichproben parametrische Analysen für die Ergebnisvariable durchgeführt werden, betrachten Lechner (1998a, 1999a) und Fitzenberger und Prey (1998) innerhalb der gematchten Stichprobe die durchschnittlichen Unterschiede zwischen Teilnehmern und Kontrollbeobachtungen in der Beschäftigung als Erfolgskriterium.

**Ergebnisse (vgl. Tabelle 2):** Aus den hier dargestellten empirischen Analysen der Weiterbildung in Ostdeutschland läßt sich zunächst keine klare Aussage über deren Wirksamkeit ziehen. Für die Weiterbildung außerhalb von Betrieben findet Pannenberg (1995a) positive Effekte, die jedoch mit der Länge der Maßnahme abnehmen, während bei Lechner



(1999a) kein signifikanter Effekt gefunden wird.

Bei der engeren Eingrenzung auf Maßnahmen, während der Unterhaltsgeld bezogen wurde, zeigt sich ein noch widersprüchlicheres Bild: Hier finden Fitzenberger und Prey (1998, 2000) in ihren Probit-Schätzungen positive Effekte für die Periode 1990–94, wenn die Beschäftigungsunterschiede vor der Aufnahme der Maßnahme berücksichtigt werden. Auch Pannenberg (1995b) findet eine erhöhte Abgangsrate nach Teilnahme in einer Maßnahme, während Staat (1997) mit Ausnahme von Teilnehmern im Alter von über 45 Jahren keinen signifikanten Effekt findet. In einzelnen Spezifikationen von Hübler (1997) finden sich ebenfalls positive Effekte, die jedoch in gematchten Stichproben nur noch für Männer zu beobachten sind, während Frauen eine deutlich reduzierte Beschäftigung aufweisen. In der Auswertung der gematchten Stichproben können Hujer und Wellner (1999) eine Wirksamkeit in bezug auf die Abgangsrate ebensowenig nachweisen wie Lechner (1998a, 1999a) hinsichtlich der Unterschiede in der Arbeitslosigkeit von Teilnehmern und zugeordneten Kontrollpersonen. Fitzenberger und Prey (1998) zeigen in gematchten Stichproben, daß die Unterschiede in der Beschäftigung zwischen Teilnehmern und zugeordneten Nichtteilnehmern nicht signifikant von null verschieden sind. Kraus, Puhani und Steiner (1999) erzielen unterschiedliche Ergebnisse für die zwei Perioden 1990–92 und 1993–94. Sie zeigen, daß die Teilnahme an Maßnahmen zur Weiterbildung im Zeitraum 1990–92 negative und in der späteren Periode positive Effekte auf den Übergang in Beschäftigung hatte. Hier, wie auch in einigen Spezifikationen bei Hübler (1997, 1998) wird deutlich, daß der Zeitpunkt der Maßnahmen bedeutsam für die Wirksamkeit der FuU in Ostdeutschland gewesen sein kann und insbesondere die weitreichende Implementation der Maßnahmen zu Beginn des Vereinigungsprozesses aufgrund mangelnder Zielgruppenorientierung oder des noch stattfindenden Institutionenaufbaus in den Neuen Ländern nicht zu einer Verbesserung der Arbeitsmarktsituation der Teilnehmer gegenüber Nichtteilnahme geführt hat. Zwar werden in den Evaluationsstudien (z.B. Pannenberg 1995b, Fitzenberger und Prey 1998, 2000) einzelner Autoren auch positive Effekte deutlich; diese zeigen sich aber nicht durchgängig über alle Spezifikationen. Eine Verbesserung der Beschäftigungschancen der Teilnehmer kann mit den vorhandenen Datensätzen nicht zweifelsfrei nachgewiesen werden.

(Tabelle 2 etwa hier)

#### 4.1.2 Weiterbildung in Westdeutschland

**Datensätze und Evaluationsperioden:** Ähnlich wie in Evaluationsstudien zur Weiterbildung in Ostdeutschland wird bei der Evaluation der AAMP Westdeutschlands vor allem die monatsgenaue Information zum Beschäftigungsstatus der Personen aus dem GSOEP genutzt. Hinsichtlich der Zeiträume, die in den einzelnen Evaluationsstudien betrachtet werden, variieren die hier zusammengefaßten Studien: Hujer, Maurer und Wellner (1997a, 1997b, 1998) nutzen für die Analyse den gesamten verfügbaren Zeitraum von 1984 bis 1994, während Pannenberg (1995a) Weiterbildungsmaßnahmen im Zeitraum von 1984 bis 1991 betrachtet, Prey (1997) bis 1993 und Prey (1999) von 1985 bis 1994. Die meisten Evaluationen reflektieren damit den Einsatz der Weiterbildung in einer Zeit relativ stetigen Beschäftigungswachstums. Die aktuelle Situation nach der Zunahme der Arbeitslosigkeit

seit 1993 und der Restriktion der Förderung auf Problemgruppen im Jahr 1994 wird durch die meisten Studien noch nicht abgebildet.

Der genaue Zeitraum der Weiterbildung läßt sich in GSOEP in der Regel sehr gut identifizieren. Um auszuschließen, daß bei der Weiterbildung auch nicht staatlich geförderte Maßnahmen einbezogen werden, werden in einigen Evaluationen (Prey 1997, 1999, Staat 1997, Bender und Klose 1999) die Maßnahmen durch einen zeitgleichen Bezug von Unterhaltsgeld als Teil der AAMP identifiziert. Dabei kann, wie bereits für Ostdeutschland beschrieben, von einer Unterschätzung des Anteils der staatlich geförderten Weiterbildung ausgegangen werden. Diese Einschränkung der Daten schließt damit z.B. Berufsrückkehrerinnen aus, die nach einer Zeit außerhalb des Erwerbssystems keine Ansprüche auf Unterhaltsgeld haben, und könnte zu einer Fehleinschätzung der Wirkung von Maßnahmen führen.

Die enge Definition führt jedoch auch zu einer starken Einschränkung der verfügbaren Daten auf Teilnehmerstichproben, die selten mehr als 100 Beobachtungen umfassen. In diesem Zusammenhang ist sicherlich die Entscheidung einiger Autoren zu sehen, eine Analyse für die Weiterbildung insgesamt durchzuführen (Pannenberg 1995a, Hujer, Maurer und Wellner 1997a, 1997b, 1998), wodurch sich zwar mehr Fälle ergeben, aber eine klare Abgrenzung, ob bei der Weiterbildung eine Förderung durch die Arbeitsmarktpolitik vorlag, nicht mehr möglich ist. Die Ergebnisse der empirischen Analysen für Westdeutschland, die im folgenden dargestellt werden, sollten unter dieser Einschränkung bewertet werden.

