

Leo Wangler

Evaluation von Forschungs-, Entwicklungs- und Innovationsbeihilfen: Zu einer praktischen Umsetzung von Vergleichsgruppenansätzen

Evaluationen können erheblich dazu beitragen, die Transparenz von Entscheidungsprozessen zu erhöhen und vorhandene Risiken in der Entscheidungsfindung zu reduzieren. Vor allem bei der Förderung von Forschungs-, Entwicklungs- und Innovationsprojekten (FuEul-Projekte) greifen politische Akteure und Berater immer häufiger auf systematische empirische Untersuchungen zurück. In der Praxis wird ein solches Vorgehen als „evidenzbasierte Politik“ bezeichnet. Die zentrale Herausforderung von Evaluationen besteht darin, den bestehenden Erfahrungsschatz in Form von Daten mit der richtigen theoretischen Annahme zu kombinieren, um so belastbare empirische Belege abzuleiten. Gelingt eine logische und sinnvolle Verknüpfung zwischen „Empirie“ und „Theorie“, lassen sich – trotz einer oftmals vorhandenen Komplexität der zu evaluierenden Themen – Ergebnisse generieren, die wertvolle Hilfestellungen zur Steuerung oder Begründung politischer Fördermaßnahmen liefern.

Treatment-Effekte mit Kontroll- oder Vergleichsgruppenansätzen messen

Wie aber lassen sich Erfolge und Misserfolge bei der Projektförderung angemessen beurteilen? Oder konkret gefragt: Zu welchem tatsächlichen Effekt hat eine bestimmte förderpolitische Maßnahme geführt? Diese Frage lässt sich aus evaluatorischer Sicht zumeist nur dann exakt beantworten, wenn eine geeignete **Kontroll- oder Vergleichsgruppe** zur Verfügung steht, mittels derer der „wahre Effekt“ einer politischen Maßnahme bzw. eines „Treatment“ herausgefiltert werden kann.

Während der **Kontrollgruppenansatz** in einer engen Verbindung zu randomisierten Experimenten (Losentscheidung über die Zuteilung einer politischen Maßnahme aus einem Pool potenzieller Förderkandidaten) steht, versucht der **Vergleichs-**

gruppenansatz anhand der Spiegelung mit einer möglichst ähnlichen Gruppe den „Treatment-Effekt“ zu erfassen.

Kontrollgruppenansätze werden häufig für wissenschaftliche Experimente – wie sie beispielsweise in der Medizin oft durchgeführt werden – eingesetzt und sind aufgrund der hohen Merkmalsübereinstimmung zwischen Treatmentempfängern und Nicht-Treatmentempfängern als sogenannte „First-Best-Lösung“ zu verstehen. Im Idealfall lässt der Kontrollgruppenansatz allein aus dem Vergleich beider Gruppen eine exakte Zurechnung des Treatment-Effektes zu.

In der Praxis zeigt sich jedoch, dass der evidenzbasierten Politikberatung in vielen Fällen keine statistisch verwertbaren Kontrollgruppen zur Verfügung stehen. Zwar wird zum Beispiel im Bereich der Innovationspolitik umfassend diskutiert, die Mittel randomisiert bzw. per Losentscheidung¹ zu vergeben, doch konnte sich die Politik, zumindest in Deutschland, bisher nicht zu einer praktischen Umsetzung durchringen.

Die Gründe sind vielschichtig. Einer könnte darin bestehen, dass die randomisierte Mittelvergabe mit einem gewissen politischen Kontrollverlust einhergeht. Demgegenüber dürften die Widerstände bei den Förderempfängern eher gering ausfallen. Denn rechnen potenzielle Förderempfänger bereits bei der Antragstellung damit, dass die Mittelvergabe randomisiert umgesetzt wird, so kann dies durchaus auf Akzeptanz stoßen. Dies setzt jedoch voraus, dass die Antragsteller in den nächsten Förderrunden mit einer ähnlichen Prozedur und damit mit einer vergleichbaren Förderwahrscheinlichkeit rechnen können.

