

Das TrEffeR-Projekt der Bundesagentur für Arbeit: Die Wirkung von Maßnahmen aktiver Arbeitsmarktpolitik*

*Gesine Stephan, Susanne Rässler und Torben Schewe***

Der Beitrag stellt erstens das Projekt TrEffeR „Treatment Effects and Prediction“ (TrEffeR) der Bundesagentur für Arbeit (BA) vor. Eine Aufgabe des Projektes ist die zeitnahe und an Steuerungszwecken orientierte Untersuchung der Effektivität der Arbeitsförderung. Die Ergebnisse einer regional differenzierten retrospektiven Wirkungsanalyse sollen die Arbeitsagenturen laufend dabei unterstützen, ex post Stärken und Schwächen bei der Vergabe von Maßnahmen zu identifizieren. Der für das Projekt aufgebaute Mikrodatensatz lässt sich auch dazu nutzen, einen Gesamteindruck von der Wirkung arbeitsmarktpolitischer Instrumente zu gewinnen: Der Beitrag untersucht zweitens auf Basis der neuen Datenbank, ob die Förderung mit arbeitsmarktpolitischen Maßnahmen bei Personen, die im September 2002 in eine Maßnahme eingetreten sind, zur Vermeidung von Arbeitslosigkeit beitragen konnte.

Gliederung

- 1 Einleitung
- 2 Wirkungsanalyse in der Bundesagentur für Arbeit
 - 2.1 Der Bedarf
 - 2.2 Die Datenbank
 - 2.3 Weitere Entwicklungsmöglichkeiten
- 3 Matching-Analysen für die wichtigsten Instrumente aktiver Arbeitsmarktpolitik
 - 3.1 Instrumente und Literaturüberblick
 - 3.2 Untersuchungsansatz
 - 3.3 Ergebnisse
- 4 Fazit

Literatur

Anhang

* Der Beitrag wurde November 2006 von der/dem betreuenden Herausgeber/in zur Publikation freigegeben.

** Sven Schütt hat den Aufbau des TrEffeR-Projekts in der BA initiiert und steht dem Projekt mit Rat und Tat zur Seite. Donald B. Rubin hat das TrEffeR-Projekt mit entwickelt und wissenschaftlich begleitet. Wichtige Akzente setzte Heinrich Alt. Der Aufbau der TrEffeR-Datenbank aus den Daten der BA wäre nicht möglich gewesen ohne die Mitarbeiter des Bereichs PP4 Produkt- und Programmanalyse – aus dem insbesondere Michael Irskens, Rolf Lehmann und Britta Ullrich wertvolle Beiträge geleistet haben – sowie die Datenbank-Spezialisten Jörg Bernlöhner und Roman Ruckerl. Hilfreiche Hinweise zu früheren Fassungen des Manuskripts gaben zwei anonyme Gutachter, Susanne Koch, Dirk Oberschachtsiek und Hermann Roß sowie die Teilnehmerinnen und Teilnehmer des Autorenworkshops zum ZAF-Schwerpunktheft 2006 „Evaluation der Arbeitsmarktpolitik in Deutschland“ am 26./27. 4. 2006 in Nürnberg. Johanna Rapp unterstützte uns bei der Tabellerstellung. Ihnen allen sei herzlich gedankt.

1 Einleitung

Das sozialpolitische System der Bundesrepublik Deutschland befindet sich in einem umfassenden Reformprozess. Im Bereich der Arbeitsförderung wurde bereits durch das Job-AQTIV-Gesetz im Jahr 2002 eine Neuausrichtung der Arbeitsvermittlung eingeleitet. Fortgesetzt wurde die Entwicklung auf Basis der Empfehlungen der Hartz-Kommission, deren Aufgabe es war, Strategien für eine effizientere Gestaltung der Arbeitsmarktpolitik in Deutschland zu entwickeln. Das erste und zweite Gesetz für moderne Dienstleistungen am Arbeitsmarkt traten zu Beginn des Jahres 2003 in Kraft. Beide Gesetze zielen auf eine Verbesserung der Vermittlung sowie der Anreize zur Aufnahme einer Beschäftigung ab. Zudem wurden viel beachtete neue Instrumente – wie der Existenzgründungszuschuss und die Personal-Service-Agentur – geschaffen und bestehende Instrumente reformiert. Mit dem dritten Gesetz für moderne Dienstleistungen am Arbeitsmarkt 2004 begann der Umbau der ehemaligen Bundesanstalt für Arbeit in die serviceorientierte Bundesagentur für Arbeit (BA). Schließlich wurden 2005 durch das vierte Gesetz für moderne Dienstleistungen am Arbeitsmarkt Arbeitslosenhilfe und Sozialhilfe zusammengelegt. Einhergehend mit diesen Arbeitsmarktreformen sind – mit der Einführung der neuen Steuerungslogik über regionale Ziele im Jahr 2004 – auch die Anforderungen an die operative Steuerung der BA gestiegen.

Der vorliegende Beitrag gibt einen Ausblick auf das System der quantitativen Wirkungsanalyse arbeitsmarktpolitischer Maßnahmen, das derzeit in der BA entwickelt wird und in Zukunft die operative Steuerung der Agenturen unterstützen soll. Einen Gesamteindruck von den im Rahmen der Wirkungsanalyse untersuchten Instrumenten wie auch von der neu entwickelten Datenbasis vermitteln Effektivitätsanalysen für Personen, die im September 2002 in Maßnahmen der aktiven Arbeitsmarktpolitik eingetreten sind.

Inzwischen liegen sorgfältige und differenzierte Evaluationen für viele der in der Bundesrepublik eingesetzten arbeitsmarktpolitischen Instrumente vor (s. für einen selektiven Überblick Abschnitt 3.1). Dies ist nicht zuletzt darauf zurückzuführen, dass sich die Datengrundlage für eine Evaluation aktiver Arbeitsmarktpolitik in Deutschland deutlich verbessert hat. In den letzten Jahren wurde es möglich, Prozessdaten zu Zeiten der Arbeitssuche, des Leistungsbezugs und von Maßnahmeteilnahmen sowie zu Zeiten sozialversicherungspflichtiger Beschäftigung auf individueller Basis zusammenzuführen.

Pionierarbeit bei der Erschließung der administrativen Daten leisteten dabei Forschergruppen um die Professoren Hujer und Fitzenberger (beide Universität Frankfurt/Main) sowie Lechner (Universität St. Gallen), die in Kooperation mit dem Institut für Arbeitsmarkt- und Berufsforschung (IAB) umfangreiche Studien zur Effektivität von Arbeitsbeschaffungsmaßnahmen bzw. Bildungsmaßnahmen durchgeführt haben.

Im Sommer 2004 erging durch das damalige Bundesministerium für Wirtschaft und Arbeit ein umfassender Forschungsauftrag zur Evaluierung der Hartz-Gesetze I–III; die Zwischenberichte aus dem Jahr 2005 wurden im Januar des Jahres 2006 dem Deutschen Bundestag vorgelegt (Bundesministerium für Arbeit und Soziales 2006). Als eine zentrale Datengrundlage für die Hartz-Evaluation entstand am IAB der Forschungsdatensatz „Integrierte Erwerbsbiografien“ (IEB), der seit dem Jahr 2004 für Forschungszwecke zur Verfügung gestellt werden kann (Hummel et al. 2005).

Damit ist die Frage berechtigt, welchen zusätzlichen Beitrag eine BA-eigene Wirkungsanalyse überhaupt leisten kann. Wie im Folgenden deutlich wird, sind die spezifischen Informationsbedarfe der Steuerung ein maßgeblicher Grund für die Implementierung eines Systems der quantitativen Wirkungsanalyse in der BA. Ziel ist die Unterstützung des operativen Geschäfts der BA; einerseits durch Beiträge zum Monitoring, andererseits durch das Angebot direkter Hilfen für die Vermittler. Aufgabe des Monitorings ist es dabei, einen ersten Überblick über die Ergebnisse des Instrumenteneinsatzes der Arbeitsmarktpolitik bereitzustellen und Zielerreichungsgrade zu dokumentieren (Brinkmann 2000). Wirkungsanalyse in der BA ist damit notwendigerweise als „Rundumschlag“ angelegt, der alle Instrumente gleichermaßen – und zudem kleinräumig – analysieren muss. Hingegen gehen mikroanalytische Beiträge zur Evaluationsforschung in der Regel deutlich differenzierter vor, führen die Untersuchungen großräumiger durch und kombinieren sowohl quantitative als auch qualitative Aspekte.

Abschnitt 2 geht im Folgenden genauer auf das neue System der Wirkungsanalyse in der BA ein. Abschnitt 3 zeigt die Möglichkeiten der für die Wirkungsanalyse neu geschaffenen Datenbasis anhand von Matching-Analysen für die wichtigsten arbeitsmarktpolitischen Instrumente auf. Abschließend zieht Abschnitt 4 ein Fazit.

2 Wirkungsanalyse in der Bundesagentur für Arbeit

2.1 Der Bedarf

Die gesetzlich definierten Ziele der Arbeitsförderung – aus denen sich die geschäftspolitischen Ziele der BA ableiten – sind das Erreichen eines hohen Beschäftigungsstands, die Verbesserung der Beschäftigungsstruktur, die Vermeidung der Entstehung von Arbeitslosigkeit und die Verkürzung der Arbeitslosigkeitsdauer. Im Jahr 2004 wendeten die BA und der Bund für Leistungen der aktiven Arbeitsförderung 19,5 Mrd. Euro auf, davon 9,1 Mrd. Euro für Ermessensleistungen. Zum Vergleich: Im selben Jahr flossen 29,0 Mrd. Euro an Arbeitslosengeld und 18,8 Mrd. Euro an Arbeitslosenhilfe.

Ob sich die oben genannten Ziele eher durch den Einsatz aktiver Arbeitsmarktpolitik oder aber durch Änderungen der institutionellen Rahmenbedingungen des Arbeitsmarktes erreichen lassen, ist eine Frage, der hier nicht nachgegangen werden soll. Für die Steuerung der BA ist unter den gegebenen Rahmenbedingungen von Interesse, ob die beträchtlichen Mittel wirksam und wirtschaftlich eingesetzt wurden. Denn Umfang, inhaltliche Ausgestaltung und Anwendungsbereiche arbeitsmarktpolitischer Maßnahmen werden ausgesprochen kontrovers diskutiert. In diesem Sinne verfolgt auch die Reform der BA eine neue Steuerungslogik, die sich verstärkt an den Wirkungen und dem Aufwand von Maßnahmen orientiert.

Die von der Statistik der BA erstellten Verbleib- und Eingliederungsanalysen im Rahmen der Eingliederungsbilanz nach § 11 SGB III untersuchen den Verbleib in Arbeitslosigkeit oder Folgeförderung wie auch den Übergang in sozialversicherungspflichtige Beschäftigung im Anschluss an eine Maßnahme (Bundesagentur für Arbeit 2005). Dieses sind zwar ausgesprochen wichtige Informationen, sie sind in Hinsicht auf die Effektivität der Maßnahmen aber nicht aussagekräftig (siehe z. B. Caliendo/Jahn 2004): Findet ein Teilnehmer im Anschluss an eine arbeitsmarktpolitische Maßnahme eine Beschäftigung, so sagt dies noch nichts über den Beitrag der Maßnahme hierzu aus. Es fehlt eine adäquate Vergleichssituation – wie hätte sich der Verbleib der Geförderten ohne die Förderung entwickelt? Zudem werden beide Quoten umso höher liegen, desto geringer der Anteil der Maßnahmen Teilnehmer ist, der Problemgruppen des Arbeitsmarktes zuzurechnen ist. Für eine erste Messung der Effektivität arbeitsmarktpolitischer Maßnahmen wurde in der BA in der Folge ein Maß eingeführt,

das anzeigt, inwieweit der Einsatz von Fördermaßnahmen innerhalb von Personengruppen, die mit Hilfe von Data-Mining-Techniken abgegrenzt wurden, mit einer Verkürzung der Zeiten von Arbeitslosigkeit und Maßnahmenteilnahme einhergeht. Auch dieses Maß kann aufgrund seiner Konstruktionsweise jedoch noch nicht direkt als kausaler Effekt einer Maßnahme interpretiert werden.

Vom Zentralbereich Produkte und Programme (PP) der BA, gemeinsam mit dem IAB und der Harvard University (Cambridge), wurde daher im Jahr 2004 der Aufbau einer umfassenden quantitativen Wirkungsanalyse für die operative Steuerung eingeleitet. Diese stützt sich – unter strikter Wahrung des Datenschutzes – auf die Prozessdaten, die durch die BA erhoben und aufbereitet werden. Dieser Schritt war zwingend erforderlich, da die im IAB entwickelten IEB für die operative Steuerung der BA nicht direkt nutzbar sind – der Datenschutz verbietet die Verwendung der Forschungsdaten des IAB für operative Zwecke.

