



Brüssel, den XXX
SWD(2014) XXX

ARBEITSUNTERLAGE DER KOMMISSIONSDIENSTSTELLEN

Gemeinsame Methodik für die Evaluierung staatlicher Beihilfen

ARBEITSUNTERLAGE DER KOMMISSIONSDIENSTSTELLEN

Gemeinsame Methodik für die Evaluierung staatlicher Beihilfen

Inhaltsverzeichnis

1	Einführung.....	2
2	Die Ziele der Evaluierung staatlicher Beihilfen.....	3
3	Der Evaluierungsplan.....	5
3.1	Ziele der zu evaluierenden Beihilferegelung.....	5
3.2	Die Evaluierungsfragen.....	6
3.3	Ergebnisindikatoren	7
3.4	Methoden: Ermittlung einer geeigneten Vergleichsgrundlage.....	7
3.5	Datenerhebung: Nutzung der bestmöglichen Quellen	10
3.6	Zeitlicher Rahmen für die Evaluierung	11
3.7	Das Evaluierungsgremium: Sicherstellung von Unabhängigkeit und Fachkompetenz	12
3.8	Veröffentlichung: Erleichterung der Einbeziehung von Interessenträgern	13
4	Kriterien für die Auswahl der zu evaluierenden Beihilferegelungen	14
4.1	Umfangreiche Beihilferegelungen einschließlich solcher, die unter die Allgemeine Gruppenfreistellungsverordnung fallen	14
4.2	Neuartige Beihilferegelungen	15
4.3	Beihilferegelungen, bei denen wesentliche Veränderungen abzusehen sind	15
4.4	Sonstige Beihilferegelungen.....	16
Anhang I:	Technischer Anhang zu den einschlägigen Methoden zur Ermittlung kausaler Effekte ...	18
Anhang II:	Aufstellung möglicher Ergebnisindikatoren	41
Anhang III:	Glossar	46
Anhang IV:	Literaturhinweise.....	47

1 Einführung

Die Mitgliedstaaten gewähren staatliche Beihilfen, um zur Verwirklichung eines breiten Spektrums politischer Ziele beizutragen, zum Beispiel, um das regionale Entwicklungsgefälle innerhalb eines Landes abzubauen, um Forschung, Entwicklung und Innovation zu fördern oder um ein hohes Umweltschutzniveau herbeizuführen.

Für die Feststellung, welche Arten von Beihilfen mit dem Binnenmarkt vereinbar sind, sehen die EU-Vorschriften für staatliche Beihilfen ein System der Ex-ante-Kontrolle vor: Beihilferegulungen¹ werden auf der Grundlage vorab festgelegter Prüfungskriterien genehmigt, wobei davon ausgegangen wird, dass ihre positiven Effekte etwaige negative Effekte überwiegen, wenn sie diese Kriterien erfüllen. Im Allgemeinen wird bei der Prüfung von Beihilferegulungen nicht ausreichend evaluiert, welche tatsächlichen Auswirkungen sie im Laufe der Zeit auf die Märkte haben.

Bei der Anwendung der EU-Beihilfavorschriften wurde Ex-post-Belegen dafür, was mit den öffentlichen Mitteln tatsächlich erreicht wurde oder wie sich die staatliche Beihilfe auf den Wettbewerb ausgewirkt hat, bisher relativ wenig Bedeutung beigemessen. Für die Entscheidungsträger in den Mitgliedstaaten und auf EU-Ebene ist es jedoch unerlässlich, die messbaren Ergebnisse der in der Vergangenheit gewährten staatlichen Beihilfen zu prüfen und die daraus gewonnenen Erkenntnisse zu berücksichtigen. Dies wird dazu beitragen, dass aus staatlichen Mitteln finanzierte Beihilferegulungen wirksamer werden und weniger Marktverzerrungen verursachen, und auch die Effizienz künftiger Regelungen und möglicherweise künftiger Vorschriften für die Gewährung staatlicher Beihilfen verbessern.

Zahlreiche Länder evaluieren bereits ihre Fördermaßnahmen, wenn auch nicht immer systematisch.² Auch die Ausgaben der EU (einschließlich der Finanzierungen aus EU-Struktur- und Investitionsfonds wie EFRE, ESF oder ELER) werden nach Maßgabe der geltenden Verordnungen und der von der Kommission veröffentlichten Anleitungsdokumente³ Ex-ante-, laufenden und Ex-post-Evaluierungen unterzogen. Zur Vermeidung von Doppelarbeit bei den von den Mitgliedstaaten durchgeführten Evaluierungen wird im Anleitungsdokument für Monitoring und Evaluierung „*Concepts and*

1 Beihilferegulungen machen den Großteil aller gewährten Beihilfen aus: Wie die Daten des Beihilfenanzeigers für 2013 zeigen, entfallen 23 % aller Beihilfemaßnahmen und 55 % der Beihilfebeträge auf genehmigte Beihilferegulungen sowie 63 % aller Beihilfemaßnahmen und rund 32 % der Beihilfebeträge auf Regelungen, die unter eine Gruppenfreistellung fallen. In der Verordnung (EG) Nr. 659/1999 des Rates ist „Beihilferegulung“ definiert als „eine Regelung, wonach Unternehmen, die in der Regelung in einer allgemeinen und abstrakten Weise definiert werden, ohne nähere Durchführungsmaßnahmen Einzelbeihilfen gewährt werden können, beziehungsweise eine Regelung, wonach einem oder mehreren Unternehmen nicht an ein bestimmtes Vorhaben gebundene Beihilfen für unbestimmte Zeit und/oder in unbestimmter Höhe gewährt werden können“.

2 So werden in mehreren Mitgliedstaaten regelmäßig Berichte über die Evaluierung staatlicher Beihilfen für den Rechnungshof oder das Parlament erstellt.

3 Die einschlägigen Konzepte und Empfehlungen sind in den Anleitungsdokumenten der Kommission 2014-2020 ausführlich dargelegt (abrufbar hier: http://ec.europa.eu/regional_policy/information/evaluations/guidance_de.cfm).

Recommendations“ klargestellt, dass die Evaluierungspflichten für die europäischen Struktur- und Investitionsfonds dadurch erfüllt werden können, dass die nach den Beihilfavorschriften erforderlichen Evaluierungen vorgenommen werden.

Mit der Initiative zur Modernisierung der Beihilfenkontrolle⁴ soll erreicht werden, dass sich die Kommission bei der Durchsetzung des Beihilferechts auf umfangreichere Beihilferegelungen konzentriert, die wahrscheinlich die größten Auswirkungen auf den Binnenmarkt haben. Gleichzeitig soll die Prüfung von Fällen eher lokaler Art, die sich nur in geringem oder begrenztem Maße auf den Handel auswirken, unter anderem dadurch vereinfacht werden, dass den Mitgliedstaaten durch Ausweitung des Anwendungsbereichs der neuen Allgemeinen Gruppenfreistellungsverordnung⁵ mehr Flexibilität bei der Durchführung dieser Beihilfemaßnahmen eingeräumt wird. Um sicherzustellen, dass die positiven Effekte staatlicher Beihilfen (bei der Erreichung ihres ursprünglichen Zwecks) insgesamt weiterhin die potenziellen negativen Effekte auf Handel und Wettbewerb überwiegen, und um übermäßige Marktverzerrungen zu verhindern, soll die stärkere Vereinfachung mit mehr Transparenz, einer verstärkten Kontrolle der Einhaltung der Beihilfavorschriften auf nationaler und europäischer Ebene und einer wirksamen Evaluierung einhergehen.⁶

In diesem Papier ist eine gemeinsame Methodik für die Evaluierung von Beihilferegelungen dargelegt. Es soll den an der Planung und Durchführung von Evaluierungen beteiligten Behörden als Orientierungshilfe dienen.

2 Die Ziele der Evaluierung staatlicher Beihilfen

Im Rahmen der Evaluierung staatlicher Beihilfen sollen vor allem die relativen positiven und negativen Effekte einer Regelung, d. h. der öffentliche Zweck der Beihilfe im Verhältnis zu deren Auswirkungen auf den Wettbewerb und den Handel zwischen Mitgliedstaaten geprüft werden. Durch Evaluierung staatlicher Beihilfen kann ermittelt werden, ob und inwieweit der ursprüngliche Zweck einer Beihilferegelung erreicht wurde (Prüfung der positiven Effekte) und welche Auswirkungen die Regelung auf die Märkte und den Wettbewerb hatte (Prüfung der negativen Effekte). Die Evaluierung unterscheidet sich daher in ihrem Zweck von Monitoring⁷ und Berichterstattung⁸, den beiden Ex-post-Prüfungen, die die Kommission bisher bei Beihilferegelungen vornimmt.

4 Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen – Modernisierung des EU-Beihilfenrechts (COM(2012) 209 final vom 8.5.2012).

5 Verordnung (EU) Nr. 651/2014 der Kommission vom 17. Juni 2014 zur Feststellung der Vereinbarkeit bestimmter Gruppen von Beihilfen mit dem Binnenmarkt in Anwendung der Artikel 107 und 108 des Vertrags über die Arbeitsweise der Europäischen Union (ABl. L 187 vom 26.6.2014, S. 1).

6 Siehe auch die Schlussfolgerungen des Rates zur Reform der Beihilfenkontrolle vom 13. November 2012.

7 Im Rahmen des Monitorings prüft die Kommission in regelmäßigen Abständen eine Stichprobe von Beihilfemaßnahmen, die von den Mitgliedstaaten durchgeführt wurden, auf ihre Rechtmäßigkeit. Dies soll gewährleisten, dass die Mitgliedstaaten die Beschlüsse der Kommission korrekt umsetzen und die einschlägigen Rechtsvorschriften (der Allgemeinen Gruppenfreistellungsverordnung) einhalten. Bei einer repräsentativen Stichprobe von Fällen überprüft die Kommission auch die Einhaltung der Ex-ante-Vorschriften und Voraussetzungen.

Die Evaluierung staatlicher Beihilfen sollte insbesondere eine Bewertung des unmittelbaren Anreizeffekts der Beihilfe auf den Empfänger ermöglichen (d. h. ob die Beihilfe den Empfänger zu einer Verhaltensänderung veranlasst hat und wie groß die Auswirkungen der Beihilfe waren). Sie sollte auch Angaben zu den allgemeinen positiven und negativen Effekten der Beihilferegulung auf die Verwirklichung des angestrebten politischen Ziels sowie auf Wettbewerb und Handel liefern und könnte eine Prüfung der Geeignetheit und Angemessenheit des gewählten Beihilfeinstruments umfassen.

Auf dieser Grundlage kann die Evaluierung bestätigen, ob die bei der Ex-ante-Genehmigung der Beihilferegulung zugrunde gelegten Annahmen noch zutreffen, und dazu beitragen, die Gestaltung künftiger Beihilferegulungen und -vorschriften zu verbessern. Sie könnte die Basis für eine Anpassung künftiger staatlicher Interventionen bilden, um die Wirksamkeit und Effizienz der Beihilfen so weit zu verbessern, dass gewährleistet ist, dass die positiven Auswirkungen ausreichen, um die Inkaufnahme der interventionsbedingten Marktverzerrung zu rechtfertigen. Solche Verbesserungen könnten von der Anpassung der Ausgestaltung, einschließlich einer Änderung der Auswahlkriterien und einer umfassenderen Bewertung des Anreizeffekts, bis zu tiefer greifenden Änderungen wie der Förderung des Einsatzes einer anderen Form der Beihilfe, der Neufestlegung der Ziele oder Beihilfeempfänger oder der Prüfung beihilfefreier Optionen für die Verwirklichung derselben politischen Ziele reichen.

Es ist wichtig, einen angemessenen zeitlichen Rahmen für die Evaluierung mit genügend Zeit für die Sammlung ausreichender Belege festzulegen, gleichzeitig jedoch die Ergebnisse den politischen Entscheidungsträgern so bald wie möglich zur Verfügung zu stellen, damit potenzielle Verbesserungen zeitnah eingeführt werden können.⁹ Die Evaluierung staatlicher Beihilfen sollte daher nicht ausschließlich als Ex-post-Evaluierung nach Abschluss der Durchführung der Beihilferegulung vorgenommen werden, sondern in der Regel als laufende Evaluierung während ihrer Durchführung. Dabei ist zu berücksichtigen, dass in bestimmten Fällen die Auswirkungen der Intervention erst längerfristig in ihrem vollen Ausmaß erkennbar sind und die Evaluierung nur die ersten Auswirkungen erfassen und messen kann.

Die Evaluierung staatlicher Beihilfen sollte letztlich ein Lernprozess sowohl für die Kommission als auch die Mitgliedstaaten sein. Zu diesem Zweck muss die Evaluierung bestimmte Mindestqualitätsstandards erfüllen. Die Kommission sollte deshalb dafür sorgen, dass eine geeignete Qualitätskontrolle der Evaluierungen sichergestellt ist. Insbesondere wird die Kommission die Gesamtzuverlässigkeit der Evaluierung eingehend prüfen und auf mögliche Schwachstellen in den beiden entscheidenden Phasen (Evaluierungsplan und

8 Die Jahresberichte der Mitgliedstaaten dienen hauptsächlich dazu, die Transparenz der von den Mitgliedstaaten gewährten staatlichen Beihilfen zu erhöhen. Sie sind auch eine Quelle verlässlicher Statistiken für die Politikgestaltung und für Monitoringzwecke. Die in den Jahresberichten enthaltenen Daten liefern in erster Linie Informationen quantitativer Art (zum Beispiel, welche Ziele mit den staatlichen Beihilfen verfolgt wurden und mit welcher Mittelausstattung). Anhand der Berichte der Mitgliedstaaten erstellt die Kommission den Anzeiger für staatliche Beihilfen.

9 In einigen Leitlinien für staatliche Beihilfen ist für die Evaluierung von Beihilferegulungen ein Regelzeitrahmen von vier Jahren vorgesehen.

Abschlussbericht) hinweisen. In geeigneten Fällen könnte sich die Kommission von externen unabhängigen Sachverständigen bei der Qualitätskontrolle unterstützen lassen.

Die Kommission könnte auch Schulungen und Workshops zum Thema Evaluierungsmethoden und -techniken für die nationalen Verwaltungen organisieren. Ferner könnten positive Erfahrungen und bewährte Methoden aus den Mitgliedstaaten ausgetauscht und genutzt werden, um Beihilferegulungen in Zukunft wirksamer zu gestalten.

Der Nutzen der Evaluierung wird sich in einigen Jahren zeigen, wenn die ersten Evaluierungsberichte vorliegen und die entsprechenden Feststellungen und Empfehlungen zur Verfügung stehen. Diese können dann genutzt werden, um die Ausgestaltung künftiger Beihilferegulungen und möglicherweise Vorschriften für staatliche Beihilfen zu verbessern. Mittel- bis langfristig könnte die Evaluierung schrittweise zu grundlegenden Veränderungen in der allgemeinen Haltung zu staatlichen Beihilfen führen.

3 Der Evaluierungsplan

Es ist unerlässlich, dass parallel zur Ausarbeitung einer Beihilferegelung möglichst früh ein umfassender Plan für deren Evaluierung entworfen wird. Die Genehmigung des Evaluierungsplans durch die Kommission ist für die Sicherstellung der Gleichbehandlung von entscheidender Bedeutung. Der genehmigte Plan muss konsequent umgesetzt werden.

Denn es wird allgemein anerkannt, dass Evaluierungen wirksamer sind, wenn sie sorgfältig geplant und rechtzeitig vorbereitet werden, insbesondere da dies die Erhebung geeigneter Daten erleichtert. Eine frühzeitige Planung dürfte den Ressourcenbedarf für die Evaluierung erheblich verringern und deren Qualität letztlich verbessern.

Der Evaluierungsplan, der von den Mitgliedstaaten nach den einschlägigen Vorschriften anzumelden ist, sollte mindestens Folgendes enthalten.

3.1 Ziele der zu evaluierenden Beihilferegelung

Bei der Evaluierung einer Beihilferegelung ist zunächst die zugrundeliegende „Interventionslogik“ der Regelung deutlich darzulegen und zu beschreiben, welche Bedürfnisse und Probleme mit der Regelung angegangen werden sollen, auf welche Beihilfeempfänger und Investitionen sie ausgerichtet ist, welche allgemeinen und besonderen Ziele verfolgt werden und welche Auswirkungen erwartet werden. Auch die wichtigsten Annahmen in Bezug auf externe Faktoren, die Einfluss auf die Regelung haben könnten, sollten erwähnt werden.

3.2 Die Evaluierungsfragen

Im Evaluierungsplan sollte der Gegenstand der Evaluierung festgelegt werden, d. h., er sollte genaue Fragen enthalten, auf die quantitative, belegbare Antworten gegeben werden können. Die Evaluierungsfragen, die sich auf die Auswirkungen der Beihilferegulung konzentrieren sollten, können die folgenden drei Ebenen betreffen:

1. Direkte Auswirkungen der Beihilfe auf die Empfänger, z. B.:

- Hat sich die Beihilfe in nennenswertem Maße auf das Verhalten der Beihilfeempfänger ausgewirkt? (Anreizeffekt)
- Hat sich die Beihilfe auf die Lage der Beihilfeempfänger ausgewirkt? (Hat sie beispielsweise ihre Wettbewerbsposition oder Ausfallwahrscheinlichkeit verändert?)
- Inwieweit hat die Beihilfe die erwarteten Auswirkungen konkret gezeigt?
- Hat sich die Beihilfe in unterschiedlicher Weise auf die Beihilfeempfänger ausgewirkt? (Zum Beispiel je nach Größe, Standort oder Wirtschaftszweig?)

2. Indirekte Auswirkungen der Beihilferegulung, z. B.:

- Hat die Regelung Spillover-Effekte auf die Tätigkeit anderer Unternehmen oder auf andere Regionen gehabt? Hat die Beihilfe Investitionen von Wettbewerbern verdrängt oder Tätigkeiten von benachbarten Standorten abgezogen?
- Hat die Regelung zur Verwirklichung des betreffenden politischen Ziels beigetragen?
- Kann der Gesamteffekt der Regelung auf Wettbewerb und Handel gemessen werden?

3. Geeignetheit und Angemessenheit der Beihilferegulung, z. B.:

- Stand die Beihilferegulung in einem angemessenen Verhältnis zu dem Problem, das mit ihr bewältigt werden sollte? Hätten die gleichen Effekte auch mit weniger Beihilfen oder einer anderen Beihilfeform erzielt werden können? (Zum Beispiel mit Darlehen statt Zuschüssen?)
- Wurde das wirksamste Beihilfeinstrument gewählt? Wären andere Beihilfeinstrumente oder Arten von Maßnahmen besser geeignet gewesen, um das angestrebte Ziel zu erreichen?

Im Rahmen der Evaluierung sollten die Auswirkungen der Beihilferegulung möglichst auf allen drei Ebenen bewertet und die einschlägigen Fragen zu den Zielen der Regelung beantwortet werden. Am zuverlässigsten lassen sich jedoch im Allgemeinen die direkten Auswirkungen der Beihilfe auf die Empfänger messen. In der Praxis sind die meisten

Evaluierungsmethoden für die Bewertung dieser Auswirkungen entwickelt worden. Zudem ist die Evaluierung der direkten Auswirkungen der Beihilfe, einschließlich des Anreizeffekts, von größter Bedeutung, da durch sie wertvolle Erkenntnisse über die zu erwartenden indirekten Effekte und Verzerrungen gewonnen werden können. Insbesondere kann, wenn die Beihilfe keinen Anreizeffekt aufweist, davon ausgegangen werden, dass sie insofern verzerrend wirkt, als sie den betreffenden Empfängern unerwartete Gewinne beschert.