Aufgrund der fehlenden praktischen Umsetzung der randomisierten Mittelvergabe ist es weniger offensichtlich, wie sich ein bestimmter Sachverhalt entwickelt hätte, wäre eine bestimmte

¹ Die Entscheidung Förderung ja oder nein obliegt damit nicht mehr dem Ministerium mit seinen Fachreferaten, vielmehr entscheidet das Los über die Verteilung der Gelder.

Maßnahme nicht durchgeführt worden. Für Evaluatoren stellt sich damit die Aufgabe, Lösungen zu entwickeln, die dem Kontrollgruppenansatz möglichst nahe kommen und ein sogenanntes „Second-Best-Ergebnis“ zulassen.² Hier könnten **Vergleichsgruppenansätze** eine Alternative sein, doch spielen dabei auch die entsprechenden Methoden zur Auswertung der Daten eine wichtige Rolle.

Vor- und Nachteile von Vergleichsgruppenansätzen

Durch die fehlende Umsetzung randomisierter Experimente entsteht bei Evaluationen in der Regel die Anforderung, alternativ zum Kontrollgruppenansatz einen **Vergleichsgruppenansatz** zu wählen. Der Vergleichsgruppenansatz zeichnet sich vor allem dadurch aus, dass die Daten, die bei der Evaluation eines spezifischen Programms erhoben werden, mit **externen Daten** abgeglichen werden. Eine wichtige Voraussetzung dabei ist, dass die Gruppe, bei der die externen Daten erhoben werden, den Förderempfängern in möglichst vielen Merkmalsausprägungen entspricht, jedoch in keinem direkten Kontext zu der zu evaluierenden Fördermaßnahme steht.

Naheliegender wäre, gezielt abgelehnte Antragsteller zu befragen, die sich ebenfalls auf das zu evaluierende Förderprogramm beworben haben, aber nicht mit einer Förderung bedacht wurden. Doch auch diese Vorgehensweise dürfte in der Praxis aus mehreren Gründen schwierig sein:

- ▶ Es kann durchaus vorkommen, dass die Projekte aus einer vergleichsweise geringen Anzahl an Skizzen ausgewählt werden.
 - ▶ Oftmals wird ein hoher Anteil aufgrund mangelnder Qualität aussortiert. Dies führt in vielen Fällen dazu, dass die Anzahl jener abgelehnten Projektnehmer, die sich tatsächlich für eine Vergleichsgruppe eignen, relativ gering ist.
 - ▶ Die Motivation der abgelehnten Projektnehmer, an einer Befragung zum Programm teilzunehmen, kann durchaus gering sein, so dass mit einem niedrigen Rücklauf zu rechnen ist.
- ▶ Es ist nicht unwahrscheinlich, dass einige der abgelehnten Projektnehmer ihre Projektskizzen in modifizierter Form erfolgreich in anderen Förderprogrammen eingereicht haben. Dies ist besonders dann zu erwarten, wenn die Forschungsförderung einen besonders hohen Durchdringungsgrad hat. Dies hätte zwangsläufig zur Folge, dass keine „echte“ Kontrollgruppe existiert bzw. die in anderen Programmen geförderten Kandidaten zumindest nicht in der Vergleichsgruppe berücksichtigt werden dürften, da eine anderweitige Förderung die ursächliche Zuordnung der Effekte dem zu evaluierenden Programm erschwert.
- ▶ Als weiterer Aspekt ist bei der Befragung abgelehnter Antragsteller zu berücksichtigen, dass die Anträge aus einem spezifischen Grund zurückgewiesen wurden und sich die Förderempfänger im Vergleich zu den Abgelehnten in einzelnen Charakteristika grundlegend unterscheiden.

Vor dem Hintergrund der genannten Probleme stellt sich die Befragung abgelehnter Antragsteller teilweise als schwierig und nicht praktikabel heraus.

