



Evaluation of Labor Market Programs: Quality Assurance and Quality Development

Dr. Ulf Rinne

IZA – Institute of Labor Economics

Mainz, 18. Januar 2018

Über das IZA

Forschungsinstitut zur Zukunft der Arbeit (IZA)

- Das IZA analysiert Probleme und Organisation der Arbeit in einer sich rasch verändernden ökonomischen Umwelt. Als **privates, unabhängiges Institut** wird es seit seiner Gründung im Jahr 1998 von der Deutschen Post-Stiftung gefördert.
- Das IZA betreibt **internationale Grundlagenforschung und Politikberatung**. Es kooperiert mit einem der größten Forschungsnetzwerke der Welt, dem rund 1.500 Spitzenforscherinnen und -forscher aus aller Welt angehören.
- Zu den renommierten Schriftenreihen des Instituts zählen die **IZA Diskussionspapiere**, eigene Open-Access Fachzeitschriften und das Online-Kompendium **IZA World of Labor**.
- Unsere Mission: **Wissenschaft in Dienst der Gesellschaft**.

Kausale
Wirkungsforschung

Das fundamentale Evaluationsproblem

- Um kausale Ursache-Wirkungs-Zusammenhänge zu identifizieren, muss jede mikroökonomische Evaluation das **fundamentalen Evaluationsproblem** lösen:

Was ist der Unterschied in der Ergebnisvariable der Teilnehmer zwischen der Situation, in der sie teilnehmen, und der Situation, in der sie nicht teilnehmen?

- Beide Situationen können für dieselbe Person jedoch nicht gleichzeitig beobachtet werden – die **kontrafaktische Situation** muss deshalb in geeigneter Weise approximiert werden.

Kontrafaktische Situation und Selektionsverzerrung

- Die **kontrafaktische Situation** der Teilnehmer z.B. durch den Mittelwert der Ergebnisvariable der Nichtteilnehmer zu approximieren ist in der Regel nicht ratsam.
- Es resultiert eine verzernte Schätzung des Teilnahmeeffektes, falls Teilnehmer und Nichtteilnehmer sich nicht nur im Teilnahmezustand unterscheiden (**Selektionsverzerrung**).
- Es ist erforderlich, **Selektionseffekte und kausale Teilnahmeeffekten** zu trennen.
- **Beispiel:** Qualifizierte Individuen nehmen mit einer höheren Wahrscheinlichkeit an einem Weiterbildungsprogramm teil und haben gleichzeitig bessere Perspektiven auf dem Arbeitsmarkt – auch ohne ihre Teilnahme am Weiterbildungsprogramm.

Verpasste Chancen

Durch „richtige“ Evaluationen lernen

- Es werden jedes Jahr massenweise „Evaluationsstudien“ produziert, in denen die offensichtliche Frage nach **kausalen Effekten** unbeantwortet bleibt.
- Auf diese Weise wird der **Begriff „Evaluation“** missbraucht und in Misskredit gezogen.
- Es werden **öffentliche und privatwirtschaftliche Mittel** verschwendet, ohne neues Wissen zu generieren.
- So werden Chancen vertan, auf Grundlage **evidenzbasierter Erkenntnisse** künftig bessere Entscheidungen zu treffen.
- Die **zentrale Kritik** ist eine unzureichende Annäherung an die kontrafaktische Situation.

Elemente einer „richtigen“ Evaluation

- **Wirkungsanalyse:** Identifizierung der kausalen Wirkung einer Politikmaßnahme bzw. einer Intervention (z.B. Einführung des Mindestlohns, Weiterbildungsmaßnahme).
- Dies erfordert die Identifizierung der **kontrafaktischen Situation als Referenz**.
- Die Wirkungsmessung erfolgt **auf Grundlage von Ergebnisvariablen** und nicht anhand von Eingangsgrößen (z.B. Investitionen, Mittelabfluss).
- **Effizienzanalyse:** Eine hilfreiche Kosten-Nutzen-Analyse basiert auf den Ergebnissen der Wirkungsanalyse und somit auf kausalen Ursache-Wirkungs-Zusammenhängen.

Lösungsansätze

Lösungsansätze in der Theorie

- **Randomisierte Experimente** (zufällige Variation der Programmteilnahme).
- **Natürliche Experimente** (Diskontinuitäten in der Wahrscheinlichkeit der Programmteilnahme, z.B. exogene „Schocks“ oder regionale und zeitliche Variation).
- **Quasi Experimente** (Anwendung ökonometrischer Verfahren, um ex post eine geeignete Approximation der kontrafaktischen Situation herzustellen).

Übersetzung in die Praxis

- **Systematische bzw. kontrollierte Variation** der Politikmaßnahme bzw. der Teilnahme-wahrscheinlichkeit an einer Intervention erleichtert eine kausale Wirkungsanalyse.
- So sollte die landesweite Einführung einer Politikmaßnahme oder Interventionen zum selben Zeitpunkt wenn möglich vermieden werden, denn **sowohl zeitliche als auch regionale Variation ist bei einer Evaluation hilfreich** (Negativbeispiel: Einführung des flächendeckenden gesetzlichen Mindestlohns in Deutschland zum 1. Januar 2015).
- Ein geeignetes Evaluationsdesign (inklusive Datenerhebung bzw. -verfügbarkeit) sollte **bereits integraler Bestandteil der Programmkonzeption** sein.

Ziel: Vermeidung von Selektionsprozessen

Ex ante:

- Programmteilnahme ist tatsächlich eine **zufällige Entscheidung** (und darüber wird nicht durch Sachbearbeiter oder Teilnehmende selbst entschieden, d.h. der Teilnahmestatus sollte nicht von unbeobachteten Merkmalen beeinflusst werden).
- Ausnutzung von **Kapazitätsrestriktionen** als Zufallselement.
- **Verzögerte Implementierung** einer Maßnahme in zufällig ausgewählten Regionen.