**Methodisches Vorgehen und Erfolgskriterien:** In den vorliegenden Arbeiten erfolgen mehrheitlich Schätzungen der Abgangsrate der Abgänge aus Arbeitslosigkeit nach Beschäftigung als Erfolgskriterium (Pannenberg 1995a, Hujer, Maurer und Wellner 1997a, 1997b, 1998). Einen anderen Weg geht Staat (1997), der mit einem geordneten Probitmodell die Auswirkungen der Weiterbildung auf die Dauer der Arbeitsuche und die Beschäftigungsstabilität untersucht. Wie in ihrer Evaluation der Weiterbildung in Ostdeutschland werden bei Prey (1997, 1999) Schätzungen in Form simultaner Probitmodelle für Teilnahme und Auswirkungen der Maßnahmen durchgeführt.

Zur Problematik der Selektion hinsichtlich beobachtbarer Charakteristika nutzen Hujer, Maurer und Wellner (1997a) zunächst einen Instrumentvariablenansatz, in dem die Teilnahme an einer Maßnahme in der Ergebnisschätzung durch eine in einem Probitmodell geschätzte Teilnahmeneigung ersetzt wird. Dieser Ansatz findet sich im wesentlichen auch bei Staat (1997). In der zweiten und dritten Analyse nutzen Hujer, Maurer und Wellner (1997b, 1998) hingegen einen Matchingansatz, worin auf der Basis der geschätzten Teilnahmeneigung und weiterer zeitvariabler Informationen ein "Oversampling" (Matching der Teilnehmer mit einem Teil der Nichtteilnehmer) von Teilnehmern und einer Gruppe der Kontrollbeobachtungen (1997b) oder ein striktes Zuordnen eines Teilnehmers zu einem adäquaten Nichtteilnehmer (1998) implementiert wurde. Prey (1997, 1999) kontrolliert für Selektion aufgrund beobachtbarer Charakteristika analog zu ihrer Analyse für Ostdeutschland durch simultane Probitschätzungen für Teilnahme und Beschäftigung und berechnet einen DvD-Schätzer, der auch für Selektionsverzerrung durch unbeobachtbare Charakteristika kontrolliert.

**Ergebnisse (vgl. Tabelle 3):** In Evaluationsstudien ergeben sich signifikant negative Effekte der Teilnahme an öffentlich geförderten Weiterbildungsmaßnahmen: Pannen-

berg (1995a) findet diese Verschlechterung der Arbeitsmarktsituation für alle Teilnehmer an Weiterbildungsmaßnahmen, während der Unterhaltsgeld bezogen wird, während Prey (1997) dieses Ergebnis für Männer findet, hingegen bei Frauen keine signifikanten Effekte aus Weiterbildung entstehen. In der Analyse von Staat (1997) zeigen sich für einzelne Gruppen unter den Teilnehmern positive Effekte durch Weiterbildung, insbesondere für die gering qualifizierten Teilnehmer und Frauen. Hujer, Maurer und Wellner (1997a, 1997b) zeigen mehrheitlich positive Effekte: In der ersten Analyse finden sie eine signifikant erhöhte Abgangsrate für die Teilnehmer an kurzen Maßnahmen, wohingegen Teilnehmer an langen Maßnahmen keine signifikant veränderte Abgangsrate aufweisen. In der zweiten Analyse auf Basis gematchter Stichproben können positive Maßnahmeeffekte sowohl für kurze als auch für lange Maßnahmen gefunden werden. In einer Analyse der kurz- und langfristigen Effekte von Weiterbildung kommen die Autoren schließlich zu dem Ergebnis, daß kurzfristig nach einer Maßnahme positive Effekte zu finden sind, die jedoch langfristig nicht mehr signifikant sind (Hujer, Maurer und Wellner 1998). Prey (1999) findet für die Periode 1985–94 eine langfristig verschlechterte Arbeitsmarktsituation für männliche Teilnehmer, während für Frauen kein signifikanter Effekt festzustellen ist.

Unter den bereits genannten Restriktionen der Daten und der Vergleichbarkeit der hier dargestellten empirischen Analysen ergibt sich ein heterogenes Bild der Wirkung von Weiterbildung. Die für Westdeutschland einschlägigen Schätzungen auf Basis des GSOEP zeigen für Weiterbildung insgesamt ein eher positives Bild, das sich aber für die Gruppe der Teilnehmer an staatlich geförderter Weiterbildung nicht aufrechterhalten läßt. Hier sind in der Tat eher negative Effekte einer Maßnahmenteilnahme zu beobachten.

(Tabelle 3 etwa hier)

### 4.1.3 Arbeitsbeschaffungsmaßnahmen

**Datensätze und Evaluationsperioden:** Eine Analyse der Wirksamkeit von Arbeitsbeschaffungsmaßnahmen (ABM) reduziert sich in unserem Überblick auf die Neuen Bundesländer, da für die alten Bundesländer entsprechende Individualdaten aus der Erhebung des GSOEP nicht verfügbar sind. In den Studien wird mehrheitlich der AMM Ost genutzt (Steiner und Kraus 1995, Hübler 1997, Kraus, Puhani und Steiner 1998), der den Untersuchungszeitraum auf die Periode 1990–94 eingrenzt. Die erste Evaluation zu ABM in den Neuen Bundesländern (Steiner und Kraus 1995) nutzt die ersten zwei Jahre der Erhebung. Lediglich Eichler und Lechner (1999) führen mit dem Datensatz Arbeitsmarktmonitor Sachsen–Anhalt eine Analyse für den Zeitraum 1991 bis 1997 durch, mit dem längerfristige Effekte von ABM deutlich werden können.

**Methodisches Vorgehen und Erfolgskriterien:** Steiner und Kraus (1995) und Kraus, Puhani und Steiner (1998) konzentrieren sich in ihrer Analyse auf die Auswirkungen der Teilnahme an einer ABM auf die Abgangsrate von Arbeitslosigkeit oder Programmteilnahme in Beschäftigung. Einen anderen Ansatz wählt Hübler (1997), der die Beschäftigung zu einem festen Zeitpunkt (November 1994) in Abhängigkeit von Programmteilnahme mißt und multinominale Probit- und Logitschätzungen durchführt. Eichler und Lechner (1999) evaluieren die Maßnahmen mit DvD-Schätzungen in gematchten Stichproben von Teilnehmern und Nichtteilnehmern.

Hinsichtlich der Selektionsverzerrung bilden Steiner und Kraus (1995) unterschiedliche Referenzgruppen nach beobachtbaren sozioökonomischen Merkmalen. In spezifischen Preprogrammtests werden die Unterschiede in der Beschäftigung zwischen Teilnehmern und Nichtteilnehmern überprüft; es zeigen sich keine signifikanten Unterschiede, so daß von einer ausreichenden Kontrolle für unbeobachtbare Heterogenität ausgegangen wird. Auf der Basis dieser Stichprobe werden Unterschiede in den Abgangsraten geschätzt. In der zweiten Studie (Kraus, Puhani, Steiner 1999) werden weitere Korrekturen für Selektionsverzerrung aufgrund beobachtbarer Charakteristika vorgenommen, insbesondere durch eine Schätzung der Teilnahme an einer ABM.

