

Discussion Paper No. 02-13

Stand und Perspektiven der Evaluation der aktiven Arbeitsmarktpolitik in Deutschland

Bernd Fitzenberger und Reinhard Hujer

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 02-13

Stand und Perspektiven der Evaluation der aktiven Arbeitsmarktpolitik in Deutschland

Bernd Fitzenberger und Reinhard Hujer

Download this ZEW Discussion Paper from our ftp server:

<ftp://ftp.zew.de/pub/zew-docs/dp/dp0213.pdf>

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von
neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung
der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

Discussion Papers are intended to make results of ZEW research promptly available to other
economists in order to encourage discussion and suggestions for revisions. The authors are solely
responsible for the contents which do not necessarily represent the opinion of the ZEW.

Das Wichtigste in Kürze

Diese Arbeit gibt einen Überblick der Konzeption und der Evaluation der Aktiven Arbeitsmarktpolitik (AAMP) in Deutschland. Die rechtliche Grundlage der AAMP stellte von 1969 bis 1997 das Arbeitsförderungsgesetz (AFG) dar. 1998 wurde dieses durch das Sozialgesetzbuch (SGB) III abgelöst. Während das AFG noch unter Bedingungen der Vollbeschäftigung eingeführt wurde und auch eine generelle Verbesserung der Funktionsweise des Arbeitsmarktes vorsah, erfolgte im Zeitablauf eine stärkere Ausrichtung der AAMP auf die Wiedereingliederung von Problemgruppen in den Arbeitsmarkt. Das SGB III stellt die Eingliederung von Arbeitslosen in reguläre Beschäftigung in den Mittelpunkt. Obwohl das SGB III in Ansätzen eine Erfolgskontrolle der AAMP vorsieht und trotz der hohen fiskalischen Kosten (43 Mrd DM im Jahr 2001) fehlt bisher eine umfassende Evaluation der Wirkungen der AAMP unter Berücksichtigung der Kosten.

Die Arbeit stellt die grundlegenden methodischen Probleme einer aussagekräftigen Evaluation dar. Die mikroökonomische Evaluation untersucht, ob die Teilnahme an einem arbeitsmarktpolitischen Programm zu einem Erfolg im Hinblick auf individuelle Zielgrößen wie Beschäftigung oder Verdienst führt. Um den durchschnittlichen Erfolg zu bestimmen, muss geschätzt werden, wie die individuelle Zielerreichung ausgefallen wäre, hätte die Person nicht an dem Programm teilgenommen. Eine solche Schätzung erfolgt typischerweise auf Basis der Informationen der Nichtteilnehmer, wobei dem möglichen Problem der Nichtvergleichbarkeit von Teilnehmern und Nichtteilnehmern (Selektionsverzerrung) Rechnung zu tragen ist. Die Selektionsverzerrung kann auf beobachteten und nicht beobachteten Variablen beruhen. Die Arbeit skizziert die wichtigsten methodischen Ansätze, einen adäquaten Vergleichsmaßstab zur Erfolgsmessung zu schätzen. Zur umfassenden mikroökonomischen Erfolgsmessung ist der geschätzte individuelle Erfolg den Kosten der Maßnahme gegenüberzustellen. Eine makroökonomische Evaluation muss auch die Rückwirkungen der AAMP auf die Nichtteilnehmer und die Funktionsweise des Arbeitsmarktes mit einbeziehen. Außerdem ist zu berücksichtigen, dass AAMP meist gerade dann ausgeweitet wird, wenn die Arbeitsmarktprobleme zunehmen (Endogenität der AAMP).

Im weiteren Teil der Arbeit wird die Ausgestaltung und Umfang der AAMP in Deutschland nach dem AFG und dem SGB III beschrieben. Die wichtigsten Maßnahmen sind die verschiedenen Formen der Förderung der beruflichen Weiterbildung und der öffentlich geförderten Beschäftigung. Weiterhin werden neuere arbeitsmarktpolitische Ansätze und die damit einhergehenden Anstrengungen zur Evaluation ihrer Wirkungen dargestellt.

Es zeigt sich, dass auf der politischen Seite in der Vergangenheit offenbar kein großes Interesse an einer wissenschaftlichen Evaluation der gelegentlich erratischen AAMP bestand. Dementsprechend wurden der Wissenschaft keine angemessenen Datensätze für eine wissenschaftliche Evaluation der mikroökonomischen Wirkungen der AAMP zur Verfügung gestellt. Die bisher von wissenschaftlicher Seite vorgelegten Evaluationsstudien zeichnen ein eher negatives Bild der Wirkungen der AAMP. Allerdings beruhen diese Studien auf Umfragedaten, die wegen geringer Fallzahlen und zu geringer Detailinformationen keine verlässlichen Schlussfolgerungen ermöglichen. In der jüngeren Vergangenheit wächst auf politischer Seite das Interesse an einer aussagekräftigen Evaluation der AAMP (siehe Job-Aktiv-Gesetz). Allerdings besteht das Risiko, dass sich einige in jüngster Zeit von der politischen Seite unternommenen Anstrengungen zur Förderung der Evaluation als kontraproduktiv erweisen können.

Stand und Perspektiven der Evaluation der Aktiven Arbeitsmarktpolitik in Deutschland

Bernd Fitzenberger, Universität Mannheim, ZEW und IFS*

Reinhard Hujer, J. W. Goethe Universität Frankfurt, IZA Bonn*

Dezember 2001

Beitrag zum Sonderband der Perspektiven der Wirtschaftspolitik
zum Thema "Evaluation der Aktiven Arbeitsmarktpolitik"

*Wir danken Annette Bergemann, Marco Caliendo, Stefan Speckesser und Stephan Lothar Thomsen für ihre Unterstützung und für hilfreiche Kommentare. Außerdem danken wir Christian Brinkmann, Wolfgang Franz, Michael Lechner und Alexander Spermann für wertvolle Hinweise. Alle Unzulänglichkeiten dieses Beitrags liegen allein in unserer Verantwortung.

Korrespondenzadressen: Prof. Bernd Fitzenberger, Ph.D., Lehrstuhl für VWL, Universität Mannheim, D-68131 Mannheim, Tel.: 0621-1811846, FAX: 0621-1811841, E-Mail: Bernd.Fitzenberger@vwl.uni-mannheim.de
Prof. Dr. Reinhard Hujer, Lehrstuhl für Statistik und Ökonometrie, Universität Frankfurt, D-60054 Frankfurt, Tel.: 069-79828115, FAX: 069-79823673, E-Mail: hujer@wiwi.uni-frankfurt.de

“Recently, interest in evaluation has been elevated as many economies with modern welfare states have floundered, and as the costs of running welfare states have escalated.”

Heckman *et al.* (1999)

1. Zur Bedeutung der aktiven Arbeitsmarktpolitik

Trotz leicht sinkender Tendenz der Anzahl der registrierten Arbeitslosen in Deutschland in den vergangenen zwei Jahren ist die nach wie vor hohe, persistente Arbeitslosigkeit ein zentrales wirtschafts- und gesellschaftspolitisches Problem. So ist es nicht verwunderlich, dass die staatlichen Ausgaben für die verschiedenen Maßnahmen der Arbeitsmarktpolitik außerordentlich hoch sind: Die Ausgaben der Bundesanstalt für Arbeit und des Bundes betragen im Jahre 2000 97,9 Mrd. DM bzw. 28,1 Mrd. DM. Dazu kommen noch die Ausgaben anderer öffentlicher Körperschaften, wie Länder (z.B. für Kombi-Lohn-Modelle), Gemeinden (z.B. Maßnahmen zur Wiederbeschäftigung von Sozialhilfeempfängern), und des Europäischen Sozialfonds, der sich beispielsweise an dem Sofortprogramm für Jugendliche beteiligt. Für Maßnahmen der passiven Arbeitsmarktpolitik (Arbeitslosengeld, Arbeitslosenhilfe, Insolvenzgeld) wurden insgesamt 73,9 Mrd. DM und für Maßnahmen der aktiven Arbeitsmarktpolitik (AAMP), wie beispielsweise berufliche Weiterbildung oder Arbeitsbeschaffungsmaßnahmen, 43 Mrd. DM verwendet. Das trotz der enormen Aufwendungen hohe Niveau der Arbeitslosigkeit und die Budgetrestriktionen öffentlicher Haushalte begründen die dringende Notwendigkeit, die arbeitsmarktpolitischen Maßnahmen im Hinblick auf die gesetzlich formulierten Zielvorstellungen zu evaluieren (vgl. auch Sachverständigenrat (1998), Ziffer 432). Eine besondere Bedeutung erhält unter diesem Aspekt die aktive Arbeitsmarktpolitik, die als wichtigste Zielsetzung die (Wieder-)Eingliederung von Arbeitslosen in eine reguläre Beschäftigung anstrebt.

Die rechtliche Grundlage der aktiven Arbeitsmarktpolitik in der Bundesrepublik Deutschland ist das Sozialgesetzbuch III (SGB III). Während das bis 1997 gültige Arbeitsförderungsgesetz (AFG), das 1969 noch unter den Bedingungen der Vollbeschäftigung eingeführt wurde, auch eine Verbesserung der generellen Funktionsweise des Arbeitsmarktes vorsah und insbesondere einen Mangel an qualifizierten Arbeitskräften verhindern sollte (vgl. §§1 und 2, AFG), wird nun eine stärkere Ausrichtung der aktiven Arbeitsmarktpolitik auf besondere Problemgruppen, wie z.B. Langzeitarbeitslose, Ältere, Behinderte, Geringqualifizierte, angestrebt. Ein weiterer Kernpunkt der Reform war die Zusammenführung der Fach- und Ressourcenverantwortung bei der Umsetzung der Arbeitsmarktpolitik, um eine Dezentralisierung sowie eine regionale und lokale Schwerpunktsetzung zu ermöglichen. Sie soll zu einer größeren Flexibilisierung und Leistungskontrolle in der Arbeitsmarktpolitik beitragen (vgl. Sell (1998), Hagen/Steiner (2000), Fitzenberger/Speckesser (2000)).

