

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages



Deutsch-Ukrainischer
Agrarpolitischer Dialog

beim Nationalen Verband der Landwirtschaftlichen Beratungsdienste der Ukraine NAPAMA

Agrarpolitischer Bericht

APD/APB/02/2021

Evidenzbasierte agrarpolitische Entscheidungsfindung – deutsche Erfahrungen und Optionen für die Ukraine

Prof. Dr. Thomas Herzfeld
PD Dr. Mirko Titze

Kiew, Mai 2021

Durchgeführt von



Ansprechpartner:
APD Ukraine
wul. Reytarska 29-b,
01030 Kiew
info@apd-ukraine.de
www.apd-ukraine.de

Über das Projekt „Deutsch-Ukrainischer Agrarpolitischer Dialog“ (APD)

Das Projekt Deutsch-Ukrainischer Agrarpolitischer Dialog (APD) wird vom Bundesministerium für Ernährung und Landwirtschaft (BMEL) seit 2006 zunächst bis Ende 2021 gefördert und in dessen Auftrag über den Mandatar GFA Consulting Group GmbH sowie eine Arbeitsgemeinschaft bestehend aus der IAK AGRAR CONSULTING GmbH (IAK), dem Leibniz-Institut für Agrarentwicklung in Transformationsökonomien (IAMO) und der AFC Agriculture and Finance Consultants GmbH durchgeführt. Projektträger ist der Nationale Verband der Landwirtschaftlichen Beratungsdienste der Ukraine „Dorada“. Der APD kooperiert mit der BVVG Bodenverwertungs- und -verwaltungs GmbH bei der Umsetzung wichtiger Komponenten zur Entwicklung einer effektiven und transparenten Bodenverwaltung in der Ukraine. Benefiziar ist das Ministerium für Agrarpolitik und Ernährung der Ukraine.

In Übereinstimmung mit marktwirtschaftlichen und ordnungspolitischen Grundsätzen und unter Berücksichtigung der sich aus dem EU-Ukraine-Assoziierungsabkommen ergebenden Entwicklungspotentiale soll das Projekt die Ukraine bei der Entwicklung einer nachhaltigen Landwirtschaft, einer effektiven Verarbeitungsindustrie und bei der Steigerung der internationalen Wettbewerbsfähigkeit sowie bei Schutz der nützlichen Ressourcen unterstützen. Dazu sollen vor allem deutsche, hier u.a. ostdeutsche, aber auch internationale, insbesondere EU-Erfahrungen bei der Gestaltung agrar- und forstpolitischer Rahmenbedingungen sowie bei der Organisation von entsprechenden Institutionen bereitgestellt werden.



www.apd-ukraine.de

Autor

Prof. Dr. Thomas Herzfeld
PD Dr. Mirko Titze

Disclaimer

Dieser Beitrag wird unter der Verantwortung des Deutsch-Ukrainischen Agrarpolitischen Dialogs (APD) veröffentlicht. Jegliche Meinungen und Ergebnisse, Schlussfolgerungen, Vorschläge und Empfehlungen beziehen sich auf die Autoren und müssen nicht den Ansichten des APD entsprechen.

INHALTVERZEICHNIS

ABBILDUNGSVERZEICHNIS	4
DIE WICHTIGSTEN ERGEBNISSE IM ÜBERBLICK.....	5
1 ZIELE UND MOTIVATION.....	6
2 EVIDENZBASIERTE ENTSCHEIDUNGSFINDUNG - DEUTSCHE ERFAHRUNGEN.....	8
2.1 INSTITUTIONELLE VORAUSSETZUNGEN.....	8
2.1.1 Rechtsstaatsprinzip.....	9
2.1.2 Standards in der Politik	9
2.1.3 Standards in der Wissenschaft.....	11
2.1.4 Standards in der Datengenerierung und Bereitstellung.....	12
2.1.5 Zusammenfassung zu den institutionellen Rahmenbedingungen	13
2.2 EINBETTUNG VON EVALUATIONEN IN EINEN POLITIKZYKLUS	13
2.2.1 Politische Themensetzung	14
2.2.2 Politikformulierung	15
2.2.3 Politische Entscheidungsfindung	16
2.2.4 Implementierung der Politik.....	17
2.2.5 Politikevaluation	18
2.2.6 Zusammenfassung zur Einbettung von Evaluationen in einen Politikzyklus	19
2.3 AUFGABEN IN EINEM SYSTEM EVIDENZBASIERTER EVALUATIONEN	19
2.3.1 Evaluationsforschung.....	20
2.3.2 Datenmanagement	21
2.3.3 Wissenstransfer.....	24
2.3.4 Zusammenfassung zu den Aufgaben in einem System evidenzbasierter Evaluationen.....	27
2.4 EIGNUNG KONTRAFAKTISCHER ANALYSETOOLS ZUR EVALUIERUNG WIRTSCHAFTSPOLITISCHER MASSNAHMEN	28
2.4.1 Experimente	29
2.4.2 Instrumentvariablen-Ansatz (IV).....	31
2.4.3 Regressions-Diskontinuitäts-Design (RDD).....	31
2.4.4 Differenz-von-Differenzen-Ansatz (DiD)	32
2.4.5 Zusammenfassung zur Eignung kontrafaktischer Analysetools.....	33

2.5 AKTEURE/INSTITUTIONEN, MECHANISMEN UND INSTRUMENTE ZUR EVIDENZBASIERTEN POLITIKGESTALTUNG IM AGRARBEREICH IN DEUTSCHLAND...	34
2.6 PRAXISBEISPIELE FÜR EVIDENZBASIERTE POLITIKGESTALTUNG UND KRITISCHE DISKUSSION DES ZUSAMMENWIRKENS VON AGRARPOLITIK UND -WISSENSCHAFT IN DEUTSCHLAND.....	35
2.6.1 Praxisbeispiele für evidenzbasierte Politikgestaltung in Deutschland	35
2.6.2 Diskussion des Zusammenwirkens von Agrarpolitik und -wissenschaft in Deutschland	37
3 ABLEITUNG VON OPTIONEN FÜR EINE EVIDENZBASIERTE AGRARPOLITISCHE ENTSCHEIDUNGSFINDUNG IN DER UKRAINE.....	38
LITERATUR.....	41

ABBILDUNGSVERZEICHNIS

Abb. 1: Modell eines Politikzyklus.....	14
Abb. 2: Ineinander greifende Aufgaben bei evidenzbasierten Evaluationen	20
Abb. 3: Einbindung der Wissenschaft in den Politikzyklus	24
Abb. 4: Systematisierung von evidenzbasierten Evaluationsverfahren.....	29

DIE WICHTIGSTEN ERGEBNISSE IM ÜBERBLICK

- (1) Das Leitbild der sogenannten „evidenzbasierten“ Politik hält zunehmend Einzug in das praktische Handeln der politisch Verantwortlichen in zahlreichen Ländern. Evidenzbasiertes Wissen kann politische Entscheidungen sowohl vor der Entscheidung über bestimmte Eingriffe und entsprechende Maßnahmen (ex-ante) als auch nach Einführung von Interventionen (ex-post) unterstützen. Damit kann evidenzbasierte Politikberatung einen substantiellen Beitrag zu politischen Lernprozessen liefern.
- (2) Den institutionellen Rahmenbedingungen fällt ein hohes Gewicht zu für ein Gelingen oder Scheitern evidenzbasierter Politikgestaltung. Anregungen für eine stärkere Evidenzbasierung kommen sowohl aus der Politik selbst als auch aus der Wissenschaft. Hier haben sich heute eine Reihe von Standards etabliert hinsichtlich Mindestanforderungen an die Qualität von Politikberatung. Viele Regierungen haben den Nutzen von evidenzbasierter Politikberatung erkannt und diese institutionell im politischen Prozess verankert.
- (3) Die Entwicklung und Anwendung staatlicher Interventionen lässt sich in einem idealisierten Politikzyklus abbilden. Den Ausgangspunkt in diesem Zyklus bildet die Definition von konkreten Zielen und die Analyse, ob der Zielwert in der Volkswirtschaft erreicht wird. Nur wenn eine nicht tolerierbare Diskrepanz zwischen dem Zielwert und dem Ausgangszustand besteht, liegt überhaupt eine Rechtfertigung für einen staatlichen Eingriff vor. Da das Wirtschaftsgeschehen komplexen Abläufen folgt, lassen sich theoretische Wirkungszusammenhänge nicht immer exakt prognostizieren. Aus diesem Grund ist eine Evaluation der Maßnahmen dringend geboten.
- (4) Die moderne Wirkungsforschung hält eine ganze Reihe an geeigneten Methoden bereit, um zuverlässige Aussagen zum Zusammenhang zwischen einer wirtschaftspolitischen Maßnahme und dem Erreichen einer Zielgröße treffen zu können. Den wissenschaftlichen Standard stellt heute das Kontrafaktum dar. Hier geht es um die Frage, was passiert wäre, wenn es die Intervention nicht gegeben hätte. Die Anwendung von modernen Methoden der Wirkungsforschung gelingt umso eher, je früher die Wissenschaft in den Politikzyklus eingebunden wird. Hier lässt sich noch Einfluss nehmen auf die Interventionsregeln, in die man das Kontrafaktum implementieren könnte, sowie auf die Generierung der administrativen Daten über die Anwendung der Maßnahme.
- (5) Evidenzbasierte Evaluationen umfassen weit mehr Aufgaben als die Wirkungsforschung im engeren Sinne. Diese sollte eingebettet sein in ein umfangreiches Management von Daten sowie mehrdimensionalen Formen kontinuierlichen Wissensaustausches. Um ein glaubwürdiges Evaluationsdesign entwickeln zu können, ist zudem umfangreiche Kenntnis über die Interventionsregeln erforderlich, was nur in engem Wissensaustausch mit der Politik gelingt. Evidenzbasierte Evaluationen sind nur dann glaubwürdig, wenn Transparenz besteht:

über die Evaluationsfragen, das Evaluationsdesign, die verwendeten Daten und Methoden sowie die Ergebnisse. Hier können internetbasierte Plattformen einen wichtigen Beitrag leisten.

- (6) In die evidenzbasierte Politikgestaltung im Agrarbereich in Deutschland sind drei Akteursgruppen involviert. Die erste Gruppe umfasst Wissenschaftler, von denen die Ministerien und Parlamente in mehr oder weniger regelmäßigen Stellungnahmen erbitten. Dies geschieht etwa im Rahmen von formellen Anhörungen sowie der Mitgliedschaft in Wissenschaftlichen Beiräten oder dem Sachverständigenrat Ländliche Entwicklung. Die zweite Gruppe besteht aus sogenannten Ressortforschungseinrichtungen, die Finanzierung durch das Bundesministerium für Ernährung und Landwirtschaft (BMEL) erhalten. Forscherinnen und Forscher an den Hochschulen sowie den außeruniversitären Forschungsinstituten der großen Wissenschaftsverbände (Max-Planck-Gesellschaft, Helmholtz-Gemeinschaft, Leibniz-Gemeinschaft, Fraunhofer-Gesellschaft) stellen die dritte Gruppe von Akteuren im System evidenzbasierter Politikberatung dar.
- (7) Für die Etablierung eines Systems evidenzbasierter agrarpolitischer Entscheidungsfindungen in der Ukraine lassen sich drei Empfehlungen ableiten. Erstens bedarf es guter institutioneller Arrangements. Die Überprüfung der Wirkung staatlicher Maßnahmen sollte gesetzlich verankert werden – idealerweise mit Verfassungsrang. Zweitens müssen die Interventionen sauber dokumentiert werden – eingebunden in ein funktionierendes System der amtlichen Statistik, welches sich an den internationalen Standards orientieren sollte. Drittens ist zu empfehlen, die Evaluationen nicht nur über die Wirkungsforschung im engeren Sinne zu verstehen, sondern als System ineinandergreifender Aufgaben zwischen Methodenkenntnis, professionellem Datenmanagement sowie umfangreichen Transferaktivitäten zwischen Wissenschaft, Politik und breiter Öffentlichkeit.

1 ZIELE UND MOTIVATION

Die vorliegende Kurzexpertise beschreibt Voraussetzungen für eine auf Fakten beruhende Politikgestaltung. Das Leitbild dieser sogenannten „evidenzbasierten“ Politik hält zunehmend Einzug in das praktische Handeln der politisch Verantwortlichen in zahlreichen Ländern. Auf der Angebotsseite für die Bereitstellung evidenzbasierten Wissens haben sich eine ganze Reihe verschiedener Institutionen etabliert, spezialisiert auf diverse Politikfelder (Buch, Patzwald, Riphahn und Vogel 2019). Dieses evidenzbasierte Wissen kann politische Entscheidungen sowohl vor der Entscheidung über bestimmte Eingriffe und entsprechende Maßnahmen (ex-ante) als auch nach Einführung von Interventionen (ex-post) unterstützen. In den letzten Jahrzehnten hat sich vor allem das methodische Instrumentarium der ex-post Bewertung von politischen Eingriffen (Evaluationen) weiterentwickelt und verbessert.

Wirtschaftsprozesse folgen komplexen Abläufen, die sich selten präzise vorhersagen lassen (vgl. hierzu etwa die Literatur um Dörner 1990; Dörner und Schaub 1994; Vester

2019, 1999; Forrester 1961). Politik kann daher nur über Lernprozesse nach einem Trial-and-Error-Verfahren funktionieren – eingebettet in einen Politikzyklus (Howlett, Ramesh und Perl 2020, S. 250). Die Fähigkeit von Gesellschaften, Fehler im politischen Prozess zu erkennen und zu korrigieren, hängt unmittelbar von der Qualität ihrer Institutionen zusammen. Gute Institutionen können Volkswirtschaften kräftige Impulse für Wachstumsprozesse verleihen (etwa North 1990; Acemoglu und Robinson 2010).

Evidenzbasierte Evaluationen können einen substantiellen Beitrag zu politischen Lernprozessen liefern. Dabei ist es nicht ausreichend, nur zu prüfen, ob die Politik die Maßnahmen nach den vorgegebenen Regeln durchführte. Bei dieser Art von Analyse handelt es sich um Vollzugskontrollen. Aussagen zur Effektivität und Effizienz staatlicher Interventionen können solche Vollzugskontrollen jedoch nicht liefern. Vielmehr ist es wichtig, einen ursächlichen Zusammenhang zwischen der wirtschaftspolitischen Maßnahme und dem Erreichen von Zielgrößen nachzuweisen. Die internationale Wirkungsforschung verfügt heute über ein ausgefeiltes Methodenspektrum, um eben jenen kausalen Zusammenhang aufzuzeigen. Die Grundidee dieser Verfahren ist die Schaffung einer „kontrafaktischen Situation“. Es wird dabei die Frage gestellt: Was wäre geschehen, wenn es die staatliche Intervention nicht gegeben hätte? Der Unterschied zwischen tatsächlicher und kontrafaktischer Situation ist der Effekt der staatlichen Intervention (bspw. Beirat BMWi 2013, 2015; Kugler, Schwerdt und Wößmann 2014; Boockmann, Buch und Schnitzer 2014; Schmidt 2014; Expertenkommission 2015; EFI 2017, Abschnitt B5-3; Buch und Riphahn 2019; Brachert, Dettmann und Titze 2015).

Für viele Politikbereiche liegt bislang wenig Evidenz darüber vor, welche tatsächlichen Folgen Interventionen für die Gesellschaft haben. Eine ganze Reihe von Maßnahmen etwa sind mit der Umverteilung finanzieller Mittel verbunden, und es ist nicht abschließend bekannt, ob die staatlichen Ausgaben dem Gebot der Sparsamkeit und Wirtschaftlichkeit tatsächlich genügen (Beirat BMWi 2013). Anstöße für eine Änderung dieser Praxis kommen sowohl aus der Politik/Administration (EU KOM 2014a; NKR 2015) als auch aus der Wissenschaft (Beirat BMWi 2013, 2015; VfS 2015), was – streng genommen – schon den Ausgangspunkt für institutionelle Reformen darstellt.

Das trifft auch für Maßnahmen im Rahmen von Regional- und Agrarpolitik in Deutschland zu, welche eine besondere Form von Umverteilung repräsentieren. Diese sollen als Beispiel in dieser Kurzexpertise dafür dienen, wie sich Evidenz in systematischen Evaluationen nach internationalen Qualitätsstandards bereitstellen und in die Praxis transferieren lässt. Für die Effekte der Agrar- und Regionalpolitik in Deutschland liegen Evidenzen vor. Ob aber alle Maßnahmen im Rahmen der Agrar- und Regionalpolitik tatsächlich die an sie gestellten Ziele erfüllen, ist bislang längst nicht abschließend erforscht. Es handelt sich damit um einen geeigneten Fall aus der Praxis, anhand dessen sich demonstrieren lässt, wie man die Evaluationsprozesse organisieren könnte und welche institutionellen und technischen Rahmenbedingungen vorliegen müssten, um Grundlagen für kontinuierliche Lernprozesse in der Politik zu schaffen (vgl. hierzu etwa Brachert, Dettmann und Titze 2015, 2020; Brachert, Brauttsch, Dettmann, Giebler,

Schneider und Titze 2020).

Die Kurzexpertise gliedert sich wie folgt. Das *zweite* Kapitel befasst sich mit deutschen Erfahrungen zu evidenzbasierten Entscheidungen. Im Mittelpunkt stehen zunächst die institutionellen Rahmenbedingungen und wie diese evidenzbasierte Analysen fördern oder behindern können. Ein zweiter Schwerpunkt liegt auf der Organisation der Generierung von evidenzbasierten Analysen. Gegenstand der Ausführungen ist die Beschreibung eines Systems, wie sich evidenzbasierte Analysen in der praktischen Arbeit umsetzen lassen. Es zeigt sich, dass neben der Evaluationsforschung im engeren Sinne Aspekte des Wissenstransfers und einer umfassenden Datenarbeit notwendige Voraussetzungen darstellen, um zuverlässige evidenzbasierte Analysen durchführen zu können. Das *dritte* Kapitel beschreibt Optionen, wie sich Erfahrungen evidenzbasierter Evaluierungen auf Entscheidungsfindungen im Bereich von Agrarpolitik in der Ukraine ableiten lassen.

2 EVIDENZBASIERTE ENTSCHEIDUNGSFINDUNG - DEUTSCHE ERFahrungen

Dieses Kapitel befasst sich mit deutschen Erfahrungen bei der Generierung von Evidenz für politische Entscheidungen. Abschnitt 0 schildert zunächst die institutionellen Rahmenbedingungen für evidenzbasierte ökonomische Evaluationsforschung in Deutschland. Abschnitt 0 argumentiert, dass Evaluationen Bestandteil eines umfangreichen Politikzyklus sind und die Wissenschaft bereits in frühen Phasen eingebunden sein sollte. Abschnitt 0 diskutiert, dass Evaluationsforschung heute breiter gedacht werden muss, eingebettet in ein System verschiedener Aufgaben. Evidenzbasierte Evaluationsforschung ist nicht mehr vorstellbar ohne umfangreiches Datenmanagement sowie intensiven Wissensaustausch zwischen Politik, Wissenschaft und breiter Öffentlichkeit. Methoden evidenzbasierter Forschung sowie ihre Stärken und Schwächen sind Gegenstand von Abschnitt 0. Die evidenzbasierte Politikgestaltung im Agrarbereich ist Thema des Abschnitts 0. Schließlich zeigt Abschnitt 0, wie ein System evidenzbasierter Evaluationsforschung in der Praxis etabliert werden kann.