3.3 Ergebnisindikatoren

Die Evaluierungsfragen bestimmen die Wahl spezifischer Ergebnisindikatoren, die quantifizierte Angaben zu den mit der Beihilferegulung erzielten Ergebnissen erfassen sollen. Anhang II enthält eine beispielhafte, nicht erschöpfende Aufstellung von Ergebnisindikatoren für die direkten und die indirekten Auswirkungen einer Regelung, einschließlich der möglichen Effekte auf Wettbewerb und Handel. Die Ergebnisindikatoren hängen von den Zielen der zu evaluierenden Beihilfe ab. Im Evaluierungsplan sollte erläutert werden, warum die gewählten Indikatoren am besten für die Messung der Auswirkungen der betreffenden Beihilferegulung geeignet sind.

3.4 Methoden: Ermittlung einer geeigneten Vergleichsgrundlage

Mit der Evaluierung staatlicher Beihilfen sollte der von der Regelung selbst verursachte (kausale) Effekt ohne Verzerrung durch andere Variablen ermittelt werden können, die sich auf das beobachtete Ergebnis ausgewirkt haben könnten, z. B. die allgemeinen makroökonomischen Bedingungen oder Unterschiede zwischen den Unternehmen (wie Größe, Standort, Finanzmittel oder Managementfähigkeiten). Im Evaluierungsplan sollte dargelegt werden, welche Methoden im Wesentlichen angewandt werden, um die Wirkung der Beihilfe zu ermitteln, und warum sich diese Methoden am ehesten für die betreffende Regelung eignen.

Der kausale Effekt ist der Unterschied zwischen dem Ergebnis mit der Beihilfe und dem Ergebnis ohne die Beihilfe. Während das Ergebnis mit der Beihilfe bei den Unternehmen, die die Beihilfe erhalten haben, beobachtet wird, wird das Ergebnis ohne die Beihilfe nur bei den Unternehmen gemessen, die keine Beihilfe erhalten haben. Per definitionem kann bei den Unternehmen, die die Beihilfe erhalten haben, nicht beobachtet werden, welches Ergebnis ohne die Beihilfe vorgelegen hätte. Um die Wirkung der Beihilfe auf die Beihilfeempfänger schätzen zu können, muss daher anhand der mit diesen am ehesten vergleichbaren Unternehmen oder Kontrollgruppen eine kontrafaktische Fallkonstellation konstruiert werden.

Die Qualität der Kontrollgruppe ist für die Aussagekraft der Evaluierung von entscheidender Bedeutung.

Die Unternehmen, die Beihilfen erhalten, befinden sich möglicherweise in einer anderen Lage als die Unternehmen, die keine Beihilfen erhalten. So können sie unterschiedlichen lokalen Angebots- und Nachfragebedingungen ausgesetzt sein, weniger leicht Zugang zu Krediten haben oder auch effizienter bzw. ineffizienter arbeiten. All diese Faktoren können Auswirkungen auf die Leistung oder auf den Umfang der Tätigkeit der Unternehmen haben,

unabhängig davon, ob sie Beihilfen erhalten oder nicht. Ein Vergleich der Leistung von Beihilfeempfängern und Nichtempfängern dürfte diese Gegebenheiten eher widerspiegeln als die Wirkung der Beihilfe selbst. Eine Evaluierung der Beihilferegelung kann sich daher nicht auf einen einfachen Vergleich zwischen Beihilfeempfängern und Nichtempfängern stützen, sondern muss die unterschiedlichen Merkmale der beiden Gruppen von Unternehmen berücksichtigen, sowohl diejenigen, die beobachtet werden können, als auch diejenigen, die nicht beobachtet werden können.

So sind im Falle von Regionalbeihilfen die Ergebnisse der Beihilfeempfänger in Gebieten mit ungünstigen Marktbedingungen (d. h. schwachen lokalen Produkt-, Arbeits- oder Kapitalmärkten) typischerweise schlechter als die von Nichtempfängern in wohlhabenderen Gebieten. Dies gibt jedoch in keiner Weise die Wirkung der Beihilfe selbst wieder. Die entscheidende Frage ist, ob ihre Ergebnisse besser waren, als sie ohne die Beihilfe gewesen wären, und nicht, ob ihre Ergebnisse besser waren als die von Nichtempfängern in anderen Gebieten.

Auch allgemeine Branchentrends müssen bei der Ermittlung der Wirkung der Beihilfe berücksichtigt werden. Die Beihilfe kann selbst dann wirksam gewesen sein, wenn Empfänger von Regionalbeihilfen Personal entlassen. Wenn sich beispielsweise die Bedingungen in einer Branche insgesamt verschlechtern und alle Unternehmen Arbeitsplätze abbauen, kann es sein, dass Beihilfeempfänger weniger Stellen streichen, als sie es sonst getan hätten. Dies wird in der nachstehenden Grafik veranschaulicht, in der ein negativer Trend bei der Zahl der von Beihilfeempfängern bereitgestellten Arbeitsplätze vor und nach Gewährung der Beihilfe dargestellt ist. Der Trend ist weniger negativ, nachdem das Unternehmen die Beihilfe erhalten hat. Die Differenz zwischen der verlängerten Linie für den Trend ohne Beihilfe und der Linie für die Zahl der Arbeitsplätze, die von dem Unternehmen nach Erhalt der Beihilfe tatsächlich angeboten werden, zeigt den positiven Einfluss der Beihilfe.

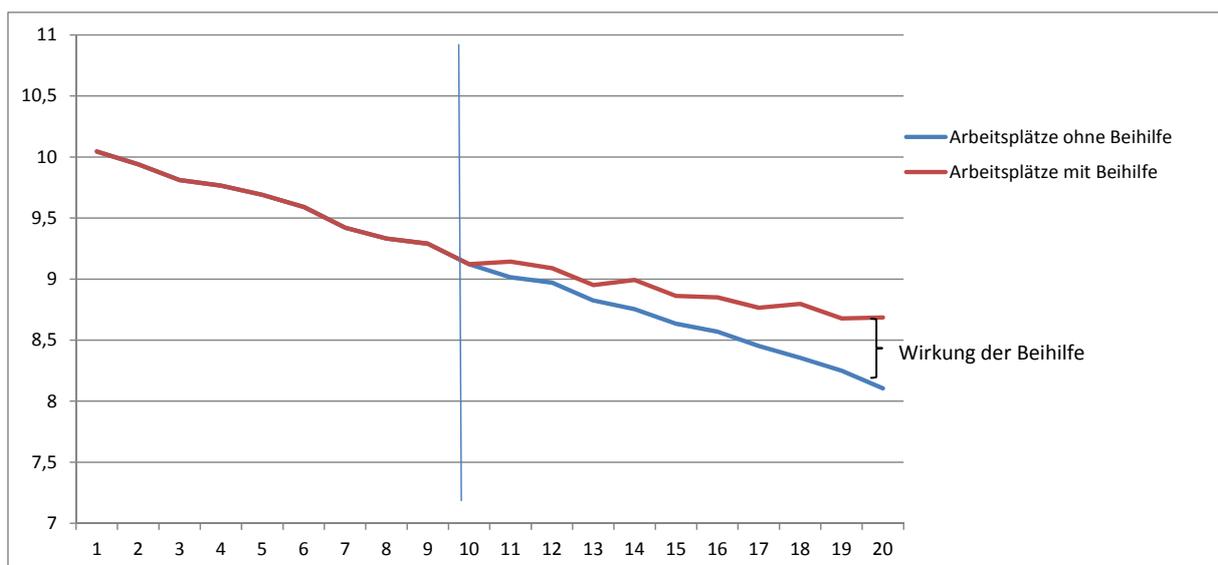


Schaubild 1 – Positiver Einfluss der Beihilfe bei aktuell negativem Trend

Ein besonderes Problem bei der Ermittlung einer Kontrollgruppe entsteht, wenn Nichtempfänger selbst beschlossen haben, eine Beihilfe zu beantragen bzw. keine Beihilfe zu beantragen. Wenn zum Beispiel alle Unternehmen für eine Beihilfe in Betracht kommen (d. h. alle Unternehmen, die ein Vorhaben vorschlagen und eine Beihilfe beantragen, erhalten tatsächlich Beihilfen), sind die Unternehmen, die keine Beihilfe beantragen, wahrscheinlich diejenigen, die über kein entsprechendes Vorhaben verfügen. Die Ergebnisse der Unternehmen zeigen möglicherweise, dass die Leistung der Unternehmen, die keine Beihilfen erhielten, absolut und relativ gesehen schlechter war als die der Unternehmen, denen Beihilfen gewährt wurden. Dies könnte jedoch ausschließlich auf die bloße Tatsache zurückzuführen sein, dass die eine Gruppe im Gegensatz zu der anderen kein Vorhaben hatte (d. h. nicht über genügend Interesse oder Kreativität verfügte). Von entscheidender Bedeutung ist daher, dass die Unternehmen in der Kontrollgruppe (denen keine Beihilfen gewährt wurden) dieser Gruppe aus Gründen angehören, die keinen Einfluss auf die gemessenen Ergebnisse haben. Insbesondere wenn Unternehmen von sich aus freiwillig entschieden haben, keine Beihilfe zu beantragen, ist diese Voraussetzung möglicherweise nicht erfüllt.

Systematische Unterschiede zwischen Beihilfeempfängern und Nichtempfängern sollten in geeigneter Weise bei der Konzipierung der Evaluierung berücksichtigt werden, um eine Verfälschung der Ergebnisse (auswahlbedingte Verzerrung) zu verhindern. In den letzten Jahrzehnten sind hierzu mehrere zuverlässige Methoden entwickelt worden. Für welche Methode man sich entscheidet, hängt von der Ausgestaltung der betreffenden Beihilferegulung und den verfügbaren Daten ab. Jede Methode hat ihre Grenzen und funktioniert nur, wenn bestimmte Annahmen zutreffen. Für die Glaubwürdigkeit einer Studie ist von entscheidender Bedeutung, dass diese Grenzen und Annahmen offen benannt und erörtert werden.

Eine Möglichkeit, eine unverfälschte Evaluierung zu gewährleisten, ist die Randomisierung der Auswahl der Beihilfeempfänger. Wenn die Beihilfeempfänger ausschließlich nach dem Zufallsprinzip ausgewählt werden, können beobachtete systematische Unterschiede in der Leistung der Unternehmen der Beihilfe zugeschrieben werden. Diese Methode kann jedoch in der Praxis schwer anzuwenden sein, insbesondere im Falle umfangreicher bestehender Regelungen. Bei anderen Methoden werden bestehende Quellen exogener Variation in dem Umfeld, in dem die Unternehmen tätig sind (d. h. einer Variation, die nicht durch Parameter und Variablen im Modell bestimmt wird), herangezogen, um die Kausalität zu ermitteln.¹⁰ In Anhang I dieses Leitfadens werden die wichtigsten Methoden ausführlich vorgestellt und vor allem die praktischen Aspekte ihrer Anwendung behandelt. Es wird darauf eingegangen, wie die Kausalität mit der betreffenden Methode ermittelt wird; besonders wichtig ist dies bei Evaluierungen staatlicher Beihilfen, bei denen die Ex-ante-Konzipierung der Evaluierung dazu dient sicherzustellen, dass eine ordnungsgemäße Evaluierung der Auswirkungen der Beihilfe möglich ist.

¹⁰ Die gebräuchlichsten Methoden sind das Difference-in-Difference-Verfahren, die Regressions-Diskontinuitäts-Analyse und die Verwendung von Instrumentvariablen.

Außerdem sollte geprüft werden, wie sich mehrere Beihilfen auswirken, die auf der Grundlage einer Regelung, mehrerer Regelungen oder als Ad-hoc-Beihilfen gewährt wurden. Wenn Unternehmen keine Beihilfen auf der Grundlage der betreffenden Regelung, aber auf der Grundlage anderer Regelungen erhalten haben oder wenn Unternehmen Beihilfen auf der Grundlage der betreffenden Regelung und zusätzlich auch auf der Grundlage anderer Regelungen erhalten haben, ist eine Verzerrung der Evaluierung der Auswirkungen der betreffenden Regelung wahrscheinlich.

3.5 Datenerhebung: Nutzung der bestmöglichen Quellen

Sowohl zu den Beihilfeempfängern als auch zur Kontrollgruppe sind in ausreichendem Maße konsistente Daten zu erheben. Im Rahmen der Planung der Evaluierung müssen die benötigten Daten ermittelt und der Zugang zu den Datenquellen gewährleistet werden.

Ein wirksames Monitoring der Maßnahme und eine genaue Erhebung und Verarbeitung der Daten sind für die Qualität der Evaluierung von entscheidender Bedeutung. Sobald die Beihilferegulierung genehmigt ist, sollte daher ein Mechanismus für das Monitoring der Maßnahme und die Erhebung und Verarbeitung geeigneter Daten eingerichtet werden. Dadurch dürften sich die Kosten der Evaluierung erheblich verringern.

Die Erhebung der erforderlichen Daten zu den Antragstellern und den Beihilfeempfängern sicherzustellen, ist ein entscheidender Schritt bei der Ausarbeitung des Evaluierungsplans, wenn für die Gewährung der Beihilfe die Verfügbarkeit dieser Daten als Voraussetzung festgelegt werden kann.

Außer den Daten zu den Beihilfeanträgen (einschließlich der abgelehnten Antragsteller, wenn vorhanden) müssen die Datenquellen für die Beihilfeempfänger und für die Kontrollgruppe dieselben sein, damit die Daten vergleichbar sind. Sehr wahrscheinlich müssen mehrere Quellen genutzt und zum Beispiel auch Daten aus Datenbanken mit Informationen über die Beihilfeempfänger mit Daten aus Unternehmensregistern kombiniert werden. Möglicherweise müssen bei der Evaluierung vorhandene Datenquellen (z. B. Steuerbehörde, Handelsregister, Innovationserhebungen oder Patentamt) herangezogen werden. Im Evaluierungsplan ist daher zu prüfen, welche Datenquellen bestehen, um zu entscheiden, ob sie ausreichende Informationen für die Evaluierung liefern, und um zu gewährleisten, dass innerhalb des betreffenden Zeitrahmens auf sie zugegriffen werden kann.

Daten aus Verwaltungsquellen, z. B. Daten der nationalen statistischen Ämter, werden den Evaluierern wahrscheinlich nur unter bestimmten Bedingungen in Bezug auf den Datenschutz und die Vertraulichkeit von Geschäftsdaten zur Verfügung gestellt. Die Bedingungen für den Zugang zu diesen Daten sind im Evaluierungsplan zu nennen. Die Behörde, die den Zugang zu den Daten gewährt, muss erforderlichenfalls dafür sorgen, dass die Sachverständigen, die die Evaluierung vornehmen, Zugang zu diesen Daten erhalten.

Wenn Daten aus mehreren Quellen verwendet werden, ist von größter Bedeutung, dass sie in einem Format erhoben werden, das eine konsistente Gegenüberstellung der Variablen ermöglicht. Es könnte notwendig sein, eindeutige Kennzeichner für die

Beobachtungseinheiten in jedem verwendeten Datensatz zu finden. So müssen die Unternehmens- und Betriebskennzeichner in allen Datensätzen eindeutig sein, die Anschriften müssen in einem Format erhoben werden, das eine geografische Ortung ermöglicht, usw. Der genaue Ursprung des Kennzeichners kann sich von Mitgliedstaat zu Mitgliedstaat unterscheiden. Er könnte beispielsweise steuerlichen Ursprungs sein (wie die Umsatzsteuer-Identifikationsnummer) oder direkt von den statistischen Ämtern stammen (wie SIREN und SIRET in Frankreich; dies sind die Identifikationsnummern für Unternehmen und Einrichtungen, die vom Nationalen Institut für Statistik und Wirtschaftsplanung (INSEE) vergeben werden).

Die Evaluierung staatlicher Beihilfen könnte durch Informationen aus Erhebungen bei den Beihilfeempfängern und/oder Befragungen der Verwalter der Regelungen ergänzt werden. Qualitative Informationen dieser Art sind natürlich subjektiv, und die Antworten entsprechen möglicherweise eher den strategischen Interessen der Beihilfeempfänger als einer echten Bewertung der Wirkung der Beihilfe. Diese Gefahr ist besonders groß, wenn der Befragte davon ausgeht, dass eine positive Aussage seine Chancen verbessert, in Zukunft Beihilfen zu erhalten. Wenn Informationen aus qualitativen Verfahren wie Befragungen und Fallstudien mit der gebotenen Vorsicht behandelt werden, können sie jedoch eine nützliche ergänzende Quelle sein und bei der Auswertung der Ergebnisse der Evaluierung helfen.

Für die Verarbeitung personenbezogener Daten im Rahmen der Evaluierungen gelten die Rechtsvorschriften der Union über den Schutz personenbezogener Daten, insbesondere die Richtlinie 95/46/EG zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr und die zu ihrer Umsetzung erlassenen nationalen Rechtsvorschriften sowie die Verordnung (EG) Nr. 45/2001 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten durch die Organe und Einrichtungen der Gemeinschaft und zum freien Datenverkehr.

3.6 Zeitlicher Rahmen für die Evaluierung

Der Evaluierungsplan sollte Angaben zum genauen zeitlichen Rahmen für die Evaluierung, der im Einklang mit der genehmigten Laufzeit der Regelung festzulegen ist, sowie feste Termine für die Erhebung der Daten, die Durchführung der Evaluierung und die Vorlage des Abschlussberichts enthalten. Der zeitliche Rahmen richtet sich nach der jeweiligen Regelung und sollte daher im Einzelfall mit der Kommission erörtert und vereinbart werden. Den an der Verwaltung von Regelungen Beteiligten wird empfohlen, möglichst auch informelle Gespräche über den Inhalt des Plans zu führen, bevor sie ihre offizielle Anmeldung bei der Kommission einreichen.

Damit eine geplante Verlängerung einer Beihilferegelung geprüft werden kann, sollte der Kommission der abschließende Evaluierungsbericht rechtzeitig (z. B. sechs Monate vor dem ursprünglich vorgesehenen Ende der Laufzeit der Regelung) vorgelegt werden. Ist keine Verlängerung geplant, so kann der Bericht nach Auslaufen der Regelung vorgelegt werden.