Hübler (1997) implementiert in seiner Analyse der ostdeutschen Arbeitsmarktpolitik eine ganze Anzahl von Verfahren, die für Selektionskorrektur aufgrund beobachtbarer und unbeobachtbarer Charakteristika kontrollieren. Der Autor nimmt eine Korrektur hinsichtlich beobachtbarer Charakteristika in der Ergebnisgleichung vor, implementiert einen PPT innerhalb einer Random Effects Probit Schätzung oder schränkt die Gruppe der Nichtteilnehmer soweit ein, daß keine Unterschiede mehr in der Beschäftigung zwischen Teilnehmern und Nichtteilnehmern vor dem Beginn des Programms bestehen, und bildet damit gematchte Stichproben (auf Basis des PPT).

Eichler und Lechner (1999) setzen einen Matchingansatz um, der über die Zuordnung aufgrund der Teilnahmeneigung und weiterer beobachtbarer Charakteristika gebildet wird und für Selektionsverzerrung aufgrund beobachtbarer Charakteristika kontrolliert. In der gematchten Stichprobe wird eine DvD-Schätzung implementiert.

**Ergebnisse (vgl. Tabelle 4):** Für den Zeitraum 1990–92 konnten Steiner und Kraus (1995) für Männer signifikant positive Beschäftigungseffekte schätzen, wenn die Maßnahme mehr als ein Jahr zurücklag. Kurzfristig ergeben sich für Teilnehmer jedoch signifikant negative Effekte. Für Frauen wird für die in dieser Studie betrachtete Periode sowohl kurz- als auch langfristig nach Ende der Maßnahme ein signifikant negativer Effekt geschätzt. In der zweiten Analyse (Kraus, Puhani und Steiner 1998) werden für beide Perioden (1990–92 und 1993–94) bei Männern ebenso wie bei Frauen negative Effekte festgestellt.

Hübler (1997) findet in verschiedenen Spezifikationen widersprüchliche Aussagen über die Wirksamkeit der ABM: Auf der Basis multinominaler Logitschätzungen für die Beschäftigung zum Zeitpunkt November 1994 zeigt sich ein positiver Maßnahmeneffekt aus ABM. Bei der Einbeziehung einer Korrektur für Selektionsverzerrung auf Basis beobachtbarer Charakteristika können weiterhin langfristig positive Effekte der Teilnahme an einer ABM nachgewiesen werden, die sich jedoch nicht mehr zeigen, wenn die Analyse auf Basis gematchter Stichproben nach dem PPT stattfindet: Danach haben Teilnehmerinnen gegenüber Nichtteilnehmerinnen keinen signifikanten Unterschied in der Beschäftigungswahrscheinlichkeit, Männer hingegen kurz- wie langfristig verringerte Beschäftigungschancen aufgrund der Programmteilnahme gegenüber der zugeordneten Kontrollbeobachtungen.

Eichler und Lechner (1999) schließen, daß ABM einen positiven Effekt auf die Teilnehmer aufweisen. Sowohl für Männer als auch für Frauen zeigt der DvD-Schätzer eine deutlich niedrigere Arbeitslosigkeit nach einer Programmteilnahme. Diesem positiven Effekt steht jedoch entgegen, daß bei einem beträchtlichen Anteil der Frauen nach einer Maßnahme ein Abgang in die Inaktivität zu beobachten ist, so daß ein eindeutig

positiver Effekt nicht nachgewiesen werden kann.

(Tabelle 4 etwa hier)

## 4.2 Makroökonomische Evaluationsstudien

Makroökonomische Analysen von Arbeitsmarktpolitik sind oft Gegenstand international vergleichender Forschung (OECD 1993, Kraft 1994, Bellmann und Jackman 1996), wobei sehr häufig die Auswirkung der Arbeitsmarktpolitik auf das Verhältnis von offenen Stellen und Arbeitslosen ("Beveridge-Kurve") oder Langzeitarbeitslosigkeit im Zentrum der Analyse steht. Hier ergeben sich jedoch zumeist nur grobe Anhaltspunkte für die Wirkung der Arbeitsmarktpolitik, da die internationale Vergleichbarkeit nur zuläßt, hoch aggregierte Indikatoren der Arbeitsmarktpolitik (z.B. die Ausgaben für AAMP als Anteil am BIP) in Wirkungsanalysen aufzunehmen. Informationen zum Design der AAMP und den unterschiedlichen Wirkungsweisen einzelner Programme werden bei diesen Analysen nach unserer Kenntnis nicht berücksichtigt, so daß keine Aussagen über die Funktionsweise einzelner Programme möglich sind. Wir konzentrieren uns in unserem Überblick von makroökonomischen Evaluationen daher auf Analysen, die aggregierte Auswirkungen der AAMP auf regionale Arbeitsmärkte in Deutschland schätzen, vgl. Tabelle 5.

Bellmann und Lehmann (1991) messen die Auswirkungen einzelner Programme der AAMP auf die Abgänge aus Kurz- und Langzeitarbeitslosigkeit für Deutschland und Großbritannien. Pannenberg und Schwarze (1996) wenden eine Lohnkurvenschätzung an, die die Auswirkungen der Weiterbildung auf die regionalen aggregierten Löhne abbildet. Büttner und Prey (1998) und Prey (1999) schätzen die Auswirkungen der AAMP auf den regionalen Mismatch und zeigen, wie regionale Ungleichgewichte durch AAMP beeinflußt werden. Schmid, Speckesser und Hilbert (2000b) schließlich schätzen die Auswirkungen der AAMP auf ihre wesentlichen Zielgruppen, vor allem Langzeitarbeitslose, und die regionalen Abgänge aus Arbeitslosigkeit.

Da die Bereitstellung der Maßnahmen der AAMP sehr stark vom Niveau und der Entwicklung der regionalen Arbeitslosigkeit abhängt, ist entscheidend, für diese Endogenität in den Schätzansätzen zu kontrollieren. Dafür werden in den einzelnen Studien unterschiedliche Ansätze verfolgt: Bellmann und Lehmann (1991), Pannenberg und Schwarze (1996) und Schmid, Speckesser und Hilbert (2000b) führen Schätzungen mit regionalen Fixed Effects aus, während Büttner und Prey (1998) und Prey (1999) neben Fixed Effect-Schätzungen auch Instrumentvariablenansätze verwenden.

