

Evaluation aktiver Arbeitsmarktpolitik in Deutschland – eine Einführung

Christian Brinkmann, Reinhard Hujer, Susanne Koch

1 Zur Entwicklung der Evaluationsforschung in Deutschland

Mit dem vorliegenden Schwerpunktheft der ZAF wird der aktuelle Forschungsstand in einem Themenfeld aufbereitet, das in den letzten Jahren in Wissenschaft und Politik zunehmend an Bedeutung gewonnen hat, dessen Ursprünge aber bis in die zwanziger und dreißiger Jahre des letzten Jahrhunderts zurückreichen. Zur Bekämpfung der damaligen Wirtschaftskrise wurden bereits von der 1927 gegründeten Reichsanstalt für Arbeitsvermittlung und Arbeitslosenversicherung als Instrumente Arbeitsbeschaffung, Fortbildung und Umschulung sowie Kurzarbeit eingesetzt, die Anfänge der damaligen „Produktiven Erwerbslosenfürsorge“ lassen sich bis zur Zeit unmittelbar nach dem Ersten Weltkrieg zurückverfolgen (Tiefenbacher 1931: 43 ff.; Schmuhl 2003: 122 ff.).

Praxis und Wissenschaft bemühten sich damals schon um Erkenntnisse zu den Auswirkungen dieses Maßnahmeinsatzes, so zum Beispiel auch hinsichtlich der indirekten Beschäftigungseffekte von Straßenbauarbeiten (Arbeit und Beruf 1931: 364 f.). Eine damalige Dissertation beschäftigte sich mit der volkswirtschaftlichen Bedeutung der Produktiven Erwerbslosenfürsorge (Tiefenbacher 1931). Ansprüche hinsichtlich einer systematischen wissenschaftlichen Fundierung der Arbeitsmarkt- und Beschäftigungspolitik existierten damals aber noch nicht – ganz zu schweigen von den damals ganz unzulänglichen Realisierungsmöglichkeiten.

Auch in den 1950er und 1960er Jahren gab es an den Hochschulen der Bundesrepublik Deutschland noch keine Basis für Arbeitsmarktforschung im Allgemeinen und wissenschaftliche „Erfolgskontrollen“ des Instrumenteneinsatzes nach dem von 1952 bis 1969 geltenden Gesetz über Arbeitsvermittlung und Arbeitslosenversicherung (AVAVG) im Besonderen. Ein entsprechendes Problembewusstsein bildete sich erst Mitte der 1960er Jahre heraus und führte 1967 zur Gründung des IAB (Schmuhl 2003: 467 ff.; Mertens 1972). Erwartet wurden in einem ersten von der Selbstverwaltung der BA formulierten „Aufgabenkatalog für die Arbeitsmarkt- und Berufsforschung in der Bundesanstalt“ vom 3.3.1966 u.a. Untersuchungen „für arbeitsmarktpolitische Aspekte wie z.B. Berufsbildung und Umschulung, berufliche Rehabilitation, Anpassungshilfen für freigewordene und von Freisetzungen bedrohte Arbeitskräfte ...“ (IAB 1972: 2** f.), ohne dass hier allerdings „Erfolgskontrollen“ explizit angesprochen wurden. Letzteres gilt auch für die Formulierungen des Arbeitsförderungsgesetzes (AFG) von 1969, mit denen Arbeitsmarkt- und Berufsforschung gesetzlich festgeschrieben wurde (§ 6).

Ende der 1960er und Anfang der 1970er Jahre wurden den Erwartungen der politischen Akteure entsprechend – nach damaligem „state-of-the-art“ – erste Analysen zur Förderung der beruflichen Weiterbildung durchgeführt (Bolte 1969; Brinkmann et al. 1970), die später auf weitere Instrumente ausgeweitet wurden.¹ Darüber hinaus gab es zu dieser Zeit Untersuchungen zu den Wirkungen der Beschäftigungspolitik im weiteren Sinn (Reyher 1968; Ernst 1974).

Aber erst Ende der 1970er und Anfang der 1980er Jahre wurden hierzulande die Fäden der Evaluationsforschung systematisch aufgenommen und das Forschungsfeld weiter

¹ Seit 1973 z. B. regelmäßiger Ausweis der Beschäftigungs- und Entlastungseffekte der größeren arbeitsmarktpolitischen Instrumente im Rahmen der kurzfristigen Arbeitsmarktanalysen des IAB (erstmalig Autorengruppe 1973).

entwickelt (Mertens/Reyher/Kühl 1981). Mit dem Aufbau des Wissenschaftszentrums Berlin (für Sozialforschung) wurden zudem die einschlägigen Forschungskapazitäten erweitert und eine internationale Perspektive in die Diskussion eingeführt (Scharpf/Schmid 1981). Hintergrund waren die ersten Arbeitsmarktkrisen nach der Bewältigung der Nachkriegsarbeitslosigkeit in Deutschland und die damit verbundene „massive“ Ausweitung der aktiven Arbeitsförderung (von 3,8 Mrd. DM im Jahr 1973 auf immerhin 11,4 Mrd. DM im Jahr 1982).²

1979 wurde ein für damalige Verhältnisse groß dimensioniertes arbeitsmarktpolitisches Sonderprogramm des Bundes – das so genannte „500-Millionen-Programm“ – durchgeführt. Für eine vielschichtige Evaluation dieses Programms wurden in beachtlichem Umfang zusätzliche Ressourcen zur Verfügung gestellt.³ Parallel dazu wuchs die Erkenntnis, dass zur Steuerung der Arbeitsmarktpolitik effiziente Strukturen der Arbeitsverwaltung benötigt werden – auch der Beginn der Dezentralisierungsdebatte lässt sich bis in diese Zeit zurückverfolgen (Blankenburg/Krautkrämer 1977; Blankenburg 1978; Lauritzen 1980).