Als Vergleichsgruppe lassen sich jedoch oftmals auch externe Indikatoren heranziehen – sei es mit Bezug zu Indikatoren der amtlichen Statistik oder Indikatoren aus anderen Befragungskontexten. Voraussetzung ist, dass bei der zu evaluierenden Gruppe die gleichen Indikatoren abgefragt werden. Gerade bei der Evaluation von politischen Programmen mit starkem FuEul-Bezug bieten sich zum Beispiel Indikatoren an, die im Rahmen der jährlichen CIS-Erhebung der EU für den sogenannten „Community Innovation Survey“ abgefragt werden.³ Die Daten werden jährlich erhoben und zielen auf die Innovationskraft der Unternehmen ab.⁴

Unabhängig davon, ob ein Vergleichsgruppenansatz durch eine Befragung nicht geförderter Antragsteller durchgeführt wird oder auf den Informationen externer Datenquellen basiert, besteht bei Vergleichsgruppen immer das Problem, dass allein aus der Gegenüberstellung beider Gruppen (Indikator aus der Evaluationserhebung mit dem Vergleichsindikator) nicht zwingend abgeleitet werden kann, ob der zu beobachtende Unterschied auf eine allgemeine (exogene) Entwicklung oder auf

² Gemeint ist ein möglichst optimales Ergebnis, das aus praktischen Gründen des Kontrollgruppensdesigns nicht möglich ist.

³ Diese Daten werden für Deutschland durch das Zentrum für Europäische Wirtschaftsforschung (ZEW) in Zusammenarbeit mit dem Institut für Angewandte Sozialwissenschaft (infas) und dem Fraunhofer Institut für System- und Innovationsforschung ISI erhoben. Die Umfrage erfolgt im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF). Bestandteil der CIS-Erhebung bzw. des sogenannten Mannheimer Innovationspanels (MIP) sind Indikatoren zur Einführung neuer Produkte, Dienstleistungen und Verfahren in Unternehmen. Auch werden Informationen zu den Aufwendungen für Innovationen und zum Erfolg abgefragt, den Unternehmen mit neuen Produkten, neuen Dienstleistungen und verbesserten Verfahren erzielen.

⁴ Die Einordnung der Daten des MIP erfolgt nach der gängigen Systematik der Wirtschaftszweige der Europäischen Gemeinschaft (sog. NACE-Klassifikation). NACE bezieht sich auf statistische Einheiten, die eine wirtschaftliche Gesamtheit, also ein Unternehmen oder einen Wirtschaftszweig, bilden. Zu bedenken ist jedoch, dass sich die Daten der CIS-Erhebungen aufgrund der methodischen Vorgaben von Eurostat ausschließlich auf Unternehmen mit zehn oder mehr Beschäftigten beziehen und beispielsweise im Dienstleistungsbereich nur bestimmte Branchen umfassen (nur Großhandel, Transportgewerbe und Post, Banken und Versicherungen, EDV und Telekommunikation und technische Dienstleister). Siehe hierzu auch vgl. Licht, G., Harhoff, D. (1993): Das Mannheimer Innovationspanel. In: ZEW Discussion Papers, No. 93–21. Online unter: <http://www.econstor.eu/bitstream/10419/29459/1/257047077.pdf> [01.07.2014]; Schmidt, T., Aschhoff, B. (2007): Die Nutzung der Innovationsdaten des Mannheimer Innovationspanels für die Politikberatung. In: Vierteljahreshefte zur Wirtschaftsforschung 76 (3), S. 17–28. Online unter: <http://ejournals.duncker-humboldt.de/doi/pdf/10.3790/vjh.76.3.17> [01.07.2014].

das Programm zurückzuführen ist – und damit als sogenannter Treatment-Effekt dargestellt werden kann. Ein Ausweg aus diesem „Dilemma“ ist der sogenannte „Differences in Differences-Ansatz“. In bestimmten Fällen erweist sich dieser Ansatz als praktikabel, um Treatment-Effekte – unabhängig von einer randomisierten Mittelvergabe – herauszufiltern.

