

Komplexe Methoden verständlich machen - Das Beispiel der Synthetic Control Method

Session: Mit Verwaltungen kommunizieren: Herausforderungen für
Evaluationsforschung und -praxis

Steffen Zabler

DUV Speyer

17.09.2020

Einführung

Argumentationslinie

- Spezifische Herausforderungen von Evaluationen im politisch-administrativen Bereich...
- ...erfordern teils methodisch komplexe Lösungen,...
- ...die ihrerseits Schwierigkeiten für die Kommunikation/Vermittlung von Ergebnissen (besonders in die Praxis hinein) nach sich ziehen.
- Daher ist Offenheit für kreative und innovative methodische Verfahren unabdingbar,...
- ...wobei die Synthetic Control Method (Synthetisches Matching) analytische Probleme löst...
- ...und zugleich eine gelungene Kombination aus wissenschaftlichem State-of-the-art und Vermittelbarkeit darstellt.

Einführung - Herausforderungen von Evaluationen

Typische Herausforderungen von Evaluationen

Per se für alle Evaluationsdesigns zutreffend

- Zweck der Evaluation, geeigneter Akteur, passender Evaluationsgegenstand, geeignete Kriterien, Verwendung von Ergebnissen

Zentrale Herausforderung: Kausalität

- 'impact problem' (Vedung 2000): Klärung von Ursache und Wirkung
- Ist eine bestimmte Intervention tatsächlich für ein zu evaluierendes Outcome verantwortlich?
- Vergleiche bzw. Varianz zwischen Einheiten oder über Zeit sehr annahmenstark
- soweit nichts Besonderes, klassisches Problem aller (Sozial-)Wissenschaften

Einführung - Herausforderungen von Evaluationen

Kausalität in den Sozialwissenschaften

- 'fundamental problem of causal inference' (Holland 1986): Unmöglichkeit der Beobachtung einer kontrafaktischen Situation
- 'scientific solution' (ibd.)
 - eine Einheit zu zwei Zeitpunkten (Annahmen: temporal stability, causal transience)
 - zwei Einheiten zum gleichen Zeitpunkt (Annahme: unit homogeneity)
- 'statistical solution' (ibd.)
 - experimentelles Setting (Annahme: independence)
 - Analyse mit Beobachtungsdaten ex post (Annahme: conditional independence)

Einführung - Herausforderungen von Evaluationen

Möglichkeiten der Kausalanalyse in der PA-Evaluation

- Annahmen von simplen Vergleichen unrealistisch
- Ausschluss von (randomisierten) Experimenten aufgrund praktischer und ethischer Erwägungen; externe Validität
- was (meist) bleibt: ex post Beobachtungsdaten inkl. Kontrolle von Drittvariablen

Probleme der notwendigen analytischen Komplexität

- hohe Anforderungen an Daten und/oder Untersuchungssettings
- komplexe statistische Modelle
- Interpretationshindernisse

Einführung - Herausforderungen von Evaluationen

Probleme in der Praxis

- Kommunikation mit Stakeholdern (diverse Hintergründe)
- Bewertung der Qualität der Ergebnisse
- Bereitstellung passender Materialien (Berichte, Handreichungen etc.)
- "natürliche Abwehrhaltung"

⇒ "stiefmütterliches" Dasein in der praxisnahen Evaluation

Zusätzliche Herausforderungen

- politisch-administrative Programme mit "Seltenheitswert"
- oft wenige Treatmentfälle, daher kaum Varianz (auch in Drittvariablen)
- **Welcher Ansatz ist (auch in seiner Vermittlung) praxistauglich und genügt dennoch wissenschaftlichem Anspruch?**

Analytischer Ansatz

Standardlösung: qualitativ-vergleichende Fallstudie

- Vergleichslogik intuitiv
- greifbare reale Fälle
- verschiedene Probleme:
 - quantitatives Interesse nicht im Vordergrund
 - Fallauswahl
 - kausale Kontrolle

Alternativ: Synthese eines Vergleichsfalls

'arguably the most important innovation in the policy evaluation literature in the last 15 years' (Athey/Imbens 2017, S. 9)

- statt der Suche nach einem Vergleichsfall...
 - ...Konstruktion eines passenden Falls...
 - ...auf der Basis objektiver Kriterien
- ⇒ Simple Vergleichslogik zweier Fälle bleibt bestehen!