Vor dem Hintergrund der transformationsbedingten Arbeitsmarktprobleme in Ostdeutschland und des anhaltenden enormen Arbeitsmarktungleichgewichts in Gesamtdeutschland wurden in den 1990er Jahren die Fördermittel nochmals erheblich ausgeweitet. Mit der Verabschiedung des Dritten Buches des Sozialgesetzbuchs (SGB III) anstelle des AFG im Jahre 1997 wurde zudem die Förderung neu strukturiert – und es entstanden neue Erwartungen an die Wirkungsforschung. Die Weiterentwicklung der Forschungsmethoden und wesentliche Fortschritte bei der Aufbereitung von Prozessdaten der BA⁴ für Evaluationszwecke eröffneten neue Möglichkeiten für methodisch anspruchsvolle, namentlich auch ökonometrische Evaluationen (Hujer/Bellmann/Brinkmann 2000; Stephan/Brinkmann 2006). Das Forschungsdatenzentrum (FDZ) der BA im IAB ermöglicht nunmehr einen transparenten, standardisierten und gleichberechtigten Datenzugang für die Forschung (Allmendinger/Kohlmann 2005; Kohlmann 2005).⁵

Schließlich setzten vor allem das Job-AQTIV-Gesetz (2002; Wirkungsforschung zur Arbeitsförderung wurde nunmehr erstmals ausdrücklich gesetzlich verankert)⁶ und die jüngsten Gesetze für moderne Dienstleistungen am Arbeitsmarkt („Hartz I–IV“) neue Akzente für die Evaluationsforschung in der Bundesrepublik Deutschland.

Mit diesen Entwicklungen wurden für die Evaluationsforschung auch Qualitätsstandards festgeschrieben, die zu früheren Zeiten keineswegs selbstverständlich waren. „Es ist noch nicht so lange her, da galt die Forderung von Wissenschaftlern, bei Wirkungsanalysen auch Kontrollgruppen einzubeziehen, bei den Förderern und Auftraggebern als ein völlig überzogenes Verlangen, an dem man sehr gut sähe, wie sehr die Wissenschaft geneigt sei, mit öffentlichem Geld zu prassen. Inzwischen hat sich die Einbeziehung von Kontrollgruppen weitgehend als Standard durchgesetzt“ (Lutz 2006: 360).

Zur Erfolgskontrolle im Hinblick auf politisch formulierte Ziele, z. B. Eingliederung in reguläre Beschäftigung, können vier Ansätze verwendet werden:

- Mikroökonomische Evaluation
- Makroökonomische Evaluation

² Vgl. Mertens/Reyher/Kühl (1988: 359).

³ Vgl. dazu die verschiedenen im Kontext der Begleitforschung entstandenen Veröffentlichungen (u. a. Peters/Schmid 1982; Scharpf et al. 1982; Bosch/Seifert/Spies 1983; Hellmich 1983; Mendius et al. 1983).

⁴ Früher Bundesanstalt für Arbeit, nunmehr Bundesagentur für Arbeit.

⁵ Vgl. auch <http://fdz.iab.de>.

⁶ Als Daueraufgabe der BA bzw. des IAB (§ 282 ff. SGB III; vgl. auch Brinkmann 2000).

- Effizienzanalyse durch einen Vergleich von Kosten und Nutzen
- Implementationsanalyse.

In der mikroökonomischen Analyse steht die Ermittlung der direkten Effekte für die teilnehmenden Individuen im Vordergrund des Interesses. Mit diesem Ansatz können individual-, maßnahmen- und regionspezifische Effekte in differenzierter Form ermittelt werden. Die makroökonomische Analyse bezieht dagegen auch die indirekten Effekte auf Nichtteilnehmer bzw. die Gesamtwirtschaft ein. Dies ist insbesondere dann von Bedeutung, wenn die Programme ein großes Volumen haben (gemessen an der Teilnehmerzahl und/oder dem Mitteleinsatz) und von Spillover-Effekten auf Nichtteilnehmer ausgegangen werden muss. Berücksichtigt man, dass die indirekten Effekte die direkten Auswirkungen stärken oder schwächen können, so kann sich beim Übergang von einer mikro- zu einer makroanalytischen Betrachtung nicht nur die Stärke, sondern auch das Vorzeichen der Politikwirkung ändern. Calmfors (1994) diskutiert in diesem Zusammenhang Mitnahme-, Substitutions- und Verdrängungseffekte.