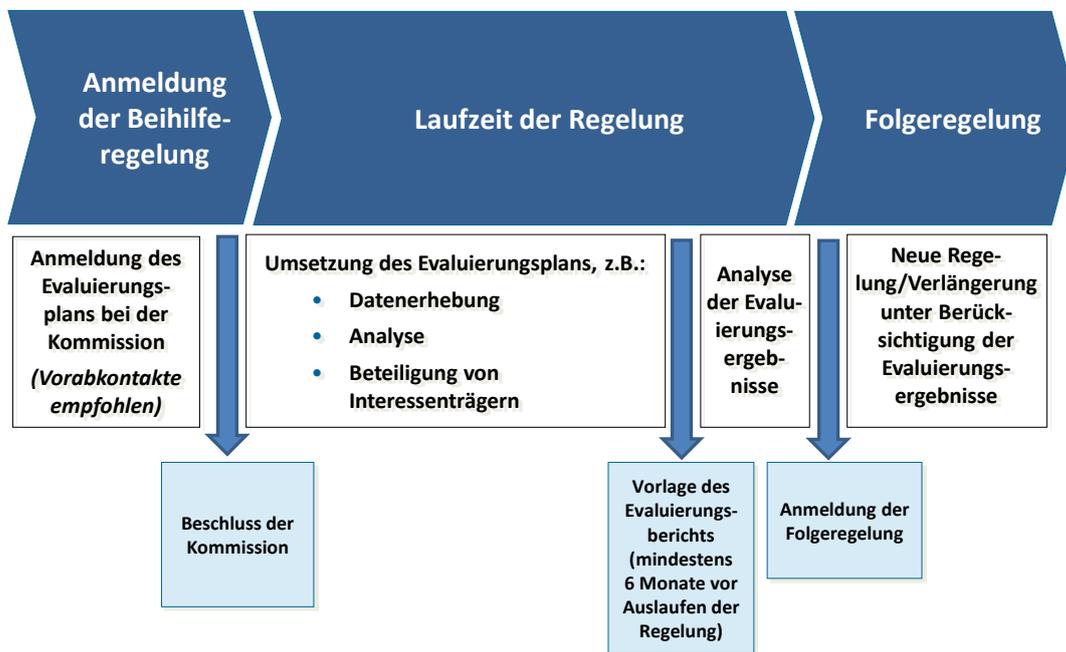


Schaubild 2 – Überblick über den Evaluierungsprozess im Falle einer angemeldeten Regelung

3.7 Das Evaluierungsgremium: Sicherstellung von Unabhängigkeit und Fachkompetenz

Die Auswirkungen von Beihilferegelungen sollten objektiv, gründlich, unparteiisch und transparent evaluiert werden.¹¹ Jede Evaluierung sollte von Sachverständigen, die über die nachgewiesene einschlägige Erfahrung und die Methodenkenntnis verfügen, die für die Wahrnehmung dieser Aufgabe erforderlich sind, auf der Grundlage solider Methoden vorgenommen werden.

Die Evaluierungen sollten von einem Gremium vorgenommen werden, das von der die Beihilfen gewährenden Behörde zumindest funktional unabhängig ist und über die für die Vornahme solcher Evaluierungen erforderliche nachgewiesene Kompetenz und angemessen qualifiziertes Personal verfügt. Die funktionale Unabhängigkeit des Evaluierers von der Bewilligungsbehörde ist für die Qualität und die Glaubwürdigkeit der Evaluierung von entscheidender Bedeutung. Dies heißt aber nicht zwangsläufig, dass ein neues Gremium eingerichtet oder die Evaluierung an Evaluierungsunternehmen vergeben werden müsste. Je

¹¹ Siehe zum Beispiel die Evaluierungsnormen der Europäischen Kommission, die *Evaluation Norms and Standards* der OECD, die *Evaluation Standards* der Vereinten Nationen und *Independent Evaluation: Principles, Guidelines and Good Practice* der Weltbank.

nachdem, welche Organisationen in einem Mitgliedstaat existieren, könnten beispielsweise die Unabhängigkeit und Kompetenz des statistischen Amtes, der Zentralbank, des Rechnungshofs, öffentlicher oder privater Hochschulen oder Forschungszentren oder ähnlicher Organisationen genutzt werden. Dies kann im Einzelfall für jede Regelung entschieden werden.

Die frühzeitige Einbeziehung des Evaluierungsgremiums, zum Beispiel in die Ausarbeitung der Regelung, ist für den Erfolg einer Evaluierung wichtig. Sie gewährleistet, dass die Beihilferegelung wie vorgesehen evaluiert werden kann und dass die notwendigen Daten erhoben werden. Daher sollte der Evaluierungsplan möglichst von oder zumindest in sehr enger Zusammenarbeit mit dem vorgesehenen Evaluierer entworfen werden. Der Evaluierungsplan sollte Angaben – gegebenenfalls auch nur in Form von Richtwerten – zu den notwendigen personellen und finanziellen Ressourcen enthalten, die für die Evaluierung bereitgestellt werden sollten. Besonders wichtig sind Informationen über die Identität und Rolle aller an der Evaluierung beteiligten Hauptsachverständigen sowie eine Schätzung des Umfangs ihrer Beteiligung.

Im Evaluierungsplan sollte das Evaluierungsgremium genau beschrieben werden; falls dieses noch nicht eingesetzt wurde, sind die Kriterien für seine Auswahl ausführlich darzulegen, insbesondere in Bezug auf Unabhängigkeit, Erfahrung und Kompetenz. Auch bestehende Alternativen sollten nach Möglichkeit angegeben werden. Wenn der Evaluierer noch nicht ausgewählt wurde oder zwar ausgewählt wurde, aber nicht aktiv an der Ausarbeitung des Evaluierungsplans teilgenommen hat, sind die Gründe hierfür klar anzugeben. Auch in diesem Fall muss der Evaluierungsplan so detailliert sein, dass die Aussagekraft der Evaluierung ordnungsgemäß geprüft werden kann.

3.8 Veröffentlichung: Erleichterung der Einbeziehung von Interessenträgern

Die Evaluierung sollte veröffentlicht werden. Dies bedeutet, dass sowohl der Evaluierungsplan als auch der abschließende Evaluierungsbericht nach ihrer Genehmigung in geeigneter Weise öffentlich bekanntgemacht und an den im Evaluierungsplan angegebenen Stellen, zum Beispiel auf einer Website, zur Verfügung gestellt werden sollten. Die Kommission könnte diese Unterlagen ebenfalls veröffentlichen.¹²

Wenn für die Evaluierung personenbezogene und/oder vertrauliche Daten verwendet werden, muss ihre Vertraulichkeit im Einklang mit den Artikeln 8, 16 und 17 der Charta der Grundrechte der Europäischen Union während des ganzen Evaluierungsprozesses garantiert sein. Die Vertraulichkeit erstreckt sich jedoch nicht auf die Ergebnisse der Evaluierung. Insbesondere darf in den Vertrag über die Evaluierung keine Vertraulichkeitsklausel

¹² Mit Ausnahme von Geschäftsgeheimnissen und sonstigen vertraulichen Informationen in hinreichend begründeten Fällen (Mitteilung der Kommission C(2003) 4582 zum Berufsgeheimnis in Beihilfeentscheidungen, ABl. C 297 vom 9.12.2003, S. 6). Bei einer Veröffentlichung personenbezogener Daten müssen die Rechtsvorschriften der Union über den Schutz personenbezogener Daten, insbesondere die Richtlinie 95/46/EG und die zu ihrer Umsetzung erlassenen nationalen Rechtsvorschriften sowie die Verordnung (EG) Nr. 45/2001 eingehalten werden.

aufgenommen werden, abgesehen von 1. Geheimhaltungspflichten für personenbezogene und/oder vertrauliche Daten und 2. Pflichten zur Einhaltung der allgemeinen Bestimmungen des nationalen Statistikrechts und der statistischen Geheimhaltung zum Beispiel in Bezug auf die Darstellung der Ergebnisse.

Die während der Evaluierung erhobenen Daten sollten zugänglich gemacht werden, damit die Ergebnisse nachvollzogen oder für weitere Studien genutzt werden können, ohne dass die Bedingungen restriktiver sind als für das ursprüngliche Evaluierungsgremium.

Die Bewilligungsbehörde könnte für eine angemessene Einbeziehung relevanter Interessenträger sorgen, die während der Umsetzung des Evaluierungsplans zumindest einmal gehört werden sollten. So könnten Interessenträger zu einer Erörterung erster Ergebnisse der Evaluierung auf der Grundlage eines Zwischenberichts eingeladen werden. Entsprechende Vorkehrungen sollten in den Evaluierungsplan aufgenommen werden.

4 Kriterien für die Auswahl der zu evaluierenden Beihilferegelungen

Grundsätzlich kommt jede Beihilferegelung für eine Evaluierung in Betracht; die Evaluierung wird zwar als bewährte Praxis angesehen, ist aber nicht in allen Fällen beihilferechtlich vorgeschrieben. Bei der Evaluierung staatlicher Beihilfen sollte die Verhältnismäßigkeit gewahrt bleiben; evaluiert werden sollten im Allgemeinen Regelungen mit potenziell erheblichen Auswirkungen auf den Binnenmarkt, bei denen die Gefahr erheblicher Verfälschungen besteht, wenn ihre Durchführung nicht rechtzeitig überprüft wird. Die einschlägigen Leitlinien für staatliche Beihilfen konzentrieren sich daher auf 1) umfangreiche Beihilferegelungen einschließlich solcher, die unter die Allgemeine Gruppenfreistellungsverordnung fallen, 2) neuartige Beihilferegelungen und 3) Beihilferegelungen, bei denen in naher Zukunft wesentliche Veränderungen (hinsichtlich Markt, Technologie oder Regulierung) möglich sind, die eine Überprüfung der Würdigung der Regelung erfordern könnten. In den einzelnen Leitlinien für staatliche Beihilfen sind weitere Arten von Beihilfen aufgeführt, bei denen eine Evaluierung von Vorteil wäre.

4.1 Umfangreiche Beihilferegelungen einschließlich solcher, die unter die Allgemeine Gruppenfreistellungsverordnung fallen

Im Einklang mit der Mitteilung über die Modernisierung des EU-Beihilferechts könnte die Kommission aus folgenden Gründen eine Evaluierung der umfangreichsten Beihilferegelungen verlangen: 1) Diese Regelungen können die stärksten Auswirkungen auf den Binnenmarkt haben, wenn sie nicht gut konzipiert sind, 2) mit diesen Regelungen können wegen ihrer hohen Mittelausstattung die größten Effizienzsteigerungen erzielt werden, und 3) bei umfangreichen Regelungen mit vielen verschiedenen Arten von Beihilfeempfängern liegen genügend Daten für eine Evaluierung vor.

Wenn bestimmte Beihilferegelungen trotz ihres Umfangs keine besonderen problematischen Aspekte aufweisen, müssen sie allerdings nicht evaluiert werden (z. B. Routinefälle, Fälle, in denen eine hohe Zahl von Beihilfeempfängern jeweils einen geringen Beihilfebetrag erhält,

oder Fälle, bei denen nicht die Gefahr wesentlicher Veränderungen besteht oder nicht mit erheblichen Verfälschungen zu rechnen ist).

In der neuen Allgemeinen Gruppenfreistellungsverordnung (AGVO) ist der Begriff „umfangreiche Beihilferegelung“ auf der Grundlage der Mittelausstattung definiert (durchschnittliche jährliche Mittelausstattung von mehr als 150 Mio. EUR) und für bestimmte Gruppen von Beihilfen¹³ eine Evaluierung vorgesehen.

Damit sich das Inkrafttreten dieser umfangreichen Regelungen nicht verzögert, gleichzeitig jedoch eine wirksame Evaluierung gewährleistet ist, sieht die AGVO eine Freistellung von der Anmeldepflicht für einen Zeitraum von höchstens sechs Monaten vor, der von der Kommission bei Genehmigung des Evaluierungsplans verlängert werden kann.¹⁴ Der Evaluierungsplan sollte so bald wie möglich, spätestens aber 20 Arbeitstage nach Inkrafttreten der Regelung angemeldet werden.

Nach der neuen AGVO müssen auch Änderungen und Folgeregelungen zu diesen evaluierungspflichtigen umfangreichen Regelungen angemeldet werden, es sei denn, es handelt sich um rein formale Änderungen, administrative Änderungen oder Änderungen, die im Rahmen der von der Union kofinanzierten Maßnahmen vorgenommen werden.

4.2 Neuartige Beihilferegelungen

Die „Neuartigkeit“ kann in den verschiedenen Beihilfeinstrumenten und Mitgliedstaaten unterschiedlich definiert sein. Für die Neuartigkeit wird grundsätzlich die Art der Beihilferegelung oder der Zielmarkt maßgebend sein, z. B. aufstrebende Märkte, die sich in einem sehr frühen Stadium der Marktentwicklung befinden. Diese Regelungen verfügen über das Potenzial, Wirtschaftszweige dauerhaft und grundlegend zu prägen. Sowohl die Vorteile als auch die Verfälschungen können daher besonders groß ausfallen. Als Neuartigkeit käme zum Beispiel Folgendes in Betracht: die Einführung eines neuen Kapazitätsmechanismus im Energiesektor, Beihilfen für neue Arten von Technologien oder eine neuartige Form der Förderung erneuerbarer Energiequellen im Rahmen der Umweltbeihilfen. Die Evaluierung neuartiger Regelungen hilft auch denjenigen, die neue Regelungen ausarbeiten, da sie auf diese Weise die jüngsten Entwicklungen auf dem Markt berücksichtigen können.

4.3 Beihilferegelungen, bei denen wesentliche Veränderungen abzusehen sind

Ob in naher Zukunft wesentliche Veränderungen (hinsichtlich Markt, Technologie oder Regulierung) möglich sind, muss im Einzelfall geprüft werden. Wesentliche Veränderungen könnten zum Beispiel die bevorstehende Überarbeitung einer geltenden Verordnung oder Beihilfen für sehr dynamische Wirtschaftszweige sein, in denen sich das Marktumfeld und

13 Regionalbeihilfen (ausgenommen regionale Betriebsbeihilfen), Beihilfen für KMU, Beihilfen zur Erschließung von KMU-Finanzierungen, Beihilfen für Forschung und Entwicklung und Innovation, Umweltschutzbeihilfen (ausgenommen Beihilfen in Form von Umweltsteuerermäßigungen nach der Richtlinie 2003/96/EG) und Beihilfen für Breitbandinfrastrukturen.

14 Die Kommission kann ausnahmsweise auch beschließen, dass wegen der Besonderheiten des Falles keine Evaluierung notwendig ist.

das Technologieangebot rasch weiterentwickeln. Wenn Regelungen nicht an die Auswirkungen dieser wesentlichen Veränderungen angepasst werden, besteht die Gefahr, dass öffentliche Mittel nicht wirksam eingesetzt werden (und z. B. Mittel für die Behebung eines potenziellen „Marktversagens“ bereitgestellt werden, das in Zukunft nicht mehr bestehen würde) oder dass erhebliche Verfälschungen auftreten, die sich auf neue Marktteilnehmer anders auswirken als auf etablierte Unternehmen oder zu ungleichen Bedingungen für neue und alte Technologien führen. Die Überarbeitung eines bestehenden Regulierungsrahmens (z. B. im Bereich der elektronischen Kommunikation), starke Schwankungen der Input- oder Outputpreise (z. B. bei Solarpaneelen) oder die Markteinführung einer neuen Technologie (z. B. die Bereitstellung eines Mobilfunknetzes der vierten Generation für Breitbanddienste) sind anschauliche Beispiele dafür, in welchen Fällen eine Evaluierung gerechtfertigt sein könnte, damit neue Marktentwicklungen in künftigen Regelungen berücksichtigt werden können.

4.4 Sonstige Beihilferegulungen

In den einschlägigen Leitlinien für die einzelnen Bereiche, in denen staatliche Beihilfen gewährt werden, sind auch Beihilferegulungen genannt, bei denen eine Evaluierung besonders sinnvoll wäre.

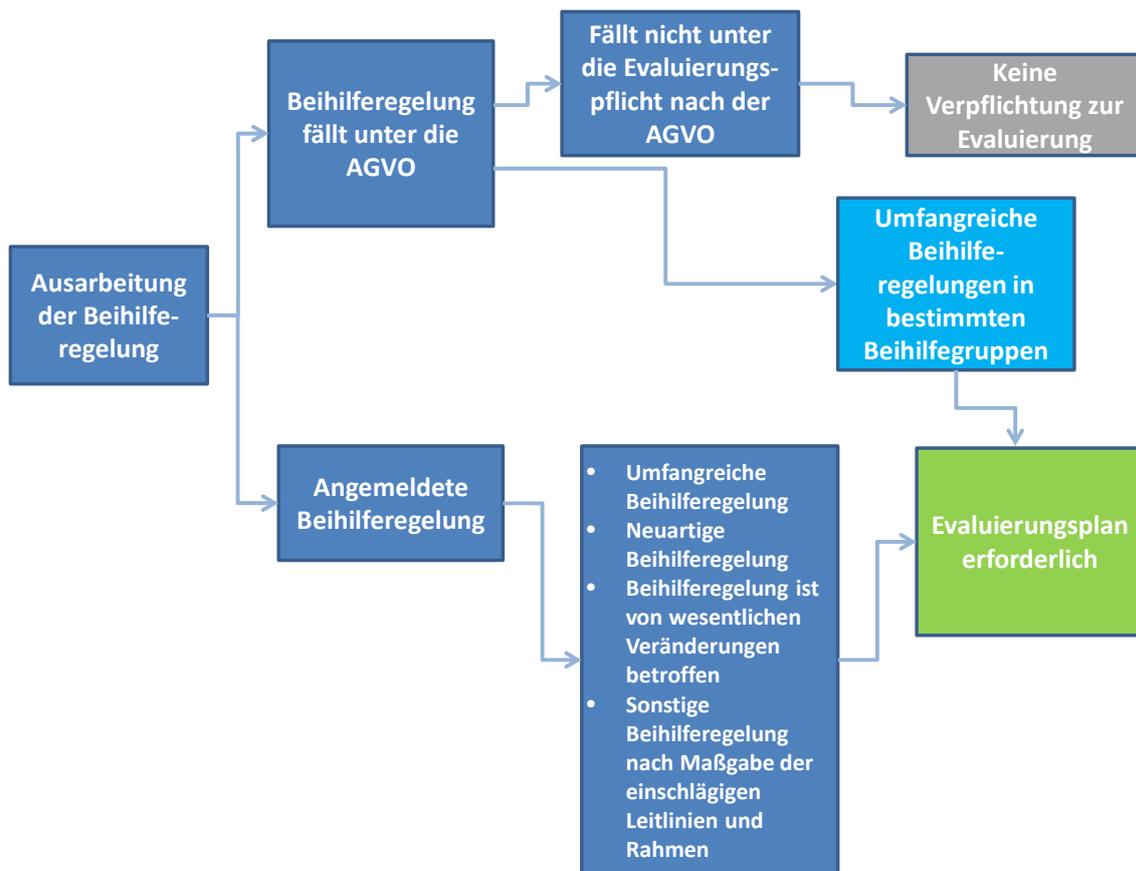


Schaubild 3 – Auswahl von Beihilferegelungen für die Evaluierung

Anhang I: Technischer Anhang zu den einschlägigen Methoden zur Ermittlung kausaler Effekte

Eine Beihilferegulation kann auf sehr unterschiedlichen Ebenen Auswirkungen zeigen. Normalerweise wird davon ausgegangen, dass sie sich direkt auf der Ebene der Beihilfeempfänger auswirkt. Für die Beurteilung der Effizienz und Wirksamkeit einer staatlichen Maßnahme ist es deshalb entscheidend, das Ausmaß dieser Auswirkungen zu erfassen. Da die Beihilfen jedoch direkt an Unternehmen gehen, die in Märkten oder Regionen tätig sind, die alle versuchen, möglichst viele Wirtschaftstätigkeiten für sich zu gewinnen, haben staatliche Beihilfen in der Regel auch indirekte Auswirkungen. Hierzu zählen zum Beispiel Spillover-Effekte auf andere Unternehmen (z. B. positive Spillovers aus der FuE oder eine Verdrängung von Investitionen anderer konkurrierender Unternehmen) und Verlagerungseffekte (z. B. Verlagerung einer wirtschaftlichen Aktivität von einer Region in eine andere). Diese indirekten Auswirkungen sind die Grundlage sowohl für eine potenzielle Beeinträchtigung als auch für potenzielle Vorteile, die einer Wirtschaft aus staatlicher Intervention erwachsen können. Deshalb muss eine Evaluierung öffentlicher Maßnahmen auch mit einer Beurteilung des Ausmaßes dieser indirekten Auswirkungen einhergehen.