Die wichtigsten Ziele der aktiven Arbeitsmarktpolitik lassen sich aus §7 SGB III ableiten:

- Effizienter Einsatz der Instrumente (Grundsatz der Wirtschaftlichkeit und Sparsamkeit);
- Förderung bestimmter Problemgruppen, insbesondere Langzeitarbeitslose, Geringqualifizierte, Ältere, Behinderte;
- Eingliederung von Arbeitslosen in reguläre Beschäftigung.

Soll eine Erfolgskontrolle von Maßnahmen der aktiven Arbeitsmarktpolitik im Hinblick auf diese Ziele durchgeführt werden, so können zwei grundlegende Ansätze verwendet werden. Die mikroökonomische Evaluation untersucht die Effekte der Programmteilnahme für das einzelne teilnehmende Individuum. Die makroökonomische Evaluation analysiert gesamtwirtschaftliche Auswirkungen. Insbesondere stellt sich die Frage, ob die Maßnahmen einen Netto-Gewinn für die Volkswirtschaft darstellen, oder ob sie lediglich den Teilnehmern ermöglichen, ihre relative Wettbewerbssituation gegenüber Nichtteilnehmern zu verbessern oder sogar nur ein Transferinkommen während der Teilnahme zu sichern.

Unser Beitrag beschränkt sich auf die AAMP in Deutschland. Die Beiträge von Dolton/Balfour (2002), Lechner (2002) sowie Hagen/Steiner (2002) in diesem Sonderband sowie Heckman et al. (1999) und Schmidt et al. (2001) gehen auf internationale Erfahrungen ein.

2. Methodische Aspekte der Evaluation der AAMP

Es existiert bereits eine umfangreiche Literatur über methodische Aspekte der Wirkungsanalyse von AAMP.¹ Dieser Abschnitt kann keinen umfassenden Überblick der Methoden geben. Stattdessen wird in intuitiver Weise in die wichtigsten methodischen Aspekte eingeführt, die für eine wissenschaftliche Evaluation der AAMP in Deutschland notwendig sind.

2.1. Mikroökonomische Evaluation

Eine mikroökonomische Evaluation untersucht, ob die Programmteilnahme einen Erfolg für die Teilnehmer im Hinblick auf relevante Zielvariablen aufweist und ob dieser Erfolg die direkten Kosten des Programms rechtfertigt. Zunächst muss daher die Zielsetzung der Maßnahme definiert werden, wobei hier in der öffentlichen Diskussion auch schwierig zu messende Größen wie „die Vermeidung unterwertiger Beschäftigung“ und „die Erhöhung der sozialen Stabilität“

¹Vgl. Heckman *et al.* (1999) als aktuellen international orientierten Überblick und Calmfors (1994) im Hinblick auf makroökonomische Evaluationen. Im Hinblick auf die Evaluation der AAMP in Deutschland seien hier auch die Arbeiten von Lechner (1998), Schneider *et al.* (2000), Fitzenberger/Speckesser (2000), Hagen/Steiner (2000), Hujer/Caliendo (2001) und Schmidt *et al.* (2001) genannt. Hübler (2001) gibt einen aktuellen Überblick der Methoden zur empirischen Evaluation von Politikinterventionen.

oder auch allgemeine Gerechtigkeitsziele genannt werden. Die ökonomische Analyse beschränkt sich meist auf die meßbaren Ziele Einkommen und Beschäftigungschancen. Im Hinblick auf die Zielsetzungen des AFG und des SGB III sind für Deutschland vor allem die Auswirkungen einer Maßnahmenteilnahme auf die Beschäftigungschancen und nur sekundär die Verdiensteffekte zu untersuchen. Obwohl die Kosten für eine umfassende Bewertung von zentraler Bedeutung sind, wurde dieser Aspekt in den empirischen Studien für Deutschland bislang nicht berücksichtigt. Die vorliegenden Studien analysieren, ob selbst bei Vernachlässigung der Kosten überhaupt ein positiver Effekt der Maßnahmen der aktiven Arbeitsmarktpolitik festgestellt werden kann.

Der Ansatz potenzieller Ergebnisse

Die mikroökonomische Evaluation betrachtet den Maßnahmeneffekt als die Differenz zwischen dem Wert der Ergebnisvariablen (Zielgröße) YT für einen Maßnahmenteilnehmer und dem Wert der Ergebnisvariablen YC , der sich für dieselbe Person bei Nichtteilnahme ergeben hätte. Der Ansatz potenzieller Ergebnisse („Potential Outcome Approach“) unterstellt, dass die Höhe der Ergebnisse YT und YC für eine Person gegeben sind, unabhängig davon, ob die Person tatsächlich teilnimmt. Diese Annahme besagt unter anderem, dass die Ergebnisse bei Teilnahme und Nichtteilnahme nicht davon abhängen, wieviele Personen insgesamt an den Maßnahmen teilnehmen. Rückwirkungen der Größe eines Programms auf diese Ergebnisse sollen im Rahmen von makroökonomischen Evaluationen berücksichtigt werden. Aus dem Ansatz potenzieller Ergebnisse ergibt sich, dass YC für Teilnehmer und YT für Nichtteilnehmer kontrafaktisch, d.h. nicht beobachtbar sind.

Wollte man den Maßnahmeneffekt, d.h. die Differenz $YT - YC$, für einen Maßnahmenteilnehmer bestimmen, wäre daher eine Schätzung des nicht beobachtbaren Ergebnisses bei Nichtteilnahme YC erforderlich. Umgekehrt, wenn man den Maßnahmeneffekt für einen Nichtteilnehmer prognostizieren wollte, wäre eine Schätzung des nicht beobachtbaren Ergebnisses bei Teilnahme YT notwendig. Typischerweise ist es jedoch nicht möglich, individuelle Maßnahmeneffekte sondern nur beispielsweise durchschnittliche Maßnahmeneffekte zu schätzen. Wir beschränken uns im Folgenden auf die Schätzung des durchschnittlichen Maßnahmeneffektes für die Teilnehmer. Das Grundproblem der mikroökonomischen Evaluation besteht dann darin, ein adäquates durchschnittliches YC als Vergleichsmaßstab zu ermitteln, wobei die Schätzung auf dem Ergebnis für vergleichbare Nichtteilnehmer („Kontrollgruppe“) und auf dem Ergebnis für den Teilnehmer vor Eintritt in die Maßnahme („Vorher-Nachher-Vergleich“: vor der Maßnahme sind sie Nichtteilnehmer) beruhen kann. Der Vergleichsmaßstab wird mit statistisch-ökonomischen Methoden geschätzt.

Die Evaluationsforschung unterscheidet a priori die experimentelle und die nichtexperimentelle Vorgehensweise. In einem Experiment wird in einer Gruppe von potenziellen Teilnehmern zufällig bestimmt, wer an der Maßnahme teilnimmt und wer nicht. Der durchschnittliche Maß-

nahmeneffekt ergibt aus dem Vergleich des Ergebnisses für die Teilnehmer und die Nichtteilnehmer. Zur Evaluation der AAMP in Deutschland – im Gegensatz zu den USA (Heckman *et al.* (1999)) – wurden in der Vergangenheit jedoch keine Experimente durchgeführt, so dass wir uns auf die nichtexperimentelle Vorgehensweise konzentrieren. Auf aktuelle Modellversuche, die zum Teil experimentellen Charakter haben, werden wir im Abschnitt 3.3 eingehen.

Nichtexperimentelle Evaluation und Selektionsverzerrung

Das Grundproblem der nichtexperimentellen Evaluation besteht darin, dass ein adäquater Vergleichsmaßstab zu schätzen ist. Hierbei ist dem Problem einer möglichen Selektionsverzerrung Rechnung zu tragen, das darin besteht, dass Teilnehmer gerade deshalb an einer Maßnahme teilnehmen, *weil* sie sich von Nichtteilnehmern im Hinblick auf die Zielgröße unterscheiden. Wenn Teilnehmer aus Problemgruppen am Arbeitsmarkt stammen, liegt meist eine negative Selektion vor, so dass der adäquate Vergleichsmaßstab deutlich niedriger liegt als das durchschnittliche Ergebnis für Nichtteilnehmer. Umgekehrt kann es sein, dass nur Personen an einer Maßnahme teilnehmen, weil für sie ein besonders hoher Ertrag zu erwarten ist. Dies kann eine positive Selektion ergeben, wenn in der Gruppe der potenziellen Teilnehmer (beispielsweise alle Arbeitslosen in einem Arbeitsamtsbezirk) diejenigen mit den besten Arbeitsmarktchancen durch die Arbeitsvermittler ausgewählt werden („Teilnehmercreaming“). Auch kann dann der durchschnittliche Maßnahmeneffekt für die Teilnehmer höher sein als der durchschnittliche Maßnahmeneffekt, wenn alle potenziellen Teilnehmer an der Maßnahme teilnehmen. Wenn sich die Evaluation auch auf den Vorher–Nachher–Vergleich für Teilnehmer stützt, ist schließlich zu beachten, dass die antizipierte Teilnahme an einer Maßnahme in der nahen Zukunft mit einer Verschlechterung des Arbeitsmarktergebnisses der Teilnehmer vor Beginn der Maßnahme einhergeht („Ashenfelter’s Dip“, vgl. Heckman *et al.* (1999) und Fitzenberger/Speckesser (2000)). Ein Grund hierfür kann in der verminderten Suchintensität der Arbeitslosen nach einem neuen Arbeitsplatz liegen, wenn klar ist, dass bald mit einem Maßnahmeneintritt zu rechnen ist.