Überlegungen zu einem praktischen Umgang mit dem Differences in Differences-Ansatz

Ein Differences in Differences-Ansatz (DiD-Ansatz) lässt sich immer dann umsetzen, wenn im Rahmen einer Evaluation mindestens zu zwei Zeitpunkten Daten erhoben vorliegen (z. B. zum Beginn der Evaluationsphase und vom Ende der Evaluationsphase). Damit eignet sich der Ansatz besonders für **begleitende Evaluationen**, ggf. auch für summative bzw. ex-post Evaluationen, insofern zwei Messzeitpunkte zur Verfügung stehen. Bei einer Vielzahl an Beobachtungen bietet sich das Verfahren auch für ökonometrische Schätzmodelle an.⁵

Bei einer reinen Datengegenüberstellung der Programmevaluation mit externen Vergleichsdaten besteht das grundsätzliche Problem, dass sich Förderempfänger (gefördert durch die Maßnahme X) und der Vergleichsindikator (im Folgenden als Vergleichsgruppe Y bezeichnet) in bestimmten Charakteristika

grundlegend unterscheiden (Differenz \overline{AB}). Lassen sich allgemeine Unterschiede beobachten, kann sich die Entwicklung bei den Förderempfängern (X) zum Betrachtungszeitpunkt anders vollzogen haben als in der Vergleichsgruppe (Y), unabhängig von der Förderung, die Gruppe X erhalten hat. Werden die Werte (ohne Berücksichtigung dieser Diskrepanz) gegenübergestellt, würden unter Umständen Effekte der Fördermaßnahme X zugerechnet, die etwa auf externe Effekte (z. B. unterschiedliche konjunkturelle Entwicklungen) zurückzuführen sind. Eine Korrektur kann durch den DiD-Ansatz erfolgen.

Wie Abbildung 1 zeigt, unterscheiden sich Vergleichsgruppe und Förderempfänger bereits anhand bestimmter Charakteristika. In der Grafik wird unterstellt, dass die Förderempfänger der Maßnahme X bereits vor der Maßnahme insgesamt „besser“ aufgestellt sind als diejenige Gruppe, die nicht gefördert wird (Messzeitpunkt 1, vgl. Strecke \overline{HA} und Strecke \overline{HB}). Zusätzlich ist jedoch auch ein Treatment-Effekt in Bezug auf die Förderung zu erwarten (Messzeitpunkt 2). Aus der Differenz zwischen beiden Messzeitpunkten und der unterschiedlichen Entwicklung zwischen Förderempfänger X und Vergleichsgruppe Y lässt sich dann auf den Treatment-Effekt schließen. Dieser ergibt sich aus dem Abgleich der Entwicklung zwischen Gruppe X und Gruppe Y (Vgl. Strecke \overline{IC} und Strecke \overline{IF}) minus der „Differenz zwischen Fördernehmer und Vergleichsgruppe“, die unabhängig vom Treatment existiert (minus Strecke \overline{AB} bzw. Strecke \overline{DF}).⁶