Das Projekt „Treatment Effects and Prediction“ (TrEffeR) hat sich zwei zentrale Ziele gesetzt: Erstens erfolgt eine nach Arbeitsagenturen differenzierte retrospektive Wirkungsanalyse der zahlreichen durch die BA administrierten Maßnahmen und Programme. Den verwendeten statistischen Matching-Verfahren liegt die Idee zugrunde, eine Vergleichsgruppe von Nicht-Teilnehmern zu bilden, die die gleichen relevanten Merkmale wie die Teilnehmer aufweisen (z. B. Rubin 1977; Heckman et al. 1999). Zweitens soll die Wirkungsanalyse in der Zukunft den Vermittlern im Rahmen eines Targeting-Systems (s. für andere Länder Eberts/O’Leary 2002; Colpitts 2002; Frölich et al. 2004) statistisch valide Prognosen der kausalen Effekte von verschiedenen Maßnahmen für individuelle Arbeitssuchende anbieten. Diese Prognosen können dann als Orientierungshilfe für die Zuweisung in Maßnahmen dienen. Beide Vorhaben greifen auf das Konzept der potenziellen Ergebnisse zurück (Rubin 1977), da sie eine Prognose darüber erfordern, wie sich die individuelle Erwerbsbiografie bei einer Nichtteilnahme entwickelt hätte.

Bei der Retrospektivanalyse geht es um eine zentrale Fragestellung der mikroökonomischen Evaluationsforschung: Haben sich die Arbeitsmarktchancen der Geförderten durch die Teilnahme an Maßnahmen der aktiven Arbeitsmarktpolitik im Mittel verbessert? Diese Frage wurde – wie erwähnt – für viele arbeitsmarktpolitische Instrumente bereits im Detail untersucht. Die in diesen Untersuchungen ermittelten Wirkungen gelten aber stets nur für den untersuchten Personenkreis und den untersuchten

Zeitraum; zudem ist von einer erheblichen regionalen Heterogenität der Wirkungen auszugehen. Die Leistung des TrEffeR-Projektes ist es, die Frage nach der Wirkung einer Maßnahme für die Geförderten kleinräumig für jeden Agenturbezirk und jedes Instrument (weiterhin differenziert nach Geschlecht sowie der Maßnahmendauer) einzeln zu beantworten und hierdurch wichtige Informationen für das Monitoring der BA bereitzustellen.

In Hinsicht auf die Entwicklung der Erwerbshistorie ist aus Sicht der BA einerseits eine Vermeidung von künftiger Arbeitslosigkeit, andererseits die Integration in sozialversicherungspflichtige Beschäftigung von Interesse. Dies soll im Folgenden kurz erläutert werden.

Ein zentrales geschäftspolitisches Ziel der BA ist die nachhaltige Verkürzung der „Dauer der faktischen Arbeitslosigkeit“. Die „Dauer der faktischen Arbeitslosigkeit“ ist eine im Rahmen der neuen Steuerungslogik der BA geschaffene ergänzende Definition der Arbeitslosigkeitsdauer. Sie folgt einem ähnlichen Konzept wie die schon bisher von der Statistik der BA verwendete Dauer der registrierten Arbeitslosigkeit gemäß § 18 (2) SGB III und umfasst die Dauer registrierter Arbeitslosigkeit zuzüglich der Dauer von Maßnahmeteilnahmen. Hintergrund ist, dass Teilnehmer an den meisten arbeitsmarktpolitischen Maßnahmen nach der gesetzlichen Definition nicht arbeitslos sind. Sie sind aber auch nicht ungefördert im ersten Arbeitsmarkt beschäftigt, sondern der Versicherungsgemeinschaft entstehen entsprechende Kosten. Unterschieden werden generell zwei Varianten der „Dauer der faktischen Arbeitslosigkeit“: In der einen Variante werden den Maßnahmedauern auch beschäftigungsbegleitende Leistungen (BBL) wie Eingliederungszuschüsse zugerechnet, während dies in der anderen Variante nicht der Fall ist. Im Zentrum des Interesses steht insbesondere die Zielgröße „Dauer der faktischen Arbeitslosigkeit inklusive BBL“. Damit wird eine strengere Definition der Integration von Arbeitslosen in Beschäftigung verfolgt und die Kostenwirksamkeit des Maßnahmeinsatzes unterstrichen. Die weitere Verwendung des Begriffs der „Dauer der faktischen Arbeitslosigkeit“ in diesem Beitrag bezieht die beschäftigungsbegleitenden Leistungen ebenfalls mit ein.

Eine Vermeidung von Arbeitslosigkeit oder Förderung impliziert dabei nicht notwendigerweise eine Integration in ungeforderte sozialversicherungspflichtige Beschäftigung, sondern kann auch daraus resultieren, dass Personen z. B. in die Nichterwerbstätigkeit eintreten oder sich selbständig machen. Aus Sicht der Arbeitslosenversicherung sind sowohl

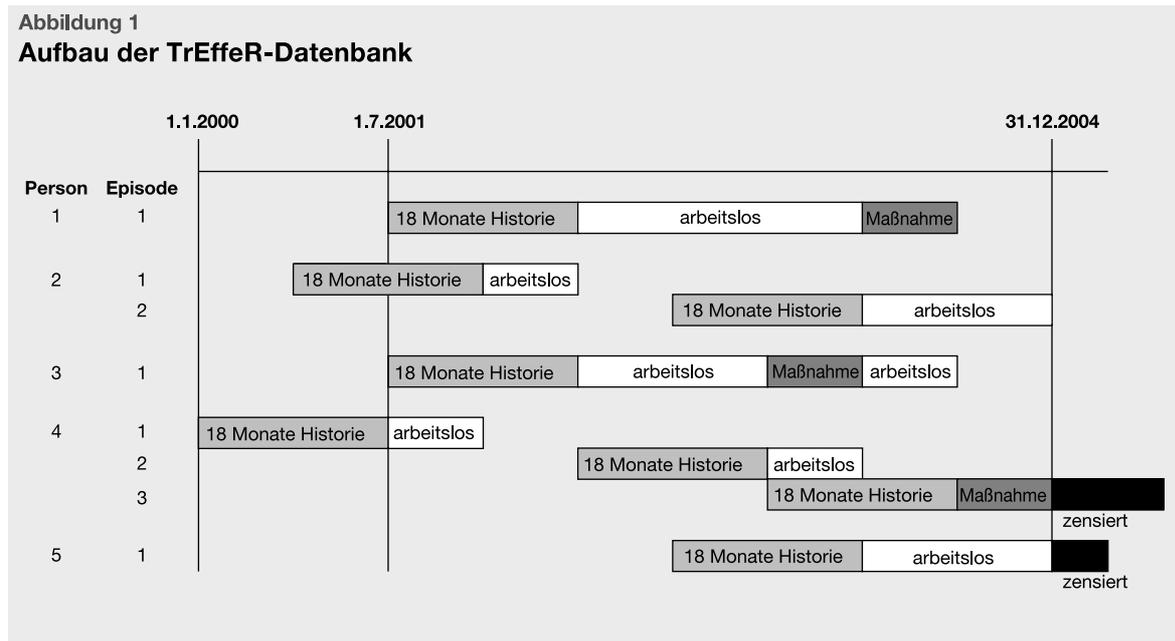
eine möglichst kurze „Dauer der faktischen Arbeitslosigkeit“ als auch die Integration in sozialversicherungspflichtige Beschäftigung, die neben der Vermeidung von Ausgaben auch zusätzliche Einnahmen für die Sozialversicherung bedeutet, sowie das dort erzielte Einkommen von Interesse. Hieraus kann durchaus ein Zielkonflikt bei der Steuerung der Agenturen folgen: Einerseits ist eine Verkürzung der „Dauer faktischer Arbeitslosigkeit“ anzustreben; dies impliziert einen Verzicht auf arbeitsmarktpolitische Maßnahmen längerer Dauer. Andererseits könnten sich solche Maßnahmen in der langen Frist auszahlen, wenn sie die Beschäftigungschancen und das Einkommenspotenzial stark genug erhöhen.

Gegenwärtig beschränkt sich das TrEffeR-Projekt (noch) auf den erstgenannten Verbleibsnachweis. Geprüft wird also, inwieweit eine Förderung bei den Geförderten zu einer Vermeidung von Arbeitslosigkeit und weiteren Maßnahmezeiten beitragen konnte.

2.2 Die Datenbank

Wie bereits erwähnt wurde, dürfen die Forschungsdaten des IAB für operative Zwecke der BA nicht verwendet werden; daher musste für die Zwecke der Wirkungsanalyse in der BA eine neue Datenbank aufgebaut werden. Bei der TrEffeR-Datenbank handelt es sich um einen Auszug aus den Datenbeständen der BA, die für die Zwecke der Wirkungsanalyse strukturiert und aufbereitet wurden. Die erste Version der Datenbank umfasst Informationen für die Jahre 2000 bis 2004. In den dem IAB zur Verfügung stehenden IEB ist der für eine Evaluation zur Verfügung stehende Zeitraum hingegen länger; Meldungen aus BeH und LeH sind bereits 1990 integriert. Ab 2005 kann – aufgrund der Zusammenlegung von Arbeitslosenhilfe und Sozialhilfe zum Arbeitslosengeld II, das nur noch zum Teil durch die BA administriert wird – zur Zeit nicht davon ausgegangen werden, dass die Erwerbsbiografien der betrachteten Personen in den Datenbeständen der BA vollständig verfügbar sind.

Die TrEffeR-Datenbank kombiniert dabei Informationen aus den Fachverfahren der computerunterstützten Arbeitsvermittlung coArb, der computerunterstützten Leistungsgewährung coLei und der computerunterstützten Sachbearbeitung coSach. Ausgangspunkt der Datensatzerstellung sind Personen, die mindestens eine Arbeitslosigkeitsmeldung bei der BA aufweisen. Abbildung 1 verdeutlicht die Struktur der Datenbank. Grundsätzlich ist die Datenbank in Form von Episoden aufgebaut. Diese be-



ginnen mit der Meldung einer Person als arbeitslos und/oder einem Maßnahmen eintritt und enden, wenn die Person nicht mehr arbeitslos ist bzw. gefördert wird.

An dieser Stelle sei kurz auf ein mögliches Selektivitätsproblem hingewiesen, dem sich generell alle auf Prozessdaten basierenden Wirkungsanalysen gegenübersehen: Insbesondere erwerbslose Arbeitssuchende ohne Ansprüche aus der Arbeitslosenversicherung könnten auf eine Arbeitslosigkeitsmeldung verzichten oder sich erst dann arbeitslos melden, wenn ihnen eine Maßnahme in Aussicht gestellt wird. Dann würden erfasste Zeiten der Arbeitslosigkeit erstens nicht die wahre Arbeitslosigkeitsdauer widerspiegeln. Zweitens lassen sich bei der Bildung von Vergleichsgruppen für Teilnehmer arbeitsmarktpolitischer Maßnahmen nur solche Personen berücksichtigen, die sich auch tatsächlich arbeitslos gemeldet haben. Die Ergebnisse des TrEffeR-Projektes beschränken sich damit auf die Gruppe registrierter Arbeitsloser.

Auf Basis der derzeit verfügbaren Informationen lassen sich zwei Verbleibsnachweise konstruieren: Dies ist erstens der stichtagsbezogene Verbleib in Arbeitslosigkeit oder Maßnahmen und zweitens die kumulierte „Dauer faktischer Arbeitslosigkeit“ nach Eintritt in eine Förderung. Förderzeiten werden anhand der tatsächlichen, nicht der geplanten Maßnahmedauer gemessen. Die tatsächliche Förderdauer ist insofern endogen, als dass der Abschluss einer Maßnahme mit persönlichen Merkmalen und dem Fördererfolg korreliert ist. Sie ist aus Sicht der

BA aus Kostengründen aber die interessantere Variable.

Die Datenbank enthält folgende personenbezogene Merkmale, jeweils gemessen zum Episodenbeginn: Alter, Geschlecht, Familienstand, Schulbildung, Berufsausbildung, Beschäftigungsstatus vor Zugang, Auswirkungen der Gesundheit auf die Vermittlung, Behinderungsgrad, Art des Leistungsbezugs, Nationalität, der Berufsbereich, gewünschte Arbeitszeit sowie Beginn der Arbeitslosigkeit. Hinzu werden auf der Regionalebene folgende Variablen auf der Berufsebene zugespielt: Anzahl Arbeitslose, Anzahl offener Stellen sowie Anzahl Einstellungen. Hintergrund ist, dass durch diese Variablen für die regionale Arbeitsmarktlage in bestimmten Berufsfeldern kontrolliert werden kann.