2.1 INSTITUTIONELLE VORAUSSETZUNGEN

Wie eingangs geschildert, stellt die Qualität von Institutionen eine Determinante langfristigen Wachstums dar. Allerdings erweisen sich Institutionen als recht starre Gebilde – sie sind das Ergebnis komplexer und langwieriger kollektiver Entscheidungsprozesse, die wiederum vom politischen System und der Verteilung von Macht abhängen. Dadurch kann es passieren, dass sich notwendige Reformen verzögern oder auch ganz ausbleiben, obwohl sie objektiv geboten scheinen (etwa North 1990; Acemoglu und Robinson 2010). Dieser Abschnitt befasst sich mit institutionellen Gegebenheiten und wie diese evidenzbasierte Evaluationen befördern oder behindern können. Letztendlich geht es um die Frage, ob die Gesellschaft als Ganzes bereit ist, Fehler im politischen Handeln zuzulassen, diese als Teil von Lernprozessen zu akzeptieren und damit eine Korrektur von Fehlern überhaupt erst zu ermöglichen.

2.1.1 Rechtsstaatsprinzip

Ihre ökonomische Rechtfertigung erfahren staatliche Interventionen aus Fehlallokationen von Ressourcen (d. h. das Versorgungsniveau mit bestimmten Gütern liegt ober- oder unterhalb des gesamtwirtschaftlichen Optimums) infolge von Marktmängeln (vgl. Fritsch 2018 für einen umfassenden Überblick) sowie distributiven Erwägungen (d. h. zu starke Einkommensdisparitäten können soziale Spannungen auslösen) (vgl. Kilper und Rosenfeld 2007). In demokratischen Gesellschaften liegt die Entscheidung für oder gegen staatliche Maßnahmen in den Händen staatlicher Gewalt (Legislative, Exekutive, Judikative), kontrolliert durch die Wähler aber auch beeinflusst durch Interessengruppen.

Nahezu jede staatliche Maßnahme außerhalb des ordnungspolitischen Instrumentariums ist – wenn nicht direkt, dann zumindest indirekt – mit öffentlichen Ausgaben verbunden, die sich durch öffentliche Einnahmen finanzieren müssen, was erst einmal auch einen Entzug von Finanzkraft für den privaten Sektor bedeutet. Für jede Gesellschaft ist es von erheblicher Bedeutung, ob die unter „Abgabenzwang“ generierten öffentlichen Einnahmen auf eine breite Akzeptanz in der Bevölkerung stoßen. Gelingt dies nicht, befördert dies Schattenwirtschaft mit all ihren negativen Konsequenzen (vgl. hierzu etwa die Diskussion in Schneider und Enste 2000). Insofern ist es nur folgerichtig, systematisch zu überprüfen, ob mit den staatlichen Maßnahmen tatsächlich die Marktmängel korrigiert wurden. Dies gilt nicht nur für die Politik, die durch systematische Wirkungsanalysen Lerneffekte realisieren kann, sondern auch für die breite Öffentlichkeit (und hier insbesondere die Steuerzahler), bei der dadurch die Akzeptanz für die Erhebung der Zwangsabgaben zur Finanzierung der öffentlichen Haushalte steigen sollte.

Die Kontrolle staatlicher Ausgaben in der Bundesrepublik Deutschland fußt auf einem umfangreichen gesetzlichen Regelwerk, beginnend mit dem Grundgesetz (Artikel 114 [2]), über das Haushaltsgrundsätzegesetz (§7 [2]), die Bundeshaushaltsordnung bis zu den verschiedenen Gesetzen und Verordnungen auf Ebene der Länder und Gemeinden. Ein wichtiger Punkt ist, dass die mit der Überprüfung der öffentlichen Finanzen betrauten Rechnungshöfe nach Artikel 114 [2] Grundgesetz institutionell unabhängig von den übrigen Gewalten sind. Sie verfügen über richterliche Unabhängigkeit. Dieses Regelwerk fordert, öffentliche Ausgaben zu kontrollieren. Offen bleibt allerdings, wie dies konkret für die einzelnen Maßnahmen geschehen soll und welche Qualitätskriterien für die Prüfung der öffentlichen Finanzen anzusetzen sind.

2.1.2 Standards in der Politik

Über das im vorangegangenen Abschnitt beschriebene Regelwerk hinaus, haben sich aus der praktischen Arbeit der Politik Initiativen herausgebildet, die gewisse **Mindeststandards** für die Notwendigkeit und die Durchführung von Evaluationen festlegen. Dabei ist zu berücksichtigen, dass die Kompetenzen für die Politikfelder auf unterschiedlichen administrativen Ebenen (Europäische Union [EU], Bund, Länder, Gemeinden, Sozialversicherungsträger) festgelegt sind. Die folgenden Ausführungen stellen *drei* dieser Initiativen – eine auf der Ebene des Bundes, die anderen beiden auf der Ebene der EU – etwas näher vor.

Die *erste* Initiative betrifft den **Nationalen Normenkontrollrat (NKR)** (<https://www.normenkontrollrat.bund.de/nkr-de>), dessen Einsetzung von der Bundesregierung im Jahr 2006 per Gesetz (NKR-G) beschlossen wurde. Dieses Gremium agiert unabhängig, und ihm obliegt die Aufgabe, die Bundesregierung in Fragen von „Bürokratieabbau und besserer Rechtsetzung“ zu beraten. Zu seinen Aufgaben zählt die Prüfung:

- des Erfüllungsaufwands¹ neuer Regelungen;
- der Darstellung des Ziels und der Notwendigkeit der Regelung;
- von Erwägungen zu anderen Lösungsmöglichkeiten;
- von Erwägungen zum Zeitpunkt des Inkrafttretens, zur Befristung und Evaluierung;
- der Ausführungen zur Rechts- und Verwaltungsvereinfachung sowie
- der Umsetzung von EU-Recht in nationales Recht.

Gegenstand der Prüfung ist das gesamte Spektrum der Rechtsetzung auf Bundesebene.² Auf Empfehlung des NKR hat sich die Bundesregierung im Jahr 2013 auf eine „Konzeption zur Evaluierung neuer Regelungsvorhaben“³ verständigt, wonach Gesetze nach drei bis fünf Jahren zu evaluieren sind, wenn der Erfüllungsaufwand eine Million Euro übersteigt.

Eine andere Initiative, welche Evaluation konsequent einfordert und dafür gewisse Mindeststandards verlangt, ist auf der Ebene der EU angesiedelt. Konkret handelt es sich um sogenannte **Notifizierungsprozesse**. Die Verträge der EU verbieten grundsätzlich jede Art von staatlichen Beihilfen, da sie den Wettbewerb im EU-Binnenmarkt behindern. Daher müssen die Mitgliedsstaaten jegliche Art von nationalen Beihilfen, die den Wettbewerb beeinträchtigen könnten, bei der EU-Kommission anmelden. Beim Anmeldeverfahren handelt es sich um einen strukturierten Prozess, in dem die Mitgliedsstaaten – ähnlich wie auch vom NKR in Deutschland eingefordert – Adressaten, Umfang, Art, Ziele usw. der Maßnahmen darlegen müssen. Besonderes Augenmerk legt die EU-Kommission auf die Evaluation der zu genehmigenden Maßnahme. Hierfür sind die Mitgliedsstaaten verpflichtet, sogenannte **Evaluationspläne** zu erarbeiten (EU KOM 2014a), die wissenschaftlichen Mindeststandards entsprechen müssen. Dafür stellt die EU Kommission eine Reihe von Arbeitsdokumenten bereit (vgl. etwa EU KOM 2014b).

In die Prüfung der Evaluationspläne und der vorgelegten Evaluationsstudien ist zudem das **Joint Research Centre (JRC)** (<https://ec.europa.eu/jrc/en>) der EU Kommission

¹ Hierzu zählen die Kosten, die durch Befolgung der Vorschriften entstehen inklusive der Kosten durch Informationspflichten.

² Dazu zählen alle neuen sowie bestehenden Bundesgesetze, Änderungsgesetze, Rechts- und Verwaltungsvorschriften, Rechtsakte der EU und die durch Umsetzung von EU-Recht betroffenen Gesetze.

³ Konzeption zur Evaluierung neuer Regelungsvorhaben gemäß Arbeitsprogramm bessere Rechtsetzung der Bundesregierung vom 28. März 2012, Ziffer II. 3. In: Bundeskanzleramt, Geschäftsstelle Bürokratieabbau: Bessere Rechtsetzung 2012: Belastungen vermeiden. Bürokratischen Aufwand verringern. Wirtschaftliche Dynamik sichern. Bericht der Bundesregierung 2012 nach § 7 des Gesetzes zur Einsetzung eines Nationalen Normenkontrollrates. Mai 2013. Berlin.

eingebunden. Hierbei handelt es sich um eine Service-Einrichtung, die in der EU Kommission angesiedelt ist, welche für die Politik der EU umfassende wissenschaftliche Expertise – insbesondere auch auf dem Gebiet evidenzbasierter Evaluationen – bereitstellt.

Die EU-Kommission strebt, nicht zuletzt als Folge des Drucks vom Rechnungshof und als Rechtfertigung der Gemeinsamen Agrarpolitik (GAP), in den vergangenen zwei Jahrzehnten zunehmend nach einer **systematischeren Überprüfung der Effektivität der Agrarpolitik**. Mit dem Artikel 110 in der Verordnung 1306/2013 (Finanzierung, Verwaltung und Kontrollsystem der GAP, Horizontal Regulation) verpflichten sich die Kommission und die Mitgliedstaaten zur Etablierung mehrjähriger Evaluationspläne. Mit dem **Gemeinsamen Überwachungs- und Bewertungsrahmen** (Common Monitoring and Evaluation Framework, CMEF) wurde ein Indikatorenset für die Politikziele etabliert. Damit sollen alle Direktzahlungen, Marktstützungsmaßnahmen und Maßnahmen der ländlichen Entwicklungspolitik regelmäßig bewertet werden (vgl. EU KOM 2018).

Diese Beispiele zeigen, dass die Politik den Nutzen von evidenzbasierter Politikberatung erkannt hat und diese Evidenz Berücksichtigung in der politischen Praxis findet. Neben der EU-Ebene und der Bundesebene scheint es nur eine Frage der Zeit zu sein, bis diese Art von Politikberatung institutionell auch auf der Ebene der Länder und der Gemeinden in der Bundesrepublik Deutschland verankert wird.

2.1.3 Standards in der Wissenschaft

Die Standards für Evaluationen, die heute schon in bestimmten Feldern der Politik eingefordert werden, orientieren sich am State-of-the-Art der internationalen Wirkungsforschung. In einer Evaluation von wirtschaftspolitischen Interventionen geht es darum, zu untersuchen, ob ein ursächlicher Zusammenhang zwischen der **Maßnahme** und der Entwicklung von politischen Zielgrößen besteht. Hierbei offenbart sich das Problem, dass nicht beobachtbar ist, wie sich ein Untersuchungsobjekt (Person, Betrieb, Region usw.), oder auch **Teilnehmer** genannt, entwickelt hätte, wenn es die Intervention nicht gegeben hätte. Zu einem bestimmten Zeitpunkt kann ein Teilnehmer entweder eine Maßnahme erhalten oder eben nicht. Beides gleichzeitig ist nicht möglich. Zur Identifikation eines ursächlichen Effekts der Maßnahme ist es daher notwendig, ein **Kontrafaktum** zu bilden. Dieses Kontrafaktum soll dabei helfen, eben jene Frage zu beantworten - Wie wäre die Entwicklung verlaufen, wenn es die Maßnahme nicht gegeben hätte? Der Unterschied zwischen der tatsächlichen, messbaren Entwicklung der Ergebnisgrößen der Teilnehmer in der Maßnahme, d. h. der **Behandlungsgruppe**, und jener der Vergleichsgruppe, oder auch **Kontrollgruppe** genannt, die nicht dieser Maßnahme ausgesetzt war, ermöglicht dann Rückschlüsse auf den **kausalen Effekt** der Intervention. Orientiert sich eine Untersuchung am Leitbild des Kontrafaktums, dann wird diese Analyse in der Literatur gemeinhin als **evidenzbasierte Evaluation** bezeichnet.

Das Kontrafaktum lässt sich über verschiedene ökonometrisch-statistische Methoden abbilden. Eine detaillierte Diskussion dieser Methoden erfolgt in Abschnitt 0. Naturgemäß besitzen alle Methoden Stärken und Schwächen. Für eine **Beurteilung der Eignung**

bestimmter Methoden auf konkrete praktische Fragestellungen hat sich etwa in der internationalen ökonomischen Evaluationsforschung ein Bewertungssystem auf Basis einer Punkteskala etabliert. Dieses **Punktesystem** der sogenannten „Maryland Scientific Methods Scale“ (Sherman, Gottfredson, MacKenzie, Eck, Reuter und Bushway 1998) reicht von 1 (einfache Korrelationen, keine Aussagen zu Kausalität möglich) bis 5 (randomisierte Kontrollstudien, sehr zuverlässige Aussagen zu Kausalität möglich). Madaleno und Waights (2016) haben dieses Punktesystem weiterentwickelt. Nach diesem Bewertungsbogen muss mindestens die Stufe 3 erreicht werden, um Kausalität glaubhaft bewerten zu können.

2.1.4 Standards in der Datengenerierung und Bereitstellung

Evidenzbasierte Evaluationen benötigen gute Daten.⁴ Die Daten der Amtlichen Statistik in Deutschland stellen aufgrund ihrer hohen Qualität eine sehr gute Quelle für die Forschung dar. Um ihr – unter Wahrung des Datenschutzes – Zugriff auf diese Daten zu ermöglichen, wurden deutschlandweit **Forschungsdatenzentren (FDZ)** gegründet, die häufig an die Amtlichen Statistik angegliedert sind. Prominente Beispiele sind etwa die FDZ der Statistischen Ämter des Bundes und der Länder sowie der Bundesagentur für Arbeit (in ihrer Eigenschaft als Verwaltungsträgerin für Aufgaben in der deutschen Sozialversicherung). Die FDZ organisieren den Zugriff auf die Daten der Amtlichen Statistik für Forschungszwecke über standardisierte Antragsverfahren.

Die Einhaltung der Qualitätsstandards an den FDZ wiederum überwacht der **Rat für Sozial- und Wirtschaftsdaten (RatSWD)** (<https://www.ratswd.de>). Er repräsentiert einen unabhängigen Beirat der Bundesregierung. Zu seinen Aufgaben gehört neben der Akkreditierung und Evaluation der Arbeit der FDZ auch die strategische Weiterentwicklung der Forschungsdateninfrastruktur sowie ihre Vernetzung. Darüber hinaus sieht sich der RatSWD als Interessenvertretung von Datenproduzenten und –nutzenden und übt Beraterfunktionen für Politik und Wissenschaftsinstitutionen aus.

Im Agrarbereich ist das **Informationsnetz landwirtschaftlicher Buchführungen** der Europäischen Union (INLB, englisch: Farm Accountancy Data Network, FADN) von besonderer Bedeutung. Mit der Verordnung Nr. 79/65/EWG wurde kurz nach Begründung einer gemeinsamen europäischen Agrarpolitik die Grundlage für eine harmonisierte Erfassung von produktionstechnischen und wirtschaftlichen Betriebsdaten gelegt. Die Umsetzung in der Bundesrepublik Deutschland erfolgte im Rahmen des

⁴ Je disaggregierter diese Daten vorliegen, um so präzisere Wirkungsanalysen lassen sich durchführen. Dies kollidiert jedoch in vielen Fällen mit dem Recht auf informationelle Selbstbestimmung, welches sich ableitet aus dem Allgemeinen Persönlichkeitsrecht im Grundgesetz der Bundesrepublik Deutschland (Artikel 2 [1] in Verbindung mit Artikel 1 [1]). Die systematische Erhebung von Informationen auf individueller Ebene (Personen, Betriebe) geschieht in Deutschland im Rahmen der Amtlichen Statistik, etwa auf Basis des Bundesstatistikgesetzes (BStatG) oder bei den Meldungen zur Sozialversicherung auf Basis der Sozialgesetzbücher (SGB). Hier sind die Befragten verpflichtet, teilzunehmen und wahrheitsgemäß zu antworten. Ein Verstoß zieht empfindliche Strafen nach sich. Gleichzeitig garantiert der Gesetzgeber den Schutz persönlicher Daten, der in Deutschland einen sehr hohen Stellenwert innehat.

Landwirtschaftsgesetzes von 1955 (§§2-7). Die Stichprobe von 6.000 bis 8.000 Betrieben soll oberhalb einer Grenze von 25.000 EUR Standardoutputs (kalkulatorische Erfassung des Werts der Produktion ohne Direktzahlungen) die Agrarstruktur in Deutschland repräsentativ abbilden. Teilnehmende Betriebe an den jährlichen Erhebungen erhalten eine Aufwandsentschädigung (derzeit 160 EUR pro Betriebsbogen). Die Erfassung der Aufwands- und Ertragspositionen folgt den Vorgaben des Handelsgesetzbuches. Ein Zugang zu den Daten für wissenschaftliche Untersuchungen ist möglich, aber wird relativ restriktiv gehandhabt (vgl. BMEL 2018).

Seitens der wissenschaftlichen Politikberatung wird untersucht, in wieweit die FADN-Datenerhebung für die Erfassung weiterer Indikatoren auf Betriebsebene, vor allem umweltbezogene, erweitert werden kann, um den sich ändernden politischen Zielen Rechnung zu tragen (Poppe und Vrolijk 2016).

2.1.5 Zusammenfassung zu den institutionellen Rahmenbedingungen

Die Ausführungen in Abschnitt 0 haben verdeutlicht, dass den institutionellen Rahmenbedingungen ein hohes Gewicht zufällt für ein Gelingen oder Scheitern evidenzbasierter Politikgestaltung. Anregungen für eine stärkere Evidenzbasierung kommen sowohl aus der Politik selbst als auch aus der Wissenschaft. Hier haben sich heute eine Reihe von Standards etabliert hinsichtlich Mindestanforderungen an die Qualität von Politikberatung. Viele Regierungen haben den **Nutzen von evidenzbasierter Politikberatung** erkannt und diese institutionell im politischen Prozess verankert. Diese institutionellen Arrangements ermöglichen es, dass Politik lernt und sich dadurch verbessert.

Im Umkehrschluss bedeutet dies aber auch, dass eine schlechte Qualität oder ein gänzlich fehlendes dieser Standards die Generierung evidenzbasierten Wissens behindern, ja sogar unmöglich machen. Der berechtigte Wunsch nach Bereitstellung von Evidenz ist nicht frei von **Zielkonflikten** (und kann es womöglich auch nie sein). Dies zeigt sich etwa bei der Bereitstellung guter Daten (im Sinne von Repräsentativität, Aktualität, Konsistenz, frei von Messfehlern usw.), ohne die evidenzbasierte Analysen überhaupt nicht möglich wären.