Matching von Teilnehmern und Nichtteilnehmern

Die wissenschaftlichen Analysen zur mikroökonomischen Evaluation in Deutschland tragen den genannten Problemen zunehmend Rechnung. Die Schätzung des Vergleichsmaßstabes erfolgt auf Basis der Nichtteilnehmer, die mit den Teilnehmern möglichst exakt im Hinblick auf die beobachteten Charakteristika übereinstimmen, um dem Problem einer Selektionsverzerrung zu begegnen.² Prinzipiell wird dies durch konventionelle ökonometrische Modelle geleistet, die

²Diese Charakteristika sollten einen Einfluß auf die Teilnahmewahrscheinlichkeit aufweisen und idealerweise sollte die Ergebnisvariable bei Nichtteilnahme YC im Durchschnitt für alle Personen mit denselben beobachteten Charakteristika unabhängig vom Teilnehmerstatus gleich sein (Annahme der bedingten Unabhängigkeit), vgl.

Unterschiede in den beobachteten Erklärungsvariablen berücksichtigen. Allerdings besteht das Problem, dass die Approximation, wie sie beispielsweise durch eine lineare Regression erfolgt, sehr unpräzise ist, oder dass der Vergleichsmaßstab auf Basis des geschätzten Modells für die Nichtteilnehmer prognostiziert wird, obwohl die Gruppe der Nichtteilnehmer keine zu den Teilnehmern vergleichbare Personen aufweist. Eine robuste Schätzung des Vergleichsmaßstabes, die nur Nichtteilnehmer mit ähnlichen beobachteten Charakteristika berücksichtigt, kann durch nichtparametrische Verfahren („Matching–Verfahren“) erzielt werden (Heckman *et al.* (1999), Lechner (1998)). Die einfachste Variante besteht darin, als Referenz für einen Teilnehmer das Ergebnis desjenigen Nichtteilnehmers zu verwenden, der dem Teilnehmer in den beobachteten Charakteristika am ähnlichsten ist („Matching von Zwillingen“).³ In diesem Fall repräsentieren die Zwillinge die Kontrollgruppe für die Teilnehmer. Es kann dann präzise unterschieden werden, ob für Teilnehmer mit bestimmten Charakteristika überhaupt ein vergleichbarer Zwilling vorliegt; wenn dies nicht der Fall ist, dann ist eine Evaluation für Teilnehmer mit diesen Charakteristika nicht möglich. Dieses Problem liegt häufig vor, wenn nur kleine Datensätze aus Umfragen zur Verfügung stehen und wenn bestimmte Charakteristika nahezu automatisch zu einer Programmteilnahme führen. Letzteres ist bei umfassenden staatlichen Programmen gegeben, bei denen ein Rechtsanspruch besteht.⁴

Umso größer die Anzahl der beobachteten Charakteristika X beim Matchen, umso so schwieriger wird es bei einer gegebenen Gruppe von Nichtteilnehmern, noch eine in allen Dimensionen vergleichbare Person zu finden. Abhilfe für dieses Problem ergibt sich durch das Ergebnis von Rosenbaum/Rubin (1983), die zeigen, dass es ausreicht, eine Zuordnung von Teilnehmer und Nichtteilnehmer auf Basis der eindimensionalen Wahrscheinlichkeit für die Maßnahmenteilnahme $P(X)$ vorzunehmen: Im ersten Schritt wird dabei die Wahrscheinlichkeit der Maßnahmenteilnahme als Funktion der beobachteten Variablen X geschätzt. Im zweiten Schritt werden dann Teilnehmer und Nichtteilnehmer auf Basis der in der ersten Stufe geschätzten Teilnahme-wahrscheinlichkeit „gematcht“ und der Maßnahmeneffekt wird als durchschnittliche Differenz in der Ergebnisvariable geschätzt.⁵

Heckman *et al.* (1999), Lechner (1998) oder Fitzenberger/Speckesser (2000) für eine präzise Darstellung. Diese Charakteristika können sehr wohl die Ergebnisvariablen YT und YC beeinflussen und unterscheiden sich damit von ökonomischen Instrumentvariablen, vgl. Fußnote 7.

³Häufig wird unter Matching nur das „Matching von Zwillingen“ verstanden. Es existiert jedoch eine Fülle von nichtparametrischen Matchingverfahren (Heckman *et al.* (1999)).

⁴Beispielsweise hat das Jugendsofortprogramm (JUMP) einen solchen Charakter, vgl. Abschnitt 3.3. Alle Jugendlichen mit bestimmten Arbeitsmarktproblemen haben einen Rechtsanspruch auf die Maßnahme. Es ist daher sehr schwierig, geeignete Vergleichspersonen für die JUMP–Teilnehmer zu finden.

⁵Alternativ kann auf Basis der gemachten Stichprobe ein komplexeres ökonomisches Modell verwendet werden. Beispielsweise kann der Effekt auf die Wiederbeschäftigungschancen von Arbeitslosen oder das Entlassungsrisiko mit Hilfe von Hazardratenmodellen geschätzt werden, um der Dynamik in der gematchten Stichprobe Rechnung zu tragen, vgl. Studien von einem der Autoren (Hujer), die in den Übersichten Fitzenberger/Speckesser

Über die große Bedeutung der adäquaten Zuordnung (Matching) von Nichtteilnehmern mit vergleichbaren beobachteten Charakteristika besteht in der Literatur spätestens seit den Studien des Nobelpreisträgers James J. Heckman und seiner Koautoren (Heckman *et al.* (1999)) für die USA Konsens. Dissens besteht jedoch darin, ob dies im Einzelfall ausreicht (Heckman *et al.* (1999), Lechner (1998), Fitzenberger/Speckesser (2000)). Klar ist, dass eine Zuordnung auf Basis beobachteter Variablen umso eher ausreicht, je mehr Informationen (umso größer die Anzahl der Variablen und der Beobachtungen) die verfügbaren Datensätze beinhalten.⁶ Eine wichtige Ausnahme besteht darin, dass Informationen über die Teilnehmer kurz vor Beginn der Maßnahme schon durch die Antizipation der Maßnahme beeinflusst sein können (Ashenfelter's Dip) und deshalb als Matchingvariablen nicht geeignet sind (Fitzenberger/Speckesser (2000)).

Selektionsverzerrung auf Basis unbeobachteter Größen

Neben dem Problem der Selektionsverzerrung im Hinblick auf beobachtete Variablen ist jedoch auch eine Selektionsverzerrung im Hinblick auf unbeobachtete Variablen möglich. Dies ist Gegenstand der traditionellen mikroökonomischen Analyse von Problemen der Selbstselektion (Heckman *et al.* (1999)), die in der Vergangenheit jedoch auf zu starken Annahmen hinsichtlich der Spezifikation der ökonomischen Zusammenhänge beruhten. Für Deutschland wurden bisher keine Evaluationsstudien vorgelegt, die allgemeine nicht- oder semiparametrische Modelle der Selbstselektion umsetzen.⁷

Ein häufig praktizierter Ausweg ergibt sich bei Vorliegen von Paneldaten, wenn für ein Individuum der Einfluß der unbeobachteten Charakteristika auf die Zielgröße bei Nichtteilnahme YC im Zeitablauf konstant ist. Dann ermöglicht ein sogenannter konditionaler Differenz-von-Differenzen (DvD) Ansatz die Schätzung des Maßnahmeneffektes. Der DvD-Ansatz erweitert den einfachen Vorher-Nachher-Vergleich dahingehend, dass die zeitliche Veränderung der Ergebnisvariablen für die Teilnehmer der entsprechenden Veränderung für die Nichtteilnehmer im gleichen Zeitraum gegenüberzustellen ist. Der konditionale DvD-Schätzer (Heckman *et al.* (1999)) matcht zunächst Teilnehmer und Nichtteilnehmer hinsichtlich ihrer beobachteten Charakteristika und berechnet dann für diesen Match die Differenz der Differenzen in der Ergebnis-

⁶Mehr Variablen sind nur sinnvoll, wenn sie die Teilnahmewahrscheinlichkeit beeinflussen und wenn die Annahme der bedingten Unabhängigkeit erfüllt ist, siehe Fußnote 2 und die dort zitierte Literatur für eine präzise Darstellung.

⁷Ein solcher Ansatz sollte ebenso flexibel wie nichtparametrische Matchingverfahren den Einfluß der beobachteten Charakteristika berücksichtigen und keine allzu starken Annahmen machen. Klar ist, dass ein solcher Ansatz Instrumentvariablen erfordert, d.h. Variablen, die autonom die Teilnahmewahrscheinlichkeit beeinflussen, ohne dass sie einen Einfluß auf die Ergebnisvariablen aufweisen. Wie solche Modelle zu schätzen sind, ist der Gegenstand einer intensiv geführten Diskussion in der ökonomischen Fachliteratur (siehe Verweise in Heckman *et al.* (1999)).

variablen. Beobachten wir beispielsweise für die Teilnehmer die Differenz in der Beschäftigungsquote vor und nach der Maßnahme $DT = YT_n - YC_v$, wobei n für „nach der Maßnahme“ und v für „vor der Maßnahme“ steht. Dann werden Nichtteilnehmer mit vergleichbaren beobachteten Charakteristika gesucht und für diese auch die Differenz $DC = YC_n - YC_v$ zwischen den beiden Zeitpunkten berechnet. Der Maßnahmeneffekt kann nun durch die Differenz $DT - DC$ geschätzt werden, da sich konstante individuelle Charakteristika bei der Differenzen-Bildung „herauskürzen“.

