

Wissenschaftlicher Beirat
beim Bundesministerium für Wirtschaft und Technologie (BMWi)

**Evaluierung wirtschaftspolitischer Fördermaßnahmen
als Element einer evidenzbasierten Wirtschaftspolitik**

Stand: 20.9.2013

1. Anlass	2
2. Hintergrund	5
2.1 Der gesetzliche und verwaltungsrechtliche Rahmen	5
2.2 Fördermaßnahmen des BMWi	7
2.3 Evaluierungserfahrungen der deutschen Arbeitsmarkt-, Sozial- und Familienpolitik.	8
2.4 Internationale Erfahrungen der Evaluierung wirtschaftspolitischer Fördermaßnahmen	9
3. Elemente einer zielführenden Evaluierung wirtschaftspolitischer Fördermaßnahmen	12
3.1 Ziele der Evaluierung	12
3.1.1 <i>Wirkungsanalyse als Instrument eines dynamischen Lernprozesses</i>	12
3.1.2 <i>Definition und Operationalisierung wirtschaftspolitischer Ziele</i>	13
3.2 Anforderungen an die Evaluierung	14
3.2.1 <i>Vergleich von Behandlungs- und Vergleichsgruppe</i>	14
3.2.2 <i>Der Nachweis der Wirksamkeit</i>	15
3.2.3 <i>Randomisierte Feldexperimente</i>	17
3.2.4 <i>Quasi-experimentelle Evaluierungsmethoden</i>	20
3.3 Grenzen der Evaluierung.....	26
3.4 Evaluierungen von einzelnen Fördermaßnahmen sowie Maßnahmenbereichen	28
4. Einsatz in der wirtschaftspolitischen Praxis	29
4.1 Datenerhebung und -zugang	29
4.1.1 <i>Existierende Datensätze und Eigenerhebungen von Evaluierungsdaten</i>	29
4.1.2 <i>Offener Datenzugang für Forschungszwecke in Forschungsdatenzentrum</i>	32
4.2 Rahmenbedingungen für unabhängige und ergebnisoffene Forschung	34
4.2.1 <i>Organisation der Evaluierung in einer spezifischen Einheit des Ministeriums</i>	34
4.2.2 <i>Ausschreibung und Öffnung für den wissenschaftlichen Wettbewerb</i>	36
4.2.3 <i>Vorbehaltlose Veröffentlichung</i>	37
5. Empfehlungen.....	38
Literatur	40
Mitgliederverzeichnis.....	45
Anhang: Fördermittel im Haushalt des Bundesministeriums für Wirtschaft und Technologie	48

Der Wissenschaftliche Beirat beim Bundesministerium für Wirtschaft und Technologie hat sich in mehreren Sitzungen, zuletzt am 20. September 2013, mit dem Thema

Evaluierung wirtschaftspolitischer Fördermaßnahmen als Element einer evidenzbasierten Wirtschaftspolitik

befasst und ist dabei zu der nachfolgenden Stellungnahme gelangt:

1. Anlass

Im Jahre 2013 gewährt der Bund 21 Milliarden Euro an Finanzhilfen und Steuervergünstigungen.¹ Allein im Budget des Bundesministeriums für Wirtschaft und Technologie (BMWi) werden jährlich über 5 Milliarden Euro (aus dem Gesamtbudget von gut 6 Milliarden Euro) für Wirtschaftsförderung veranschlagt,² darunter etwa jeweils über eine halbe Milliarde Euro für das „Zentrale Innovationsprogramm Mittelstand“ sowie für Zuweisungen im Rahmen der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“. In anderen Ressorts werden zum Teil weit höhere Förderausgaben getätigt, etwa für Maßnahmen der Sozialpolitik, der Gesundheitspolitik, der Arbeitsmarktpolitik, der Verkehrsförderung, der Familienpolitik, der wirtschaftlichen Zusammenarbeit und der Forschungs- und Innovationspolitik.³ Auch wenn eine objektive und sachgerechte Evaluierung für alle Ressorts von Bedeutung ist, konzentriert sich der Beirat in diesem Gutachten auf ein wichtiges Politikfeld aus dem Geschäftsbereich des BMWi, nämlich die Maßnahmen der Wirtschaftsförderung.

Inwiefern diese Maßnahmen unterstützenswerte Ziele verfolgen, diskutiert der Beirat in diesem Gutachten nicht. Ihm geht es allein um die abgeleitete Frage, ob der Bund die von ihm selbst gewählten Ziele erreicht, und ob er die gleichen Ziele auch mit geringerem finanziellem Aufwand erreichen könnte, so dass jedenfalls keine Steuergelder verschwendet werden. Daran

¹ Bundesministerium der Finanzen (2013), 24. Subventionsbericht der Bundesregierung, S.11. In der erweiterten Abgrenzung des Subventionsbegriffs des Kieler Instituts für Weltwirtschaft lagen die Subventionen in Deutschland im Jahre 2011 bei 166,7 Milliarden Euro (Laaser und Rosenschon (2013)).

² Werte aus dem Haushaltsplan für das Jahr 2013 für Wirtschaftsförderungsmaßnahmen in den Bereichen Innovation, Technologie und Neue Mobilität; Energie und Nachhaltigkeit; Mittelstand; und Chancen der Globalisierung; siehe Abschnitt 2.2.

³ Darüber hinaus gibt es viele wirtschaftspolitische Maßnahmen wie Gesetzesregelungen oder Festpreissetzungen, die zwar keine direkten Förderströme auslösen, gleichwohl aber weitreichende finanzwirksame oder anderweitig in die Lebenssituation der Bürger eingreifende Konsequenzen haben. Die Darstellung dieses Gutachtens beschränkt sich der Überschaubarkeit halber auf die Evaluierung von Fördermaßnahmen. Gleichwohl sollten aus den hier angeführten Gründen auch sonstige wirtschaftspolitische Maßnahmen einer überzeugenden Evaluierung unterzogen werden.

besteht in der Öffentlichkeit hohes Interesse. Es geht darum, die Effektivität (erreicht eine Maßnahme ihr Ziel?) und die Effizienz (welche Maßnahme erreicht ein Ziel zu geringstmöglichen Kosten?) wirtschaftspolitischer Maßnahmen zu steigern. So sollen etwa Mitnahmeeffekte vermindert werden, die entstehen, wenn öffentliche Fördermittel Ausgaben ersetzen, die ansonsten privat getätigt worden wären.

Der Beirat betrachtet die Wirkungsanalyse staatlicher Fördermaßnahmen als ein notwendiges Element der Ordnungspolitik.⁴ Die Wirtschaftlichkeit des Mitteleinsatzes und der damit verbundene Lernprozess im Verlaufe von Fördermaßnahmen sind nicht nur aus Gründen der Haushaltskonsolidierung in Zeiten von Schuldenbremsen und knappen Kassen geboten. Sie sind auch erforderlich, um gesetzte wirtschaftspolitische Ziele bestmöglich zu erreichen.

Die Bundeshaushaltsordnung schreibt für alle finanzwirksamen Maßnahmen eine Wirtschaftlichkeitsuntersuchung vor (§ 7 Abs. 2 BHO). Die Leitlinien der Subventionspolitik, die seit 2006 als Selbstbindung der Bundesregierung für die von ihr zu verantwortenden Maßnahmen bestehen, schreiben vor, dass die Ziele neuer Finanzhilfen in einer Weise bestimmt werden, die eine Erfolgskontrolle erlaubt.⁵ Das Bundesfinanzministerium dringt darauf, möglichst viele Finanzhilfen zu evaluieren. Tatsächlich ist das aktuell aber erst bei 44 von 62 berücksichtigten Finanzhilfen gelungen.⁶ Auch fehlt es in der allgemeinen Diskussion an Einigkeit über Methoden und Beurteilungskriterien.⁷ Der Beirat unterstützt die Bestrebungen zu umfassenderer Evaluierung. Er will mit diesem Gutachten einen Beitrag zur Weiterentwicklung der Evaluierungskultur in Deutschland leisten und Vorschläge zur Lösung von Problemen machen, die bislang oft der wirksamen Evaluierung wirtschaftspolitischer Maßnahmen im Wege stehen.

Mit der Anwendung von randomisierten Feldexperimenten und quasi-experimentellen Analysemethoden hat sich das wissenschaftliche Instrumentarium zur kausalen Evaluierung politischer Maßnahmen in den letzten drei Jahrzehnten deutlich verbessert, es wird in diesem Zusammenhang gar von einer „Glaubwürdigkeitsrevolution in der empirischen Wirtschaftswissenschaft“ gesprochen.⁸ Dadurch ist das Ziel einer evidenzbasierten Wirtschaftspolitik in weit größere Nähe gerückt. Das belegen auch die Erfahrungen aus anderen Politikfeldern und

⁴ Dabei geht es im vorliegenden Gutachten um Ex-post-Evaluierungen in dem Sinne, dass die Ergebnisse bereits implementierter Fördermaßnahmen begleitend oder abschließend empirisch überprüft werden, wobei das Evaluierungsdesign immer auch schon ex ante mitgedacht werden muss.

⁵ Bundesministerium der Finanzen (2013), S. 10.

⁶ Bezogen auf den Berichtszeitraum 2011 bis 2014; Bundesministerium der Finanzen (2013), S. 35.

⁷ Bundesministerium der Finanzen (2011), S. 34; Präsident des Bundesrechnungshofes (2013), S. 5.

⁸ Angrist und Pischke (2010), „The Credibility Revolution in Empirical Economics“.

aus dem Ausland. So ist die Evaluierung arbeitsmarktpolitischer Maßnahmen bereits wesentlich weiter entwickelt als die Evaluierung anderer wirtschaftspolitischer Maßnahmen, sowohl bei der Datenverfügbarkeit als auch bei der gesetzlichen Verankerung der empirisch-kausal-analytischen Wirkungsforschung (siehe Abschnitt 2.3). Auch in internationalen Erfahrungen lassen sich Vorbilder erkennen (siehe Abschnitt 2.4). Diese Erfahrungen belegen nicht nur das große Potential der Evaluierung, die Wirksamkeit politischer Maßnahmen zu erhöhen. Sie zeigen auch, dass Evaluierungen, die mit unzureichender Methodik durchgeführt werden, leicht zu irreführenden Schlussfolgerungen führen können. Insofern besteht großer Bedarf an einer Weiterentwicklung der bisherigen Evaluierungspraxis.

Der Beirat kommt zu folgenden Empfehlungen:

- Zur Sicherung der Wirtschaftlichkeit wirtschaftspolitischer Fördermaßnahmen sollten diese ab einer gewissen Größe standardmäßig einer Evaluierung unterzogen werden.⁹ Von Zeit zu Zeit sollten die verschiedenen Maßnahmen eines Politikbereiches vergleichend evaluiert werden.
- Damit eine objektive und sachgerechte Evaluierung möglich ist, sollte vor Einführung einer Fördermaßnahme nicht nur festgelegt werden, welche Ziele sie verfolgt, sondern auch, anhand welcher Zielgrößen ihre Wirksamkeit gemessen werden soll.
- Die Evaluierungen müssen Mindestanforderungen der empirisch-kausalanalytischen Wirkungsforschung genügen, um Aussagen über die ursächliche Wirkung wirtschaftspolitischer Fördermaßnahmen ableiten zu können; randomisierte Feldexperimente und quasi-experimentelle Methoden sollten soweit wie möglich Anwendung finden. Insbesondere Pilotmaßnahmen sollten in Form von randomisierten Experimenten evaluiert werden.
- Da objektive Evaluierung erst durch Verfügbarkeit und Zugang zu entsprechenden Daten möglich ist, sollte mit einem Forschungsdatenzentrum für Evaluierungsdaten eine Dateninfrastruktur aufgebaut werden, die den evaluierenden Einrichtungen und der Wissenschaft insgesamt Zugang zu evaluierungsrelevanten Mikrodaten ermöglicht.
- Um die Güte der Evaluierungsstudien zu steigern, müssen Rahmenbedingungen für eine unabhängige und ergebnisoffene Evaluierung geschaffen werden. Eine für die

⁹ Der Begriff „Fördermaßnahme“ wird im vorliegenden Gutachten allgemein für Förderprogramme verwendet, die in der Regel zahlreiche Einzelprojekte etwa in verschiedenen geförderten Unternehmen beinhalten.

Initiierung, Steuerung und Beaufsichtigung der Evaluierungen zuständige Organisationseinheit des Ministeriums sollte die wissenschaftliche Qualität und die fachliche Unabhängigkeit der Studien sicherstellen. Darüber hinaus muss die wettbewerbliche Ausschreibung garantiert sein. Bei anspruchsvollen Vorhaben sollte der wissenschaftliche Wettbewerb zur Qualitätssicherung genutzt werden.

- Um Transparenz über Vorgehen und Ergebnisse der Evaluierungen herzustellen, sind die Evaluierungsstudien ohne Publikationsvorbehalt spätestens sechs Monate nach Abschluss zu veröffentlichen.

Das Gutachten ist wie folgt aufgebaut: In Abschnitt 2 werden der gesetzliche Hintergrund, aktuelle wirtschaftspolitische Förderprogramme sowie Evaluierungserfahrungen aus der in- und ausländischen Wirtschaftspolitik dargestellt. Abschnitt 3 stellt Elemente einer zielführenden Politikevaluierung dar. Abschnitt 4 diskutiert, welche Rahmenbedingungen notwendig sind, damit diese Elemente konkret in der wirtschaftspolitischen Praxis zum Einsatz kommen können.

2. Hintergrund

2.1 Der gesetzliche und verwaltungsrechtliche Rahmen

Die Evaluierung wirtschaftspolitischer Fördermaßnahmen ist in Deutschland gesetzlich vorgeschrieben. Artikel 114 Abs. 2 des Grundgesetzes legt fest: „Der Bundesrechnungshof [...] prüft die Rechnung sowie die Wirtschaftlichkeit und Ordnungsmäßigkeit der Haushalts- und Wirtschaftsführung“. § 7 Abs. 2 BHO schreibt vor: „Für alle finanzwirksamen Maßnahmen sind angemessene Wirtschaftlichkeitsuntersuchungen durchzuführen.“¹⁰ Die Verwaltungsvorschriften zur BHO definieren Wirtschaftlichkeit als „die bestmögliche Nutzung von Ressourcen“; es „ist die günstigste Relation zwischen dem verfolgten Zweck und den einzusetzenden Mitteln (Ressourcen) anzustreben“. Wirtschaftlichkeitsuntersuchungen sind „Instrumente zur Umsetzung des Grundsatzes der Wirtschaftlichkeit“ und sollen sowohl als „Planungsinstrument“ als auch als „Instrument der [begleitenden sowie abschließenden] Erfolgskontrolle“ eingesetzt werden. Als letztere sollen sie feststellen, „ob und in

¹⁰ Für Details siehe die „Anforderungen an Wirtschaftlichkeitsuntersuchungen finanzwirksamer Maßnahmen nach § 7 Bundeshaushaltsordnung“, die der Präsident des Bundesrechnungshofes (2013) als Bundesbeauftragter für Wirtschaftlichkeit in der Verwaltung als Empfehlungen herausgegeben hat.

welchem Ausmaß die [mit einer Maßnahme] angestrebten Ziele erreicht wurden, ob die Maßnahme ursächlich für die Zielerreichung war und ob die Maßnahme wirtschaftlich war.“¹¹

Auch bei den Bestrebungen zum Abbau von Bürokratiebelastungen wird die systematische Evaluierung staatlicher Maßnahmen zum Thema. Nach der am 23. Januar 2013 beschlossenen „Konzeption zur Evaluierung neuer Regelungsvorhaben“ des Staatssekretärsausschusses Bürokratieabbau ist „für alle wesentlichen Regelungsvorhaben“ (mit einem Erfüllungsaufwand von mindestens 1 Mio. Euro) eine „Evaluierung“ vorzusehen, die „einen Zusammenhang [...] zwischen Ziel und Zweck einer Regelung und den tatsächlichen erzielten Wirkungen sowie den damit verbundenen Kosten“ herstellt.¹² „Wichtigstes Evaluierungskriterium ist die Zielerreichung, da Regelungen, die das im Regelungsentwurf formulierte Ziel verfehlen, meist unnötigen Erfüllungsaufwand verursachen.“¹³ Allerdings sind „Tiefe“ (Teile vs. Regelungsvorhaben insgesamt, Vollzug), „Methodik (von interner bis hin zu einer wissenschaftlichen Evaluierung)“ und „Umfang (von ‚Zwei-Seiten-Bericht‘ bis hin zu einem ausführlichen Bericht ...)“ der Evaluierung explizit „nicht vorbestimmt“ und dem federführenden Ressort anheimgestellt.¹⁴

Generell ist zu konstatieren, dass in der Praxis bisher nicht für alle existierenden Förderprogramme angemessene Wirtschaftlichkeitsuntersuchungen vorliegen. Nach Angaben des Bundesrechnungshofes lagen sogar „für nahezu 85 Prozent der von Bundesministerien und nachgeordneten Behörden gemeldeten finanzwirksamen Maßnahmen [...] keine Wirtschaftlichkeitsuntersuchungen vor.“¹⁵ Wo diese vorliegen, erfüllen sie häufig nicht Mindestanforderungen an eine überzeugende Evaluierung wie etwa den Vergleich von geförderten und nicht geförderten Objekten anhand objektivierbarer Kennzahlen. Von einer Ursächlichkeitsanalyse im Sinne des aktuellen Standes der Wissenschaft kann in einer Vielzahl der Fälle nicht die Rede sein.

¹¹ Weitere Erläuterungen und Hinweise zur Durchführung von Wirtschaftlichkeitsuntersuchungen finden sich in der als Empfehlung abgefassten Arbeitsanleitung Einführung in Wirtschaftlichkeitsuntersuchungen des Bundesministeriums der Finanzen (siehe Präsident des Bundesrechnungshofes (2013)).

¹² Bundesregierung (2013), S. 64. Der Vorsitzende des Nationalen Normenkontrollrats, Johannes Ludewig, bezeichnet die Einführung der systematischen Ex-post-Evaluierung als „historisch“: „Es ist eigentlich erstaunlich, dass es eine solche systematische Ex-post-Evaluierung von Gesetzen bisher in Deutschland nicht gegeben hat Insofern kann der entsprechende Beschluss vom Januar 2013, eine solche systematische Ex-post-Evaluierung für alle Gesetze mit Folgekosten ab einer Million Euro erstmals in Deutschland einzuführen, ohne Übertreibung historisch genannt werden.“ (Handelsblatt, 15. Mai 2013, S. 48)

¹³ Bundesregierung (2013), S. 64.

¹⁴ Bundesregierung (2013), S. 65.

¹⁵ Präsident des Bundesrechnungshofes (2013), S. 16; Aussage bezieht sich auf das Prüfwahljahr 2006 (S. 73).