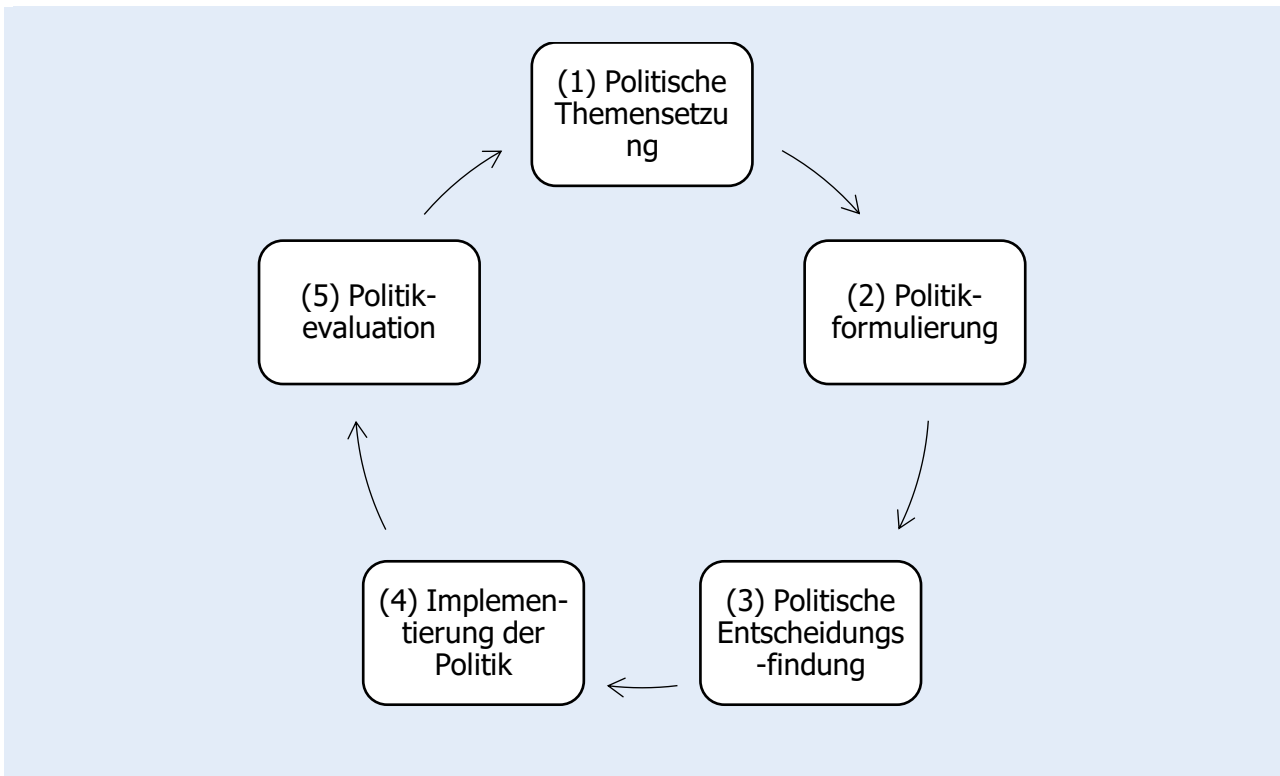
Hier ist abzuwägen zwischen dem Recht auf informationelle Selbstbestimmung und den Erfordernissen, welche sich aus den praktischen Forschungsaktivitäten ergeben. Im Zweifel lassen sich mit den zur Verfügung stehenden Daten relevante Forschungsfragen nicht beantworten, wenn die Politik Datenschutzaspekten Vorrang einräumt.

2.2 EINBETTUNG VON EVALUATIONEN IN EINEN POLITIKZYKLUS

Evaluationen kann man nicht isoliert betrachten, vielmehr sind sie Bestandteil eines ‚**Politikzyklus**‘ (Abb. 1). Der hier vorgestellte Politikzyklus basiert auf dem Fünf-Phasen-Modell von Howlett, Ramesh und Perl (2020). Darüber hinaus finden Ideen aus Buch, Patzwald, Riphahn und Vogel (2019, S. 3-4) sowie NKR (2017, S. 31) Berücksichtigung. In dem hier präsentierten Modell handelt es sich um einen idealisierten Politikzyklus. In

der Praxis dürften die Übergänge zwischen den einzelnen Phasen fließend sein.

Abb. 1: Modell eines Politikzyklus



Quelle: Eigene Darstellung in Anlehnung an Howlett, Ramesh und Perl (2020).

Die Beschreibung der einzelnen Phasen greift auf das eingangs angesprochene Beispiel der *Regionalpolitik in Deutschland* zurück. Über dieses Instrument vergibt die öffentliche Hand Subventionen an Betriebe und Gemeinden in strukturschwachen Gebieten in Deutschland. Mit Hilfe dieses konkreten Falls aus der Praxis lässt sich demonstrieren, welche Arbeitsschritte die Politik in jeder Phase vollzieht. Hierbei ist jedoch zu anmerken, dass die deutschen Behörden die Regionalpolitik nicht nach diesem Schema entwickelt haben. Es handelt sich um eine idealisierte Darstellung der Verfasser auf Basis der veröffentlichten Dokumente zur Regionalpolitik.

2.2.1 Politische Themensetzung

Die *erste Phase* beinhaltet die Erarbeitung einer politischen Agenda. Hier geht es zunächst darum, ein Zielsystem für die Gesellschaft als Ganzes zu entwickeln. Dafür müssen die meist allgemein vorgegebenen **Oberziele** (bspw. Wohlstand) **in konkrete Teilziele** (bspw. Gerechtigkeit, Wirtschaftswachstum, Zufriedenheit, Lebenserwartung, Gesundheit usw.) herunter gebrochen werden.

An diesem Punkt ist es sehr wichtig, dass die Politik die meist sehr abstrakt formulierten Ziele in objektiv **messbare Indikatoren** übersetzt – idealerweise solchen, über die bereits in amtlichen Statistiken Informationen vorliegen.

Für jeden Indikator definiert die Politik einen Wert, bei dem das Ziel als erfüllt betrachtet werden kann. Eine weitere wichtige Frage ergibt sich für die Politik dahingehend, welche **Diskrepanzen** sie **zwischen angestrebten Zielwert** (Soll-Zustand) und der

aktuellen Situation (Ist-Zustand) als noch **tolerabel** ansieht. Mit anderen Worten, die Politik muss einen **Schwellenwert** festlegen, ab wann sie die **Notwendigkeit für Interventionen** sieht.

Bezogen auf das Beispiel der *deutschen Regionalpolitik* lässt sich diese Phase wie folgt beschreiben. Das Oberziel der Politik ist die Maximierung der gesamtwirtschaftlichen Wohlfahrt. Das Erreichen dieses Ziels bestimmt sich in erster Linie dadurch, inwiefern die Volkswirtschaft in der Lage ist, die Bedürfnisse der Bevölkerung zu befriedigen, d. h. Güter (Waren und Dienstleistungen) zu produzieren. Der Entscheidung über den Einsatz für die zur Produktion benötigten Ressourcen fällt auf individueller Ebene. Der Marktmechanismus sorgt für eine Koordination der Pläne der zahlreichen Marktakteure. In einigen Fällen funktioniert dieser Mechanismus nicht perfekt, sodass die Volkswirtschaft entweder zu viel oder zu wenig von bestimmten Gütern produziert. In solchen Fällen liegen Marktmängel vor (Fritsch 2018). Diese finden sich auch im räumlichen Kontext⁵, woraus sich wirtschaftspolitische Ziele für die Korrektur von Marktmängeln begründen, um eine effiziente Allokation der Ressourcen zu ermöglichen (bspw. Neumark und Simpson 2015).

Neben den allokativen Zielen verpflichtet sich die deutsche Politik auch dem normativen Ziel der Verteilungsgerechtigkeit, woraus sich die Rechtfertigung für die Begründung des Sozialstaats ableitet (Grundgesetz, Artikel 20 Absatz 1 Satz 1). Das Prinzip der Verteilungsgerechtigkeit ist im Grundgesetz sehr weit gefasst und versteht sich im Sinne einer regionalen Daseinsvorsorge, was sich im Ziel der „Herstellung gleichwertiger Lebensverhältnisse im Bundesgebiet“ (Artikel 72 Absatz 2 Satz 3 Grundgesetz) manifestiert (Kilper und Rosenfeld 2007).

Zusammengefasst verfolgt die deutsche Regionalpolitik das Ziel, Ungleichgewichte zwischen den Regionen innerhalb Deutschlands abzubauen. Hierfür orientiert sie sich an einem Set ökonomischer Regionalindikatoren, insbesondere Bruttoinlandsprodukt, Bruttowertschöpfung; Arbeitslosenquote; Löhne, Gehälter; Beschäftigung, Beschäftigtenentwicklung; Ausstattung mit wirtschaftsnahen Infrastrukturen (Alm und Fisch 2014). Einen konkreten Wert, wann das Ziel der Verringerung regionaler Disparitäten erreicht ist, nennt die deutsche Regionalpolitik nicht.

2.2.2 Politikformulierung

Die *zweite Phase* befasst sich mit der konkreten Formulierung der Politik. Hier geht es zunächst um eine umfassende Analyse des Problems. Dafür ist eine grundlegende Kenntnis möglicher **Wirkungszusammenhänge** erforderlich. Eine detaillierte Analyse sollte mögliche Konflikte aufzeigen (vgl. hierzu etwa die Literatur um Dörner 1990; Dörner

⁵ Die deutsche Politik zählt in diesem Zusammenhang den Ausgleich natürlicher Standortnachteile (etwa geografische Randlagen), nicht ausgeschöpfte Agglomerationsvorteile, immobile Produktionsfaktoren (insbesondere Arbeit) sowie Friktionen an den Faktormärkten (etwa regionale Verfügbarkeit von Risikokapital) als Argumente für die Notwendigkeit regionalpolitischer Interventionen auf (Alm und Fisch 2014).

und Schaub 1994; Vester 2019, 1999; Forrester 1961). So kann es passieren, dass mit dem Erreichen eines Ziels ein anderes negativ betroffen ist. Die Ein- oder Fortführung einer Politikmaßnahme führt dann zu unerwünschten Nebenwirkungen. Ein typisches Beispiel ist etwa der **Zielkonflikt** zwischen Wirtschaftswachstum und Verteilungsgerechtigkeit. Der Marktmechanismus kann eine Verteilung der Einkommen herbeiführen, die die Gesellschaft als ungerecht empfindet. Eine Korrektur dieses Ergebnisses gelingt nur über Umverteilung: Produktiven Bereichen der Volkswirtschaft werden Ressourcen entzogen und an ökonomisch Benachteiligte transferiert. Letztendlich geht es in dieser Phase darum, für mögliche Politikoptionen **Chancen und Risiken abzuwägen**. So können Interventionen, die das Ziel nicht erreichen oder unerwünschte Nebenwirkungen aufweisen, vorab ausgeschlossen werden. Erweisen sich mehrere Optionen als grundsätzlich geeignet für die Politik, können **Kosten-Nutzen-Abwägungen** die Auswahl weiter einengen.

Die *deutsche Regionalpolitik* stellt ein diskretionäres Förderinstrument dar. Antragsberechtigt sind nur Betriebe und Gemeinden in strukturschwachen Regionen. Akteure außerhalb dieser Gebiete haben keinen Zugang zu dieser Art der Förderung. Über die Förderung sollen die Akteure in strukturschwachen Gebieten Hilfe zur Selbsthilfe erhalten. Durch die Sicherung bestehender und Schaffung neuer Beschäftigungsverhältnisse sollen zusätzliche Einkommensquellen in den betroffenen Regionen generiert werden (Alm und Fisch 2014).

Bezüglich des Fördergegenstands hat sich die Politik auf Investitionsvorhaben festgelegt. Nach der mikroökonomischen Theorie reduziert eine Förderung von Investitionen die Kapitalnutzungskosten, weshalb – unter der Annahme gewinnmaximierender Unternehmen und unter sonst gleichen Bedingungen – die Unternehmen ihren Kapitalbestand ausdehnen. Wenn die Nachfrage hinreichend elastisch reagiert, schaffen die Unternehmen zusätzliche Arbeitsplätze (Crisuolo, Martin, Overman, van Reenen 2019).

Für die geförderten Investitionsobjekte in den Betrieben und Gemeinden gilt eine von der Politik vorgegebene Mindestnutzungsdauer. Für diesen Zeitraum müssen auch die Arbeitsplätze in den geförderten Betrieben bestehen bleiben. Als zentrale Indikatoren für die Messung der Zielerreichung betrachtet die Politik die Anzahl an Arbeitsplätzen sowie die Einkommen (Alm und Fisch 2014).

Neben dem immanenten Zielkonflikt bei Umverteilungsmaßnahmen besteht ein weiterer Zielkonflikt der Regionalpolitik dahingehend, dass diese Art von Förderung den Strukturwandel behindern könnte, da nicht wettbewerbsfähige Wirtschaftsstrukturen künstlich aufrechterhalten werden. Die Politik versucht diesen Zielkonflikt etwa dadurch zu adressieren, dass sie Regionalförderung für Unternehmen in Schwierigkeiten nach den Förderregularien ausschließt (vgl. Bundesanzeiger 2018 für die aktuell gültige Fassung).

2.2.3 Politische Entscheidungsfindung

In der *dritten Phase* geht es darum, eine Entscheidung für eine konkrete Politikmaßnahme

zu treffen. Dies ist häufig mit **formellen Prozessen** in demokratischen Gesellschaften verbunden, etwa dem Entwurf, der Diskussion und der Verabschiedung von Gesetzen in den Parlamenten. Die Gesetze, Verordnungen und Richtlinien werden auf den einschlägigen Plattformen bereitgestellt, womit die Regelungen für jedermann öffentlich zugänglich sind. Diese formalen Schritte schaffen **Rechtssicherheit** in Bezug auf die Anwendung der Politikmaßnahme.

Für die *deutsche Regionalpolitik* beinhaltet diese Phase die Schaffung des rechtlichen Rahmens für die Anwendung des Förderprogramms und zwar durch das Gesetz über die Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRWG). Hier sind folgende Aspekte geregelt (vgl. im Folgenden Alm und Fisch 2014):

- Wer soll über die Förderregeln entscheiden?
- Welche Gebiete sollen die Förderung erhalten?
- Welche Unternehmen sollen Förderung erhalten?
- Welche Projekte sollen Unterstützung erhalten?
- Wie soll die Förderung ausgereicht werden (Zulage, Zuschuss, Darlehen)?
- In welcher Höhe sollen Betriebe und Gemeinden Förderung erhalten?

Aufgrund verfassungsrechtlicher Besonderheiten sind in der deutschen Regionalpolitik neben dem Bund auch die Länder involviert. Die Abstimmung zwischen Bund und Ländern erfolgt im Rahmen eines sogenannten Koordinierungsausschusses. Nach außen muss sich die Bundesregierung mit der EU abstimmen im Rahmen des Notifizierungsprozesses für die deutsche Regionalpolitik, da es sich um ein Instrument handelt, das den Wettbewerb innerhalb des Binnenmarktes verfälschen könnte. Im Falle der Unbedenklichkeit der staatlichen Beihilfen, erlaubt die EU die Anwendung dieser Politik für einen Zeitraum von mehreren Jahren. Danach müssen die Mitgliedsstaaten die Verlängerung der Maßnahme bei der EU beantragen. Der Koordinierungsprozess mit der EU beinhaltet auch die Abstimmung eines Evaluationsplans.

Die Festlegung der Fördergebiete im Rahmen der deutschen Regionalpolitik erfolgt anhand eines Strukturschwächeindikators. Der Umfang des Fördergebiets ist gedeckelt. Die Förderung erfolgt über Zuschüsse für Investitionen von Betrieben und Investitionen in wirtschaftsnahe Infrastrukturen von Gemeinden. Die Förderung ist grundsätzlich technologieoffen, richtet sich aber an Betriebe, die ihre Produkte überregional absetzen.⁶ Eine Förderung ist nur auf Antrag möglich, wobei die Länder für die endgültige Auswahl der zu fördernden Projekte verantwortlich sind.

2.2.4 Implementierung der Politik

Die *vierte* Phase beinhaltet die Umsetzung der Politik und die konkrete **Anwendung geltenden Rechts** in der Praxis sowie ihre **Dokumentation**. Involviert sind die

⁶ Ferner sind einige Branchen von der Förderung aufgrund besonderer Regelungen in der EU ausgeschlossen.

Behörden der verschiedenen administrativen Ebenen.

Die Implementierung der *deutschen Regionalpolitik* erfolgt über die Landesregierungen. Sie organisieren die Antragsbearbeitung, -prüfung und -bewilligung sowie die Auszahlung der Fördergelder. Nach Abschluss der Förderung findet eine Prüfung statt, ob die Mittel entsprechend der Festlegungen in der Bewilligung verwendet wurden. Häufig beauftragen die Landesministerien öffentliche Förderbanken für die Durchführung des operativen Geschäfts des Förderprozesses. Zudem müssen die Länder flexibel auf Änderungen im Fördergeschehen reagieren, wenn etwa mehr Anträge gestellt als bewilligt werden können. In diesen Fällen müssen die Länder eigene Regeln entwickeln, welche Anträge Vorrang erfahren sollen.

Die Behörden sind verpflichtet, den Förderprozess genau zu dokumentieren, was die Rechnungshöfe (vgl. Abschnitt 2.1.1) umfassend prüfen. In dieser Phase wird eine ganze Reihe an Daten generiert, die den Ausgangspunkt für Evaluationen in Phase 5 darstellen. Die Daten des Förderprozesses (insbesondere Empfänger der Förderung; Daten für die Antragstellung, Bewilligung, Projektbeginn und -ende; Höhe der Förderung) übermitteln die Länder in einem standardisierten Prozess mit Hilfe elektronischer Meldesysteme an die Bundesregierung, die sie entsprechend archiviert.

2.2.5 Politikevaluation

In der *fünften* Phase erfolgt idealerweise eine **Reflektion der Ergebnisse der Politik**. Im Mittelpunkt steht die Analyse, ob sich ein ursächlicher **Zusammenhang zwischen der Intervention und dem Erreichen wirtschaftspolitischer Ziele** nachweisen lässt. Ein wichtiger Punkt ist zudem, welche **Nebenwirkungen** die Politik hatte und ob dadurch andere Ziele beeinträchtigt wurden. Kamen mehrere Politikoptionen zur Anwendung, sollte geprüft werden, welche dieser Maßnahmen das beste **Kosten-Nutzen-Verhältnis** aufweist. Hat die Intervention die an sie gestellten Erwartungen nicht erfüllt, dann kann die Politik nachsteuern oder sich gegebenenfalls für ein Auslaufen der Maßnahme entscheiden.

Den kritischen Punkt in dieser Phase stellt die Analyse des Zusammenhangs zwischen der Intervention und dem Erreichen wirtschaftspolitischer Ziele dar. Um glaubwürdige Aussagen aus diesen Analysen ableiten zu können, müssen überzeugende Untersuchungsdesigns entwickelt werden, die dem Stand der internationalen Wirkungsforschung genügen.

Die Methodik der Evaluation der *deutschen Regionalpolitik* ist in einem sogenannten Evaluationsplan detailliert beschrieben (vgl. Bundesanzeiger 2018). Zur Anwendung kommen kontrafaktische Untersuchungsdesigns, die dem internationalen Stand der Evaluationsforschung entsprechen und damit zuverlässige Aussagen über die Wirkungen von Regionalpolitik auf verschiedene wirtschaftspolitische Zielgrößen erlauben.

Die Bundesregierung vergibt in einem standardisierten wettbewerblichen Ausschreibungsverfahren einen wissenschaftlichen Dienstleistungsauftrag für die Evaluationsstudie an externe Gutachter. Die Qualität des Evaluationsberichts wird von

der EU begutachtet. Die Ergebnisse der Evaluation müssen Bund und Länder für die Verlängerung des Programms bei der EU berücksichtigen. Die Evaluationsstudien müssen der breiten Öffentlichkeit in geeigneter Form bereitgestellt werden.