Die Messung der direkten und indirekten Auswirkungen einer bestimmten Politik erfordert in der Regel die Verwendung verschiedener Instrumente. In den letzten Jahrzehnten wurden zahlreiche Methoden und Verfahren zur Bewertung der direkten Auswirkungen politischer Maßnahmen auf die Begünstigten entwickelt. Auf diese Verfahren wird in diesem Abschnitt ausführlich eingegangen. Leider können diese Verfahren nur sehr selten auch für die Bewertung der indirekten Auswirkungen von Beihilferegulationen auf Unternehmen oder Regionen herangezogen werden. Für die Evaluierung der indirekten Auswirkungen von Beihilferegulationen sind in der Regel andere Arten von Belegen erforderlich als für die Evaluierung der direkten Auswirkungen auf die Beihilfeempfänger, und die Auswertung erfolgt eher anhand ökonomischer Theorien und Modelle. Für diese Art der Evaluierung ist es schwieriger, präzise Hilfestellungen zu geben, da sie möglichst auf die potenziellen und zu erwartenden positiven und negativen Auswirkungen der jeweiligen Politik abstellen sollten. Deshalb kann eine solche Evaluierung erst nach sorgfältiger und gründlicher Analyse der indirekten Auswirkungen einer Politik, die am ehesten eintreten könnten, vorgenommen werden. Aus einer solchen Analyse können die Evaluierer anhand von Mikrodaten von Unternehmen, die keine Beihilfen erhalten haben, aber in derselben Region, im selben Cluster oder in derselben Branche oder aber in benachbarten Regionen tätig sind, Maßnahmen ableiten. Dies sollte den Kern für die Bewertung der indirekten Auswirkungen der Beihilferegulation bilden. Falls erforderlich können weitere, makroökonomische Daten, vor allem aber auch sorgfältig ausgewählte Fallstudien herangezogen werden.

Die Evaluierung der direkten Auswirkungen ist eine unverzichtbare und entscheidende erste Stufe. Eine gründliche Bewertung der indirekten Auswirkungen einer Beihilfe ist eine wichtige Komponente in der Bewertung der allgemeineren Auswirkungen einer Regelung. Wenn die Tatsache, dass Beihilfeempfänger zusätzlich zur Beihilfe keine weiteren

Investitionen tätigen, ganz allgemein ein Indiz dafür ist, dass die Politik fehlgeschlagen ist, dann lässt auch ein positiver Effekt nicht den Schluss zu, dass die gesetzten Ziele erreicht wurden. Sollte es sich zudem erweisen, dass sich eine Beihilfe kaum oder gar nicht auf die Empfänger ausgewirkt hat, so dürfte das Ergebnis der Evaluierung aller Wahrscheinlichkeit nach lauten, dass die Regelung ihren Zweck nicht erfüllt hat, es sei denn, es können überzeugende Argumente dafür vorgebracht werden, dass erhebliche und sehr positive indirekte Auswirkungen festzustellen sind. Der gegenteilige Fall ist ebenfalls denkbar: Selbst wenn bei der Evaluierung festgestellt wird, dass die direkten Auswirkungen der Beihilfe positiv sind, bleibt die Frage, ob möglicherweise negative indirekte Auswirkungen bestehen, die die positiven Auswirkungen neutralisieren oder sogar überwiegen könnten.

Zudem lassen sich die direkten und indirekten Auswirkungen nicht immer klar voneinander trennen. So hat ein Unternehmen eventuell mehr investiert (eigentlich eine direkte Auswirkung), weil seine eigenen Investitionen die Investitionen konkurrierender Unternehmen verdrängt haben (damit verbundene indirekte Auswirkung). Ein Unternehmen kann aber auch mehr investiert haben, da es Spillover-Effekte und Investitionen seitens anderer Unternehmen erwartet. Zudem könnte die Beihilfe selbst oder sogar die reine Gewährung eine Beihilfe ebenfalls einen dieser beiden Auswirkungen haben. In der Evaluierung der direkten Auswirkungen sollte deshalb auch ausführlich auf die Wahrscheinlichkeit indirekter Auswirkungen, deren Richtung und deren Umfang eingegangen werden. Ferner sollte ausdrücklich die Wirtschaftstheorie genannt werden, nach der die indirekten Auswirkungen mit der Beihilfe in Verbindung gesetzt werden. Dabei sollten zusätzliche Informationen, die diese Theorie untermauern helfen, fester Bestandteil der Evaluierung sein.¹⁵

Kausale Inferenz

Der Kausaleffekt ist der Unterschied zwischen dem Ergebnis mit Beihilfe und dem Ergebnis ohne Beihilfe. Während das Ergebnis mit Beihilfe bei den Unternehmen, die eine Beihilfe erhalten haben, beobachtet wird, wird das Ergebnis ohne Beihilfe nur bei den Unternehmen gemessen, die keine Beihilfe erhalten haben. So wird per definitionem bei den Unternehmen, die die Beihilfe erhalten haben, nicht beobachtet, welches Ergebnis ohne Beihilfe vorgelegen hätte. Um abschätzen zu können, wie sich die Beihilfe auf die Beihilfeempfänger auswirken könnte, muss folglich eine *kontrafaktische Fallkonstellation* konstruiert werden, d. h., es wird untersucht, was wahrscheinlich geschehen wäre, wenn die Beihilfeempfänger keine Beihilfe erhalten hätten. Hierfür muss eine *Kontrollgruppe* gefunden werden, d. h. eine Gruppe von Unternehmen, die der Gruppe von Unternehmen, die die Beihilfe erhalten hat, möglichst ähnlich ist, mit Ausnahme der Tatsache, dass sie keine Beihilfe erhalten hat.

15 Auch wenn der Hauptfokus dieser Arbeitsunterlage auf den direkten Auswirkungen von Beihilfen liegt, so stellt die Tatsache, dass möglicherweise auch indirekte Auswirkungen vorliegen, gewisse analytische Herausforderungen an die Bewertung der direkten Auswirkungen; den Auswirkungen auf das Marktgeschehen ist dabei besondere Aufmerksamkeit zu widmen.

Die Qualität der Kontrollgruppe ist für die Aussagekraft der Evaluierung von entscheidender Bedeutung. Die Unternehmen, die Beihilfen erhalten, unterscheiden sich in ihren Merkmalen in der Regel von jenen, die keine Beihilfen erhalten. Sie sind möglicherweise in einer wirtschaftlich schwächeren Region mit geringerem Marktpotenzial tätig, sind mehr von Krediten abhängig, mehr oder weniger effizient oder haben ein bestimmtes Vorhaben, das es durchzuführen gilt. Ein einfacher Vergleich von Beihilfeempfängern mit Nichtempfängern dürfte somit eher diese Aspekte widerspiegeln, aber nicht die Wirkung der Förderpolitik.

Damit eine aussagekräftige Evaluierung erfolgen kann, muss sichergestellt werden, dass dieser systematische Unterschied zwischen Beihilfeempfängern und Nichtempfängern die Ergebnisse nicht verfälscht (sogenannter Auswahleffekt). In den letzten Jahrzehnten sind hierzu mehrere zuverlässige Methoden entwickelt worden. Die Methodenwahl richtet sich nach der zu evaluierenden Politik und den verfügbaren Daten. Jede Methode hat ihre Grenzen und ist nur unter bestimmten Annahmen angemessen. Eine Studie gewinnt an Glaubwürdigkeit, wenn auch ausführlich dargelegt wird, was sie leisten und was sie nicht leisten kann. In diesem technischen Anhang werden die einschlägigen Methoden vorgestellt. Dabei wird vor allem auf die wichtigsten praktischen Aspekte und die Bedeutung einer guten Identifikationsstrategie eingegangen.¹⁶

A. Randomisierte Experimente

Die Ermittlung einer geeigneten Kontrollgruppe ist Voraussetzung für eine gute (d. h. unverzerrte) Schätzung der Auswirkungen einer bestimmten Politik. Im günstigsten Falle bestünde kein Auswahleffekt, da es sich bei den Beihilfeempfängern um eine Zufallsstichprobe handelt.¹⁷ In diesem Falle besteht außer dem Beihilfetatbestand kein systematischer Unterschied zwischen Beihilfeempfängern und Nichtbeihilfeempfängern, so dass die Ergebnisse der jeweiligen Förderpolitik zugeschrieben werden können.

Die Zufallsauswahl von Beihilfeempfängern wird jedoch gelegentlich kritisiert, da sie dem Ziel vieler Regelungen entgegensteht, auf der Grundlage objektiver Kriterien die am besten geeigneten Empfänger auszuwählen. Dennoch kann es unter bestimmten Umständen zulässig sein, in Bezug auf die Förderfähigkeit der Beihilfeempfänger oder die Anreize für die Teilnahme der Beihilfeempfänger Zufallselemente einzubringen. Als Beispiel sei hier die Festlegung eines festen Budgets für eine bestimmte Regelung angeführt. Wenn die von den Beihilfeempfängern beantragten Summen das Budget übersteigen und sie sich in ihren Merkmalen mehr oder weniger ähneln, dann könnte es angebracht sein, eine gewisse Zufälligkeit bei der Behandlung zu etablieren. Eine andere Möglichkeit bestünde darin, die potenziellen Beihilfeempfänger in unterschiedlichem Umfang über die Regelung zu informieren.

¹⁶ Dieser Anhang gibt einen kurzen und allgemeinen Überblick über ökonometrische Methoden für die Evaluierung bestimmter politischer Maßnahmen. Dieser Überblick stützt sich in vielen Aspekten auf Givord (2010); weitere aufschlussreiche Darstellungen befinden sich bei Imbens und Wooldridge (2009) sowie Angrist und Pischke (2008).

¹⁷ Bei der Bewertung der Wirkung von Drogen und medizinischen Behandlungen sind randomisierte Experimente seit Jahrzehnten die einzige akzeptable Vorgehensweise.

Pilotvorhaben bieten weitere Möglichkeiten, mit einer zufälligen Zuweisung von Beihilfen zu arbeiten. Im Falle innovativer Maßnahmen könnte es ratsam sein, zunächst ein Pilotvorhaben von geringerem Umfang zu evaluieren. Ein solches Vorhaben hätte den Vorteil, dass sein Umfang begrenzt ist und die Beihilfeempfänger leichter nach dem Zufallsprinzip ausgewählt werden können. Eine andere Option wäre, für eine Regelung eine „Ramp-up“-Phase vorzusehen, d. h. eine Anlaufphase, in der zum Beispiel zunächst 25 % nach dem Zufallsprinzip ausgewählte Unternehmen für beihilfefähig erklärt werden und anschließend dieser Anteil im zweiten, dritten und vierten Jahr auf 50, 75 bzw. 100 % erhöht wird (oder aber die Regelung einer immer größeren Zielgruppe anzubieten). Für eine neue politische Maßnahme ist eine solche Ramp-up-Phase in vielen Fällen eine administrative Notwendigkeit.

Diese Vorgehensweisen eignen sich wahrscheinlich eher für die Durchführung völlig neuer Regelungen oder für eine große Gruppe bestehende Regelungen. Im Falle der Fortführung einer bestehenden Regelung dürfte sich eine (direkte oder indirekte) Randomisierung der Beihilfefähigkeit als relativ schwierig erweisen. Dies bedeutet jedoch nicht, dass sich randomisierte Experimente nicht für die Evaluierung einzelner Teile von bestehenden Regelungen eignen würden. So ist es weiterhin möglich, für potenziell effizientere, stärker zielgerichtete und/oder weniger wettbewerbsverzerrende Varianten einer Regelung Beihilfeempfänger nach dem Zufallsprinzip auszuwählen. Im Falle einer Zuschussregelung wäre es denkbar, stattdessen nach dem Zufallsprinzip eine neu eingeführte Kreditregelung vorzuschlagen.

B. Quasi-experimentelle Methoden

Auch wenn sich die Auswirkungen bestimmter politischer Maßnahmen am besten im Rahmen randomisierter Experimente evaluieren lassen, ist deren Durchführung nicht immer möglich. Es gibt zudem andere Methoden, mit denen die Auswirkungen politischer Maßnahmen aus der Ex-post-Perspektive evaluiert werden können. Auch sie stellen auf die exogenen Veränderungen des Umfelds ab, in dem Unternehmen tätig sind, um Konstellationen zu schaffen, die einem Experiment sehr nahe kommen (sogenannte natürliche Experimente oder Quasi-Experimente).

Für Ex-post-Bewertungen ist es generell schwierig festzustellen, ob natürliche Experimente oder Quasi-Experimente angezeigt sind. Eine gründliche Analyse der Ausgestaltung einer politischen Maßnahme kann jedoch eine Analyse des Vorliegens einer ausreichenden exogenen Variabilität beinhalten. Falls erforderlich kann das anfängliche Untersuchungsdesign angepasst werden, um mehr Elemente aufzunehmen, anhand derer die Auswirkungen der jeweiligen Maßnahme ermittelt werden können.

Kontrolle in Bezug auf beobachtbare Unterschiede

Wie weiter oben erläutert, bestehen in der Regel Unterschiede zwischen Beihilfeempfängern und Nichtempfängern. Bei der Gegenüberstellung der Untersuchungsergebnisse für diese beiden Gruppen muss diesen Unterschieden Rechnung getragen werden.

Viele der Merkmalsunterschiede sind in der Regel beobachtbar. Die gängigste Methode zur Erfassung dieser Unterschiede ist die lineare Regression. Mit der linearen Regression wird versucht zu erklären, wie sich beobachtbare Merkmale auf die Ergebnisse ausgewirkt haben. Dabei wird von einem linearen Zusammenhang zwischen dem Ergebnis, z. B. einer FuE-Investition, und anderen Merkmalen eines Unternehmens, z. B. Wirtschaftszweig, Alter oder Größe einschließlich der Gewährung der Beihilfe, ausgegangen. Es ist möglich, die lineare Regression als lineare Näherung komplexerer Zusammenhänge zu betrachten.¹⁸ Lineare Regressionen sind vielseitig verwendbar und kommen in ganz unterschiedlichen Evaluierungsumgebungen zum Einsatz.

Eine Alternative zur linearen Regression ist das Matching-Verfahren, bei dem für jeden Beihilfeempfänger ein „statistischer Zwilling“ gesucht wird, d. h. ein anderes Unternehmen, das möglichst ähnlich „aussieht“, aber keine Beihilfe erhalten hat. Als beobachtbare Variablen für das Matching können Merkmale der Unternehmen oder die geschätzte Wahrscheinlichkeit einer Beihilfe (*Propensity Score Matching – PSM*) verwendet werden. Matching kann sich als hilfreiche Methode zur Kontrolle der beobachtbaren Variablen im Kontext einer validen empirischen Strategie erweisen.

Aber sowohl die einfache lineare Regression als auch das Matching sind in ihren Möglichkeiten begrenzt. Beide sind nur unter Annahme der bedingten Unabhängigkeit (*conditional independence assumption*) aussagekräftig. Diese besagt, dass sobald die Auswirkung eines beobachtbaren Merkmals berücksichtigt worden ist, das Ergebnis von dem beobachtbaren Merkmal unabhängig ist. Für die Praxis bedeutet dies, dass jede Variable, die sowohl Einfluss auf das Ergebnis als auch die Auswahl hat, beobachtbar sein und in der angemessenen funktionalen Form berücksichtigt werden muss. Ist dies nicht der Fall, spiegelt der reine Umstand, dass ein Unternehmen teilnimmt, bestimmte (unbeobachtete) Merkmale eines Unternehmens wider, die ebenfalls dessen Entwicklung beeinflussen. Sowohl die lineare Regression als auch die Matching-Methode würden dann keine gültige Evaluierung gewährleisten. Wenn ein Unternehmen ein „vielversprechendes Vorhaben“ hat, wirkt sich dies sowohl auf die Wahrscheinlichkeit, dass sich das Unternehmen um eine Beihilfe bemühen (und diese erhalten) wird als auch auf die Wahrscheinlichkeit, dass das Unternehmen geschäftlich erfolgreich sein wird. Dies nicht zu berücksichtigen, würde die Ergebnisse verfälschen.

Insbesondere beim Matching könnte durch einen Vergleich der Ergebnisse bei einem Beihilfeempfänger und dessen gematchtem Zwilling, der keine Beihilfe erhalten hat, ein Auswahleffekt nur dann vermieden werden, wenn die Gewährung der Beihilfe in keinem Zusammenhang mit unbeobachteten Variablen steht, die das Ergebnis ebenfalls beeinflussen. In der Realität ist diese Annahme selten erfüllt. Die Messung aller Variablen, die sich auf die Beantragung oder den Erhalt einer Beihilfe auswirken, ist selten möglich. Des Weiteren setzt die Verwendung des Matching-Verfahrens voraus, dass Unternehmen, die eine Beihilfe

¹⁸ Außerdem ist es möglich, Merkmale miteinander zu verknüpfen (z. B. Umsatz und Branche) und für die Merkmale Funktionen einzuführen (z. B. Variablen im Quadrat).

erhalten, in ihren beobachtbaren Merkmalen sehr stark mit jenen übereinstimmen, die keine Beihilfe erhalten haben. Wenn sich die gematchten Unternehmen in Bezug auf jeden beobachtbaren Aspekt wirklich sehr ähneln, dann sind die Gründe, warum einige Unternehmen eine Beihilfe erhalten haben und andere nicht, unbeobachtet. Die Begründung für die Aussagekraft einer auf Matching beruhenden Evaluierung oder einer einfacheren klassischen linearen Regression liegt folglich nicht im reinen Vorliegen eines sehr vollständigen Datensatzes mit vielen beobachteten Merkmalen.