2.2 Fördermaßnahmen des BMWi

Der Beirat konzentriert sich in diesem Gutachten auf ein Politikfeld aus dem Geschäftsbereich des Bundeswirtschaftsministeriums, die Maßnahmen der Wirtschaftsförderung. Obgleich für etliche dieser Fördermaßnahmen Evaluierungsstudien vorliegen, entsprechen diese zumeist nicht den im vorliegenden Gutachten vorgestellten methodischen Standards. Die Wirtschaftsförderung durch das BMWi ist umfangreich und vielfältig. Das Volumen beträgt im Jahr 2013 über 5 Mrd. Euro. Die Förderung verteilt sich laut Haushaltsplan auf vier Politikbereiche: Innovation, Technologie und Neue Mobilität; Mittelstand; Energie und Nachhaltigkeit; und Chancen der Globalisierung. Die Tabelle im Anhang gibt einen Komplettüberblick über die einzelnen Posten der Fördermittel im Haushalt des BMWi. Nachfolgend werden beispielhaft existierende Wirtschaftsfördermaßnahmen aufgezählt, die das Anwendungsgebiet möglicher Evaluierungsmaßnahmen verdeutlichen und auf die im weiteren Verlauf für Anwendungsbeispiele zurückgegriffen werden kann.¹⁶

Im Politikbereich „Innovation, Technologie und Neue Mobilität“ bildet das technologieoffene Zentrale Innovationsprogramm Mittelstand (ZIM) mit einem Fördervolumen von über einer halben Milliarde Euro den Schwerpunkt der Förderung des innovativen Mittelstandes. Seit Programmstart Mitte 2008 gab es im ZIM über 30.000 Anträge und über 21.000 Förderzusagen (durchschnittlich rund 4.500 Neubewilligungen pro Jahr) mit insgesamt 2,7 Mrd. Euro bewilligten Mitteln (AiF Projekt GmbH (2013)). Für die Unterstützung der Forschungsinfrastruktur für den Mittelstand stehen knapp 200 Mio. Euro zur Verfügung. Unter die Förderung des Technologie- und Innovationstransfers (TuIT) mit knapp 30 Mio. Euro fallen beispielsweise die Patentinitiative SIGNO sowie die Förderung des Normenwesens. Das Beratungsprogramm go-Inno unterstützt kleine und mittlere Unternehmen (KMU) mit Gutscheinen von über 14 Mio. Euro für externe Beratungen zur Produkt- und Prozessinnovation. Die Förderung der Luft- und Raumfahrtindustrie beläuft sich auf insgesamt 1,4 Milliarden Euro. Für Anwendungen der Informations- und Kommunikationstechnologien (IKT) sind rund 72 Mio. Euro angesetzt.

Im Politikbereich „Mittelstand: Gründen, Wachsen, Investieren“ stehen 583 Mio. Euro für die Investitionsförderung im Rahmen der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRW) zur Verfügung. Gefördert werden auch innovative („Gründerland Deutschland“, 83 Mio. Euro) und technologieorientierte (High-Tech-Gründerfonds) Unternehmensgründungen, Ausgründungen aus wissenschaftlichen Einrichtungen (EXIST) sowie

¹⁶ Die folgende Darstellung beruht direkt auf der eigenen Darstellung der Politikbereiche des Haushalts durch das BMWi (Stand 22. November 2012).

die Stärkung des Wagniskapitalmarktes (Investitionszulage für Business Angel, 150 Mio. Euro bis 2016). Das Programm „Förderung unternehmerischen Know-hows“ unterstützt KMU bei der Inanspruchnahme von externem Know-how zu Fragen der Unternehmensführung (40 Mio. Euro). Darüber hinaus gibt es Mittel zur Unterstützung überbetrieblicher Lehrgänge (74 Mio. Euro), zur Unterstützung der Fachkräftesicherung (13 Mio. Euro) sowie für eine Kampagne zur Anwerbung von Fachkräften aus dem Ausland.

Im Politikbereich „Energie und Nachhaltigkeit“ wird die Energieforschung mit 150 Mio. Euro gefördert. Für Maßnahmen zur Steigerung der Energieeffizienz, insbesondere Energieberatung für KMU und private Verbraucher, stehen 30 Mio. Euro zur Verfügung. Im Rahmen der Beendigung des subventionierten Steinkohlenbergbaus sind 1,2 Mrd. Euro vorgesehen, für die Wismut-Sanierung 148 Mio. Euro. Hinzu kommen 421,5 Mio. Euro aus dem Energie- und Klimafonds (EKF), unter anderem 284 Mio. Euro für Energieeffizienz und 129 Mio. Euro für Elektromobilität.

Im Politikbereich „Chancen der Globalisierung“ fördert das „Programm zur Erschließung von Auslandsmärkten“ (81 Mio. Euro) Exportinitiativen, Messebeteiligungen und Managerfortbildungen. Die Außenwirtschaftsförder- und Standortmarketinggesellschaft des Bundes, Germany Trade and Invest (GTAI), und das Netzwerk der Auslandshandelskammern werden mit 57 Mio. Euro gefördert. Darüber hinaus werden die Entsorgung russischer Atom-U-Boote (65 Mio. Euro), die Stadtbahn in Ho Chi Minh-Stadt (30 Mio.), die Teilnahme an Weltausstellungen (9 Mio. Euro) sowie die Deutsche Zentrale für Tourismus (28 Mio. Euro) finanziert.

2.3 Evaluierungserfahrungen der deutschen Arbeitsmarkt-, Sozial- und Familienpolitik

In Deutschland hat die empirisch-kausalanalytische Evaluierungsforschung im letzten Jahrzehnt vor allem in der Arbeitsmarktpolitik Einzug gehalten. Vor dem Hintergrund eines weltweiten Trends der Methodenentwicklung, der besseren Verfügbarkeit von Mikrodaten sowie der Notwendigkeit eines wirksamen und wirtschaftlichen Einsatzes der verfügbaren Mittel wurden und werden die einzelnen Instrumente in den verschiedenen Bereichen der aktiven Arbeitsmarktpolitik umfassend evaluiert.¹⁷

Ein wichtiger Ausgangspunkt der Weiterentwicklung der arbeitsmarktpolitischen Evaluierungsforschung in Deutschland waren wesentliche Fortschritte bei Aufbereitung, Verfügbarkeit und Zugang zu den statistischen Daten der Bundesagentur für Arbeit. Dies wurde seit 2004 vor allem mit der Einrichtung und Weiterentwicklung des Forschungsdatenzentrums der

¹⁷ Siehe den Sachstandsbericht der Evaluation der Instrumente (Bundesministerium für Arbeit und Soziales und Institut für Arbeitsmarkt- und Berufsforschung (2011)).

Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung vorangebracht, das der interessierten Fachöffentlichkeit einen institutionalisierten Datenzugangsweg bereitstellt.¹⁸ Einen wesentlichen Schub erfuhr die Evaluierungsforschung auch durch die explizite gesetzliche Verankerung der zeitnahen Wirkungsforschung der Arbeitsförderung als ständige Aufgabe des Instituts für Arbeitsmarkt- und Berufsforschung in § 282 Sozialgesetzbuch III im Rahmen des Job-AQTIV-Gesetzes 2002 sowie durch die Hartz-Gesetze 2003-2004, die ebenfalls umfassende Evaluierungsmaßnahmen vorsahen.¹⁹

Auf der Basis der durchgeführten Wirksamkeitsanalysen kam es nicht zuletzt mit dem 2009 in Kraft getretenen Gesetz zur Neuausrichtung der arbeitsmarktpolitischen Instrumente zu einer deutlichen Umorientierung der aktiven Arbeitsmarktpolitik, mit der ineffiziente²⁰ Arbeitsmarktinstrumente abgeschafft und wirksame Instrumente weiterentwickelt werden sollten.²¹ In der Folge wurden etwa die Arbeitsbeschaffungsmaßnahmen aufgrund von negativen Evaluierungsergebnissen mit dem Gesetz zur Verbesserung der Eingliederungschancen am Arbeitsmarkt 2012 gänzlich abgeschafft. So konnten die staatlichen Fördermaßnahmen in der Arbeitsmarktpolitik insgesamt effektiver und effizienter ausgerichtet werden.

Mit der Gesamtevaluierung der ehe- und familienbezogenen Leistungen hat die breit angelegte empirisch-kausalanalytische Evaluierungsforschung auch in der Familienpolitik Einzug gehalten.²² Als Teil des Koalitionsvertrages ist sie vom Bundesministerium für Familie, Senioren, Frauen und Jugend zusammen mit dem Bundesministerium der Finanzen als vierjähriges Forschungsprogramm angelegt. Die von verschiedenen Forschungsinstituten durchgeführten 11 Teilstudien evaluieren seit 2009 die Wirkung von zentralen ehe- und familienbezogenen Leistungen auf familienpolitische Ziele. Im Rahmen einzelner Module werden zudem wissenschaftlich nutzbare Datenbestände erhoben.

2.4 Internationale Erfahrungen der Evaluierung wirtschaftspolitischer Fördermaßnahmen

In den Vereinigten Staaten gehört die Politikevaluierung seit den 1980er Jahren zum Standardrepertoire in Forschung und Lehre. Das U.S.-amerikanische National Bureau of Economic Research publiziert jährlich Hunderte von Studien, die die Wirkung von staatlichen Eingriffen in das Verhalten von Haushalten und Unternehmen mit ökonometrischen Verfahren auf der Basis von Mikrodatensätzen untersuchen. Die vornehmlich mit Längsschnittdaten

¹⁸ Allmendinger und Kohlmann (2005).

¹⁹ Brinkmann, Hujer und Koch (2006).

²⁰ Vgl. etwa Wunsch und Lechner (2008).

²¹ Steinke u.a. (2012).

²² <http://www.bmfsfj.de/BMFSFJ/familie,did=195944.html> (30.6.2013).

arbeitenden Studien sind zum Teil sehr präzise auf begrenzte Einzelmaßnahmen fokussiert. Die untersuchten Eingriffe stammen praktisch aus der gesamten Palette staatlicher Maßnahmen, so etwa Steuern (z.B. die negative Einkommenssteuer), Preissetzungen (z.B. den gesetzlichen Mindestlohn) und Mengensetzungen (z.B. das Renteneintrittsalter).

Wie weit die Möglichkeiten zur Evaluierung wirtschaftspolitischer Maßnahmen reichen, sollen einige ausländische Studien zur Evaluierung von Maßnahmen der Regional-, Innovations- und Mittelstandsförderung deutlich machen – Anwendungsgebiete, die, wie gesehen, auch in den Fördermaßnahmen des BMWi eine große Rolle spielen. In Großbritannien, Frankreich und den Vereinigten Staaten wurden Maßnahmen evaluiert, die darauf abzielen, Beschäftigung und Wirtschaftskraft in wirtschaftlich rückständigen Regionen zu fördern.²³ Um den kausalen Effekt der Maßnahmen zu prüfen, analysieren die drei Studien kleingliedrige Regionaldaten mit quasi-experimentellen Methoden. Die Ergebnisse liefern ein relativ einheitliches Bild: Die Maßnahmen führten in der Tat zu zusätzlichen Wirtschaftsansiedlungen in den geförderten Regionen. Diese gingen allerdings auf Kosten von Ansiedlungen in nicht geförderten Regionen (Verlagerungseffekte), so dass der gesamtwirtschaftliche Effekt nur in einem der Fälle als positiv angesehen wird.

Das U.S.-amerikanische Small Business Innovation Research (SBIR) Programm fördert seit 1982 Forschungs- und Entwicklungs-Vorhaben (F&E) von KMU mit bisher über 27 Milliarden Dollar.²⁴ Eine Studie, die die gegenseitige Verursachung von Mittelempfang und F&E-Tätigkeit berücksichtigt, findet, dass Firmen mit mehr Beschäftigten und mehr F&E zwar mehr SBIR-Mittel erhalten, dass die SBIR-Mittel aber keine positiven Beschäftigungseffekte auslösen.²⁵ Darüber hinaus verringert jeder Dollar Förderung die privaten F&E-Ausgaben um einen Dollar (Mitnahmeeffekte). Einfache Korrelationen wie auch Studien, die lediglich geförderte Unternehmen ohne Vergleich mit nicht geförderten Unternehmen befragten, hatten den Anschein erweckt, dass die Fördermaßnahmen deutlich positivere Effekte hatten.²⁶

Ebenfalls im Bereich der innovationsorientierten Mittelstandsförderung gibt es in den Vereinigten Staaten über 1.000 geförderte Small Business Development Centers (SBDC), die

²³ Einiö und Overman (2012) untersuchen die „Local Enterprise Growth Initiative“ (LEGI) in Großbritannien. Mayer, Mayneris und Py (2012) untersuchen das „Zones Franches Urbaines“ (ZFU) Programm in Frankreich. Kline und Moretti (2013) untersuchen das „Tennessee Valley Authority“ (TVA) Programm in den Vereinigten Staaten.

²⁴ Vgl. <http://www.sbir.gov> (30.6.2013) oder auch Link und Scott (2012).

²⁵ Wallsten (2000). In einer weiteren Evaluierungsstudie kommt Lerner (1999) mithilfe eines Matching-Ansatzes zu ähnlichen Befunden; mögliche positive Effekte waren auf Regionen mit hoher Venture Capital Aktivität und Hochtechnologieindustrien beschränkt. Siehe hierzu auch Falck und Wiederhold (2013).

²⁶ Wessner (2008).

Beratung und Weiterbildung für Unternehmensgründungen anbieten. Das U.S. Department of Labor hat zusammen mit der U.S. Small Business Association die bisher größte randomisierte Feldevaluierung solcher Maßnahmen durchgeführt, das Projekt „Growing America through Entrepreneurship“ (Project GATE).²⁷ Ein Vergleich der geförderten mit den nicht geförderten Antragstellern hat ergeben, dass die Maßnahme einerseits für Arbeitslose einen bedeutsamen kurzfristigen Effekt auf die Aufnahme einer selbständigen Tätigkeit hatte, dass sich aber andererseits für andere Zielgruppen und Zeithorizonte keine starken Effekte auf Selbständigkeit, Unternehmensergebnis und weiter gefasste Ergebnisvariablen zeigen.

Zusammengenommen zeigen die Beispiele, dass wissenschaftliche Evaluierungsstudien helfen können zu lernen, welche Aspekte wirtschaftspolitischer Fördermaßnahmen besonders erfolgreich sind und wo nachgesteuert werden sollte. Gleichzeitig wird auch deutlich, dass Evaluierungen auf der Basis unzureichender Methoden zu durchaus irreführenden Schlussfolgerungen führen können.

Das letztgenannte Beispiel, das auf einem von einem Ministerium initiierten randomisierten Evaluierungsverfahren basiert, belegt, dass Evaluierungen nicht nur in Einzelstudien der wissenschaftlichen Forschung umgesetzt, sondern auch systematisch bei der Weiterentwicklung wirtschaftspolitischer Fördermaßnahmen eingesetzt werden können. In der internationalen Praxis geschieht dies zunehmend. So werden etwa in den skandinavischen Ländern regelmäßig umfassende Registerdaten zur Evaluierung sozial- und wirtschaftspolitischer Maßnahmen ausgewertet.²⁸ In den Niederlanden ist das Centraal Planbureau neben der Kosten-Nutzen-Analyse wirtschaftspolitischer Vorschläge auch dafür zuständig, die Effekte wirtschaftspolitischer Maßnahmen in weiten Anwendungsbereichen mit aktuellen wissenschaftlichen Methoden zu evaluieren.²⁹

²⁷ Vgl. Benus u.a. (2009) und Fairlie, Karlan und Zinman (2012).

²⁸ In Dänemark werden darüber hinaus systematisch Feldexperimente in der Evaluierung arbeitsmarktpolitischer Maßnahmen eingesetzt; vgl. Graversen und van Ours (2008) und Pedersen, Rosholm und Svarer (2012).

²⁹ Das Centraal Planbureau (CPB) hat sich den englischen Namen CPB Netherlands Bureau for Economic Policy Analysis gegeben, da es explizit keine Planung macht, sondern die Effekte derzeitiger und zukünftiger Regierungspolitik analysiert (siehe <http://www.cpb.nl/en/about-cpb> (30.6.2013)). Als eines von unzähligen Anwendungsbeispielen untersuchen etwa Cornet, Vroomen und Steeg (2006) mithilfe eines randomisierten Experiments die Effekte von Innovationsgutscheinen für KMU.

3. Elemente einer zielführenden Evaluierung wirtschaftspolitischer Fördermaßnahmen

3.1 Ziele der Evaluierung

3.1.1 Wirkungsanalyse als Instrument eines dynamischen Lernprozesses

Unter „Evaluierung“ wird im vorliegenden Gutachten die Wirkungsanalyse verstanden. Wirkungsanalyse im Sinne der Aufdeckung eines Ziel-Mittel-Zusammenhangs hat ein abgeleitetes Anliegen. Sie fragt nicht, ob der Staat bestimmte Ziele verfolgen sollte. Es geht der Evaluierung stattdessen darum, ob die Mittel effektiv und effizient eingesetzt werden. Deshalb beschränkt sie sich auf die Fragen, ob (a) der Staat mit dem Einsatz der Mittel die von ihm selbst gewählten Ziele überhaupt erreicht (Effektivität) und ob (b) er die gleiche Wirkung auf kostengünstigerem Weg erreichen könnte (Effizienz). Effizienz bedeutet sowohl, dass beim Einsatz einer bestimmten Maßnahme keine Mittel verschwendet werden, als auch, dass es keinen alternativen, weniger kostspieligen Weg gibt, dasselbe Ziel zu erreichen.

Um diese Fragen beantworten zu können, bedarf es der empirischen Überprüfung der Ergebnisse politischer Maßnahmen. Deshalb fokussiert das vorliegende Gutachten auf Ex-post-Evaluierungen in dem Sinne, dass die Wirkungen einer bereits implementierten Maßnahme im Rahmen einer begleitenden oder abschließenden Erfolgskontrolle empirisch untersucht werden, um damit eine Kosten-Nutzen-Analyse zu ermöglichen. Damit thematisiert es nicht hypothetische Ex-ante-Evaluierungen, die in der Planungsphase vor Implementierung einer Maßnahme deren mögliche Wirkungen abschätzen sollen. Allerdings haben, wie im Folgenden diskutiert, Ex-post-Evaluierungen auch immer ein Ex-ante-Element in dem Sinne, dass es auch schon vor Implementierung der Maßnahme, etwa bei ihrer Ausschreibung, einer Verständigung darüber bedarf, was später evaluiert werden soll. Darüber hinaus sollten sie sich selbstverständlich auf die Anpassung laufender Fördermaßnahmen sowie die Konzeption zukünftiger Maßnahmen auswirken.