2.2.6 Zusammenfassung zur Einbettung von Evaluationen in einen Politikzyklus

Die meisten Volkswirtschaften sind nach marktwirtschaftlichen Gesichtspunkten organisiert. Die Entscheidungen über den Einsatz von Ressourcen fallen auf individueller Ebene. In manchen Fällen funktioniert der Marktmechanismus jedoch nicht vollkommen. Von bestimmten Gütern produziert die Volkswirtschaft zu viel oder zu wenig. Zudem kann die Gesellschaft die durch den Marktmechanismus zustande gekommene Verteilung der Güter als ungerecht empfinden. Diese Umstände rechtfertigen politische Interventionen.

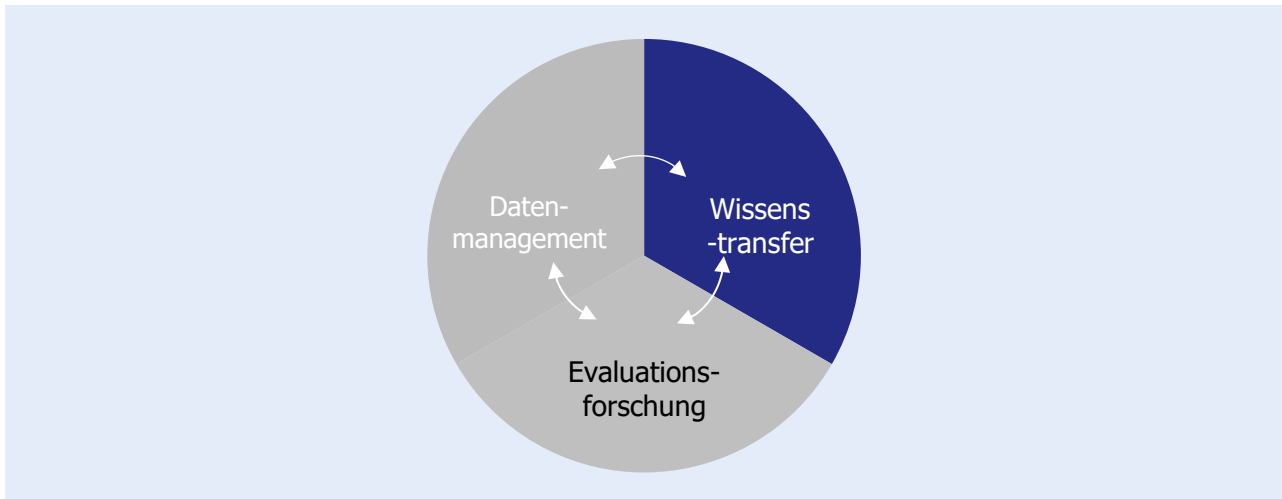
Die Entwicklung und Anwendung staatlicher Interventionen lässt sich in einem idealisierten Politikzyklus abbilden. Den Ausgangspunkt in diesem Zyklus bildet die Definition von konkreten Zielen und die Analyse, ob der Zielwert in der Volkswirtschaft erreicht wird. Nur wenn eine nicht tolerierbare Diskrepanz zwischen dem Zielwert und dem Ausgangszustand besteht, liegt überhaupt eine Rechtfertigung für einen staatlichen Eingriff vor. Für die Lösung des Problems stehen der Politik häufig verschiedene Optionen zur Verfügung. Eine umfassende Analyse theoretisch fundierter Wirkungszusammenhänge offenbart, welche der Optionen sich am besten zur Lösung des Problems eignen. Sie sollte auch aufzeigen, wo sich Zielkonflikte entwickeln könnten. Daran schließt sich die Entscheidung über die konkrete Ausgestaltung des Regelwerks für die Intervention an. Rechtssicherheit erlangen politische Entscheidungen über parlamentarische Prozesse. Am Ende dieser Kette steht der Erlass von Gesetzen, Verordnungen und Richtlinien. Die praktische Umsetzung des Rechts obliegt den Behörden auf den verschiedenen administrativen Ebenen. Diese müssen – im Einklang mit dem Rechtsstaatsprinzip – die Anwendung und Umsetzung der Politiken dokumentieren. Da das Wirtschaftsgeschehen komplexen Abläufen folgt, lassen sich theoretische Wirkungszusammenhänge nicht immer exakt prognostizieren.

Aus diesem Grund ist eine Evaluation der Maßnahmen dringend geboten. Diese sollte sich am internationalen Stand der Wirkungsforschung orientieren, um zuverlässige Aussagen über den Zusammenhang zwischen einer Intervention und dem Erreichen wirtschaftspolitischer Ziele zu generieren. Auf Basis der Erkenntnisse aus der Evaluation kann die Politik die Entscheidung über die Fortführung der Intervention oder über mögliche Anpassungen der Maßnahme treffen.

2.3 AUFGABEN IN EINEM SYSTEM EVIDENZBASIERTER EVALUATIONEN

Dieser Abschnitt argumentiert, dass die Umsetzung von Phase (5) im Politikzyklus, d. h. die Politikevaluation, nur dann gelingt, wenn drei Aufgaben eng ineinandergreifen: die Evaluationsforschung, das Datenmanagement sowie der Wissenstransfer (Abb. 2).

Abb. 2: Ineinander greifende Aufgaben bei evidenzbasierten Evaluationen



Quelle: Eigene Darstellung

Sobald auch nur eine Aufgabe in diesem System evidenzbasierter Evaluationen fehlt, können Aussagen über die Wirkungen wirtschaftspolitischer Maßnahmen nicht mehr umfassend getroffen werden. Politische Lernprozesse geraten ins Stocken. Dadurch arbeiten Institutionen ineffizient, was sich letztendlich in der langen Frist in Wohlfahrtseinbußen niederschlägt. Die folgenden Ausführungen skizzieren diese drei Aufgabenbereiche in ihren Grundzügen und demonstrieren am Beispiel der *deutschen Regionalpolitik*, wie dieses System in der praktischen Arbeit an Evaluationsstudien zur Anwendung kommt.

2.3.1 Evaluationsforschung

Der Abschnitt 2.1.3 beschrieb, dass das **Kontrafaktum** das Leitbild internationaler Wirkungsforschung darstellt. Der kausale Effekt einer politischen Maßnahme ist die Differenz in der Zielgrößen zwischen einer **Behandlungs-** und einer **Kontrollgruppe**. Eine wichtige Voraussetzung für die Aussagekraft der Ergebnisse ist, dass sich die Teilnehmer in der Behandlungs- und der Kontrollgruppe in allen Merkmalen gleichen, die sowohl die Wahrscheinlichkeit für den Erhalt der Maßnahme als auch die Ergebnisgröße selbst beeinflussen könnten. Systematische Unterschiede zwischen einer Behandlungs- und Kontrollgruppe sind aus wissenschaftlicher Perspektive für die exakte Bestimmung der Effekte einer Politikmaßnahme sehr problematisch. So können Merkmale, welche die Selektion in die Maßnahme beeinflussen, auch einen Einfluss auf Ergebnisgrößen ausüben und damit den Effekt der Maßnahme verzerren. Es lässt sich dann nicht mehr auflösen, ob das Erreichen eines Ziels auf die Maßnahme oder die spezifischen Merkmale der Teilnehmer zurückzuführen ist.

Dieses Problem tritt insbesondere dann zutage, wenn es sich bei den Interventionen um Maßnahmen handelt, an denen die Teilnehmer freiwillig mitwirken. Hier kann es Mechanismen geben, die eine **Selbstselektion von Teilnehmern in eine Maßnahme** befördern. Dies trifft auch bei der Subventionierung von Betrieben (oder Gemeinden) im Rahmen der *deutschen Regionalpolitik* zu, da es sich hier um ein nachfragegetriebenes Programm handelt.

Systematische Unterschiede und damit verbundene mögliche Verzerrungen bei der Analyse der Effekte können sich beispielsweise aus Differenzen der räumlichen Lage der Betriebe ergeben. Heckman, LaLonde und Smith (1999) zeigen etwa, dass das wirtschaftliche Umfeld die ökonomische Leistungsfähigkeit von Personen oder Betrieben beeinflusst und bei der Analyse von Politikmaßnahmen berücksichtigt werden sollte. Pellegrini und Centra (2006) weisen auf die verzerrende Wirkung von Unterschieden in der sektoralen Zusammensetzung von geförderten und nicht-geförderten Betrieben hin. Existieren beispielsweise positive (oder negative) temporäre Schocks in bestimmten Wirtschaftszweigen, können Betriebe in einem Sektor mehr (oder weniger) bereit sein, zu investieren und sich unabhängig von der Inanspruchnahme der Förderung positiver (oder negativer) entwickeln. Gleiches ist für eine Vielzahl weiterer beobachtbarer (bspw. Unternehmensgröße) und nicht-beobachtbarer (bspw. Managementfähigkeiten) betrieblicher Variablen denkbar (Kugler, Schwerdt und Wößmann 2014). Die Möglichkeit der Existenz von Selektionsverzerrungen muss daher zwingend in einer Evaluation beachtet werden, damit die Höhe und Richtung der Effekte einer Fördermaßnahme glaubwürdig sind und nicht verzerrt wiedergegeben werden (siehe auch Bernini und Pellegrini 2011 für eine Diskussion der Problematik im Rahmen der Evaluation der italienischen Investitionsförderung). Für die betriebliche Investitionsförderung im Rahmen der deutschen Regionalpolitik bestehen ebenfalls sehr deutliche Hinweise für eine Selbstselektion von Betrieben in die Förderung (Brachert, Brautzsch, Dettmann, Giebler, Schneider und Titze 2020, Kapitel 2 und 3 sowie die dort angegebene Literatur). Anliegen aller modernen Verfahren der Wirkungsforschung ist es, das **Selektionsproblem** – so gut es geht – **auszuschließen**, um eine Verzerrung der Richtung und der Höhe der Effekte zu vermeiden. Gelingt dies nicht ausreichend, dann würde die Politik ihre Entscheidungen auf einer unzureichenden Informationsbasis treffen. Eine detaillierte Diskussion der Verfahren erfolgt in Abschnitt 0.

2.3.2 Datenmanagement

Gute Evaluationen brauchen gute Daten. Für die Durchführung der Evaluation ist es von Vorteil, wenn der gesamte Politikzyklus dokumentiert ist und die Politik die Frage der Daten in jeder Phase mitdenkt. Die Initiativen der EU sowie des NKR weisen den Weg in die richtige Richtung (vgl. Abschnitt 2.1.2). Dennoch besteht weiterer Verbesserungsbedarf. Im Besonderen betrifft das die **Daten** zu den **Zielvariablen**, **Interventionsregeln** sowie zur **Anwendung der Interventionen**. Der letztgenannte Punkt berührt die systematische und umfassende Erhebung von Informationen über die Zuordnung der Teilnehmer in eine Maßnahme – insbesondere die **eineindeutige Identifikation des Teilnehmers**; Ort, Zeitpunkt, Dauer sowie die Intensität der Maßnahme.

Der vorangegangene Abschnitt verdeutlichte, dass es für ein glaubwürdiges kontrafaktisches Design wichtig ist, das **Problem der Selbstselektion von Teilnehmern in eine Maßnahme** adäquat zu adressieren. Das Untersuchungsdesign muss sicherstellen, dass sich die Teilnehmer der Behandlungs- und der Kontrollgruppe in

allen Merkmalen gleichen, die sowohl die Wahrscheinlichkeit für die Teilnahme in der Maßnahme als auch die Zielvariable beeinflussen könnten. Im Idealfall bilden die Daten einen Zeitraum vor der Einführung der Maßnahme ab, was Aussagen über die Entwicklung der Teilnehmer vor der Einführung der Maßnahme zulässt. Evidenzbasierte Evaluationen benötigen daher sehr **umfangreiche Daten** über die **Teilnehmer**, und zwar sowohl über jene der **Behandlungs-** als auch über diejenigen der **Kontrollgruppe**. Die gilt nicht nur für die individuellen Charakteristika, sondern auch für die Zielvariable(n).

Eine Möglichkeit für die Generierung dieser Informationen besteht darin, die Daten im Rahmen von **Befragungen** neu zu erheben. Dies ist häufig mit erheblichen Kosten, sowohl für die Befragten als auch für die Befragenden, verbunden. Darüber hinaus kann niemand in diesem Rahmen verpflichtet werden, an einer Befragung teilzunehmen. Im Besonderen trifft dies auf die Teilnehmer der Kontrollgruppe zu. Diese Vorgehensweise kann zu einer **Verzerrung der Antworten** führen, wenn nur ganz bestimmte Teilnehmer an der Befragung teilnehmen. Eine weitere Quelle für Verzerrungen liegt in den Antworten selbst, da die Befragungsteilnehmer nicht verpflichtet sind, die Wahrheit zu offenbaren. Letztendlich schafft man sich über diese Art der Informationsgewinnung weitere Probleme für Verzerrungen durch eine Selbstselektion der Teilnehmer in die Befragung.⁷

Die kostengünstigere und qualitativ **bessere Alternative** liegt im Zugriff auf Daten, die die **amtliche Statistik** bereits erhoben hat. Hier sind die Individuen verpflichtet, an der Befragung teilzunehmen und wahrheitsgemäß zu antworten (vgl. hierzu auch den Abschnitt 2.1.4).

Die Daten der amtlichen Statistik würden dann **verknüpft** mit den **administrativen Daten** über die **Durchführung der Maßnahmen**. Im besten Fall erheben die Behörden geeignete **Identifikatoren**, d. h. eindeutige Schlüsselnummern in den amtlichen Statistiken (z. B. Sozialversicherungsnummern, Betriebsnummern, Steuernummern, Handelsregisternummern usw.), während der Durchführung der Maßnahmen. Über diese Identifikatoren lassen sich die administrativen Daten sehr zuverlässig mit den Daten der amtlichen Statistik verknüpfen. Wenn diese Identifikatoren nicht vorliegen, dann funktioniert das Zusammenführen der Daten auch über **Record-Linkage-Techniken**. Diese Methoden berechnen Ähnlichkeiten in Namensbestandteilen von Personen oder Betrieben zwischen den unterschiedlichen Datensätzen. Überschreiten diese Ähnlichkeiten einen vom Evaluator festgelegten Schwellenwert, dann kann man die Datenfelder verknüpfen. Diese Methode produziert naturgemäß falsch-negative und falsch-positive Matches, weshalb sie auch nur eine Alternative zu den Verknüpfungen auf Basis amtlicher Identifikatoren sein kann (vgl. Brachert, Giebler, Heimpold, Titze und Urban-Thielicke 2018, Abschnitt 3).

⁷ So dürfte etwa – gerade bei nachfragegetriebenen Programmen – die Bereitschaft unter den Teilnehmern der Behandlungsgruppe die Bereitschaft höher sein, erstens an der Befragung teilzunehmen und zweitens den Grad der Zielerreichung bewusst zu überschätzen, damit das Programm fortgeführt wird.

Ein zuverlässiges Datenmanagement im Rahmen eines Systems evidenzbasierter Evaluationen muss zudem eine **systematische Archivierung** aller verwendeten Datensätze, Aufbereitungsprozeduren sowie Berechnungscodes enthalten, was die Grundlagen schafft für Reproduzierbarkeit, Replizierbarkeit sowie Verallgemeinerung der Ergebnisse (vgl. etwa Bollen, Cacioppo, Kaplan, Krosnick und Olds 2015).

Die Regeln für die Anwendung für die *deutsche Regionalpolitik* lagen vor in Form von Politikdokumenten, sogenannten Koordinierungsrahmen (früher: Rahmenplänen) (vgl. bspw. Bundesanzeiger 2018), Dokumenten aus den Notifizierungsprozessen des Programms bei der EU sowie sonstigen Veröffentlichungen (bspw. Eckey 2008; Alm und Fisch 2014). Die Förderregeln wurden von Brachert, Brautzsch, Dettmann, Giebler, Schneider und Titze (2020) digitalisiert und in geeignete kontrafaktische Untersuchungsdesigns übersetzt.

Die Länder erheben die administrativen Daten zum Förderprozess (insbesondere Name und Ort des investierenden Betriebs, Daten zum Antragsprozess, Beginn und Ende des geförderten Investitionsprojekts, Fördersatz und Fördersumme) und melden diese an die Bundesregierung, welche auch die Archivierung übernimmt. Diese Daten enthalten zudem einen Identifikator der amtlichen Statistik. Konkret handelt es sich um die Betriebsnummer, welche die Bundesagentur für Arbeit im Rahmen der Meldungen zur sozialversicherungspflichtigen Beschäftigung vergibt. Über die Betriebsnummer lassen sich die Förderdaten mit der Statistik der sozialversicherungspflichtig Beschäftigten verknüpfen, die im Grunde eine Vollerhebung für diesen Beschäftigungstypus darstellt und die Biographie der betrieblichen Entwicklung beinhaltet (vgl. Abschnitt 2.1.4).⁸ Dieser Weg stößt allerdings dort an seine Grenzen, wo es um die Erfassung der Aktivitäten von selbstständig Erwerbstätigen (bspw. in Familienbetrieben, wie sie in der Landwirtschaft vorrangig anzutreffen sind) geht.

Die administrative Förderstatistik können noch mit einem zweiten Datensatz der amtlichen Statistik verknüpft werden, und zwar den Erhebungen für das Verarbeitende Gewerbe, welche die FDZ der Statistischen Ämter des Bundes und der Länder bereitstellen (vgl. Abschnitt 2.1.4). Eine Zusammenführung mit diesem Datensatz ist über einen Identifikator, die Handelsregisternummer, möglich. Diese Nummer erheben die Bewilligungsbehörden allerdings nicht. Eine Verknüpfung lässt sich dennoch organisieren, und zwar über eine Zwischenstufe, wo die administrativen Förderdaten zunächst mit den Amadeus-Daten des privaten Anbieter Bureau van Dijk zusammengespielt werden. Über diesen Datensatz erhält man Auskunft über die Handelsregisternummer, welche nunmehr als Identifikator für die Verknüpfung mit den Daten der amtlichen Statistik zum Verarbeitenden Gewerbe dient. Naturgemäß lassen sich über diesen Umweg weniger Informationen zusammenspielen, als wenn die Behörden die entsprechende Nummer von

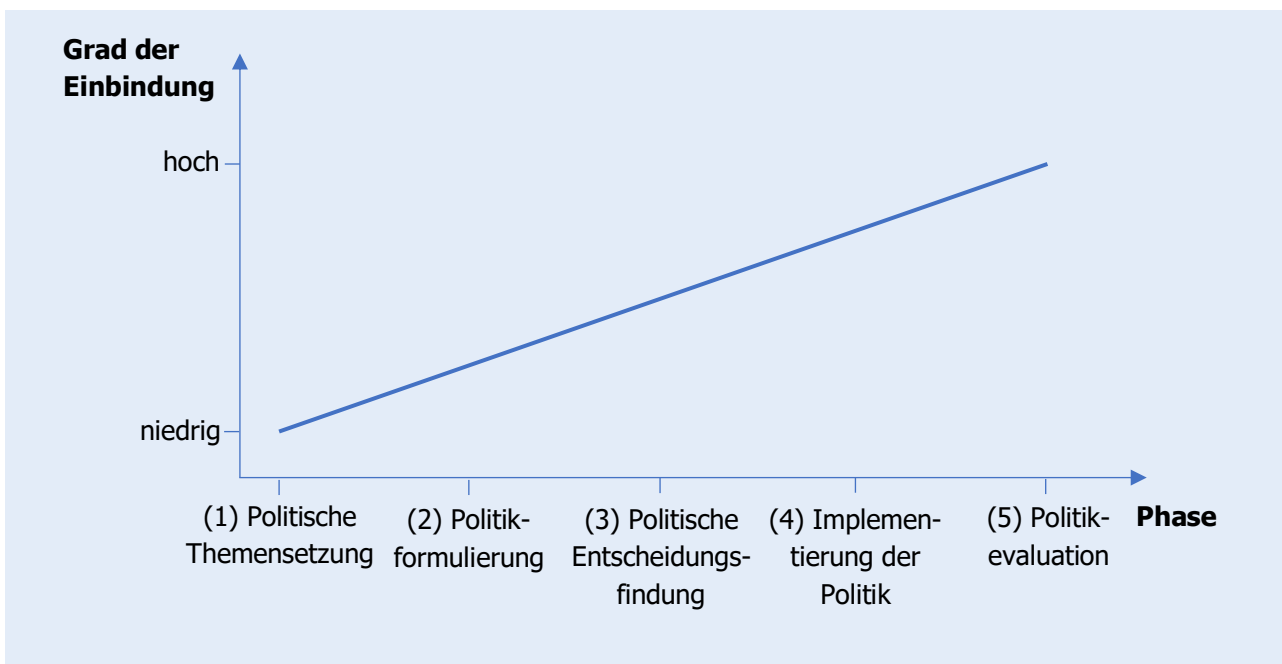
⁸ Der Anteil sozialversicherungspflichtig Beschäftigter an der gesamten Beschäftigung beträgt in Deutschland nach Auskunft des Statistischen Bundesamts (www.destatis.de => Arbeit (Arbeitsmarkt) => Erwerbstätigkeit) etwa 75% (Stand: 12.02.2021).