In den Evaluationen werden die wesentlichen Programme, ABM und FuU, hinsichtlich ihrer Auswirkungen auf regionale Arbeitsmärkte gemessen. Pannenberg und Schwarze (1996) führen allerdings nur Schätzungen für den Einsatz der Weiterbildung durch. Neben FuU und ABM wird in den Evaluationsstudien von Schmid, Speckesser und Hilbert (2000b) und Bellmann und Lehmann (1991) auch die Wirkung von Lohnkostenzuschüsse (LKZ) auf die Zielgruppe der Langzeitarbeitslosen einbezogen.

Bellmann und Lehmann (1991) zeigen, daß die Abgänge aus regionaler Arbeitslosigkeit für verschiedene Gruppen von Kurz- oder Langzeitarbeitslosen durch den Einsatz der ABM gefördert werden können. Mit diesem Ergebnis decken sich auch die Schätzungen

von Büttner und Prey (1998), die einen auf das regionale Ungleichgewicht am Arbeitsmarkt mindernden Einfluß der ABM finden. Die Folgestudie (Prey 1999) zeigt jedoch, daß dieser Einfluß für Frauen nicht nachzuweisen ist. Die Wirkung der Weiterbildung auf den regionalen Mismatch wird für die Männer als insignifikant eingeschätzt, bei Frauen deuten die Ergebnisse auf eine erhöhte Mismatch–Arbeitslosigkeit als Folge von Weiterbildung hin.

Schmid, Speckesser und Hilbert (2000b) finden in unterschiedlichen Spezifikationen zumeist eine reduzierende Wirkung der Weiterbildung auf das regionale Niveau der Langzeitarbeitslosigkeit und die regionale Struktur der Arbeitslosigkeit. Bei ABM wird hingegen kein signifikant positiver Effekte nachgewiesen. In Hinblick auf die regionalen Abgänge aus Arbeitslosigkeit insgesamt finden sich keine positiven Effekte aufgrund des Einsatzes von Arbeitsmarktpolitik: keine der Programmvariablen zeigt positive Effekte.

(Tabelle 5 etwa hier)

## 5 Schlußfolgerungen

Der vorliegende Beitrag vermittelt einen Überblick über den Stand der wissenschaftlichen Evaluation der Arbeitsmarktpolitik in Deutschland Ende der neunziger Jahre. Dabei sind folgende Punkte deutlich geworden:

1. Die AAMP in Deutschland verfügt über einen ganzen Katalog an Maßnahmen, von denen einige gleichwertig, andere aufeinander aufbauend für die Integration von Problemgruppen in den ersten Arbeitsmarkt betrachtet werden können. Seit der Reform der Gesetzgebung haben die Arbeitsämter zudem die Freiheit gewonnen, Maßnahmen gezielter an regionalen Bedürfnissen auszurichten. Im Gegensatz zur bisherigen Vorgehensweise muß die Evaluation der AAMP daher in Zukunft die gleichzeitige Evaluation heterogener Maßnahmen und die Evaluation einzelner Maßnahmeneffekte innerhalb einer “Maßnahmenkarriere” berücksichtigen.
2. Hinsichtlich der in Evaluationsstudien angewandten Methoden wird deutlich, daß für verschiedene Aspekte der Selektionsverzerrung als dem wesentlichen Problem bei mikroökonomischer Evaluation mit nichtexperimentellen Daten in der Literatur unterschiedliche Lösungen vorgeschlagen werden: Es gibt keine “best practice” Lösung der Selektionsverzerrung aufgrund beobachtbarer Merkmale, wenn auch in vielen Fällen die geschätzte Teilnahmeneigung als Zuordnungsvariable für den besten Match zwischen Teilnehmern und Kontrollbeobachtungen verwendet wird. Zusätzlich besteht jedoch auch Selektionsverzerrung aufgrund unbeobachtbarer Charakteristika, für die in der Literatur DvD–Schätzer und verschiedene Spezifikationstests vorgeschlagen werden.
3. Der Überblick über die vorhandenen Studien zeigt folgende **Ergebnisse**:  
Hinsichtlich des Einsatzes der geförderten Weiterbildung in Ostdeutschland kann kein eindeutiges Fazit gezogen werden. In einzelnen Arbeiten werden positive Wirkungen deutlich, die sich jedoch nicht konsistent in allen Spezifikationen finden. In den meisten Evaluationsstudien ergibt sich keine signifikante Wirkung.

Für Westdeutschland sind die Ergebnisse diffuser, weil zumeist öffentliche und private Weiterbildung zusammengefaßt werden. Drei Arbeiten, die ausschließlich die Weiterbildung innerhalb der AAMP betrachten (Prey 1997, 1999 und Staat 1997), zeigen, von einer positiven Wirkung auf gering qualifizierte Teilnehmer abgesehen, negative Effekte.

Für ABM liegen bisher noch keine Evaluation für Westdeutschland vor. In Ostdeutschland werden deutlich positive Effekte nur in der Arbeit von Eichler und Lechner (1999) gefunden, drei andere Studien finden signifikant negative Effekte.

In den makroökonomischen Analysen finden sich dagegen für ABM zumeist positive Effekte auf die Verminderung struktureller Arbeitslosigkeit, während FuU die regionale Langzeitarbeitslosigkeit senkt.

Die Schlußfolgerungen, die aus diesem Vergleich in Hinblick auf die Wirksamkeit der AAMP gezogen werden, sollten allerdings mit Vorsicht interpretiert werden. Es ergeben sich insbesondere folgende Einschränkungen:

1. Wie der zweite Teil gezeigt hat, besteht die AAMP in Deutschland aus zahlreichen Maßnahmen. Von diesen waren bisher fast ausschließlich ABM und FuU Gegenstand wissenschaftlicher Evaluationen und dies auch hauptsächlich für Ostdeutschland. Für eine ganze Fülle von Lohnkostenzuschüssen, Mobilitätshilfen, Ausbildungsförderungen, usw. läßt sich bislang überhaupt keine Aussage treffen.
2. Der Überblick über die vorhandenen Evaluationsstudien zeigt, daß die Studien sich bei der Lösung von Selektionsverzerrung aufgrund beobachtbarer und unbeobachtbarer Charakteristika unterscheiden. Insbesondere die Nutzung der zeitvariablen Information etwa für Matching-Ansätze, führt zu unterschiedlichen Referenzwerten, auf denen die Schätzung des kausalen Effekts der Maßnahme für Teilnehmer basiert. Ein Teil der Unschärfe der Ergebnisse könnte als von diesen unterschiedlichen Vorgehensweisen verursacht interpretiert werden.
3. Mindestens so wichtig für die Widersprüche der empirischen Ergebnisse dürfte jedoch das zugrunde liegende Datenmaterial sein, das im wesentlichen aus sehr kleinen Ausschnitten der Umfragedaten von GSOEP und AMM besteht. Wegen der wenigen Beobachtungen können die Maßnahmen, die evaluiert werden, nicht adäquat abgebildet werden. In einer Kategorie wie Fortbildung verbergen sich daher manchmal privat finanzierte und öffentlich geförderte Maßnahmen. Und selbst wenn diese voneinander deutlich abgegrenzt sind, bleiben die Maßnahmen heterogen und bestehen aus kurzen und langen, spezifischen und allgemeinen, Auffrischungs- und Grundbildungskursen. Nach unserer Ansicht sind die politikrelevanten Folgerungen auf Basis dieser Maßnahmekategorien unklar.