Das Gutachten beschäftigt sich auch nicht mit der – gerade in der Praxis des BMWi bereits weit entwickelten – Evaluierung des Förderprozesses in Sinne einer Befragung der geförderten Unternehmen, wie die Förderpraxis zu bewerten und gegebenenfalls zu verbessern wäre („formative“ Evaluierung). Stattdessen geht es hier um eine Evaluierung, die die Wirkung und das Ergebnis einer Maßnahme in der Gesamtheit der geförderten Unternehmen bestimmen möchte und damit Argumente für ihre Legitimität liefern kann („summative“ Evaluierung). In der Evaluierungspraxis wird es darum gehen, eine Mischung aus Implementations- und

Wirkungsanalyse, aus qualitativen Interviews und Fallstudien einerseits und quantitativer empirisch-kausalanalytischer Wirkungsforschung andererseits zum Einsatz zu bringen.

In der wirtschaftspolitischen Praxis ist die Wahl der besten Maßnahmen zur Zielerreichung kein statischer Akt. Oftmals ist nicht von vornherein offensichtlich, welche Maßnahmen am besten geeignet sind, ein politisches Ziel zu erreichen. Evaluierungen können dabei helfen, dies zu lernen. Bei der Umsetzung von Maßnahmen können Lerneffekte erzielt werden, die in einem dynamischen Lernprozess dazu führen, Mittel effektiver einzusetzen. Insofern findet die auf eine empirische Analyse der Ergebnisse zielende Evaluierung auch bereits im Verlaufe einer Fördermaßnahme und nicht erst nach Abschluss eines Förderprogramms statt. In diesem Sinne wirken Evaluierungen von Fördermaßnahmen als Element einer evidenzbasierten Wirtschaftspolitik, die bestrebt ist, aufgrund immer neuer Evidenz zu lernen, dem bestmöglichen Erreichen ihrer Ziele näher zu kommen.

3.1.2 Definition und Operationalisierung wirtschaftspolitischer Ziele

Obwohl Wirkungsanalyse den Primat der Politik zur Definition wirtschaftspolitischer Ziele akzeptiert, setzt sie Anforderungen an die Wirtschaftspolitik. Die Politik ist dazu aufgerufen, zu entscheiden. Damit Wirkungsanalyse möglich wird, muss die Politik ihre Entscheidungen explizit machen. Sie muss sich offen zu ihren normativen Anliegen bekennen.

Darüber hinaus muss sie diese Anliegen in einer Weise operationalisieren, die es möglich machen, zwei Fragen zu beantworten: Hat die politische Maßnahme das selbst gesetzte Ziel erreicht? Welche Mittel waren dafür erforderlich? Operationalisierung bedeutet praktisch: Für die jeweilige politische Maßnahmen ist zu bestimmen, anhand welcher Parameter über Erfolg oder Misserfolg entschieden werden soll. Die Definition dieser Parameter – üblicherweise in Form einer Serie von Zielindikatoren – bleibt Sache der politisch Verantwortlichen. Aber Wirtschaftspolitik als Ordnungspolitik erfordert, dass diese Parameter im Vorhinein formuliert werden.³⁰

Gleichwohl lässt sich die Evaluierung auch jenseits der formulierten Ziele einer wirtschaftspolitischen Fördermaßnahme auf mögliche unintendierte Effekte ausweiten.³¹ Dies scheint

³⁰ So sehen auch die Allgemeinen Verwaltungsvorschriften zu § 7 BHO vor, dass der Haushaltsplan „in aller Regel die Aufgaben (Ergebnis, Ziele) bereits formuliert“ und dass Wirtschaftlichkeitsuntersuchungen bereits in der Planungsphase Aussagen über „Ziele, Prioritätsvorstellungen und mögliche Zielkonflikte“ sowie über „Kriterien und Verfahren für Erfolgskontrollen“ enthalten müssen. Auch in den Leitlinien der Subventionspolitik der Bundesregierung heißt es: „Die Ziele der Finanzhilfen werden in einer Form festgehalten, die eine Erfolgskontrolle ermöglicht.“ (Bundesministerium der Finanzen (2013), S. 10).

³¹ Auch die Verwaltungsvorschriften zu § 7 BHO sehen vor, dass die Wirkungskontrolle „alle beabsichtigten und unbeabsichtigten Auswirkungen der durchgeführten Maßnahme“ ermitteln soll.

insbesondere dann geboten, wenn die Maßnahme möglicherweise mit anderen expliziten Zielen der Wirtschaftspolitik in Konflikt stehen könnte. Dies scheint nicht zuletzt auch deshalb ratsam, weil die jeweiligen Ziele einer Einzelmaßnahme durchaus durch partikuläre Interessengruppen beeinflusst sein können.

3.2 Anforderungen an die Evaluierung

Aufgabe der Evaluierung ist es, den Zusammenhang zwischen den Maßnahmen und den von der Politik gesetzten Zielparametern zu untersuchen. Überzeugende Evaluierung kann sich also nicht darin erschöpfen, qualitative Beurteilungen der Prozesse und subjektive Einschätzungen der Beteiligten zusammenzutragen. Sie muss untersuchen, inwieweit die zu untersuchenden Maßnahmen zu objektiv operationalisierten Ergebnissen beigetragen haben. Um zu entscheiden, ob eine wirtschaftspolitische Maßnahme ihr Ziel erreicht hat, genügt es auch nicht, den geregelten Lebensbereich vor und nach dem Inkrafttreten der Maßnahme zu vergleichen. Denn die Maßnahme war ja nur dann erfolgreich, wenn die beobachtete Veränderung gerade auf der Maßnahme beruht, wenn also ein kausaler Zusammenhang zwischen der Maßnahme und der Veränderung besteht.

3.2.1 Vergleich von Behandlungs- und Vergleichsgruppe

Die entscheidende Herausforderung der Evaluierung besteht darin zu bestimmen, was geschehen wäre, wenn die zu evaluierende Maßnahme nicht durchgeführt worden wäre.³² Deshalb muss neben der „Behandlungsgruppe“ von (beispielsweise) Unternehmen, die einer Maßnahme ausgesetzt wurden, immer auch eine Vergleichsgruppe („Kontrollgruppe“) solcher Unternehmen betrachtet werden, die sich von der Kontrollgruppe nur dadurch unterscheiden, dass sie nicht an der Maßnahme teilgenommen haben. Ein Vergleich der beiden Gruppen kann dann (unter weiteren, im Folgenden diskutierten Bedingungen) Aufschluss über die Wirkung der Maßnahme geben.

Als wesentliche Anforderung einer Wirkungsanalyse müssen geförderte Objekte also immer einer Vergleichsgruppe von Objekten, die (zunächst) keine Förderung erhalten haben, gegenübergestellt werden. Deshalb können Studien, die lediglich die geförderten Unternehmen befragen und keinen Vergleich mit nicht geförderten Unternehmen durchführen, auch keine Evaluierungen im Sinne einer kausalen Wirkungsanalyse sein.

³² Holland (1986) bezeichnet es als das fundamentale Problem der kausalen Inferenz, dass wir niemals dasselbe Objekt sowohl im Zustand der „Behandlung“ als auch im Zustand der Nicht-Behandlung beobachten können, sondern immer höchstens in einem dieser Zustände. Gleichwohl zu belastbaren Aussagen über kausale Effekte zu gelangen ist die Aufgabe der Evaluierungsforschung.

Beispielsweise könnten im Rahmen eines Förderprogramms, das Beratungsleistungen für die Innovationstätigkeit von KMU finanziert, die geförderten Unternehmen über die Art und Weise befragt werden, wie sie die Fördermittel eingesetzt haben und inwiefern sie denken, dass dies ihre unternehmerischen Prozesse verbessert hat. Solche Informationen sind durchaus nützlich, um mögliche Wirkungsmechanismen besser zu verstehen. Sie erlauben aber keine kausale Wirkungsanalyse. Denn möglicherweise haben sich etwa gerade solche Unternehmen für die Maßnahme beworben, die ein größeres Bewusstsein für die Notwendigkeit gezielter Innovationsmaßnahmen haben und deshalb ihre Innovationsanstrengungen auch ohne die Förderung gesteigert hätten. Offensichtlich ist es bei einer solchen Vorgehensweise auch naheliegend, dass Unternehmen, die in den Genuss von Fördermaßnahmen gekommen sind, keinen Anreiz haben zu erklären, diese Maßnahmen hätten keine entsprechende Wirkung gehabt. Insofern ist ein Vergleich objektiver Ergebnisindikatoren zwischen geförderten und nicht geförderten Unternehmen für eine überzeugende Evaluierung unabdingbar.

3.2.2 Der Nachweis der Wirksamkeit

Darüber hinaus zielt die Evaluierungsforschung darauf ab, dass mögliche Unterschiede in betrachteten Ergebnissen zwischen den beiden Gruppen tatsächlich von der Politikmaßnahme verursacht wurden und nicht aus anderen Gründen entstanden sind. Häufig vorkommende Gründe, die zu unterschiedlichen Ergebnissen führen können, ohne dass ein ursächlicher Effekt der Maßnahme vorliegt, sind umgekehrte Kausalitätsbeziehungen, Selektion in Maßnahmen hinein oder heraus und generell Eigenschaften, die sowohl den Erhalt der Maßnahme als auch das Ergebnis beeinflussen. Für diese drei Kategorien seien hier kurze Beispiele angeführt, die jeweils letztlich unbeobachtete Unterschiede in den Erfolgsaussichten der beiden Gruppen widerspiegeln.

Es ist beispielsweise denkbar, dass die Förderung von Energieberatung in besonders energieintensiven Sektoren eingesetzt wird. Dann ist es möglich, dass die geförderten Sektoren auch nach der Maßnahme eine geringere Energieeffizienz aufweisen als nicht geförderte Sektoren – aber nicht, weil die Maßnahme negative Auswirkungen gehabt hätte, sondern weil die geringe Energieeffizienz zur Einführung der Maßnahme geführt hat (umgekehrte Kausalität).

Selbst wenn nicht im Vorhinein Unterschiede in dem Ergebnismaß bestehen, können bewusste Entscheidungen, an einer Maßnahme teilzunehmen, darauf zurückzuführen sein, dass Unternehmen mit Teilnahmeinteresse bereits andere Eigenschaften aufweisen, die das spätere Ergebnis beeinflussen. Beispielsweise mag es zwei ansonsten ähnliche kleine Unternehmen geben, von denen das eine beabsichtigt, in das Exportgeschäft einzutreten, das andere nicht.

Das erste Unternehmen nimmt deshalb an einer geförderten Exportmesse teil. Wenn dieses Unternehmen nach der Maßnahme eine höhere Exporttätigkeit aufweist, so ist dies nicht notwendigerweise die Konsequenz der Maßnahme, sondern spiegelt die Tatsache wider, dass das Unternehmen mit Exportabsicht sich für diese Maßnahme entschieden hat (selektive Teilnahme).

Auch ohne bewusste Selbstselektion können sich teilnehmende Unternehmen von anderen Unternehmen in ihren Erfolgsaussichten unterscheiden. Wenn etwa das Ministerium die Förderung von Ausgründungen aus wissenschaftlichen Einrichtungen auf solche Personen fokussiert, von denen es annimmt, dass sie eine höhere Managementkompetenz aufweisen, dann könnten die geförderten Personen am Ende allein aufgrund ihrer höheren Managementkompetenz bessere Erfolgsindikatoren aufweisen als nicht geförderte Personen. Die Managementkompetenz wäre hier eine unbeobachtete Eigenschaft, die sowohl das Ergebnis als auch die Teilnahme an der Maßnahme bestimmt (unbeobachtete Faktoren).

Ziel der empirisch-kausalanalytischen Wirkungsforschung ist es, aus solchen Zusammenhängen den kausalen Effekt der Politikmaßnahme herauszukristallisieren: Was wäre das Ergebnis des geförderten Unternehmens, wenn es die Förderung nicht erfahren hätte?

In der empirischen Forschung hat sich die Erkenntnis durchgesetzt, dass eine überzeugende Kausalitätsüberprüfung am einfachsten dann möglich ist, wenn der Zufall bestimmt, wer einer bestimmten Maßnahme ausgesetzt wird und wer nicht. Sind Behandlungs- und Vergleichsgruppe groß genug, so stellt die zufällige Aufteilung sicher, dass sich die beiden Gruppen ohne die Maßnahme nicht signifikant voneinander unterscheiden würden. Ein einfacher Vergleich der Ergebnisindikatoren zwischen den beiden Gruppen lässt sich in diesem Fall als die Wirkung der Maßnahme interpretieren. Wie im Folgenden dargelegt lässt sich eine solche Wirkungsanalyse mithilfe direkter randomisierter Feldexperimente und verschiedener quasi-experimenteller Evaluierungsmethoden durchführen und kausal interpretieren. Mittlerweile ist das wissenschaftliche Instrumentarium zur kausalen Evaluierung politischer Maßnahmen ausreichend ausgereift, um in der praktischen Anwendung nutzbringend eingesetzt zu werden.³³

³³ Für methodische Grundlagen und Details vgl. etwa Angrist und Pischke (2009, 2010), Imbens und Wooldridge (2009), Manski (1995) und Stock und Watson (2012). Eine deutschsprachige Einführung liefern Bauer, Fertig und Schmidt (2009). Schlotter, Schwerdt und Woessmann (2011) bieten eine nicht-technische Einführung mit zahlreichen Anwendungsbeispielen aus der Bildungspolitik. Für eine Diskussion der Möglichkeiten und Grenzen experimenteller und quasi-experimenteller Evaluierungsmethoden für die Politikanalyse siehe Heckman (2010), Imbens (2010) und Deaton (2010). Zum Umfeld der Evaluierung wirtschaftspolitischer Maßnahmen und der wissenschaftlichen Politikberatung in Deutschland siehe etwa Schmidt (2009) und Kirchgässner (2013).

3.2.3 Randomisierte Feldexperimente

Die methodisch einfachste Weise, eine zufällige Aufteilung in Behandlungs- und Vergleichsgruppe zu erreichen, besteht darin, per Zufall zu entscheiden, welche Objekte (z.B. Unternehmen) an einer wirtschaftspolitischen Fördermaßnahme teilnehmen und welche nicht. So könnte unter allen Unternehmen, die die Förderkriterien erfüllen, eine randomisierte Zuteilung stattfinden. Bei Aufteilung nach dem Zufallsprinzip lässt sich die Wirkung der Maßnahme – ähnlich wie in der medizinischen Medikamentenentwicklung – durch einen einfachen Vergleich der Ergebnisindikatoren zwischen der Behandlungs- und Vergleichsgruppe untersuchen (vgl. Kasten 1). Auch wenn die Aussagekraft randomisierter Experimente von ihrer jeweiligen konkreten Anwendung abhängt, macht die Zufallsaufteilung sie in Bezug auf die Kausalitätsüberprüfung zu einem Vorbild, an dem sich andere Methoden ausrichten.³⁴

Kasten 1: Randomisierte Feldexperimente

Bei einem randomisierten Feldexperiment werden die Unternehmen – wie bei der experimentellen Untersuchung der Wirksamkeit neuer Medikamente – zufällig in eine „Behandlungsgruppe“, die die „Behandlung“ der wirtschaftspolitischen Fördermaßnahmen erhält, und in eine Vergleichsgruppe, die nicht in den Genuss der Maßnahme kommt, aufgeteilt. Bei ausreichender Größe der Gruppen stellt die zufällige Aufteilung sicher, dass sich die beiden Gruppen nicht systematisch unterscheiden. Deshalb kann nach Durchführung der Maßnahme der durchschnittliche Unterschied in den Ergebnisindikatoren zwischen den beiden Gruppen direkt auf die kausale Auswirkung der Maßnahme zurückgeführt werden.

Beispielsweise untersucht die Studie von Fairlie, Karlan und Zinman (2012) das in Abschnitt 2.4 angesprochene, als randomisiertes Feldexperiment angelegte Projekt „Growing America through Entrepreneurship“ (Project GATE). In dem zwischen 2003 und 2005 in sieben Städten durchgeführten Projekt kam ein per Zufall ausgewählter Teil der über 4.000 Interessenten in den Genuss einer kostenlosen Beratungsmaßnahme für Unternehmensgründungen. Die Interessenten wurden vorab darüber informiert, dass GATE nicht für alle Interessenten Plätze hat und daher eine Lotterie oder Zufallsziehung über die Teilnahme entscheidet. Bei einem

³⁴ Grundlagen und weiterführende Diskussionen zum Einsatz von Feldexperimenten in der Wirtschaftsforschung bieten etwa Harrison und List (2004), List (2011) und Ludwig, Kling und Mullainathan (2011). Eine politikorientierte Einführung bieten Haynes u.a. (2012). Für eine anwendungsorientierte Diskussion mit Fokus auf die Bildungspolitik vgl. Bouguen und Gurgand (2012). Für eine kurze Einführung mit Fokus auf innovationspolitische Maßnahmen vgl. Expertenkommission Forschung und Innovation (2013), S. 39-45. Zum Einsatz von Experimenten in der Rechtswissenschaft vgl. Engel (2013).

Vergleich zwischen geförderten und nicht geförderten Individuen, die 6, 18 und 60 Monate nach Start des Experiments befragt wurden, zeigen sich außer bei der kurzfristigen Selbständigkeit von zuvor Arbeitslosen keine signifikanten Unterschiede für längere Zeithorizonte, andere Zielgruppen und andere Ergebnisindikatoren.³⁵

Generell bietet sich gerade bei Maßnahmen, bei denen die Nachfrage förderungswürdiger Unternehmen die verfügbaren Mittel übersteigen, der Einsatz einer Lotterie an, die ebenfalls eine zufällige Aufteilung sicherstellt und so ein randomisiertes Forschungsdesign sicherstellt. Eine Grenze der Verallgemeinerung von experimentellen Untersuchungen besteht darin, dass sich die Teilnehmer möglicherweise anders als sonst verhalten, wenn sie wissen, dass ihr Verhalten oder ihre Ergebnisse beobachtet werden (sogenannter Hawthorne-Effekt).

Gleichwohl hat auch die Wirksamkeitsanalyse mithilfe randomisierter Feldexperimente Grenzen. Zunächst einmal ist die Untersuchung auf den Effekt der jeweils per Zufall zugewiesenen „Behandlung“, die die Behandlungs- von der Vergleichsgruppe unterscheidet – also etwa die Teilnahme an einer Fördermaßnahme –, beschränkt und lässt sich für gewöhnlich nicht auf Effekte anderer Einflüsse ausweiten. Auch muss jeweils im Einzelfall abgewogen werden, inwieweit sich die im konkreten experimentellen Anwendungsfall erlangten Erkenntnisse auf andere relevante Anwendungsfälle verallgemeinern lassen.³⁶ Schließlich können sich die Ergebnisse einer Fördermaßnahme unterscheiden, je nachdem, ob sie in einzelnen Firmen oder in der breiten Fläche umgesetzt wird. In diesem Fall können in eher geringem Umfang umgesetzte Experimente die Rückwirkungen auf den Gesamtmarkt (allgemeine Gleichgewichtseffekte) nicht abbilden.