Analytischer Ansatz - Synthetic Control Method

Vorgehen

1. Auswahl des Analysefalls, Festlegung des Zeitpunkts des Treatments
2. Definition Pool potentieller Vergleichsfälle
3. Sammlung von Paneldaten: zentrales Outcome, Vorhersagevariablen ("untheoretisch")
4. Berechnung von Gewichten für potentielle Vergleichsfälle
 - in Abhängigkeit der Ähnlichkeit des Outcomes vor dem Treatment
 - und der Ähnlichkeit der Vorhersagevariablen
5. Berechnung gewichteter Mittelwert als synthetischer Vergleichsfall
6. Vergleich der Verläufe vor und nach dem Treatment

Analytischer Ansatz - Synthetic Control Method

Stärken

- objektive Fallauswahl
- "Einfangen" von unbeobachteten Drittvariablen
- quantitatives Ergebnis
- Angabe von Unsicherheit
- Kriterien zur Eignung der Methode
- zahlreiche Erweiterungen verfügbar
- intuitive (Vergleichsfall-)Logik

Anwendungsfall - Sparkommissar in Hagen

Beratender Sparkommissar

- Einsatz 2008 in Hagen (davor zwei andere Fälle)
- aufgrund einer problematischen Haushaltssituation
- lediglich beratend tätig (in Ggs. zum Staatskommissar)

Analyse aus Zabler (2016)

- im Fokus: Kassenkreditverschuldung
- verschiedene Determinanten der Verschuldung als Vorhersagevariablen
- restliche krfr. Städte in NRW als Vergleichspool

Anwendungsfall - Sparkommissar in Hagen

Tabelle 2: Mittelwerte und Gewichte der Variablen zur Vorhersage der Höhe der Kassenkredite

	Hagen	Synthetischer Vergleichsfall	Durchschnitt Kontrollgruppe	Gewicht der Variablen
Bevölkerung	202759,538	374964,502	341294,873	0,000
Bevölkerungsentwicklung	-1521,250	-370,134	-557,192	0,008
Arbeitslosenquote	12,010	11,059	12,082	0,051
Seniorenquote	0,198	0,187	0,185	0,000
Einkommen	16914,750	16943,537 ←	16848,963	0,238
Gewerbesteuer	0,364	0,467	0,436	0,032
Zentralität	8,000	9,113	8,500	0,000
Personalstand	0,052	0,087	0,053	0,000
Personalausgaben	0,665	0,641 ←	0,589	0,177
Wettbewerbspotential	1,000	1,291	1,800	0,04
Wettbewerbsintensität	2,000	1,818 ←	1,200	0,21
Durchsetzungsstärke	2,000	2,280	2,750	0,1
Anteil SPD-Sitze	0,428	0,373	0,394	0,008
Kassenkredite 1995	61,056	55,927	51,888	0,059
Kassenkredite 2000	20,344	51,787	81,306	0,074

Anmerkung: Rundung der Gewichte auf die dritte Nachkommastelle.

Quelle: Eigene Darstellung auf Basis eigener Berechnungen.