Ex post:

- **Ökonometrische Methoden** („Reparaturmaßnahmen“).

Ein kluges Programmdesign ist überlegen

- Mit **ökonometrischen Methoden** kann die kontrafaktische Situation häufig auch noch ex post approximiert werden und für Selektionsprozesse korrigiert werden.
- Ein **kluges Programmdesign** ist dennoch überlegen, denn es ermöglicht:
 - a) Verlässliche und statistisch **signifikante Resultate** auch bei einer relativ überschaubaren Zahl von Teilnehmenden bzw. Beobachtungen.
 - b) Glaubhafte und **überzeugende Kommunikation** an Entscheider und Verantwortliche, wenn die Methodik leicht nachvollziehbar ist.
 - c) Antworten auf tatsächlich **relevante Fragen**.

Zusätzliche Herausforderung

- Ein kausaler Teilnahmeeffekt betrachtet ausschließlich Wirkung einer Maßnahme auf die Teilnehmenden und damit lediglich die **mikroökonomische Ebene**.
- Darüber hinaus kann es auf **makroökonomischer Ebene** weitere wichtige Effekte geben (z.B. Substitutionseffekte, Verdrängungseffekte und Wohlfahrtseffekte), die die **Gesamtbilanz einer Maßnahme** substantziell beeinflussen können.
 - Es ist sehr unwahrscheinlich, dass bei negativen nicht signifikanten kausalen Teilnahmeeffekten die Gesamtbilanz einer Maßnahme positiv ausfällt.
 - Umgekehrt kann die Gesamtbilanz einer Maßnahme z.B. aufgrund von starken Verdrängungseffekten auch bei positiven Teilnahmeeffekten noch negativ ausfallen.

Praxisbeispiel:
„Beautiful Serbia“

„Beautiful Serbia“ im Überblick

- **Eine der ersten Maßnahmen der aktiven Arbeitsmarktpolitik**, die in Serbien vor dem Hintergrund hoher Arbeitslosigkeit implementiert wurde (im Zeitraum 2004/2005).
- Das Programm war eine **Kombination von Aus- bzw. Weiterbildung und anschließender temporärer Beschäftigung** im Bausektor.
- **Forschungsziel:** Ermittlung der kausalen Effekte auf a) den Erwerbsstatus und b) das subjektive Wohlbefinden der Teilnehmenden.
- **Methodik:** Vergleich von Teilnehmern und Nichtteilnehmern, die sich hinsichtlich beobachtbarer Charakteristika nicht (mehr) unterscheiden; Anwendung eines Matching- bzw. Vergleichsgruppenansatzes bereits bei der Datenerhebung.

Schwächen im Evaluationsdesign

- **Fehlgeschlagener Ansatz**, eine spätere Evaluation schon bei der Datenerhebung zu berücksichtigen: Teilnehmende und (intendierte) Kontrollgruppe weisen sehr unterschiedliche beobachtete Charakteristika auf.
- Vergleichsweise **geringe Fallzahlen in der Kontrollgruppe** erschweren die (ex post) Herstellung eines quasi-experimentelles Designs durch ökonometrische Methoden.
- **Insgesamt geringe Fallzahlen** führen zu Ergebnissen, die eine geringe Belastbarkeit und eine hohe statistische Unsicherheit aufweisen.
- Erhebung der **Daten mit geringen zeitlichem Abstand** vor und nach der Programmteilnahme erschweren die Analyse (Erwerbshistorie? Langfristige Effekte?).

Aus Fehlern lernen...

RICHTIGE ANSÄTZE:

- ✓ Spätere Evaluation des Programms bereits vor Programmbeginn (d.h. schon bei der Projektgestaltung und -konzeption sowie bei der Datenerhebung) berücksichtigen.
- ✓ Bei Unsicherheit über Teilnahmeeffekte Pilotprojekte mit geringer Teilnehmerzahl durchführen und anschließend evaluieren.

ABER MANGELHAFTE UMSETZUNG:

- ✗ Herstellung eines experimentellen Designs ist misslungen.
- ✗ Hohe Datenanforderungen, die ein Vergleichsgruppenansatz stellt, nur unzureichend erfüllt.
- ✗ Zu kurzer Zeithorizont der Daten (vor und nach Programmteilnahme).
- ✗ Geringe Teilnehmerzahlen von Pilotprojekten machen soziale Experimente erforderlich, um dennoch belastbare Ergebnisse zu erhalten.

Fazit

Abschließende Bemerkungen

- Ein „**evidenzorientiertes**“ **Programmdesign** (das sinnvolle Überlegungen beinhaltet, um die kontrafaktische Situation in geeigneter Weise zu approximieren) sollte ein **zentrales Kriterium** bei Entscheidungen über die Finanzierung bzw. Durchführung von Maßnahmen und Programmen sein.
- Der **Erfolg einer Maßnahme** lässt sich nicht mit Inputgrößen (z.B. Mittelabfluss) oder nur allein durch Ergebnisvariablen der Teilnehmer messen – um kausale Effekte zu ermitteln, ist ein Vergleich mit der **kontrafaktischen Situation** erforderlich.
- Spart man zu Beginn an einem klugen Programmdesign, spart man am falschen Ende: Die **Reparaturkosten übersteigen die Anfangersparnis um ein Vielfaches** – und eventuell ist eine Reparatur (mit ökonometrischen Methoden) nicht möglich.



Dr. Ulf Rinne
rinne@iza.org

www.iza.org