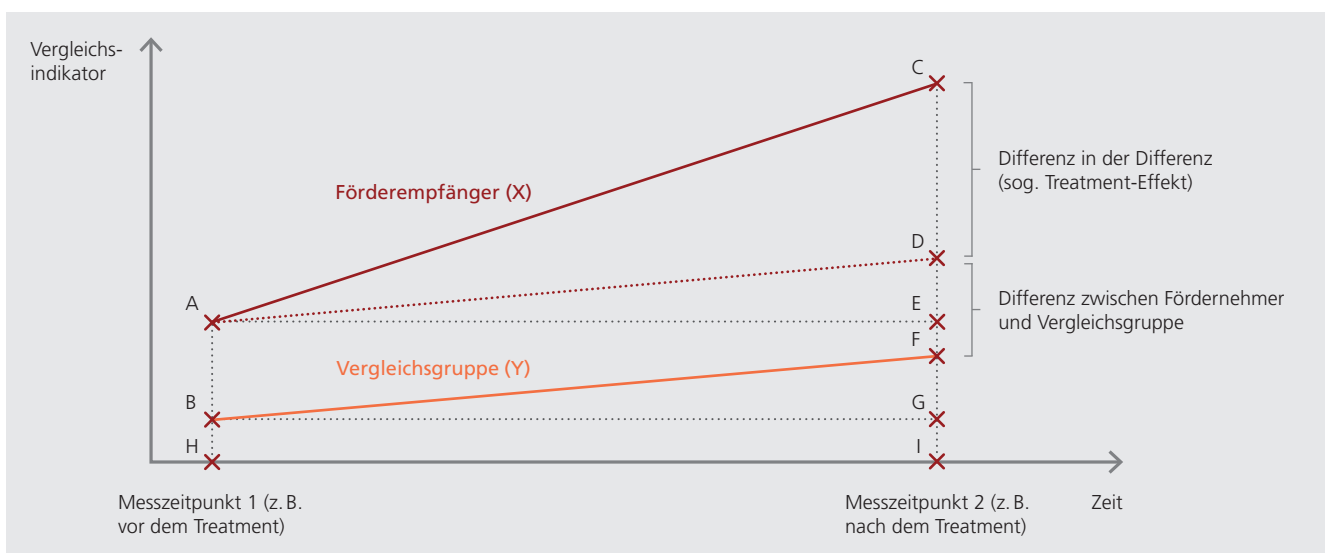


Abb. 1: Beispiel einer Messung des zusätzlich bewirkten Effekts bei Förderempfängern durch den Differences in Differences-Ansatz (eigene Darstellung).

- 5 Vgl. hierzu Abadie, A. (2005): „Semiparametric difference-in-differences estimators“. In: Review of Economic Studies 72 (1), S. 1–19; Bertrand, M., Duflo, E., Mullainathan, S. (2004): „How Much Should We Trust Differences-in-Differences Estimates?“. In: Quarterly Journal of Economics 119 (1), S. 249–275; Card, D., Krueger, A. B. (1994): „Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania“. In: American Economic Review 84 (4), S. 772–793.
- 6 Es wird bereits deutlich, dass bei randomisierten Experimenten – abgesehen vom Treatment – zwischen Treatmentempfänger und Nicht-Treatmentempfänger kein Unterschied existiert, weshalb mit Bezug zu Abbildung 1 beide Gruppen in Punkt A starten und der Vergleich beider Gruppen (Strecke \overline{CD}) bereits den Treatment-Effekt wiedergibt. Deshalb reichen bei randomisierten Experimenten in der Regel zwei Beobachtungszeitpunkte aus, wodurch der Differences in Differences-Ansatz hinfällig wird.

Der Treatment-Effekt lässt sich also wie folgt berechnen:

$$\begin{aligned} \text{Treatment-Effekt} &= \overline{CD} = ((\overline{CE}) - (\overline{FG})) = ((\overline{CF}) - (\overline{AB})) \\ &= (\overline{C} - \overline{F}) - \overline{AB} \end{aligned}$$

Methodisch lässt sich immer dann ein DiD-Ansatz realisieren, wenn für beide Gruppen (Förderempfänger und Vergleichsgruppe) mindesten zwei Beobachtungszeitpunkte zur Verfügung stehen (sog. 2x2-Design). Die vier Variablen unterscheiden sich dann jeweils in Bezug auf zwei Charakteristika: einmal die Befragungszeit und einmal in Bezug auf das Treatment (Teilnahme am Programm X Ja oder Nein).

Die folgende Matrix zeigt anhand eines einfachen Beispiels die Berechnung der Differenz anhand fiktiver Umsatzzahlen.