Anwendungsfall - Sparkommissar in Hagen

Tabelle 3: Gewichte der Kommunen aus der Kontrollgruppe zur Bildung des synthetischen Vergleichsfalls

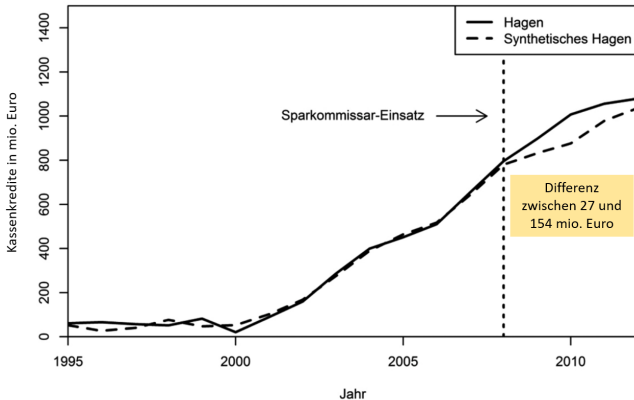
Kommune	Gewicht	Kommune	Gewicht
Bielefeld	0,000	Herne	0,025
Bochum	0,006	Krefeld	0,000
Bonn	0,410	Köln	0,000
Bottrop	0,001	Leverkusen	0,126
Dortmund	0,000	Mönchengladbach	0,000
Duisburg	0,101	Münster	0,000
Düsseldorf	0,000	Oberhausen	0,000
Essen	0,246	Remscheid	0,000
Gelsenkirchen	0,084	Solingen	0,000
Hamm	0,000	Wuppertal	0,000

Anmerkung: Rundung der Gewichte auf die dritte Nachkommastelle.

Quelle: Eigene Darstellung auf Basis eigener Berechnungen.

Anwendungsfall - Sparkommissar in Hagen

Abbildung 1: Verlauf der Kassenkredithöhe über den Zeitraum von 1995 bis 2012 der kreisfreien Stadt Hagen und ihrem synthetisch gebildeten Vergleichsfall



Quelle: Eigene Darstellung auf Basis eigener Berechnungen.

Implikationen und Praxiskommunikation

Spezifische Praxisrelevanz der SCM

- "harte Zahlen" für seltene Ereignisse
- qualitative Fallstudien nicht die einzige Strategie
- Matching-Logik intuitiver als vergleichsweise (für den Praktiker) intransparente stat. Modelle
- grafischer Ansatz

Implikationen und Praxiskommunikation

"Richtige" Vergleiche schaffen!

- Probleme von Vorher-Nachher-Vergleichen
- ...auch wenn diese nicht per se "falsch" sind.
- Bewusstsein für Kausalität/kontrafaktische Situationen
- Erweiterung des wissenschaftlichen Angebots

Transparenz und Vertrauen schaffen!


- gute Vermittelbarkeit \neq triviale Methode ("Mut zur Komplexität")
- Offenlegung der Analyseschritte
- Diskussion der Gültigkeit/potentieller Unsicherheiten der Methode
- im Ggs. zu "auf den Knopf drücken"
- Begrenztheit der Erkenntnisse

Implikationen und Praxiskommunikation

Keynote Monika Taddicken vom 17.9.20:

Beispiel: Evaluation der Qualität als Begleitforschung des Wissenschaftskommunikations-Projekts „Die Debatte“

Ableitung verschiedener Qualitätsdimensionen von Wissenschaftskommunikation (Wicke & Taddicken, 2020):



- **Verständlichkeit** komplexer wissenschaftlicher Themen
- **Erklärung** wissenschaftlicher Forschungsprozesse und Methoden
- Thematisierung von **Unsicherheiten und Widersprüchen** von wissenschaftlichen Ergebnissen
- Darstellung der **Vielfalt** (wissenschaftlicher) Perspektiven
- Hervorheben der **Alltagsrelevanz** wissenschaftlicher Fragen
- Berücksichtigung von **Neuigkeits- und Unterhaltungswert**

Technische Universität
17. September 2020 | Monika Taddicken | Eine Frage der Erwartungen? Eine kommunikationswissenschaftliche Perspektive auf Qualitätskriterien in | Kanone vor 74 | Jahresthema der DfP (Paul Fehrer) | Seite 16