Anfang an mit erhoben hätten. Mit Hilfe der beschriebenen Daten sind Evaluierungen zu den Effekten der deutschen Regionalpolitik möglich, die die Anforderungen des State-of-the-Art der internationalen Wirkungsforschung vollumfänglich erfüllen (vgl. Brachert, Giebler, Heimpold, Titze und Urban-Thielicke 2018 sowie Brachert, Brautzsch, Dettmann, Giebler, Schneider und Titze 2020). Arbeiten zur Archivierung der Datensätze, Aufbereitungsprozeduren sowie Programmcodes für die Evaluierung der deutschen Regionalpolitik hat das Forschungsdatenzentrum am Leibniz-Institut für Wirtschaftsforschung (www.iwh-halle.de/fdz) begonnen.

2.3.3 Wissenstransfer

Der Transfer umfasst mehrere Dimensionen bilateralen Wissensaustausches zwischen Politik, Wissenschaft und der breiten Öffentlichkeit. Die Dimension des **kontinuierlichen gegenseitigen Austauschs von Wissen zwischen der Politik und der Wissenschaft** sollte sich nicht nur auf die Phase (5) Politikevaluation erstrecken, sondern schon in den frühen Phasen beginnen (Abb. 3).

Abb. 3: Einbindung der Wissenschaft in den Politikzyklus



Quelle: Eigene Darstellung

Vorstellbar ist etwa der folgende **interaktive Prozess**, wobei der Grad der Einbindung der Wissenschaft anfangs gering ist und mit späteren Phasen zunimmt: In *Phase (1) Politische Themensetzung* könnte die Politik die Wissenschaft über ihre politischen Ziele informieren und die Wissenschaft Unterstützung bereitstellen, wie sich die Ziele in objektiv messbare Indikatoren operationalisieren lassen und wo Möglichkeiten und Grenzen der Indikatorik bestehen. In *Phase (2) Politikformulierung* könnte die Wissenschaft Beiträge zu theoretischen Wirkungszusammenhängen liefern und auf mögliche Zielkonflikte hinweisen. Die Politik müsste der Wissenschaft gegenüber darlegen, welche Abwägungen eine Rolle bei der Auswahl der Politikmaßnahme spielen. In *Phase (3) Politische Entscheidungsfindung* kann die Wissenschaft wertvolle Impulse

liefern, wie die Politik bei der Ausgestaltung der Interventionsregeln bereits kontrafaktische Untersuchungsdesigns implementieren könnte.⁹ Aufgabe der Politik wäre es, der Wissenschaft umfassende Informationen über die Interventionsregeln zu liefern. *Phase (4) Implementierung der Politik* beinhaltet die konkrete Umsetzung der Politik. Hier wäre es enorm wichtig, die Wissenschaft (und die Datenproduzenten, d. h. die FDZ) bei der Organisation der Generierung von administrativen Daten zur Anwendung der Maßnahme einzubeziehen. Das betrifft insbesondere die Erfassung von Identifikatoren für die Teilnehmer, was das Zusammenspielen mit externen Daten, etwa der amtlichen Statistik, ermöglichen würde. Versäumnisse in der Erhebung notwendiger Daten lassen sich – wenn die Maßnahme erst einmal eingeführt wurde – kaum mehr korrigieren. Die Politik müsste die so generierten administrativen Daten über die Anwendung der Maßnahmen dann der Wissenschaft zur Verfügung stellen. Die *Phase (5) Politikevaluation* befasst sich mit der Erarbeitung der Evaluation im engeren Sinne. Hier kommt es vor allem auf den Wissensaustausch innerhalb des Wissenschaftssystems an. Die Evaluationsstudie sollte den in der Wissenschaft üblichen Begutachtungsprozess durchlaufen, um eine hohe Qualität sicher zu stellen. Nach erfolgreichem Abschluss des Begutachtungsprozesses können Handlungsempfehlungen an die Politik gegeben werden. Der Transfer der Ergebnisse erfolgt in der Regel über schriftliche Berichte, Anhörungen, Stellungnahmen usw. Ob die Politik den Empfehlungen folgt, hat sie letztendlich selbst zu verantworten. Es gibt keinerlei Verpflichtungen, dass die Politik die Empfehlungen umsetzen muss.

Die Dimension des **kontinuierlichen gegenseitigen Austauschs von Wissen zwischen der Wissenschaft und der breiten Öffentlichkeit** könnte zwei Elemente enthalten. *Erstens* empfehlen Buch, Patzwald, Riphahn und Vogel (2019), evidenzbasierte Evaluationen der breiten Öffentlichkeit über **Datenbanken** – etwa **internetbasierte Plattformen**, die den Charakter von **Bibliotheken** haben – frei zugänglich zu machen. Die Autoren verweisen auf zwei prominente Beispiele, wie sich der Transfer in die Praxis organisieren lässt, und zwar das *Cochrane-Netzwerk* im Bereich der Medizin sowie das *What-Works-Netzwerk* in Großbritannien. In beiden Fällen handelt es sich um **institutionell unabhängige Plattformen**, die die Ergebnisse in kondensierter Form für Forscher und Praktiker in den jeweiligen Disziplinen bereitstellen. Die Unabhängigkeit ist wichtig, um eine Einflussnahme auf die Bereitstellung von Evidenz durch Interessengruppen zu verhindern. Im Fall medizinischer und pharmazeutischer Forschung

⁹ So könnten die Interventionen etwa zeitliche Regeln beinhalten, sodass eine Gruppe von zufällig ausgewählten Teilnehmern eher die Maßnahme erhält (Behandlungsgruppe) und die übrigen Teilnehmer mit Zeitverzug in die Maßnahme aufgenommen werden (Behandlungsgruppe). Vorstellbar wäre auch die Zuteilung von Maßnahmen anhand von kontinuierlichen Bewertungsskalen und eines exogen gegebenen Schwellwertes, wo Teilnehmer auf der einen Seite des Schwellwertes in den Genuss der Maßnahme kommen (Behandlungsgruppe), während die Teilnehmer mit einer fast gleichen Bewertung auf der anderen Seite des Schwellwertes von der Maßnahme ausgeschlossen sind. Die Details dieser Designs bespricht Abschnitt 2.4.

etwa erfolgt ein Großteil von Forschungsaktivitäten in privaten Unternehmen.

Das *Cochrane-Netzwerk* (www.cochrane.org) nimmt daher keine Gelder aus der Privatwirtschaft an und achtet streng darauf, dass kein Interessenkonflikt besteht. Die Finanzierung erfolgt vor allem aus Mitteln, die die Regierungen der Nationalstaaten bereitstellen. Über sogenannte Review Groups erfolgt eine Begutachtung der vorliegenden Studien nach vorab festgelegten wissenschaftlichen Qualitätsstandards. Die Vorgehensweise und Ergebnisse jeder Studie sowie ihre wissenschaftliche Bewertung im Rahmen des Review Verfahrens werden in eine elektronische Bibliothek eingepflegt, die für jedermann frei zugänglich ist. Mit Hilfe dieser Bibliothek können Übersichtsstudien, sogenannte Metaanalysen, erstellt werden, die die Evidenz aus den Einzelstudien in kompakter Form zusammenfassen. Damit steht Wissenschaftlern, Praktikern sowie Politikern eine evidenzbasierte Informationsquelle zur Verfügung, anhand der Entscheidungen nach dem aktuellen Stand der Forschung vorbereitet werden können.

Ein anderes Beispiel stellt das *What-Works-Netzwerk* in Großbritannien dar (<https://www.gov.uk/guidance/what-works-network>). Finanziert durch die Regierung und weitere Quellen verfolgt diese Initiative – vergleichbar dem *Cochrane-Netzwerk* – das Ziel, evidenzbasierte Informationen für die Regierungen und andere öffentliche Institutionen bereit zu stellen. Aktuell besteht das *What-Works-Netzwerk* aus 13 Einrichtungen verschiedener Politikfelder.¹⁰

Diese Datenbanken stellen zudem eine Voraussetzung dafür dar, die Evidenz aus der Vielzahl an verschiedenen Studien in sogenannten **Meta-Studien** weiter zusammen zu fassen. Praktiker verfügen in der Regel nicht über die zeitlichen Ressourcen, jede einzelne Studie intensiv durchzuarbeiten. Meta-Studien können dabei helfen, in kurzer Zeit einen fundierten Überblick über ein Thema zu erhalten und darauf aufbauend fundierte Entscheidungen zu treffen.

Ein *zweites* Element gegenseitigen Wissensaustauschs zwischen der Wissenschaft und der breiten Öffentlichkeit könnten **Programme der Aus- und Weiterbildung** sein. Die Idee evidenzbasierter Evaluationen hat heute schon weitgehend Eingang in die **Lehrpläne** der **Studierendenausbildung** gefunden (bspw. im Master of Policy Advice Programm an der Martin-Luther-Universität Halle-Wittenberg [MLU] in Kooperation mit dem Leibniz-Institut für Wirtschaftsforschung Halle [IWH] [<http://master-policy-advice-halle.de>] oder der **Graduiertenausbildung** (bspw. im International Graduate Program Evidence-Based Economics an der Ludwig-Maximilians-Universität München [LMU], der Friedrich-Alexander-Universität Erlangen-Nürnberg [FAU] sowie der Universität Regensburg im Rahmen des Bayrischen Elitenetzwerks [<https://www.ebe.econ.uni-muenchen.de>])).

¹⁰ Zu den Politikfeldern gehören aktuell (Stand 15.02.2021) Health and social care, Educational achievement, Crime reduction, Early intervention, Local economic growth, Improved quality of life for older people, Wellbeing, Homelessness, Children's social care, Youth offending, Youth employment, Higher Education.

Darüber hinaus könnten **Weiterbildungsangebote** für **Mitarbeitende** in den **Behörden** und **Journalisten** sinnvoll sein. Mitarbeitende in den Behörden stellen deswegen eine interessante Zielgruppe dar, da die Fachreferate in den Bundes- und Landesministerien weiter für die Durchführung von Evaluationen verantwortlich sein werden, zumindest in der mittleren Frist. Wie in Abschnitt 2.1.2 erläutert, haben die Ausführungen des Normenkontrollrats allenfalls empfehlenden Charakter für den Gesetzgebungsprozess auf der Bundesebene. Soweit bekannt, befinden sich derartige Initiativen auf der Ebene der Länder, wenn überhaupt, in sehr frühen Phasen. Regelungen zu Mindestanforderungen für Evaluationen auf der Seite der Rechnungshöfe gibt es nicht. Aus diesen Gründen sind Schulungen für die Mitarbeitenden in den Behörden sehr vielversprechend, um die öffentliche Verwaltung weiter mit der Idee evidenzbasierter Evaluationen zu durchdringen.

Neben den Mitarbeitenden in den Behörden sind auch Journalisten eine relevante Zielgruppe von Weiterbildungsaktivitäten für evidenzbasierte Evaluationen. Über die Arbeit der Journalisten erhält die breite Öffentlichkeit Zugang zu kritischen Reflektionen des wirtschaftlichen Geschehens. Die Bereitstellung evidenzbasierten Wissens über unabhängige Medien würde zudem zu einer Erhöhung von Transparenz beitragen.

2.3.4 Zusammenfassung zu den Aufgaben in einem System evidenzbasierter Evaluationen

Abschnitt 0 hat argumentiert, dass evidenzbasierte Evaluationen weit mehr Aufgaben umfassen müssen als die Wirkungsforschung im engeren Sinne. Vielmehr sollte diese eingebettet sein in ein umfangreiches Management von Daten sowie mehrdimensionalen Formen kontinuierlichen Wissensaustausches.

Die internationale Evaluationsforschung orientiert sich heute am Standard des Kontrafaktums. Hier geht es um die Frage, was passiert wäre, wenn es die Intervention nicht gegeben hätte. Dieser Zustand ist in der Praxis nicht beobachtbar. Das Kontrafaktum lässt sich aber „simulieren“, indem man die Entwicklung von Teilnehmern in einer Behandlungsgruppe mit derjenigen der Kontrollgruppe vergleicht. Der Unterschied in der Zielvariablen zwischen beiden Gruppen lässt sich als Effekt der Maßnahme interpretieren. Zuverlässige Aussagen sind nur möglich, wenn sich die Teilnehmer der Behandlungs- und Kontrollgruppe in allen Merkmalen gleichen, die sowohl die Wahrscheinlichkeit für den Erhalt der Maßnahme als auch die Zielgröße beeinflussen könnten. Findet Selbstselektion in die Maßnahme statt, wird der Effekt der Maßnahme verzerrt geschätzt. Fundierte Aussagen zu den Wirkungen der Maßnahme sind dann nicht mehr möglich.

Gute Evaluationsforschung ist nur auf Basis guter Daten möglich. Um ein glaubwürdiges Evaluationsdesign entwickeln zu können, ist eine umfangreiche Kenntnis der Interventionsregeln erforderlich, was nur in engem Wissensaustausch mit der Politik gelingt. Idealerweise wird die Wissenschaft bereits in frühen Phasen des Politikzyklus eingebunden. Damit kann es etwa gelingen, Interventionsregeln zu erarbeiten, in denen das Kontrafaktum bereits „enthalten“ ist.

Eine frühe Einbindung der Wissenschaft in den Politikzyklus ist auch deswegen ratsam, um mit der Einführung der Maßnahme die Grundlagen für die Erhebung der Daten zu schaffen. Dies trifft zu auf die Zielvariablen und Informationen über die Merkmale der Teilnehmer, um das Selektionsproblem bestmöglich zu adressieren. Für diesen Prozess müssen nicht alle Daten aufwendig neu erhoben werden – vielfach liegen diese bereits in der amtlichen Statistik vor. Diese könnten dann mit den administrativen Daten aus der Anwendung der Maßnahme und den Interventionsregeln verknüpft werden. Dafür ist es allerdings notwendig, entsprechende Identifikatoren für das Zusammenführen von Datensätzen im administrativen Prozess der Anwendung der Maßnahme mit zu erheben und die datenschutzrechtlichen Grundlagen für die Arbeit mit derartig erzeugten Datensätzen zu organisieren.

Evidenzbasierte Evaluationen sind nur dann glaubwürdig, wenn Transparenz besteht: über die Evaluationsfragen, das Evaluationsdesign, die verwendeten Daten und Methoden sowie die Ergebnisse. Hier können internetbasierte Plattformen einen Beitrag leisten. Meta-Studien würden es erlauben, Praktiker sowie die breite Öffentlichkeit darüber zu informieren, was wirkt und was am besten funktioniert. Abschließend sollte es auch Aufgabe eines Systems evidenzbasierter Evaluationen sein, Weiterbildung für Praktiker und die breite Öffentlichkeit anzubieten. Nur so kann es gelingen, dass sich Ideen und Nutzen evidenzbasierter Evaluationen tief in der Gesellschaft verankern.

2.4 EIGNUNG KONTRAFAKTISCHER ANALYSETOOLS ZUR EVALUIERUNG WIRTSCHAFTSPOLITISCHER MASSNAHMEN¹¹

Kontrafaktische Analysetools stehen im Mittelpunkt der *Evaluationsforschung* im System evidenzbasierter Evaluationen (Abbildung 2). Gegenstand dieser Verfahren ist es, das Problem der Selektionsverzerrung so gut wie möglich in den Griff zu bekommen. Hierfür bieten sich unterschiedliche wissenschaftliche Methoden an. Diese lassen sich grundsätzlich unterscheiden in experimentelle und sogenannte quasi-experimentelle Methoden (Abbildung 4).

Während im Fall von Experimenten (die sich wiederum in Labor- und Feldexperimente differenzieren lassen) die Zuordnung der Untersuchungsobjekte zur Behandlungs- und Kontrollgruppe unter kontrollierten Bedingungen stattfindet, geschieht dies im Fall von Quasi-Experimenten durch äußere Umstände, anhand derer sich ein Experiment „simulieren“ lässt. Die letztere Gruppe umfasst insbesondere den Instrumentvariablen-Ansatz (IV-Ansatz), das Regressions-Diskontinuitäts-Design (RDD) sowie den Differenz-von-Differenzen-Ansatz (DiD). Die folgenden Ausführungen basieren auf den umfassenden Darstellungen in Alm (2013, Abschnitt C III); Kugler, Schwerdt und Wößmann (2014); Madaleno und Waight (2016); Khandker, Koolwal und Samad (2010) sowie Storey (2017).