2.2. Makroökonomische Evaluation

Mikroökonomische Studien sind partialanalytisch, d.h. sie untersuchen nur die (direkten) Effekte auf die teilnehmenden Individuen. Angesichts des hohen Umfangs der aktiven Arbeitsmarktpolitik in Deutschland ist es aber wahrscheinlich, dass die Maßnahmen auch (indirekte) Effekte auf die Nichtteilnehmer bzw. die gesamte Volkswirtschaft haben. Insbesondere stellt sich die Frage, ob die Maßnahmen einen Netto-Gewinn für die Volkswirtschaft darstellen, oder ob sie lediglich den Teilnehmern ermöglichen, ihre relative Wettbewerbssituation gegenüber Nichtteilnehmern zu verbessern. Calmfors (1994) unterscheidet dabei zwischen Verdrängungseffekten (Teilnehmer besetzen die Arbeitsplätze, die sonst von Nichtteilnehmern besetzt worden wären), Substitutionseffekten (Veränderung der Lohnrelationen reduziert die Nachfrage nach anderen Typen von Nichtteilnehmern), Mitnahmeeffekten (Förderung von Maßnahmen, die auch ohne Förderung durchgeführt worden wären) und Steuereffekten (Finanzierung der aktiven Arbeitsmarktpolitik über Steuern wirkt auch auf Nichtteilnehmer).

Die methodischen Probleme einer makroökonomischen Evaluation bestehen darin, die Rückwirkungen auf den gesamten Arbeitsmarkt oder die gesamte Volkswirtschaft zu identifizieren. Hier können allgemeine Gleichgewichtsmodelle oder Modelle für den gesamten (Teil-)Arbeitsmarkt verwendet werden (vgl. die in Heckman *et al.* (1999) diskutierten diesbezüglichen Arbeiten), die die Auswirkungen der AAMP auf die Struktur der Beschäftigung, den Ausgleich zwischen Angebot und Nachfrage am Arbeitsmarkt oder das Lohnverhandlungsergebnis berücksichtigen. Zum anderen ist der möglichen Endogenität im Ausmaß der AAMP Rechnung zu tragen: Wenn AAMP vor allem dann eingesetzt wird, wenn große Arbeitsmarktprobleme bestehen, dann kann das Arbeitsergebnis negativ mit der Ausgabenhöhe der AAMP korreliert sein, obwohl möglicherweise ein positiver Effekt der AAMP besteht.

3. Aktive Arbeitsmarktpolitik in Deutschland seit 1969

Die durch das AFG und das SGB III geregelte aktive Arbeitsmarktpolitik in Deutschland hat seit 1969 eine Vielzahl von Änderungen erfahren. Als Hauptursache für die „Politische Dis-

kontinuität und Instabilität“ (Lampert (1989)) in der Arbeitsmarktpolitik sind vor allem die anhaltend hohe und im langfristigen Trend steigende Arbeitslosigkeit und die damit einhergehenden budgetären Probleme zu nennen. Im folgenden wollen wir den Rahmen der AAMP in Deutschland seit 1969 skizzieren, wie sie durch die Bundesanstalt für Arbeit durchgeführt wurde. Soweit möglich, erfolgt eine Einschätzung der bisher von der Bundesanstalt für Arbeit durchgeführten Erfolgskontrolle. Schließlich wollen wir auf die in jüngster Zeit verstärkt vorgenommenen Modellversuche in der AAMP eingehen. Dieser Abschnitt kann keinen umfassenden Überblick der AAMP in Deutschland geben, aber wir können auf einer Fülle von Überblicksstudien aufbauen, vgl. Hagen/Steiner (2000) und Schneider *et al.* (2000) sowie die dort zitierten Arbeiten.

Für die ordnungspolitische Einordnung der AAMP sind folgende Aspekte von Bedeutung: Erstens besteht ein Zielkonflikt zwischen der Zielgruppenorientierung der Maßnahmen und den möglichen Mitnahmeeffekten. Je offener der Zugang in Maßnahmen ist, umso größer sind vermutlich die möglichen Mitnahmeeffekte. Wird der Zugang in eine Maßnahme stärker auf Problemgruppen am Arbeitsmarkt begrenzt, umso geringer sind im Allgemeinen die a priori-Fähigkeiten der Teilnehmer („negative Selektion“), die sich in einem niedrigeren zu erwartenden Maßnahmeneffekt niederschlagen. Zweitens sind mögliche negative Anreizeffekte der Maßnahmen auf das individuelle Verhalten zu beachten. Die Möglichkeit der Maßnahmenteilnahme, die mit einer Einkommenszahlung einhergeht, kann zu einer reduzierten Intensität der Suche nach einem Arbeitsplatz durch die Arbeitslosen führen. Auch kann AAMP den Anreiz für Arbeitsmarktakteure vermindern, die wahren Ursachen der Arbeitslosigkeit zu bekämpfen. Drittens ist zu untersuchen, welche Konsequenzen sich daraus ergeben, dass die Finanzierung der AAMP im Wesentlichen als Versicherungsleistung aus den Beiträgen der sozialversicherungspflichtigen Beschäftigten erfolgt.

3.1. Rahmen der AAMP nach AFG und SGB III

In Deutschland wird der überwiegende Teil der AAMP durch die Bundesanstalt für Arbeit (BA) organisiert. Die Konzeption des noch unter Vollbeschäftigung konzipierten AFG von 1969 unterstellt, dass eine rationale gesamtwirtschaftliche Arbeitsmarktpolitik Probleme am Arbeitsmarkt beseitigen oder sie sogar vorausschauend verhindern kann, vgl. Lampert (1989). Dies entsprach dem vorherrschenden wirtschaftspolitischen Denken am Ende der 60er Jahre, als man davon ausging, mit Hilfe der Globalsteuerung rezessive Entwicklungen und die damit einhergehende Arbeitslosigkeit schnell erfolgreich bekämpfen zu können.

Trotz der umfassenden Zielsetzung der Verbesserung der Funktionsfähigkeit des Arbeitsmarktes im AFG konnte jedoch der schubweise Anstieg der Arbeitslosigkeit seit Mitte der 70er Jahre nicht verhindert werden. Die nach dem AFG konzipierten Maßnahmen, die oft sehr großzügig ausgestaltet (man wollte unter anderem Anreize zur Teilnahme schaffen) waren, führten

regelmässig zu Finanzierungsproblemen, wenn die Arbeitslosigkeit anstieg. Zudem wurde mit steigender Arbeitslosigkeit regelmässig vor der „Gefahr des Leistungsmissbrauchs“ gewarnt. Die Novellierungen des AFG hinsichtlich der AAMP seit 1975 beinhalten meist die Abschaffung von bestimmten Maßnahmen für bestimmte Gruppen und die Reduktion von Leistungen während der Maßnahmenteilnahme. Da hier der Raum fehlt, um alle Änderungen des AFG zu beschreiben, beschränken wir uns auf zwei wichtige Beispiele. Erstens erfolgte eine kontinuierliche Reduktion der Höhe des Unterhaltsgelds (Transferzahlung an einen Teilnehmer einer notwendigen Maßnahme der beruflichen Weiterbildung), das 1975 noch 90% des letzten Nettoentgeltes betrug und das seit 1994 67% bei Teilnehmern mit Kindern und 60% bei Teilnehmern ohne Kinder beträgt. Ausserdem ergab sich bis zur Einführung des SGB III durch eine solche Maßnahme ein neuer Anspruch auf Arbeitslosengeld, da Zeiten mit Unterhaltsgeldbezug einer beitragspflichtigen Beschäftigung gleichstanden (§ 107 Nr. 5d AFG). Zweitens ist seit 1994 die Möglichkeit der sogenannten zweckmässigen Förderung der beruflichen Fortbildung („Aufstiegsförderung“), bei der Unterhaltsgeld als Darlehen gezahlt wurde, nicht mehr gegeben.

Neben restriktiven Änderungen wurden jedoch genauso regelmässig neue Instrumente der AAMP eingesetzt. Auch hierzu zwei wichtige Beispiele: Erstens wurde 1979 die Förderung nach § 41a AFG (kurze Maßnahmen zur Verbesserung der Vermittlungsaussichten) eingeführt. Diese Maßnahmen wurden in den neuen Bundesländern kurz nach der Wende sehr intensiv eingesetzt. Die Förderung nach § 41a AFG wurde 1993 abgeschafft und stattdessen war nach § 53 I AFG eine erweiterte Arbeitsberatung vorgesehen. Zweitens wurden 1993 für die neuen Bundesländer die Strukturanpassungsmaßnahmen (SAM) nach § 242h AFG als subventionierte Beschäftigung im zweiten Arbeitsmarkt mit längeren Förderlaufzeiten als übliche Arbeitsbeschaffungsmaßnahmen eingeführt. 1994 wurde ein ähnliches Programm nach § 249s AFG auch auf Westdeutschland übertragen. Schließlich wurden 1997 Strukturanpassungsmaßnahmen Ost für Wirtschaftsunternehmen (SAM OfW) als subventionierte Beschäftigung im ersten Arbeitsmarkt eingeführt, obwohl es mit den Arbeitsbeschaffungsmaßnahmen, Einarbeitungszuschüssen und anderen Formen von Lohnkostenzuschüssen bereits mehrere Maßnahmen der subventionierten Beschäftigung nach dem AFG gab.