Für die Wirkungsanalyse ist es entscheidend, Informationen über die Erwerbshistorie der Geförderten zu berücksichtigen. Ergebnisse aus der Literatur weisen darauf hin, dass die Erwerbsbiografie bzw. Historie einer der wichtigsten Faktoren zur Erklärung von Teilnahmeentscheidung und Maßnahmeerfolg ist (Heckman et al. 1999; Lüdemann et al. 2006). Für den zur Verfügung stehenden begrenzten Zeitraum wurde zunächst entschieden, für jede Person eine 18 Monate lange Erwerbshistorie zu konstruieren und dann Arbeitslosenmeldungen und Maßnahmenzugänge ab dem 1.7.2001 zu analysieren (s. Abbildung 1). Die Historie enthält Informationen über: die Tage in Arbeitslosigkeit, in Maßnahmen, in Krankheit und des Leistungsbezugs sowie die Anzahl der Arbeitslosigkeitsperioden, Maß-

nahmen, Krankmeldungen, Ortswechsel, Leistungsbezugsperioden sowie Sperrzeiten in den 18 Monaten vor Beginn der untersuchten Episode.

Bei dem Aufbau der Datenbank waren dabei häufig pragmatische Entscheidungen über den Umgang mit verschiedenen Datenqualitätsproblemen zu treffen (s. Jaenichen et al. 2005 für Beispiele von Problemfällen, die bei dem Zusammenspielen von Prozessdaten auftreten können). Die Konstruktion von Kundenhistorien auf Basis der Meldungen aus den einzelnen computerunterstützten Fachverfahren erfordert eine Vielzahl von Plausibilitätsprüfungen. Die dabei vorzunehmenden Korrekturen beziehen sich einerseits auf zeitliche Überschneidungen sowie Lücken einzelner Meldungen von Arbeitslosigkeit, Maßnahmeteilnahme und Leistungsbezug. Dies betrifft vorwiegend die Dauern der einzelnen Meldungen, aber auch die Kombinationen von Meldungen, die nicht mit den zugrunde liegenden Gesetzestexten konform sind. Zudem ist mit dem Problem zum Teil fehlender oder variierender Werte der relevanten Hintergrundvariablen zu einzelnen Meldungszeiträumen umzugehen. Dabei bedarf es häufig einer Vielzahl von Annahmen, die eine abschließende fehlerfreie Abbildung der Realität nicht erlauben. Es verbleiben Konstellationen, bei denen sich lediglich fallweise Überprüfungen auf Basis von Stichproben vornehmen lassen, die anschließend indizienbasiert eine Verdichtung zu Bereinigungsregeln ermöglichen.

Dabei sind als wesentliche und häufig vorgenommene Bereinigungsschritte für die Datensatzerstellung hervorzuheben: a) Einzelne Arbeitslosigkeitsmeldungen werden zu Episoden zusammengefasst. b) Bei variierenden oder unsystematisch fehlenden Hintergrundvariablen werden generell die zum jeweiligen Zeitpunkt aktuell verfügbaren Informationen genutzt. c) Informationen zu gebuchten Maßnahmezeiträumen werden aufgrund der höheren Validität gegenüber Arbeitslosigkeitszeiträumen priorisiert. d) Leistungsbezugsräume werden im Falle von Unplausibilitäten zu Korrekturzwecken herangezogen. e) Sich überschneidende Verläufe von individuellen Maßnahmeteilnahmen wurden gezielt aus der Analyse ausgeschlossen. f) Kurze Unterbrechungen zwischen den Meldungen (z.B. aufgrund von Krankmeldungen) sind unschädlich für den Fortgang und somit die Dauer einer Episode.

2.3 Weitere Entwicklungsmöglichkeiten

Das TrEffeR-Projekt befindet sich in weiten Teilen noch in der Entwicklungsphase. Zukünftige Themen sind die Datenvalidierung, Sensitivitätsanalysen, die

Weiterentwicklung der Retrospektivanalyse und des Targeting-Tools, eine Trägerevaluation zur Unterstützung der Einkaufsplanung der BA sowie eine Makroanalyse. Einige dieser Punkte sollen im Folgenden kurz angesprochen werden.

Sensitivitätsanalysen haben sich im Rahmen des TrEffeR-Projektes bisher auf die Frage konzentriert, inwieweit sich die Ergebnisse der retrospektiven Wirkungsanalyse verändern, wenn alternative Matching-Techniken angewendet werden. Für die agenturspezifischen Auswertungen wurde zunächst flächendeckend ein „Neighest-Neighbour-Matching“ durchgeführt. Für ausgewählte Agenturen wurden die Ergebnisse mit denen eines genetischen Matching-Algorithmus' abgeglichen, der von der Harvard University direkt auf die Zwecke der BA zugeschnitten entwickelt wurde, dessen praktischer Einsatz bisher aber durch sehr viel längere Rechenzeiten verhindert wurde. Weitere Sensitivitätsanalysen sind geplant. So soll zum Beispiel geprüft werden, ob sich die Ergebnisse von Wirkungsanalysen verändern, wenn die Historie einen längeren Zeitraum umfasst.

Der Verbleib in sozialversicherungspflichtiger Beschäftigung sowie die Höhe des Arbeitseinkommens lassen sich mit der derzeitigen TrEffeR-Datenbank noch nicht untersuchen. Es ist wünschenswert, in Zukunft die – allerdings nur deutlich zeitverzögert verfügbaren – Meldungen sozialversicherungspflichtiger Beschäftigungsverhältnisse durch die Betriebe in die TrEffeR-Datenbank zu integrieren. Hierdurch ließen sich auch alternative Definitionen von Arbeitslosigkeit untersuchen (Fitzenberger/Wilke 2004). Ob und wieweit diese Erweiterungen datenschutzrechtlich zulässig sind, muss allerdings noch geprüft werden.

Der Nutzen der bisher vorliegenden Ergebnisse der Retrospektivanalyse ist dadurch begrenzt, dass die geschätzte Förderwirkung für die Geförderten immer nur für den jeweils untersuchten Personenkreis gilt. Ein paarweiser Vergleich der Ergebnisse über Maßnahmen und Agenturen ist auf Basis der vorliegenden Schätzungen noch nicht möglich – gerade dieser wäre aber aus Sicht der BA ausgesprochen interessant.

Für die Retrospektivanalyse sind Matching-Verfahren ein adäquates und vergleichsweise einfaches Verfahren. Anders stellt sich die Problematik in Hinsicht auf ein Targeting-System dar; hier muss die Prognose individuell zugeschnitten sein. Wünschenswert ist zudem, dass die konkrete Prognose vor Ort in den Arbeitsagenturen erstellt werden kann, was voraussetzt, dass die Erstellung der Prognose rechentechnisch nicht zu aufwändig gestaltet ist. Ein

möglicher Ansatz für ein Targeting-System (dessen Einsatztauglichkeit derzeit umfassend geprüft wird) besteht in der Schätzung von Treatment-Effects-Modellen (z.B. Greene 2003). Die so geschätzten Koeffizienten können dann vor Ort von Vermittlern zur Prognose von Maßnahmeeffekten herangezogen werden. Mit ersten Konzepttests in einer ausgewählten Agentur wurde bereits begonnen. Als eine mögliche Alternative wurde von der Harvard University kürzlich ein Ansatz vorgeschlagen, der eine Kombination von Matching-Verfahren und parametrischen Verfahren vorsieht.

Die Zielerreichung von Maßnahmen der aktiven Arbeitsmarktpolitik wird durch diverse indirekte Effekte unterstützt oder auch konterkariert (Calmfors 1994). Von der Öffentlichkeit mit besonderer Aufmerksamkeit verfolgt werden Mitnahme-, Substitutions- und Verdrängungseffekte arbeitsmarktpolitischer Maßnahmen. Ein Mitnahmeeffekt liegt vor, wenn eine Einstellung auch ohne die Maßnahme zustande gekommen wäre. Substitutions- bzw. Verdrängungseffekte sind zu beobachten, wenn Teilnehmer an einer arbeitsmarktpolitischen Maßnahme ungeforderte Erwerbspersonen im selben oder in anderen Unternehmen verdrängen. Bisher konzentriert sich das Projekt auf die Effektivität arbeitsmarktpolitischer Programme auf der Mikroebene. Um eine Messung der Effektivität arbeitsmarktpolitischer Maßnahmen unter Berücksichtigung indirekter Effekte zu erhalten, wäre eine ergänzende Makroanalyse unabdingbar. Dabei wären für Steuerungszwecke insbesondere die „Nettoeffekte“ von Maßnahmen auf der Agenturebene von Interesse.

Abschließend sei angemerkt, dass die Daten derzeit nur für Mitarbeiter der BA zugänglich sind. Es soll allerdings geprüft werden, ob es in Zukunft möglich ist, die Daten – unter strikter Wahrung des Datenschutzes – einem breiteren Nutzerkreis zugänglich zu machen. Dies würde mit einer ausführlichen Dokumentation der Datensatzerstellung verbunden sein. Der größte Vorteil der TrEffeR-Datenbank wäre dabei die weitgehende Aufbereitung der Daten, die dem Anwender viele Mühen spart. Dies ist jedoch auch gleichzeitig ihr größter Nachteil, da der Anwender nicht die Möglichkeit hat, eigene Entscheidungen über die Aufbereitung der Daten zu treffen.

3 Matching-Analysen für die wichtigsten Instrumente aktiver Arbeitsmarktpolitik

3.1 Instrumente und Literaturüberblick

In TrEffeR erfolgen agenturspezifische Auswertungen, denn nur hierdurch ist eine Unterstützung der

Vermittler in ihrer Tätigkeit vor Ort möglich. Die Ergebnisse für die einzelnen Agenturen werden dabei nicht BA-extern veröffentlicht. Jedoch lässt sich die Datenbank auch für eine Gesamtschau der Wirkungen der Maßnahmenteilnahme für die Geförderten nutzen. Dies vermittelt zudem einen Eindruck davon, was die gegenwärtig agenturspezifisch aufbereiteten Wirkungsanalysen in der BA leisten können bzw. was sie nicht leisten können. Bevor dies in der Folge geschieht, seien zunächst die untersuchten Maßnahmen skizziert und ausgewählte Befunde aus der neueren Literatur dargestellt. Zwar sind die Ergebnisse von Studien, die unterschiedliche Methoden und unterschiedliche Stichproben nutzen, nicht unmittelbar vergleichbar. Dennoch ist eine wichtige Frage, ob Wirkungsanalysen der BA in der Tendenz zu ähnlichen Ergebnissen kommen.

Das *Überbrückungsgeld* (ÜG) soll für Arbeitslose den Schritt in die Selbständigkeit erleichtern, indem während der Startphase der Gründung für sechs Monate eine Unterstützung in Höhe der Lohnersatzleistung (und zusätzlicher pauschalierter Sozialversicherungsbeiträge) gezahlt wird. Der Zwischenbericht der Hartz-Evaluation kommt insgesamt zu einer positiven Einschätzung der Förderwirkung von Existenzgründungen (IAB et al. 2005). Besonders wirksam scheint eine Förderung mit dem Überbrückungsgeld dabei für ältere Personen und Langzeitarbeitslose zu sein. Allerdings weist der Bericht auch auf mögliche Mitnahmeeffekte hin, die sich mit Hilfe von Matching-Techniken jedoch nicht quantifizieren lassen. Zudem ist gerade beim Überbrückungsgeld zu vermuten, dass – in den Prozessdaten nicht beobachtbare – motivationale Variablen eine wichtige Rolle für den Erfolg der Förderung spielen könnten. Angemerkt sei schließlich, dass die Überlebenschancen von Existenzgründungen aus der Arbeitslosigkeit heraus geringer zu sein scheinen als die ungeförderter Gründungen (Pfeiffer/Reize 2000).

Arbeitsbeschaffungsmaßnahmen (ABM) sind geförderte, sozialversicherungspflichtige Beschäftigungsverhältnisse. Noch bis Anfang 2004 wurden sie mit dem Ziel der Verbesserung der Eingliederungsaussichten vergeben. ABM werden durch Träger durchgeführt; sie müssen einerseits „zusätzliche“ Arbeiten sein, andererseits müssen sie im öffentlichen Interesse liegen. Vor 2002 – als diese Bedingung aufgehoben wurde – konnten nur Langzeitarbeitslose oder Arbeitslose, die innerhalb der letzten zwölf Monate mindestens sechs Monate lang arbeitslos waren, in eine ABM zugewiesen werden. Befunde liegen hier vor allem aus einem großen Kooperationsprojekt der Universität Frankfurt/Main und des IAB (s. Einleitung) zu den Eingliederungseffekten

von ABM vor: Caliendo et al. (2004a) ermitteln für eine Eintrittskohorte Geförderter, die im Februar des Jahres 2000 eine Maßnahme aufnahmen, dass ABM in Ostdeutschland noch drei Jahre nach ihrem Beginn im Durchschnitt einen negativen Effekt auf die Beschäftigungschancen der Geförderten hatten; in Westdeutschland waren die geschätzten Förderwirkungen nach gut zwei Jahren zumindest nicht mehr signifikant negativ. Allerdings zeigen Caliendo et al. (2004b, 2005) im Rahmen von Analysen der Teilnehmerheterogenität auch, dass ABM bei – den hier nicht mit untersuchten – Langzeitarbeitslosen wie auch bei Personen mit gesundheitlichen Einschränkungen tendenziell zu einer Verbesserung der Wiedereingliederungschancen beitragen. Hujer/Thomsen (2006) prüfen, ob die Effektivität der Förderung mit der Dauer der Arbeitslosigkeit vor Förderbeginn variiert. Dabei finden sie dreißig Monate nach Förderbeginn nur für sehr wenige Teilgruppen von Geförderten in Westdeutschland positive Förderwirkungen auf die Integration in ein ungeförderetes Beschäftigungsverhältnis.