In besonderer Weise bietet sich die explizite randomisierte Zuteilung bei Pilotprojekten an, bei denen nicht alle Bewerber, die die Fördervoraussetzungen formal erfüllen, die begrenzten Mittel erhalten können.³⁷ Generell sollte eine größere Zahl von Antragstellern gegeben sein, damit die Wirksamkeitsanalyse statistisch mit ausreichender Sicherheit durchgeführt werden

³⁵ Als weitere Beispiele für die Anwendung von Feldexperimenten untersuchen Schwerdt u.a. (2012) die Wirksamkeit von Gutscheinen (im Wert von bis zu 1500 Schweizer Franken) für Erwachsenenbildungsmaßnahmen in der Schweiz und Rosendahl Huber, Sloof und Van Praag (2012) ein Schülerfirmenprogramm in den Niederlanden.

³⁶ Die „interne Validität“ – also die Tatsache, dass die Kausalität im konkreten Anwendungsfall gegeben ist – ist bei randomisierten Feldexperimenten sehr hoch. Gleichzeitig kann die „externe Validität“ – also die Möglichkeit, den Befund auf weitere Anwendungsfälle zu verallgemeinern – begrenzt sein.

³⁷ Arni (2012) diskutiert die kausale Evaluierung von Pilotprojekten in der Praxis und demonstriert die Einsatzmöglichkeit anhand einer Reihe europäischer Beispiele. Siehe auch Expertenkommission Forschung und Innovation (2013), S. 39-45.

kann. Bei solchen überzeichneten Förderprogrammen mit begrenztem Budget kann im Rahmen eines Losverfahrens der Zufall entscheiden, welche der gleichermaßen qualifizierten Bewerber die Förderung im Pilotprojekt tatsächlich erhalten. Im Vergleich zu bisher angewandten Zuweisungsverfahren, die aufgrund der beschränkten Fördermittel ebenfalls üblicherweise nicht alle geeigneten Antragsteller fördern konnten, entstehen hier keine zusätzlichen „Verlierer“. Im Gegenteil kann die zufallsgesteuerte Vergabe unter den förderberechtigten Antragstellern als die fairste Methode angesehen werden.

Die Information über die Randomisierung der Vergabe der Fördermittel sollte in die Ausschreibung der Fördermaßnahme aufgenommen werden. Unternehmen sollten sich nur dann um Fördermittel bewerben können, wenn sie sich ausdrücklich mit dem Vergabeverfahren einverstanden erklären. Damit könnte Konkurrentenklagen abgewiesener Unternehmen vorgebeugt werden. Vielmehr kann die Ausschreibung die Teilnahme am Verfahren zugleich von der Selbstverpflichtung abhängig machen, die für die Evaluierung erforderlichen Daten zu erheben und zu übermitteln – und zwar gerade unabhängig davon, ob ein Unternehmen bei der Vergabe zum Zuge kommt.³⁸ Auf diese Weise erhält das Ministerium auch für die Kontrollgruppe verlässliche Daten. Die Kontrollgruppe und die Gruppe der geförderten Unternehmen müssen im Übrigen nicht gleich groß sein, aber groß genug, damit statistische Verfahren Anwendung finden können.

Soll keiner der Antragsteller ohne Förderung bleiben, so ist zumindest bei relativ kurzfristigen Ergebnisindikatoren auch eine Analyse im „Phase-in Design“ möglich, in dem zunächst nur eine zufällig ausgewählte Teilgruppe der Antragsteller die Förderung erhält und die andere Teilgruppe die Förderung zum Beispiel mit einem Jahr Verspätung erhält. In diesem Fall kann die zweite Gruppe im ersten Jahr als Vergleichsgruppe für die erste Gruppe fungieren.³⁹

Ein weiteres Anwendungsbeispiel für randomisierte Feldexperimente besteht darin, die Wirksamkeit verschiedener Fördermaßnahmen und alternativer Ausgestaltungen spezifischer Förderinstrumente gegeneinander zu testen, um so aus verschiedenen Alternativen die wirksamste auswählen zu können. Eine solche Analyse wäre gegebenenfalls sogar ohne Berücksichtigung einer Kontrollgruppe, die gar keine Förderung erfährt, möglich, wobei dann allerdings die Wirksamkeit der Maßnahmen insgesamt nicht überprüft werden kann, sondern nur die relative Wirksamkeit der alternativen Maßnahmen. Auch bei diesem Design müssen die

³⁸ Eine solche Selbstverpflichtung kann Unternehmen davon abhalten, sich auf eine solche Fördermaßnahme zu bewerben, was die Verallgemeinerungsmöglichkeit der Aussage (externe Validität) einschränken kann.

³⁹ Ganz unproblematisch ist dieses Design allerdings nicht, da sich die zweite Gruppe gegebenenfalls bewusst ist, dass sie im folgenden Jahr auch gefördert werden wird. Es liegt nahe, dass sie sich mit ihrem Verhalten darauf einstellt, wodurch der Effekt der Maßnahme bei diesem Design tendenziell überschätzt werden dürfte.

verschiedenen Ausgestaltungen den Unternehmen zufällig zugewiesen werden. Die Unternehmen selbst dürfen keinen Einfluss auf die Auswahl der konkreten Maßnahme haben.

Vor dem Hintergrund der vielversprechenden Erkenntnismöglichkeiten empfiehlt die von der Bundesregierung eingerichtete Expertenkommission Forschung und Innovation, randomisierte Experimente „als eines der Standardinstrumente in das Evaluationsportfolio staatlicher FuE-Förderung aufzunehmen“.⁴⁰ Konkret sieht die Expertenkommission gute Anwendungsfelder randomisierter Evaluierungsverfahren etwa in der Mittelstands- und Gründerforschung, so zum Beispiel beim ZIM-Programm, bei KMU-innovativ und beim EXIST-Gründerstipendium (vgl. Abschnitt 2.2). Der Beirat schließt sich diesen Empfehlungen an und regt an, diese auf weitere wirtschaftspolitische Handlungsfelder auszuweiten.

3.2.4 Quasi-experimentelle Evaluierungsmethoden

Auch jenseits der kontrollierten Randomisierung ist eine überzeugende Evaluierung möglich. Dazu haben sich in der evidenzbasierten Politikberatung sogenannte quasi-experimentelle Methoden etabliert, die unter bestimmten Umständen eine Wirkungsanalyse auch auf der Basis von nicht randomisiert erhobenen Daten ermöglichen. Zentraler Punkt dieser Methoden ist die überzeugende Definition einer Vergleichsgruppe, die aufzeigt, welche Ergebnisse (beispielsweise) geförderte Unternehmen ohne die Förderung erzielt hätten. Es geht also darum, mit Daten, die nicht im Rahmen einer explizit zufälligen Aufteilung erhoben wurden, möglichst nah an eine solche zufällige Aufteilung in Behandlungs- und Vergleichsgruppe heranzukommen. In dem Maße, wie etwa natürliche Gegebenheiten, institutionelle Regelungen oder politische Entscheidungen zu einer „so gut wie zufälligen“ Aufteilung führen, kann hier von „natürlichen“ oder „Quasi“-Experimenten gesprochen werden.

Die gängigsten dieser quasi-experimentellen Evaluierungsmethoden sind der Vergleich von Behandlungs- und Vergleichsgruppe vor und nach Einführung einer Maßnahme (sogenannter Differenzen-in-Differenzen-Ansatz); der Vergleich von Objekten ober- und unterhalb eines Schwellenwerts für den Zugang zu einer Fördermaßnahme (sogenannter Regressions-Diskontinuitäten-Ansatz); und die Identifikation exogener Politikvariation durch Drittvariablen (sogenannter Instrumentvariablen-Ansatz).⁴¹ Diese Methoden werden in den Kästen 2-4 kurz beschrieben und anhand von Beispielen verdeutlicht. Die Darstellungen können allerdings nur

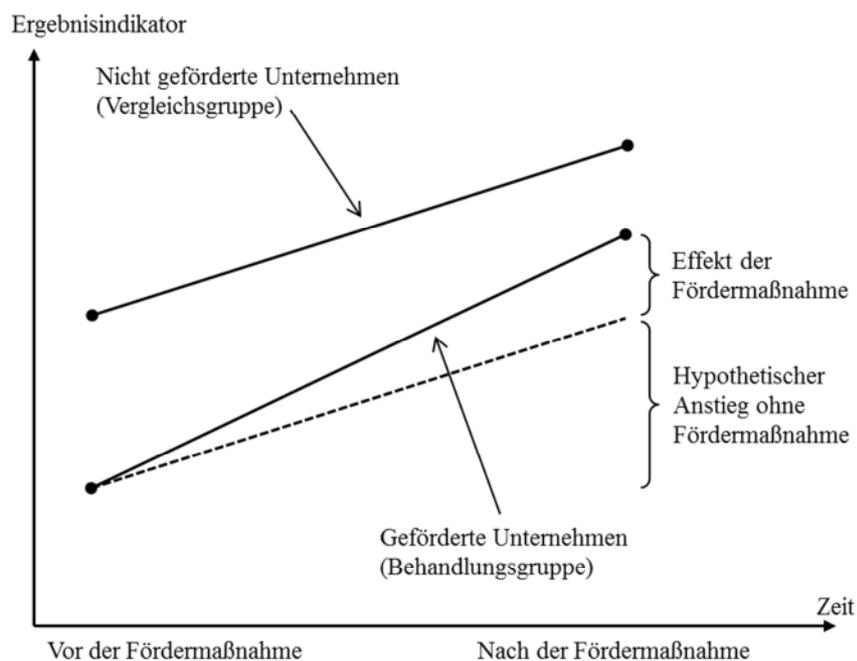
⁴⁰ Expertenkommission Forschung und Innovation (2013), S. 45.

⁴¹ Neben den Literaturverweisen in Abschnitt 3.2.2 siehe etwa Besley und Case (2000) und Bertrand, Duflo und Mullainathan (2004) zum Differenzen-in-Differenzen-Ansatz, Lee und Lemieux (2010) zum Regressions-Diskontinuitäten-Ansatz und Angrist und Krueger (2001) zum Instrumentvariablen-Ansatz.

die generelle Intuition der Methoden vermitteln; die konkrete Ausgestaltung ist auf den jeweiligen Anwendungsfall auszurichten.

Kasten 2: Differenzen-in-Differenzen-Ansatz

Ein in der Politikevaluierung häufig verwendetes Verfahren ist der sogenannte Differenzen-in-Differenzen-Ansatz. Dabei werden die Ergebnisindikatoren sowohl für geförderte als auch für nicht geförderte Objekte nicht nur nach der Maßnahme, sondern auch bereits vor der Maßnahme erhoben. Die Wirkung der Maßnahme ergibt sich nun durch einen Vergleich der *Veränderung* der Ergebnisindikatoren zwischen den beiden Gruppen (vgl. Abbildung): Wie unterscheidet sich die Veränderung der Ergebnisindikatoren (erste Differenz, zwischen den beiden Zeitpunkten) zwischen denjenigen Objekten, die zwischen den beiden Zeitpunkten an einer Maßnahme teilgenommen haben, und denjenigen Objekten, die nicht teilgenommen haben (zweite Differenz, zwischen den beiden Gruppen)? Wenn die Annahme zu rechtfertigen ist, dass sich die beiden Gruppen ohne die Maßnahme parallel entwickelt hätten – dass also andere relevante unbeobachtete Merkmale sich nicht systematisch über die Zeit verändern –, gibt der Differenzen-in-Differenzen-Ansatz den ursächlichen Effekt der Maßnahme wieder. Dies ist auch dann der Fall, wenn sich die beiden Gruppen (etwa aufgrund von nicht beobachteten Merkmalen) im Ausgangsniveau unterscheiden.



Beispielsweise nutzen Falck, Heblich und Kipar (2010) den Differenzen-in-Differenzen-Ansatz, um die Auswirkung einer bayerischen Politikmaßnahme, die die Bildung von Clustern – regionaler Kooperationen von Unternehmen mit einer Vielzahl von Akteuren aus dem Unternehmens-, Wissenschafts- und Finanzsektor – fördert, auf die Innovationstätigkeit der Unternehmen zu untersuchen. Bayern hat 1999 die „High-Tech-Offensive“ eingeführt, die Cluster in fünf Hochtechnologiefeldern fördert. Die Studie vergleicht, wie sich die Innovationstätigkeit von Unternehmen in diesen Technologiefeldern innerhalb und außerhalb Bayerns von vor bis nach der Einführung der Politikmaßnahme entwickelt hat. Es zeigt sich, dass bayerische und nicht-bayerische Unternehmen in den fünf Technologiefeldern vor Einführung der Maßnahme ein ähnliches Niveau an Innovationstätigkeit aufwiesen, dass die bayerischen Unternehmen nach der Einführung aber häufiger innovieren, was auf einen positiven Effekt der Politikmaßnahme hindeutet.⁴²

Veränderungen in anderen beobachteten Merkmalen können im Rahmen des Differenzen-in-Differenzen-Ansatzes herausgerechnet werden. Der Ansatz lässt sich auch auf viele Perioden vor und nach der Einführung einer Maßnahme sowie auf Maßnahmen, die zu verschiedenen Zeitpunkten eingeführt werden, verallgemeinern (Panelmethoden mit fixen Effekten).⁴³ Auch lässt sich der Ansatz in bestimmten Fällen auf Unterschiede über andere Dimensionen als die Zeit anwenden, etwa beim Vergleich zwischen mehreren Bereichen eines Unternehmens.

Eine weitere Evaluierungsmethode ist der sogenannte Matching-Ansatz, der darauf abzielt, „Paare“ in Behandlungs- und Vergleichsgruppe zu finden, die sich in Bezug auf beobachtete Merkmale so ähnlich wie möglich sind. Entscheidend dabei ist, dass das Matching-Verfahren beobachtete Werte der Ergebnisindikatoren vor Einführung der Maßnahme berücksichtigt.⁴⁴

Kasten 3: Regressions-Diskontinuitäten-Ansatz

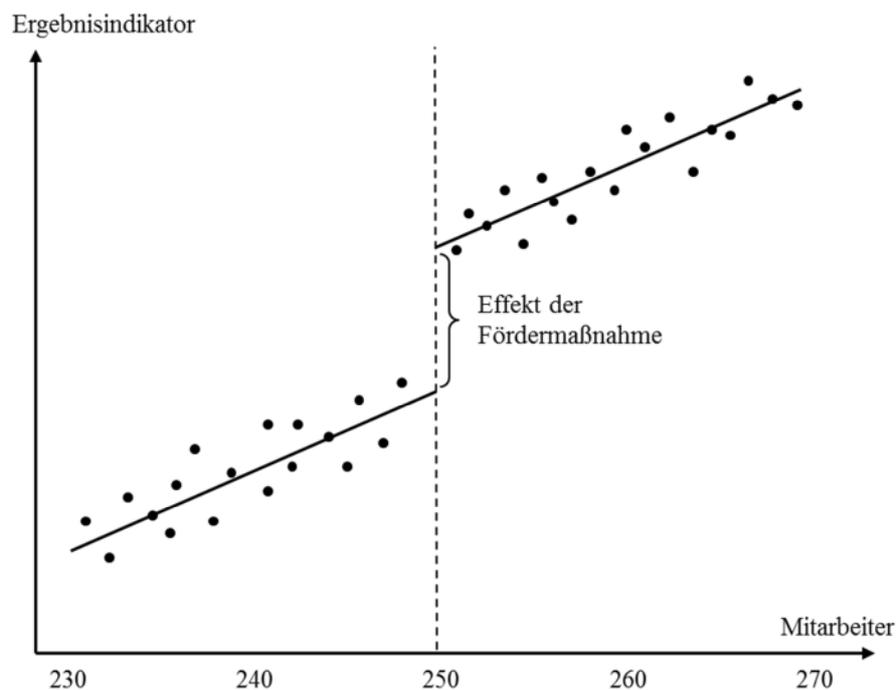
Wenn die Teilnahmemöglichkeit an einer Fördermaßnahme daran geknüpft ist, dass bestimmte Eigenschaften eines Unternehmens einen vorgegebenen Schwellenwert unter- oder überschreiten – wenn beispielsweise eine Forschungsförderung daran geknüpft ist, dass ein Unter-

⁴² Die Studie erweitert die Analyse noch auf einen „dreifachen Differenzen-Ansatz“, der zusätzlich zur regionalen und zeitlichen Dimension noch Unternehmen in förderfähigen und nicht-förderfähigen Technologiefeldern vergleicht.

⁴³ Beispielsweise nutzen Felbermayr und Yalcin (2013) solche Panelmethoden mit fixen Effekten, um Effekte der Hermes-Bürgschaften auf die Exporttätigkeit zu evaluieren.

⁴⁴ Beispielsweise nutzen Lechner, Miquel und Wunsch (2011) und Biewen u.a. (2013) solche Matching-Verfahren, um öffentliche Weiterbildungsprogramme zu evaluieren.

nehmen weniger als 250 Beschäftigte hat oder dass eine Person eine bestimmte Altersgrenze unter- oder überschreitet –, dann bietet sich häufig die Nutzung des sogenannten Regressions-Diskontinuitäten-Ansatzes an. Dabei werden letztlich die Unternehmen just unter und just über dem Schwellenwert miteinander verglichen. Die zugrunde liegende Annahme besteht darin, dass sich zum Beispiel Unternehmen mit 240-249 Mitarbeitern kaum von Unternehmen mit 250-259 Mitarbeitern unterscheiden (bzw. genau genommen, dass der Zusammenhang zwischen Mitarbeiterzahl und Ergebnisindikator in diesem Intervall stetig ist) – außer eben im Zugang zu der Fördermaßnahme. Diese Annahme lässt sich etwa daran testen, dass vor Einführung der Maßnahme keine beobachtbaren Unterschiede zwischen den beiden Gruppen bestanden haben. Solange es keine anderen Sprünge („Diskontinuitäten“) an dieser Stelle gibt, erlaubt der Ansatz, den kausalen Effekt der Maßnahme wiederzugeben: Wenn die Ergebnisindikatoren an dieser Stelle einen Sprung aufweisen, ist dies auf die Maßnahme zurückzuführen (vgl. Abbildung).