Die potenzielle Begründung für den Rückgriff auf die Matching-Methode oder die einfachere lineare Regression beruht vielmehr auf der Tatsache, dass die unbeobachteten Gründe für die Beihilfefähigkeit oder die Beihilfegewährung keinen direkten oder indirekten Einfluss auf die Ergebnisse haben (sobald sie in Bezug auf die beobachtbaren Variablen kontrolliert wurden). Wenn die Validität einer Evaluierung, für die ein einfaches Matching oder eine einfache lineare Regression verwendet wurde, gewährleistet sein soll, muss die Gewissheit bestehen, dass die Gruppe von Unternehmen, die keine Beihilfe erhalten haben, exogen bestimmt wurde. Dies setzt voraus, dass sobald die beobachtbaren Variablen kontrolliert worden sind, es keinen unbeobachteten Faktor für die Beihilfefähigkeit oder Beihilfegewährung mehr gibt, der sich auch direkt oder indirekt auf die Ergebnisse auswirken könnte. Dieses Kriterium ist in der Regel beim Matching von Unternehmen, die in gleicher Weise für eine Beihilfe in Frage kommen, nicht erfüllt. Wenn zum Beispiel alle Unternehmen beihilfefähig sind, dann ist die Wahrscheinlichkeit, dass Unternehmen, die eine Investitionsbeihilfe erhalten, ein Vorhaben haben, größer als bei Unternehmen, die keine Beihilfe erhalten haben (denn ansonsten hätten sie ebenfalls eine Beihilfe beantragt und eine Beihilfe erhalten). Insgesamt sind die Umsatz- und Beschäftigungsperspektiven von Unternehmen mit einem Vorhaben besser, was jedoch nicht auf die Beihilfe zurückzuführen ist; ein Matching nach beobachtbaren Variablen hilft nicht, diese beiden Aspekte voneinander zu trennen (es sei denn, es wird das Vorliegen eines vergleichbaren Investitionsvorhabens gemessen).

In vielen Fällen ist die Annahme der bedingten Unabhängigkeit zum Scheitern verurteilt. Es könnte deshalb erforderlich sein, neben der reinen linearen Regression oder des Matching-Verfahrens andere Methoden heranzuziehen, um bei der Behandlung dem Vorliegen unbeobachteter Auswahl Rechnung zu tragen.

Im Folgenden werden die gängigsten Methoden, die in diesem Kontext für die Evaluierung der Auswirkungen bestimmter politischer Maßnahmen verwendet werden, ausführlicher erläutert. Hierbei handelt es sich um das Difference-in-Difference-Verfahren, die Regression-Diskontinuität-Analyse (engl. *Regression Diskontierungssatz Design – RDD*) und den Ansatz der Instrumentvariablen (*IV*). Die Validität dieser Methoden beruht auf unterschiedlichen Annahmen, und die Wahl richtet sich in der Regel nach dem Kontext der politischen Maßnahme und der Verfügbarkeit von Daten. Nachstehend werden die Stärken und Schwächen jedes einzelnen Verfahrens dargestellt. Es gibt nachweislich kein besseres Verfahren als das der weiter oben beschriebenen randomisierten kontrollierten Studien (*randomised controlled trials – RCT*). Grundlage für die Wahl eines bestimmten Verfahrens

muss eine sorgfältige Analyse des Kontextes, in dem die Maßnahme vorgenommen werden soll, und der verfügbaren Daten sein.

An dieser Stelle sei besonders darauf hingewiesen, dass nicht die Verwendung eines bestimmten ökonomischen Verfahrens es ermöglicht, die Auswirkungen einer bestimmten politischen Maßnahme zu ermitteln, sondern vielmehr die Exogenität der Kontrollgruppe und folglich die Qualität der kontrafaktischen Fallkonstellation. Die Qualität der Evaluierungsstudie hängt folglich entscheidend davon ab, wie überzeugend der Forscher die Exogenität der Kontrollgruppe gewährleisten kann. In Fällen, in denen noch letzte Verzerrungen verbleiben, müssen diese einschließlich ihrer Ursprünge sowie die Richtung und das voraussichtliche Ausmaß ihrer Auswirkungen unbedingt erörtert werden.

a) Das Difference-in-Difference-Verfahren

Ansatz und Identifikation

Wie bereits weiter oben angesprochen, ist es unwahrscheinlich, dass ein einfacher Vergleich zwischen Beihilfeempfängern und einer sei sie noch so gut ausgewählten Gruppe von Nichtempfängern zu einem aussagekräftigen Evaluierungsergebnis führt. Es ist nicht möglich, das Vorliegen unbeobachteter Unterschiede zwischen zwei Gruppen auszuschließen, was selbst in der beihilfefreien Konstellation bedeutet, dass es einen dauerhaften Unterschied gibt. Ein einfacher Vergleich der Ergebnisse vor und nach der Gewährung einer Beihilfe dürfte ebenfalls keine verlässlichen Evaluierungsergebnisse erbringen. Eine solche Vorgehensweise ermöglicht es nicht, zwischen den Auswirkungen der Beihilfe und anderen Faktoren, die die Ergebnisse bei den beiden Gruppen beeinflussen könnten (z. B. allgemeine konjunkturelle Entwicklung, geänderte rechtliche Rahmenbedingungen oder steigende Arbeitskosten), zu trennen.

Bei dem Difference-in-Difference-Ansatz handelt es sich um eine Kombination der beiden vorgenannten Ansätze, so dass es möglich ist, den Kausaleffekt der Beihilfe zu bewerten. Die Grundidee dieses Ansatzes besteht darin, über einen längeren Zeitraum die unterschiedliche Entwicklung von Unternehmen zu erfassen. Bisherige Unterschiede würden dann anderen Faktoren als der staatlichen Beihilfe zugeschrieben. Nur die Abweichung bei diesen Unterschieden (die „Difference-in-Difference“) würde dann auf die staatliche Beihilfe zurückgeführt werden. Mit anderen Worten, bei diesem Verfahren wird der Unterschied in der Entwicklung der Beihilfeempfänger und der Kontrollgruppe *vor* Erhalt der Beihilfe und *nach* Erhalt der Beihilfe erfasst und die Differenz wiederum der staatlichen Beihilfe zugeschrieben. Das Verfahren funktioniert, wenn im Laufe der Zeit sowohl die Beihilfeempfänger als auch die Kontrollgruppe unter dem Einfluss der anderen Faktoren stehen, die ihre Entwicklung in derselben Weise beeinflussen. Dies lässt dann den Schluss zu, dass die Beihilfe der einzige relevante Faktor ist, der die beobachtete Weiterentwicklung bei den Beihilfeempfängern gegenüber der Kontrollgruppe bedingt hat.

Die entscheidende Annahme ist, dass die Differenz zwischen den Beihilfeempfängern und der Kontrollgruppe über die Zeit stabil ist und dass beide Gruppen in derselben Weise in dem

betreffenden Zeitraum von allgemeinen Schocks betroffen sind (Abweichungen vom Mittel). Diese Annahme muss in der Praxis nicht unbedingt zutreffen. Wenn es sich zum Beispiel bei den Beihilfeempfängern um weniger robuste Unternehmen handelt, dürften diese stärker unter einem Konjunkturrückgang leiden und vom allgemeinem Wirtschaftsklima abhängig sein. Deshalb muss die Kontrollgruppe ebenfalls aus weniger robusten Unternehmen bestehen. Die Wahl der Kontrollgruppe ist für die Validität des Verfahrens von entscheidender Bedeutung. Die Identifikation ergibt sich nicht aus dem Difference-in-Difference-Verfahren, bei dem es sich um die rein technische Durchführung handelt, sondern in der richtigen Wahl der Kontrollgruppe.

Bei der Zusammenstellung der Kontrollgruppe ist besondere Vorsicht geboten, wenn sich die Nichtempfänger selbst gegen die Beantragung einer Beihilfe entschieden haben. Es ist davon auszugehen, dass es bei der Entscheidung für oder gegen die Beantragung einer Beihilfe eine Rolle spielt, welche Vorteile sich der Antragsteller von einer Beihilfe verspricht. Es gibt deshalb Anhaltspunkte für die Annahme, dass sich die erwarteten Ergebnisse der Unternehmen, die keinen Antrag auf Beihilfe stellen (Ergebnisse in Bezug auf Beschäftigung, Produktivität, Umsatz usw.), sich von jenen der Beihilfeempfänger unterscheiden. Wenn zum Beispiel alle Unternehmen, die eine Beihilfe beantragt haben, auch eine Beihilfe erhalten, dann sind die einzigen beihilfefähigen Unternehmen, die keinen Beihilfeantrag stellen, jene ohne ein Vorhaben (unter der Annahme niedriger Kosten für die Antragstellung). Diese Unternehmen werden wahrscheinlich nicht nur absolut, sondern auch relativ über die Zeit schlechter abschneiden, während andere Unternehmen ihre Vorhaben durchführen und wachsen. Es kann nicht erwartet werden, dass sich Beschäftigung, Produktivität oder Umsatz parallel entwickeln, und eine Doppeldifferenzierung ist in der Regel nicht die Lösung des Problems.

Deshalb müssen die Unternehmen der Kontrollgruppe, denen keine Beihilfe gewährt wurde, dieser Gruppe aus Gründen angehören, die keinen Einfluss auf die gemessenen Ergebnisse haben. Es darf sich nicht um Unternehmen handeln, die von sich aus freiwillig entschieden haben, nicht teilzunehmen. Am überzeugendsten ist eine Zusammensetzung, bei der als Folge eines natürlichen Experiments ein Zusammenhang zwischen Nichtbeteiligung und Nichtförderfähigkeit besteht. In diesem Falle ist es unwahrscheinlich, dass die Nichtförderfähigkeit auf unbeobachtete Faktoren zurückzuführen ist, die ebenfalls die Ergebnisse beeinflussen. Die Kontrollgruppen könnten zum Beispiel aus in der Region ansässigen Unternehmen gebildet werden, die nicht mehr für Beihilfen in Frage kommen (solange deren Förderfähigkeit nicht von ihrer eigenen Entwicklung, sondern eher von einem exogenen Ereignis abhängt).

Durchführung

Technisch gesehen können Difference-in-Difference-Verfahren entweder innerhalb eines linearen Regressionsmodells oder in Verbindung mit einem Matching durchgeführt werden. Bei einer linearen Regression wird die Kontrollgruppe unabhängig von den beobachtbaren Merkmalen ausgewählt; sie ist deshalb insgesamt mit der gesamten Gruppe der Beihilfeempfänger vergleichbar. Die beobachtbaren Unterschiede werden dann im Rahmen

der klassischen linearen Regression berücksichtigt. Beim Matching wird die Kontrollgruppe aus Unternehmen gebildet, die einzeln in Bezug auf beobachtbare Faktoren mit jedem geförderten Unternehmen der Stichprobe vergleichbar sind. Das Ergebnis für jedes Unternehmen wird mit dem Ergebnis der Unternehmen, die ihm am stärksten ähneln, verglichen und die Ergebnisse werden aggregiert. Die beiden Methoden stellen zwei unterschiedliche Vorgehensweisen dar, um beobachtbare Unterschiede zu berücksichtigen, es bestehen allerdings keine fundamentalen Unterschiede in Bezug auf die Identifikation des Kausaleffekts der jeweiligen politischen Maßnahme.

Es könnte sich gegebenenfalls lohnen, die Variationen in Bezug auf die Ergebnisse der Beihilfeempfänger und jene der Kontrollgruppe vor Gewährung der Beihilfe zu vergleichen. Wenn sich die Unternehmen schon vor der eigentlichen Gewährung der Beihilfe in ihrer Entwicklung systematisch unterscheiden, ist es wahrscheinlich, dass sich die Kontrollgruppe und die Gruppe der Beihilfeempfänger aus nicht beihilfebedingten Gründen in ihrer Entwicklung voneinander entfernen, so dass die Methode keine aussagekräftige Schätzung des Kausaleffekts der Beihilfe zulässt. Hierbei handelt es sich nicht um eine rigorose Prüfung der Validität der Annahme: einen solchen Test gibt es nicht. Es ist aber zumindest eine Art erster Plausibilitätstest.

Wenn es mehrere potenzielle Kontrollgruppen gibt, die a priori zuverlässig sind, können zusätzliche Methoden und Robustheitstests eingesetzt werden. Der erste und natürlichste Robustheitstest besteht darin, mehrere Difference-in-Difference-Schätzer einzuführen und die Ergebnisse zu vergleichen. Darüber hinaus können diese unterschiedlichen Kontrollgruppen dazu dienen, eine zuverlässigere Schätzung vorzunehmen. Als Beispiel sei hier eine Regelung zur Förderung von KMU in einer bestimmten Region angeführt. Die möglichen Kontrollgruppen wären dann Unternehmen in dieser Region, die keine KMU sind, oder KMU in einer angrenzenden Region. Keines dieser Unternehmen hat sich freiwillig gegen eine Teilnahme an der Regelung entschieden, sondern sie kommen gar nicht für eine Förderung in Betracht. Dennoch ist keine dieser Kontrollgruppen perfekt: die allgemeinen konjunkturellen Entwicklungen dürften sich auf größere Unternehmen derselben Region anders auswirken als auf die KMU, während die KMU einer benachbarten Region möglicherweise anderen regionalen Schocks unterliegen. Anstatt zwischen diesen beiden möglichen Difference-in-Difference-Schätzern zu wählen, besteht die Möglichkeit, diese beiden zu kombinieren und ein Dreifachmodell, das sogenannte *Difference-in-Difference-in-Differences*-Modell (DDD-Modell) zu verwenden: Ausgangspunkt ist dabei die klassische „Difference-in-Difference“ zwischen KMU und Nicht-KMU in der betreffenden Region; dieselbe Difference-in-Difference kann von der benachbarten Region subtrahiert werden, um die Variation bei den Ergebnissen zwischen KMU und Nicht-KMU in der geförderten Region auszugleichen.¹⁹

19 Als Beispiel sei hier eine regionale KMU-Förderregelung zur Schaffung neuer Arbeitsplätze genannt, bei der sich zum Ende der Förderregelung zeigt, dass die Ergebnisse der KMU in der Region in Bezug auf die Schaffung neuer Arbeitsplätze 20 % besser sind als jene der größeren Unternehmen. Wenn in einer vergleichbaren angrenzenden Region (in der keine Beihilfen gewährt wurden) KMU ebenfalls bessere Ergebnisse erzielt haben als größere Unternehmen (z. B. um 15 %), dann kann die Wirkung der Beihilfe mit rund 5 % veranschlagt werden.

Stattdessen könnte man auch systematisch versuchen, eine synthetische Kontrollgruppe aus KMU aus mehreren benachbarten Regionen und Nicht-KMU derselben Region zu etablieren, um das Entwicklungsmuster bei den Beihilfeempfängern vor der Beihilfe besser abzubilden (zu weiteren Einzelheiten, siehe Abadie, Diamond und Hainmuller, 2010).

Inferenz

Zusätzlich zu einer sorgfältigen Zusammensetzung und Auswahl der Kontrollgruppe muss die Frage der Inferenz angesprochen werden. Der Begriff der Inferenz bezieht sich in diesem Kontext auf die Frage, ob die geschätzten Auswirkungen wirklich signifikant sind. Statistische Signifikanz und wirtschaftliche Signifikanz sind zwei unterschiedliche Aspekte. Letzterer bezieht sich auf den Umfang der geschätzten Wirkungen gegenüber den anderen relevanten Parametern aus wirtschaftstheoretischer Sicht. Wie bereits weiter oben erläutert, ist die wirtschaftliche Signifikanz entscheidend. Diese Diskussion ist jedoch nur dann relevant, wenn – aus statistischer Sicht – die Auswirkungen genau genug geschätzt wurden, d. h., es kann ausgeschlossen werden, dass es keine Auswirkungen gab.

Es gibt Gründe für die Auffassung, dass eine einfache Inferenz bei Standardannahmen (wie Homoskedastizität-Annahme und Annahme der Autokorrelation) wahrscheinlich zu einer Überschätzung der statistischen Signifikanz der Auswirkungen führt.²⁰

Das erste Problem entsteht in Verbindung mit der Bündelung (*clustering*) von Daten. Wenn es sich bei der Kontrollgruppe wie auch bei der Gruppe der Beihilfeempfänger jeweils um sehr homogene Gruppen handelt (selbst wenn sie sich selbst klar voneinander unterscheiden), dann dürften in jeder Gruppe alle Unternehmen in ähnlicher Weise von Abweichungen vom Mittel (Schocks) betroffen sein. Statistisch bedeutet dies, dass der Fehlerterm eine gemeinsame Komponente hat. Wenn die Varianz dieser gemeinsamen Komponente gegenüber dem Spektrum der beobachteten Ergebnisse für einzelne Unternehmen sehr groß ist, dann ist die Inferenz verzerrt. Bei zwei Zeiträumen und zwei Gruppen könnte sich dieses Problem zuspitzen und dann zu einem Identifikationsproblem werden, wenn es nicht mehr möglich ist, die Auswirkungen der Schocks innerhalb einer Gruppe von den Auswirkungen der politischen Maßnahme zu trennen. Dieses Problem muss nicht unbedingt genauso gravierend sein, wenn die Gruppen nicht so homogen sind. Es ist allerdings immer notwendig, der Tatsache Rechnung zu tragen, dass homogene Teile der Gruppe von denselben Schocks betroffen sein könnten. Wenn es sich zum Beispiel um eine lokale Nachfrage handelt, wird es in der Regel notwendig sein, auf Ebene der örtlichen Gemeinschaften Korrekturen bei der gebündelten Struktur des Fehlerterms vorzunehmen. Dasselbe könnte bei einzelnen Industrien oder Branchen erforderlich sein.

Das zweite Problem tritt auf, wenn Paneldaten verwendet werden. Die Fehlerterme der meisten Daten, die sich auf das Unternehmen beziehen (z. B. Beschäftigung, Produktivität und Investitionen), sind in der Regel autokorreliert. Dies bedeutet, dass sich Abweichungen

²⁰ Dieser Aspekt wurde in Bezug auf Difference-in-Difference-Verfahren besonders unterstrichen, dieselben Probleme können allerdings bei anderen hier beschriebenen Verfahren auftauchen.

vom Mittel in einem bestimmten Zeitraum wahrscheinlich im nächsten Zeitraum fortsetzen werden. Eine Ausklammerung dieses Aspekts würde zu einer Überschätzung der Genauigkeit der Schätzung der Auswirkungen und – allzu häufig – zur Abweisung der Hypothese, dass die politische Maßnahme keine Wirkung hatte, führen. Dies kann, wie von Bertrand, Duflo und Mullainathan (2004) dargelegt, ein schwerwiegendes Problem sein.

b) Instrumentvariablen

Ansatz und Identifizierung

Die Methode der Instrumentvariablen („IV“) ist ein klassisches Verfahren zur Behandlung der Endogenität erklärender Variablen. Da die Inanspruchnahme von Beihilfen im Kontext einer linearen Regression als eine endogene erklärende Variable in der Weiterentwicklung eines Unternehmens betrachtet werden kann, erscheint es naheliegend, Instrumentvariablen zu verwenden, um die Auswirkung der Beihilfe zu evaluieren.

Eine Variable ist endogen, wenn sie mit einem unbeobachteten Element korreliert, das ebenfalls das Ergebnis bestimmt. So könnte zum Beispiel versucht werden, durch Regression der Beschäftigung auf die Beteiligung am Förderprogramm und andere beobachtbare Variablen zu identifizieren, wie sich staatliche Zuschüsse auf die Beschäftigungspolitik von Unternehmen auswirken. Ein Beispiel: Ein Förderprogramm ist für leistungsschwache Unternehmen bestimmt, für die es schwer sein könnte, den lokalen Marktbedingungen standzuhalten. Marktbedingungen sind nicht durch einen Evaluierer zu beobachten und folglich nicht direkt kontrollierbar. Würde diese Variable außer Acht gelassen werden, könnte der Effekt dieses Zuschusses aufgrund des endogenen Charakters der Beteiligung an dem Förderprogramm allzu leicht vom Evaluierer unterschätzt werden. Ob das Unternehmen unter günstigen oder schwierigen Marktbedingungen tätig ist, wirkt sich sowohl auf dessen Beteiligung an dem Förderprogramm als auch die Beschäftigung aus, d. h., die Programmbeteiligung korreliert mit dem Fehlerterm, der die Beschäftigung erklärt. Der Einfluss der Marktbedingungen auf die Programmbeteiligung bedeutet, dass es nicht möglich ist, die gesamte Korrelation zwischen Programmbeteiligung und Beschäftigung dem kausalen Effekt der Beihilfe zuzuschreiben.