Durch den DiD-Ansatz ist es also möglich, Unterschiede, die ohnehin existieren, von Treatment-Effekten zu unterscheiden. In dem vorgestellten Beispiel ließe sich schlussfolgern, dass der Treatment-Effekt der **Maßnahme X** zu einem Umsatzplus von **50.000 Euro** ($100.000 - 50.000 = 50.000$) geführt hat.

Differences in Differences-Ansatz mit unterschiedlichen Datenbezugsjahren

Nachdem nun gezeigt werden konnte, dass der DiD-Ansatz bei Vergleichsgruppen eine Identifikation von Treatment-Effekten zulässt, stellt sich allerdings die Frage, ob sich der beschriebene Ansatz auch für einen Vergleich mit Indikatoren der amtlichen Statistik eignet.⁷

Werden nämlich die Daten einer Evaluation (z. B. aus einer Online-Befragung) mit Daten der amtlichen Statistik verglichen,

entsteht ein weiteres Problem: Die aktuellen Daten der amtlichen Statistik werden in der Regel zwei Jahre zeitverzögert veröffentlicht (Datenbezugsjahr X-2). So beziehen sich beispielsweise die Daten der aktuellen CIS-Erhebung 2014 auf das Jahr 2012. Demgegenüber werden bei Online-Erhebungen im Jahr 2014 in der Regel die Umsatzzahlen des Jahres 2013 abgefragt (Datenbezugsjahr X-1).

Hat z. B. eine begleitende Evaluation eine Laufzeit von mehreren Jahren so ist das beschriebene Problem weniger relevant, da die veröffentlichten Daten der amtlichen Statistik ein Jahr später das gleiche Bezugsjahr haben, wie die erste Online-Befragung. Doch bei kürzeren Laufzeiten wird das Problem für die zweite Befragung offensichtlich, da z. B. aus der Online-Erhebung im Jahr X Indikatoren für Jahr X-1 vorliegen, während aus der amtlichen Statistik nur Indikatoren für das Jahr X-2 vorhanden sind. Die folgenden beiden Optionen bieten sich an, um das Problem zu umgehen:

1. In der Online-Befragung werden gegen Ende der Projektlaufzeit – zusätzlich zu den Daten, die sich auf das Vorjahr beziehen – auch Werte abgefragt, die zwei Jahre zurückliegen. Hat die begleitende Evaluation jedoch nur eine kurze Laufzeit (z. B. drei Jahre), so tritt das Problem auf, dass die Effekte des Förderprogramms in diesem Fall kaum zum Tragen kommen können. Demnach ist diese Vorgehensweise bei kürzeren Zeiträumen kaum sinnvoll.
2. Als zweite Option besteht die Möglichkeit, die aktuellen Zahlen der Online-Befragung jeweils mit den in diesen Jahren veröffentlichten Zahlen der amtlichen Statistik zu vergleichen, ohne dass die Datenbezugsjahre identisch sind.

Die zweite Option soll in der Folge einer näheren Betrachtung unterzogen werden. Die Frage ist, ob mit den unterschiedlichen Datenbezugsjahren trotz der offensichtlichen Unschärfe den-

	fiktive Umsatzzahlen (Durchschnittswerte)			
Beispiele Indikatoren:	Treatment	Messzeitpunkt 1	Messzeitpunkt 2	Differenz 2-1
▶ Umsatz	Ja (X)	500.000	600.000	100.000
▶ Mitarbeiter	Nein (Y)	300.000	350.000	50.000
▶ Patente	Differenz X-Y	-200.000	-250.000	50.000
▶ Publikationen				
▶ Kontakte zu Partnern				
▶ FuE-Ausgaben				

Tabelle 1: Matrix-Darstellung des Differences in Differences-Ansatzes (eigene Darstellung).

⁷ Voraussetzung ist, dass im Rahmen der Evaluation bei den Förderempfängern grundsätzlich die gleichen Indikatoren abgefragt werden, die auch der CIS-Erhebung zugrunde liegen.