Eingliederungszuschüsse (EGZ) sind Lohnkostenzuschüsse, die Arbeitgeber zeitlich befristet bei der Einstellung von Personen mit Vermittlungshemmnissen erhalten können. Sie sollen tatsächliche oder vermutete Minderleistungen des Arbeitnehmers finanziell ausgleichen. Mit dem dritten Hartz-Gesetz wurden im Januar 2004 die drei „alten“ Varianten des EGZ [a) bei Einarbeitung, b) bei erschwelter Vermittlung, c) für Ältere] zusammengefasst und mit neuen Förderkonditionen versehen. Insgesamt weisen Ergebnisse von Jaenichen (2002, 2005) sowie auch der Hartz-Evaluation (ZEW et al. 2005; Bernhard et al. 2006; Heinemann et al. 2006) darauf hin, dass Eingliederungszuschüsse bewirkt haben, dass die Geförderten häufiger sozialversicherungspflichtig beschäftigt und seltener arbeitslos waren als vergleichbare nicht geförderte Personen. Dabei zeigt sich auch, dass die Fördereffekte höher ausfallen, wenn als Erfolgskriterium der Verbleib in ungeförderter Beschäftigung – anstelle der Vermeidung von Arbeitslosigkeit – herangezogen wird. Ein höherer Anteil von Vergleichspersonen als von Geförderten wechselt im Zeitablauf in die Nichterwerbstätigkeit. Anzumerken ist aber, dass beim EGZ erhebliche Mitnahme- und Substitutionseffekte auftreten könnten.

Die *Förderung beruflicher Weiterbildung* (FbW), die 1998 die Maßnahmekategorie „Fortbildungs- und Umschulungsmaßnahmen“ (FuU) ablöste, kann in unterschiedlichen Varianten erfolgen. Grob unterscheiden lässt sich zwischen a) Maßnahmen zur Qualifikationserweiterung für Arbeitnehmer, die be-

reits einen Berufsabschluss besitzen, b) beruflichen Weiterbildungen mit einem Abschluss in einem anerkannten Ausbildungsberuf sowie c) Maßnahmen in Übungseinrichtungen. Ergebnisse auf Basis der neuen Prozessdaten liegen hier vor allem aus einem Kooperationsprojekt zur Evaluation von FuU vor, an dem die Universität St. Gallen, die Universität Frankfurt/Main sowie das IAB beteiligt sind. Lechner et al. (2005a, 2005b) untersuchen für Fördereintritte in den Jahren 1993/1994 die Wirksamkeit verschiedener Varianten im Vergleich untereinander und gegenüber einer Nichtteilnahme, und zwar über einen Zeitraum von mehr als sieben Jahren. Ihr zentrales Ergebnis ist, dass die Teilnahme an FuU längerfristig überwiegend positive Effekte auf die Wiedereingliederungschancen von Teilnehmern in sozialversicherungspflichtige Beschäftigung haben. Zu ähnlichen Resultaten für FuU kommen auch Fitzenberger et al. (2006). Biewen et al. (2006) zeigen für Fördereintritte in FbW in den Jahren 2000 bis 2002, dass die Teilnehmer an Maßnahmen zur Qualifikationserweiterung oder in Übungsfirmen zwei Jahre nach Förderbeginn in Westdeutschland – aber nicht in Ostdeutschland – tendenziell häufiger beschäftigt waren als vergleichbare nicht geförderte Personen; ihre Ergebnisse weisen zudem darauf hin, dass die Förderung für solche Personen effektiver war, die vor dem Förderbeginn bereits länger arbeitslos waren. Analysen im Rahmen der Hartz-Evaluation (IZA et al. 2005) lassen ebenfalls vermuten, dass sich FbW tendenziell positiv auf die Integration in ungeförderter sozialversicherungspflichtige Beschäftigung auswirken (Bundesministerium für Arbeit und Soziales 2006: Anhang 1, S. I).

Diese positiven Einschätzungen modifizieren sich jedoch, wenn die untersuchte Ergebnisvariable nicht der Verbleib in sozialversicherungspflichtiger Beschäftigung, sondern die Vermeidung von Arbeitslosigkeit ist: Arbeitslosigkeit wird nicht nur vermieden, wenn eine Person sozialversicherungspflichtig beschäftigt ist, sondern z. B. auch, wenn sie in die Stille Reserve abwandert oder sich selbständig macht. Lechner et al. (2005a: 36, 2005b: 36) zeigen, dass keines der von ihnen untersuchten Weiterbildungsprogramme bei den Geförderten zu einer Vermeidung von Arbeitslosigkeit beigetragen hat (wiederrum für Fördereintritte der Jahre 1994/1995 und für einen Beobachtungszeitraum von mehr als sieben Jahren). Sie führen dies darauf zurück, dass FuU Personen in den Arbeitsmarkt integrieren, die ansonsten in die Nichterwerbstätigkeit gewechselt wären. Auch Hujer et al. (2006a) kommen mit Hilfe von Ereignisdatenanalysen zu dem Ergebnis, dass die Teilnahme an Weiterbildungsmaßnahmen für Personen, die im letzten Quartal 1999 in die Förderung eintraten, zwei Jahre nach Förderbeginn keine

positiven Effekte auf den Austritt auf Arbeitslosigkeit gehabt hat.

Ergebnisse zu *Trainingsmaßnahmen* (TM) sind vorsichtig zu interpretieren. Durch ihren Einsatz werden sehr unterschiedliche Zielsetzungen verfolgt (Kurtz 2003): Sie können der Eignungsfeststellung dienen, die Suche nach einem Arbeitsplatz durch Bewerbungstraining unterstützen oder Kenntnisse und Fähigkeiten vermitteln, aber auch die Verfügbarkeit des Arbeitslosen und seine Arbeitsbereitschaft überprüfen. Durchgeführt werden sie in Gruppenmaßnahmen bei Trägern oder als Praktika in Betrieben. Ihre Dauer beträgt nur wenige Wochen; gelegentlich dienen sie als Vorbereitung einer anderen Maßnahme aktiver Arbeitsmarktpolitik. Damit ist bei TM – wie auch bei den zuvor diskutierten FbW – von einer hohen Maßnahmenheterogenität auszugehen, die durch das hier gewählte Vorgehen nicht abgebildet wird. Zu TM liegen bisher nur wenige Wirkungsanalysen vor; sie wurden auch im Rahmen der Hartz-Evaluation nicht untersucht. Ausnahmen bilden die Studien von Biewen et al. (2006) und Hujer et al. (2006b). Beide Untersuchungen kommen für Westdeutschland – einmal auf der Basis eines Vergleichsgruppenansatzes, einmal unter Anwendung ereignisanalytischer Verfahren – zu einer tendenziell positiven Einschätzung. Die untersuchte Ergebnisvariable ist dabei bei Biewen et al. (2006) der Verbleib in sozialversicherungspflichtiger Beschäftigung, bei Hujer et al. (2006) hingegen die Dauer bis zum Übergang in ein solches Beschäftigungsverhältnis. Allerdings kann in beiden Untersuchungen aufgrund der vergleichsweise geringen Fallzahlen nicht zwischen unterschiedlichen Varianten von Trainingsmaßnahmen differenziert werden.

Insgesamt weisen die Ergebnisse aus der Literatur darauf hin, dass die zielgruppenspezifische Förderung von Problemgruppen – zu denen gerade auch Langzeitarbeitslose zählen – bei bestimmten Maßnahmen vergleichsweise effektiv sein könnte; dies gilt etwa für Lohnkostenzuschüsse und Beschäftigungssubventionen (s. für einen Überblick Konle-Seidl 2005). Jedoch ist eine strenge Beschränkung auf die Förderung von Problemgruppen generell als problematisch anzusehen, da sie das Risiko einer Stigmatisierung der Teilnehmer verstärkt.

3.2 Untersuchungsansatz

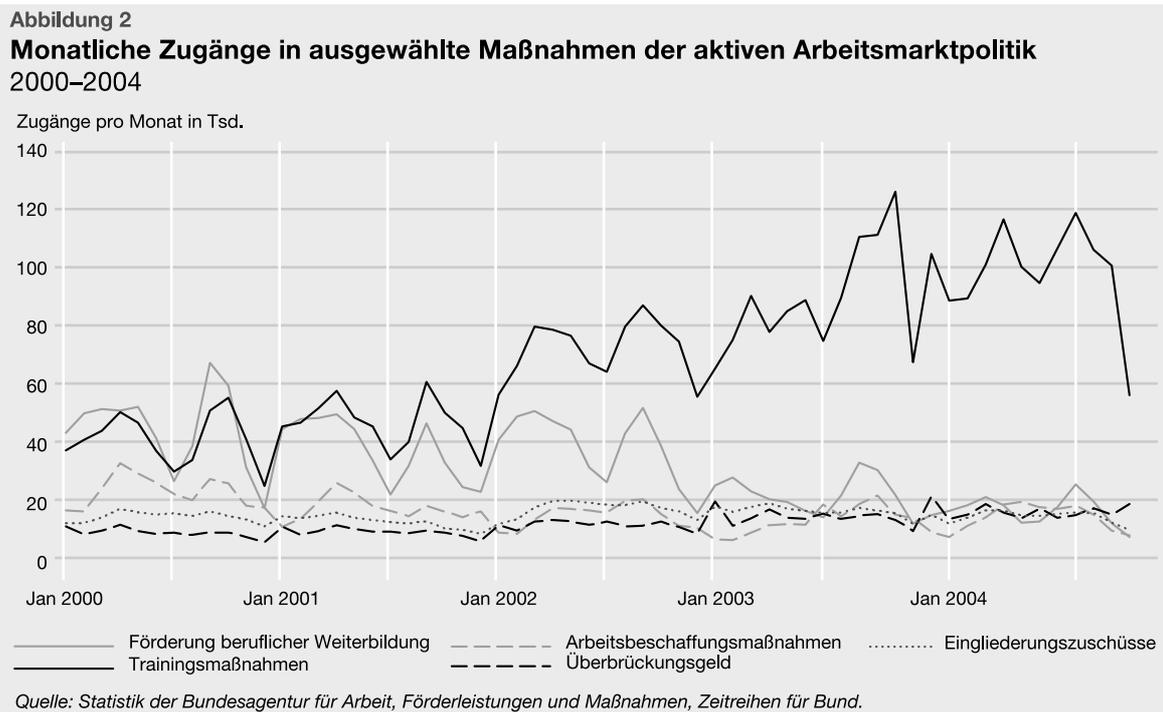
Die im Folgenden durchgeführte Analyse soll zum einen das Potenzial der TrEffeR-Datenbank verdeutlichen. Zum anderen wird auf dieser Basis erstmals eine Auswertung aller Instrumente auf Basis

einer einheitlich aufbereiteten Datenbank und mit derselben Methodik vorgenommen. Der einheitliche Ansatz hat allerdings den Preis, dass keine Feinabstimmung von Methodik und Variablenauswahl auf die einzelnen Instrumente erfolgt. In diesem Abschnitt wird das Vorgehen kurz skizziert; Details des Untersuchungsaufbaus finden sich in Anhang A.

Untersucht wird auf Basis eines Vergleichsgruppenansatzes, ob ein Fördereintritt im September 2002 für die geförderten Personen in der Folge mit einer Vermeidung von weiteren Zeiten der Arbeitslosigkeit einhergegangen ist. Dabei begrenzt sich die Analyse auf Personen im Alter von 25 bis unter 60 Jahren, die bis zum Maßnahmeeintritt nicht länger als ein Jahr arbeitslos waren, und bei diesen auf den ersten Maßnahmeeintritt in einer Arbeitslosigkeitsepisode. Der Verbleib der untersuchten Personen nach Förderbeginn lässt sich für gut zwei Jahre beobachten.