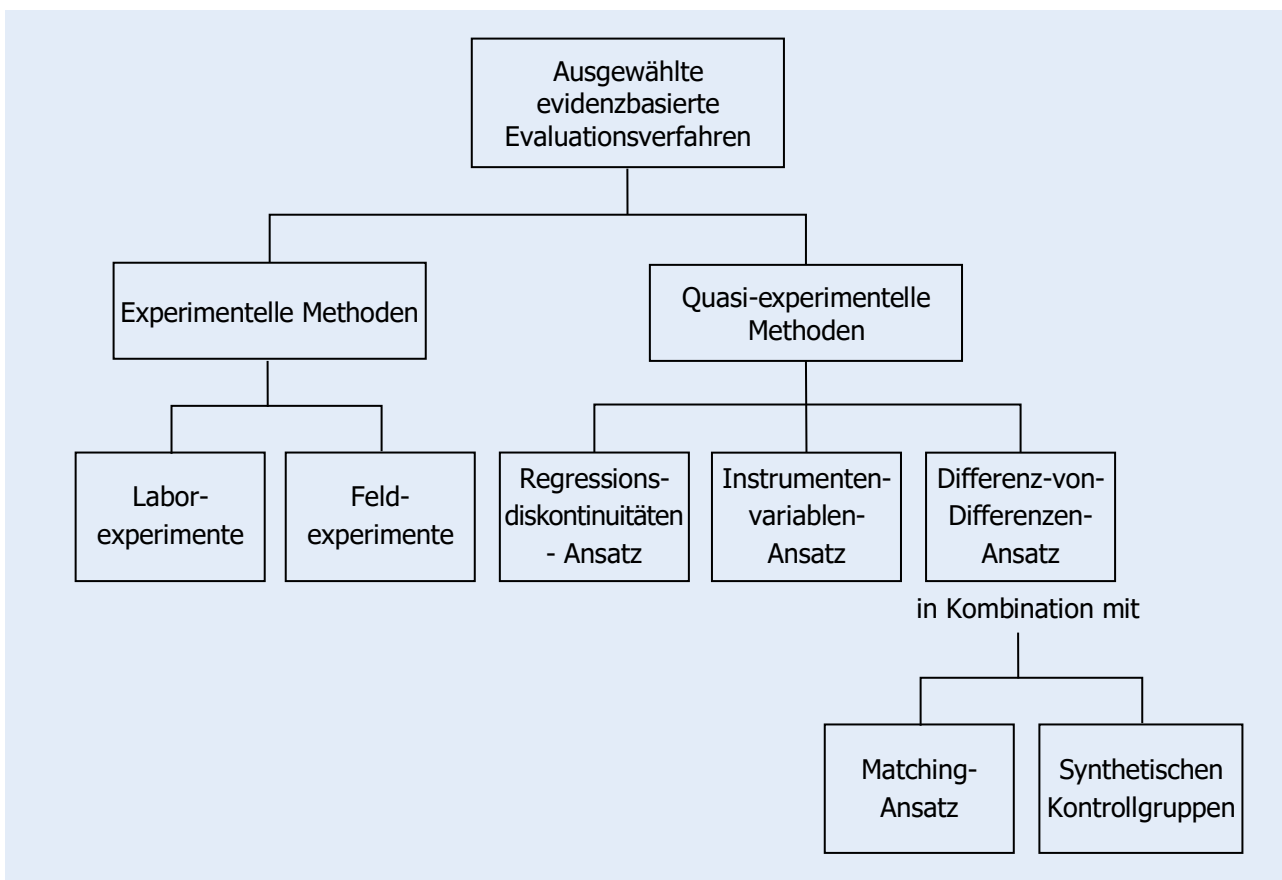
¹¹ Die Textpassagen aus diesem Abschnitt sind – bis auf den Abschnitt 2.4.1 – weitgehend dem Abschnitt 3.1 aus Brachert, Brautzsch, Dettmann, Giebler, Schneider und Titze (2020) entnommen.

Auch in diesem Abschnitt sollen die Evaluationen der deutschen Agrar- und Regionalpolitik als Beispiel dienen. Anhand dieser konkreten Fälle soll es darum gehen, die Vor- und Nachteile der einzelnen Methoden zu diskutieren.

2.4.1 Experimente

In Experimenten erfolgt die Zuteilung von Teilnehmern in die Behandlungs- und Kontrollgruppe anhand eines zufälligen Prozesses. Um zu verhindern, dass sich die Teilnehmer in der Behandlungs- und Kontrollgruppe nicht doch in relevanten Merkmalen unterscheiden, die zu einer Selektionsverzerrung führen könnten, fällt dem Randomisierungsprozess eine essentielle Bedeutung in experimentellen Untersuchungsdesigns zu. Die Literatur hält eine Reihe von Verfahren bereit, die von einfacher Randomisierung (z. B. dem Wurf einer Münze), über restringierte Randomisierung (der Auswahlprozess kontrolliert dafür, dass die Behandlungs- und Kontrollgruppe ausbalanciert sind hinsichtlich der Merkmale, die zu Selektionsverzerrung führen könnten) bis zu einer Randomisierung auf Basis vorgelagerter Matchingprozesse (wo zunächst Gruppen von Teilnehmern mit identischen Merkmalen gebildet werden und danach innerhalb dieser Gruppen eine Randomisierung erfolgt) reichen.

Abb. 4: Systematisierung von evidenzbasierten Evaluationsverfahren



Quelle: Eigene Darstellung in Anlehnung an Buch, Dettmann, Heimpold und Titze (2018)

Jedes dieser Verfahren weist naturgemäß Vor- und Nachteile auf, und die endgültige Auswahl eines Randomisierungsprozesses wird sich immer an der konkreten Fragestellung sowie dem speziellen Untersuchungsdesign orientieren müssen (Edovald

und Firpo 2016).

Experimentelle Untersuchungsdesigns weisen generell den Vorteil auf, dass eine Auswertung der Ergebnisse ohne komplexe statistisch-ökonomische Methoden möglich ist. Es müssen lediglich Mittelwerte zwischen der Behandlungs- und der Kontrollgruppe verglichen werden, um Einblicke in den Erfolg der Maßnahme zu erhalten. Ein weiterer Vorteil besteht darin, dass diese Untersuchungsansätze eine sehr hohe *interne Validität* aufweisen. Da die Zuordnung der Untersuchungseinheiten in Behandlungs- und Kontrollgruppe unter kontrollierten Bedingungen stattfindet, lassen sich störende Einflüsse, die die Messung des tatsächlichen ursächlichen Zusammenhangs zwischen einer Maßnahme und einer Ergebnisgröße verfälschen können, nahezu vollständig ausschließen.

Nachteile experimenteller Ansätze bestehen vor allem darin, dass die Ergebnisse nicht unbedingt verallgemeinerbar sind, da die Untersuchung auf einer vergleichsweise kleinen Sample beruht. In einem solchen Fall ist *externe Validität* nur bedingt gegeben. Für Laborexperimente trifft dies noch stärker zu als für Feldexperimente.

Auch wenn Behandlungs- und Kontrollgruppe absolut identisch sind hinsichtlich der relevanten Merkmale, so bedeutet das nicht, dass sich Schwierigkeiten im weiteren Verlauf des Experiments ergeben können. So kann es etwa passieren, dass Teilnehmer das Experiment nicht beenden, was zu einer Verzerrung der Ergebnisse führen könnte. Abschließend sei noch auf die generelle Machbarkeit von Experimenten hingewiesen.

In Deutschland gibt es nur wenige Beispiele für eine randomisierte Zuordnung der Teilnehmer in Maßnahmen der Wirtschaftspolitik. Für die Evaluation der *deutschen Regionalpolitik* kam dieses Verfahren bislang nicht zur Anwendung. Ebenso sind im Rahmen der GAP Feldexperimente bisher noch nicht umgesetzt worden, obwohl vor allem hinsichtlich der Ausgestaltung von Instrumenten diese Methode wichtige Erkenntnisse liefern könnte (Behaghel, Macours und Subervie 2019). Diese Art von Untersuchungen müsste in der Gesellschaft auf breiten Konsens stoßen. Es ist ex ante nicht unbedingt zu erwarten, dass die Teilnehmer die Bereitschaft mitbringen, sich einer Randomisierung auszusetzen (Edovald und Firpo 2016).

Andererseits stellen Experimente in der Medizin und der pharmazeutischen Forschung quasi das Standardverfahren dar (Buch, Patzwald, Riphahn und Vogel 2019). Auch gibt es mit der Evaluation von Politikmaßnahmen in Entwicklungsländern Beispiele, wo sich Experimente sehr wohl umsetzen ließen und dadurch bessere Erkenntnisse über die Effekte einzelner Maßnahmen gewonnen werden konnten, als wenn die Zuteilung der Teilnehmer in die Maßnahmen nach üblichen Vergabekriterien erfolgen würde (Crépon, Duflo, Gurgand, Rathelot und Zamora 2012). Im Fall der Förderung von Investitionen im Rahmen der *deutschen Regionalpolitik* etwa geschieht dies nach wie vor entlang von diskretionären Vergabeentscheidungen über komplexe Antragsverfahren, was per se eine potenzielle Quelle für Selektionsprobleme darstellt.

2.4.2 Instrumentvariablen-Ansatz (IV)

Die Grundidee dieses Ansatzes besteht darin, von der Variable, welche die Teilnahme an der Maßnahme in dem Schätzmodell abbildet, nur denjenigen Teil der Variation zu nutzen, der nicht von Endogenitätsproblemen (z. B. aufgrund von Selbstselektion in die Maßnahme) betroffen ist. Dafür kommt eine zusätzliche Variable (das Instrument) zum Einsatz, die mit der Behandlungsvariable korreliert ist, aber sonst in keinem Zusammenhang mit der interessierenden Ergebnisvariable steht.

Angewendet auf den Fall der Förderung von Investitionen in Betrieben im Rahmen der *deutschen Regionalförderung* müsste eine Variable gefunden werden, die mit der Wahrscheinlichkeit korreliert, dass der Betrieb gefördert wird und gleichzeitig unkorreliert mit dem Fehlerterm ist. Die Variable darf nur über den Förderstatus auf die Zielgröße wirken, nicht direkt.

Eine aktuelle Arbeit von Criscuolo, Martin, Overman und van Reenen (2019) verwendet den IV-Ansatz und nutzt die Veränderung der auf EU-Ebene festgelegten Politikparameter als Instrument für die Änderung der Förderhöchstsätze in den betreffenden Regionen Großbritanniens, um die Wirkungen des dortigen Investitionsförderprogramms, der Regional Selective Assistance (RSA), zu analysieren. Bezüglich der Verwendung des Instruments argumentieren die Autoren, dass die Politikparameter für den Einsatz der Förderung auf der Ebene der EU festgelegt werden und damit exogen sind. Dadurch gelingt es den Autoren, den Teil des Förderhöchstsatzes zu isolieren, der auf die durch die EU bestimmte Politikänderung zurückzuführen ist.

2.4.3 Regressions-Diskontinuitäts-Design (RDD)

Der Regressions-Diskontinuitäts-Ansatz nutzt Situationen aus, in denen die Zuteilung zu einer Interventionsmaßnahme anhand eines kontinuierlich verlaufenden Merkmals (die sogenannte Laufvariable) sowie eines (exogen festgelegten) Schwellenwertes erfolgt. Typische Beispiele sind feste Altersgrenzen, Stichtage, Grenzwerte für Gewichte und Abmessungen oder ähnliches. Wenn die Laufvariable am Schwellenwert einen kontinuierlichen Verlauf aufweist, dann vergleicht man mit Hilfe des Designs Beobachtungseinheiten, die mit Blick auf die Laufvariable keine signifikanten Unterschiede aufweisen, jedoch bei Über- bzw. Unterschreiten des Grenzwertes an der Maßnahme teilnehmen oder eben nicht. Wichtig für eine glaubwürdige Identifikation ist, dass die Beobachtungseinheiten die Laufvariable nicht manipulieren können und keine genaue Kontrolle über die Festlegung des Schwellenwerts haben.

Ein Beispiel für eine Anwendung des Regressions-Diskontinuitäts-Ansatzes findet sich in der Arbeit von Bronzini und Iachini (2014). Sie analysieren die Effekte eines Programms zur Förderung von Forschungs- und Entwicklungsaktivitäten, welches die Regierung der Region Emilia-Romagna in Italien aufgelegt hat. Interessierte Unternehmen haben Projektanträge an die Bewilligungsbehörde geschickt, die sodann von einer Expertengruppe anhand eines Punktesystems bewertet wurden. Anträge mit einem Punktwert oberhalb eines Schwellenwerts erhielten die Förderung, die anderen gingen

leer aus. Diese institutionelle Regelung nutzen die Autoren als Identifikationsstrategie und vergleichen die Entwicklung von Unternehmen, deren Anträge kurz über und kurz unter dem Schwellenwert liegen. Diese haben annähernd identische Projekte zur Begutachtung eingereicht und bilden folglich eine sehr gut vergleichbare Behandlungs- und Kontrollgruppe. Die Autoren finden dadurch heraus, dass das Programm einen positiven Effekt auf die Investitionsaktivitäten der geförderten Unternehmen (am Schwellenwert) hat. Bei diesem Verfahren ist zu bemerken, dass der Regressions-Diskontinuitätsansatz die Effekte über die Beobachtungseinheiten knapp über und unter dem Schwellenwert identifiziert. Das Verfahren bestimmt daher einen lokalen durchschnittlichen Treatment-Effekt (LATE).

2.4.4 Differenz-von-Differenzen-Ansatz (DiD)

Die Idee des Differenz-von-Differenzen-Ansatzes basiert darauf, für die Gruppe der Beobachtungseinheiten mit und ohne Intervention jeweils eine Differenz der interessierenden Zielgrößen vor Erhalt der Förderung und danach zu bestimmen. Als Effekt der Förderung wird die Differenz der Gruppendifferenzen interpretiert. Der Ansatz beruht auf der Annahme, dass sich die von der Intervention betroffenen Beobachtungseinheiten in gleicher Weise wie diejenigen ohne Intervention entwickelt hätten, wenn sie nicht betroffen gewesen wären. Diese Annahme ist nicht direkt testbar. Ein möglicher Weg, die Validität dieser Annahme zu stützen, ist es, nur diejenigen nicht-geförderten Unternehmen für die Bildung der Kontrollgruppe zu berücksichtigen, die den geförderten Unternehmen in wichtigen, die Zielgröße betreffenden, Charakteristika soweit wie möglich ähneln. Hierbei können Matching-Verfahren helfen, sogenannte „statistische Zwillinge“ zu den Beobachtungseinheiten mit Intervention unter denjenigen ohne Intervention herauszufiltern.

Die beschriebene Vorgehensweise hat sich in jüngeren Studien für die Evaluierung der Investitionsförderung im Rahmen der deutschen Regionalpolitik (Bade und Alm 2010; Kölling 2015; Dettmann, Weyh und Titze 2017; Dettmann, Titze und Weyh 2018; Brachert, Dettmann und Titze 2018) als praktikabel und wissenschaftlichen Standards entsprechend erwiesen. Die genannten Studien untersuchen die Wirkungen der deutschen Regionalpolitik auf betriebliche Zielgrößen für Deutschland (Bade und Alm 2010) und ausgewählte Bundesländer (Sachsen-Anhalt, Dettmann, Weyh und Titze 2017; Dettmann, Titze und Weyh 2018; Brachert, Dettmann und Titze 2018) bzw. ein Bündel an Unternehmenssubventionen (zu dem auch die Regionalförderung gehört, Kölling 2015). Alle genannten Studien finden einen positiven Effekt des Programms auf die Entwicklung der betrieblichen Beschäftigung.

Eine Anwendung der Kombination von Matching und Differenz-von-Differenzen-Ansatz erfolgte vor allem auch für die Evaluation der Regionalförderung in Italien. Hier existiert mit dem Gesetz 488/92 ein der deutschen Regionalförderung vergleichbares Programm. Die Untersuchungen von Bernini und Pellegrini (2011) sowie Cerqua und Pellegrini (2017) finden positive Effekte der Regionalförderung auf die Beschäftigung in den geförderten Betrieben. Bondonio und Greenbaum (2014) untersuchen für die Region Piemont ein

breites Bündel an Unternehmensförderprogrammen und weisen ebenfalls einen positiven Effekt für die Beschäftigung nach.

Michalek (2012) bewertet mittels einer Kombination von Matching-Ansätzen mit einem Differenz-von-Differenzen-Schätzer das Agrarinvestitionsprogramm in Schleswig-Holstein. Die Ergebnisse belegen zwar eine positive Wirkung auf das Betriebsergebnis der geförderten Landwirte, aber gleichzeitig verweist Michalek auf die sehr hohe Relevanz von Mitnahmeeffekten. Viele Betriebsinhaber würden die Investitionen auch ohne Förderung durch das Agrarinvestitionsprogramm durchgeführt haben.

In der Praxis gibt es allerdings auch Fälle, in denen es nur eine sehr geringe Anzahl an Teilnehmern gibt, die die Maßnahme erhalten haben. Ein solcher Fall liegt etwa vor, wenn öffentliche Einrichtungen (Ministerien, Behörden, öffentlich finanzierte Wissenschaftseinrichtungen u. a.) eingerichtet werden. Die Ansiedlung betrifft dann häufig nur eine einzelne Gemeinde oder Region. Für solche Konstellationen gibt es in der Literatur den Ansatz synthetischer Kontrollgruppen. Der eigentlichen Prozedur ist häufig eine Vorauswahl der Beobachtungseinheiten vorgeschaltet. Diese sollten dem Teilnehmer in der Maßnahme in groben Merkmalen ähneln. Die synthetische Kontrollgruppe besteht dann aus einem gewichteten Durchschnitt der Teilnehmer dieser „Spendergruppe“, wobei die Gewichte derart bestimmt werden, dass diese synthetische Beobachtungseinheit den gleichen Trend in der Ergebnisgröße vor der Maßnahme aufweist wie der Teilnehmer der Behandlungsgruppe. Auf diese Weise etwa haben Becker, Heblich und Sturm (2020) den Effekt des Aufbaus von Bonn als Bundeshauptstadt in der Zeit der deutschen Teilung auf die Beschäftigung untersucht. Sie finden heraus, dass Bonn im Vergleich zu seiner synthetischen Kontrollregion einen substanziellen Anstieg der Beschäftigung im öffentlichen Sektor verzeichnet hat. Da jedoch jeder zusätzliche Arbeitsplatz im öffentlichen Sektor Arbeitsplätze in privaten Sektor verdrängt, fällt der Effekt auf die Gesamtbeschäftigung zwar immer noch positiv aber allenfalls moderat aus.

2.4.5 Zusammenfassung zur Eignung kontrafaktischer Analysetools

Die moderne Wirkungsforschung hält eine ganze Reihe an geeigneten Methoden bereit, um zuverlässige Aussagen zum Zusammenhang zwischen einer wirtschaftspolitischen Maßnahme und dem Erreichen einer Zielgröße treffen zu können. Alle in diesem Abschnitt diskutierten Methoden erfüllen die Qualitätskriterien der internationalen Wirkungsforschung. Die Anwendung von Methoden der Wirkungsforschung gelingt umso eher, je früher die Wissenschaft in den Politikzyklus eingebunden wird. Hier lässt sich noch Einfluss nehmen auf die Interventionsregeln, in die man das Kontrafaktum implementieren könnte, sowie auf die Generierung der administrativen Daten über die Anwendung der Maßnahme.

Dieser Abschnitt macht aber auch deutlich, dass die Anwendung bestimmter Methoden (wie etwa Experimente oder Regressions-Diskontinuitäts-Designs) in der Gesellschaft auf Akzeptanz stoßen muss. Schließlich findet bei diesen Untersuchungsansätzen Diskriminierung statt in dem Sinne, dass nicht alle Teilnehmer, die wollen, am Ende die

Maßnahme auch in Anspruch nehmen können.¹² Es gibt also „Verlierer“ in diesem Prozess. Für die Politik könnte sich hier ein Zielkonflikt zwischen hoher interner Validität und gesellschaftlicher Akzeptanz für bestimmte Untersuchungsmethoden offenbaren.

2.5 AKTEURE/INSTITUTIONEN, MECHANISMEN UND INSTRUMENTE ZUR EVIDENZBASIERTEN POLITIKGESTALTUNG IM AGRARBEREICH IN DEUTSCHLAND

Die Akteure einer evidenzbasierten Politikberatung im Agrarbereich in Deutschland lassen sich in Abhängigkeit ihrer Beziehung zu politischen Entscheidungsträgern in drei Gruppen einteilen.