Gegenüber dem AFG ist die Zielsetzung des seit 1998 gültigen SGB III wesentlich eingeschränkter. Es wird primär der Ausgleich am Arbeitsmarkt (§1, SGB III) angestrebt, d.h. insbesondere sollen Zeiten der Arbeitslosigkeit vermieden oder verkürzt werden (Vorrang der Vermittlung in Ausbildung und Arbeit). Problemgruppen sollen durch Beratung, berufliche Aus- und Weiterbildung und befristete Schaffung oder Teilfinanzierung von Arbeitsplätzen in reguläre Beschäftigung integriert werden, so dass insbesondere Leistungen zur Einkommenssicherung bei Arbeitslosigkeit reduziert werden. Während diese rechtliche Verengung des Zielkanons nur die Richtung fortschreibt, in die sich die Ausgestaltung der AAMP spätestens mit den 1994 umgesetzten Reformen entwickelt hat, ergeben sich mit dem SGB III zwei wichtige Änderungen im Hinblick auf die Erfolgskontrolle. Zum einen werden Flexibilitätsspielräume für die AAMP auf der regionalen und lokalen Ebene möglich und spiegelbildlich dazu sieht das SGB III eine

explizite Evaluation des Erfolges auf regionaler Ebene vor (Brinkmann (2000) und Vollkommer (2000)). 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. Nach § 11 SGB III sollen die Arbeitsämter über ihre Ermessensleistungen eine Eingliederungsbilanz erstellen, die den Mitteleinsatz und den Erfolg der Förderung festhält. Als erste quantitative Kennziffer zur Bewertung des Erfolges wird von der Bundesanstalt für Arbeit die sogenannte Verbleibsquote⁸ erhoben (Verbleib in registrierter Arbeitslosigkeit). Es wird angestrebt, einen entsprechenden Indikator zum Übergang in sozialversicherungspflichtige Beschäftigung zu konstruieren, und es ist vorgesehen, auf Basis dieser Zahlen ein Monitoringsystem für einen Leistungswettbewerb unter den regionalen Arbeitsämtern zu entwickeln. Das grundlegende methodische Problem (siehe Abschnitt 2) besteht darin, dass meßbare Erfolgsindikatoren, wie die Verbleibsquote, stark von der Selektion der Maßnahmenteilnehmer abhängen. Je weniger die Maßnahmen auf Problemgruppen konzentriert werden, umso größer ist vermutlich der gemessene Maßnahmenerfolg. Daher besteht für die nach dem SGB III nun eigenverantwortlicher handelnden Arbeitsämter ein Anreiz, Personen mit vergleichsweise guten Arbeitsmarktchancen in Maßnahmen („Teilnehmercreaming“) zu schicken, um möglichst positive Eingliederungsergebnisse zu erzielen (vgl. Vollkommer (2000) und Zängle (2000)). Um einen effizienten Leistungswettbewerb zu organisieren, müssten Eingliederungsbilanzen jedoch diesem Problem von möglichen Fehlanreizen (Selektionsproblem) Rechnung tragen.

Außer der Zunahme regionaler Flexibilität in der Implementation von AAMP sind 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.

3.2. Wichtigste Maßnahmen und deren Umfang

Insgesamt wurden im Jahre 2000 für die Maßnahmen der aktiven Arbeitsmarktpolitik 43,04 Mrd. DM aufgewendet. Dabei ist die Förderung der beruflichen Weiterbildung (FbW) die

⁸Die Verbleibsquote ist definiert als das Verhältnis von Personen, die sechs Monate nach dem individuellen Austritt aus der Maßnahme arbeitslos gemeldet sind, zu der Gesamtzahl der Austritte aus der Maßnahme.

wichtigste Form der beruflichen Bildung im Rahmen der AAMP. Die Ausgaben 2000 betragen 13,31 Mrd. DM. 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. Im Rahmen dieser Maßnahmen werden Weiterbildungskosten übernommen und Unterhaltsgeld in Höhe von 60% (67% bei mindestens einem Kind) des letzten Nettoarbeitsentgelts gezahlt. Dieses Programm zielt darauf ab, Arbeitslose wiederzubeschäftigen und drohende Arbeitslosigkeit bei Beschäftigung abzuwenden. Das AFG sah als Fortbildung auch den Einarbeitungszuschuss als Lohnsubvention an den Arbeitgeber bei Arbeitnehmern vor. Diese Förderung ist im SGB III eine Form der Eingliederungszuschüsse nach §§ 217–224.

Das wichtigste Programm im Bereich der subventionierten Beschäftigung sind Arbeitsbeschaffungsmaßnahmen (ABM). Im Jahr 2000 wurden dafür insgesamt 7,20 Mrd. DM aufgewendet. Sie ermöglichen eine befristete Beschäftigung von Langzeitarbeitslosen in Projekten, die im öffentlichen Interesse liegen und zusätzlich sind, d.h. ohne die Subvention nicht durchgeführt werden können. Die Zuschüsse betragen zwischen 30% und 75% bezogen auf 80% des tariflichen Arbeitsentgelts für entsprechende ungeforderte Arbeit. In Ausnahmefällen können auch weitere Leistungen in Anspruch genommen werden, so dass diese Förderungsrichtlinie nur eine grobe Abgrenzung ist.

Vor allem in Ostdeutschland spielen neben ABM die Strukturanpassungsmaßnahmen (SAM) eine erhebliche Rolle. In Ostdeutschland wurden im Jahr 2000 insgesamt 2,39 Mrd. DM ausgegeben. Ebenso wie ABM zielen traditionelle SAM auf eine Integration über den „zweiten Arbeitsmarkt“, allerdings liegen andere Förderungsgrundsätze vor: Zielgruppen für SAM sind neben Arbeitslosen auch Beschäftigte, die von Arbeitslosigkeit bedroht sind. Unter SAM können nur Projekte gefördert werden, die der Verbesserung der regionalen Infrastruktur dienen. Seit Einführung des SGB III⁹ können SAM auch von Wirtschaftsunternehmen in Ostdeutschland (SAM OfW) in Anspruch genommen werden. Im Gegensatz zu traditionellen SAM zielen SAM OfW auf die direkte Integration in den ersten Arbeitsmarkt. Neben der Förderung des zweiten Arbeitsmarkts (und der SAM OfW) unterstützt die BA auch die Aufnahme einer regulären Beschäftigung für bestimmte Problemgruppen am Arbeitsmarkt über gezielte Lohnsubventionen in Form von Eingliederungszuschüssen und Beschäftigungshilfen für Langzeitarbeitslose.

Mit dem SGB III wurden zusätzlich Eingliederungsverträge eingeführt. In diesem Programm können Arbeitgeber, die Langzeitarbeitslose eingestellt haben, eine hundertprozentige Lohnsubvention für die Zeit erhalten, in der 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

⁹Mit Verabschiedung des Arbeitsförderungsreformgesetzes (AFRG) vom 1. April 1997 wurden SAM OfW unter dem Namen Produktive Lohnkostenzuschüsse Ost für Wirtschaftsunternehmen (§249 Abs. 4b AFG) und der Eingliederungsvertrag (§54a AFG) in das AFG aufgenommen.

Tabelle 1: Teilnahme an Maßnahmen der AAMP (Bestände in vH der zivilen Erwerbspersonen)

		1970	1975	1980	1985	1990	1993	1995	1998	1999
Berufliche	West	0,37	0,70	0,63	0,85	1,15	1,12	0,98	0,60	0,66
Weiterbildung	Ost	–	–	–	–	–	5,17	3,44	1,85	1,76
ABM¹ und SAM²	West	0,01	0,06	0,15	0,30	0,27	0,16	0,23	0,21	0,27
	Ost	–	–	–	–	–	3,53	4,20	2,43	2,66
EGZ/LKZ³	West	–	–	–	–	0,08	0,15	0,14	0,23	0,29
	Ost	–	–	–	–	–	0,16	0,47	1,83	2,09
Überbrückungs-	West	–	–	–	–	0,04	0,04	0,15	0,21	0,20
geld (Eintritte)⁴	Ost	–	–	–	–	–	0,19	0,32	0,39	0,39

¹ Arbeitsbeschaffungsmaßnahmen; ² Strukturanpassungsmaßnahmen (ohne SAM-OfW); ³ Eingliederungs- und Lohnkostenzuschüsse für regulär Beschäftigte (einschließlich SAM-OfW); ⁴ Eintritte in vH der zivilen Erwerbspersonen, da Bestände nicht verfügbar.

Quelle: Amtliche Nachrichten der Bundesanstalt für Arbeit, verschiedene Jahrgänge; eigene Berechnungen.

bildet einen weiteren Bereich der AAMP in Deutschland. Hierbei sind vor allem zwei Programme zu nennen: Mobilitätshilfen sind darauf ausgerichtet, dass die Aufnahme einer Beschäftigung in einer anderen Region erleichtert wird. Überbrückungsgeld ist eine Hilfe für die Aufnahme einer selbständigen Beschäftigung durch ehemalige Arbeitslose.