Die Beschränkung der Analyse auf Personen, die (noch) nicht langzeitarbeitslos sind, resultiert hier einerseits aus dem derzeit noch recht kurzen verfügbaren Beobachtungszeitraum. Andererseits besteht derzeit aus Sicht der BA aber auch ein verstärktes Erkenntnisinteresse an den Auswirkungen ihres Handelns auf Personen, die noch nicht langzeitarbeitslos sind. Seit Februar 2006 beträgt die maximale Anspruchsdauer im Rechtskreis SGB III auf den Bezug von Arbeitslosengeld I für Arbeitslose bis 54 Jahre zwölf Monate und für Ältere 18 Monate. Für den Fall des Übertritts von Arbeitslosen aus dem beitragsfinanzierten Rechtskreis SGB III in den steuerfinanzierten Rechtskreis SGB II hat die Bundesagentur für Arbeit pro Fall einen Aussteuerungsbetrag zu zahlen. Die Intention des Gesetzgebers war es, einen finanziellen Anreiz für die Bundesagentur für Arbeit zu schaffen, das Risiko der Langzeitarbeitslosigkeit frühzeitig und effektiv zu bekämpfen.

Die untersuchten Instrumente sind die Förderung beruflicher Weiterbildung (FbW), Trainingsmaßnahmen (TM), Eingliederungszuschüsse (EGZ), Arbeitsbeschaffungsmaßnahmen (ABM) sowie das Überbrückungsgeld (ÜG). Abbildung 2 stellt die monatlichen Zugänge in diese Maßnahmen im Überblick dar; der September ist regelmäßig einer der Monate mit den meisten Förderzugängen. Dabei ist – wie in Abschnitt 3.1. kurz dargestellt wurde – zu beachten, dass sich hinter FbW, TM und EGZ sehr heterogene Einzelmaßnahmen, auch mit unterschiedlicher Dauer, verbergen können. Dies wird in der Auswertung berücksichtigt, indem weiter nach Unterkategorien dieser Maßnahmen differenziert wird. Bei einer starken Streuung der Förderdauern wird bei der Auswertung weiterhin nach



Maßnahmedauern differenziert. Diese Unterscheidung ist dabei in gewissem Sinne endogen, da die Förderdauer unter anderem durch die Vermittlungshemmnisse der Geförderten bestimmt wird. Anzumerken ist, dass bei der Analyse des EGZ die Nachbeschäftigungsfrist in der vorliegenden Auswertung nicht mehr als Teil des Förderzeitraums interpretiert wurde.

Für die Bildung der Vergleichsgruppen wird jeder geförderten Person mit Hilfe von Matching-Methoden eine Person mit möglichst ähnlicher Förderwahrscheinlichkeit zugeordnet, die nie oder später gefördert wurde. Eine wichtige Voraussetzung bei diesem Vorgehen ist, dass alle Merkmale, die die Teilnahmewahrscheinlichkeit sowie den Teilnahmeerfolg beeinflussen, bekannt sind und berücksichtigt werden („Conditional Independence Assumption“ CIA). Zudem darf die Förderung bestimmter Personen keinen Einfluss auf die Beschäftigungschancen anderer Personen haben; es dürfen also keine indirekten Effekte auf Dritte vorliegen („Stable Unit Treatment Assumption“ SUTVA).

3.3 Ergebnisse

Tabelle 1 dokumentiert im Überblick die stichtagsbezogenen und kumulierten Verbleibsergebnisse zwei Jahre nach Förderbeginn, die geschätzte Förderwirkung auf die Geförderten sowie den Mittel-

wert des standardisierten Bias (MSB) vor und nach dem Matching. Die dort dargestellten Ergebnisvariablen sind erstens der Verbleib am Stichtag (operationalisiert in Form einer Dummy-Variable, die den Wert Eins annimmt, wenn eine Person am Stichtag arbeitslos oder in einer Maßnahme war) und zweitens die kumulierte Dauer der tatsächlichen Arbeitslosigkeit, gemessen jeweils zwei Jahre nach Förderbeginn. Die ausgewiesene geschätzte Förderwirkung (ATT) gibt an, inwieweit die geförderten Personen infolge der Förderung Zeiten der Arbeitslosigkeit vermeiden konnten (s. Anhang A für die genaue Definition).

Die Gütetests weisen darauf hin, dass insgesamt – vor allem unter Berücksichtigung der vergleichsweise geringen Fallzahlen – durch das Matching ein guter Ausgleich zwischen den Merkmalen der Geförderten und der Vergleichspersonen gelingt. Einen detaillierten Einblick in den Verlauf der Förderwirkungen vermitteln die grafischen Abbildungen der durchschnittlichen Förderwirkung der Instrumente auf die Geförderten in Anhang B. Generell gilt für so gut wie alle untersuchten Instrumente, dass die Förderung in Ostdeutschland für die Geförderten vergleichsweise effektiver als in Westdeutschland gewesen ist.

Insbesondere Personen, die im September 2002 in die Förderung mit dem Überbrückungsgeld oder mit Eingliederungszuschüssen bei Einarbeitung

Tabelle 1

Verbleibe zwei Jahre nach Förderbeginn, geschätzte Förderwirkung und Gütetests

Anteil am Stichtag bzw. kumulierte Tage in Arbeitslosigkeit oder Maßnahmen für Geförderte (G) und Vergleichspersonen (V), geschätzte Förderwirkung auf die Vermeidung von Arbeitslosigkeit bei den Geförderten (ATT), mittlerer standardisierter Bias (MSB) vor und nach dem Matching sowie Anzahl der Geförderten

Maßnahme und Dauer	Gruppe	Am Stichtag			Kumulierte Tage			MSB		Anzahl
		G	V	ATT	G	V	ATT	vor	nach	
Überbrückungsgeld	Männer West	0,16	0,35	0,19 **	261,94	386,96	125,02 **	15,82	1,59	2.106
	Frauen West	0,12	0,29	0,18 **	243,84	339,30	95,45 **	16,52	3,17	805
	Männer Ost	0,15	0,41	0,26 **	256,48	434,92	178,44 **	17,87	3,00	502
	Frauen Ost	0,18	0,45	0,27 **	252,31	428,84	176,53 **	19,25	3,86	211
Arbeitsbeschaffungsmaßnahmen 7 bis 12 Monate	Männer West	0,61	0,55	-0,06	606,46	497,94	-108,52 **	16,61	5,29	227
	Frauen West	0,62	0,43	-0,19 **	592,67	434,77	-157,90 **	19,92	5,37	207
	Männer Ost	0,64	0,52	-0,12 **	619,38	505,43	-113,95 **	17,26	3,80	413
	Frauen Ost	0,69	0,53	-0,15 **	628,04	499,15	-128,88 **	16,02	4,47	312
EGZ bei Einarbeitung bis 3 Monate	Männer West	0,22	0,42	0,20 **	229,92	397,38	167,46 **	13,46	4,55	242
	Frauen West	0,17	0,35	0,18 **	204,86	372,23	167,36 **	16,30	5,41	162
	Männer Ost	0,28	0,39	0,11 *	241,98	440,76	198,79 **	16,67	3,93	170
	Frauen Ost	0,23	0,42	0,19 *	218,01	407,91	189,90 **	21,07	4,55	78
EGZ bei Einarbeitung 4 bis 6 Monate	Männer West	0,23	0,42	0,20 **	291,43	412,78	121,35 **	14,53	3,63	220
	Frauen West	0,22	0,33	0,11 *	261,89	380,36	118,47 **	17,40	3,38	148
	Männer Ost	0,19	0,40	0,21 **	281,46	428,27	146,81 **	17,86	3,37	385
	Frauen Ost	0,13	0,47	0,33 **	240,43	465,41	224,98 **	17,72	4,76	233
EGZ bei erschwelter Vermittlung 7 bis 12 Monate	Männer West	0,35	0,52	0,17 **	465,31	470,92	5,60	13,72	7,85	154
	Frauen West	0,24	0,37	0,13	422,12	392,90	-29,22	18,43	8,31	78
	Männer Ost	0,30	0,50	0,20 **	445,20	485,86	40,65	16,78	5,71	132
	Frauen Ost	0,22	0,59	0,37 **	418,92	515,30	96,38 **	19,38	7,77	110
EGZ für Ältere ab 12 Monate	Männer West	0,89	0,43	-0,46 **	696,27	460,71	-235,57 **	21,62	6,00	146
	Frauen West	0,87	0,34	-0,53 **	697,32	423,29	-274,04 **	24,42	5,45	77
	Männer Ost	0,95	0,56	-0,39 **	714,06	486,30	-227,76 **	26,12	4,40	317
	Frauen Ost	0,98	0,56	-0,42 **	713,48	506,99	-206,49 **	22,74	5,45	161
FbW zur Qualifikations- erweiterung 4 bis 6 Monate	Männer West	0,44	0,40	-0,04	459,30	420,50	-38,81 **	9,72	2,96	797
	Frauen West	0,33	0,29	-0,04	404,04	368,04	-36,01 **	11,73	3,81	1.003
	Männer Ost	0,41	0,41	0,00	464,58	455,58	-9,00	10,96	4,86	258
	Frauen Ost	0,50	0,56	0,06	498,23	492,78	-5,45	10,95	4,21	161
FbW zur Qualifikations- erweiterung 7 bis 12 Monate	Männer West	0,43	0,38	-0,05 *	522,27	395,44	-126,82 **	12,74	2,94	1.039
	Frauen West	0,34	0,32	-0,01	480,55	367,10	-113,45 **	14,39	2,26	2.038
	Männer Ost	0,46	0,44	-0,03	549,69	475,77	-73,92 **	11,34	2,01	872
	Frauen Ost	0,60	0,54	-0,07	592,32	503,69	-88,62 **	12,52	3,55	672
FbW in Übungsfirma 4 bis 6 Monate	Männer West	0,47	0,41	-0,06	483,28	420,13	-63,15 **	13,59	6,04	211
	Frauen West	0,38	0,30	-0,08	452,82	355,02	-97,79 **	15,91	3,74	239
	Männer Ost	0,52	0,41	-0,11	513,62	452,51	-61,11	16,22	5,84	82
	Frauen Ost	0,62	0,55	-0,07	563,90	507,50	-56,39	15,85	4,11	105
Nicht-betriebliche FbW Berufsabschluss ab 12 Monate	Männer West	0,69	0,44	-0,25 **	695,89	410,75	-285,14 **	15,42	3,13	532
	Frauen West	0,78	0,31	-0,47 **	705,04	346,76	-358,27 **	15,38	2,81	842
	Männer Ost	0,85	0,49	-0,36 **	705,44	473,74	-231,70 **	18,38	5,48	126
	Frauen Ost	0,85	0,47	-0,38 **	706,72	488,56	-218,16 **	20,14	4,00	186
Betriebliche TM zur Vermittlung von Kenntnissen	Männer West	0,35	0,39	0,04	336,54	409,46	72,92 **	10,24	2,57	946
	Frauen West	0,34	0,35	0,01	344,83	374,78	29,95	11,14	3,91	525
	Männer Ost	0,34	0,43	0,09 *	366,07	462,10	96,02 **	10,16	3,17	373
	Frauen Ost	0,33	0,51	0,17 **	403,09	467,37	64,28 *	15,31	5,10	195
Nicht-betriebliche TM zur Vermittlung von Kenntnissen	Männer West	0,44	0,41	-0,04	456,13	412,80	-43,33 **	9,08	2,05	1.495
	Frauen West	0,33	0,32	-0,01	400,99	369,66	-31,33 **	11,33	1,77	2.631
	Männer Ost	0,48	0,48	0,00	460,40	466,33	5,93	7,80	3,39	648
	Frauen Ost	0,52	0,53	0,02	514,73	493,57	-21,16	10,78	2,91	786
TM zur Überprüfung der Verfügbarkeit	Männer West	0,43	0,39	-0,04 *	432,26	399,55	-32,70 **	13,28	2,02	3.228
	Frauen West	0,33	0,31	-0,02	388,93	358,02	-30,91 **	13,00	2,16	2.832
	Männer Ost	0,53	0,49	-0,04	473,27	450,73	-22,54	13,86	3,37	657
	Frauen Ost	0,55	0,53	-0,02	511,33	472,69	-38,64 *	13,14	3,43	669

* $\alpha = 0,05$, ** $\alpha = 0,01$

oder erschwerter Vermittlung (Förderdauern bis zwölf Monate) eingetreten sind, waren zwei Jahre nach Förderbeginn seltener arbeitslos oder in Maßnahmen als Vergleichsgruppen nicht oder später geförderter Personen. So gilt z. B. für Männer in Westdeutschland, die im September 2002 eine Förderung mit dem Überbrückungsgeld aufnahmen, dass zwei Jahre später der Arbeitslosenanteil unter den Geförderten um 19 Prozentpunkte geringer war als unter den Vergleichspersonen. Zugleich waren die Geförderten in diesem Zeitraum 125 Tage weniger in Arbeitslosigkeit oder Maßnahmen. Die positive Förderwirkung folgt auf einen anfänglichen „Lock-in“-Effekt der Förderung und nimmt im Zeitablauf ab (siehe Anhang B). Bei Förderdauern bis zu sechs Monaten, wie sie beim Überbrückungsgeld und bei den hier untersuchten EGZ bei Einarbeitung der Fall sind, ist auch der kumulative Effekt der Förderung zwei Jahre nach Förderbeginn signifikant positiv. Bei den längeren Förderdauern des EGZ bei erschwerter Vermittlung ist der Verbleibszeitraum hingegen zu kurz, um durchgehend signifikant positive kumulierte Förderwirkungen zu beobachten.