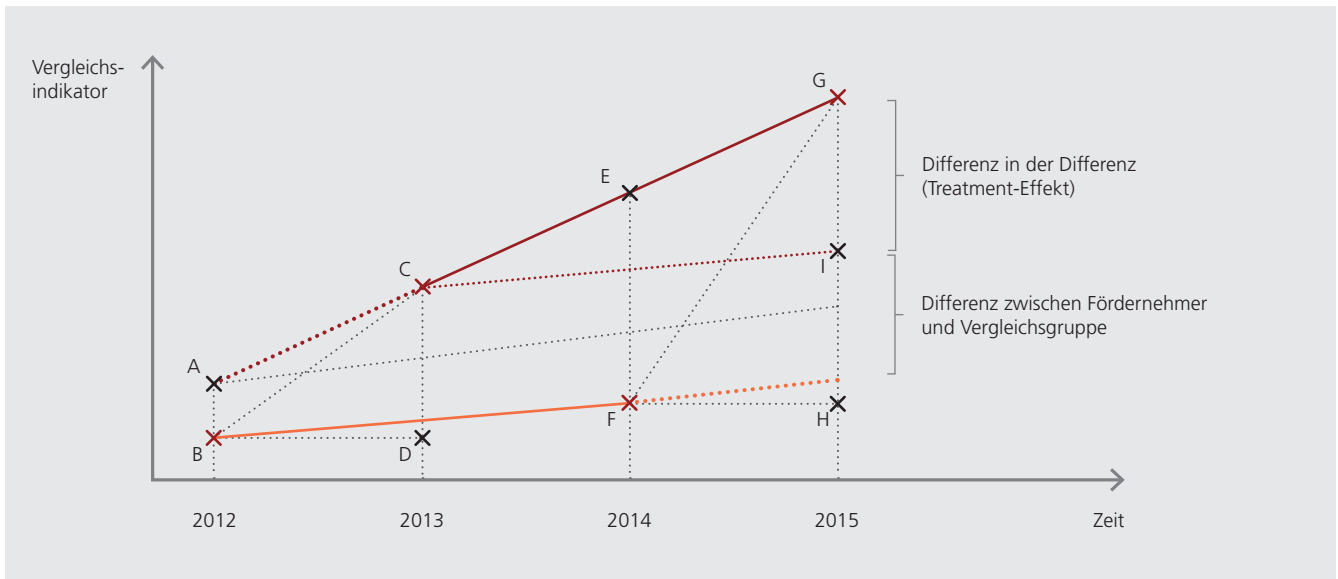


Abbildung 2: Beispiel einer Messung des zusätzlich bewirkten Effekts bei Förderempfängern durch den Differences in Differences-Ansatz unter Berücksichtigung der unterschiedlichen Datenbezugsjahre (eigene Darstellung).

noch sinnvolle Ergebnisse erzielt werden können. Diese Frage lässt sich am besten anhand einer leichten Modifikation von Abbildung 1 verdeutlichen.

Unter Berücksichtigung unterschiedlicher Datenbezugsjahre (z. B. 2012/2013 sowie 2014/2015) verändert sich das Schaubild im Vergleich zur Abbildung 1. Die Abbildung 2 zeigt hingegen den DiD-Ansatz, in dem Daten mit unterschiedlichen Bezugsjahren verglichen werden.

Anhand Abbildung 2 wird deutlich, dass unter der **Annahme eines linearen Trends**, sowohl für die geförderten Unternehmen wie für die Vergleichsgruppe (z. B. die CIS-Daten), die zeitliche Verzögerung im DiD-Ansatz zu keinen systematischen Verzerrungen führen sollte, insofern sich beide Gruppen in einer parallelen Dynamik weiter entwickeln. Unter dieser Prämisse ist der durch den DiD-Ansatz gemessene Unterschied tatsächlich auf die Programmteilnahme zurückzuführen, was die nachfolgende Gleichung noch einmal veranschaulicht:

$$\begin{aligned} \text{Treatment-Effekt} &= \overline{GI} = (G_{2015} - F_{2014}) - (C_{2013} - D_{2012}) \\ &= ((GH) - (CD)) = ((EF) - (AB)). \end{aligned}$$

Verzerrungen sind hingegen immer dann zu erwarten, wenn die Entwicklung zwischen beiden Gruppen (Förderempfänger und Vergleichsgruppe) sehr unterschiedlich verläuft und damit kein linearer Trend unterstellt werden kann. Dies gilt etwa für Strukturbrüche wie im Rahmen der Finanzkrise 2008/2009.