Im Rahmen einer Effizienzanalyse wird analysiert, welche Kosten einerseits durch die unterschiedlichen arbeitsmarktpolitischen Programme entstehen, andererseits Einsparungen durch Unterstützungsleistungen und Verkürzung der Arbeitslosigkeitsdauer zu erzielen sind.

Schließlich dient eine Implementationsanalyse dazu, insbesondere mit Hilfe von Befragungen die konkrete Umsetzung zu untersuchen und auch qualitative Aspekte der Erfolgskontrolle zu betrachten.

Mit der nun verfügbaren integrierten Datenbasis wird vor allem für mikroökonomische Evaluationen ein Qualitätssprung erreicht. Sie haben nunmehr auch von der Datenseite her – anders noch als Ende der 1990er Jahre – eine Reife und Differenziertheit erreicht, wie sie für eine verantwortliche Politikberatung unerlässlich erscheint.

Insgesamt gilt, dass je nach Fragestellung, namentlich bei neu eingeführten Instrumenten, ein Methoden-Mix erforderlich ist, der von qualitativen Untersuchungsbausteinen (Implementationsforschung, Erfolgsmessung hinsichtlich quantitativ – noch – nicht messbarer Zieldimensionen wie „Beschäftigungsfähigkeit“) über mikro- und makroökonomische Untersuchungen bis hin zu (Ansätzen von) Kosten-Nutzen-Untersuchungen reicht. Eine erste derart vielschichtig angelegte Evaluation (über den Vermittlungsgutschein und die Beauftragung Dritter nach § 37 SGB III) wurde in Deutschland nach längerer Vorbereitung Anfang 2003 vom IAB an ein Konsortium vergeben, bestehend aus dem Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim, dem Institut für angewandte Wirtschaftsforschung (IAW), Tübingen, dem sinus-Institut, München, sowie dem Lehrstuhl für Statistik und Ökonometrie der Johann Wolfgang Goethe-Universität, Frankfurt/M. (vgl. den Beitrag von Pfeiffer und Winterhager in diesem Schwerpunktheft).⁷ Ein entsprechender Methoden-Mix liegt auch der Konzeption der Hartz-Evaluationen zugrunde, aus denen im zweiten Teil dieses Heftes Zwischenergebnisse vorgestellt werden (vgl. den hierzu einführenden Beitrag von Heyer in diesem Heft).

Wie hilfreich Implementationsuntersuchungen z. B. auch zur Interpretation von ökonomischen Befunden sein können, verdeutlicht der Beitrag von Brussig u. a. über Lohnkostenzuschüsse und Entgeltsicherung (in diesem Heft).

Anhand der makroökonomischen Studie von Fertig, Kluge und Schmidt in diesem Heft lässt sich einerseits besonders deutlich erkennen, welche großen Methodenfort-

⁷ Vgl. auch IAB 2005.

schritt es auch auf dieser Analyse-Ebene gegeben hat. Die nunmehr bestehende Möglichkeit, die regionale Varianz beim Instrumenteneinsatz⁸ als erklärende Variable für die Entwicklung von Bewegungs- oder Bestandsgrößen des Arbeitsmarktes zu nutzen, kommt einem Durchbruch auch für solche Evaluationen gleich. Andererseits ist zu bedenken, dass bei der grundsätzlich begrenzten Anzahl verfügbarer regionaler Einheiten und der vergleichsweise kurzen Zeit seit Einführung wichtiger Reformelemente das methodisch bedingte Risiko groß ist, in der Realität vorhandene Einflüsse (noch) nicht statistisch signifikant nachweisen zu können. Solche Untersuchungen sind – namentlich wenn sie zu Negativbefunden hinsichtlich der Wirkungen von Maßnahmen auf dem Arbeitsmarkt führen – von den Akteuren als Warnsignale durchaus auch jetzt schon ernst zu nehmen. Die Robustheit solcher Ergebnisse bedarf jedoch noch weiterer wissenschaftlicher Überprüfung.

2 Zu den Beiträgen

Im *ersten Teil dieses Schwerpunktheftes* werden zum einen (Teil-)Ergebnisse aus drei längerfristig angelegten Evaluationsprojekten vorgestellt, die neben den inhaltlichen Zielsetzungen auch ausdrücklich auf die Entwicklung von Prozessdaten in der BA für Forschungszwecke hin ausgelegt waren: die mikroökonomischen Analysen von ABM (Hujer und Thomsen) und beruflicher Qualifizierung (Lechner/Wunsch und Biewen et al.) sowie die Begleitforschung zum ESF-BA-Programm (Deeke/Kruppe). Weiterhin stellen Pfeiffer und Winterhager Befunde vor, die auf der neu geschaffenen Datenbasis basieren und auf den Vergleich zweier Instrumente der Arbeitsmarktpolitik abstellen. Stephan, Rässler und Schewe gehen einen weiteren wichtigen Schritt in Richtung instrumentenübergreifender Evaluationen, die im Spannungsfeld von hochgradig differenziertem Monitoring (für aktuelle Steuerungszwecke in der BA) und längerfristig angelegter Wirkungsforschung angelegt sind.⁹

Hujer und Thomsen diskutieren eine Reihe neuerer mikroökonomischer Studien zur Wirksamkeit von ABM und analysieren auf der Basis der neu verfügbaren Datensätze (einschließlich Beschäftigtendatei) u. a. die Beschäftigungseffekte von ABM in einem dynamischen Evaluationsansatz. Sie kommen zu dem Ergebnis, dass diese Maßnahmen nur einigen Zielgruppen nutzen, während sie die Chancen der Mehrzahl der teilnehmenden Personen am Arbeitsmarkt reduzieren.

