

Meier, Kristina

Research Report

Wirkungsevaluierung in der Entwicklungszusammenarbeit

DIW Roundup: Politik im Fokus, No. 27

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Meier, Kristina (2014) : Wirkungsevaluierung in der Entwicklungszusammenarbeit, DIW Roundup: Politik im Fokus, No. 27, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/111806>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

DIW Roundup

Politik im Fokus

Deutsches Institut für Wirtschaftsforschung

2014

Wirkungsevaluierung in der Entwicklungszusammenarbeit

Kristina Meier

Wirkungsevaluierung in der Entwicklungszusammenarbeit

Kristina Meier | kmeier@diw.de | Abteilung Entwicklung und Sicherheit am DIW Berlin

Nachdem jahrzehntelang Projekte in der Entwicklungszusammenarbeit entweder gar nicht oder bestenfalls intern nach uneinheitlichen Standards hinsichtlich ihrer Wirkung bewertet wurden, rückt nun seit einiger Zeit die methodisch fundierte, quantitative Impactevaluierung immer mehr in den Vordergrund. Auch Deutschland hat 2012 mit der Gründung des unabhängigen Deutschen Evaluierungsinstituts der Entwicklungszusammenarbeit (DEval) einen wichtigen Schritt in diese Richtung unternommen. Es leuchtet ein, dass eine objektive Erfolgsmessung von Projekten nötig ist um eine effiziente Allokation begrenzter finanzieller Mittel zu ermöglichen. Darüber hinaus ist es ebenfalls wichtig sicher zu stellen, dass eine Intervention auch tatsächlich die erwarteten positiven Effekte aufweist und nicht etwa unvorhergesehene negative Auswirkungen auf die Zielbevölkerung hat (Terberger (2011)). Die Schwierigkeiten, die es hierbei zu beachten gilt, sollen im Folgenden kurz erörtert werden.

Das Problem der Verzerrung

Um kausale Zusammenhänge zwischen einer Intervention und beobachteten Ergebnissen methodisch fundiert nachzuweisen, muss mit Hilfe geeigneter Analyseansätze ausgeschlossen werden, dass unbeachtete Faktoren die Resultate zusätzlich zur evaluierenden Intervention beeinflussen.

Die Problematik lässt sich an folgendem Beispiel veranschaulichen: Ein Dorf erhält durch die Bereitstellung eines Brunnens erstmalig direkten Zugang zu sauberem Trinkwasser. Nun soll ermittelt werden, ob hierdurch die Zahl der Magen-Darm-Erkrankungen in der Dorfbevölkerung zurückgegangen ist. Ein einfacher Vorher-nachher-Vergleich kann einen guten Anhaltspunkt liefern, ist aber nicht ausreichend, um Kausalität zwischen Intervention und Ergebnis statistisch valide nachzuweisen. Dies liegt daran, dass es in der Regel unmöglich ist, alle parallel zur Intervention ablaufenden Veränderungen in die Berechnungen miteinzubeziehen. So könnte beispielsweise zeitgleich mit der Inbetriebnahme des Brunnens zufällig eine nicht primär durch Trinkwasser übertragene Infektionskrankheit ausgebrochen sein. Dies würde zu einer Unterschätzung des möglicherweise positiven Effektes der Maßnahme auf die Gesundheit der Dorfbevölkerung führen. Ebenso wenig ausreichend wäre ein einfacher Vergleich der Situation mit einem benachbarten Dorf, in welchem kein Brunnen gebaut wurde. Hier muss man davon ausgehen, dass die Projektteilnahme nicht zufällig geschah, sondern beispielsweise durch Faktoren wie geographische Lage/Erreichbarkeit des Dorfes, oder aktive Lobbyarbeit einflussreicher Bewohner beeinflusst wurde. In diesem Fall ist zu erwarten, dass sich die Bevölkerung in den beiden Dörfern unterscheidet, und dass diese Unterschiede nicht in allen Fällen quantitativ messbar und somit mit den zur Evaluierung vorliegenden Daten kontrollierbar