Auch die Teilnehmer an Trainingsmaßnahmen zur Vermittlung von Kenntnissen, die in Betrieben durchgeführt wurden, sind in der Folge seltener arbeitslos als die Mitglieder einer nicht oder später geförderten Vergleichsgruppe; signifikant sind die Ergebnisse allerdings nur für Ostdeutschland. Für Teilnehmer an nicht-betrieblich durchgeführten TM zur Vermittlung von Kenntnissen sowie TM zur Überprüfung der Verfügbarkeit sind hingegen in der Folgezeit insignifikante oder signifikante (aber geringe) negative Fördereffekte festzustellen.

Teilnehmer an beruflichen Weiterbildungsmaßnahmen sind zwei Jahre nach dem Antritt der Maßnahme im September 2002 häufiger arbeitslos oder in Maßnahmen als vergleichbare Personen, die nicht oder später gefördert wurden. Allerdings sind für Teilnehmer von Maßnahmen mit einer Förderdauer bis zu zwölf Monaten die Effekte auf den Verbleib am Stichtag nach dem anfänglichen „Lock-in“-Effekt der Förderung kaum noch signifikant; bei Maßnahmen bis zu sechs Monaten Dauer gilt dies auch für die kumulierte Förderwirkung.

Von „Lock-in“-Effekten dominiert wird die Schätzung der Förderwirkung für Personen, die an langen Maßnahmen von über zwölf Monaten (die häufig mindestens zwei Jahre andauern) teilgenommen haben. Bei den vorliegenden Auswertungen betrifft dies einerseits Personen, die mit dem EGZ für Ältere gefördert wurden, andererseits Personen, die an einer Umschulung teilgenommen haben. Einen Extremfall stellen Frauen in Westdeutschland dar, die

im September 2002 mit einer Umschulung begonnen haben: Im Zeitraum von zwei Jahren waren sie in der kumulierten Betrachtung etwa ein Jahr länger „Kunden“ der Arbeitsagentur als ihre Vergleichspersonen. Auch für ABM mit einer Förderdauer von sieben bis zwölf Monaten weisen die Ergebnisse darauf hin, dass die Förderung bei den Geförderten nicht zur Vermeidung von Arbeitslosigkeit beitragen konnte.

Die Ergebnisse dieser Gesamtschau sind – bei allen Vorbehalten gegenüber unterschiedlichen Stichproben und Methoden – in der Tendenz mit den in Abschnitt 3.1 zitierten Befunden aus der Literatur vereinbar. Sie zeigen aber auch sehr deutlich, welchen Restriktionen sich die bisher im Rahmen des TrEffeR-Projektes durchgeführten Wirkungsanalysen gegenübersehen:

Erstens hängt die Einschätzung der Effektivität einer Maßnahme ganz entscheidend vom betrachteten Vergleichshorizont ab. Vergleichsweise kurze Förderdauern können sich in dem betrachteten Verbleibszeitraum von zwei Jahren durchaus „amortisieren“. Bei längeren Förderdauern ist dies hingegen nicht möglich, zum Teil sind die Maßnahmen im hier verfügbaren Beobachtungszeitraum noch gar nicht abgeschlossen. Dass es durchaus Jahre dauern kann, bis Maßnahmen ihre Wirkung entfalten, haben aber kürzlich die in Abschnitt 3.1 erwähnten neueren Untersuchungen zu FbW gezeigt. Schneider/Uhlendorff (2006) diskutieren – ebenfalls für das Instrument der FbW – ausführlich, dass die verbreitete Evaluationstechnik auf Basis einer Vergleichsgruppenbildung dazu führt, dass es bei „Lock-in“-Effekten und schwachen Maßnahmeeffekten mehrere Jahre dauert, bis ein Programmeffekt von Maßnahmen sichtbar wird. Generell gilt: Maßnahmen, die mit einer direkten Integration in den ersten Arbeitsmarkt einhergehen (dies sind die Eingliederungszuschüsse und das Überbrückungsgeld), bewirken nach Maßnahmeabschluss zunächst einen „Vorsprung“ der Geförderten gegenüber der Vergleichsgruppe. Hingegen weist bei Weiterbildungsmaßnahmen – aufgrund der unvermeidlichen „Lock-in“-Effekte – die Vergleichsgruppe unmittelbar nach Maßnahmeabschluss einen „Vorteil“ auf.

Zweitens kann gerade bei den Eingliederungszuschüssen und dem Überbrückungsgeld – also den Maßnahmen, bei denen bereits zwei Jahre nach Förderbeginn ein signifikant positiver Effekt der Förderung auf die Geförderten ermittelt wird – nicht ausgeschlossen werden, dass Mitnahme-, Substitutions- und Verdrängungseffekte auftreten. Aus den geschätzten Förderwirkungen auf individueller Ebene kann damit nicht auf die Vermeidung von Arbeitslosigkeit im Aggregat geschlossen werden.

Drittens ist eine zentrale Annahme, die dem Vergleichsgruppenansatz zugrunde liegt, die CIA (s. Abschnitt 3.2). Sie setzt voraus, dass alle relevanten Merkmale beobachtet werden, die sowohl den Zugang in die Förderung – als Ergebnis der Entscheidungen der Vermittler sowie der Geförderten selbst – als auch den Fördererfolg beeinflussen. Der umfangreiche Variablenbestand in den Prozessdaten der BA gewährleistet, dass zentrale Einflussfaktoren – sozio-demografische Merkmale, Erwerbsbiografie, regionaler Arbeitsmarktkontext sowie Dauer der bisherigen Arbeitslosigkeit bis zum Förderbeginn – berücksichtigt werden (s. Anhang A). Dabei ist davon auszugehen, dass sich auch „weiche“ Faktoren wie Motivation und Leistungsbereitschaft in der Erwerbshistorie von Arbeitslosen widerspiegeln. Unseres Erachtens bilden die Prozessdaten Faktoren, die in der Person der Arbeitslosen liegen und die für die Selektion in Maßnahmen sowie deren Fördererfolg entscheidend sind, für die Mehrzahl der Maßnahmen damit weitgehend ab. Bei den Eingliederungszuschüssen liegt zwar die Überlegung nahe, dass neben den individuellen Eigenschaften von Arbeitslosen auch die Verhandlungsstrategien von Betrieben und Arbeitsvermittlern eine wichtige Rolle für die Aufnahme eines geförderten Beschäftigungsverhältnisses spielen. Aus methodischer Sicht wäre dies jedoch nur dann problematisch, wenn diese Verhandlungsstrategien systematisch mit Eigenschaften vermittelter Arbeitsloser korreliert sind, die durch die – oben erläuterten – umfangreichen Prozessdaten der BA nicht erfasst werden. Kritisch zu hinterfragen ist die CIA aus unserer Sicht insbesondere bei der Gründungsförderung: Beim Überbrückungsgeld wäre eine Überschätzung der Förderwirkung für die Geförderten zu erwarten, wenn systematisch gerade solche Arbeitslose mit dem Überbrückungsgeld gefördert werden, die eine marktfähige Gründungsidee hatten, während die Vergleichspersonen eine solche Idee möglicherweise nicht gehabt hätten.

Viertens lassen bisherige Untersuchungen vermuten, dass arbeitsmarktpolitische Maßnahmen auch dazu beitragen, Personen zu aktivieren, die sonst in die Nichterwerbstätigkeit abgewandert wären – die geschätzten Förderwirkungen dürften höher ausfallen, wenn nicht wie hier die Vermeidung von Arbeitslosigkeit, sondern die Eingliederung in sozialversicherungspflichtige Beschäftigung betrachtet wird.

4 Fazit

Die wissenschaftliche Evaluation arbeitsmarktpolitischer Maßnahmen hat in den letzten Jahren in

Deutschland enorm an Bedeutung gewonnen. Dass diese Evaluation möglich wurde, ist nicht zuletzt auf die verbesserten Datenzugänge zu den Prozessdaten der BA zurückzuführen. Auch innerhalb der BA wächst das Bewusstsein dafür, dass eine quantitative Wirkungsanalyse unverzichtbarer Bestandteil der operativen Steuerung sein sollte.

Um diesem Bedarf nachzukommen, wird in der BA gegenwärtig die TrEffeR-Datenbank aufgebaut, die – als ein erstes Ergebnis des TrEffeR-Projekts – retrospektiv kleinräumig Informationen zur Effektivität des Instrumenteneinsatzes für die Geförderten bereitstellt. Hierdurch wird nicht zuletzt der Forderung von Wissenschaft und Politik Rechnung getragen, die Wirksamkeit arbeitsmarktpolitischer Maßnahmen laufend und vergleichsweise zeitnah systematisch zu überprüfen (Franz 2005).

Das Potenzial der Datenbasis verdeutlicht eine empirische Auswertung für Personen, die im September 2002 nach höchstens einem Jahr Arbeitslosigkeit in eine Förderung eintraten. Insbesondere Personen, die mit Eingliederungszuschüssen oder dem Überbrückungsgeld gefördert wurden, sind demnach zwei Jahre nach Beginn der Förderung seltener arbeitslos oder in Maßnahmen als eine Vergleichsgruppe nicht oder erst später geförderter Personen.

Diskutiert wurden aber auch die Grenzen des bisherigen Datensatzes und des verwendeten Matching-Ansatzes. Als ein letzter, aber nicht unwichtiger Punkt sei erwähnt, dass auch die Vielzahl gesetzlicher Änderungen im Bereich der Arbeitsförderung sowie laufende Anpassungen im Zuge der Reform der BA die Aufdeckung kausaler Zusammenhänge erschweren. Entsprechend können retrospektive Wirkungsuntersuchungen zwar als Anhaltspunkt für die gegenwärtige Wirkung arbeitsmarktpolitischer Maßnahmen herangezogen werden – der laufende Veränderungsprozess der „Förderlandschaften“ macht eine systematische Weiterführung und Weiterentwicklung der bisherigen Arbeiten jedoch unabdindbar.

Literatur

- Abadie, A./Imbens, W. (2006): Large Sample Properties of Matching Estimators for Average Treatment Effects, *Econometrica* 74, 235–267.
- Bernhard, S./Jaenichen, U./Stephan, G. (2006): Einarbeitungszuschüsse bei Einarbeitung und erschwelter Vermittlung: Matching-Analysen auf der Basis von Prozessdaten, *Vierteljahreshefte zur Wirtschaftsforschung* 75, 67–84.