Lechner und Wunsch beschäftigen sich mit der Frage, inwieweit unterschiedliche Zielvariablen (Beschäftigung, Einkommen, Arbeitslosigkeit) zu unterschiedlichen Aussagen über die Wirksamkeit von Qualifizierungsmaßnahmen führen. Sie kommen zu dem Befund, dass sich längerfristig nachweisbare positive Beschäftigungs- und Einkommenseffekte hinsichtlich eines Rückgangs der Arbeitslosigkeit nicht in entsprechenden Effekten spiegeln. Sie plädieren für eine differenzierte Betrachtungsweise und längerfristig angelegte Beobachtungszeiträume.

Biewen, Fitzenberger, Osikominu, Völter und Waller berichten sowohl Ergebnisse anderer als auch eigener Studien zur Wirksamkeit von Fortbildung und Umschulung nach dem AFG bzw. beruflicher Weiterbildung nach SGB III. Die Ergebnisse zeigen, dass die Beschäftigungswirkungen von Maßnahmen der beruflichen Weiterbildung weitaus positiver sind als dies Studien in der Vergangenheit auf Basis von Umfragedaten nahelegen. Auch sie betonen die Notwendigkeit einer längerfristigen Betrachtungsweise.

Deeke und Kruppe untersuchen im Rahmen der ESF-Begleitforschung die über den ESF finanzierte ergänzende Weiterbildung (ESF-FbW) im Vergleich zur gesetzlichen

⁸ Beim Beitrag von Mauer in diesem Heft geht es analog um Wirkungen des neuen Kundenzentrums insgesamt.

⁹ Zu diesem Spannungsverhältnis vgl. Kruppe 2006.

Regelförderung („added value“). Ihre Wirkungsanalysen mit Vergleichsgruppen beruhen sowohl auf Prozess- als auch auf Befragungsdaten, die mittels eines sequentiellen Matching selektiert wurden. Als Erfolgskriterien dienen der spätere Verbleib in ungeförderter sozialversicherungspflichtiger Beschäftigung sowie der Verbleib in geförderter und ungeförderter Beschäftigung inklusive geförderter Selbständigkeit. Weiterhin wird analysiert, ob eine Gruppe den Zustand der registrierten Arbeitslosigkeit (inkl. Maßnahme-Teilnahme) schneller beendet. Im Ergebnis zeigen die ESF-Geförderten nach sechs Monaten eine leicht geringere Beschäftigungswahrscheinlichkeit als Teilnehmer an der Regelförderung, sie sind in dieser Zeit mit deutlich höherer Wahrscheinlichkeit arbeitslos gemeldet bzw. nehmen an Maßnahmen teil und sie haben in der Tendenz die gleiche Beschäftigungswahrscheinlichkeit wie die Gruppe der vergleichbaren, aber nicht teilnehmenden Arbeitslosen.

Pfeiffer und *Winterhager* untersuchen ergänzend zu bereits vorliegenden Studien (mit einem Vergleich von Teilnehmern und Nichtteilnehmern) die relative Wirksamkeit von Vermittlungsgutscheinen und der Beauftragung Dritter mit der gesamten Vermittlung im direkten Vergleich. Bezogen auf Arbeitslose, die beide Instrumente nutzen können, kommen sie zu dem Ergebnis, dass Vermittlungsgutscheine vergleichsweise mehr Arbeitslose in Beschäftigung gebracht haben.

Stephan, *Rässler* und *Schewe* stellen ein in der Entwicklung befindliches System der quantitativen Wirkungsanalyse vor, das gegenwärtig in der BA zur Unterstützung der operativen Steuerung entwickelt wird (TrEffeR = Treatment Effects and Prediction). Die bislang vorliegenden Ergebnisse deuten darauf hin, dass bei instrumentenübergreifender Betrachtung im Hinblick auf die Zielvariable Verbleib arbeitslos innerhalb eines Zeitraums von zwei Jahren insbesondere Eingliederungszuschüsse sowie das Überbrückungsgeld erfolgreich eingesetzt wurden. Sie weisen allerdings ausdrücklich darauf hin, dass gerade bei FbW ein längerer Zeithorizont entscheidend sein kann.