sind. Hier kommen nicht oder nicht ausreichend genau erhobene Variablen wie Einkommenshöhe oder Bildungsgrad in Frage, aber eben auch schwerer greifbare Faktoren, wie zum Beispiel politischer Einfluss. All dies sind jedoch Eigenschaften, die sich höchstwahrscheinlich auch auf die Ergebnisvariable, in diesem Fall die Erkrankungsrate, auswirken werden. Wohlhabendere und einflussreiche Individuen werden erfahrungsgemäß im Durchschnitt mit ihren Familien in besseren hygienischen Verhältnissen leben und demnach auch ohne den Zugang zu einem Brunnen eine niedrigere Inzidenz von Magen-Darm-Erkrankungen aufweisen. Im hier skizzierten Beispiel würde dies zu einer Überschätzung des Interventionseffektes führen.

Um den wahren Effekt einer Maßnahme ohne die geschilderte Verzerrung zu berechnen, müsste man die Situation **desselben** Individuums zum **selben** Zeitpunkt einmal mit Erhalt der Intervention und einmal ohne vergleichen. Da dies offensichtlich unmöglich ist, kann man sagen, dass jeder Wirkungsevaluierung ein Problem fehlender Daten zugrunde liegt, welches mit Hilfe angemessener Methoden so weit wie möglich abgemildert werden muss. Im Kern wird mit jeder dieser Methoden versucht, die fehlende Beobachtung, also die hypothetische Situation des behandelten Individuums, wenn es nicht an der Intervention teilgenommen hätte (kontrafaktisches Szenario), durch die Ergebnisse einer Kontrollgruppe so gut wie möglich zu approximieren. Hierbei besteht die Kontrollgruppe aus Individuen, die nicht an der Intervention teilnehmen, ansonsten aber soweit wie möglich mit den Teilnehmern vergleichbar sind.

Quasi-experimentelle Methoden vs. Randomisierung

Als „Königsweg“ in der Wirkungsevaluierung wird häufig die sogenannte Randomisierung bezeichnet. Hier wird ausschließlich per Zufall über die Projektteilnahme entschieden. In diesem Fall kann davon ausgegangen werden, dass sich sämtliche Unterschiede zwischen Teilnehmern und Nichtteilnehmern im statistischen Mittel verlieren, es kann also durch einen einfachen Vergleich der Ergebnisse der beiden Gruppen vor und nach Durchführung der Maßnahme statistisch valide der Interventionseffekt gemessen werden. Es liegt allerdings auf der Hand, dass eine solche Randomisierung nicht immer durchführbar ist. Zum einen setzt sie voraus, dass bereits vor Beginn der Projektimplementierung eine spätere Evaluierung gezielt geplant wird, was häufig nicht der Fall ist. Zum anderen ist eine rein zufällige Verteilung der Interventionsteilnahme oft auch aus ethischen Gründen fragwürdig. Dies wird deutlich, wenn man sich beispielsweise den Ausschluss bedürftiger Kinder von einer Impfkation vorstellt, welcher mit der besseren Evaluierbarkeit der Maßnahme begründet wird. Allerdings wird dieser Kritikpunkt in der Realität häufig dadurch abgemildert, dass aufgrund begrenzter Projektmittel ohnehin nur ein Teil der Bevölkerung an einer Intervention teilnehmen kann. Hier wäre eine rein zufällige Zuteilung wiederum am gerechtesten. Darüber hinaus gibt es auch Maßnahmen, wie zum Beispiel Infrastrukturprojekte, welche aus technischen Gründen nicht randomisierbar sind – dies betrifft etwa netzwerkbasierte Technologien wie Leitungswasser, Elektrizitätsversorgung oder Straßenbau.