Die Ergebnisse, die in diesem Überblick dargestellt wurden, relativieren sich unter den oben genannten Einschränkungen weiter: Nach unserer Einschätzung sind positive Beschäftigungseffekte von ABM und FuU allenfalls gering, was nicht mit vernachlässigbar gleichzusetzen ist. Selbst bei gegebener Datenstruktur und methodischer Vielfalt müßten sich

positive Effekte sonst deutlicher abzeichnen. Natürlich würden auch in diesem Fall die unterschiedlichen methodischen Ansätze zu quantitativ unterschiedlichen Ergebnissen kommen, es sei denn die in Abschnitt 3 skizzierten methodischen Probleme wären nicht relevant und es käme dann auch nicht darauf an, wie ihnen Rechnung getragen wird.

Die Aussage über die Wirksamkeit der in diesen Aggregaten zusammengefaßten einzelnen Maßnahmen und die Identifikation der erfolgreichsten Maßnahmen erfordert jedoch wesentlich umfangreichere und differenziertere Daten, die unserer Ansicht nach unter Mithilfe der Arbeitsverwaltung der wissenschaftlichen Evaluationsforschung bereitgestellt werden sollten. Diese Forderung wird unter anderem vom Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung geteilt, vgl. SVR (1998), Ziffer 432.

## Literatur

**Bellmann, L. und R. Jackman** (1996). The Impact of Labour Market Policy on Wages, Employment and Labour Market Mismatch. In: Schmid, G., o'Reilly, J., und K. Schömann (Hrsg.), *International Handbook of Labour Market Policy and Evaluation* (S. 725–744). Cheltenham: Edward Elgar Publishing.

**Bellmann, L. und H. Lehmann** (1991). Active Labour Market Policy in Britain and Germany and Long-term Unemployment: An Evaluation. *Paper presented at 3rd EALE Conference, Lund, September 1991.*

**Bender, S. und C. Klose** (1999). Back to Work? Comparing the Employment History of Participants and Non-Participants in Publicly Financed Off-the-Job Training. Beitrag zum Workshop Evaluation of Active Labour Market Policy in Germany, Nürnberg, Bundesanstalt für Arbeit, 28. und 29. Oktober 1999.

**Berthold, N. und R. Fehn** (1997). Aktive Arbeitsmarktpolitik – wirksames Instrument der Beschäftigungspolitik oder politische Beruhigungspille. *ORDO Jahrbuch für die Ordnung von Wirtschaft und Gesellschaft*, 48, 411–435.

**Büttner, T. und H. Prey** (1998). Does Active Labour Market Policy Affect Structural Unemployment? An Empirical Investigation for West German Regions, 1986 to 1993. *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 118, 389–413.

**Bundesanstalt für Arbeit** (1999a). *Amtliche Nachrichten der Bundesanstalt für Arbeit (ANBA).*

**Bundesanstalt für Arbeit** (1999b). *Geschäftsbericht 1998.*

**Bundesministerium für Arbeit und Sozialordnung** (1997). *Sozialgesetzbuch III, Stand 16. Dezember 1997.* Bonn: Bundesministerium für Arbeit und Sozialordnung.

**Calmfors, L.** (1994). Active labour market policy and unemployment - a framework for the analysis of crucial design features. *OECD Economic Studies*, 22(1), 7–47.

**Calmfors, L. und P. Skedinger** (1995). Does Active Labour Market Policy Increase Employment? Theoretical Considerations and Some Empirical Evidence from Sweden. *Oxford Review of Economics*, 11(1), 91–109.



- Deutscher Städtetag** (1999). [http://www.staedtetag-nw.de/staedtetag/schlagz/s\\_0803](http://www.staedtetag-nw.de/staedtetag/schlagz/s_0803). 8. März 1999.
- Eichler, M. und M. Lechner** (1999). An Evaluation of Public Employment Programs in the East German State of Sachsen–Anhalt. Unpublished manuscript, St.Gallen: University of St. Gallen.
- Fitzenberger, B. und H. Prey** (1998). Beschäftigungs- und Verdienstmöglichkeiten von Weiterbildungsteilnehmern im ostdeutschen Transformationsprozeß: Eine Methodenkritik. In: Pfeiffer, F. und W. Pohlmeier (Hrsg.), *Qualifikation, Weiterbildung und Arbeitsmarkterfolg*, volume ZEW Wirtschaftsanalysen Bd. 31 (S. 39–96). Baden-Baden: Nomos Verlagsgesellschaft.
- Fitzenberger, B. und H. Prey** (2000). Evaluating public sector sponsored training in East Germany. *Oxford Economic Papers*, im Erscheinen.
- Heckman, J. J. und V. Hotz** (1989). Choosing Among Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training. *Journal of the American Statistical Association*, 84, 862–880.
- Heckman, J. J., Ichimura, H., Smith, J. und P. Todd** (1998). Characterising Selection Bias using Experimental Data. *Econometrica*, 66(5), 1017–1098.
- Heckman, J. J., LaLonde, R. J. und J. A. Smith** (1999). The Economics and Econometrics of Active Labor Market Programs. In: Ashenfelter, O. und D. Card (Hrsg.), *Handbook of Labor Economics*, volume 3a (S. 1865–2097). Amsterdam: North Holland.
- Hübler, O.** (1994). Weiterbildung, Arbeitsplatzsuche und individuelle Beschäftigung, Eine ökonometrische Untersuchung für Ostdeutschland. *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 114, 419–447.
- Hübler, O.** (1997). Evaluation beschäftigungspolitischer Maßnahmen in Ostdeutschland. *Jahrbücher für Nationalökonomie und Statistik*, 216, 21–44.
- Hübler, O.** (1998). Betriebliche Weiterbildung und Umschulung in Ostdeutschland – Erfahrungen und Perspektiven. In: Pfeiffer, F. und W. Pohlmeier (Hrsg.), *Qualifikation, Weiterbildung und Arbeitsmarkterfolg*, volume ZEW Wirtschaftsanalysen Bd. 31 (S. 97–132). Baden-Baden: Nomos Verlagsgesellschaft.
- Hujer, R., Maurer, K. O. und M. Wellner** (1997a). Auswirkungen beruflicher Weiterbildung auf die individuelle Arbeitslosigkeitsdauer in Westdeutschland: Eine mikroökonomische Analyse. *Beitrag zur Jahrestagung des Vereins für Socialpolitik*.
- Hujer, R., Maurer, K. O. und M. Wellner** (1997b). Estimating the Effect of Training on Unemployment Duration in West Germany, A Discrete Hazard-Rate Model with Instrumental Variables. *Frankfurter Volkswirtschaftliche Diskussionsbeiträge*, Frankfurt: Johann Wolfgang Goethe Universität.
- Hujer, R., Maurer, K. O. und M. Wellner** (1998). Kurz- und langfristige Effekte von Weiterbildungsmaßnahmen auf die Arbeitslosigkeitsdauer in Westdeutschland. In: Pfeiffer, F. und W. Pohlmeier (Hrsg.), *Qualifikation, Weiterbildung und Arbeitsmarkterfolg*, volume ZEW Wirtschaftsanalysen Bd. 31 (S. 197–222). Baden-Baden: Nomos Verlagsgesellschaft.