Beispielsweise untersuchen Becker, Egger und von Ehrlich (2010, 2013) mit dem Regressions-Diskontinuitäten-Ansatz die Auswirkungen der EU-Strukturfonds auf die regionale Entwicklung.⁴⁵ Die Regel, dass die auf Konvergenz zielenden Fördermaßnahmen solchen („Ziel-1“-) Regionen vorbehalten sind, deren Bruttoinlandsprodukt pro Kopf unter 75% des europäischen Durchschnitts liegen, führt an dieser Schwelle zu einem abrupten Sprung in der

⁴⁵ In Deutschland ist das BMWi verantwortlich für die Koordinierung der EU-Strukturpolitik (<http://www.bmwi.de/DE/Themen/Europa/eu-strukturpolitik.html> (30.6.2013)).

Förderwahrscheinlichkeit. Es zeigt sich, dass das Wirtschaftswachstum an dieser Schwelle ebenfalls einen Sprung macht, nicht aber das Beschäftigungswachstum, was auf Wachstums-, nicht aber Beschäftigungseffekte hindeutet. Allerdings sind die positiven Effekte auf rund ein Viertel der geförderten Regionen beschränkt, die genügend Humankapital und genügend gute Institutionen aufweisen.⁴⁶

Wichtig beim Regressions-Diskontinuitäten-Ansatz ist, dass die Förderempfänger nicht in der Lage sind, die Variable, die den Schwellenwert bestimmt, präzise zu manipulieren (nicht-präzise Manipulierbarkeit ist generell kein Problem, da dann die Variation in der Nähe des Schwellenwertes wiederum als zufällig angesehen werden kann). Eine solche Manipulationsmöglichkeit ist etwa dann nicht gegeben, wenn das Datum, auf das sich der Schwellenwert bezieht, in der Vergangenheit liegt und der Schwellenwert nicht zuvor bekannt war.

Kasten 4: Instrumentvariablen-Ansatz

Der ursächliche Effekt einer Maßnahme auf eine Ergebnisvariable lässt sich auch dann von anderen Ursachen eines Zusammenhangs zwischen Maßnahme und Ergebnis trennen, wenn es eine dritte Variable gibt, die zwar die Wahrscheinlichkeit, an der Maßnahme teilzunehmen, beeinflusst, ansonsten aber nicht mit der Ergebnisvariable zusammenhängt. In diesem Fall erlaubt diese dritte Variable (die sogenannte „Instrumentvariable“), die Teilnahme an der Maßnahme zu erklären. Die Analyse nutzt dann nur den Teil der Teilnahmewahrscheinlichkeit, der mit der Drittvariable erklärt werden kann und daher nicht aus anderen Gründen zustande gekommen ist. Zeigt dieser Teil des Unterschieds in der Maßnahmenteilnahme einen Zusammenhang mit der Ergebnisvariable, so ist dies auf einen ursächlichen Effekt der Maßnahme zurückzuführen.

Beispielsweise nutzen Criscuolo u.a. (2012) den Instrumentvariablen-Ansatz, um Effekte des britischen Investitionsförderungsprogramms „Regional Selective Assistance“ zu untersuchen. Die Studie nutzt nur solche Änderungen in der Teilnahmemöglichkeit an dieser britischen Förderpolitik (die zu untersuchende Maßnahme), die aufgrund von regionalspezifischen Änderungen in den Ausnahmeregelungen der EU für staatliche Beihilfen (die Instrumentvariable) zustande gekommen sind. Für Firmen mit mehr als 150 Beschäftigten findet sie

⁴⁶ Lalive (2008), Schmieder, von Wachter und Bender (2012) und Caliendo, Tatsiramos und Uhlendorff (2013) bieten Beispiele für die Anwendung des Regressions-Diskontinuitäten-Ansatzes auf Wirkungen der Bezugsdauer von Arbeitslosengeld in Österreich und Deutschland.

keine Effekte der Fördermaßnahme, für kleine Unternehmen hingegen positive Effekte auf Investitionen und Beschäftigung, nicht aber auf Produktivität.⁴⁷

Der Instrumentvariablen-Ansatz kann sehr unterschiedliche Anwendungsfelder haben, etwa durch die Nutzung von Politikänderungen, rein „natürlicher“ Variationen oder auch Unterschieden aufgrund historischer Zusammenhänge. Die zentrale Herausforderung beim Instrumentvariablen-Ansatz liegt darin, eine überzeugende dritte Variable zu finden, die einerseits nah genug mit der Maßnahmenteilnahme verbunden ist, andererseits aber keinen sonstigen Zusammenhang mit der Ergebnisvariable aufweist.

Die wichtigste Herausforderung bei der Wahl einer überzeugenden quasi-experimentellen Evaluierungsmethode besteht darin, plausibel zu machen, dass die jeweilige Aufteilung in Behandlungs- und Vergleichsgruppe tatsächlich „so gut wie zufällig“ ist. Dies ist im Vergleich zu kontrolliert randomisierten Feldexperimenten eine größere Herausforderung. Jedoch haben quasi-experimentelle Evaluierungsmethoden die Vorteile, dass sich die untersuchten Objekte nicht der Teilnahme an einer experimentellen Untersuchung bewusst sind und dementsprechend ihr Verhalten anpassen, dass in bestimmten Anwendungsfällen auch Rückwirkungen auf den Gesamtmarkt abgebildet werden können und dass sie einige der Herausforderungen der Implementierung expliziter Experimente in der Praxis umgehen.

Eine konkrete Anwendungsmöglichkeit des Regressions-Diskontinuitäten-Ansatzes bei der Evaluierung wirtschaftspolitischer Fördermaßnahmen besteht dann, wenn die förderberechtigten Antragsteller auf einer Liste hierarchisiert werden und dann vom am meisten priorisierten Unternehmen angefangen so viele Unternehmen gefördert werden, bis die Mittel verausgabt sind. In diesem Fall lässt sich der Effekt der Fördermaßnahme an dem Sprung der interessierenden Ergebnisindikatoren an der Stelle zwischen den geförderten und den nicht geförderten Antragstellern ablesen (wobei lineare Effekte des Kriteriums der Priorisierung wiederum herausgerechnet werden können).⁴⁸

Letztlich ist die konkrete Anwendbarkeit und Ausgestaltung der quasi-experimentellen Evaluierungsmethoden für jeden Anwendungsfall spezifisch und muss daher im Einzelfall

⁴⁷ Als weiteres Beispiel nutzen Frölich und Lechner (2010) regionale Umsetzungsintensitäten als Instrumentvariable zur Evaluierung von Maßnahmen der aktiven Arbeitsmarktpolitik in der Schweiz. Aichele und Felbermayr (2012) nutzen nationale Ratifizierungen des Internationalen Strafgerichtshofs als Instrumentvariable, um den Effekt der Ratifizierung des Kyoto-Protokolls auf die CO₂-Bilanz zu evaluieren.

⁴⁸ Für ein Anwendungsbeispiel aus dem Bereich der Evaluierung von Effekten der Wagniskapitalfinanzierung durch Business Angels siehe Kerr, Lerner und Schoar (2013).

ausgearbeitet werden. So bedarf es für den Einsatz des Regressions-Diskontinuitäten-Ansatzes etwa des Vorhandenseins von Schwellenwerten beim Zugang zu einer Maßnahme. Beim Differenzen-in-Differenzen-Ansatz ist es etwa von zentraler Bedeutung, geförderte und nicht geförderte Objekte schon vor Einführung der Maßnahme zu beobachten.

Eine überzeugende Kausalitätsanalyse kann häufig auch über die Kombination mehrerer Methoden erreicht werden. So kann etwa bei der Analyse eines randomisierten Feldexperimentes, bei dem Unternehmen zwischen Behandlungs- und Vergleichsgruppe wechseln können, die ursprüngliche Zuteilung als Instrumentvariable genutzt werden. Ebenso findet die Instrumentvariablenschätzung zumeist bei der Umsetzung des Regressions-Diskontinuitäten-Ansatzes Anwendung, indem der Schwellenwert als Instrumentvariable für die Förderteilnahme verwendet wird. Auch kann der herkömmliche Differenzen-in-Differenzen-Ansatz durch den Instrumentvariablen-Ansatz ergänzt werden, um die Annahme der parallelen Entwicklung ohne Maßnahme stichhaltiger zu machen.

Aus dem Kontinuum an experimentellen und quasi-experimentellen Evaluierungsmethoden wird jeweils im einzelnen Anwendungsfall zu entscheiden sein, welche Methode(n) praktikabel und erfolgversprechend eingesetzt werden können. Insgesamt muss bei der Wahl der am besten passenden Evaluierungsmethodik häufig abgewogen werden zwischen einer abschließenden Stringenz in der Überprüfung der Ursächlichkeit und der möglichst einschlägigen Relevanz des Anwendungsfalls für die jeweilige wirtschaftspolitische Fragestellung (Tradeoff zwischen interner und externer Validität). Nicht zuletzt deshalb kann es hilfreiche Hinweise für die Wirtschaftspolitik liefern, wenn mit alternativen Instrumenten konvergierende Ergebnisse erzielt werden.

3.3 Grenzen der Evaluierung

Die empirische Evaluierung wirtschaftspolitischer Maßnahmen hat Grenzen. Insbesondere die externe Validität, das heißt die Übertragbarkeit der gefundenen Ergebnisse, muss im Einzelfall kritisch hinterfragt werden. So lassen Evaluierungen einzelner Maßnahmen zumeist nur Partialaussagen für die Einführung im jeweils gegebenen Rahmen zu. Aufgrund von Anpassungsmaßnahmen auf anderen Märkten (allgemeine Gleichgewichtseffekte) können sich diese durchaus von den Auswirkungen bei flächendeckender Einführung unterscheiden. Auch ist es möglich, dass Lerneffekte von geförderten Unternehmen auf nicht geförderte Unternehmen überspringen (Spillover-Effekte) oder dass in einem engen Markt nicht geförderte Unternehmen verdrängt werden. In diesen Fällen ist es wichtig, nicht nur Effekte auf einzelne

Unternehmen, sondern nach Möglichkeit auch auf Industrien oder Regionen insgesamt zu untersuchen.

Evaluierungsstudien beziehen sich zumeist auf kurz- bis mittelfristige Auswirkungen einer Fördermaßnahme, die nicht unmittelbar Urteile über langfristige Wirkungen zulassen. In bestimmten Anwendungsfeldern wie etwa in der Luftfahrtförderung oder bei Großtechnologieprojekten kann es sich aufgrund der geringen Anzahl von einschlägigen Unternehmen auch als schwierig erweisen, überzeugende Vergleichsgruppen für empirische Analysen zu finden.

Generell können einzelne empirische Studien keine „letzten Antworten“ liefern, sondern sind auf den konkreten Anwendungsfall mit den entsprechenden Randbedingungen bezogen. Metastudien können hier helfen, allgemeinere Aussagen zu ermöglichen.

Bei einigen Fördermaßnahmen kann die empirische Forschung auch an rechtliche Grenzen stoßen. Die Evaluierung einer Fördermaßnahme mag in bestimmten Fällen die Auswertung persönlicher Daten von Individuen oder den Einblick in interne Geschäftsunterlagen von Unternehmen erfordern. Den berechtigten Interessen an Datenschutz und Geschäftsgeheimnissen kann zumeist aber durch die Anonymisierung von Daten, durch Vertraulichkeitsvereinbarungen bei der Datenanalyse sowie durch aggregierte Darstellung der Analyseergebnisse Rechnung getragen werden (siehe auch Abschnitt 4.1.2).

Auch der Gleichbehandlungsgrundsatz wird einer empirisch-kausalanalytischen Wirkungsforschung von Fördermaßnahmen normalerweise nicht entgegenstehen. Zwar darf der Staat Fördermaßnahmen nicht willkürlich vergeben. Bei einem randomisierten Feldexperiment kann der Staat jedoch objektive Förderkriterien benennen. Das ist offensichtlich, wenn die Anzahl der Unternehmen, die diese Förderkriterien erfüllt, den Umfang der Fördermaßnahme übersteigt. Wenn der Staat aus dieser Gruppe zufällig Unternehmen auswählt, die durch die Maßnahme gefördert werden, verwirklicht das sogar den Gleichheitssatz. Denn die zufällige Auswahl bei objektiv beschränkter Menge verhindert, dass (versteckt oder offen) sachwidrige Kriterien in die Auswahlentscheidung einfließen.

Auch wenn es solche objektiven Grenzen nicht gibt, bezweckt der Staat mit der Randomisierung keine zufällig induzierte Ungleichbehandlung von Unternehmen, sondern will Informationen über die Effektivität und Effizienz von Fördermaßnahmen gewinnen. Verfassungsrechtlich genügt das zur Rechtfertigung der unvermeidlichen Ungleichbehandlung. Eine verfassungsrechtlich relevante Grenze wäre erst dann erreicht, wenn der Staat das Ausmaß der Fördermaßnahme künstlich beschränkt, um randomisieren zu können. Regelmäßig wird diese Grenze bei Fördermaßnahmen schon deshalb nicht überschritten sein, weil auf die Förderung

kein gesetzlicher Anspruch besteht. Bei der Auswahl der Adressaten von Fördermaßnahmen, auf die dieses Gutachten beschränkt ist, stellt das Grundgesetz die Verwaltung besonders frei. Sie darf nur keine Willkür walten lassen. Bei quasi-experimentellen Methoden tritt dieser Aspekt ohnehin in den Hintergrund, wird die Evaluierungsmethode doch meistens keine zufällige Zuweisung von Unternehmen zu Behandlungs- und Vergleichsgruppen durch den Staat erfordern.

Solange der Wunsch nach einer Verbesserung der Wirksamkeit und Treffsicherheit staatlicher Fördermaßnahmen nicht nur vorgeschoben ist, sprechen gegen die Randomisierung auch keine durchgreifenden ethischen Einwände. Der Staat verlangt den Unternehmen, die nicht ausgewählt werden, ja kein dauerhaftes Opfer ab. Die Fördermaßnahme bleibt ihnen nur für die Dauer eines zeitlich begrenzten staatlichen Lernvorgangs vorenthalten; die gleiche Chance auf Förderung hatten sie ohnehin. Dem steht der Vorzug gegenüber, dass die von allen aufgebrachtten Steuergelder die wirtschaftspolitisch intendierte Wirkung auch wirklich erreichen.

Schließlich ist die Frage der Wirtschaftlichkeit der Evaluierung selbst zu bedenken. Durch eine Evaluierung entstehen Kosten, sowohl direkte Kosten der Durchführung als auch indirekte Bürokratiebelastungen bei den Evaluierten. Auch für die Evaluierung selbst gilt also, sie so wirksam und kostengünstig wie möglich einzusetzen. Gleichwohl können die Mittel für die Evaluierung gut eingesetzt sein, wenn durch sie ineffektive Fördermaßnahmen ausgesetzt und zukünftige Politikmaßnahmen wirkungsvoller ausgestaltet werden können.

3.4 Evaluierungen von einzelnen Fördermaßnahmen sowie Maßnahmenbereichen

Letztlich lassen sich in allen Politikbereichen des BMWi – der Innovations- und Mittelstandsförderung, der Energie- und Exportförderung – zahlreiche Fördermaßnahmen finden, die auch in Anbetracht dieser Grenzen einer überzeugenden Evaluierung mit experimentellen oder quasi-experimentellen Methoden zugänglich sind. Die Evaluierung wirtschaftspolitischer Fördermaßnahmen sollte sowohl die regelmäßige Evaluierung einzelner Fördermaßnahmen als auch in unregelmäßigen Abständen die Durchführung von Evaluierungspaketen ganzer Maßnahmenbündel umfassen. Vor dem Hintergrund der Vielzahl der gezielten Maßnahmen im BMWi-Förderportfolio können beide Ansätze komplementär zu einer Verbesserung der Wirksamkeit wirtschaftspolitischer Fördermaßnahmen beitragen.

Im Rahmen der Einzelevaluierungen findet eine regelmäßige begleitende Evaluierung einzelner Fördermaßnahmen statt. Jede Fördermaßnahme ab einer festzusetzenden Größenordnung sollte soweit möglich standardmäßig einer empirisch-kausalen Wirkungsanalyse unterzogen werden. Im Verlaufe der Maßnahme sollte diese durch qualitative Interviews und Fallstudien

ergänzt werden. Durch eine solche Kombination aus quantitativen und qualitativen Methoden kann häufig ein vertieftes Verständnis der Wirkungsmechanismen entstehen. Die empirisch-kausale Wirkungsanalyse sollte durchgeführt werden, sobald die festgelegten Ergebnisindikatoren erhoben werden können, so dass die Evaluierungsergebnisse in die Weiterentwicklung der Maßnahme einfließen können. Dabei sollte der ermittelte Effekt einer Maßnahme im Rahmen einer Kosten-Nutzen-Analyse auch ihren Kosten gegenübergestellt werden. Auch nach Abschluss der Maßnahme kann die Evaluierung dieser Maßnahme für die Gestaltung zukünftiger Fördermaßnahmen relevantes Wissen hervorbringen.

Darüber hinaus sollte von Zeit zu Zeit im Rahmen von Gesamtevaluierungen ganzer Maßnahmenbereiche eine vergleichende Evaluierung aller Maßnahmen in einem Politikfeld durchgeführt werden. Gerade bei der vergleichenden Evaluierung ganzer Bereiche sollte jenseits der reinen Wirksamkeitsanalyse für die einzelnen Maßnahmen eine Nutzen und Kosten vergleichende Effizienzanalyse durchgeführt werden, um bestimmen zu können, welche Maßnahmen die jeweiligen Ziele zu den geringstmöglichen Kosten erreichen. Solche Evaluierungspakete helfen bei der Entscheidung, mit welchen Maßnahmenbündeln in Zukunft die Förderung etwa von Innovationstätigkeit, Mittelstand, Energie oder Exporten bestmöglich ausgestaltet werden sollte. Erfahrungen von Evaluierungspaketen aus der Arbeitsmarkt- oder Familienpolitik verdeutlichen die großen Erkenntnispotentiale solcher vergleichender Evaluierungen ganzer Maßnahmenbereiche.

4. Einsatz in der wirtschaftspolitischen Praxis

4.1 Datenerhebung und -zugang

Eine Grundvoraussetzung für überzeugende Evaluierungen ist die Verfügbarkeit der für die Evaluierung benötigten Daten. Sowohl die internationalen Erfahrungen mit der Evaluierung wirtschaftspolitischer Fördermaßnahmen als auch die nationalen Erfahrungen mit der Evaluierung arbeitsmarktpolitischer Maßnahmen belegen, dass die Qualität der Evaluierung typischerweise spürbar zunimmt, sobald umfangreiche Mikrodaten verfügbar werden.