- Biewen, M./Fitzenberger, B./Osikominu, A./Waller, M. (2006): Employment Effects of Short and Medium Term Training Programs in Germany in the Early 2000s, mimeo.
- Blien, U./Hirschenauer, F./Arendt, M./Braun, H.-J./Gunst, D.-M./Kilcioglu, S./Kleinschmidt, H./Musati, M./Roß, H./Vollkommer, D./Wein, J. (2004): Typisierung von Bezirken der Agenturen für Arbeit, Zeitschrift für ArbeitsmarktForschung 37, 146–175.
- Brinkmann, C. (2000): Zur Evaluation der aktiven Arbeitsförderung nach dem SGB III – Monitoring und Wirkungsforschung im Umbruch. In: Mitteilungen aus der Arbeitsmarkt- und Berufsforschung 33, S. 483–499.
- Bundesagentur für Arbeit (2005): Daten zu den Eingliederungsbilanzen, Sondernummer der Amtlichen Nachrichten der Bundesagentur für Arbeit, Februar 2005.
- Bundesministerium für Arbeit und Soziales (2006): Die Wirksamkeit moderner Dienstleistungen am Arbeitsmarkt – Bericht 2005 der Bundesregierung zur Wirkung der Umsetzung der Vorschläge der Kommission Moderne Dienstleistungen am Arbeitsmarkt, im Internet unter: <http://www.bmas.bund.de/BMAS/Navigation/root, did=109764.html>
- Caliendo, M./Hujer, R./Thomsen, S. (2004a): Evaluation der Eingliederungseffekte von Arbeitsbeschaffungsmaßnahmen in reguläre Beschäftigung für Teilnehmer in Deutschland, Zeitschrift für ArbeitsmarktForschung 37, 211–238.
- Caliendo, M./Hujer, R./Thomsen, S. (2004b): New Evidence on the Effect of Job Creation Schemes in Germany – A Matching Approach with Threefold Heterogeneity, Research in Economics 58, 257–302.
- Caliendo, M./Hujer, R./Thomsen, S. (2005): Identifying Effect Heterogeneity to Improve the Efficiency of Job Creation Schemes in Germany, IZA Discussion Paper 1512.
- Caliendo, M./Jahn, E. J. (2004): Verbleibsquote ein Controlling-Indikator für den Eingliederungserfolg von ABM!?, Zeitschrift für Evaluation 2, 51–69.
- Calmfors, L. (1994): Active Labour Market Policy and Unemployment – A Framework for the Analysis of Crucial Design Features, OECD Labour Market and Social Policy Occasional Paper 15, Paris.
- Colpitts, T. (2002): Targeting Reemployment Services in Canada: The Service and Outcome Measurement System (SOMS) Experience. In: Eberts, R. W., O’Leary, C. J., Wandner, S. A. (Hrsg.): Targeting Employment Services, Kalamazoo, Michigan, 283–301.
- Eberts, R. W./O’Leary, C. J. (2002): A Frontline Decision Support System for Georgia Career Centers, W. E., Upjohn Institute for Employment.
- Fitzenberger, B./Osikominu, A./Völter, R. (2006): Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany, ZEW Discussion Paper 06–39.
- Fitzenberger, B./Wilke, R. (2004): Unemployment Durations in West-Germany Before and After the Reform of the Unemployment Compensation System During the 1980s, ZEW Discussion Paper 04–24.
- Franz, W. (2005): Arbeitsmarktforschung und Arbeitsmarktstatistik aus der Sicht der wirtschaftlichen Beratung: Erfahrungen und Perspektiven, Allgemeines Statistisches Archiv 89, 141–158.
- Frölich, M./Lechner, M./Steiger, H. (2004): Statistically Assisted Programme Selection – International Experiences and Potential Benefits for Switzerland, IFAU Working Paper 2004/1.
- Greene, W. H. (2003, 5th edition): Econometric Analysis, New York etc.
- Heckman, J. J./LaLonde, R. J./Smith, J. A. (1999): The Economics and Econometrics of Active Labor Market Programs. In: Ashenfelter, O., Card, D. (Hrsg.): Handbook of Labor Economics, Vol. IIIA, 1865–2097.
- Heinemann, S./Jaenichen, U./Stephan, G. (2006): Eingliederungszuschüsse: Förderumfang, Strukturen und Effektivität, Bundesarbeitsblatt 3–2006, 4–10.
- Hujer, R./Thomsen, S. (2006): How Do Employment Effects of Job Creation Schemes Differ with Respect to the Foregoing Unemployment Duration?, ZEW Discussion Paper 06–47.
- Hujer, R./Thomsen, S./Zeiss, C. (2006a): The Effects of Vocational Training Programmes on the Duration of Unemployment in Eastern Germany, Allgemeines Statistisches Archiv 90, 299–322.
- Hujer, R./Thomsen, S./Zeiss, C. (2006b): The Effects of Short-Term Training Measures on the Individual Unemployment Duration in West Germany, ZEW Discussion Paper 06–65.
- Hummel, E./Jacobebbinghaus, P./Kohlmann, A./Oertel, M./Wübbecke, C./Ziegerer, M. (2005): Stichprobe der Integrierten Erwerbsbiografien IEBS 1.0, FDZ Datenreport 6/2005.
- IAB, DIW, sinus, GfA, Infas (2005): Zwischenbericht zum „Modul 1e: Existenzgründungen“ im Rahmen der Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission, Arbeitspaket 1 (Wirksamkeit der Instrumente), im Internet unter: <http://www.bmas.bund.de/BMAS/Redaktion/Pdf/Publikationen/forschungsbericht-1e-existenzgruendungen-bericht,property=pdf,bereich=bmas,sprache=de,rwb=true.pdf>
- IZA, DIW, infas (2005): Zwischenbericht zum „Modul 1b: Förderung beruflicher Weiterbildung und Transferleistungen“ im Rahmen der Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission, Arbeitspaket 1 (Wirksamkeit der Instrumente), im Internet unter:

- http://www.iza.org/de/webcontent/news/Hartz-Evaluierung-Modul1b_Haupttext.pdf
- Jaenichen, U.* (2002): Lohnkostenzuschüsse und individuelle Arbeitslosigkeit, *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 35, 327–351.
- Jaenichen, U.* (2005): Lohnkostenzuschüsse und individuelle Beschäftigungschancen. In: Bellmann, L. Hübler, O. Meyer, W. und Stephan, G. (Hrsg.): *Institutionen, Löhne und Beschäftigung, Beiträge zur Arbeitsmarkt- und Berufsforschung* 294, Nürnberg, 137–156.
- Jaenichen, U./Kruppe, T./Stephan, G./Ullrich, B./Wießner, F.* (2005): You can split it if you really want. Korrekturvorschläge für ausgewählte Inkonsistenzen in IEB und MTG, FDZ Datenreport 4/2005.
- Konle-Seidl, R.* (2005): Lessons Learned: Internationale Evaluierungsergebnisse zu Wirkungen aktiver und aktivierender Arbeitsmarktpolitik, IAB-Forschungsbericht 9/2005.
- Kurtz, B.* (2003): Trainingsmaßnahmen – was verbirgt sich dahinter?, IAB-Werkstattbericht 08/2003.
- Lechner, M.* (2001): Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption. In: Lechner, M., Pfeiffer, F. (Hrsg.): *Econometric Evaluation of Active Labour Market Policies*, Heidelberg, 43–58.
- Lechner, M./Miquel, R./Wunsch, C.* (2005a): Long-Run Effects of Public Sector Sponsored Training in West Germany, IAB Discussion Paper 03/2005.
- Lechner, M./Miquel, R./Wunsch, C.* (2005b): The Curse and Blessing of Training the Unemployed in a Changing Economy: The Case of East Germany after Unification, IAB Discussion Paper 14/2005.
- Leuven, E./Sianesi, B.* (2003): psmatch2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing, im Internet unter: <http://ideas.repec.org/c/boc/bocode/s432001.html>
- Lüdemann, E./Wilke, R. A./Zhang, X.* (2006): Censored Quantile Regressions and the Length of Unemployment Periods in West Germany, erscheint in: *Empirical Economics*.
- Pfeiffer, F./Reize, F.* (2000): Business Start-ups by the Unemployed – an Econometric Analysis Based on Firm Data, *Labour Economics* 7, 629–663.
- Rosenbaum, P. R./Rubin, D. B.* (1983): The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika* 70, 41–55.
- Rubin, D. B.* (1977): Assignment to Treatment Group on the Basis of a Covariate, *Journal of Educational Statistics* 2, 1–26, Printer's correction note 3, 384.
- Schneider, H./Uhlendorff, A.* (2006): Die Wirkung der Hartz-Reform im Bereich der beruflichen Weiterbildung, IZA Discussion paper 2255.
- Sianesi, B.* (2004): An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s, *The Review of Economics and Statistics* 86, 133–155.
- ZEW, IAB, IAT* (2005): Zwischenbericht zum „Modul 1d: Eingliederungszuschüsse und Entgeltsicherung“ im Rahmen der Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission, Arbeitspaket 1 (Wirksamkeit der Instrumente), im Internet unter: <http://www.bmas.bund.de/BMAS/Redaktion/Pdf/Publikationen/forschungsbericht-1d-eingliederungszuschuesse-und-entgeltsicherung.property=pdf,bereich=bmas,sprache=de,rwb=true.pdf>

Anhang A: Untersuchungsaufbau

Fragestellung der empirischen Untersuchung ist: Hat ein Fördereintritt im September 2002 bei den geförderten Personen zu einer Vermeidung von Arbeitslosigkeit und weiteren Maßnahmezeiten beigetragen? Es sei $D^{(9/2002)}$ ein Indikator für den Fördereintritt im September 2002, $Y^{1(9/2002)}$ der Verbleib nach Fördereintritt im September 2002 und $Y^{0(9/2002)}$ der Verbleib ohne Fördereintritt im September 2002. Der Verbleib wird für $t = 30, 60 \dots 720$ Tage nach Förderbeginn erstens am Stichtag – in Form einer Dummy-Variable, die den Wert Eins annimmt, wenn eine Person arbeitslos oder in einer Maßnahme ist – gemessen. Zweitens wird für jeden Zeitpunkt t die kumulierte Dauer der faktischen Arbeitslosigkeit berechnet. Die geschätzte Förderwirkung auf die Geförderten ist für jeden Zeitpunkt t nach Maßnahmebeginn formal (negativ) definiert als

$$\tau_t^{9/2002} = E(Y_t^{0(9/2002)} | D^{(9/2002)} = 1) - E(Y_t^{1(9/2002)} | D^{(9/2002)} = 1).$$

Eine positive geschätzte Förderwirkung besagt damit, dass die Geförderten t Tage nach Förderbeginn im Mittel zu einem geringeren Anteil bzw. weniger kumulierte Tage arbeitslos (oder in Maßnahmen) waren als die Mitglieder der Vergleichsgruppe.

Als Matching-Technik wird ein „Nearest-Neighbour-Matching“ mit Zurücklegen auf Basis der geschätzten Förderwahrscheinlichkeiten verwendet: Zur Schätzung von $E(Y_t^{0(9/2002)} | D^{(9/2002)} = 1)$ wird jeder geförderten Person diejenige nicht geförderte Person zugeordnet, deren Förderwahrscheinlichkeit den geringsten Abstand zur Förderwahrscheinlichkeit der geförderten Person aufweist; das Ausbalancieren der Verteilungen der Förderwahrscheinlichkeiten entspricht einem Ausbalancieren der Verteilungen der einzelnen Personenmerkmale zwischen den Gruppen geförderter und nicht geförderter Personen (Rosenbaum/Rubin 1983). Eine nicht geförderte Person kann dabei auch mehrmals als Vergleichsperson für unterschiedliche geförderte Personen herangezogen werden. Die Varianz der Fördereffekte wird auf Basis einer einfachen Varianzformel geschätzt, der Schätzfehler der Förderwahrscheinlichkeiten wird hierbei vernachlässigt (Lechner 2001). Methodisch wird damit auf einen einfachen Ansatz zurückgegriffen, der in Hinsicht auf Konsistenz und Effizienz durchaus kritisch diskutiert wird (Abadie/Imbens 2006). Genutzt wird das Stata-Modul *psmatch2* (Leuven/Sianesi 2003).

Potenzielle Vergleichspersonen „dürfen“ nach ihrem potenziellen Förderbeginn im September 2002

an arbeitsmarktpolitischen Maßnahmen (sowohl der jeweils untersuchten Maßnahme als auch an anderen Maßnahmen) teilgenommen haben. Hierdurch wird vermieden, bei der Schätzung der Förderwirkung bereits auf deren erwartetes Ergebnis zu konditionieren (Sianesi 2004). Denn wenn eine Person ex post nie in eine Maßnahme eingetreten ist, liegt dies möglicherweise daran, dass sie bereits einen ungeförderten Arbeitsplatz gefunden hatte, bevor ihr ein Programm angeboten wurde. Daher würde bei Vernachlässigung später geförderter Personen die Förderwirkung tendenziell unterschätzt. Ermittelt wird also der Effekt der Förderung mit einer bestimmten Maßnahme im September 2002 im Vergleich zu a) keiner Förderung oder aber b) einer späteren Förderung mit einer beliebigen Maßnahme. Sianesi (2004) charakterisiert den so geschätzten Fördereffekt als die Wirkung eines „Joining versus Waiting“.

Die Vergleichsgruppenbildung erfolgt auf Basis einer 10-Prozent-Stichprobe der TrEffeR-Datenbank. Für die Konstruktion der Verbleibsnachweise wird den Vergleichspersonen als hypothetischer Startzeitpunkt für jedes Instrument separat der durchschnittliche Förderbeginn der Geförderten im Laufe des September 2002 zugeordnet. Um die Güte der Vergleichsgruppenauswahl – die für die Validität der Ergebnisse kritisch ist – zu überprüfen, werden vor und nach dem Matching der standardisierte Bias der erklärenden Variablen zwischen ungeförderten und geförderten Personen berechnet (Rosenbaum/Rubin 1983). Ergänzend werden – hier nicht ausgewiesene – t -Tests auf Gleichheit der Mittelwerte der erklärenden Variablen von ungeförderten und geförderten Personen durchgeführt. Da lediglich Dummy-Variablen als erklärende Variablen herangezogen werden, erübrigt sich die Durchführung von Tests auf Gleichheit der Verteilungen der Variablen nach dem Matching.