Fazit

Vor allem in der praktischen Umsetzung von Evaluationen ist die Suche nach geeigneten Kontroll- oder Vergleichsgruppenansätzen häufig ein konkretes Problem. Im Gegensatz zu Vergleichsgruppenansätzen, die mindestens zwei Messzeitpunkte erfordern, liefern Kontrollgruppenansätze schon bei der einfachen Gegenüberstellung der Indikatoren (Förderempfänger vs. Kontrollgruppe) Erkenntnisse zum Treatment-Effekt. Kontrollgruppenansätze sind daher durchaus als „First-Best-Lösung“ zu verstehen. Die praktische Umsetzung scheitert jedoch in der Regel daran, dass die Mittelvergabe nicht randomisiert erfolgt. Damit lassen sich Effekte der Förderung nur mit einem Vergleichsgruppenansatz sinnvoll identifizieren. In dieser sogenannten „Second-Best-Welt“ wird häufig in Erwägung gezogen, abgelehnte Förderempfänger als Vergleichsgruppe zu befragen. Bei einer näheren Betrachtung stellt sich jedoch auch diese Option in vielen Fällen als nicht praktikabel heraus. Als zusätzliche Option können schließlich relevante Indikatoren der amtlichen Statistik zum Vergleich herangezogen werden, wie zum Beispiel die FuEul-Indikatoren der CIS-Erhebung.

Unabhängig davon, welche Daten bzw. welche Gruppe für einen Vergleich herangezogen wird, bietet sich der DiD-Ansatz an, um im Rahmen von Vergleichsgruppenansätzen Treatment-Effekte von externen Effekten zu unterscheiden. Der Vorteil ist, dass hierdurch auch ohne einen großen methodischen Aufwand Datensätze unterschiedlicher Bezugsjahre sinnvoll miteinander verglichen werden können. Eine wichtige Voraussetzung ist jedoch, dass für die Umsetzung des beschriebenen Ansatzes jeweils mindestens zwei Beobachtungszeitpunkte und

verfügbare Vergleichsindikatoren vorhanden sind, die in keinem direkten Zusammenhang mit den zu evaluierenden Förderempfängern stehen. Unter diesen Voraussetzungen ist der DiD-Ansatz als sinnvolles Instrument anzusehen – der sich besonders für begleitende Evaluationen und/oder ex-post bzw. summative Evaluationen eignet, vorausgesetzt es stehen zwei Messzeitpunkte zur Verfügung.

Die Diskussion zur praktischen Umsetzung von Kontroll- und Vergleichsgruppenansätzen zeigt schließlich, wie wichtig es ist, eine mögliche Evaluation bereits ex-ante – im Idealfall also bereits bei der Ausschreibung des Programms – in alle Überlegungen miteinzubeziehen. Dieser Aussage liegt die Erkenntnis zu Grunde, dass die Messbarkeit bestimmter Effekte entscheidend davon abhängt, zu welchem Zeitpunkt eine Evaluation erfolgt (ex-ante, begleitend und/oder ex-post) und ob die praktische Umsetzung der Mittelvergabe randomisiert oder selektiv umgesetzt wurde.

Kontakt:

Institut für Innovation und Technik (iit)
Steinplatz 1, 10623 Berlin

Dr. Leo Wangler

Institut für Innovation und Technik (iit)
Tel.: 030 310078-434
E-Mail: wangler@iit-berlin.de

iit perspektive Nr. 22

September 2014
Layout: Anne-Sophie Piehl