Ist eine Randomisierung nicht möglich oder nicht erwünscht, kommen sogenannte quasi-experimentelle Methoden zum Einsatz. Hierzu gehören vor allem das Propensity Score Matching (Rosenbaum und Rubin (1983)), die Methode der

doppelten Differenzen (siehe z.B. Pitt und Khandker (1998)), sowie auch der Instrumentalvariablenansatz (siehe Angrist und Krueger (2009) für einen Überblick über wissenschaftliche Impactevaluierungen, die diese Methode verwenden.) Auf eine detaillierte Diskussion der einzelnen Ansätze soll hier verzichtet werden (siehe hierzu [Khandker et al. \(2010\)](#)). Verallgemeinert lässt sich sagen, dass sie alle das Ziel haben, die Unterschiede zwischen dem hypothetischen kontrafaktischen Szenario und der Kontrollgruppe soweit wie möglich zu minimieren. Hierzu treffen sie unterschiedliche Annahmen, deren Erfüllung in der Regel nicht verlässlich nachweisbar ist, weshalb die resultierenden Ergebnisse im Vergleich zur Randomisierung als weniger belastbar einzustufen sind.

Quantitative vs. qualitative Evaluierung

Je nach Art der zu bewertenden Maßnahme kann allerdings überhaupt jegliche quantitative Wirkungsevaluierung, egal ob randomisiert oder mit Hilfe quasi-experimenteller Methoden, ungeeignet sein. Dies kann zum einen an den hohen Kosten liegen, die durch die Datenerhebung hervorgerufen werden und für kleinere Investitionen schnell prohibitiv werden können. Nicht immer steht auch, selbst bei ausreichenden finanziellen Mitteln, der Erkenntniswert einer solchen Evaluierung in einem sinnvollen Verhältnis zu den Kosten. Ebenso gibt es Maßnahmen, vor allem im Politikbereich, wo aufgrund der niedrigen Beobachtungszahl quantitative Methoden nicht anwendbar sind. In diesen Fällen ist eine qualitative Evaluierung vorzuziehen.

Die Entwicklung hin zu unabhängigen, methodisch fundierten Impactanalysen ist ein wichtiger Schritt in der Bewertung von Entwicklungszusammenarbeit. Ebenso wichtig ist es allerdings, dass die hierdurch gewonnenen Erkenntnisse bezüglich Wirkungsmechanismen der einzelnen Interventionen auch aktiv in der Planung und Steuerung weiterführender Projekte eingesetzt werden.

Quellen

Angrist, Joshua D. ; Krueger, Alan B. (2009): " *Instrumental variables and the search for identification: From supply and demand to natural experiments*". Journal of Economic Perspectives, 15(4):69{85, 2009.

Khandker et al (2010): " *Handbook on impact evaluation: quantitative methods and practices.*" World Bank Publications, 2010. <http://disde.minedu.gob.pe/xmlui/handle/123456789/1419>

Pitt, Mark M.; Khandker, Shahidur R. (1998): " *The impact of group-based credit programs on poor households in bangladesh: Does the gender of participants matter?*" Journal of Political Economy, 106(5):958{996, 1998.

Rosenbaum, Paul R.; Rubin, Donald B. (1983): "The central role of the propensity score in observational studies for causal effects." Biometrika, 70(1):41{55, 1983.

Terberger, Eva (2011): "Evaluierung in der Entwicklungszusammenarbeit. Das Beispiel der finanziellen Zusammenarbeit", *Nachhaltigkeit in der Entwicklungszusammenarbeit*. VS Verlag für Sozialwissenschaften, 2011. 219-238.

Impressum

DIW Berlin – Deutsches Institut
für Wirtschaftsforschung
Mohrenstraße 58, 10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
<http://www.diw.de>

ISSN 2198-3925

Alle Rechte vorbehalten
© 2014 DIW Berlin

Abdruck oder vergleichbare
Verwendung von Arbeiten
des DIW Berlin ist auch in
Auszügen nur mit vorheriger
schriftlicher Genehmigung
gestattet.