Im zweiten Teil dieses Schwerpunktheftes werden Konzeption und ausgewählte Zwischenergebnisse aus den laufenden „Hartz-Evaluationen“ behandelt, die nicht nur für die aktuelle politische Diskussion der Arbeitsmarktpolitik von Bedeutung sind, sondern auch einen weiteren Entwicklungsschub für die Evaluationsforschung zur Arbeitsmarktpolitik in Deutschland bedeuten, wobei allerdings weiterhin viele alte und neue datentechnische, methodische, inhaltliche und forschungsorganisatorische Probleme zu lösen bleiben, auf die hier im Einzelnen nicht näher eingegangen werden soll.¹⁰ Die Beiträge zielen nicht nur auf den mit den Gesetzen für Moderne Dienstleistungen am Arbeitsmarkt reformierten Einsatz einzelner Eingliederungsinstrumente (*Schneider/Uhlendorff*, *Brussig et al.* und *Caliendo/Wießner/Kritikos*), sondern auch auf die Reform der Arbeitsvermittlung insgesamt (*Mauer*) und die mit den Instrumenten angestrebten Makro-Wirkungen auf dem Arbeitsmarkt (*Fertig/Kluve/Schmidt*). Mit dem Beitrag von *Bernhard/Wolff/Jozwiak* werden erste Ergebnisse aus der Begleitforschung zum Instrumenteneinsatz nach dem Sozialgesetzbuch II (SGB II) für erwerbsfähige Hilfebedürftige vorgestellt.

Heyer erläutert zunächst Hintergründe und Konzeption der „Hartz-Evaluationen“, die auf einem Auftrag des Deutschen Bundestags beruhen, und sich auf die drei ersten Gesetze für Moderne Dienstleistungen am Arbeitsmarkt beziehen. Sie können als das bisher größte eigenständige Projekt im Bereich der Arbeitsmarktforschung in Deutschland angesehen werden, das nunmehr durch eine ebenfalls sehr umfangreiche Begleitforschung zum Sozialgesetzbuch II ergänzt wird, welches mit dem Vierten Gesetz für Moderne Dienstleistungen 2005 eingeführt wurde.

Schneider und *Uhlendorff* untersuchen für die Phasen vor und nach der Reform die Wirkung der Teilnahme an berufsbezogener und -übergreifender Weiterbildung. Sie

¹⁰ Vgl. hierzu *Brinkmann/Koch/Mendius* (2006) und insbesondere den dort enthaltenen Beitrag von *Lutz* (2006).

greifen auf die Prozessdaten der BA zurück und schätzen ausgehend von einem Propensity-Score-Matching Zugangswahrscheinlichkeiten und Übergangsraten in Beschäftigung. Die Ergebnisse zeigen, dass die Teilnahme bereits vor der Reform zu einem Anstieg der Beschäftigungswahrscheinlichkeit geführt hat. Für die noch relativ kurze Beobachtungsperiode nach der Reform deutet sich an, dass die Effektivität der Maßnahme zugenommen hat.

Brussig, Heinemann, Jaenichen und *Zwick* untersuchen die 2003 neu eingeführte Entgeltssicherung für Arbeitnehmer ab 50 Jahren, durch die ein Anreiz zur raschen Aufnahme einer Beschäftigung trotz Lohneinbußen gegeben werden soll. Die mikroökonomische Studie lässt keine solche Wirkung erkennen, was sich z. T. schon durch die geringe Inanspruchnahme erklärt. Der zugleich durchgeführten Implementationsstudie folgend, sind die geringe Inanspruchnahme und die ausbleibende Wirkung der Entgeltssicherung auf einen geringen Bekanntheitsgrad, einen zurückhaltenden Instrumenteneinsatz durch die Vermittler, unklare und widersprüchliche Förderbestimmungen und die Tatsache zurückzuführen, dass die Entgeltssicherung bestehende Akzeptanzprobleme für die Aufnahme der neuen Beschäftigung bei potenziellen Nutzern mehrheitlich nur teilweise beseitigt.

Caliendo, Kritikos und *Wießner* untersuchen die Existenzgründungsförderung von Arbeitslosen, die neben dem traditionellen Überbrückungsgeld den mit den Hartz-Gesetzen neu eingeführten Existenzgründungszuschuss umfasst.¹¹ Erste Ergebnisse deuten darauf hin, dass beide Programme z. T. unterschiedliche Zielgruppen erreichen und hinsichtlich der Vermeidung einer Rückkehr in die Arbeitslosigkeit als erfolgreich angesehen werden können. Wegen der vergleichsweise langen Förderdauer von bis zu drei Jahren sind endgültige Aussagen zum Existenzgründungszuschuss allerdings erst zu einem späteren Zeitpunkt möglich.

Bernhard, Wolff und *Jozwiak* untersuchen die Zugangswahrscheinlichkeit von arbeitslosen Personen im Rechtskreis des SGB II in zwei ausgewählte Maßnahmebereiche (Beauftragung Dritter mit der Vermittlung und Trainingsmaßnahmen). Sie stützen sich hierbei auf die um erwerbsfähige Hilfebedürftige ergänzte integrierte Datenbasis der BA. Die nunmehr vorliegenden ersten Ergebnisse sprechen nicht generell dafür, dass Personengruppen, die bereits in der Vergangenheit als schwer vermittelbar galten, mit diesen Instrumenten gezielt gefördert wurden. Festgestellt wurde allerdings, dass insbesondere junge Erwachsene einbezogen wurden.