Wissenschaftliche Expertise, die direkt in den politischen Willensbildungsprozess einfließt, wird üblicherweise in der Phase der Politikformulierung durch Ministerien und Parlamente in Form von Gutachten oder im Rahmen von Anhörungen abgerufen. Eher im Bereich der Politischen Themenfindung ist die Arbeit von wissenschaftlichen Beiräten angesiedelt. So wird das Bundesministerium für Ernährung und Landwirtschaft von mehreren Wissenschaftlichen Beiräten (Agrarpolitik, Ernährung und gesundheitlichen Verbraucherschutz; Biodiversität und genetische Ressourcen; Waldpolitik; Düngungsfragen) und dem Sachverständigenrat Ländliche Entwicklung beraten. In diesen Gremien sind Forschende aus Universitäten und Forschungseinrichtungen Deutschlands, aber auch Sachverständige aus der Praxis vertreten. Die Mitglieder werden vom Ministerium berufen.

Sogenannte Ressortforschungseinrichtungen sind Einrichtungen, die direkt dem Bundesministerium für Ernährung und Landwirtschaft (BMEL) unterstehen (nichtrechtsfähige Anstalten des öffentlichen Rechts und selbständige Bundesoberbehörden) und über dessen den Haushalt finanziert werden. Hierzu zählen das Johann Heinrich von Thünen Institut (TI), das Friedrich-Loeffler-Institut (FLI), das Max Rubner-Institut und das Julius Kühn-Institut (JKI). Zusätzlich zählen die beiden rechtlich selbständigen Einrichtungen, das Bundesinstitut für Risikobewertung (BfR) und das Deutsche Biomasseforschungszentrum (DBFZ), zur Ressortforschung. Diese Einrichtungen bilden durch ihre Anbindung an Wissenschaft und Politik eine Brückenfunktion. Neben der Forschungsaufgabe und der Entscheidungsunterstützung werden auch gesetzlich zugewiesene Aufgaben übernommen. Die wissenschaftliche Expertise soll zum Teil kurzfristig abrufbar sein und trägt damit zur Phase der Politikformulierung bei. Im Zuge des zunehmenden Anspruchs an Evaluierungen beteiligt sich bspw. auch das TI an Ausschreibungen für die Erstellung entsprechender Studien durch Ministerien des Bundes und der Bundesländer. Neben strukturellen Veränderungen, u. a. eine starke Bündelung der vielen Fachinstitute in den heute bestehenden Strukturen, sind die Einrichtungen gefordert, „ihre Forschungstätigkeit stärker an den Grundsätzen

¹² Denkbar ist auch der umgekehrte Fall, dass unter den Teilnehmern, die nicht wollen, einige in die Maßnahme müssten.

des Wissenschaftssystems zu orientieren“ (Wissenschaftsrat 2017; S. 6). So wurde durch das BMEL 2008 eine wissenschaftliche Qualitätssicherung, z. B. mittels regelmäßiger Evaluationen und die Etablierung von wissenschaftlichen Beiräten, für die Einrichtungen eingeführt. Beispielsweise wurde das TI durch den Wissenschaftsrat 2014/15 evaluiert (vgl. Wissenschaftsrat 2017).

Forscherinnen und Forscher an wissenschaftlichen Einrichtungen sind im Vergleich zu den vorher genannten Akteuren am unabhängigsten. Hierzu zählen auch Einrichtungen der Leibniz-Gemeinschaft, in deren Aufsichtsgremien das BMEL vertreten ist, die aber in ihrer strategischen Forschungsplanung unabhängiger als die Ressortforschungseinrichtungen agieren können. Die Wahl von Themen und Methoden richtet sich im Allgemeinen an eher wissenschaftsinternen Standards und Diskursen aus. Damit sind die Arbeiten eher in der Phase der Politischen Themenfindung oder der Evaluierung angesiedelt. Das BMEL kann über die externe Vergabe von Forschungsaufträge diese Expertise für verschiedene Phasen des Politikzyklus nutzen.

2.6 PRAXISBEISPIELE FÜR EVIDENZBASIERTE POLITIKGESTALTUNG UND KRITISCHE DISKUSSION DES ZUSAMMENWIRKENS VON AGRARPOLITIK UND -WISSENSCHAFT IN DEUTSCHLAND

Die folgenden beiden Beispiele zeigen, wie sich evidenzbasierte Politikberatung institutionalisieren lässt. Hier geht es vor allem darum, diese Initiativen auf Dauer anzulegen, um evidenzbasierte Expertise kontinuierlich und ohne großen Zeitverzug für die Politik und die breite Öffentlichkeit im Sinne einer Bibliothek vorzuhalten.

2.6.1 Praxisbeispiele für evidenzbasierte Politikgestaltung in Deutschland

Politikgestaltung, die auf systematischer evidenzbasierter Evaluationsforschung basiert, ist möglich. In Großbritannien, den Niederlanden und den Vereinigten Staaten gibt es eine ganze Reihe von Beispielen, wie breit angelegte **Evaluationsforschung institutionalisiert** werden kann und so zu einer Verbesserung von Politikberatung beiträgt (Buch, Patzwaldt, Riphahn und Vogel 2019).

Neben den Beispielen aus dem internationalen Raum, gibt es auch in Deutschland Initiativen, evidenzbasierte Evaluationen in der Forschungslandschaft stärker zu institutionalisieren. Eine dieser Initiativen ist das Zentrum für evidenzbasierte Politikberatung am Leibniz-Institut für Wirtschaftsforschung Halle (IWH-CEP, www.iwh-halle.de/iwh-cep). Das IWH-CEP wurde im Jahr 2014 als Infrastruktureinheit gegründet und ist organisatorisch dem Bereich des Präsidenten des Leibniz-Instituts für Wirtschaftsforschung Halle (IWH, www.iwh-halle.de) zugeordnet. Das IWH gehört zu den führenden wirtschaftswissenschaftlichen Forschungsinstituten in Deutschland und ist Mitglied der Leibniz-Gemeinschaft (www.leibniz-gemeinschaft.de). Der institutionelle Haushalt der Leibniz-Institute setzt sich je zur Hälfte aus Mitteln des Bundes und der Länder zusammen.

Die Arbeit im IWH-CEP orientiert sich an dem in Abschnitt 0 vorgestellten System

evidenzbasierter Evaluationen, welches neben der Wirkungsforschung im engeren Sinne auch umfangreiche Datenarbeit sowie bilateralen Wissenstransfer mit der Politik und der breiten Öffentlichkeit einschließt. Das IWH-CEP koordiniert alle Arbeiten im IWH, die einen Bezug zu Evaluationen wirtschaftspolitischer Interventionen aufweisen. Entsprechend des Institutsprofils handelt es sich hier vor allem um Evaluationen auf den Gebieten der Subventionierung von Unternehmen und Regionen sowie von Regulierungsmaßnahmen für Finanzmärkte. Die inhaltliche Klammer für die beiden Gebiete ergibt sich aus dem Umstand, dass viele Förderprogramme für Unternehmen mit Restriktionen an den Finanzmärkten gerechtfertigt werden. Die Politik argumentiert, dass Marktmängel eine ausreichende Versorgung der Unternehmen (v. a. in strukturschwachen Regionen) mit Finanzmitteln verhindern. Diese Mängel solle der Staat über seine Förderpolitik beheben. Ungeklärt ist bislang, ob diese Marktmängel tatsächlich vorliegen und ob die Förderprogramme die an sie gestellten Ziele bewirken.

Unmittelbar nach seiner Gründung hat das IWH-CEP damit begonnen, Studien zu den Wirkungen von Regionalpolitik zu verfassen (IWH 2018; Brachert, Dettmann und Titze 2018). Hierzu ist das IWH-CEP in einen intensiven Dialog mit der Regierung des Landes Sachsen-Anhalt getreten. Die Landesregierung hat dem IWH-CEP umfangreichen Daten der administrativen Förderstatistik zur Verfügung gestellt. In diesen Pilotstudien wurden darüber hinaus Analysetools für ein Statistikprogramm entwickelt, um systematisch Wirkungsanalysen durchzuführen, die dem internationalen Stand der Evaluationsforschung entsprechen (Dettmann, Giebler und Weyh 2019). Ein weiteres Ergebnis der Arbeit an den Pilotstudien ist der Aufbau einer Subventionsdatenbank (Brachert, Giebler, Heimpold, Titze und Urban-Thielicke 2018). Schließlich erhalten die Ideen zum System evidenzbasierter Evaluierungen Eingang in forschungsbezogene Lehre an der Martin-Luther-Universität Halle-Wittenberg (MLU). Hervorzuheben ist in diesem Zusammenhang das Seminar Evidence-based Policy Advice, in welchem das IWH-CEP die Studierenden mit den Herausforderungen bei der Erstellung von Evaluationsstudien in der praktischen Arbeit vertraut macht. Das IWH-CEP nahm darüber hinaus Einladungen für die Durchführung von Weiterbildungsaktivitäten im Bundesministerium für Wirtschaft und Energie sowie vom Finanzministerium des Landes Mecklenburg-Vorpommern wahr.

Aufgrund dieser umfangreichen Vorarbeiten konnte sich das IWH-CEP im Jahr 2018 erfolgreich auf die Durchführung der Evaluation der Regionalpolitik in Deutschland bewerben. Das Bundesministerium für Wirtschaft und Energie hatte diesen Dienstleistungsauftrag in einem wettbewerblichen Verfahren vergeben. Das Vorhaben wurde im Jahr 2020 abgeschlossen (Brachert, Brautzsch, Dettmann, Giebler, Schneider und Titze 2020). Die Untersuchungen haben gezeigt, dass die geförderten Betriebe eine bessere Beschäftigtenentwicklung aufwiesen als die Betriebe der Kontrollgruppe. Darüber hinaus hat die deutsche Regionalförderung einen positiven Effekt auf die Entwicklung des Umsatzes in den geförderten Betrieben – ein Effekt auf die Entwicklung der Produktivität war allerdings nicht nachweisbar.

Die wissenschaftliche Prüfung der EU Kommission bestätigte die korrekte Durchführung

der Evaluation. Der Bericht wurde auf der Homepage des Bundesministeriums für Wirtschaft und Energie (BMWi) sowie des Leibniz-Instituts für Wirtschaftsforschung Halle (IWH) bereitgestellt.¹³ Die breite Öffentlichkeit wurde über das Gutachten über Pressemitteilungen des BMWi sowie des IWH informiert.¹⁴ Im Jahr 2022 wird es eine Aktualisierung der Evaluationsstudie geben.

Die umfangreichen administrativen Förderdaten sowie die Kenntnis der institutionellen Regelungen über die Anwendung von Regionalpolitik gaben den Startschuss für Forschungsprojekte mit renommierten Kooperationspartnern auf nationaler und internationaler Ebene.

2.6.2 Diskussion des Zusammenwirkens von Agrarpolitik und -wissenschaft in Deutschland

Seit Bestehen der Gemeinsamen Agrarpolitik (GAP) findet auch seitens der Wissenschaft eine Auseinandersetzung mit der Effektivität und Effizienz des eingesetzten Instrumentariums statt (z. B. Tangermann 1974; Koester und Tangermann 1976). Die praktische Agrarpolitik weist eine hohe Persistenz auf und wissenschaftliche Erkenntnisse werden nur sehr zögerlich bzw. nur teilweise aufgegriffen. Seit der letzten GAP-Reform 2013 postuliert die EU-Kommission neun Ziele, die sowohl ökonomische als auch Naturschutzaspekte abdecken. Naturgemäß konzentrieren sich wissenschaftliche Untersuchungen zur Effektivität der GAP vor allem auf diese Ziele, können hier aber nicht umfassend dargestellt werden.

Exemplarisch sei auf die Arbeiten zu den Arbeitsmarkteffekten der GAP-Instrumente verwiesen. In diesem Kontext sind sowohl die Auswirkungen der Direktzahlungen als auch der Zahlungen in der ländlichen Entwicklungspolitik auf die Beschäftigung in der Landwirtschaft mit verschiedenen Methoden untersucht worden. Die empirischen Ergebnisse unterstützen keine abschließende Bewertung, da die Maßnahmen teilweise gegenläufige Effekte verursachen und die Messung der landwirtschaftlichen Beschäftigung Erfassungsproblemen ausgesetzt ist (vgl. Zier 2013; Olper, Raimondi, Cavicchioli und Vigani 2014).

Ein weiteres Beispiel ist die Evaluierung der landwirtschaftlichen Investitionsförderung im Rahmen der ländlichen Entwicklungspolitik. Michalek (2012) präsentiert einen detaillierten Überblick über die bestehenden Studien vor allem aus einer methodischen Perspektive. Zusätzlich illustriert er unter Verwendung eines Propensity Score Matching-

¹³Vgl. <https://www.bmwi.de/Redaktion/DE/Downloads/E/evaluierung-der-gemeinschaftsaufgabe-verbesserung-der-regionalen-wirtschaftsstruktur-durch-einzelbetriebliche-erfolgskontrolle-enderbericht.pdf> sowie <https://www.iwh-halle.de/publikationen/detail/evaluation-der-gemeinschaftsaufgabe-verbesserung-der-regionalen-wirtschaftsstruktur-grw-durch/>.

¹⁴Vgl. <https://www.bmwi.de/Redaktion/DE/Pressemitteilungen/2020/20200702-regionale-wirtschaftsfoerderung-schafft-zusaetzliche-arbeitsplaetze-und-erhoeht-einkommen-in-den-regionen.html> sowie <https://www.iwh-halle.de/nc/presse/pressemitteilungen/detail/iwh-gutachten-regionalfoerderung-wirkt-sollte-aber-weiter-verbessert-werden/>.

Ansatzes in Kombination mit einem Differenz-von-Differenzen-Ansatz am Beispiel der Investitionsförderung in Schleswig-Holstein zwar eine positive Wirkung auf das Betriebsergebnis der geförderten Landwirte aber gleichzeitig die sehr hohe Relevanz von Mitnahmeeffekten. Viele Betriebsinhaber würden die Investitionen auch ohne Förderung durch das Agrarinvestitionsprogramm durchgeführt haben.

Die in den vorgestellten Beispielen untersuchten Politikmaßnahmen bestehen aber weiterhin. Eine direkte Politikänderung auf Basis wissenschaftlicher Erkenntnisse ist im Bereich der Agrarpolitik, nicht zuletzt aufgrund der langwierigen Reformprozesse, nicht zu erwarten. Dennoch ist unbestreitbar, dass sich eine evidenzbasierte Entscheidungsfindung etabliert hat. Schließlich stellt auch die EU Kommission selber fest, dass das Common Monitoring and Evaluation Framework (CMEF) zu viele Indikatoren enthält (derzeit 146) und einige der Indikatoren nicht zweckdienlich sind (EU KOM 2018). Mit einem auch weiterhin zunehmenden Druck zur Rechtfertigung der GAP muss das Evaluationsinstrumentarium weiterentwickelt werden, um besser Auskünfte über die Auswirkungen der GAP geben zu können, aber gleichzeitig auch für die damit betrauten Akteure umsetzbar zu sein.

3 ABLEITUNG VON OPTIONEN FÜR EINE EVIDENZBASIERTE AGRARPOLITISCHE ENTSCHEIDUNGSFINDUNG IN DER UKRAINE

Vor dem Hintergrund der Erfahrungen mit evidenzbasierter Politikberatung auf internationaler sowie auf deutscher Ebene lassen sich folgende Empfehlungen für die Ukraine ableiten.

Erstens bedarf es **guter institutioneller Arrangements**. Alle **Interventionen** sollten **zeitlich begrenzt** sein. Damit lässt sich verhindern, dass sich eine Maßnahme automatisch verlängert und es womöglich nie einen Ausstieg aus den Regelungen – auch wenn sie sich in der Praxis als völlig wirkungslos oder sogar gefährlich erweisen – gibt. Die **Überprüfung der Wirkung staatlicher Maßnahmen** sollte **gesetzlich verankert** werden – idealerweise mit Verfassungsrang. Die Gesetze sollten nicht nur einfordern, dass die Politik die Wirkungen ihrer Maßnahmen überprüft, sondern auch Eckpunkte zum „Wie?“ der Evaluation. Diese sollten sich an den in **der internationalen Wirkungsforschung** geltenden **Qualitätsstandards** orientieren. Schließlich ist es erforderlich, dass die Kontrolle der Wirkungen der Maßnahmen von **unabhängigen Gutachtern** durchgeführt wird und die **Daten, Methoden sowie Ergebnisse** für die breite **Öffentlichkeit transparent** sind. Evaluationen benötigen gute Daten. Hierfür ist der Aufbau eines **zuverlässigen Systems der amtlichen Statistik** erforderlich.

Internationale Organisationen haben der Ukraine bestätigt, dass ihre Statistik den internationalen Standards, beispielsweise denen der Europäischen Union, schon sehr weitgehend entsprechen. Herausforderungen bestehen aktuell darin, die hohen Statistikanforderungen, wie sie etwa in EU-Richtlinien bestehen, in der Ukraine vollständig

zu implementieren. Problematisch sei, dass die Anzahl erhobener Kennziffern verringert wurde und die statistischen Ämter in der Ukraine ihren Personalbestand reduzieren. Auch gebe es großen Nachholbedarf bei der technischen Ausstattung der Datenerfassung und -verarbeitung. Für die Abstellung dieser Defizite bedürfe es der Entwicklung und Verabschiedung eines Aktionsplans für die Ukraine, der eine breite Akzeptanz in der Gesellschaft haben sollte (Chodakiwska und Besduschna 2021).

Für jedes Statistiksyste**m** besteht zweifelsohne ein Zielkonflikt zwischen der Genauigkeit sowie dem Umfang der zu erhebenden Daten und den mit der Erhebung verbundenen Kosten (Belastungen auf Seiten der Befragenden und Befragten, Aspekte des Datenschutzes usw.). Wenn sich einige Bereiche des Wirtschaftslebens bzw. der Umfang für die Bewertung bestimmter Maßnahmen zu erhebender Daten als zu groß herausstellen, um sie zu angemessenen Kosten über Vollerhebungen abzubilden, könnte es sinnvoll sein, repräsentative Stichproben zu erheben. Ein Vorbild könnten hier etwa die Strukturen im Rahmen des FADN sein. Vor dem Hintergrund einer nur bedingten statistischen Erfassung der Produktion und wirtschaftlichen Situation kleinere landwirtschaftlicher Hauswirtschaften, spielt die Repräsentativität der Agrarstruktur der Ukraine bei der Stichprobenentwicklung eine besonders bedeutsame Rolle.

Zweifelsohne besteht ein Zielkonflikt zwischen der Genauigkeit sowie dem Umfang der zu erhebenden Daten und den mit der Erhebung verbundenen Kosten (Belastungen auf Seiten der Befragenden und Befragten, Aspekte des Datenschutzes usw.). Wenn sich einige Bereiche des Wirtschaftslebens bzw. der Umfang für die Bewertung bestimmter Maßnahmen zu erhebender Daten als zu groß herausstellen, um sie zu angemessenen Kosten über Vollerhebungen abzubilden, könnte es sinnvoll sein, repräsentative Stichproben zu erheben. Ein Vorbild könnten hier etwa die Strukturen im Rahmen des FADN sein.