Tabelle 1 beschreibt die Entwicklung der durchschnittlichen Bestände in den wichtigsten Maßnahmen der AAMP als Anteil an allen zivilen Erwerbspersonen für den Zeitraum 1970 bis 1999. Es ergeben sich deutliche Veränderungen in Umfang und Struktur der AAMP. Für Westdeutschland zeigt sich ein deutlicher Anstieg der Bedeutung der AAMP, der mit dem Anstieg der Arbeitslosigkeit einhergeht. In den 70er und 80er Jahren dominieren die Maßnahmen der beruflichen Weiterbildung, was der ursprünglichen Konzeption des AFG entspricht. Im Zeitablauf steigt jedoch absolut und relativ die Bedeutung der subventionierten Beschäftigung und in den 90er Jahren geht die Bedeutung der beruflichen Weiterbildung sogar absolut zurück. Hier zeigt sich die Verschiebung hin zur direkten Förderung der Beschäftigung, die auch im Wechsel vom AFG zum SGB III deutlich wird.

Für Ostdeutschland zeigt sich die ungleich höhere Bedeutung der AAMP. Das Ausmaß war zu Beginn der 90er Jahre am stärksten, und auch dort dominierte zunächst die Förderung der beruflichen Weiterbildung mit dem Ziel einer Anpassung der Qualifikationen an das westliche Niveau. Ende der 90er Jahre weisen jedoch die Formen subventionierter Beschäftigung im ersten und zweiten Arbeitsmarkt einen wesentlich höheren Anteil auf.

Als wichtigste Form der Mobilitätsförderung hat sich in den 90er Jahren das Überbrückungsgeld zur Förderung der Selbständigkeit aus Arbeitslosigkeit entwickelt. Für weitere Mobilitätshilfen liegen uns zwar keine Teilnehmerzahlen vor, aber Informationen über die Ausgaben aus den Geschäftsberichten der Bundesanstalt für Arbeit. Danach erhielten Arbeitslose, die eine versicherungspflichtige Beschäftigung aufgenommen haben, Mobilitätshilfen (§§ 53 und 54 SGB III) im Jahre 2000 in Höhe von 39 Mill. DM in Westdeutschland und 95 Mill. DM in Ostdeutschland, was nur 0,32% der gesamten Ausgaben der aktiven Arbeitsförderung der Bundesanstalt für Arbeit entspricht.

3.3. Neuere arbeitsmarktpolitische Ansätze

Nach dem Regierungswechsel in Deutschland 1998 setzte die neue Bundesregierung mit Beginn des Jahres 1999 das „Sofortprogramm der Bundesregierung zum Abbau der Jugendarbeitslosigkeit“ – auch JUMP (\equiv „Jugend mit Perspektive“) genannt – um (Dietrich 2001). Vorbild war der sogenannte „New Deal“ der britischen Labour-Regierung (Beitrag von Dolton/Balfour (2002) in diesem Sonderband). Anlaß für die Einführung von JUMP war die Feststellung, dass Mitte der 90er Jahre die in Deutschland im internationalen Vergleich niedrige Arbeitslosenquote von Jugendlichen unter 25 Jahren insbesondere unter den Jugendlichen ohne abgeschlossene Berufsausbildung und in den neuen Bundesländern angestiegen war. Dies ging einher mit einer etwas abnehmenden Bedeutung des dualen Ausbildungssystems.

Ähnlich dem britischen Vorbild besteht JUMP aus verschiedenen Teilprogrammen: Weiterbildungsmaßnahmen, Förderung der Ausbildung in Betrieben und überbetrieblichen Einrichtungen, verstärkte Beratung und soziale Betreuung, Lohnkostenzuschüsse, Qualifizierungs-Arbeitsbeschaffungsmaßnahmen sowie Nachholen des allgemeinen Schulabschlusses. Fast alle diese Maßnahmen gab es schon nach dem AFG und dem SGB III. Daher stellt JUMP im Wesentlichen eine Bündelung von Beratung und Maßnahmenteilnahme auf eine bestimmte Zielgruppe dar, für die der Gesetzgeber schnell entsprechend hohe finanzielle Mittel bereitstellte.

In den Jahren 1999 und 2000 wurden insgesamt 268.205 Personen mit 308.112 Maßnahmen eintritten gefördert, und die Ausgaben für das Programm betragen ca. 3,7 Mrd. DM. JUMP verstärkt den überproportionalen Einsatz der AAMP in den neuen Bundesländern, und im Zeitablauf erfolgte eine stärkere Gewichtung beschäftigungsbezogener Maßnahmen gegenüber qualifizierungsorientierten Maßnahmen (Dietrich (2001), S. 4 und S. 26).

Im Hinblick auf die Evaluation der AAMP ist bemerkenswert, dass der Gesetzgeber – wie bei der Einführung der Eingliederungsbilanzen im SGB III – eine wissenschaftliche Erfolgskontrolle des Programmes vorgesehen hat. Diese Absicht wurde dadurch umgesetzt, dass das Institut für Arbeitsmarkt- und Berufsforschung mit der Erforschung des Maßnahmen Erfolges beauftragt wurde und hierzu detaillierte prozessproduzierte Daten über alle JUMP-Teilnehmer während

der Maßnahme verfügbar sind (Dietrich (2001)). Für eine aussagekräftige Evaluation müssen jedoch auch Informationen nach Teilnahme und entsprechend valide Informationen für Nichtteilnehmer vorliegen. Diese werden für 5000 Teilnehmer durch mehrere telefonische Befragungen ermittelt und entsprechend ist auch geplant, zufällig eine Stichprobe von vergleichbaren Nichtteilnehmern zu befragen. Dietrich (2001) kann nur eine Analyse zum Verbleib der befragten Teilnehmer vorlegen. Es zeigt sich, dass sechs Monate nach Abschluß der Maßnahme 24,5% arbeitslos, 24,1% beschäftigt, 19,4% in einer schulischen oder betrieblichen Ausbildung und 21,6% an weiteren Maßnahmen der AAMP teilnehmen. Ohne Kenntnis über den Status der Teilnehmer zum gleichen Zeitpunkt bei Nichtteilnahme ist jedoch keine Aussage über Erfolg oder Mißerfolg von JUMP möglich. Es ist weiter fraglich, ob sich durch Interviews eine vergleichbare Kontrollgruppe im gleichen Zeitraum finden läßt, da JUMP prinzipiell allen Jugendlichen in Problemgruppen offenstand. Zudem stehen die Teilnehmer typischerweise am Anfang ihres Erwerbslebens, so dass es nicht möglich sein wird, die Evaluation auf Basis eines individuellen Vorher-Nachher-Vergleichs durchzuführen.

Neben JUMP sind als neuere Ansätze die in einzelnen Regionen versuchsweise umgesetzten Projekte zur Stärkung der Arbeitsanreize zu nennen, bei denen eine wissenschaftliche Begleitforschung von politischer Seite vorgesehen ist. Angesichts der zunehmenden Arbeitslosigkeit im Bereich niedrig qualifizierter Arbeit in den 90er Jahren besteht die Befürchtung, dass die Arbeitsanreize für Langzeitarbeitslose und Sozialhilfeempfänger zu gering sind. Es gibt Bereiche des Transfersystems, bei der die impliziten Grenzsteuersätze im Niedriglohnbereich, die sich aus der Anrechnung von Arbeitseinkommen auf Transferzahlungen ergeben, nahe oder exakt 100% sind. Als neuere Ansätze sind hier zum einen die Modellansätze zur Förderung der Beschäftigung von Geringqualifizierten und Langzeitarbeitslosen, kurz CAST genannt (Bittner *et al.* (2001)), und zum anderen das Projekt zum Einstiegsgeld für Langzeitarbeitslose zu nennen, vgl. Dann *et al.* (2001), Spermann (2001) und Schneider *et al.* (2000).

Das CAST-Programm, das unter anderem auf einen Beschluß des „Bündnisses für Arbeit, Ausbildung und Wettbewerbsfähigkeit“ zurückgeht, besteht aus zwei Modellansätzen. Nach dem „Mainzer Modell“ (MZM), das in einigen Arbeitsamtsbezirken in Rheinland-Pfalz und Brandenburg umgesetzt wird, und nach dem „Modell der Saar-Gemeinschaftsinitiative“ (SGI), das im Saarland und im Arbeitsamtsbezirk Chemnitz umgesetzt wird, erfolgt eine Subventionierung der Sozialversicherungsbeiträge bei geringverdienenden Arbeitnehmern, die zuvor arbeitslos waren (oder geringfügig beschäftigt waren – nur MZM). SGI sieht darüber hinaus eine Möglichkeit der Weiterqualifikation vor. Obwohl eine wissenschaftliche Begleitforschung vorgesehen ist, kann nach unserem Kenntnisstand bei den CAST-Modellprojekten keine zufällige Auswahl der Programmteilnehmer unter den potenziellen Teilnehmern in den betreffenden Regionen vorgenommen werden. Deshalb müsste für eine aussagekräftige Evaluation eine Kontrollgruppe lediglich auf Basis nichtexperimenteller Methoden gebildet werden, wobei die Nichtteilnehmer in der Kontrollgruppe aus dem gleichen regionalen Arbeitsmarkt stammen sollten, vgl. Heckman *et al.* (1999). Im September 2001 wurde von Bittner *et al.* (2001) der 1. Zwischenbericht zu den

CAST-Projekten vorgelegt. Eine aussagekräftige Evaluation ist zum jetzigen Zeitpunkt noch nicht möglich, aber die Zahl der bisher 592 Förderfälle (Stand September 2001) blieb hinter den Erwartungen zurück. Von den Förderfällen entfielen 87% auf MZM und nur 13% auf SGI. Bittner *et al.* (2001) streben eine Evaluation der CAST-Programme auf Basis von Vergleichsregionen an. Die Evaluation muss möglichen Veränderungen im Programmdesign aufgrund der geringen Zahl an Förderfällen Rechnung tragen.