Mauer untersucht die Frage, ob die Neuausrichtung der Arbeitsvermittlung im neuen Kundenzentrum (KuZ) Abgänge aus Arbeitslosigkeit in eine Beschäftigung des ersten Arbeitsmarktes erhöht. Aufgrund der stufenweisen Einführung des Kundenzentrums konnten – analog zur Bildung „statistischer Zwillinge“ auf der Individualebene – auf Basis zentraler Arbeitsmarktindikatoren ähnliche Kontrollagenturen gefunden werden, bei denen die Umstellung noch nicht stattgefunden hatte. Mit Hilfe von zufälligen Stichproben des Arbeitslosenzugangs und von Übergangsratenmodellen wurde dann der „KuZ-Effekt“ gemessen. Im Ergebnis konnte kein Gesamteffekt der zum Beobachtungszeitraum umgesetzten Elemente der reformierten Arbeitsvermittlung gefunden werden. Erkennbar waren Tendenzen, dass sich die insgesamt auf größere Effizienz des Ressourceneinsatzes ausgerichtete reformierte Arbeitsvermittlung im KuZ tendenziell negativ auf die Vermittlungschancen benachteiligter Arbeitsmarktgruppen auszuwirken scheint. Ausdrücklich wird darauf hingewiesen, dass angesichts der noch laufenden Entwicklungen im neuen Kundenzentrum eine abschließende Bewertung noch nicht möglich sei.

Fertig, Kluge und *Schmidt* untersuchen schließlich die kausale Wirkung des regional unterschiedlichen Einsatzes arbeitsmarktpolitischer Instrumente auf Stromgrößen des

¹¹ Zwischenzeitlich wurden beide Instrumente in ein einheitliches überführt.

Arbeitsmarktes (Nettoaustritte aus der Arbeitslosigkeit bzw. Langzeitarbeitslosigkeit in nicht geförderte Beschäftigung). Die vorgelegten Schätzergebnisse legen den Schluss nahe, dass nur wenige Maßnahmen im Untersuchungszeitraum von 2000 bis 2004 solche „Makro-Wirkungen“ am Arbeitsmarkt erzielt haben und dass kaum Unterschiede für die Zeit vor und nach den Hartz-Reformen existierten. Zu berücksichtigen ist allerdings, dass die Implementation wichtiger Elemente der Hartz-Reformen zur Effizienzsteigerung des Instrumenteneinsatzes im Untersuchungszeitraum noch voll im Gange war (z. B. Einführung von Kundenzentren) bzw. z. T. sogar erst begonnen hatte (z. B. die Einführung von „Handlungsprogrammen“).

3 Ein Zwischenresümee

Die vorliegenden Beiträge lassen den großen Fortschritt erkennen, den Evaluationsforschung in den letzten Jahren erzielen konnte – auf allen Ebenen, namentlich aber im Bereich der Mikro-Ökonometrie. Nach langjähriger und mühsamer Entwicklungsarbeit ermöglichen die nunmehr für Forschungszwecke verfügbaren Prozessdaten der BA differenzierte und auch für Politikberatung hinreichend zuverlässige Analysen. Korrigiert werden können damit auch überwiegend pauschale Einschätzungen zur (Un-)Wirksamkeit der Arbeitsförderung, die in der zweiten Hälfte der neunziger Jahre auf der Basis von Umfragedaten entstanden waren, welche aber für Evaluationszwecke allenfalls begrenzt tauglich waren.

Dabei bleibt herauszustellen, dass auch das jetzige Wissen noch sehr lückenhaft ist. Weitere Forschung bleibt – angesichts des hohen Mitteleinsatzes – dringend erforderlich, um diejenigen Konstellationen einzugrenzen, in denen der Maßnahmeinsatz den Teilnehmern auch wirklich hilft, in denen möglichst auch der Arbeitsmarkt insgesamt positiv beeinflusst wird und sich die hohen Kosten in gewissem Umfang auch „rechnen“ (mit gegenwärtig noch notwendigen Abstrichen hinsichtlich der Quantifizierbarkeit und „Monetarisierung“ von Erträgen).

Bei Instrumenten und Einsatzbereichen ohne nachweisbare kurzfristigen Eingliederungserfolge bleiben z. B. Fragen zu beantworten nach möglichen längerfristigen Effekten; bei nachgewiesenen Eingliederungserfolgen (Mikro-Ebene) bleiben Fragen nach den indirekten Effekten im Sinne von Calmfors (1994) zu untersuchen, die bei letztlich anzustrebenden Kosten-Nutzen-Erwägungen den Ausschlag geben können. Und hinzuweisen ist auf die Vielzahl von weiterhin zu lösenden datentechnischen, methodischen, theoretischen und forschungsorganisatorischen Problemen, auf die hier nicht näher eingegangen wurde (vgl. hierzu Brinkmann/Koch/Mendius 2006). Evaluationsforschung bleibt anstrengend und spannend, eine Herausforderung für Wissenschaft und Praxis!