- Hujer, R. und M. Wellner** (1999). The Effects of Public Sector Sponsored Training on Unemployment and Employment Duration in East Germany. Unpublished manuscript, Frankfurt: Johann Wolfgang Goethe Universität.
- Kraft, K.** (1994). An Evaluation of Active and Passive Labour Market Policy. *WZB Discussion Paper FS I 94-208*, Berlin: Wissenschaftszentrum Berlin für Sozialforschung.
- Kraus, F., Puhani, P. A. und V. Steiner** (1998). Do Public Works Programs Work? Some Unpleasant Results from the East German Experience. *ZEW Discussion Paper 98-07*, Mannheim: Zentrum für Europäische Wirtschaftsforschung.
- Kraus, F., Puhani, P. A. und V. Steiner** (1999). Employment Effects of Publicly Financed Training Programs, The East German Experience. *Jahrbücher für Nationalökonomie und Statistik*, 219, 216–248.
- Lechner, M.** (1998a). *Training the East German Labour Force, Microeconomic Evaluations of Continuous Vocational Training after Unification*. Heidelberg: Physica-Verlag.
- Lechner, M.** (1998b). Mikroökonomische Evaluationsstudien, Anmerkungen zu Theorie und Praxis. In: Pfeiffer, F. und W. Pohlmeier (Hrsg.), *Qualifikation, Weiterbildung und Arbeitsmarkterfolg*, volume ZEW Wirtschaftsanalysen Bd. 31 (S. 13–38). Baden-Baden: Nomos Verlagsgesellschaft.
- Lechner, M.** (1999a). Earnings and Employment Effects of Continuous Off-the-Job Training in East-Germany After Unification. *Journal of Business Economic Statistics*, 17, 74–90.
- Lechner, M.** (1999b). Identification and Estimation of causal Effects of multiple treatments under the conditional Independence Assumption. *Discussion Paper University of St. Gallen*.
- Lechner, M., Pfeiffer, F., Spengler, H. und M. Almus** (2000). The impact of Non-profit Temping Agencies on Individual Labour Market Success. In: Lechner, M. und F. Pfeiffer (Hrsg.), *Econometric Evaluations of Active Labour Market Policies in Europe*. Heidelberg: Physica (im Erscheinen).
- OECD** (1993). *Employment Outlook*. Paris: OECD.
- Pannenberg, M.** (1995a). *Weiterbildungsaktivitäten und Erwerbsbiographie, Eine empirische Analyse für Deutschland*. Frankfurt am Main, New York: Campus Verlag.
- Pannenberg, M.** (1995b). Zur Evaluation staatlicher Qualifizierungsmaßnahmen in Ostdeutschland: Das Instrument der Fortbildung und Umschulung (FuU). *Diskussionspapier Institut für Wirtschaftsforschung Halle 38*, Halle: IWH.
- Pannenberg, M. und J. Schwarze** (1996). Unemployment, Labor Market Training Programs and Regional Wages: An Extended Wage Curve Approach. *DIW-Diskussionspapiere 139*, Berlin: Deutsches Institut für Wirtschaftsforschung.
- Pfeiffer, F.** (1999). Existenzgründungspotentiale unter Arbeitssuchenden. Empirische Evidenz auf Basis des Mikrozensus. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 32(3), 300–314.

- Pfeiffer, F. und F. Reize** (2000). Business Start-ups by the Unemployed. An econometric Analysis on Firm Data. *Labour Economics*, 7 (im Erscheinen).
- Prey, H.** (1997). Beschäftigungswirkungen von öffentlich geförderten Qualifizierungsmaßnahmen, Eine Paneluntersuchung für Westdeutschland. *CILE Discussion Paper 73*, Konstanz: Center of International Labor Economics, Universität Konstanz.
- Prey, H.** (1999). *Wirkungen staatlicher Qualifizierungsmaßnahmen. Eine empirische Untersuchung für die Bundesrepublik Deutschland*. Bern, Stuttgart, Wien: Verlag Paul Haupt.
- Rosenbaum, P. R. und R. B. Rubin** (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41–55.
- Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (SVR)** (1998). *Jahresgutachten 1998/99: Vor weitreichenden Entscheidungen*. Stuttgart: Metzler-Poeschel.
- Schmid, G., Mosley, H., Hilbert, C. und H. Schütz** (2000a). Zur Effektivität aktiver Arbeitsmarktpolitik. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 32(4), 547–563.
- Schmid, G., Speckesser, S. und C. Hilbert** (2000b). Does Active Labour Market Policy Matter? An Aggregate Impact Analysis for Germany. In: de Koning, J. und H. Mosley (Hrsg.), *Labour Market Policy and Unemployment. Evaluation of Active Measures in France, Germany, The Netherlands, Spain and Sweden*, volume im Erscheinen. Cheltenham: Edward Elgar Publishing.
- Sell, S.** (1998). Entwicklung und Reform des Arbeitsförderungsgesetzes als Anpassung des Sozialrechts an flexible Erwerbsformen? *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, (3/1998), 532–533.
- Staat, M.** (1997). *Empirische Evaluation von Fortbildung und Umschulung*. Baden-Baden: Nomos Verlagsgesellschaft.
- Steiner, V. und F. Kraus** (1995). Haben Teilnehmer an Arbeitsbeschaffungsmaßnahmen in Ostdeutschland bessere Wiederbeschäftigungschancen als Arbeitslose. In: Steiner, V. und L. Bellmann (Hrsg.), *Mikroökonomik des Arbeitsmarktes*, volume Beiträge aus der Arbeitsmarkt- und Berufsforschung 192 (S. 387–423). Nürnberg: Bundesanstalt für Arbeit.