Es gibt jedoch auch noch andere Faktoren als die Beschäftigung, die die Programmbeteiligung erklären. Wie in Criscuolo u. a. (2012) erläutert, könnte die geografische Lage eine Rolle spielen, wenn es um die Höhe der insgesamt für die Region zur Verfügung stehenden Fördermittel geht. Auch die Zahl der Regionen, die für die Regelung in Frage kommen, kann sich im Laufe der Zeit ändern. Wenn sich die Mittelausstattung eines Programms für eine bestimmte Region im Laufe der Zeit aus externen Gründen (z. B. rückläufiges durchschnittliches BIP pro Kopf in der EU) ändert, so wirkt sich dies auf die Programmbeteiligung aus, aber nicht auf die Marktbedingungen eines Unternehmens vor Ort. Die durch die exogene Veränderung in der Programmabdeckung bedingte geänderte Beschäftigungssituation steht in keinem Bezug zu den örtlichen Marktbedingungen. Durch Fokussierung auf diesen „Teil“ der Programmbeteiligung als Variable, ist es möglich, ohne Interferenz der lokalen Marktbedingungen die tatsächliche Auswirkung der

Programmbeteiligung auf die Beschäftigung des Unternehmens zu isolieren. Dies ist die Logik von Instrumentvariablen.

Bei der Evaluierung staatlicher Beihilfen ist eine Instrumentvariable eine Variable, die die Tatsache erklären kann, dass die Beihilfe gewährt wurde, aber ansonsten keinen direkten Einfluss auf andere unbeobachtete Determinanten des zu messenden Ergebnisses hat. Instrumentvariablen ermöglichen also eine Fokussierung auf die Beteiligung an einer Förderregelung ohne Interferenz der Auswahleffekte. Zur Veranschaulichung sei die Logik der Instrumentvariablen hier erläutert.²¹ In einem ersten Schritt wird die Programmbeteiligung in Bezug auf alle exogenen Variablen (einschließlich der Instrumentvariablen) regressiert. In einem zweiten Schritt wird die Variable Beteiligung (die Variable, aus der hervorgeht, dass die Beihilfe gewährt wurde) durch die Beteiligung ersetzt, die im ersten Schritt vorhergesagt wurde: diese erwartete Beteiligung ist nicht mit dem unbeobachteten Element korreliert, das das Ergebnis ebenfalls bestimmt.

Probleme bei schwachen Instrumenten

Eine Instrumentvariable ist eine Variable, die die Tatsache erklären kann, dass die Beihilfe gewährt wurde, aber ansonsten keinen direkten Einfluss auf andere unbeobachtete Determinanten des zu messenden Ergebnisses hat. Die einfache und klassische Definition verhüllt allerdings eine Reihe praktischer Schwierigkeiten. Wenn mehr Instrumente verwendet werden, als streng genommen für die Identifikation des Modells erforderlich sind, gibt es Tests, mit denen geprüft werden kann, ob diese Instrumente konsistent sind. Es gibt allerdings keine Tests, mit denen die Aussagekraft der Instrumente untersucht werden kann. Das Hauptziel einer Studie, die Instrumentvariablen verwendet, besteht in der Regel darin, zu erläutern, warum jedes einzelne Instrument als unkorreliert mit den unbeobachteten Determinanten für die Entwicklung von Unternehmen (z. B. Beschäftigung, Produktivität, Umsatz, Investitionen usw.) erachtet werden kann. Solche auf ökonomischen und faktischen Elementen beruhenden Erklärungen sind für die Bewertung der Aussagekraft der Evaluierung erforderlich. Sie sind allerdings, insbesondere wenn mehrere Instrumente verwendet werden, nicht ausreichend.

In der Erörterung der Qualität der Instrumentvariablen sollte auch die Frage schwacher Instrumente angesprochen werden, z. B. Instrumente, die schwach mit der Ergebnisvariablen korreliert sind. Wenn Instrumentvariablen schwach mit der endogenen Variablen korreliert sind, ist es wahrscheinlich, dass die Schätzungen unpräzise sind. In einem solchen Fall besteht die Versuchung, weitere Instrumentvariablen heranzuziehen. Es ist allgemein bekannt, dass bei Verwendung einer ausreichend großen Zahl von Variablen ausreichend Material der ursprünglichen Variablen vorliegt, um statistisch signifikante Ergebnisse zu erzielen. Gleichzeitig nähert sich der 2SLS-Ansatz natürlich immer mehr der verzerrten Schätzung der

²¹ In der Praxis wird somit der Ansatz der 2SLS-Ansatz aus bekannten Inferenzgründen in einem Schritt implementiert.

Methode der kleinsten Quadrate (Fehlerquadratmethode).²² Die Gefahr einer solchen Verzerrung sollte bei jeder Evaluierung anhand der IV-Methode ausdrücklich angesprochen werden. So muss insbesondere die Glaubwürdigkeit nicht nur der individuellen, sondern auch der gemeinsamen Exogenität geprüft werden.

Ein besonderer Fall liegt vor, wenn angenommen wird, dass die endogene Variable autokorreliert ist. Wenn eine rein zeitgleiche Quelle der Endogenität zugrunde gelegt wird, dann ist es möglich, bisherige Werte als Instrumentvariablen zu verwenden. In diesem Falle müsste jedoch die Validität dieses Ansatzes erörtert werden. Wenn zum Beispiel erklärende Variablen autokorreliert sind, dann könnte dies auch auf das gemessene Ergebnis zutreffen. Dann sind auch die verzögerten Variablen endogen. Ganz allgemein bedeutet dies, dass Exogenitätsannahmen nicht zutreffen, wenn die Autokorrelation der erklärenden Variablen sehr hoch ist. Bei geringer Autokorrelation besteht die Möglichkeit, auf viele verzögerte Variablen (und potenziell künftige Werte) zurückzugreifen, aber auch das Risiko, den Fehler zu begehen, viele schwache Instrumente (siehe oben) zu verwenden. Insgesamt ist die Verwendung von früheren Werten schon eine zulässige Strategie, bei der allerdings Vorsicht geboten ist.

Um die weiter oben beschriebenen Probleme zu vermeiden, sollte unbedingt nur eine kleine Zahl überzeugender Instrumente verwendet werden. Dabei muss jedoch auch dargelegt werden, dass es sich bei den Instrumenten um gute Prädiktoren für die endogene erklärende Variable handelt.²³

Variationen bei Zweistufen-Verfahren: Das Selektionsmodell von Heckman (1979)

Handelt es sich bei der endogenen Variablen um eine Behandlungsvariable (Pseudovariablen), kann die erste Regressionsschätzung des 2SLS-Ansatzes als lineares Wahrscheinlichkeitsmodell für die zu behandelnde Wahrscheinlichkeit betrachtet werden. Das lineare Wahrscheinlichkeitsmodell ist eine lineare Näherung. Manchmal kann die zu behandelnde Wahrscheinlichkeit, selbst wenn sie auf beihilfefähige Unternehmen beschränkt wird, gering sein. In diesem Falle könnten sich lineare Näherungen als zu grob erweisen, so dass sie die zu behandelnde Wahrscheinlichkeit nicht wirklich abbilden und eine Konzentration auf die Verteilungsenden, die in diesem Falle von Interesse sind, nicht möglich ist. Dieses Problem kann auf verschiedene Weise angegangen werden. Bei allen Ansätzen

22 Eine sehr interessante praktische Erörterung der Problematik von Verzerrungen aufgrund schwacher Instrumente ist nachzulesen in Bound, Jaeger und Baker's (1995), die auf die statistischen Verzerrungen in Angrist und Krueger (1991) eingehen. Zudem sind IV-Schätzungen mit endlicher Distanz verzerrt. Deshalb können selbst bei ausreichend großen Datensätzen für die Gewährleistung offensichtlicher statistischer Signifikanz nicht-asymptotische Verzerrungen wichtig sein.

23 Dies kann durch Berechnung der Fisher-Statistik für die erste Stufe der Regressionsanalyse erfolgen. Je höher die Fisher-Statistik, desto unwahrscheinlicher ist es, dass die Instrumente schwach sind. Stock, Wright und Yogo (2002) schlagen einen förmlichen Test vor. Bei einem Instrument ist es zum Beispiel wichtig, dass die Fisher-Statistik in der ersten Regressionsschätzung unter 10 liegt.

wird das lineare Modell für die zu behandelnde Wahrscheinlichkeit durch eine nichtlineare Funktion ersetzt.²⁴

Das Selektionsmodell (Heckman, 1979) ist ein klassischer Ansatz zur Behandlung der Evaluierungsproblematik. Bei diesem Ansatz wird der Selektionseffekt als ein Problem der Ausschlussvariablen (*omitted variable*) bei linearer Regression des Ergebnisses auf beobachtbare Variablen und Beteiligung behandelt.²⁵ Bei Varianten dieser Methode wird zum Beispiel das gesamte Modell bei maximaler Wahrscheinlichkeit geschätzt oder die Gewährung der Beihilfe nach dem prognostizierten Wert der Selektionsgleichung errechnet.²⁶

Die Identifikation und insbesondere die Wahl der Variable bedürfen in jedem Fall eingehender Überlegungen. Es reicht nicht aus, in beiden Stufen der Schätzung dieselben Variablen heranzuziehen, selbst wenn die Ergebnisse präzise genug sein mögen.²⁷

Ein Selektionsmodell kann nur mit einer sogenannten Ausschlussvariablen zuverlässig geschätzt werden. Eine ausgeschlossene Variable erklärt die Auswahl des Beihilfeempfängers, aber nicht das Ergebnis. Es reicht nicht aus, aus der Hauptgleichung eine Variable zu entfernen und sie der Liste der erklärenden Variablen der Selektionsgleichung hinzuzufügen. Die Ausschlussvariable hat vielmehr die Funktion, die Auswahl zu erklären, darf aber keinen Einfluss auf das Ergebnis haben, das es zu erklären gilt. In ihrem Wesen ähnelt sie sehr stark einer gültigen Instrumentvariablen. Die Wahl einer solchen Variablen ist keine Frage der Zweckmäßigkeit; sie muss auf Wirtschaftstheorie, institutioneller Struktur und/oder Erfahrung beruhen.

24 Der folgende Abschnitt enthält eine kurze Beschreibung der einschlägigen Selektionsmodelle. Einen vollständigen Überblick enthält die bahnbrechende Veröffentlichung von Heckmann (1979) und vor allem auch Wooldridge (2002), Kapitel 17.

25 Diese Ausschlussvariable ist die Differenz bei der bedingten Erwartung des Ergebnisses für die ausgewählte Stichprobe (in diesem Falle die Beihilfeempfänger). Bei bestimmten Annahmen bezüglich des Auswahlverfahrens für die Beihilfeempfänger (z. B. Probit- oder Logit-Modell) kann diese Differenz formal abgeleitet werden („inverse Mills-Ratio“) und ist eine Funktion der Auswahlparameter. Der Effekt der politischen Maßnahme kann ermittelt werden, indem die Ausschlussvariable in die Regression aufgenommen wird. Die Auswahlparameter sind nicht bekannt, aber in einer ersten Schätzung für das Auswahlverfahren können konsistente Parameter herangezogen werden. Dies führt zu einem Schätzungsprozess, der in der Fachwelt auch manchmal als „Heckit“ bezeichnet wird. Hierfür wären zunächst die Parameter zu ermitteln, die für die Auswahl der Beihilfeempfänger von Interesse sind, z. B. eine Probit- oder eine Logit-Spezifikation. Anschließend kann der Effekt der politischen Maßnahme ermittelt werden, indem die geschätzte inverse Mills-Ratio in die lineare Regression aufgenommen wird. Statistische Softwareanwendungen haben in der Regel eine Funktion, mit der Heckman-Schätzungen vorgenommen werden können.

26 Die verschiedenen Methoden sind unter anderem in Kapitel 17 von Wooldridge (2002) ausführlich beschrieben.

27 Bei nichtlinearer Selektionsgleichung ist die inverse Mills-Ratio nicht kollinear mit den anderen erklärenden Variablen, selbst wenn die erste Gleichung einen Teil dieser erklärenden Variablen beinhaltet. Theoretisch ist das Modell bereits identifiziert. In diesem Fall weist die inverse Mills-Ratio nicht genügend Variation auf, was zu sehr unpräzisen Schätzungen führt. Besonders bei großen Stichproben dürfte die Schätzung aber dennoch signifikante Ergebnisse hervorbringen. Wenn sich alle Variablen des Selektionsmodells auch in der Hauptgleichung befinden, lässt sich das Modell dennoch nur allein aufgrund der Nichtlinearität einer bestimmten parametrischen Form identifizieren.

c) Regressions-Diskontinuitäts-Analyse

Die Regressions-Diskontinuitäts-Analyse (RD-Analyse) ist das neueste Instrument, das zur Verbesserung von Evaluierungen beitragen kann.²⁸ RD-Analysen haben in den vergangenen zehn Jahren in der Wissenschaft einen großen Erfolg verzeichnet, vor allem aufgrund ihrer Einfachheit. Bei dieser Methode wird das Bestehen einer Variablen genutzt, die unstete Auswirkungen auf die Wahrscheinlichkeit hat, behandelt zu werden. Im Falle von Beihilferegulungen können verschiedene Arten von Unstetigkeiten („Diskontinuitäten“ bzw. Kontinuitätsbrüche) herangezogen werden. Die erste Unstetigkeit ergibt sich aus geografischen Grenzen: Bestimmte Regelungen gelten nur innerhalb eines Verwaltungsgebiets, etwa einer Gemeinde oder einer NUTS-Region. Die zweite Unstetigkeit resultiert aus Bedingungen, die Unternehmen erfüllen müssen, um in den Genuss einer Regelung zu kommen, beispielsweise in Bezug auf Alter und Größe des Unternehmens.

Dazu ein Beispiel: Bei der Auswahl von Beihilfevorhaben werden diese mit Punkten bewertet (Höchstpunktzahl: 100); Vorhaben ab einer Punktzahl von 70 werden gefördert, bei einer niedrigeren Punktzahl geht das entsprechende Unternehmen leer aus. Ein Unternehmen mit einer Punktzahl von 71 hat ein nur geringfügig besseres Vorhaben vorgelegt als ein Unternehmen mit 69 Punkten. Die Auswirkungen dieses geringfügigen Unterschieds sind jedoch enorm: Ein Unternehmen erhält eine Beihilfe und das andere nicht. Ein Vergleich der Entwicklung dieser beiden Unternehmen erlaubt somit sehr aussagekräftige Rückschlüsse über die kausalen Auswirkungen der Beihilfe.

Formal ist es für eine RD-Analyse erforderlich, dass die Wahrscheinlichkeit, eine Beihilfe zu erhalten, unstet ist, also ein Bruch bzw. Sprung vorliegt, während alle anderen Variablen stet sind.²⁹ Die technische Umsetzung kann sehr ähnlich erfolgen wie bei Instrumentvariablen, wobei eine Pseudovariablen für die Schwellenüberschreitung als Instrument verwendet wird. Es gibt jedoch zwei wesentliche Unterschiede. Erstens beruht eine RD-Analyse auf schwächeren Annahmen. Insbesondere muss das Instrument nicht grundsätzlich unabhängig sein. So könnten im Falle einer Bewertung der eingereichten Vorhaben Unternehmen mit guten Vorhaben eher einen Beihilfeantrag stellen als Unternehmen mit schlechten Vorhaben. Es ist lediglich erforderlich, dass im Bereich der Schwelle die Antragswahrscheinlichkeit nicht unstet ist. Der zweite Unterschied besteht darin, dass die Schätzungen nur auf Unternehmen beruhen, die knapp über oder knapp unter der Schwelle liegen. Die schwächeren Annahmen haben somit ihren Preis: Schätzungen auf der Basis von RD-

28 Eine formale und vollständige Beschreibung von RD-Analysen findet sich in Imbens und Lemieux (2008).

29 Formal ist zwischen zwei Arten von RD-Analysen zu unterscheiden, und zwar zwischen scharfen und unscharfen RD-Analysen. Bei scharfen RD-Analysen, auf die hier abgehoben wird, werden alle Unternehmen, die über einer bestimmten Schwelle liegen, behandelt und keine anderen. Bei unscharfen RD-Analysen ist die Unstetigkeit weniger drastisch: Die Wahrscheinlichkeit der Behandlung ist unstet (es gibt einen Bruch/Sprung), sie ändert sich jedoch nicht von 0 auf 1. Folglich liegt bei Beihilferegulungen mit Förderfähigkeitsbedingungen nur dann eine scharfe RD-Analyse vor, wenn die Förderfähigkeit als die Behandlung angesehen wird. Wird hingegen der Erhalt einer Beihilfe als die Behandlung betrachtet, so liegt eine unscharfe RD-Analyse vor. Wenn die Beihilfe auf der Grundlage einer Punktebewertung gewährt wird, werden ausschließlich Unternehmen betrachtet, die einen Beihilfeantrag gestellt haben, und es liegt eine scharfe RD-Analyse vor.

Analysen sind noch stärker auf den Bereich in unmittelbarer Nähe der Schwelle fokussiert, als dies bei IV-Schätzungen im Allgemeinen der Fall ist. Wenn die Auswirkungen der Beihilfe bei Unternehmen, die weiter von der Schwelle entfernt liegen, abweichen, erlauben RD-Analysen-Schätzungen keine getreuen Rückschlüsse auf die Auswirkungen der Beihilfe auf alle Empfänger.

Die Fokussierung auf den Bereich in unmittelbarer Nähe der Schwelle kann Schätzungen dann beeinträchtigen, wenn fernab der Schwelle mit einer großen Abweichung der Auswirkungen zu rechnen ist. Zudem können einzelne Unternehmen jenseits der Grenze sehr stark durch die Maßnahme betroffen sein, etwa wenn es erhebliche Verlagerungseffekte gibt. In diesem Fall wäre die Durchführung einer RD-Analyse an der geografischen Grenze keine gute empirische Strategie. Ein großer Vorteil von RD-Analysen besteht in der Fokussierung auf eine geringe Bandbreite in unmittelbarer Nähe der Schwelle. Bei einer großen Bandbreite kann nicht davon ausgegangen werden, dass die Auswirkungen der anderen Merkmale konstant sind. Dieses Problem wird in der Regel nicht dadurch gelöst, dass die Messgrößen überprüft werden, was über eine besondere Funktionsform erfolgt.

Eine grafische Überprüfung der Daten kann Aufschluss über die Zuverlässigkeit der der RD-Analyse zugrundeliegenden Annahme geben. Insbesondere sollten drei Aspekte geprüft werden: 1. Liegt an der Schwelle eine Unstetigkeit in Bezug auf die Gewährung der Beihilfe vor? 2. Liegt bei den zu messenden Auswirkungen eine Unstetigkeit am selben Punkt vor und keine vergleichbare Unstetigkeit an einem anderen Punkt? 3. Liegt keine Unstetigkeit der anderen Parameter vor, die mit den Auswirkungen korrelieren, einschließlich der Beihilfebeantragungswilligkeit?