Trotz der erkennbaren Fortschritte wäre es verfrüht, zum jetzigen Zeitpunkt ein quasi abschließendes Urteil auch nur für Teilbereiche der Arbeitsförderung abzugeben. Nicht nur, dass die Vielzahl der in den letzten Jahren angestoßenen und weitgehend noch nicht abgeschlossenen Evaluationsprojekte erst noch zu Ende zu führen sind. Es sind namentlich die noch nicht abgeschlossenen Umstrukturierungen in der BA selbst (sie sollen ja die Effizienz des Maßnahmeinsatzes deutlich steigern), die notwendigerweise zu einem Vorbehalt hinsichtlich der künftigen Geltung der vorliegenden Forschungsergebnisse führen. Wertlos werden diese aber dadurch keinesfalls: Weisen sie doch auf Chancen und Risikobereiche hin, die es bei der weiteren Entwicklung besonders zu beachten gilt.

Angesichts des bislang vorliegenden „bunten Straußes“ von Ergebnissen mit vielen „Tupfern“ bestätigen sich Erfahrungen aus dem Ausland: In keinem Teilbereich sind Ergebnisse so durchschlagend positiv oder negativ, dass pauschale Urteile über ein-

zelne Instrumente möglich wären. Klar wird einerseits, dass Arbeitsmarktpolitik auch in Deutschland keine Wunder bewirken kann hinsichtlich der Bewältigung der Massenarbeitslosigkeit (vgl. Hemström 2006 mit entsprechenden Ergebnissen und Einschätzungen für Schweden). Andererseits bestehen aber Chancen hinsichtlich eines zielgruppenspezifischen Einsatzes von Fördermaßnahmen in bestimmten Konstellationen (z. B. auch bezogen auf die jeweilige regionale Arbeitsmarktsituation), für die Erfolgchancen nachweisbar sind und für die sich Kosten im Hinblick auf die Erträge „rechnen“. Evaluationsforschung in Deutschland ist zunehmend in der Lage, solche wissenschaftlich fundierten Nachweise zu erbringen – und genau dies ist der unschätzbare Fortschritt gegenüber der Situation noch vor einigen Jahren.

Literatur¹²

Arbeit und Beruf (1931): Halbmonatszeitschrift für Fragen des Arbeitsmarktes, der Arbeitslosenversicherung, der Berufsberatung und verwandter Gebiete im Deutschen Reich und in Österreich, 10. Jahrgang, Grüner-Verlag, Bernau bei Berlin.

Allmendinger, Jutta/Kohlmann, Annette (2005): Datenverfügbarkeit und Datenzugang am Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung. In: Allgemeines Statistisches Archiv, Jg. 89, Nr. 2, S. 159–182.

Autorengemeinschaft (1973): Alternativrechnungen zur Entwicklung des Arbeitsmarktes 1974. Annahmen, Maßnahmen und ihre Auswirkungen. In: MittAB, Nr. 4.

Blankenburg, Erhard (1978): Mehr steuern durch weniger Kontrolle. Einige Lehren aus einem Ländervergleich von Arbeitsverwaltungen. Discussion Paper des IIMV, Wissenschaftszentrum Berlin.

Blankenburg, Erhard/Krautkrämer, Uta (1977): Die Handlungsspielräume der Arbeitsverwaltung. Discussion Paper des IIMV, Wissenschaftszentrum Berlin.

Bolte, Karl Martin (1969): Untersuchung über Ergebnisse der beruflichen Fortbildung im Rahmen des „Individuellen Förderungsprogramms“. In: Mitt(IAB), Nr. 7.

Bosch, Gerhard/Seifert, Hartmut/Spies, Bernd-Georg (1983): Arbeitsmarkt und Gewerkschaftliche Interessenvertretung. Endbericht, Düsseldorf.

Brinkmann, Christian (2000): Zur Evaluation der aktiven Arbeitsförderung nach dem SGB III. Monitoring und Wirkungsforschung im Umbruch. In: MittAB, Nr. 3.

Brinkmann, Christian/Dadzio, Werner/König, Paul/Kraft, Hermine/Richter, Hans-Jürgen (1970): Ergebnisse der beruflichen Fortbildung im Rahmen des „Individuellen Förderungsprogramms“. BeitrAB, Nr. 2.

Brinkmann, Christian/Koch, Susanne/Mendijs, Hans Gerhard (Hrsg.) (2006): Wirkungsforschung und Politikberatung – eine Gratwanderung? BeitrAB 300, Nürnberg.

Calmfors, Lars (1994): Active Labour Market Policy and Unemployment. A framework for the analysis of crucial design features. In: OECD Labour Market and Social Policy Occasional Paper, Nr. 15, Paris.

Ernst, Angelika (1974): Arbeitsmarktpolitische Prioritäten für die regionale Wirtschaftsförderung. Ein quantifizierter Orientierungsrahmen. In: MittAB, Nr. 3.

Hellmich, Andrea (1983): Arbeitsbeschaffungsmaßnahmen zur Verbesserung der sozialen Dienste und der sozialen Infrastruktur. Forschungsbericht 88 der Reihe Sozialforschung des Bundesministers für Arbeit und Sozialordnung, Bonn.

¹² MittAB = Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, Jahrgänge 1970–2003.

Ab 2004: Zeitschrift für ArbeitsmarktForschung (ZAF).

Mitt(IAB) = Mitteilungen aus dem IAB, Jahrgänge 1968/1969 (Vorläufer von MittAB).

IIMV = Internationales Institut für Management und Verwaltung des Wissenschaftszentrum Berlin (für Sozialforschung).