## Tabellenanhang

**Tabelle 1: Teilnahme an Maßnahmen der AAMP (in vH der zivilen Erwerbspersonen), 1993-98**

	1993	1994	1995	1996	1997	1998
Arbeitsbeschaffungsmaßnahmen (ABM) und Strukturpassungsmaßnahmen (SAM)	West 0,20 Ost 4,12 Gesamt 0,97	West 0,31 Ost 4,74 Gesamt 1,17	West 0,30 Ost 3,74 Gesamt 0,97	West 0,31 Ost 3,80 Gesamt 0,99	West 0,26 Ost 2,54 Gesamt 0,71	West 0,35 Ost 4,39 Gesamt 1,15
Lohnkostenzuschüsse für regulär Beschäftigte (einschließlich SAM-OFW)	West 0,09 Ost 0,68 Gesamt 0,20	West 0,09 Ost 0,53 Gesamt 0,18	West 0,2 Ost 0,84 Gesamt 0,32	West 0,16 Ost 0,60 Gesamt 0,25	West 0,21 Ost 1,46 Gesamt 0,46	West 0,44 Ost 3,38 Gesamt 1,02
Fortbildung	West 1,12 Ost 3,87 Gesamt 1,66	West 0,98 Ost 3,82 Gesamt 1,53	West 1,30 Ost 3,44 Gesamt 1,72	West 1,23 Ost 3,62 Gesamt 1,70	West 0,89 Ost 2,20 Gesamt 1,15	West 0,86 Ost 2,42 Gesamt 1,17
Summe	West 1,41 Ost 8,66 Gesamt 2,84	West 1,38 Ost 9,09 Gesamt 2,88	West 1,80 Ost 8,03 Gesamt 3,01	West 1,71 Ost 8,01 Gesamt 2,94	West 1,37 Ost 7,17 Gesamt 2,51	West 1,66 Ost 12,93 Gesamt 3,87

Quelle: Bundesanstalt für Arbeit 1999a: Amtliche Nachrichten der Bundesanstalt für Arbeit, Ausgabe 1, 1999: 14ff.; eigene Berechnungen

**Tabelle 2: Übersicht über mikroökonomische Evaluationen, Weiterbildung in Ostdeutschland <sup>1</sup>**

Studie	Datensatz	Evaluationsmethode und Ergebnisvariable	Ergebnis
Hübler (1994)	AMM Ost 1990	Simultane Probit Schätzungen für Teilnahme und Ergebnisvariable j) Wahrscheinlichkeit der Arbeitssuche und ii) Arbeitszeit	j) + ii) -
Pannenberg (1995a)	GSOEP Ost 1990-1992	(i) Diskretes Hazardatenmodell: Abgänge aus Arbeitslosigkeit (ii) Lineare Panelerschätzung (Fixed Effects) für Löhne	i) - ii) 0
Pannenberg (1995b)	GSOEP Ost, 1990-94	(i) Diskretes Hazardatenmodell: Abgänge aus Arbeitslosigkeit (ii) Lineare Panelerschätzung (Fixed Effects) für Löhne	i) + ii) +
Hübler (1997)	(siehe Tabelle ABM)		i) +; ii) + iii) Männer +; Frauen -
Staat (1997)	GSOEP Ost, 1992-94	Geordnetes Probit, Ersatz des Teilnahmedummies durch Teilnahmeeignung als Instrumentvariable (Probit Schätzung). Ergebnisvariable Beschäftigungsstrahlbarkeit	0
Hübler (1998)	AMM Ost, 1990-94	i) Kontr. auf beobachtb. Charakt. in einem simultanen RE Probit für Beschäftigung ii) MN-Logit f. Beschäftigung, Arbeitslosigk., Inakt. mit Teilnahmeeignung; iii) Nach PPT eingeschränkte Kontrollgruppe (genatchte Stichproben), Beschäftigung	i) - ii) Besch. 0/-; iii) Besch. -;
Fitzenberger, Prey (1998)	AMM Ost 1990-94	Simultane i) statische und ii) dynamische random effects Probit Schätzungen für Teilnahme und Ergebnisvariable. Kurz- und langfristige Auswirkungen auf Beschäftigung iii) Genatchte Stichproben (P-Score und zeitvar. Charakterist.), Unterschiede Besch.	i) kurzfr. -; langfr. 0/+ ii) kurzfr. -; langfr. 0/+ iii) kurzfr. -; langfr. 0
Lechner (1998a)	GSOEP Ost 1990-94	Matching mit Teilnahmeeignung und zeitvariablen Kovariaten; Ergebnisvariable: Unterschiede in genatchten Stichproben f. i) Arbeitslosigkeit, ii) Vollzeitbesch., iii) Löhne	i) kurzfr. +; langfr. 0 ii) kurzfr. -; langfr. 0; iii) 0
Hujer, Wellner (1999)	GSOEP Ost 1990-94	Matching mit Teilnahmeeignung und zeitvariablen Kovariaten; Gemischt proportionales Hazardatenmodell für i) Beschäftigungs- und ii) Arbeitslosigkeitsdauer nach Maßnahme	Kurze Maßnahmen i) 0; ii) 0 Lange Maßnahmen i) 0; ii) 0
Kraus, Puhani, Steiner (1999)	AMM Ost 1990-94	Diskretes Hazardatenmodell mit Kontrolle für beobachtbare Charakteristika und PPT für Übergänge in Beschäftigung in den Perioden i) 1990-92 und ii) 1993-94	i) Frauen -; Männer - ii) Frauen +; Männer +
Lechner (1999a)	GSOEP Ost 1990-94	Matching: i) Teilnahmeeignung, zus. zeitvariable Kovariaten; ii) Teilnahmeeignung auf Basis zeitvar. und inv. Variablen; Mismatch-Korrektur mit lokaler linearer Regression Ergebnisvariable: Nichtparametr. Unterschiede in Arbeitslosenquote	i) 0 ii) 0
Fitzenberger, Prey (2000)	AMM Ost, 1990-94	Simultane Random Effects Probit und Tobit Modelle für Teilnahme und Ergebnisvari- able, Preprogram-Test (PPT), Differenz-von-Differenzen-Schätzer (DvD), Ergebnisvar. Beschäftigung	+

Legende: - signifikant negativer Effekt    i) ii) iii) jeweilige Spezifikationen oder Ergebnisvariablen  
0 nicht signifikanter Effekt    MN-Logit: Multinomialer Logit  
+ signifikant positiver Effekt    PPT: Preprogram-Test