Ferner könnten im Hinblick auf eine Evaluierung der Regulierung Unstetigkeiten bewusst geschaffen werden. Insbesondere die Anlaufphase der Maßnahmen könnte genutzt werden, um Unstetigkeiten zu erzeugen und so die Auswirkungen der Maßnahmen herauszustellen.

C. Strukturelle Schätzung

In einigen Fällen ist es möglich, einen Schritt weiter zu gehen und ein theoretisches Modell, beispielsweise für die Investitionen eines Unternehmens, mit den Daten abzugleichen, um die wichtigsten relevanten Parameter zu erhalten. Dieser Ansatz unterscheidet sich qualitativ von den oben dargelegten Ansätzen. Bei einer strukturellen Schätzung wird ein vollständig spezifiziertes theoretisches Modell des Verhaltens von Unternehmen zugrunde gelegt. Aus der Schätzung ergeben sich dann die Parameter, die das Verhalten des Unternehmens bestimmen. Dies ermöglicht eine präzise Evaluierung der Bestimmungsfaktoren für das individuelle Verhalten der Unternehmen sowie Simulationen über die Effizienz anderer Instrumente. Für strukturelle Schätzungen sind jedoch in der Regel mehr Ressourcen und Daten sowie umfangreichere Annahmen erforderlich.

Es ist nicht möglich, genauere Orientierungshilfen zur strukturellen Schätzung zu erteilen, da Ermittlung, Schätzung und Schlussfolgerung von Fall zu Fall vorzunehmen sind. Es gelten jedoch die zuvor dargelegten allgemeinen Hinweise. Erstens muss das theoretische Modell

mit den wichtigsten stilisierten Fakten des Marktes übereinstimmen. Zweitens müssen vernachlässigte Merkmale und die Auswahl ausdrücklich und angemessen berücksichtigt werden.

D. Weitere methodische Anmerkungen

Heterogenität der Behandlungsauswirkungen

In den vorstehenden Abschnitten wurde auf die Beurteilung der durchschnittlichen Auswirkungen der Behandlung auf die Behandelten (konkret: der Beihilfe auf die Empfänger) eingegangen. Es ist offenkundig, dass die Auswirkungen einer Beihilfe von Empfänger zu Empfänger unterschiedlich ausfallen. Diese Heterogenität kann viele Ursachen haben und unterschiedliche Folgen nach sich ziehen. Eine wichtige Folge kann darin bestehen, dass, wenn eine Beihilfe bei einigen Unternehmen sehr wirksam ist, bei anderen hingegen kaum, die durchschnittliche Wirkung statistisch insignifikant sein kann. Somit bedeutet der Umstand, dass eine Beihilfe keine statistisch signifikanten Auswirkungen hat, nicht, dass sie für kein Unternehmen Auswirkungen hat. Aus beihilfepolitischer Sicht ist die durchschnittliche Wirkung einer Regelung ein sehr interessanter erster Indikator. Ebenso wichtig für die bessere Ausgestaltung von Beihilferegulungen ist es jedoch, die für die Heterogenität der Auswirkungen ausschlaggebenden Faktoren zu ermitteln. Dadurch wird es möglich, die Beihilfe gezielt auf diejenigen Unternehmen auszurichten, bei denen sie am effizientesten wirkt und am wenigsten Wettbewerbsverzerrungen hervorruft.

Folglich sollten die Auswirkungen der Beihilfe, wann immer dies möglich ist, für verschiedene Arten von Unternehmen wie z. B. kleine und große Unternehmen, junge und alte Unternehmen, innovative Unternehmen, kreditabhängige Unternehmen usw. geschätzt werden.³⁰

Verzerrungen zulasten nichtgeförderter Unternehmen

Die Beurteilung der direkten oder indirekten Auswirkungen einer Regelung auf Nichtteilnehmer ist sehr wichtig für die Evaluierung einer staatlichen Beihilfe. Staatliche Beihilfen können aufgrund von Auswirkungen auf Unternehmen, die keine Beihilfen erhalten, Verzerrungen auf Märkten hervorrufen, beispielsweise wenn Know-how von Empfängern übertragen wird (Wissens-Spillover) oder wenn nichtgeförderte Unternehmen im Vergleich zu Beihilfeempfängern an Wettbewerbsfähigkeit verlieren.

Zudem können die Auswirkungen auf nichtgeförderte Unternehmen oder Standorte auch Auswirkungen auf die Gültigkeit der Evaluierung haben. Bisweilen sind die Auswirkungen

30 Ein anderer Ansatz besteht darin, für Unternehmen, die sich an unterschiedlichen Positionen der bedingten Verteilung befinden, systematisch unterschiedliche Auswirkungen der Behandlung anzunehmen. Es gibt eine wachsende Fachliteratur über die Schätzung solcher quantilsbezogener Behandlungsauswirkungen, ausgehend von Abadie, Angrist und Imbens (2002). Dies ist sehr nützlich, wenn es darum geht, die Merkmale der Heterogenität der Behandlung zu verstehen. Von einem rein beihilfepolitischen Standpunkt hingegen ist dies weniger nützlich, außer wenn es möglich ist, Maßnahmen direkt auf einzelne Unternehmen an einer bestimmten Position in der bedingten Verteilung auszurichten.

einer Regionalbeihilfe teilweise lediglich auf Verlagerungen zurückzuführen, etwa wenn Unternehmen, die knapp außerhalb einer geförderten Region ansässig waren, ihren Standort just in die Region hinein verlegen. In einem solchen Fall würde eine RD-Analyse an der Regionalgrenze hauptsächlich den Verlagerungseffekt erfassen, so dass die tatsächlichen aggregierten Auswirkungen der Beihilfe möglicherweise überbewertet würden. In einer solchen Situation muss eine weitere empirische Strategie verwendet werden. So kann es etwa sinnvoll sein, die Stimmigkeit der Evaluierung anhand größerer Regionen zu überprüfen.

E. Daten

Wenn es darum geht, die Auswirkungen von Beihilfen zu evaluieren, werden unbedingt geeignete mikroökonomische Daten benötigt. Diese Daten von Beihilfeempfängern und Nichtempfängern müssen kohärent sein. Daher müssen sie von derselben Quelle stammen, ausgenommen natürlich Informationen über die Beihilfe selbst. Die Daten sollten auf der tiefstmöglichen Gliederungsebene zugänglich sein, wenngleich in manchen Fällen später eine gewisse Aggregation erforderlich sein kann.

Erfasst werden müssen Daten zu den Ergebnisindikatoren sowohl für die Behandlungs- als auch die Kontrollgruppe, einschließlich des Zeitpunkts der Ergebnisermittlung. Ferner sollten möglichst umfassende Angaben über Faktoren gesammelt werden, die Einfluss auf die Ergebnisse sowie auf die Entscheidung des Unternehmens bzw. der Einrichtung haben, sich an dem Förderprogramm zu beteiligen. Diese Daten sind erforderlich, um zu prüfen, ob Unterschiede zwischen der Behandlungs- und der Kontrollgruppe vorliegen. Auf Ebene der Unternehmen werden möglicherweise z. B. Angaben zu Standort, Größe und Demografie sowie Produktions-Inputs benötigt.

Die naheliegendsten Datenquellen sind Verwaltungsakten wie Angaben aus der Steuerbilanz, oder Daten aus nationalen Erhebungen. Diesen Quellen lassen sich Informationen über den Standort und die Tätigkeit von Unternehmen und bisweilen auch einzelner Werke entnehmen. Diese Quellen liefern in der Regel Zahlen zu Investitionen und Umsätzen, aufgeschlüsselt nach Unternehmenssparten, und ermöglichen so die Berechnung von Finanzkennzahlen. Umfangreiche nationale oder gemeinschaftliche Erhebungen wie Innovationserhebungen der Gemeinschaft können ebenfalls von Interesse sein. Sie umfassen eine große und repräsentative Auswahl und bieten sehr umfassende Informationen zu spezifischen Themen. Schließlich stellen auch zusammengeführte Datensätze zu Arbeitgebern und Arbeitnehmern eine relevante Informationsquelle dar. Anhand dieser Angaben lassen sich in der Regel für die einzelnen Unternehmensstandorte bestimmte Beschäftigungsmerkmale ermitteln. Dies kann von entscheidender Bedeutung sein, wenn es um die geografische Dimension der Beschäftigung geht.

Abgesehen von Ergebnisindikatoren und Empfängermerkmalen sind auch Angaben über die Beihilfe und das Bewilligungsverfahren erforderlich. Diese Informationen liefert in der Regel die Bewilligungsbehörde. Benötigt werden etwa Angaben zum Umfang und zum Zeitpunkt der Gewährung der Beihilfe an die Empfänger. Auch allgemeine Angaben über den Bewilligungsprozess können überaus nützlich sein. Angaben über Antragsteller, deren

Anträge abgelehnt wurden, sind insbesondere dann von Bedeutung, wenn im Rahmen des Bewilligungsverfahrens ein Bewertungsmechanismus zur Anwendung kommt.³¹

Der Zugang zu solchen vertraulichen Angaben unterliegt in der Regel gewissen Beschränkungen. Daher sollte dafür gesorgt werden, dass alle Mitglieder des wissenschaftlichen Teams, die die Evaluierung durchführen, rechtzeitig Zugang zu diesen Daten erhalten. Solche Daten von Verwaltungseinrichtungen werden zudem meist erst nach einer bestimmten Vorlaufzeit zugänglich gemacht. Folglich muss bei der Erstellung eines Evaluierungsplans auch die Zugänglichkeit von Daten berücksichtigt werden.

³¹ Angaben über abgelehnte Anträge sind besonders wertvoll für Studien, die auf einer Regressions-Diskontinuitäts-Analyse beruhen.

F. Beispiele

Beispiel 1 (Regionalbeihilfen): Criscuolo u. a. (2012)³² haben die Regionalbeihilferegelung „Regional Selective Assistance“ (RSA) im Vereinigten Königreich im Zeitraum 1986 bis 2004 untersucht. In diesem Zeitraum wurden über die RSA-Regelung nach freiem Ermessen Unternehmen in benachteiligten Gebieten Zuschüsse gewährt. RSA war die wichtigste Regelung zur Unternehmensförderung im Vereinigten Königreich. Der Umfang der im Rahmen von RSA gewährten Beihilfen war in Regionalbeihilfleitlinien und insbesondere in der Fördergebietskarte festgelegt. Diese Fördergebietskarte wurde von Zeit zu Zeit angepasst. Das Hauptkriterium für die Förderfähigkeit einer Region war grundsätzlich deren relative Stellung in Bezug auf das Pro-Kopf-BIP und die Arbeitslosenquote. Folglich konnte sich der Status einer Region entweder aufgrund der regionalen Entwicklung oder aufgrund einer Veränderung des durchschnittlichen Pro-Kopf-BIP in der EU ändern (z. B. bei der Aufnahme neuer Mitgliedstaaten in die EU im Jahr 1995). Ferner wurden auch die Kriterien für die Förderfähigkeit bisweilen angepasst. Ob bestimmte Unternehmen förderfähig waren oder nicht, hing also nicht nur von der Lage der Unternehmen selbst, sondern auch von Ereignissen außerhalb des Vereinigten Königreichs und von den Änderungen der Förderbedingungen ab. Criscuolo u. a. (2012) analysieren vor allem, welche Auswirkungen solche Änderungen der Förderfähigkeit auf Investitionstätigkeit, Beschäftigung und Produktivität hatten, und können so die Auswirkungen der Beihilfen schlüssig ermitteln.³³

Beispiel 2 (Unterstützung von Unternehmen): Martini und Bondonio (2012)³⁴ haben zwei Förderregelungen für Unternehmen untersucht: eine Investitionszuschussregelung für Unternehmen in ganz Italien (Gesetz 488) und verschiedene KMU-Regelungen in der Region Piemont. Die erste Evaluierung ist von besonderem Interesse. Darin werden Unternehmen, deren Beihilfeantrag bewilligt wurde (d. h. Beihilfeempfänger) mit vergleichbaren Unternehmen verglichen, deren Beihilfeantrag abgelehnt wurde, da die zur Verfügung stehenden Beihilfemittel erschöpft waren. Die Heranziehung abgelehnter Antragsteller ist insofern besonders zweckmäßig, als dadurch die auswahlbedingte Verzerrung vermieden werden kann, die sich in der Regel ergibt, wenn einfach Antragsteller und Nicht-Antragsteller verglichen werden. Die in der Vergleichsgruppe enthaltenen abgelehnten Antragsteller hatten die erste Qualitätsprüfung bestanden und verfügten somit über ein geeignetes Investitionsvorhaben. Sie hatten also genau wie die Beihilfeempfänger vor, in ein realisierbares und zweckmäßiges Vorhaben zu investieren. Sie erhielten jedoch aufgrund der begrenzten zur Verfügung stehenden Haushaltsmittel keine Beihilfe. Die Unterschiede in der

32 Criscuolo, C., R. Martin, H. Overman und J. Van Reenen, 2012. ‘The causal effects of an industrial policy’, CEPR Discussion Papers 8818, C.E.P.R. Discussion Papers.

33 Vom technischen Ansatz her verwenden Criscuolo u. a. (2012), wie weiter oben in diesem technischen Anhang dargelegt, Instrumentvariablen.

34 Bericht für die GD REGIO. A. Martini, D. Bondonio: ‘Counterfactual impact evaluation of cohesion policy: impact and cost effectiveness of investment subsidies in Italy’ (2012).

Entwicklung von (grade noch) angenommenen Antragstellern und (knapp) abgelehnten Antragstellern ließen stichhaltige Rückschlüsse auf die Auswirkungen der Beihilfe zu.

Beispiel 3 (Kreditgarantien): Lelarge, Sraer und Thesmar (2010) haben die Auswirkungen eines Kreditgarantieprogramms in Frankreich untersucht. Das Programm „SOFARIS“ versichert Kreditgeber durch Garantien gegen das Ausfallsrisiko des Kreditnehmers. Die Kreditnehmer zahlen eine Versicherungsprämie, die jedoch subventioniert wird. Lelarge, Sraer und Thesmar (2010) legen die Auswahlprozesse ausführlich dar. Erstens sind Unternehmen mit rentableren Vorhaben eher bereit, das Entgelt für die Garantie zu zahlen. Zweitens wählen die Programmverwalter bevorzugt sozial wünschenswerte Vorhaben aus, die ansonsten möglicherweise keine privaten Mittel erhielten. Somit findet sowohl auf Ebene der antragstellenden Unternehmen als auch im Zuge des Bewilligungsverfahrens eine Auswahl statt. Dies dürfte Auswirkungen auf die Ergebnisse einfacher Evaluierungen haben, die beispielsweise auf klassisch-linearen Regressionen oder Vergleichen mit dem am ehesten vergleichbaren Unternehmen beruhen.³⁵ Der faktische und institutionelle Kontext des Beihilfeprogramms bietet jedoch Anhaltspunkte, wie dessen Auswirkungen ermittelt werden können. Das Programm wurde Ende der 1980er Jahre eingeführt und war zunächst auf Unternehmen des verarbeitenden Gewerbes und Unternehmensdienstleister beschränkt. Im Jahr 1995 erhielt das Programm eine höhere Mittelausstattung und wurde für weitere Wirtschaftszweige (Bau, Groß- und Einzelhandel, Transport, Hotels und Restaurants und persönliche Dienstleistungen) geöffnet. Lelarge, Sraer und Thesmar (2010) vergleichen die nunmehr förderfähigen Unternehmen mit denen, die bereits vorher für eine Förderung in Betracht kamen, um die Auswirkungen des Programms auf verschiedene Indikatoren wie Verschuldung, Beschäftigung, Kapitalwachstum, Finanzausgaben und Insolvenzwahrscheinlichkeit zu ermitteln. Grundsätzlich ist davon auszugehen, dass die Unternehmen der beiden Vergleichsgruppen unterschiedlich abschneiden. Und da die Unternehmen durch ähnliche makroökonomische Schocks beeinträchtigt werden, dürften diese Unterschiede über die Zeit bestehen bleiben – außer wenn Auswirkungen des Beihilfeprogramms die Entwicklung der Unternehmen beeinflussen.³⁶

35 In diesem Fall sind Matching-Verfahren, hier Eins-zu-eins-Vergleiche mit dem nächsten Nachbarn, nicht besser für die Lösung von Auswahlproblemen geeignet als die herkömmliche Methode der kleinsten Quadrate (Fehlerquadratmethode). Wie bereits in diesem technischen Anhang erläutert, sind Matching-Verfahren in der Regel nicht geeignet, um bei Fehlen natürlicher Experimente Aufschluss über Auswahlwirkungen zu erhalten.

36 In der Praxis verwenden die Autoren ein Heckman-Auswahlmodell mit einer Ausschlussvariablen auf Unternehmensebene und einer klassischen IV-Strategie auf Sektorebene. Nähere Angaben zu diesen Methoden sind den weiter oben in diesem technischen Anhang enthaltenen Ausführungen zu entnehmen.

Beispiel 4 (Kreativzuschüsse): Bakhshi u. a.³⁷ greifen auf ein Experiment in Form einer randomisierten kontrollierten Studie zurück, um die Auswirkungen einer Regelung zur Förderung innovativer Unternehmen zu analysieren. Die Pilotstudie, die 2009 in Manchester begann, lief folgendermaßen ab: An kleine und mittlere Unternehmen, die Anträge für Investitionen in kreative Vorhaben wie die Erstellung von Websites, Videos und Marketingkampagnen gestellt hatten, wurden nach dem Zufallsprinzip Gutscheine in Form sogenannter „Creative Credits“ (Kreativzuschüsse) vergeben, um herauszufinden, ob diese Zuschüsse tatsächlich Auswirkungen auf die Innovation hatten. Die Kreativzuschüsse brachten vollkommen neue Beziehungen zwischen KMU und kreativen Unternehmen hervor. So erhöhte sich durch die Gewährung eines Kreativzuschusses die Wahrscheinlichkeit, dass ein Unternehmen ein innovatives Vorhaben mit einem kreativen Unternehmen realisiert, mit dem es zuvor nicht zusammengearbeitet hatte, um mindestens 84 Prozent. Die Untersuchungen ergaben, dass die Unternehmen, die Kreativzuschüsse erhalten hatten, in den sechs Monaten nach Abschluss ihrer kreativen Vorhaben kurzfristige Innovations- und Umsatzanstiege verzeichneten. Diese positiven Auswirkungen waren jedoch nicht nachhaltig: Nach 12 Monaten bestand kein statistisch signifikanter Unterschied mehr zwischen den Unternehmen, die einen Zuschuss erhalten hatten, und den anderen Unternehmen. Dem Bericht zufolge wären diese Ergebnisse bei Anwendung der üblichen Evaluierungsmethoden der Regierung verborgen geblieben; daher solle bei der Evaluierung von Maßnahmen zur Förderung von Unternehmen in stärkerem Maße auf randomisierte kontrollierte Studien zurückgegriffen werden.