Die Projekte zum Einstiegsgeld in Baden-Württemberg und Hessen ("Hessischer Kombilohn") implementieren zeitlich befristet im Rahmen der Experimentierklausel des § 18 Abs. 5 Bundessozialhilfegesetzes (BSHG) eine großzügigere Anrechnung von eigenem Erwerbseinkommen der Hilfeempfänger. Damit sollen die Wirkungen von größeren Arbeitsanreizen bei langzeitarbeitslosen Sozialhilfeempfängern untersucht werden. In Baden-Württemberg haben nach 15 Monaten Laufzeit bis zum 31. März 2001 496 Personen Einstiegsgeld bezogen, vgl. Dann *et al.* (2001). Einige Städte in Baden-Württemberg – leider nicht alle – unterteilen die Zielgruppe des Einstiegsgeldes zufällig in eine Programm- und eine Kontrollgruppe oder verwenden Programm- und Vergleichsbezirke in der Stadt, so dass hier – in bisher einmaliger Weise in Deutschland – eine experimentelle Evaluation möglich sein wird. Die Fallzahlen für Freiburg und Mannheim sind jeweils so hoch, dass wissenschaftlich verwertbare Ergebnisse erwartet werden können. Der Zwischenbericht Dann *et al.* (2001) beschreibt die Einstiegsgeldbezieher in Baden-Württemberg bis zum 31. März 2001. Eine Evaluation wurde jedoch noch nicht vorgenommen. Im Hessischen Projekt wurden in fast allen Modellstädten und -kreisen Programm- und Kontrollgruppen unterteilt. Allerdings sind die Zielgruppen und Teilnehmerzahlen kleiner als in Baden-Württemberg.

Das im April 2001 vorgestellte jüngste Modellvorhaben der Bundesregierung zur Stärkung der AAMP unter dem Titel MoZArT („Modellvorhaben zur Verbesserung der Zusammenarbeit von Arbeitsämtern und Trägern der Sozialhilfe“) hat die Integration von erwerbsfähigen Arbeitslosen- sowie Sozialhilfeempfängern als Zielsetzung. Die Verbesserung der Hilfe zur Selbsthilfe, die in den unterschiedlichen Instrumentenkästen des SGB III und des (BSHG) zum Ausdruck kommt, soll durch geförderte Projekte unterschiedlicher Kategorien erreicht werden. Vier Leitkategorien lassen sich dabei identifizieren: Schaffung gemeinsamer Anlaufstellen von Arbeits- und Sozialämtern, Straffung der Verwaltungsabläufe, Durchführung gemeinsamer Qualifizierungs- oder Beschäftigungsmaßnahmen, gemeinsame Bewilligung und/oder Auszahlung von Leistungen durch eine Stelle (BMA (2001)). Der Modellversuch wird vom Bundesministerium für Arbeit und Sozialordnung finanziert und läuft zunächst bis Ende 2002 mit einem Gesamtvolumen von bis zu 30 Mill. DM (BMA (2001)). Experimentierklauseln dienen zur Prüfung der Möglichkeiten einer weiteren Flexibilisierung der Arbeitsmarktpolitik.

Die Evaluation der MoZArT-Projekte soll über die gesamte Zeit der Durchführung erfolgen. Datengrundlage sind die in Sozial- und Arbeitsämtern erhobenen Daten, ergänzt um Personenbefragungen. Ziel ist die Schaffung eines Panels, das in zwei oder drei Erhebungen und einem Umfang von 11.000 Interviews den Modellverlauf wiedergeben kann. Nettoeffekte sollen mittels

einer Kontrollgruppe aus unbeteiligten Ämtern ermittelt werden. Das methodische Konzept der Kontrollgruppenbildung wird allerdings aufgrund der bislang vorliegenden Informationen nicht klar.

Schließlich kommt dem Gesetz zur Reform der arbeitsmarktpolitischen Instrumente (Job-AQTIV-Gesetz), das vom Deutschen Bundestag am 9. November 2001 verabschiedet wurde, auch im Hinblick auf die Evaluation eine große Bedeutung zu (Deutscher Bundestag (2001), Drucksache 14/6944). AQTIV steht dabei für Aktivieren, Qualifizieren, Trainieren, Investieren und Vermitteln. Die Ziele der Arbeitsförderung werden im § 1 Absatz 1 formuliert: “Die Leistungen der Arbeitsförderung sollen dazu beitragen, dass ein hoher Beschäftigungsstand erreicht und die Beschäftigungsstruktur ständig verbessert wird. Sie sind insbesondere darauf auszurichten, das Entstehen von Arbeitslosigkeit zu vermeiden oder die Dauer der Arbeitslosigkeit zu verkürzen. Dabei ist die Gleichstellung von Frauen und Männern als durchgängiges Prinzip zu verfolgen. Die Leistungen sind so einzusetzen, dass in der beschäftigungspolitischen Zielsetzung der Sozial-, Wirtschafts- und Finanzpolitik der Bundesregierung entsprechen.” Zur Realisierung dieser Ziele sollen arbeitsmarktpolitische Massnahmen, wie beispielsweise Konzepte der Arbeitsvermittlung, Job-Rotation, berufliche Qualifizierung, dienen. Auch die Ziele und Aufgaben der Evaluationsforschung werden im § 282 neu formuliert. Die Evaluationsforschung wird dabei als “ein Schwerpunkt der Arbeitsmarktforschung” definiert, sie soll zeitnah erfolgen und ist ständige Aufgabe des Instituts für Arbeitsmarkt- und Berufsforschung. Damit sind auch gesetzliche Rahmenbedingungen für eine aktuelle und permanente Evaluationsforschung in Deutschland geschaffen. Es werden die wichtigsten Ergebnisvariablen definiert und sowohl Kosten–Nutzen–Analysen als auch makro– und mikroökonomische Studien gefordert. Dies ist sicherlich auch für die Wissenschaft ein ermutigender Ansatz.

4. Einschätzung der AAMP und Schlußfolgerungen für die gegenwärtige und zukünftige Evaluationsforschung

Für West- und Ostdeutschland wurden in den vergangenen Jahren vor allem mikroökonomische Evaluationsstudien vorgelegt, die insbesondere den Erfolg der Maßnahmen zur beruflichen Weiterbildung und der Arbeitsbeschaffungsmaßnahmen empirisch analysiert haben. Makroökonomische Evaluationsstudien sind seltener, aber auch sie behandeln lediglich die beiden Maßnahmen der beruflichen Weiterbildung und der Arbeitsbeschaffung.¹⁰ Die fiskalischen Kosten der AAMP wurden in den bisher vorgelegten Evaluationsstudien nicht berücksichtigt.

¹⁰Einen Überblick über die relevanten Studien und Ergebnisse bieten: Fitzenberger/Speckesser (2000), Hagen/Steiner (2000), Hujer/Caliendo (2001). Ein prägnantes Beispiel für makroökonomische Verdrängungseffekte ist der Garten- und Landschaftsbau in Ostdeutschland. Hier werden ABM besonders häufig durchgeführt, vgl. Schneider *et al.* (2000), Abschnitt IV.3.

Als Datengrundlagen für die mikroökonomischen Evaluationen in Deutschland wurden bislang relativ kleine Umfragedatensätze wie das Sozioökonomische Panel oder der Arbeitsmarktmonitor-Ost verwendet.¹¹ Für die makroökonomischen Studien wurden Informationen auf der Ebene der Arbeitsamtsbezirke oder Raumordnungsregionen herangezogen. Die Ergebnisse der bislang vorgelegten empirischen Studien zur Förderung der beruflichen Weiterbildung und Arbeitsbeschaffungsmaßnahmen sind widersprüchlich, jedoch tendenziell eher negativ. Bei der Interpretation der unterschiedlichen empirischen Befunde und vor allem im Hinblick auf wirtschafts- und sozialpolitische Handlungsempfehlungen muss jedoch beachtet werden, dass die Heterogenität der Maßnahmen der aktiven Arbeitsmarktpolitik wegen fehlender Informationen und zu geringer Fallzahlen nicht berücksichtigt werden konnte. Dies gilt ebenso für unterschiedliche Teilnehmerstrukturen. Obwohl große Fortschritte in der Anwendung geeigneter methodischer Ansätze erzielt werden konnten, sind die verwendeten Datengrundlagen zur Analyse der Effekte der AAMP zu wenig detailliert (vgl. auch Sachverständigenrat (2000), S. 85).¹²

Allerdings ist geplant, die umfangreichen Datensätze der Arbeitsverwaltung für mikroökonomische Evaluationsstudien nutzbar zu machen.¹³ Erst auf dieser Grundlage können politisch relevante Aussagen über Erfolg und Mißerfolg von arbeitsmarktpolitischen Maßnahmen gemacht werden.