4.1.1 Existierende Datensätze und Eigenerhebungen von Evaluierungsdaten

Essentiell für eine erfolgreiche und kosteneffiziente Evaluierung ist es, dass schon vor der Einführung einer Fördermaßnahme bedacht wird, welche Daten für eine Zielevaluierung benötigt werden. Dazu gehören üblicherweise Mikrodaten über eine zunächst zu bestimmende und dann zu erhebende Serie von Zielindikatoren. Hinzu kommen Daten über wichtige

Merkmale der Firmen, die selbst einen Einfluss auf diese Zielindikatoren haben. Diese Daten werden nicht nur für die Behandlungsgruppe der geförderten Unternehmen, sondern auch für die Vergleichsgruppe der nicht geförderten Unternehmen benötigt. Ideal ist es, wenn diese Daten nicht nur nach, sondern auch bereits vor Einführung der Fördermaßnahme erhoben werden. Die Ausschreibungsbedingungen sollten das sicherstellen.

Grundsätzlich gibt es zwei Möglichkeiten der Datenerhebung: einerseits die Nutzung anderweitig (zumeist administrativ) erhobener Daten („Sekundärdaten“), andererseits die konkrete Erhebung der Daten zum Zweck der Evaluierung („Primärdaten“). Die beiden Verfahren lassen sich in bestimmten Fällen auch verbinden, etwa indem Daten der geförderten Unternehmen im Rahmen der Fördermaßnahme konkret erhoben und mit Daten einer Vergleichsgruppe aus einer bereits existierenden Datenbank verbunden werden.

Für viele Anwendungsfälle liegen umfangreiche Unternehmensdatensätze zu Zielindikatoren wie Beschäftigung und Umsatz, aber auch darüber hinaus, vor. Vorbildfunktion haben in diesem Bereich die skandinavischen Länder, in denen die amtliche Statistik der Forschung umfangreiche Registerdaten zumeist über alle Unternehmen und alle Personen des Landes zur Verfügung stellt. Aber auch in Deutschland sind administrative Datensätze verfügbar, die für Evaluierungsstudien genutzt werden können. Hier sind in erster Linie die Betriebs- und Unternehmensdaten der Statistischen Ämter des Bundes und der Länder, des Instituts für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit und der Deutschen Bundesbank zu nennen. Das Ministerium sollte sich dafür einsetzen, dass den evaluierenden Institutionen ausreichender Zugang zu den Datensätzen dieser Einrichtungen gewährt wird.

Die Forschungsdatenzentren des Statistischen Bundesamtes und der Statistischen Landesämter machen ausgewählte Mikrodaten der amtlichen Statistik für Zwecke der wissenschaftlichen Forschung zugänglich. So werden im Rahmen des Projektes „Amtliche Firmendaten für Deutschland (AFiD)“ schrittweise nicht nur diverse Unternehmensdatensätze verfügbar gemacht, sondern es bestehen auch Bestrebungen, diese über die Zeit und über verschiedene amtliche Wirtschafts- und Umweltstatistiken hinweg zu verknüpfen. Zu den AFiD-Paneldaten gehört beispielsweise das Unternehmensregister, das mit wenigen Ausnahmen für alle wirtschaftlich aktiven Unternehmen und Betriebe Deutschlands jährlich unter anderem die Zahl der Beschäftigten und Umsätze erfasst. Weitere AFiD-Paneldatensätze stellen für einzelne Wirtschaftszweige noch weit umfangreichere Informationen zur Verfügung. Die Nutzung dieser Daten ist über Gastwissenschaftlerarbeitsplätze an den Standorten der Forschungsdaten-

zentren der Statistischen Ämter sowie über kontrollierte Datenfernverarbeitung prinzipiell möglich.⁴⁹

Das Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung stellt insbesondere im Rahmen des IAB-Betriebspanels und des Betriebs-Historik-Panels Betriebsdaten zur Verfügung.⁵⁰ Bei der Deutschen Bundesbank liegen Unternehmensmikrodaten vor allem im Bereich der Direktinvestitionsbestände und der Unternehmensbilanzen vor.^{51,52} Hinzu kommen eine ganze Reihe von weiteren Forschungsdatenzentren, die Betriebs- und Unternehmensdaten sammeln, aufbereiten und für die Wissenschaft zugänglich machen.⁵³ Darüber hinaus gibt es kommerzielle Anbieter, die Firmendatenbanken etwa über Bilanzdaten bereitstellen, welche wiederum mit weiteren Datensätzen verknüpft werden können.⁵⁴

Ein sehr gutes Beispiel für die Nutzung umfangreicher existierender Betriebsdatensätze zur Evaluierung einer wirtschaftspolitischen Fördermaßnahme liefern Bade und Alm (2010), die im Auftrag des BMWi die Fördermaßnahmen im Rahmen der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRW) evaluiert haben. Für ihre Untersuchung haben sie anhand der Betriebsnummern die Statistik der sozialversicherungspflichtig Beschäftigten der Bundesagentur für Arbeit mit der GRW-Förderstatistik des Bundesamtes für Wirtschaft und Ausfuhrkontrolle verknüpft. So erhielten sie einen Datensatz, der nicht nur mit über 23.000 geförderten Betrieben rund 90% aller von der GRW geförderten Betriebe umfasst, sondern auch einen überaus umfangreichen Datensatz nicht-geförderter Betriebe, aus dem sich entsprechende Vergleichsgruppen abbilden lassen.

Die zweite Möglichkeit der Datenerhebung besteht in der genuinen Erhebung neuer Daten im Rahmen der Evaluierung. Diese wird immer dann in Betracht zu ziehen sein, wenn wichtige Zielindikatoren wie zum Beispiel die Forschungs- und Entwicklungsausgaben in bestehenden

⁴⁹ <http://www.forschungsdatenzentrum.de/afid.asp> (30.6.2013).

⁵⁰ http://fdz.iab.de/de/FDZ_Establishment_Data.aspx (30.6.2013).

⁵¹ <http://www.bundesbank.de/Navigation/DE/Bundesbank/Forschungszentrum/Forschungsdaten/Mikrodaten/mikrodaten.html> (30.6.2013).

⁵² In beschränktem Maße stehen im Rahmen des Projekts „Kombinierte Firmendaten für Deutschland (KombiFiD)“ auch verknüpfte Unternehmensdatensätze der Statistischen Ämter des Bundes und der Länder, der Bundesagentur für Arbeit und der Deutschen Bundesbank für eine Stichprobe von Unternehmen zur Verfügung (Wagner (2012); http://fdz.iab.de/de/FDZ_Projects/kombifid.aspx (30.6.2013)).

⁵³ Hierzu gehören insbesondere das Forschungsdatenzentrum Wissenschaftsstatistik, das LMU-ifo Economics & Business Data Center (EBDC), das ZEW-Forschungsdatenzentrum, das Forschungsdatenzentrum Ruhr am RWI, das Datenservicezentrum Betriebs- und Organisations-Daten, das Forschungsdatenzentrum im Bundesinstitut für Berufsbildung sowie das Internationale Datenservicezentrum des Forschungsinstituts zur Zukunft der Arbeit (<http://ratswd.de/forschungsdaten/fdz> (30.6.2013)).

⁵⁴ Das Forschungsdatenzentrum LMU-ifo Economics & Business Data Center (EBDC) verknüpft beispielsweise solche externen Bilanzdaten mit den ifo-Umfragedaten deutscher Unternehmen.

Datensätzen nicht hinreichend verfügbar sind oder wenn sich kaum geförderte Unternehmen in den bestehenden Datensätzen befinden. Im letzteren Fall bietet sich eine Kombination von bestehenden Datensätzen mit einer Zusatzerhebung geförderter Unternehmen an.

Mit der Vergabe von Fördermitteln sollte ab einem bestimmten Volumen eine Berichtspflicht der geförderten Unternehmen über wesentliche für die Evaluierung benötigte Informationen verbunden sein. Dazu sollte bereits in Bescheiden, mit denen Unternehmen in die Auswahl einbezogen werden, festgeschrieben werden, welche Indikatoren zu welchem Zeitpunkt für die Evaluierung zu liefern sind.

Eine wesentliche Herausforderung bei einer ausschließlichen Eigenerhebung dürfte darin bestehen, die relevanten Daten auch für eine geeignete Vergleichsgruppe von nicht geförderten Unternehmen zu erhalten. Insoweit die Gruppe der Antragsteller, die bei der Fördermaßnahme nicht zum Zuge kommt, im konkreten Anwendungsfall als geeignete Vergleichsgruppe angesehen wird (so beispielsweise bei den oben beschriebenen Verfahren der randomisierten Feldexperimente im Rahmen von Pilotprojekten oder des Regressions-Diskontinuitäten-Ansatzes anhand hierarchisierter Vergabelisten), ergibt sich die Möglichkeit, dass sich die interessierten Unternehmen bereits in ihrem Förderantrag dazu verpflichten müssen, auch im Falle der Nicht-Förderung die entsprechenden Daten für Evaluierungszwecke zur Verfügung zu stellen.

4.1.2 Offener Datenzugang für Forschungszwecke in Forschungsdatenzentrum

Die Daten, die im Rahmen der Evaluierungen genuin erhoben wurden, sollten der Wissenschaft möglichst breit zugänglich gemacht werden. Der offene Datenzugang für Forschungszwecke sollte, soweit möglich, nicht nur für formal mit einer Evaluierung beauftragte Institutionen, sondern für die Wissenschaft insgesamt gelten. So wird sichergestellt, dass die Evaluierungsergebnisse replizierbar sind und somit die Kontrollfunktion durch die wissenschaftliche Gemeinschaft ausgeübt werden kann.⁵⁵ Gleichzeitig sichert dies die erhöhte Chance auf Erkenntnisfortschritt im wissenschaftlichen Wettbewerb (siehe Abschnitt 4.2.2) sowie die im öffentlichen Interesse stehende effiziente Verwendung von Fördermitteln. Zudem ist der offene Datenzugang ein Gebot der Wissenschaftsethik, die verlangt, dass Forschung transparent und nachvollziehbar sein sollte. In diesem Sinne fordert auch der Ethikkodex des Vereins für Socialpolitik (2012), dem die Mitglieder dieser wissenschaftlichen Gesellschaft der deutschsprachigen Wirtschafts- und Sozialwissenschaftler verpflichtet sind: „Bei empirischen Arbeiten sollen, im Rahmen der rechtlichen Möglichkeiten, die verwendeten Datensätze und Programme zur Replikation der Ergebnisse verfügbar gemacht werden.“

⁵⁵ Vgl. hierzu Kirchgässner (2013).

Die Regelung des Datenzugangs stellt große Anforderungen an den Datenschutz und die Sicherung von Geschäftsgeheimnissen. In Deutschland steht mit der Einrichtung der Forschungsdatenzentren ein bewährter institutioneller Rahmen mit erprobten Mechanismen zur Verfügung, der diesen Anforderungen gerecht wird.⁵⁶ Deshalb empfiehlt der Beirat dem Ministerium, ein Forschungsdatenzentrum einzurichten, das die Daten der Evaluierungen der einzelnen Fördermaßnahmen sammelt, aufbereitet, dokumentiert und der Wissenschaft zugänglich macht. In diesem Rahmen sind auch die bei der Datenverarbeitung verwendeten Programmcodes – inklusive derer, die externe Daten nutzen – zu dokumentieren.

Für die Einrichtung eines Forschungsdatenzentrums sprechen neben der Datenverfügbarkeit unter Gewährleistung der Datenschutzbelange auch, dass die Dokumentation zur Qualitätssicherung beiträgt, indem Rohdaten und Routinen dokumentiert werden. Dies verringert zusätzlich die Gefahr wissenschaftlichen Fehlverhaltens. Bei der Vielzahl institutioneller Auftragnehmer, die die einzelnen Evaluierungen durchführen, trägt die Dokumentation im Forschungsdatenzentrum darüber hinaus zu einer stärkeren Kontinuität der Evaluierung bei. Schließlich hilft die Einrichtung eines Forschungsdatenzentrums, innerhalb des Ministeriums die Evaluierungskompetenz, die für eine langfristig hohe Qualität der Evaluierung, ihrer Vorbereitung, Begleitung und Interpretation notwendig ist, aufzubauen.

Bei der Einrichtung eines Forschungsdatenzentrums für Evaluierungsdaten des BMWi sollten die Anforderungen erfüllt werden, die der Rat für Sozial- und Wirtschaftsdaten (2010) für die Einrichtung der Forschungsdaten-Infrastruktur formuliert hat.⁵⁷ Diese basieren auf langjährigen Erfahrungen mit der Entwicklung einer informationellen Infrastruktur und gewährleisten einheitliche, verlässliche Standards des Datenzugangs für die Wissenschaft, die den Bestimmungen des Datenschutzes, der Datensicherheit sowie der Forschungsethik Rechnung tragen. Sie gewähren etwa mithilfe erprobter Verfahren der Anonymisierung und des geschützten Datenzugriffs – gegebenenfalls im Rahmen von Gastwissenschaftlerarbeitsplätzen oder kontrollierter Datenfernverarbeitung – eine hohe Qualität des Datenzugangs und -managements der anonymisierten Mikrodaten. Der Zugang unterliegt dabei transparenten und standardisierten rechtlichen Regelungen, die die Gleichbehandlung der Datennutzer gewährleisten. Das Forschungsdatenzentrum sollte sich vom Rat für Sozial- und Wirtschaftsdaten akkreditieren lassen, wie dies derzeit bereits für 25 weitere Datenzentren geschehen ist. Schrittweise sollte

⁵⁶ <http://ratswd.de/forschungsdaten/info> (30.6.2013).

⁵⁷ Die Kriterien berücksichtigen explizit die Belange der Etablierung einer Forschungsdaten-Infrastruktur jenseits der Amtlichen Statistik im engeren Sinne, einschließlich der „Ressortforschung“ sowie „mit öffentlichen Mitteln geförderte ... Evaluationsstudien“ (Rat für Sozial- und Wirtschaftsdaten (2010)).

auch eine Verknüpfung der einzelnen Evaluierungsdatensätze untereinander sowie mit Betriebs- und Unternehmensdaten anderer Forschungsdatenzentren in Betracht gezogen werden. Wird im Einzelfall ein geregelter Zugang zu Evaluierungsdaten für die Wissenschaft aus legitimen Gründen für gänzlich unmöglich befunden, sollte zumindest für eine entsprechende Kontrollinstanz wie den Bundesrechnungshof der Zugang zu ausreichend dokumentierten Evaluierungsdaten sichergestellt sein.

4.2 Rahmenbedingungen für unabhängige und ergebnisoffene Forschung

Damit die Evaluierung wirtschaftspolitischer Fördermaßnahmen zu einer tatsächlichen Verbesserung der Wirtschaftlichkeit des Mitteleinsatzes beitragen kann, muss die Evaluierungsforschung unabhängig und ergebnisoffen gestaltet sein. Dies verlangt auch die Wissenschaftsethik, etwa im Ethikkodex des Vereins für Socialpolitik (2012): „Wissenschaftliche Gutachten sind unvoreingenommen und ergebnisoffen zu erstellen. Das Ergebnis der Analyse soll von der Interessenlage des Auftraggebers unbeeinflusst sein.“ Gleichwohl können in vielen Fällen Aspekte der politischen Ökonomie der Evaluierung verhindern, dass eine unvoreingenommene und ergebnisoffene Evaluierung zustande kommt.⁵⁸ In einer demokratischen Gesellschaft sind Politiker oft auch Interessenvertreter, die möglicherweise nicht an schlechten Evaluierungsergebnissen interessiert sind. Unter Umständen liegt dies auch nicht im Interesse der Administration und von Projektträgern, die häufig mit der Durchführung wirtschaftspolitischer Fördermaßnahmen beauftragt werden.⁵⁹ Gleichzeitig kann es dazu kommen, dass mit Evaluierungen beauftragte Institutionen ein Interesse daran entwickeln, willfährige Gutachten zu verfassen, um auch in Zukunft Aufträge zu erhalten. Um diese Gefahren zu minimieren, müssen entsprechende Rahmenbedingungen eine freie und ergebnisoffene Evaluierungsforschung gewährleisten.

4.2.1 Organisation der Evaluierung in einer spezifischen Einheit des Ministeriums

Bei der Frage, welche Einheit für die Organisation der Evaluierung zuständig sein sollte, muss zwischen Vor- und Nachteilen der Nähe bzw. Ferne zur maßnahmendurchführenden Einheit abgewogen werden. Einerseits sprechen die geäußerten politökonomischen Argumente sowie die Notwendigkeit von Evaluierungsexpertise dafür, dass die Entscheidung, was und wie evaluiert wird, von der Verantwortlichkeit für die wirtschaftspolitische Fördermaßnahme selbst

⁵⁸ Vgl. Kirchgässner (2013).

⁵⁹ Die zumeist mit der Umsetzung von Fördermaßnahmen betrauten sogenannten Projektträger erhalten in der Regel eine pauschale Vergütung proportional zum Umfang der Maßnahme. Aus der sehr begrenzten Zahl von Projektträgern können Anreizprobleme erwachsen, da diese wiederholt um die Durchführung von Maßnahmen konkurrieren und Eigeninteressen haben können.

getrennt werden sollte. Andererseits sollte die für die Evaluierung verantwortliche Institution genügend domänenspezifische Expertise haben, um die fachlichen Fragen im Zusammenhang mit den zu evaluierenden Maßnahmen beurteilen zu können.⁶⁰ Aufgrund der erforderlichen Abwägung empfiehlt der Beirat, die Organisation der Evaluierung der wirtschaftspolitischen Fördermaßnahmen des BMWi in einer eigenständigen, fachlich unabhängigen Organisationseinheit des Ministeriums anzusiedeln.⁶¹

Dazu empfiehlt es sich, die derzeit als Aufbaustab innerhalb des Haushaltsreferates angesiedelte Evaluierungseinheit des BMWi weiterzuentwickeln. Die bereits außerhalb der Fachabteilungen und damit unabhängig von den einzelnen für die Fördermaßnahmen zuständigen Referaten organisierte Struktur der Evaluierungszuständigkeit im BMWi ist sehr zu begrüßen. Der Aufbaustab sollte zu einer eigenständigen, für die Evaluierung wirtschaftspolitischer Fördermaßnahmen zuständigen Organisationseinheit in Form eines Referates oder außerhalb der herkömmlichen Abteilungsstruktur ausgebaut und mit der Verantwortung für die Initiierung, Steuerung und Beaufsichtigung des Evaluierungsprozesses ausgestattet werden.