- Hemström, Maria* (2006): Active Labour Market Policies – the case of Sweden. In: Brinkmann, Christian/Koch, Susanne/Mendius, Hans Gerhard (Hrsg.): Wirkungsforschung und Politikberatung – eine Gratwanderung?, BeitrAB 300, Nürnberg.
- Hujer, Reinhard/Bellmann, Lutz/Brinkmann, Christian* (2000): Evaluation aktiver Arbeitsmarktpolitik – Probleme und Perspektiven. Einführung in das Schwerpunktheft. In: MittAB, Nr. 3.
- IAB* (2005): IAB-Forschungsberichte, Nr. 1–8: <http://iab.de/iab/publikationen/onlinepub.htm>.
- IAB* (1972): Quintessenzen aus der Arbeitsmarkt- und Berufsforschung 1968–1971, Nürnberg.
- Kohlmann, Annette* (2005): The Research Data Centre of the Federal Employment Service in the Institute for Employment Research. In: Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, Jg. 125, Nr. 3, S. 327–337.
- Kruppe, Thomas* (2006): Wirkungsanalyse der Förderung beruflicher Weiterbildung im Spannungsverhältnis von Monitoring und Evaluation. In: Brinkmann, Christian/Koch, Susanne/Mendius, Hans Gerhard (Hrsg.): Wirkungsforschung und Politikberatung – eine Gratwanderung?, BeitrAB 300, Nürnberg.
- Lauritzen, Gisela* (1980): Möglichkeiten und Hemmnisse stärkerer Regionalisierung der Handlungskompetenzen von Arbeitsämtern und ihrer Verwaltungsausschüsse. Ergebnisse einer qualitativen Expertenbefragung. Gutachten im Auftrag des Bundesministers für Arbeit und Sozialordnung, Strasbourg (unveröffentlichtes Manuskript).
- Lutz, Burkart* (2006): Wirkungsforschung und Politikberatung. In: Brinkmann, Christian/Koch, Susanne/Mendius, Hans Gerhard (Hrsg.): Wirkungsforschung und Politikberatung – eine Gratwanderung?, BeitrAB 300, Nürnberg.
- Mendius, Hans Gerhard/Sengenberger, Werner/Köhler, Christoph/Maase, Mira* (1983): Qualifizierung im Betrieb als Instrument der öffentlichen Arbeitsmarktpolitik. Forschungsbericht 89 der Reihe Sozialforschung des Bundesministers für Arbeit und Sozialordnung, Bonn.
- Mertens, Dieter* (1972): 5 Jahre Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit. In: Arbeit, Beruf und Arbeitslosenhilfe – Das Arbeitsamt, Jahrgang 1972, Heft 5, 162–167.
- Mertens, Dieter/Reyher, Lutz/Kühl, Jürgen* (1988): Ziele und Möglichkeiten von Wirkungsanalysen. In: Mertens, Dieter (Hrsg.) (1988): Konzepte der Arbeitsmarkt- und Berufsforschung. BeitrAB, Nr. 70.
- Mertens, Dieter/Reyher, Lutz/Kühl, Jürgen* (1981): Ziele und Möglichkeiten von Wirkungsanalysen. In: MittAB, Nr. 3.
- Peters, Aribert B./Schmid, Günther* (1982): Aggregierte Wirkungsanalyse des arbeitsmarktpolitischen Programms der Bundesregierung für Regionen mit besonderen Beschäftigungsproblemen. Analyse der Beschäftigungswirkungen. Discussion Paper des IIMV, Wissenschaftszentrum Berlin.
- Reyher, Lutz* (1968): Über die Auswirkungen der beiden Konjunkturprogramme von 1967 auf das Arbeitsvolumen. In: Mitt(IAB), Nr. 3.
- Scharpf, Fritz W./Garlichs, Dietrich/Maier, Friederike/Maier, Hans* (1982): Implementationsprobleme offensiver Arbeitsmarktpolitik. Frankfurt/Main, New York.
- Scharpf, Fritz W./Schmid, Günther* (1981): Forschungsprogramm „Arbeitsmarktpolitik“ am Wissenschaftszentrum Berlin. Ansätze zur vergleichenden Wirkungsanalyse. In: MittAB, Nr. 3.
- Schmuhl, Hans-Walter* (2003): Arbeitsmarktpolitik und Arbeitsverwaltung in Deutschland 1871–2002, BeitrAB 270, Nürnberg.
- Stephan, Gesine/Brinkmann, Christian* (2006): Aufbruchstimmung? Frischer Wind im Evaluationsgeschäft. In: Brinkmann, Christian/Koch, Susanne/Mendius, Hans Gerhard (Hrsg.): Wirkungsforschung und Politikberatung – eine Gratwanderung?, BeitrAB 300, Nürnberg.
- Tiefenbacher, Rudolph* (1931): Die volkswirtschaftliche Bedeutung der Produktiven Erwerbslosenfürsorge unter Zugrundelegung der Badischen Verhältnisse. Inauguraldissertation zur Erlangung der Staatswissenschaftlichen Doktorwürde einer Hohen Philosophischen Fakultät der Badischen Ruprecht-Karls-Universität zu Heidelberg.