Zweitens ist es sehr wichtig, dass die **Politik** die **Wissenschaft** in den einzelnen **Phasen des Politikzyklus** mit **einbindet**. Für Evaluationen ist es sehr hilfreich, wenn die einzelnen Phasen des Politikzyklus **sauber dokumentiert** sind. Die Dokumentationspflicht könnte im Übrigen auch gesetzlich verankert werden. Durch die strukturierten Prozesse lassen sich grobe Fehler bereits zu Beginn vermeiden. So ist es beispielsweise sehr förderlich, wenn die Politik die **Ziele klar definiert** und mit objektiv messbaren Indikatoren unterlegt. **Systematische Problemanalysen** können einen Beitrag dazu leisten, früh **Zielkonflikte** sowie **unerwünschte Nebenwirkungen** zu identifizieren. Wenn man bereits vor Einführung der Maßnahme das **Untersuchungsdesign** für die Wirkungsanalyse und die damit einhergehenden **Datenerfordernisse mitdenkt**, lassen sich Evaluationen sehr zuverlässig und kostengünstig durchführen. Alle diese Punkte stellen **Transparenz** her und dürften die Akzeptanz für die Maßnahmen und eventuell damit verbundene Nebenwirkungen (bspw. Notwendigkeit der Erhebung von Steuern) steigern.

Drittens ist zu empfehlen, die **Evaluationen** nicht nur über die Wirkungsforschung im engeren Sinne zu verstehen, sondern als **System ineinandergreifender Aufgaben**

zwischen **Methodenkenntnis, professionellem Datenmanagement** sowie umfangreichen **Transferaktivitäten** zwischen Wissenschaft, Politik und breiter Öffentlichkeit. Ein wesentlicher Aspekt hierbei ist, dass die Mitarbeiter in den Behörden und die breite Öffentlichkeit mit den Grundlagen moderner Wirkungsforschung vertraut sind. Defizite lassen sich abbauen über umfangreiche Maßnahmen der **Aus- und Weiterbildung**. Hierbei kann sicherlich ein **Austausch** mit den **Partnern auf internationaler Ebene**, die sich im Feld evidenzbasierter Politikberatung etabliert haben, einen wertvollen Beitrag leisten zur Verbesserung des Know-hows in der Ukraine.

LITERATUR

- Acemoglu, D.; Robinson, J. (2010): The Role of Institutions in Growth and Development. *Review of Economics and Institutions* 1(2), 2-33.
- Alm, B., Fisch, G. (2014): Aufgaben, Instrumente und Perspektiven der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“. In: Eberstein, H.-H.; Karl, H.; Untiedt, G. (Hrsg.): *Handbuch der regionalen Wirtschaftsförderung (Grundwerk mit Fortsetzungsbezug für mindestens 2 Jahre)*, Teil C, Abschnitt III.
- Bade, F.-J.; Alm, B. (2010): Endbericht zum Gutachten Evaluierung der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRW) durch einzelbetriebliche Erfolgskontrolle für den Förderzeitraum 1999–2008 und Schaffung eines Systems für ein gleitendes Monitoring. Dortmund.
- Beirat BMWi, Wissenschaftlicher Beirat beim Bundesministerium für Wirtschaft und Energie (2013): *Evaluierung wirtschaftspolitischer Fördermaßnahmen als Element einer evidenzbasierten Wirtschaftspolitik. Gutachten*. Berlin.
- Beirat BMWi, Wissenschaftlicher Beirat beim Bundesministerium für Wirtschaft und Energie (2015): *Regionale Wirtschaftsförderung. Gutachten*. Berlin.
- Behaghel, L.; Macours, K.; Subervie, J. (2019): How can randomised controlled trials help improve the design of the Common Agricultural Policy? *European Review of Agricultural Economics* 46(3), 473-493.
- Bernini, C.; Pellegrini, G. (2011): How are growth and productivity in private firms affected by public subsidy? Evidence from a regional policy, in: *Regional Science and Urban Economics* 41(3), 253-265.
- BMEL, Bundesministerium für Landwirtschaft und Ernährung (2018): *Buchführung der Testbetriebe: Grundlagen zur BMEL-Testbetriebsbuchführung*. Bundesministerium für Ernährung und Landwirtschaft. Berlin.
- Bollen, K.; Cacioppo, J. T.; Kaplan, R. M.; Krosnick, J. A.; Olds, J. L. (2015): *Social, Behavioral, and Economic Sciences Perspectives on Robust and Reliable Science. Report of the Subcommittee on Replicability in Science. Advisory Committee to the National Science Foundation Directorate for Social, Behavioral, and Economic Sciences*.
- Bondonio, D.; Greenbaum, R. T. (2014): Revitalizing regional economies through enterprise support policies: an impact evaluation of multiple instruments. *European Urban and Regional Studies* 21(1), 79–103. <https://doi.org/10.1177/0969776411432986>.
- Boockmann, B.; Buch, C. M.; Schnitzer, M. (2014): Evidenzbasierte Wirtschaftspolitik in Deutschland: Defizite und Potentiale, in: *Perspektiven der Wirtschaftspolitik*, 15(4), 307-323.
- Brachert, M.; Brautzsch, H.-U.; Dettmann, E.; Giebler, A.; Schneider, L.; Titze, M. (2020): „Evaluation der Gemeinschaftsaufgabe ‚Verbesserung der regionalen Wirtschaftsstruktur‘ (GRW)“ durch einzelbetriebliche Erfolgskontrolle: Endbericht, *IWH Online* 5/2020, Halle Institute for Economic Research (IWH).
- Brachert, M.; Dettmann, E.; Titze, M. (2015): Die Analyse kausaler Effekte wirtschaftspolitischer Maßnahmen – Das Zentrum für evidenzbasierte Politikberatung am IWH (IWH-CEP). *IWH, Wirtschaft im Wandel* 21(5), 84-87.
- Brachert, M.; Dettmann, E.; Titze, M. (2020): Zu den betrieblichen Effekten der Investitionsförderung im Rahmen der deutschen Regionalpolitik. *IWH, Wirtschaft im Wandel*

26(1), 5-8.

Brachert, M.; Giebler, A.; Heimpold, G.; Titze, M.; Urban-Thielicke, D. (2018): IWH-Subventionsdatenbank: Mikrodaten zu Programmen direkter Unternehmenssubventionen in Deutschland. Datendokumentation. IWH Technical Reports 02/2018. Halle (Saale).

Bronzini, R.; Iachini, E. (2014): Are Incentives for R&D Effective? Evidence from a Regression Discontinuity Approach. In: *American Economic Journal: Economic Policy* 6(4), 100-134.

Buch, C. M.; Dettmann, E.; Heimpold, G.; Titze, M. (2018): Möglichkeiten einer wissenschaftlichen Evaluation der gewerblichen Förderung im Rahmen der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRW) in Sachsen-Anhalt. In: Leibniz-Institut für Wirtschaftsforschung Halle (IWH) (Hrsg.): *Expertisen zur Evaluation der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRW) in Sachsen-Anhalt*. IWH Online 2/2018. Halle (Saale).

Buch, C. M.; Riphahn, R. (2019): Evidenzbasierte Politik: Infrastrukturen für bessere Informationen. In: Buch, C. M.; Riphahn, R. (Hrsg.): *Evaluierung von Finanzmarktrefor- men - Lehren aus den Politikfeldern Arbeitsmarkt, Gesundheit und Familie*. Leopoldina-Forum Nr. 1. Nationale Akademie der Wissenschaften Leopoldina, Halle (Saale), S. 7-13.

Buch, C., K. Patzwaldt, R. Riphahn und E. Vogel (2019). Verstehen – Entwickeln – Testen – Verbessern: Rahmenbedingungen für evidenzbasierte Politik, *Wirtschaftsdienst* 99, <https://www.wirtschaftsdienst.eu/inhalt/jahr/2019/heft/2/beitrag/verstehen-entwickeln-testen-verbessern-rahmenbedingungen-fuer-evidenzbasierte-politik.html>.

Bundesanzeiger (2018): Koordinierungsrahmen für die Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“. Amtlicher Teil 05.10.2018 B2.

Cerqua, A.; Pellegrini, G. (2017): Industrial policy evaluation in the presence of spillovers. *Small Business Economics*, 49(3), 671–686. <https://doi.org/10.1007/s11187-017-9855-9>.

Chodakiwska, O.; Besduschna, J. (2021): Evidenzbasierte Politikgestaltung – derzeitige Lage, Verfahren und Probleme in der ukrainischen Agrarpolitik. *Agrarpolitischer Bericht APD/APB/09/2020*, Kiew.

Colen, L.; Gomez y Paloma, S.; Latacz-Lohmann, U.; Lefebvre, M.; Préget, R.; Thoyer, S. (2016): Economic Experiments as a Tool for Agricultural Policy Evaluation: Insights from the European CAP. *Canadian Journal of Agricultural Economics* 64, 667-694.

Crépon, B.; Duflo, E.; Gurgand, M.; Rathelot, R.; Zamora, P. (2013): Do labor market policies have displacement effects? Evidence from a clustered randomized experiment. *Quarterly Journal of Economics* 128(2), 531-580.

Criscuolo, C.; Martin, R.; Overman, H.; van Reenen, J. (2019): Some Causal Effects of an Industrial Policy. *American Economic Review* 109, 48-85.

Dettmann, E.; Weyh, A.; Titze, M. (2017): Who Benefits from GRW? Heterogeneous Employment Effects of Investment Subsidies in Saxony Anhalt. *IWH Discussion Papers* 27/2017.

Dettmann, E.; Weyh, A.; Titze, M. (2018): Heterogeneous effects of investment grants - Evidence from a new measurement approach. *IAB-Discussion Paper*, 15/2018, Nürnberg.

Dettmann, E.; Giebler, A.; Weyh, A. (2020): flexpaneldid: A Stata toolbox for causal analysis with varying treatment time and duration, *IWH Discussion Papers*, No. 3/2020, Leibniz-Institut für Wirtschaftsforschung Halle (IWH), Halle (Saale).

Dörner, D.; Schaub, H. (1994): Errors in Planning and Decision-making and the Nature of Human

Information Processing. *Applied Psychology: An International Review* 43(4), 433-453.

Dörner, D. (1990): The logic of failure. In: *Philosophical Transactions of the Royal Society of London. B Biological Sciences* 327(1241), 463-473.

Eckey, H.-F. (2008): Verbesserung der regionalen Wirtschaftsstruktur. Gesetz über die Gemeinschaftsaufgabe vom 6. Oktober 1969. In: Eberstein, H.-H.; Karl, H.; Untiedt, G. (Hrsg.): *Handbuch der regionalen Wirtschaftsförderung (Grundwerk mit Fortsetzungsbezug für mindestens 2 Jahre)*, Teil A, Abschnitt V.

Edovald, T.; Firpo, T. (2016): *Running randomised controlled trials in innovation, entrepreneurship and growth: an introductory guide*.

EFI, Expertenkommission Forschung und Innovation (2017): *Gutachten 2017. Gutachten zu Forschung, Innovation und Technologischer Leistungsfähigkeit Deutschlands*. Berlin. Online: https://www.e-fi.de/fileadmin/Gutachten_2017/EFI_Gutachten_2017.pdf. Zugriff: 18.09.2018.

EU KOM (2014a): Commission Regulation (EU) No 651/2014 of 17 June 2014 declaring certain categories of aid compatible with the common market in application of Articles 107 and 108 of the Treaty, OJ L 187 of 26.6.2014, p. 1.

EU KOM (2014b): European Commission. Commission Staff Working Document. Common methodology for State aid evaluation. SWD(2014) 179 final. Brussels, 28.05.2014. Online: http://ec.europa.eu/competition/state_aid/modernisation/state_aid_evaluation_methodology_en.pdf. Access 05.02.2021.

EU KOM (2018): European Commission. Report from the Commission to the European Parliament and the Council on the implementation of the Common Monitoring and Evaluation Framework and first results on the performance of the Common Agricultural policy. COM(2018) 790 final. Brussels, 05.12.2018. Online: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0790&from=en>. Access 05.03.2021.

Expertenkommission „Stärkung von Investitionen in Deutschland“ (Hrsg.) (2015): *Stärkung von Investitionen in Deutschland. Bericht der Expertenkommission im Auftrag des Bundesministers für Wirtschaft und Energie, Sigmar Gabriel*. Online: <https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/staerkung-von-investitionen-in-deutschland.pdf>. Access: 15.11.2018.

Forrester, J. W. (1961): *Industrial Dynamics*. The MIT Press Cambridge (Massachusetts).

Fritsch, M. (2018): *Marktversagen und Wirtschaftspolitik. Mikroökonomische Grundlagen staatlichen Handelns*. 10., überarbeitete und ergänzte Auflage. München.

Heckman, J. J.; LaLonde, R. J.; Smith, J. A. (1999): The Economics and Econometrics of Active Labor Market Programs, in O. Ashenfelter and D. E. Card (Hrsg.): *Handbook of Labor Economics*, Vol. III, Elsevier Science B.V., Amsterdam, S. 1865–2097.

Hill, B.; Bradley, D.; Vrolijk, H. (2016): Uses and Benefits of FADN Information in the EU-28. *EuroChoices* 15(3), 11-16.

Howlett, M.; Ramesh, M.; Perl, A. (2020): *Studying Public Policy. Principles and Processes*. 4th edition. Oxford University Press, Toronto.

IWH, Leibniz-Institut für Wirtschaftsforschung Halle (2018): *Expertisen zur Evaluation der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRW) in Sachsen-Anhalt*. IWH Online 2/2018. Halle (Saale).

Khandker, S.; Koolwal, G. B.; Samad, H. (2010): *Handbook on impact evaluation: quantitative*

methods and practices. The World Bank. Washington D. C.

Kilper, H.; Rosenfeld, M. T. W. (2007): Gegenstand, Wirkungsweise und Begründung der interregionalen Ausgleichspolitik. In: Rosenfeld, M. T. W.; Alecke, B.; Franz, P.; Heimpold, G.; Kilper, H.; Kunkel, K.; Untiedt, G.; Zillmer, S. (Hrsg.): Interregionale Ausgleichspolitik in Deutschland: Untersuchungen zu den Effekten ausgewählter Systeme zur Herstellung von „gleichwertigen Lebensverhältnissen“. IWH Sonderheft 2/2007, Halle (Saale), S. 32-40.

Koester, U.; Tangermann, S. (1976): Alternativen der Agrarpolitik. Landwirtschaftsverlag, Münster-Hiltrup.

Kölling, A. (2015): Does public funding work? A causal analysis of the effects of economic promotion with establishment panel data. *Kyklos* 68(3), 385–411.

Kugler, F.; Schwerdt, G.; Wößmann, L. (2014): Ökonometrische Methoden zur Evaluierung kausaler Effekte der Wirtschaftspolitik, in: *Perspektiven der Wirtschaftspolitik* 15(2), 105-132.

Madaleno, M.; Waights, S. (2016): Guide to scoring methods using the Maryland Scientific Methods Scale. Updated June 2016. Online: http://www.whatworksgrowth.org/public/files/Methodology/16-06-28_Scoring_Guide.pdf, accessed: March 31, 2017.

Michalek, J. (2012): Counterfactual Impact Evaluation of EU Rural Development Programmes: Propensity Score Matching Methodology Applied to Selected EU Member States. JRC Scientific and Policy Reports No. 25421. Joint Research Centre/ Institute for Prospective Studies. Luxembourg, European Commission. Volume 1: A Micro-level Approach.

Neumark, D., Simpson, H. (2015): Place-based policies. In: Duranton, G., Henderson, J.V., Strange, W. (Hrsg.): *Handbook of Regional and Urban Economics*, vol. 5B, S. 1198–1287.

NKR, Normenkontrollrat (2017): Bürokratieabbau. Bessere Rechtsetzung. Digitalisierung. Erfolge ausbauen – Rückstand aufholen. Jahresbericht 2017 des Nationalen Normenkontrollrates. Berlin.

NKR, Normenkontrollrat (2015): Chancen für Kostenbegrenzung verbessert. Digitale Chancen tatsächlich nutzen! Jahresbericht 2015 des Nationalen Normenkontrollrates. Berlin.

NKRG (2006): Gesetz zur Einsetzung eines Nationalen Normenkontrollrates vom 14. August 2006 (BGBl. I S. 1866), das zuletzt durch Artikel 8 der Verordnung vom 19. Juni 2020 (BGBl. I S. 1328) geändert worden ist.

North, D. C. (1990): *Institutions, Institutional Change, and Economic Performance*. Cambridge, Cambridge University Press.

North, D. C. (1955): Location Theory and Regional Economic Growth. *Journal of Political Economy* 63(3), 243-258.

Olper, A.; Raimondi, V.; Cavicchioli, D.; Vigani, M. (2014): Do CAP Payments Reduce Farm Labour Migration? A Panel Data Analysis Across EU Regions. *European Review of Agricultural Economics* 41(5), 843-873.

Pellegrini, G.; Centra, M. (2006): Growth and efficiency in subsidized firms, in 'Workshop 'The Evaluation of Labour Market, Welfare and Firms Incentive Programmes'', Instituto Veneto di Scienze, Lettere ed Arti, Venezia.

Poppe, K. J.; Vrolijk, H. (2016): Farm Sustainability Data for Better Policy Evaluation with FADN. FLINT Project. Wageningen.

Schmidt, C. M. (2014): Wirkungstreffer erzielen — Die Rolle der evidenzbasierten Politikberatung in einer aufgeklärten Gesellschaft, in: *Perspektiven der Wirtschaftspolitik* 15(3), 219-233.

- Schneider, F.; Enste, D. H. (2000): Shadow Economies: Size, Causes, and Consequences. *Journal of Economic Literature* 38(1), 77-114.
- Sherman, L. W.; Gottfredson, D. C.; MacKenzie, D. L.; Eck, J.; Reuter, P.; Bushway, S. D. (1998): Preventing Crime: What Works, What Doesn't, What's Promising. In: National Institute of Justice. *Research in Brief*.
- Storey, D. (2017): Six steps to heaven: evaluating the impact of public policies to support small businesses in developed economies. In: Sexton, D.; Landström, H. (Hrsg.): *The Blackwell handbook of entrepreneurship*, S. 176–193. <https://doi.org/10.1002/9781405164214.ch9>.
- Tangermann, S. (1974): *Gemeinsame Agrarpreispolitik und nationale Wirtschaftspolitik in der EWG*. Tübingen, J.C.B. Mohr (Paul Siebeck).
- Vester, F. (1999): *Neuland des Denkens. Vom technokratischen zum kybernetischen Zeitalter*. 11. Auflage. Deutscher Taschenbuchverlag München. Erstaussage 1980.
- Vester, F. (2019): *Die Kunst vernetzt zu denken. Ideen und Werkzeuge für einen neuen Umgang mit Komplexität. Ein Bericht an den Club of Rome*. Pantheon Verlag München. Erstaussage 1999 Deutsche Verlagsanstalt München. Englischsprachige Version 2007: *The Art of Interconnected Thinking. Ideas and Tools for a new approach to tackling complexity*. MCB Publishing House München.
- VfS, Verein für Socialpolitik (2015): *Leitlinien und Empfehlungen des Vereins für Socialpolitik für Ex post-Wirkungsanalysen*. Fassung vom 06.09.2015. Download: https://www.socialpolitik.de/docs/VfS-Leitlinien_Ex_post-Wirkungsanalysen.pdf am 18.12.2015.
- Wissenschaftsrat (2017): *Empfehlungen zur Weiterentwicklung der Ressortforschungseinrichtungen des Bundesministeriums für Ernährung und Landwirtschaft (BMEL)*. Drucksache 5916-17. Berlin.
- Zier, P. (2013): *Econometric Impact Assessment of the Common Agricultural Policy in East German Agriculture*. *Studies on the Agricultural and Food Sector in Central and Eastern Europe* Vol. 71. Halle (Saale), Leibniz-Institut für Agrarentwicklung in Mittel- und Osteuropa.