Als vorläufige Einschätzung der AAMP und deren Evaluation kann angesichts des institutionellen Designs der Maßnahmen und der Teilnehmerentwicklung festgehalten werden:

1. Eine umfassende Analyse muss vor allem die Wirkung der Maßnahmen der beruflichen Weiterbildung, der ABM und SAM sowie der Lohnkostenzuschüsse abschätzen. Da durch die Programmvielefalt den Arbeitsämtern unterschiedliche Optionen gegeben werden, Problemgruppen in Beschäftigung zu bringen, ist eine simultane Evaluation heterogener Maßnahmen erforderlich. Wegen der regionalen Variation der Arbeitsmarktpolitik sind die Unterschiede zwischen Regionen hinsichtlich der Ausrichtung auf einzelne Maßnahmen zu berücksichtigen.
2. Die aktive Arbeitsmarktpolitik darf nicht mehr Spielball einer unsystematischen budgetgesteuerten „Stop-and-Go-Policy“ sein. Bereits bei der Einführung arbeitsmarktpolitischer Maßnahmen sind methodische und empirische Konzepte einer systematischen Erfolgskontrolle zu entwickeln, denn nur eine gezielte Wirkungsforschung kann zu einer effizienten

¹¹Sowie in einem Fall ausgewählte CoArb-Daten der Bundesanstalt für Arbeit in einer Region, die mit großem Aufwand aufbereitet werden mussten, vgl. Almus *et al.* (1999).

¹²Eine andere Einschätzung vertritt der Beitrag von Steiner/Hagen in diesem Sonderband. Steiner/Hagen sind der Meinung, dass aus den bisher vorliegenden Evaluationsstudien trotz aller Probleme hinsichtlich Datenverfügbarkeit und Heterogenität der Massnahmen geschlossen werden kann, dass die AAMP in Deutschland im Durchschnitt keine oder sogar negative Beschäftigungseffekte aufweist.

¹³An entsprechenden gemeinsamen Pilotprojekten mit dem IAB sind drei der Autoren (Fitzenberger, Hujer, Lechner) in diesem Sonderband beteiligt.

(Neu-)Gestaltung von arbeitsmarktpolitischen Maßnahmen führen (vgl. die im Beitrag von Lechner in diesem Sonderband formulierte Vision für ein Zusammenspiel von AAMP und Evaluationsforschung).

3. Die Ansätze der Eingliederungsbilanzen und der Verbleibsquoten erfüllen einen Teil des Informationsbedürfnisses. Allerdings sagt die Verbleibsquote nichts über den Übergang in die reguläre Beschäftigung aus, so dass der Sachverständigenrat diesen Indikator sogar als "irreführendes Erfolgskriterium für die Evaluierung arbeitsmarktpolitischer Maßnahmen" (Sachverständigenrat (2000), S. 85) bezeichnet. Offen bleibt auch, welche (Fehl-)Anreize mit diesem Konzept für die Arbeitsämter verbunden sind, beispielsweise "Teilnehmercreaming" statt Zielgruppenorientierung.
4. Trotz der Versuche, wissenschaftliche Begleitforschungen für das Jugendsofortprogramm (JUMP), CAST und das Projekt zum Einstiegsgeld durchführen zu lassen, fehlt in der Politik oft das Verständnis dafür, wie eine aussagekräftige Evaluation durchzuführen ist („fehlende Evaluationskultur“). Das Evaluationsdesign ist typischerweise nicht optimal daran ausgerichtet, eine adäquate Kontrollgruppe zu generieren. Es wäre wünschenswert, wenn in zunehmendem Maße auch experimentelle Designs verwendet werden könnten und sich kostenaufwendige Nacherhebungen zwecks Kontrollgruppenbildung erübrigen. Diesen Aspekten sollte in der anvisierten Novellierung des SGB III Rechnung getragen werden.
5. Eine versuchsweise verstärkte Förderung der regionalen Flexibilität der Arbeitsmarktpolitik und ein zunehmender Einsatz mobilitätsunterstützender Maßnahmen erscheint empfehlenswert, um die Effizienz der aktiven Arbeitsmarktpolitik zu erhöhen. Eine solche Politik ist natürlich einer entsprechenden Evaluation zu unterziehen.

Literatur

Almus, M., Egel, J., Lechner, M., Pfeiffer, F. und H. Spengler (1999): Wirkungen gemeinnütziger Arbeitnehmerüberlassung in Rheinland-Pfalz. Beiträge aus der Arbeitsmarkt- und Berufsforschung 225, Nürnberg: Bundesanstalt für Arbeit.

Bittner, S., Holleder, A., Kaltenborn, B., Rudolph, H., Vanselow, A. und C. Weinkopf (2001): Ein Jahr Erfahrungen mit dem arbeitsmarktpolitischen Sonderprogramm CAST. Forschungsverbund IAB/IAT/Dr. Kaltenborn, 1. Zwischenbericht September 2001, Bonn.

Brinkmann, C. (2000): Zur Evaluation der aktiven Arbeitsmarktförderung nach dem SGB III – Monitoring und Wirkungsforschung im Umbruch. Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, Nr. 3, S. 483–499.

Bundesanstalt für Arbeit – ANBA: Amtliche Nachrichten der Bundesanstalt für Arbeit. Verschiedene Jahrgänge.

- Bundesministerium für Arbeit und Sozialordnung (2001):** MoZArT – Neue Strukturen für Jobs. <http://www.bma-mozart.de>
- Calmfors, L. (1994):** Active Labour Market Policy and Unemployment - a Framework for the Analysis of Crucial Design Features. OECD Economic Studies, 22(1), S. 7–47.
- Dann, S., Kirchmann, A., Spermann A. und J. Volkert (2001):** Einstiegsgeld in Baden-Württemberg – Eine Zwischenbilanz des baden-württembergischen Modellversuchs zum Kombi-Einkommen. Institut für angewandte Wirtschaftsforschung, Juli 2001, Tübingen.
- Deutscher Bundestag (2001):** Gesetz zur Reform der arbeitsmarktpolitischen Maßnahmen (Job-AQTIV-Gesetz), Drucksache 14/6944.
- Dietrich, H. (2001):** JUMP, das Jugendsofortprogramm. IAB-Werkstattbericht, Nr. 3, Nürnberg.
- Dolton, P. und Y. Balfour (2002):** Der New Deal: “Welfare to Work”-Programme in Großbritannien. Perspektiven der Wirtschaftspolitik (erscheint demnächst).
- Fitzenberger, B. und S. Speckesser (2000):** Zur wissenschaftlichen Evaluation der Aktiven Arbeitsmarktpolitik in Deutschland. Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, Nr. 3, S. 357–370.
- Hagen, T. und V. Steiner (2000):** Von der Finanzierung der Arbeitslosigkeit zur Förderung der Arbeit. Zentrum für Europäische Wirtschaftsforschung, Mannheim, ZEW Wirtschaftsanalysen, Band 51, Nomos.
- Hagen, T. und V. Steiner (2002):** Was kann die Aktive Arbeitsmarktpolitik in Deutschland aus der Evaluationsforschung in anderen europäischen Ländern lernen? Perspektiven der Wirtschaftspolitik (erscheint demnächst).
- 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, Amsterdam: North Holland, S. 1865–2097.
- Hübler, O. (2001):** Evaluation of Policy Interventions: Measurement and Problems. Allgemeines Statistisches Archiv, 85, S. 103–126.
- Hujer, R. und M. Caliendo (2001):** Evaluation of Active Labour Market Policy – Methodological Concepts and Empirical Estimates. In: Becker, I., Ott, N. und G. Rolf (Hrsg.) Soziale Sicherung in einer dynamischen Gesellschaft, Frankfurt: Campus-Verlag, S. 583–617.
- Lampert, H. (1989):** 20 Jahre Arbeitsförderungsgesetz. Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, S. 173–186.
- Lechner, M. (1998):** Training the East German Labour Force: Microeconomic Evaluations of Continuous Vocational Training after Unification. Physica-Verlag, Heidelberg.

- Lechner, M. (2002):** Eine wirkungsorientierte aktive Arbeitsmarktpolitik in Deutschland und der Schweiz: Eine Vision - zwei Realitäten. Perspektiven der Wirtschaftspolitik (erscheint demnächst).
- Rosenbaum, P.R. und R.B. Rubin (1983):** The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70, S. 41–55.
- Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Lage (1998):** Vor weitreichenden Entscheidungen. Jahresgutachten 1998/1999.
- Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2000):** Chancen auf einen höheren Wachstumspfad. Jahresgutachten 2000/2001.
- Schmidt, C.M., K.F. Zimmermann, M. Fertig und J. Kluve (2001):** Perspektiven der Arbeitsmarktpolitik – Internationaler Vergleich und Empfehlungen für Deutschland. Springer-Verlag, Heidelberg.
- Schneider, H., A. Bergemann, O. Fuchs, W. Kempe, J. Kolb, M. Rosenfeld, B. Schultz und L. Trabert (2000):** Die Effizienz der Arbeitsmarktpolitik in den neuen Ländern - Eine Bilanz der Vergangenheit und Ansätze für zukünftige Reformen. Institut für Wirtschaftsforschung Halle, Sonderheft 3/2000.
- Sell, S. (1998):** Entwicklung und Reform des Arbeitsförderungsgesetzes als Anpassung des Sozialrechts an flexible Erwerbsformen. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 31, S. 532-533.
- Spermann, A. (2001):** Negative Einkommenssteuer, Lohnsubventionen und Langzeitarbeitslosigkeit. *Finanzwissenschaftliche Schriften*, Frankfurt.
- Vollkommer, D. (2000):** Regionalisierung der Arbeitsmarktpolitik – Verbleibsquoten von ABM-Teilnehmern in Eingliederungsbilanzen. IAB-Werkstattbericht, Nr. 5, Nürnberg.
- Zängle, M. (2000):** Eingliederungsbilanzen in der Praxis. *hiba, Durchblick* Nr. 2, Berufliche Fortbildungszentren der Bayerischen Wirtschaft GmbH.