**Tabelle 3: Übersicht über mikroökonomische Evaluationen, Weiterbildung in Westdeutschland**

Studie	Datensatz	Evaluationsmethode und Ergebnisvariable	Ergebnis
Pannenberg (1995)	GSOEP West 1984-1991	Diskretes Hazardratenmodell der Abgänge aus Arbeitslosigkeit, Kontrolle für beobachtbare Charakteristika	0
Hujer, Maurer, Welher (1997a)	GSOEP West 1986-1994	Diskretes Hazardratenmodell mit unbeob. Heterogenität: Abgänge aus Arbeitslosigkeit, Ersatz des Teilnahmedummies durch Teilnahmeneigung als IV (Probit Schätzung); i) für kurze und ii) lange Maßnahmen	i) + ii) 0
Hujer, Maurer, Welher (1997b)	GSOEP West 1986-1994	Diskretes Hazardratenmodell mit unbeob. Heterogenität: Abgänge aus Arbeitslosigkeit, Matched Samples auf Basis von Teilnahmeneigung und zeitvariablen Informationen ("Oversampling"); i) für kurze und ii) lange Maßnahmen	i) + ii) 0 / +
Prey (1997)	GSOEP West 1984-1993	Simultanes Random Effects Probit-Modell mit Preprogrm-Test für Teilnahme und Beschäftigung; getrennt für i) Männer und ii) Frauen,	i) - ii) +
Staat (1997)	GSOEP West, 1984-94	Geordnetes Probit für i) Dauer der Arbeitssuche und ii) Beschäftigungsstabilität, Kontrollgr. nach Alter, Ersatz des Teilnahmedummies durch Teilnahmeneigung als IV (Probit)	i) 0 ii) 0
Hujer, Maurer, Welher (1998)	GSOEP West 1984-1994	Diskretes Hazardratenmodell mit unbeob. Heterogenität: Abgänge aus Arbeitslosigkeit, Matching auf Basis der Teilnahmeneigung und zeitvariabler Charakteristika, für i) kurz- und ii) langfristige Effekte	i) + ii) 0
Prey (1999)	GSOEP West 1984-1993	Simultane Random Effects Probit- und Tobit Modelle mit Preprogrm-Test, für langfristige Maßnahmeneffekte auf i) Beschäftigung und ii) Löhne	i) - ii) -
Bender, Klose (1999)	IAB-S 1975-1995	Matching auf Basis sozioökonomischer Charakteristika; Langfristige Maßnahmeneffekte für Beschäftigung nach Maßnahme	+

Legende: - signifikant negativer Effekt  
0 nicht signifikanter Effekt  
+ signifikant positiver Effekt

i) ii) iii) jeweilige Spezifikationen oder Ergebnisvariablen  
IV: Instrumentenvariablenansatz  
PPT: Preprogrm-Test

**Tabelle 4: Übersicht über mikroökonomische Evaluationen, ABM in Ostdeutschland**

Studie	Datensatz	Evaluationsmethode und Ergebnisvariable	Ergebnis
Steiner, Kraus 1995	AMM Ost 1990-92	Diskretes Hazardraten-Modell als geordnetes Logit-Modell mit Referenzgruppen-Ansatz für Beschäftigung und Arbeitslosigkeit; ; i) kurz- und ii) langfr. Effekte	Männer: i) 0; ii) + Frauen: i) -; ii) -
Hübler (1997)	AMM Ost 1990-94	Evaluation versch. Maßnahmen: i) MN-Logit; RE Probit mit ii) Kontrolle für beobachtb. Charakteristika, iii) Nach PPT eingeschränkte Kontrollgruppe (gematchte Stichprobe); Ergebnisvariable: Beschäftigung	i) FuU +; ABM +; ii) FuU +; ABM + iii) Männer FuU +; ABM - iii) Frauen FuU -; ABM 0
Eichler, Lechner (1999)	AMM SA 1991-97	Matching mit Teilnahmeneigung und zeitvariablen Kovariaten; Ergebnisvar. DyD-Schätzung in gematchten Stichproben zu i) kurz- und ii) langfr. Effekten auf Arbeitslosigkeit	Männer: i) -; ii) - Frauen: i) -; ii) -
Kraus, Puhani, Steiner (1999)	AMM Ost 1990-94	Übergänge nach Beschäftigung als diskretes Hazardratenmodell mit Kontrolle für beobachtbare Charakteristika für Perioden i) 1990-92 und ii) 1993-94	Männer i) -; ii) - Frauen i) -; ii) -

Legende: - signifikant negativer Effekt  
0 nicht signifikanter Effekt  
+ signifikant positiver Effekt

i) ii) iii) jeweilige Spezifikationen oder Ergebnisvariablen  
FuU: Fortbildung und Umschulung  
DyD: Differenz-vor-Differenzen

**Tabelle 5: Übersicht über makroökonomische Evaluationen**

Studie	Datensatz	Evaluationmethode und Ergebnisvariable	Ergebnis
Bellmann	Arbeitsamt-	Lineare Schätzung (FE); Hausmann-Test zur Kontrolle von	i) ABM -; FuU 0; LKZ 0
Lehmann (1991)	bezirke (West)	Endogenität; Auswirkungen auf i) Kurz- und ii) Langzeitarbeitslosigkeit	ii) ABM 0; FuU 0; LKZ 0
Pannenberg,	Arbeitsamt-	Lohnkurvenschätzung mit Kontrolle von Endogenität i) FE in OLS, ii) FGLS mit individ.	FuU i) -; ii) -
Schwarze (1996)	bezirke (Ost)	Random Effects und regionalen FE; iii) IV (2SLS) mit regionalen FE auf aggregierte Löhne	FuU iii) -
Bittner,	Planungs-	Ausw. auf Mismatch: Schätzung der Matchingeffizienz mit OLS, LSDV, 2SLS	ABM: -; FuU: 0
Prey (1998)	regionen (West)	Kontrolle für Endogenität FE für Region und Zeit	
Prey (1999)	Planungs-	zusätzlich zu Bittner, Prey (1998) Kontrolle für regionale Altersstruktur,	i) ABM -; FuU 0;
	regionen (West)	für i) Männer und ii) Frauen	i) ABM 0; FuU +;
Schmid,	Arbeitsamt-	Lineare Schätzungen mit FE für Region und Zeit; Schätzung von 1. Differenzen	i) FuU -; ABM 0; LKZ 0
Speckesser,	bezirke (West)	für Langzeitarbeitslosigkeit über i) 6 Monate, ii) über 24 Monate und iii) Abgänge	ii) FuU -; ABM 0; LKZ -
Hilbert (2000b)			iii) FuU +; ABM 0; LKZ 0

Legende: - signifikant negativer Effekt  
 0 nicht signifikanter Effekt  
 + signifikant positiver Effekt

i) ii) iii) jeweilige Spezifikationen oder Ergebnisvariablen  
 OLS: Ordinary Least Squares  
 2SLS: Two-stage least squares  
 FE: Fixed Effects  
 FGLS: Feasible Generalized Least Squares  
 LSDV: Least Squares Dummy Variables  
 IV: Instrumentvariablen-Ansatz  
 LKZ: Lohnkostenzuschüsse für regulär Beschäftigte