Diese Organisationseinheit sollte sowohl für die regelmäßige Evaluierung einzelner Fördermaßnahmen als auch für in unregelmäßigen Abständen durchzuführende Evaluierungspakete ganzer Maßnahmenbündel zuständig sein. Sie sollte auch für die Ausschreibung und die Auswahl der evaluierenden Institutionen verantwortlich sein und für die Rückkoppelung der Evaluierungsergebnisse in das für die jeweilige Fördermaßnahme zuständige Referat sorgen, damit die Ergebnisse in die nächste Runde der Gestaltung der Maßnahme einfließen. Die Etablierung einer eigenständigen Organisationseinheit für Evaluierungen würde auch dazu beitragen, die entsprechende Evaluierungskompetenz im Ministerium aufzubauen.⁶²

⁶⁰ Neben diesem Argument spricht auch die Vielzahl der zu evaluierenden Maßnahmen gegen eine Übertragung von Evaluierungen an eine ressortübergreifende Einheit wie beispielsweise den Bundesrechnungshof, zumal dieser keine exekutive Kompetenz für benötigte Ausschreibungen hat. Die Vielzahl der Maßnahmen, die lediglich als gemeinsamer Haushaltsposten vom Parlament beschlossen werden, spricht auch gegen eine Ansiedlung etwa beim Wirtschaftsausschuss des Bundestages, zumal es keine normative Formulierung der Einzelmaßnahmen durch das Parlament gibt.

⁶¹ Das Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung hat die Evaluierung der von ihm finanzierten Entwicklungsmaßnahmen kürzlich im separaten und unabhängigen Deutschen Evaluierungsinstitut der Entwicklungszusammenarbeit (DEVal) gebündelt. Aus Sicht des Beirats ist für die Evaluierung der vom BMWi geförderten wirtschaftspolitischen Maßnahmen eine Kombination aus eigenständiger Organisationseinheit innerhalb des Ministeriums mit der Vergabe der einzelnen Evaluierungsstudien nach außen im Rahmen einer wettbewerblichen Ausschreibung zielführender.

⁶² Vorbildfunktion für diese Art der Zuständigkeit und Organisation könnte etwa das Chief Evaluation Office haben, das unter der Leitung des Chief Evaluation Officers für die Implementierung, das Management und die Koordination des Evaluierungsprogramms des U.S. Department of Labor zuständig ist.

4.2.2 Ausschreibung und Öffnung für den wissenschaftlichen Wettbewerb

Für wirtschaftspolitische Fördermaßnahmen oberhalb eines zu bestimmenden Fördervolumens sollte die Entscheidung, was und wie evaluiert wird, von der Durchführung der Evaluierung getrennt werden. Darüber hinaus sollte die Durchführung der Evaluierung selbst über die beauftragende Förderinstitution hinaus geöffnet werden, da bei Selbstevaluierungen naturgemäß Interessenkonflikte auftreten können. Dazu sollten die Evaluierungsstudien öffentlich ausgeschrieben und extern durchgeführt werden. Ein wettbewerbliches Ausschreibungsverfahren, wie es für extern vergebene Studien des Ministeriums bereits die Regel ist, an dem sich sowohl kommerzielle als auch wissenschaftliche Institutionen beteiligen können, kann dabei zu Qualität und Wirtschaftlichkeit der Evaluierung selbst beitragen.⁶³ Die öffentliche Ausschreibung schützt außerdem vor zu großer Nähe zwischen der Evaluierungsstelle des Ministeriums und solchen Auftragnehmern, die das Ministerium häufig mit der Evaluierung von Fördermaßnahmen betraut.

Bereits im Zuge der Entwicklung einer Fördermaßnahme sollten neben den Zielindikatoren (vgl. Abschnitt 3.1.2) auch Mindestanforderungen für das Evaluierungskonzept festgelegt werden. Insoweit die Maßnahmen in entsprechenden Regelungsvorhaben festgelegt sind, sollten diese Mindestanforderungen an das Evaluierungskonzept in dem einschlägigen Regelungsvorhaben verankert werden.⁶⁴

Das Evaluierungskonzept muss dann der Ausschreibung der Evaluierungsstudie zugrunde liegen. Die Ausschreibung sollte nicht nur das Ziel der Evaluierung festlegen, sondern auch ein Mindestmaß der Qualität des Evaluierungsinstrumentariums sichern, etwa in Bezug auf eine überzeugende Methodik zur Bestimmung von Behandlungs- und Vergleichsgruppe.

Insbesondere bei Evaluierungsinhalten, die größere Herausforderungen als das Standardinstrumentarium der Evaluierungsforschung erwarten lassen, sollte die Unterstützung der Wissenschaft eingeholt werden. Bei größeren Evaluierungspaketen im Rahmen von vergleichenden Evaluierungen ganzer Maßnahmenbereiche bietet sich die Vergabe der einzelnen Evaluierungsstudien an verschiedene wissenschaftliche Auftragnehmer an. Dadurch kann es nicht nur zu einem fruchtbaren Austausch und gegenseitiger Kontrolle kommen, sondern es können auch die Kräfte des wissenschaftlichen Wettbewerbs genutzt werden. So lässt sich im

⁶³ Mit generellen Aspekten der Ausschreibung und Vergabe sowie der Wirtschaftlichkeit im öffentlichen Beschaffungswesen hat sich der Beirat in seinem Gutachten „Öffentliches Beschaffungswesen“ beschäftigt (Wissenschaftlicher Beirat beim Bundesministerium für Wirtschaft und Technologie (2007)).

⁶⁴ Die gesetzliche Verankerung der Evaluierungsforschung war ein wesentlicher Impuls für die Entwicklungen im Bereich der Evaluierung arbeitsmarktpolitischer Maßnahmen.

wissenschaftlichen Diskurs eine möglichst hohe Qualität der Evaluierungsstudien erreichen. Darüber hinaus kann die Evaluierungsqualität dadurch gegebenenfalls von methodischen Neuentwicklungen profitieren.

Die Wettbewerbskräfte des Marktes der Wissenschaft können zusätzlich gestärkt werden, indem – zumindest im Nachhinein – auch an der Evaluierung nicht beteiligte Wissenschaftler freien Zugang zu den für die Evaluierung notwendigen Informationen erhalten.⁶⁵ Deshalb sollte, soweit nicht ausnahmsweise zwingende Gründe dagegen sprechen, der Zugang zu allen Unterlagen und Daten für andere Wissenschaftler offen sein. Wenn im Einzelfall der Zugang für andere Wissenschaftler aus berechtigten Gründen eingeschränkt werden muss, sollte zumindest eine neutrale Instanz mit entsprechendem Zugangsrecht wie der Bundesrechnungshof ein besonderes Augenmerk auf die Qualität der Evaluierung haben.

4.2.3 Vorbehaltlose Veröffentlichung

Zu einer glaubwürdigen unabhängigen und ergebnisoffenen Evaluierungsforschung kann es nur kommen, wenn Evaluierungsergebnisse vom Auftraggeber nicht zurückgehalten werden können. Deshalb müssen alle Evaluierungen ohne längerfristigen Aufschub – zu denken ist etwa an einen Zeitraum von höchstens sechs Monaten nach Abschluss der Evaluierung – veröffentlicht werden. Es darf keinen über die kurze Frist hinaus gehenden Vorbehalt des Publikationsrechts geben. Damit ein direkter Zugang für alle Interessierten besteht, sollten die Evaluierungen frei im Internet zugänglich sein. Indem sie die Überprüfung durch die wissenschaftliche Gemeinschaft eröffnet, trägt die vorbehaltlose Veröffentlichung auch zu einer wissenschaftlichen Qualitätssicherung der Evaluierungen bei.

Damit die Evaluierungsergebnisse zu einer Weiterentwicklung der wirtschaftspolitischen Fördermaßnahmen, die die Wirksamkeit der Förderpolitik insgesamt steigern, führen kann, müssen sie schließlich in den Prozess der Maßnahmenentwicklung zurückgespielt werden. Die Evaluierungsergebnisse müssen dazu genutzt werden, die einzelnen Fördermaßnahmen effektiver auszurichten und auszugestalten, gegebenenfalls aber auch ineffektive Fördermaßnahmen zu beenden und die Förderung auf die erwiesenermaßen effektivsten Fördermaßnahmen zu fokussieren.

⁶⁵ Vgl. Kirchgässner (2013).

5. Empfehlungen

Die systematische Evaluierung wirtschaftspolitischer Fördermaßnahmen bietet die Möglichkeit zu untersuchen, wie wirksam einzelne Maßnahmen bei der Erreichung ihrer Ziele sind und welche Maßnahmen bestimmte Ziele am kostengünstigsten erreichen. Dadurch lässt sich lernen, was besonders gut und was weniger gut funktioniert. So lassen sich die Fördermaßnahmen so weiterentwickeln, dass die Wirtschaftlichkeit des Einsatzes der Fördermittel und damit letztlich die Wirksamkeit wirtschaftspolitischer Maßnahmen insgesamt deutlich erhöht werden kann. Die Erfahrungen der Arbeitsmarktevaluierungen in Deutschland belegen, dass Evaluierungsmethoden fruchtbar in den politischen Prozess eingebracht werden können.

Der Beirat stellt fest, dass die Möglichkeiten der modernen Evaluierungsforschung in der Evaluierung wirtschaftspolitischer Fördermaßnahmen in Deutschland derzeit nicht ausreichend ausgeschöpft werden. Auch ist nicht alles, was sich in der bisherigen Praxis Evaluierung nennt, eine Evaluierung im Sinne dieses Gutachtens, die Mindestanforderungen der empirisch-kausalanalytischen Wirkungsforschung wie dem Vergleich von geförderten mit vergleichbaren nicht geförderten Objekten genügen muss.

Deshalb empfiehlt der Beirat, die Möglichkeiten der Evaluierung von Fördermaßnahmen auch in der Wirtschaftspolitik besser auszuschöpfen. Konkret kommt der Beirat zu folgenden Empfehlungen:

1. Zur Sicherung der Wirtschaftlichkeit wirtschaftspolitischer Fördermaßnahmen sollten diese ab einer gewissen Größe standardmäßig einer Evaluierung unterzogen werden. Von Zeit zu Zeit sollten in Evaluierungspaketen die verschiedenen Maßnahmen eines Politikbereiches vergleichend evaluiert werden.
2. Damit eine objektive und sachgerechte Evaluierung möglich ist, sollte vor Einführung einer Fördermaßnahme nicht nur festgelegt werden, welche Ziele sie verfolgt, sondern auch, anhand welcher Zielgrößen ihre Wirksamkeit gemessen werden soll.
3. Die Evaluierungen müssen Mindestanforderungen der empirisch-kausalanalytischen Wirkungsforschung genügen, um Aussagen über die ursächliche Wirkung wirtschaftspolitischer Fördermaßnahmen ableiten zu können; randomisierte Feldexperimente und quasi-experimentelle Methoden sollten soweit wie möglich Anwendung finden. Insbesondere Pilotmaßnahmen sollten in Form von randomisierten Experimenten evaluiert werden.

4. Da objektive Evaluierung erst durch Verfügbarkeit und Zugang zu entsprechenden Daten möglich ist, sollte mit einem Forschungsdatenzentrum für Evaluierungsdaten eine Dateninfrastruktur aufgebaut werden, die den evaluierenden Einrichtungen und der Wissenschaft insgesamt Zugang zu evaluierungsrelevanten Mikrodaten ermöglicht.
5. Um die Güte der Evaluierungsstudien zu steigern, müssen Rahmenbedingungen für eine unabhängige und ergebnisoffene Evaluierung geschaffen werden. Eine für die Initiierung, Steuerung und Beaufsichtigung der Evaluierungen zuständige Organisationseinheit des Ministeriums sollte die wissenschaftliche Qualität und die fachliche Unabhängigkeit der Studien sicherstellen. Darüber hinaus muss die wettbewerbliche Ausschreibung garantiert sein. Bei anspruchsvollen Vorhaben sollte der wissenschaftliche Wettbewerb zur Qualitätssicherung genutzt werden.
6. Um Transparenz über Vorgehen und Ergebnisse der Evaluierungen herzustellen, sind die Evaluierungsstudien ohne Publikationsvorbehalt spätestens sechs Monate nach Abschluss zu veröffentlichen.

Es muss eine Evaluierungskultur entstehen, die ins Bewusstsein rückt, dass das Lernen über effektive und weniger effektive Maßnahmen zum Prozess einer wirksamen Politikgestaltung gehört. Auch die Erkenntnis, dass eine spezifische Maßnahme nicht den Erwartungen entsprechend gewirkt hat, ist ein wichtiger Gewinn, der zur Verbesserung der politischen Maßnahmen insgesamt beitragen kann. Insofern ist das oberste Anliegen der Evaluierung zu lernen, was man besser machen kann. Es muss der Politik erlaubt sein, auch innovative Fördermaßnahmen auszuprobieren, solange sie bereit ist, die Maßnahmen so zu begleiten, dass sie über ihre Effektivität lernt. Evaluierung darf dabei keine übermäßige Risikoaversion forcieren. Dazu bedarf es der Entwicklung einer Evaluierungskultur, die positive wie negative Evaluierungsergebnisse als Gewinn versteht, die Lernprozesse anstoßen und eine produktive Weiterentwicklung des Maßnahmenpaketes ermöglichen.

Berlin, den 20. September 2013

Der Vorsitzende des Wissenschaftlichen Beirats
beim Bundesministerium für Wirtschaft und Technologie

Prof. Achim Wambach, Ph.D

Literatur

- Aichele, Rahel und Gabriel Felbermayr (2012). Kyoto and the carbon footprint of nations. *Journal of Environmental Economics and Management* 63 (3): 336-354.
- AiF Projekt GmbH (2013). *Innovation im Mittelstand: 20 Jahre FuE-Kooperationsförderung durch die AiF im Auftrag des Bundes*. Berlin: AiF Projekt GmbH. <http://www.zim-bmwi.de/download/publikationen/20-jahre-kooperationsfoerderung.de> (30.6.2013).
- Allmendinger, Jutta und Annette Kohlmann (2005). Datenverfügbarkeit und Datenzugang am Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung. *Allgemeines Statistisches Archiv* 88 (2): 159-182.
- Angrist, Joshua D. und Alan B. Krueger (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives* 15 (4): 69-85.
- Angrist, Joshua D. und Jörn-Steffen Pischke (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.
- Angrist, Joshua D. und Jörn-Steffen Pischke (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives* 24 (2): 3-30.
- Arni, Patrick (2012). Kausale Evaluation von Pilotprojekten: Die Nutzung von Randomisierung in der Praxis. IZA Standpunkte 52. Bonn: Forschungsinstitut zur Zukunft der Arbeit.
- Bade, Franz-Josef und Bastian Alm (2010). Endbericht zum Gutachten Evaluierung der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRW) durch einzelbetriebliche Erfolgskontrolle für den Förderzeitraum 1999-2008 und Schaffung eines Systems für ein gleitendes Monitoring. Dortmund: Fachgebiet Volkswirtschaftslehre, insbes. Raumwirtschaftspolitik der Technischen Universität Dortmund.
- Bauer, Thomas K., Michael Fertig und Christoph M. Schmidt (2009). *Empirische Wirtschaftsforschung: Eine Einführung*. Berlin: Springer.
- Becker, Sascha O., Peter H. Egger und Maximilian von Ehrlich (2010). Going NUTS: The effect of EU Structural Funds on regional performance. *Journal of Public Economics* 94 (9-10): 578-590.
- Becker, Sascha O., Peter H. Egger und Maximilian von Ehrlich (2013). Absorptive capacity and the growth and investment effects of regional transfers: A regression discontinuity design with heterogeneous treatment effects. *American Economic Journal: Economic Policy*: im Erscheinen.
- Benus, Jacob, Theodore Shen, Sisi Zhang, Marc Chan und Benjamin Hansen (2009). *Growing America Through Entrepreneurship: Final Evaluation of Project GATE*. Columbia, MD: IMPAQ International.
- Bertrand, Marianne, Esther Duflo und Sendhil Mullainathan (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 114 (1): 249-275.
- Besley, Timothy und Anne Case (2000). Unnatural experiments? Estimating the incidence of endogenous policies. *Economic Journal* 110 (467): F672-F694.
- Biewen, Martin, Bernd Fitzenberger, Aderonke Osikominu und Marie Paul (2013). The effectiveness of public sponsored training revisited: The importance of data and methodological choices. *Journal of Labor Economics*: im Erscheinen.

- Bouguen, Adrien und Marc Gurgand (2012). Randomized controlled experiments in education. EENEE Analytical Report 11. http://www.eenee.org/doc/eenee_ar11.pdf (30.6.2013).
- Brinkmann, Christian, Reinhard Hujer und Susanne Koch (2006). Evaluation aktiver Arbeitsmarktpolitik in Deutschland - eine Einführung. *Zeitschrift für ArbeitsmarktForschung* 39 (3/4): 319-327.
- Bundesministerium der Finanzen (2011). *Dreiundzwanzigster Subventionsbericht: Bericht der Bundesregierung über die Entwicklung der Finanzhilfen des Bundes und der Steuervergünstigungen für die Jahre 2009 bis 2012*. Berlin: Bundesministerium der Finanzen.
- Bundesministerium der Finanzen (2013). *24. Subventionsbericht: Bericht der Bundesregierung über die Entwicklung der Finanzhilfen des Bundes und der Steuervergünstigungen für die Jahre 2011 bis 2014*. Berlin: Bundesministerium der Finanzen.
- Bundesministerium für Arbeit und Soziales und Institut für Arbeitsmarkt- und Berufsforschung (2011). *Sachstandsbericht der Evaluation der Instrumente*. Berlin: Bundesministerium für Arbeit und Soziales. http://www.bmas.de/SharedDocs/Downloads/DE/arbeitsmarktpol_instr_iab_studie.pdf?__blob=publicationFile (30.6.2013).
- Bundesregierung (2013). *Bessere Rechtsetzung 2012: Belastungen vermeiden, Bürokratischen Aufwand verringern, Wirtschaftliche Dynamik sichern. Bericht der Bundesregierung 2012 nach § 7 des Gesetzes zur Einsetzung eines Nationalen Normenkontrollrates*. Berlin: Bundeskanzleramt.
- Caliendo, Marco, Konstantinos Tatsiramos und Arne Uhlendorff (2013). Benefit duration, unemployment duration and job match quality: A regression-discontinuity approach. *Journal of Applied Econometrics* 28 (4): 604-627.
- Cornet, Maarten, Björn Vroomen und Marc van der Steeg (2006). Do innovation vouchers help SMEs to cross the bridge towards science? CPB Discussion Paper 58. The Hague: CPB Netherlands Bureau for Economic Policy Analysis.
- Criscuolo, Chiara, Ralf Martin, Henry Overman und John Van Reenen (2012). The causal effects of an industrial policy. NBER Working Paper 17842. Cambridge, MA: National Bureau of Economic Research.
- Deaton, Angus (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature* 48 (2): 424-55.
- Einiö, Elias und Henry Overman (2012). The effects of spatially targeted enterprise initiatives: Evidence from UK LEGI. Mimeo.
- Engel, Christoph (2013). *Legal experiments: Mission impossible?* Den Haag: Eleven International Publishing.
- Expertenkommission Forschung und Innovation (2013). *Gutachten zu Forschung, Innovation und Technologischer Leistungsfähigkeit Deutschlands*. Berlin: EFI.
- Fairlie, Robert W., Dean Karlan und Jonathan Zinman (2012). Behind the GATE Experiment: Evidence on Effects of and Rationales for Subsidized Entrepreneurship Training. NBER Working Paper 17804. Cambridge, MA: National Bureau of Economic Research.
- Falck, Oliver, Stephan Heblich und Stefan Kipar (2010). Industrial innovation: Direct evidence from a cluster-oriented policy. *Regional Science and Urban Economics* 40 (6): 574-582.