Für die empirische Analyse werden folgende Einschränkungen getroffen: 1) In den Auswertungen werden nur Personen berücksichtigt, die zum Zeitpunkt des Eintritts in Arbeitslosigkeit zwischen 25 und unter 60 Jahre alt waren. 2) Betrachtet werden nur Fördereintritte aus der Arbeitslosigkeit, die spätestens ein Jahr nach dem Eintritt in die Arbeitslosigkeit begonnen haben. Hierdurch wird eine wichtige Zielgruppe der Arbeitsmarktpolitik aus der Betrachtung ausgeschlossen – die Langzeitarbeitslosen. Da in der TrEffeR-Datenbank Episoden erst seit Mitte des Jahres 2001 vorliegen, können für den gewählten Untersuchungszeitraum Maßnahmezugänge mit längerer vorheriger Arbeitslosigkeitsdauer allerdings nicht untersucht werden. Im Rahmen der Trennung der Rechtskreise SGB II und

SGB III ist diese Abgrenzung aus Sicht der BA aber auch durchaus von Interesse, da Personen nach einem Jahr Arbeitslosigkeit aus der Verantwortlichkeit der Arbeitsagenturen in die Verantwortlichkeiten der Arbeitsgemeinschaften, geteilten Trägerchaften oder optierenden Kommunen wechseln. 3) Analysiert wird ausschließlich die erste Maßnahme innerhalb einer Episode – Geförderte können zwar vor dem September 2002 an arbeitsmarktpolitischen Maßnahmen partizipiert haben, innerhalb der untersuchten Arbeitslosigkeitsepisode ist dies jedoch ihre erste Programmteilnahme. Auf eine Untersuchung von Maßnahmesequenzen sowie Maßnahmeparallelitäten, die insbesondere bei einer Kofinanzierung durch Sonderprogramme wie z.B. dem Europäischen Sozialfonds auftreten, wird verzichtet. 4) Um sicherzustellen, dass es sich tatsächlich um die erste Maßnahme innerhalb einer Episode handelt, werden nur solche Episoden einbezogen, bei denen seit dem Ende der letzten Meldung bei der BA mindestens 60 Tage vergangen sind. Alternative Auswertungen, bei denen 30 bzw. 90 Tage

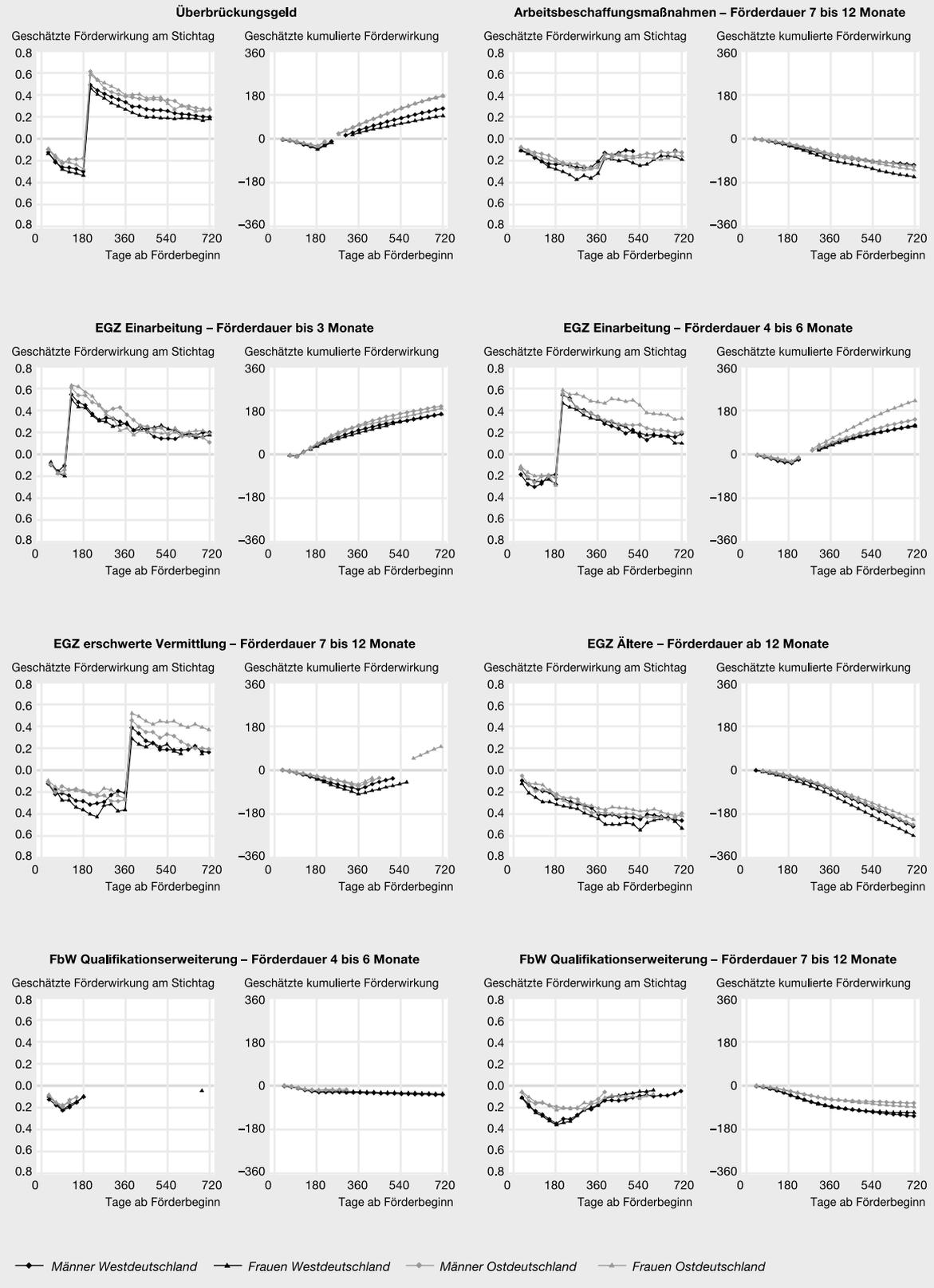
vergangen sein mussten, bewirken kaum Änderungen bei den Ergebnissen. 5) Ausgewiesen werden Ergebnisse nur für Kombinationen von Förderart und Förderdauer, bei denen für alle vier untersuchten Teilgruppen mindestens 50 Beobachtungen vorliegen.

Für die Probit-Schätzung der Förderwahrscheinlichkeiten wurden folgende Personenmerkmale zugrunde gelegt: a) allgemeine sozio-ökonomische Merkmale, gemessen zum Beginn der untersuchten Arbeitslosigkeitsepisode: Alter, Familienstand, Nationalität, Ausbildung, gesundheitliche Einschränkungen, b) 18-Monate-Historie zur Erwerbsbiografie vor Eintritt in die untersuchte Arbeitslosigkeitsepisode: Dauer der Arbeitslosigkeit in Halbjahren, Maßnahmeteilnahme, Abmeldungen wegen Krankheit oder Sperrzeiten, c) regionaler Arbeitsmarkt: Zuordnung zu einem von zwölf regionalen Vergleichstypen (Blien et al. 2004), d) Dauer der gegenwärtigen Arbeitslosigkeitsepisode bis zum Förderbeginn in Quartalen.

Anhang B

Verläufe der geschätzten Förderwirkungen (am Stichtag und kumuliert)

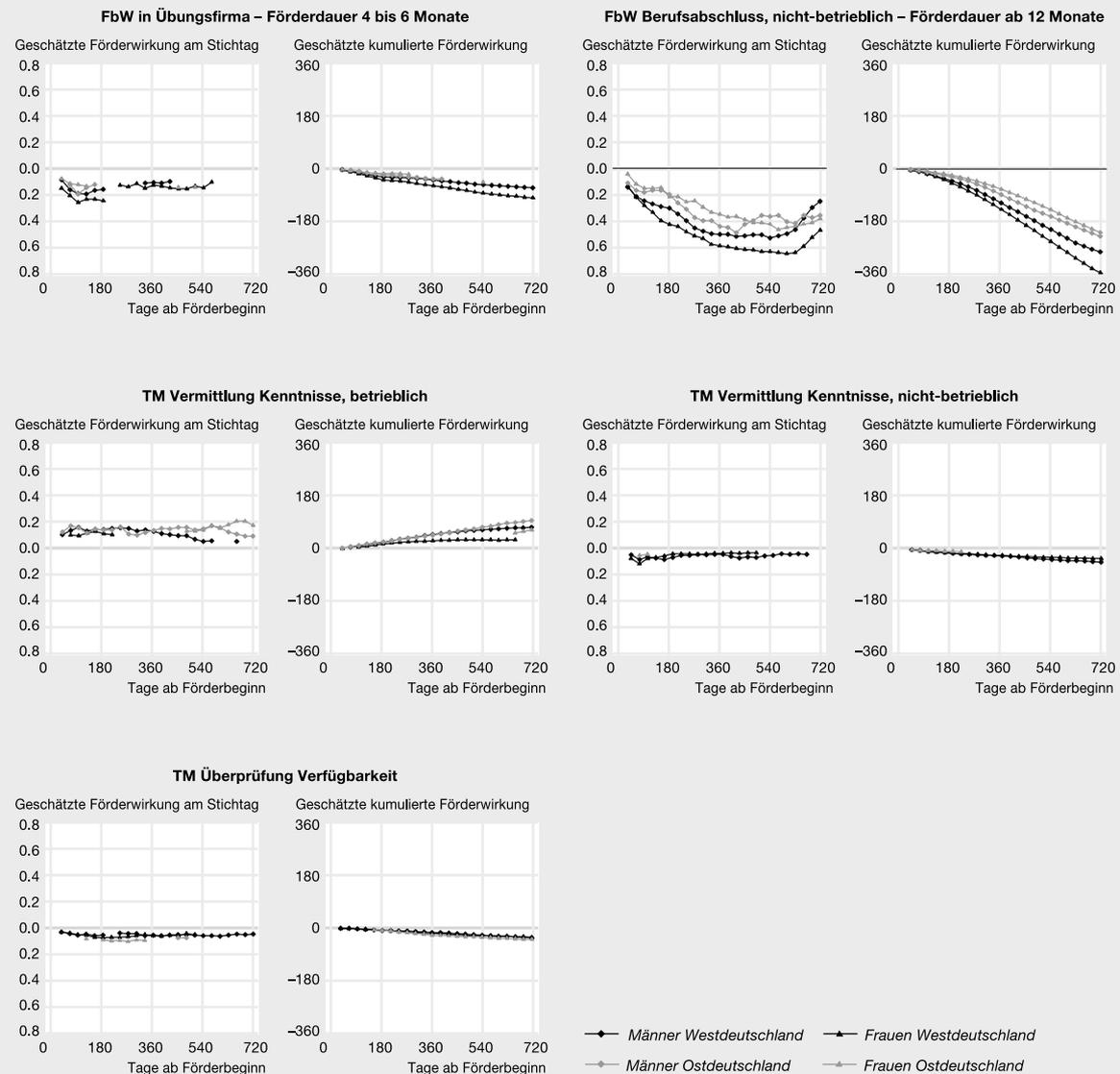
Geschätzte Förderwirkung auf die Vermeidung von Arbeitslosigkeit bei den Geförderten



Fortsetzung Anhang B

Verläufe der geschätzten Förderwirkungen (am Stichtag und kumuliert)

Geschätzte Förderwirkung auf die Vermeidung von Arbeitslosigkeit bei den Geförderten



Anmerkungen:

Ausgewiesen sind nur auf dem 5 Prozent-Niveau signifikante Förderwirkungen.

Die linke Hälfte der Abbildungen stellt die geschätzte Förderwirkung am Stichtag auf den Verbleib dar. Förderwirkungen im positiven Bereich implizieren, dass sich die Förderung positiv auf die jeweiligen Geförderten ausgewirkt hat. Beträgt zum Beispiel die geschätzte Förderwirkung 180 Tage nach Förderbeginn 0,3, so war ein halbes Jahr nach Förderbeginn die Wahrscheinlichkeit, arbeitslos oder in einer Maßnahme zu sein, in der Gruppe der Geförderten um 30 Prozentpunkte geringer als in der Vergleichsgruppe.

Die rechte Hälfte der Abbildungen weist die geschätzte kumulierte Förderwirkung aus; sie basiert auf der kumulierten „Dauer der faktischen Arbeitslosigkeit“ seit Maßnahmenbeginn. Beträgt hier zum Beispiel die geschätzte Förderwirkung 180 Tage nach Förderbeginn 90, so waren die Geförderten in dem halben Jahr nach Förderbeginn insgesamt 90 Tage weniger arbeitslos bzw. in Maßnahmen als die Vergleichsgruppenmitglieder.

Quelle: TrEffeR-Datenbank, Fördereintritte im September 2002, eigene Berechnungen.