Beispiel 5 (Förderung von FuEuI): Einiö (2013) hat die Auswirkungen von FuE-Fördermaßnahmen auf FuE-Investitionen, Beschäftigung und Produktivität im Zeitraum 2000-2006 analysiert. Die untersuchten Maßnahmen wurden von der finnischen Innovationsagentur Tekes gewährt, die für den Großteil der FuE-Förderung im Land zuständig ist. Für die Studie wird der Umstand ausgenutzt, dass die der Innovationsagentur Tekes für FuE-Förderung zur Verfügung stehenden Mittel in bestimmten Regionen höher sind, da für Teile Nord- und Ostfinnlands (Ziel-1-Gebiete) mehr EFRE-Mittel bereitgestellt werden. Die Ziel-1-Gebiete wurden ursprünglich im Beitrittsvertrag Finnlands von 1995 festgelegt, und zwar auf der Grundlage der Bevölkerungsdichte-Regel (nicht mehr als 8 Einwohner pro Quadratkilometer). Aufgrund des im Verhältnis größeren FuE-Förderbudgets war die Wahrscheinlichkeit, eine Förderung zu erhalten, in Ziel-1-Gebieten höher als in anderen Teilen des Landes. Dies führte zu regionalen Unterschieden in der Behandlung: In Ziel-1-Gebieten kam ein deutlich größerer Anteil der Unternehmen in den Genuss von Fördermitteln. Da die regionale Zuweisung auf der Grundlage der zuvor festgelegten Bevölkerungsdichte-Regel erfolgte (die auf Daten von 1993 beruhte) und nicht beispielsweise auf Basis der erwarteten künftigen FuE-Investitionen oder der erwarteten

³⁷ Report for Nesta, Creative Credits, a randomised controlled industrial policy experiment, Bakhshi, H., J. Edwards, S. Roper, J. Scully, D. Shaw, L. Morley und N. Rathbone, Juni 2013, abrufbar unter http://www.nesta.org.uk/sites/default/files/creative_credits.pdf.

künftigen wirtschaftlichen Entwicklung, wird in der Studie die Bevölkerungsdichte von 1993 überprüft, so dass Bedenken hinsichtlich der regionalen Auswahl wirksam Rechnung getragen wird. In der Praxis werden die Auswirkungen der Behandlung anhand eines IV-Ansatzes untersucht, wobei ein Indikator für die Ziel-1-Region als Instrument für die Aufnahme in das Förderprogramm dient. Auf diese Weise werden die Auswirkungen der Unterstützung auf die Unternehmen ermittelt, die aufgrund der umfangreicheren Fördermittel im Ziel-1-Gebiet in den Genuss der Förderung kamen. Die Stichhaltigkeit dieses Ansatzes wird dadurch bestätigt, dass sich die Entwicklungstrends von Unternehmen, die später im Rahmen des Programms eine Förderung erhielten, und der Kontrollgruppe im Vorfeld der Förderung nicht unterschieden. Einiö (2013) stellt positive Auswirkungen auf FuE-Investitionen, Beschäftigung und Umsatz der Unternehmen fest, die dank der zusätzlichen FuE-Fördermittel in ihrer Region eine FuE-Förderung erhielten. Wenngleich keine sofortigen Auswirkungen auf die Produktivität vorlagen, enthält die Studie Anhaltspunkte für langfristige Produktivitätsanstiege.

Anhang II: Aufstellung möglicher Ergebnisindikatoren

Die nachstehende Aufstellung dient lediglich als Beispiel zur Veranschaulichung. Die im jeweiligen Fall zugrunde zu legenden Ergebnisindikatoren sollten entsprechend den Zielen der Beihilferegelung und der Evaluierung festgelegt werden.

Direkte Auswirkungen der Beihilfe auf Ebene der Empfänger

	ART DES ERGEBNISSES	ERGEBNISINDIKATOREN
Regionalbeihilfen	Positive Auswirkungen	Private Investitionen in Ergänzung zur staatlichen Unterstützung Beschäftigungszunahme in geförderten Unternehmen
Forschungs-, Entwicklungs- und Innovationsbeihilfen	Zunahme der FuEuI-Aktivitäten	Private Investitionen in Ergänzung zur staatlichen Unterstützung Anstieg der FuEuI-Ausgaben seitens der geförderten Unternehmen Zahl der neuen Wissenschaftler in geförderten Unternehmen Zahl der neu eingetragenen Patente Zahl der Unternehmen, die bei der Markteinführung neuer Produkte unterstützt werden
Umweltbeihilfen	Positive Auswirkungen auf den Umweltschutz	Verringerung der CO ₂ -Emissionen der geförderten Unternehmen Zusätzliche Kapazität der Energieerzeugung aus erneuerbaren Quellen Verringerung des Anteils der Abfälle, die deponiert oder verbrannt werden Zahl der Sanierungen schadstoffbelasteter Standorte
	Frühzeitige Anpassung an Umweltnormen	Prozentsatz der Unternehmen, die neue Umweltnormen mindestens X Monate/Jahre vor deren Inkrafttreten erfüllen [mindestens ein Jahr zuvor, höhere Beihilfeintensität zulässig, wenn Erfüllung mehr als drei Jahre zuvor erfolgt]

Energie(infrastruktur)beihilfen	Reduzierung des Energieverbrauchs	Zahl der Haushalte mit niedrigerem Energieverbrauch Rückgang des jährlichen Primärenergieverbrauchs in öffentlichen Gebäuden Zahl der zusätzlichen, an intelligente Netze angeschlossenen Energiekunden
	Förderung erneuerbarer Energien	Anteil der Energie, die aus erneuerbaren Quellen gewonnen wird
Risikofinanzierungsbeihilfen	Positive Auswirkungen	Im Fonds erzielte Renditen Mobilisierung privater Investitionen Zahl der Unternehmen, die Risikokapital erhalten
	Fehlschläge	Im Durchschnitt schlechte Entwicklung der geförderten Unternehmen aufgrund mangelhafter Geschäftsführung/unzureichender privater Beteiligung
	Mangelnde Diversifizierung	Zu kleine/regional eingeschränkte Fonds mit begrenzten Renditeaussichten, die zu unattraktiv für private Investoren sind
Breitbandbeihilfen	Verbesserung der Breitbandversorgung	Zahl der zusätzlichen Haushalte, denen ein Breitbandanschluss von mindestens 30 Mbit/s bereitgestellt wird Zahl der zusätzlichen Haushalte, denen ein Breitbandanschluss von mindestens 100 Mbit/s bereitgestellt wird bzw. die ihn in Anspruch nehmen
	Effizienz	Investitionskosten/Beihilfe pro bereitgestelltem Anschluss (Zahl der erschlossenen Wohnungen) Zahl der Haushalte, die die neuen Dienste in Anspruch nehmen
Rettungs- und Umstrukturierungsbeihilfen	Positive Auswirkungen	Aufrechterhaltung der Beschäftigung und der Wirtschaftstätigkeit auf Ebene des Unternehmens und der Region Entwicklung der Marktanteile und der Produktivität der geförderten Unternehmen

Luftverkehrsbeihilfen	Positive Auswirkungen	<p>Zahl der Luftverkehrsunternehmen, die den Flughafen bedienen</p> <p>Private Investitionen in Ergänzung zur staatlichen Unterstützung</p> <p>Anstieg der regionalen Produktivität und/oder der Bruttowertschöpfung (BWS)</p>
	Negative Auswirkungen	<p>Verdoppelung defizitärer Infrastrukturen oder Luftverkehrsstrecken</p> <p>Rückgang des Verkehrsaufkommens an der bestehenden Infrastruktur (z. B. wegen anderer Flughäfen im Einzugsgebiet oder anderer Verkehrsmittel)</p>

Indirekte Auswirkungen der Beihilferegelung

	ART DES ERGEBNISSES	ERGEBNISINDIKATOREN
Mögliche positive Auswirkungen	Makroökonomische Vorteile	Schaffung von Arbeitsplätzen Anstieg der Produktivität und/oder der Bruttowertschöpfung (BWS)
	Diversifizierung der regionalen Wirtschaft	Zahl der Wirtschaftszweige (unterschiedliche NACE-Kodes)
	Verstärkte Zusammenarbeit zwischen privaten und öffentlichen Beteiligten	Zahl der Unternehmen, die mit Forschungseinrichtungen zusammenarbeiten
	Positive externe Effekte/Spillover-Effekte	Zahl der indirekt Begünstigten (z. B. Zahl der Dritten, die Zugang zu der Einrichtung haben) Entwicklung der Beschäftigung oder Tätigkeiten in anderen Unternehmen und Regionen (Luftverkehr) Zahl der Einwohner im Einzugsgebiet, deren Verkehrsanbindung sich verbessert hat
Mögliche negative Auswirkungen auf Wettbewerb und Handel	Sektorale Verzerrung	Im Rahmen einer sektorübergreifenden Regelung wurden Beihilfen hauptsächlich für einen einzigen Wirtschaftszweig gewährt
	Übermäßige Förderung von unrentablen Unternehmen bzw. Unternehmen mit geringer Produktivität (deren Ausscheiden verhindert wird)	Vergleich des Anteils der Unternehmen mit hoher und mit geringer Produktivität
	Übermäßige Förderung etablierter Unternehmen	Vergleich des Anteils der alten und der jungen Unternehmen
	Erhöhung der Marktmacht	Veränderung der Marktmacht eines beherrschenden Unternehmens
	Standorteffekt	Verlagerung aus einer ärmeren in eine stärker entwickelte Region
	Wahrung der Versorgungssicherheit	Festhalten an CO ₂ -intensiven Energiequellen Prüfung, ob Bedenken in Bezug auf Stromausfälle gerechtfertigt sind und fortbestehen Abschottung der nationalen Strommärkte

	Energieinfrastruktur	<p>Abschottung der nationalen Strommärkte</p> <p>Erhöhung der Marktmacht eines etablierten Betreibers</p>
	Rettung und Umstrukturierung	<p>Entwicklung der Beschäftigung oder Tätigkeiten in anderen Unternehmen und Regionen</p> <p>Entwicklung der Marktanteile und der Produktivität der geförderten Unternehmen</p>
	Luftverkehr	<p>Verdoppelung defizitärer Infrastrukturen oder Luftverkehrsstrecken</p> <p>Rückgang des Verkehrsaufkommens an der bestehenden Infrastruktur (z. B. andere Flughäfen im Einzugsgebiet oder andere Verkehrsmittel)</p>

Anhang III: Glossar

Ausgangswert	Wert des Indikators, bevor mit der in Rede stehenden Maßnahme begonnen wird.
Kontrollgruppe	Für die kontrafaktische Analyse müssen Unternehmen oder Kontrollgruppen ermittelt werden, die mit den in Rede stehenden Unternehmen am ehesten vergleichbar sind, d. h. eine Gruppe von Unternehmen, die der Gruppe von Unternehmen, die die Beihilfe erhalten hat, am ähnlichsten ist – mit Ausnahme der Tatsache, dass sie keine Beihilfe erhalten hat.
Kontrafaktische Fallkonstellation	Um abschätzen zu können, wie sich die Beihilfe auf die Beihilfeempfänger auswirken könnte, muss eine „kontrafaktische Fallkonstellation“ konstruiert werden, d. h., es wird untersucht, was wahrscheinlich geschehen wäre, wenn die Beihilfeempfänger keine Beihilfe erhalten hätten.
Evaluierung	Systematische Erfassung und Analyse von Informationen über die Regelungen und Vorhaben, ihren Zweck und ihre Durchführung, um Erkenntnisse über ihre Auswirkungen zu gewinnen, die als Grundlage für Entscheidungen dienen können. Evaluierungen sollen dazu beitragen, die Wirksamkeit der Maßnahmen zu verbessern und sachlich fundierte Entscheidungen zur laufenden oder künftigen Planung zu ermöglichen.
Auswirkungen	Veränderungen, die einer Maßnahme schlüssig zugeschrieben werden können. Dasselbe wie „Effekte“ der Maßnahme oder „Beitrag zu Verhaltensänderungen“.
Indikator	Variable, die quantitative oder qualitative Informationen zu einem Phänomen liefert. Normalerweise besteht sie aus einem Wert und einer Maßeinheit.
Methode	Methoden sind Gruppen von Evaluierungsverfahren und -instrumenten, die verschiedene Zwecke erfüllen. In der Regel bestehen sie aus Verfahren und Protokollen, die eine Systematisierung und konsistente Durchführung von Evaluierungen gewährleisten. Methoden können sich auf die Erfassung und Analyse von Informationen und Daten konzentrieren, qualitativer oder quantitativer Art sein und/oder dazu beitragen, Maßnahmen zu beschreiben, zu erläutern, vorauszusagen oder zu untermauern. Die Wahl der Methode ergibt sich aus der Art der Maßnahmen, den Fragen, die bei der Evaluierung gestellt werden, und der Art der Untersuchung – kausal, investigativ, normativ, usw.
Ergebnis	Konkrete Dimension des Wohlergehens der Menschen, die dem politischen Handeln zugrunde liegt und die durch die nach einer politischen Strategie gestalteten und durchgeführten Maßnahmen geändert werden soll. Beispiele: Mobilität in einem Gebiet oder die Kompetenz in einem bestimmten Wirtschaftszweig.
Ergebnisindikator	Indikator zur Beschreibung eines bestimmten messbaren Aspekts eines Ergebnisses. Beispiele: Fahrzeit von W nach Y bei durchschnittlicher Geschwindigkeit als Mobilitätsaspekt, Testergebnisse zu einem bestimmten Thema als Kompetenzaspekt, Anteil von Unternehmen, denen ein Kredit verweigert wurde, als Aspekt der Kreditrationierung bei Banken.

Anhang IV: Literaturhinweise

Abadie, A., J. Angrist und G. W. Imbens (2002), 'Instrumental Variables Estimates of the Effect of Subsidised Training on the Quantiles of Trainee Earnings,' *Econometrica*, 70(1), 91–117.

Abadie, A., A. Diamond und J. Hainmueller (2007), 'Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Programme,' *Journal of the American Statistical Association*, Juni 2010, Bd. 105, Nr. 490.

Angrist, J. und A. Krueger (1991), 'Does Compulsory School Attendance Affect Schooling and Earnings,' *Quarterly Journal of Economics*, 106.

Angrist, J. und J. Pischke (2008), "Mostly Harmless Econometrics: An Empiricist's Companion", Princeton University Press.

Angrist, J. und J. Pischke (2010), "Mostly Harmless Econometrics: How Better Research Design Is Taking the Con out of Econometrics.' *Journal of Economic Perspectives*, 24(2): 3-30.

Bakhshi, H., J. Edwards, S. Roper, J. Scully, D. Shaw, L. Morley und N. Rathbone (2013), "Creative credits, a randomized controlled industrial policy experiment", Bericht für Nesta, abrufbar unter: http://www.nesta.org.uk/sites/default/files/creative_credits.pdf.

Bertrand, M., E. Duflo und S. Mullainathan (2004), 'How much should we trust differences-in-differences estimates?', *The Quarterly Journal of Economics*, 119, 249–275.

Bound, J., D. Jaeger und R. Baker (1995), 'Problems with Instrumental Variable Estimation When the Correlation Between the Instruments and the Endogenous is weak,' *Journal of the American Statistical Association*, 90(430), 443–450.

Criscuolo, C., R. Martin, H. Overman und J. Van Reenen, 2012. 'The causal effects of an industrial policy', CEPR Discussion Papers 8818, C.E.P.R. Discussion Papers.

Duflo, E., R. Glennerster und M. Kremer (2007), 'Using Randomisation in Development Economics Research: A Toolkit,' CEPR Discussion Papers 6059, C.E.P.R. Discussion Papers.

Duflo, E., und M. Kremer (2005), 'Use of Randomisation in the Evaluation of Development Effectiveness,' in *Evaluating Development Effectiveness*, Hrsg. O. Feinstein, G. K. Ingram, und G. K. Pitman. New Brunswick, New Jersey und London, U.K.: Transaction Publishers, Bd. 7, S. 205{232}.

Einiö, Elias (2013), 'R&D Subsidies and Company Performance: Evidence from Geographic Variation in Government Funding Based on the ERDF Population-Density Rule', *The Review of Economics and Statistics* (angekündigt).

Evaluierungsnormen der Europäischen Kommission. Abrufbar unter: http://ec.europa.eu/dgs/secretariat_general/evaluation/docs/standards_c_2002_5267_final_de.pdf.

Garicano, L. C. Lelarge und J. Van Reenen, (2012), 'Firm Size Distortions and the Productivity Distribution: Evidence from France,' CEP Discussion Papers dp1128, Centre for Economic Performance, LSE.

Givord, P. (2010), « Méthodes économétriques pour l'évaluation de politiques publiques », WPD3E n° G2010-08.

Givord, P., Rathelot, R. und P. Sillard (2013), 'Place-based tax exemptions and displacement effects: An evaluation of the Zones Franches Urbaines programme', *Regional Science and Urban Economics*, Bd. 43, Ausgabe 1, Januar 2013, S. 151-163.

Heckman, J. J. (1979), 'Sample Selection Bias as a Specification Error,' *Econometrica* 47, 153–161.

Imbens, G. and J. Wooldridge (2009), *Recent Developments in the Econometrics of Programme Evaluation*, *Journal of Economic Literature*, 47:1, 5-86.

Imbens, G. W., und T. Lemieux (2008), 'Regression discontinuity designs: A guide to practice,' *Journal of Econometrics*, 142(2), 615–635.

C. Lelarge, D. Sraer und D. Thesmar (2010), 'Entrepreneurship and Credit Constraints: Evidence from a French Loan Guarantee Programme,' NBER Chapters, in: *International Differences in Entrepreneurship*, S. 243-273, National Bureau of Economic Research, Inc.

Keane, M. P. (2010), 'A Structural Perspective on the Experimentalist School.' *Journal of Economic Perspectives*, 24(2): 47-58.

Martini, A. und D. Bondonio (2012), 'Counterfactual impact evaluation of cohesion policy: impact and cost effectiveness of investment subsidies in Italy', Bericht für die Europäische Kommission, GD REGIO.

Nederlandse Rijksoverheid (2012), 'Durf te meten', Eindrapport Expertwerkgroep Effectmeting. Abrufbar unter <http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2012/11/23/durf-te-meten-eindrapport-expertwerkgroep-effectmeting.html>.

Nevo, A. und M. D. Whinston, (2010), 'Taking the Dogma out of Econometrics: Structural Modeling and Credible Inference.' *Journal of Economic Perspectives*, 24(2): 69-82.

Evaluation Norms and Standards der OECD. Abrufbar unter: <http://www.oecd.org/dac/evaluation/dcdndep/41612905.pdf>.

Sims, C. A. (2010), 'But Economics Is Not an Experimental Science.' *Journal of Economic Perspectives*, 24(2): 59-68.

Stock, J., J. Wright und M. Yogo (2002): 'A Survey of Weak Instruments and Weak Identification in Generalised Method of Moments,' Journal of Business and Economic Statistics, 20(4), 518-29.

United Nations Evaluation Group (2005), "Standards for Evaluation in the UN System".
Abrufbar unter: http://www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=22.

Wooldridge, J. (2002), "Econometric Analysis of Cross Section and Panel Data", Cambridge: MIT Press.

Weltbank (2003), "Independent Evaluation: Principles, Guidelines and Good Practice".
Abrufbar unter:
<http://siteresources.worldbank.org/INTDGF/Resources/Evaluation&LearningNote.pdf>.