- Falck, Oliver und Simon Wiederhold (2013). *Nachfrageorientierte Innovationspolitik*. Studien zum deutschen Innovationssystem Nr. 12-2013. Studie des ifo Instituts im Auftrag von EFI. Berlin: Expertenkommission Forschung und Innovation (EFI).
- Felbermayr, Gabriel J. und Erdal Yalcin (2013). Export credit guarantees and export performance: An empirical analysis for Germany. *The World Economy*: im Erscheinen.
- Frölich, Markus und Michael Lechner (2010). Exploiting regional treatment intensity for the evaluation of labor market policies. *Journal of the American Statistical Association* 105 (491): 1014-1029.
- Graversen, Brian Krogh und Jan C. van Ours (2008). How to help unemployed find jobs quickly: Experimental evidence from a mandatory activation program. *Journal of Public Economics* 92 (10-11): 2020-2035.
- Harrison, Glenn W. und John A. List (2004). Field experiments. *Journal of Economic Literature* 42 (4): 1009-1055.
- Haynes, Laura, Owain Service, Ben Goldacre und David Torgerson (2012). *Test, learn, adapt: Developing public policy with randomised controlled trials*. London: Cabinet Office Behavioural Insights Team.
- Heckman, James J. (2010). Building bridges between structural and program evaluation approaches to evaluating policy. *Journal of Economic Literature* 48 (2): 356-398.
- Holland, Paul W. (1986). Statistics and causal inference. *Journal of the American Statistical Association* 81 (396): 945-960.
- Imbens, Guido W. (2010). Better LATE than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature* 48 (2): 399-423.
- Imbens, Guido W. und Jeffrey M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47 (1): 5-86.
- Kerr, William R., Josh Lerner und Antoinette Schoar (2013). The consequences of entrepreneurial finance: Evidence from angel financings. *Review of Financial Studies*: im Erscheinen.
- Kirchgässner, Gebhard (2013). Zur Rolle der Ökonometrie in der wissenschaftlichen Politikberatung. *Perspektiven der Wirtschaftspolitik* 14 (1-2): 3-30.
- Kline, Patrick M. und Enrico Moretti (2013). Local economic development, agglomeration economies, and the Big Push: 100 years of evidence from the Tennessee Valley Authority. NBER Working Paper 19293. Cambridge, MA: National Bureau of Economic Research.
- Laaser, Claus-Friedrich und Astrid Rosenschon (2013). Subventionen in Deutschland in den Jahren 2000 bis 2011/2012: Der Kieler Subventionsbericht. Kieler Diskussionsbeiträge 516/517. Kiel: Institut für Weltwirtschaft.
- Lalive, Rafael (2008). How do extended benefits affect unemployment duration? A regression discontinuity approach. *Journal of Econometrics* 142 (2): 785-806.
- Lechner, Michael, Ruth Miquel und Conny Wunsch (2011). Long-run effects of public sector sponsored training in West Germany. *Journal of the European Economic Association* 9 (4): 742-784.
- Lee, David S. und Thomas Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature* 48 (2): 281-355.

- Lerner, Josh (1999). The government as venture capitalist: The long-run impact of the SBIR Program. *Journal of Business* 72 (3): 285-318.
- Link, Albert N. und John T. Scott (2012). Employment growth from the Small Business Innovation Research program. *Small Business Economics* 39 (2): 265-287.
- List, John A. (2011). Why economists should conduct field experiments and 14 tips for pulling one off. *Journal of Economic Perspectives* 25 (3): 3-16.
- Ludwig, Jens, Jeffrey R. Kling und Sendhil Mullainathan (2011). Mechanism experiments and policy evaluations. *Journal of Economic Perspectives* 25 (3): 17-38.
- Manski, Charles F. (1995). *Identification problems in the social sciences*. Cambridge, MA: Harvard University Press.
- Mayer, Thierry, Florian Mayneris und Loriane Py (2012). The impact of urban enterprise zones on establishments' location decisions: Evidence from French ZFUs. CEPR Discussion Paper 9074. London: Centre for Economic Policy Research.
- Pedersen, Jonas Maibom, Michael Rosholm und Michael Svarer (2012). Experimental evidence on the effects of early meetings and activation. IZA Discussion Paper 6970. Bonn: Institute for the Study of Labor.
- Präsident des Bundesrechnungshofes, Hrsg. (2013). *Anforderungen an Wirtschaftlichkeitsuntersuchungen finanzwirksamer Maßnahmen nach § 7 Bundeshaushaltsordnung*. Schriftenreihe des Bundesbeauftragten für Wirtschaftlichkeit in der Verwaltung, Band 18. Stuttgart: W. Kohlhammer.
- Rat für Sozial- und Wirtschaftsdaten (2010). Kriterien des Rates für Sozial- und Wirtschaftsdaten (RatSWD) für die Einrichtung der Forschungsdaten-Infrastruktur. http://ratswd.de/dl/doc/RatSWD_FDZKriterien_0.PDF (30.6.2013).
- Rosendahl Huber, Laura, Randolph Sloof und Mirjam Van Praag (2012). The effect of early entrepreneurship education: Evidence from a randomized field experiment. IZA Discussion Paper 6512. Bonn: Institute for the Study of Labor.
- Schlotter, Martin, Guido Schwerdt und Ludger Woessmann (2011). Econometric methods for causal evaluation of education policies and practices: a non-technical guide. *Education Economics* 19 (2): 109-137.
- Schmidt, Christoph M. (2009). Wirtschaftswissenschaft und Politikberatung in Deutschland – Bedeutung, Möglichkeiten und Grenzen der Kausalanalyse. In *Wirtschaftspolitik im Zeichen europäischer Integration: Festschrift für Wim Kösters anlässlich seines 65. Geburtstages*, hrsg. von Ansgar Belke, Hans-Helmut Kotz, Stephan Paul und Christoph M. Schmidt. Berlin: Duncker & Humblot.
- Schmieder, Johannes F., Till von Wachter und Stefan Bender (2012). The effects of extended unemployment insurance over the business cycle: Evidence from regression discontinuity estimates over 20 years. *Quarterly Journal of Economics* 127 (2): 701-752.
- Schwerdt, Guido, Dolores Messer, Ludger Woessmann und Stefan C. Wolter (2012). The impact of an adult education voucher program: Evidence from a randomized field experiment. *Journal of Public Economics* 96 (7-8): 569-583.
- Steinke, Joß, Susanne Koch, Peter Kupka, Christopher Osiander, Elke Dony, Detlef Güttler, Claudia Hesse und Barbara Knapp (2012). Neuorientierung der Arbeitsmarktpolitik. Die Neuausrichtung der arbeitsmarktpolitischen Instrumente aus dem Jahr 2009 im Blickpunkt: Mehr Flexibilität und größere Handlungsspielräume für die Vermittler?

- IAB-Forschungsbericht 2/2012. Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung.
- Stock, James H. und Mark W. Watson (2012). *Introduction to econometrics*. 3 ed. Boston: Pearson.
- Verein für Socialpolitik (2012). Ethikkodex des Vereins für Socialpolitik. Frankfurt: Verein für Socialpolitik.
- Wagner, Joachim (2012). Editorial: KombiFiD – Combined Firm Data for Germany. *Schmollers Jahrbuch* 132 (3): 359.
- Wallsten, Scott J. (2000). The effects of government-industry R&D programs on private R&D: The case of the Small Business Innovation Research Program. *RAND Journal of Economics* 31 (1): 82-100.
- Wessner, Charles W., Hrsg. (2008). *An Assessment of the SBIR Program*. Washington, D.C.: National Academy Press.
- Wissenschaftlicher Beirat beim Bundesministerium für Wirtschaft und Technologie (2007). *Öffentliches Beschaffungswesen*. Gutachten Nr. 2/07. Berlin: Bundesministerium für Wirtschaft und Technologie.
- Wunsch, Conny und Michael Lechner (2008). What did all the money do? On the general ineffectiveness of recent West German labour market programmes. *Kyklos* 61 (1): 134-174.

Mitgliederverzeichnis

Das Gutachten wurde vorbereitet von folgenden Mitgliedern des Wissenschaftlichen Beirats beim Bundesministerium für Wirtschaft und Technologie

Professor Dr. Ludger Wößmann (Federführung)
Professor für Volkswirtschaftslehre insb. Bildungsökonomik
an der Ludwig-Maximilians-Universität München
Bereichsleiter, Humankapital und Innovation
ifo Institut für Wirtschaftsforschung, München

Professor Dr. Stefan Bechtold
Professor für Geistiges Eigentum (Intellectual Property)
D-GESS Departement Geistes-, Sozial- und Staatswissenschaften
an der ETH Zürich

Professor Axel Börsch-Supan, Ph.D.
Direktor des Munich Center for the Economics of Aging (MEA)
am Max-Planck-Institut für Sozialrecht und Sozialpolitik, München

Professor Dr. Christoph Engel
Direktor am Max-Planck-Institut zur Erforschung von Gemeinschaftsgütern
Professor für Rechtswissenschaften an der Universität Osnabrück

Professor Dietmar Harhoff, Ph.D.
Professor für Betriebswirtschaftslehre
Vorstand des Instituts für Innovationsforschung,
Technologiemanagement und Entrepreneurship
an der Ludwig-Maximilians-Universität München

Professor Regina T. Riphahn, Ph.D.
Professor für Statistik und empirische Wirtschaftsforschung
an der Friedrich-Alexander-Universität Erlangen-Nürnberg

Professor Achim Wambach, Ph.D. (Vorsitzender)
Professor für wirtschaftliche Staatswissenschaften
an der Universität zu Köln

Das Gutachten wurde beraten von folgenden Mitgliedern des Wissenschaftlichen Beirats beim Bundesministerium für Wirtschaft und Technologie

Professor Dr. Eckhard Janeba (Stellvertretender Vorsitzender)
Professor für Volkswirtschaftslehre insbesondere Finanzwissenschaft
und Wirtschaftspolitik an der Universität Mannheim

Professor Dr. Hermann Albeck
Em. Professor für Volkswirtschaftslehre
an der Universität Saarbrücken

Professor Dr. Peter Bernholz
Em. Professor für Nationalökonomie,
insbesondere Geld- und Außenwirtschaft,
an der Universität Basel

Professor Dr. Norbert Berthold
Professor für Volkswirtschaftslehre an der
Bayerischen Julius-Maximilians-Universität in Würzburg

Professor Dr. Charles B. Blankart
Em. Professor für Wirtschaftswissenschaften
an der Humboldt-Universität zu Berlin

Professor Dr. Dres. h.c. Knut Borchardt
Em. Professor für Wirtschaftsgeschichte und
Volkswirtschaftslehre an der Universität München

Professor Dr. Friedrich Breyer
Professor für Volkswirtschaftslehre
an der Universität Konstanz

Herrn Professor Marcel Fratzscher, Ph.D.
Präsident des Deutschen Instituts für Wirtschaftsforschung (DIW Berlin) und
Professor für Makroökonomie und Finanzen an der Humboldt-Universität Berlin
Mohrenstr. 58, 10117 Berlin

Frau Professor Christina Gathmann, Ph.D.
Lehrstuhl für Arbeitsmarktökonomie und Neue Politische Ökonomik
Alfred-Weber-Institut für Wirtschaftswissenschaften
Bergheimer Str. 20, 69115 Heidelberg

Professor Dr. Hans Gersbach
Professor für Wirtschaftspolitik
CER-ETH – Center of Economic Research
at ETH Zürich, Schweiz

Professor Dr. Roman Inderst
Professor für Finanzen und Ökonomie
an der Universität Frankfurt/M.
House of Finance

Professor Dr. Günter Knieps
Direktor des Instituts für Verkehrswissenschaft
und Regionalpolitik; Wirtschaftswissenschaftliche Fakultät der
Albert-Ludwigs-Universität Freiburg

Professor Dr. Dr. h.c. Wernhard Möschel
Professor für Bürgerliches Recht, Handels- und Wirtschaftsrecht
an der Universität Tübingen

Professor Dr. Manfred J.M. Neumann
Em. Professor für Wirtschaftliche Staatswissenschaften,
insbesondere Wirtschaftspolitik,
an der Universität Bonn

Professor Dr. Manfred Neumann
Em. Professor für Volkswirtschaftslehre
an der Universität Erlangen-Nürnberg

Professor Dr. Axel Ockenfels
Professor für Wirtschaftliche Staatswissenschaften
Staatswissenschaftliches Seminar
an der Universität zu Köln

Professor Dr. Albrecht Ritschl
Professor für Wirtschaftsgeschichte
an der London School of Economics

Professor Dr. Klaus Schmidt
Professor für Volkswirtschaftslehre
an der Ludwig-Maximilians-Universität München

Professor Dr. Olaf Sievert
Präsident der Landeszentralbank in den Freistaaten Sachsen und Thüringen, Leipzig i.R.,
Honorarprofessor Universität Saarbrücken

Professor Dr. Roland Vaubel
Professor für Volkswirtschaftslehre
an der Universität Mannheim

Professor Dr. Carl Christian von Weizsäcker
Em. Professor für Volkswirtschaftslehre
an der Universität zu Köln

Professor Dr. Christian Watrin
Em. Professor für wirtschaftliche Staatswissenschaften
an der Universität Köln

Anhang: Fördermittel im Haushalt des Bundesministeriums für Wirtschaft und Technologie

Soll 2013, in Mio. Euro (Stand: 22. November 2012)

Kap. 0901	Innovation, Technologie und Neue Mobilität	2.343,4
683 01	Innovationsförderung, Zentrales Innovationsprogramm Mittelstand (ZIM)	510,1
683 02	Innovationsberatung	17,2
685 01	Technologie- und Innovationstransfer	29,7
685 02	Nationale Akkreditierungsstelle	2,0
686 01	Forschungsinfrastruktur	199,0
Tgr. 01	Neue Mobilität	107,2
546 11	Kosten der Gemeinsamen Geschäftsstelle Elektromobilität der Bundesregierung	0,5
662 11	Zinsausgleichssystem auf CIRR-Basis	1,3
683 11	Verkehrstechnologien	58,4
683 12	Maritime Technologien	32,3
683 13	Steigerung der Wettbewerbsfähigkeit der maritimen Wirtschaft	1,8
892 10	Innovationsbeihilfen zugunsten der deutschen Werftindustrie	13,0
Tgr. 02	Informations- und Kommunikationstechnologien	71,6
683 21	Entwicklung konvergenter Informations- und Kommunikationstechnik (IKT)	40,5
686 22	IKT-Anwendungen in der Wirtschaft	19,5
894 21	Förderung und Entwicklung von (IKT-)Infrastrukturen, Breitband	2,0
894 22	Zuschüsse zu den Umstellungskosten aus der Freigabe von Frequenzen ("Digitale Dividende")	9,6
Tgr. 03	Luft- und Raumfahrt	1.406,6
526 01	Gerichts- und ähnliche Kosten	1,4
662 31	Finanzierungshilfen für den Absatz von zivilen Flugzeugen inklusive Triebwerke	0,5
662 32	Ausgaben zur Absicherung des Ausfallrisikos im Zusammenhang mit Darlehen (zivile Luftfahrzeuge)	6,0
683 31	Forschungsförderung von Technologievorhaben der zivilen Luftfahrt - Förderung von Einzelvorhaben	155,5
683 32	Nationales Programm für Weltraum und Innovation - Forschungs- und Entwicklungsvorhaben	272,0
685 31	Deutsches Zentrum für Luft- und Raumfahrt e. V. - Betrieb	269,3
894 31	Deutsches Zentrum für Luft- und Raumfahrt e. V. - Investitionen	62,7
896 31	Beitrag bzw. Leistungen an die Europäische Weltraumorganisation (ESA) in Paris	639,3
Kap. 0902	Mittelstand: Gründen, Wachsen, Investieren	874,4
662 02	Zinszuschüsse im Rahmen von ERP-Förderprogrammen	64,1
686 02	Mittelstandsinstitute und Arbeitsgemeinschaft für wirtschaftliche Verwaltung e. V.	9,7
686 04	Berufliche Bildung für den Mittelstand -Lehrlingsunterweisung	45,1
686 05	Fachkräftesicherung für kleine und mittlere Unternehmen	13,0
686 06	Potenziale in der Dienstleistungswirtschaft	7,6
686 07	Innovative Unternehmensgründungen	83,0
686 08	Förderung unternehmerischen Know-hows	40,2
882 01	Zuweisungen für betriebliche Investitionen und wirtschaftsnahe Infrastrukturmaßnahmen im Rahmen der GRW	582,8
893 01	Berufliche Bildung für den Mittelstand -Fortbildungseinrichtungen	29,0
Kap. 0903	Energie und Nachhaltigkeit	1.597,5
541 01	Erstellung der Energiebilanzen sowie Bereitstellung sonstiger Daten für die Emissionsberichterstattung	1,3
683 01	Energieforschung	115,1
686 02	Sicherheitsforschung für kerntechnische Anlagen	34,1
686 03	Steigerung der Energieeffizienz	29,9
687 02	Leistungen an die Internationale Atomenergie-Organisation (IAEO) in Wien	32,4
892 01	Explorationsprogramm (Explo II)	7,5
Tgr. 01	Wismut-Sanierung / Auslaufen der Steinkohle-Subventionen	1.377,2
Kap. 0904	Chancen der Globalisierung	303,5
532 04	Kosten der Beteiligung des Bundes an Weltausstellungen im Ausland	9,3
686 01	Zuwendung an die Deutsche Zentrale für Tourismus e. V. , Frankfurt (Main)	28,3
687 01	Maßnahmen im Zusammenhang mit Twinning-Projekten der EU und Beratungshilfe für das Ausland	1,2
687 02	Wirtschaftsbeziehungen mit dem Ausland einschließlich Standortmarketing	56,9
687 03	Beiträge an internationale Organisationen mit Sitz im Ausland	31,9
687 05	Erschließung von Auslandsmärkten	81,0
687 06	Deutscher Beitrag für die Unterstützung Russlands bei Abrüstungsprojekten	65,0
896 01	Zuschuss zur anteiligen Finanzierung der Errichtung einer Stadtbahn in Ho Chi Minh-Stadt, Vietnam	30,0
Kap. 0910	Sonstige Bewilligungen (ohne Titel 972 01)	78,7
Kap. 0901 bis 0910 (Gesamtsumme der Fördermittel)		5.197,6