Andrea Pufahl

Johann Heinrich von Thünen-Institut
Institut für Ländliche Räume

Quantitative Wirkungsanalysen – Methodischer Überblick

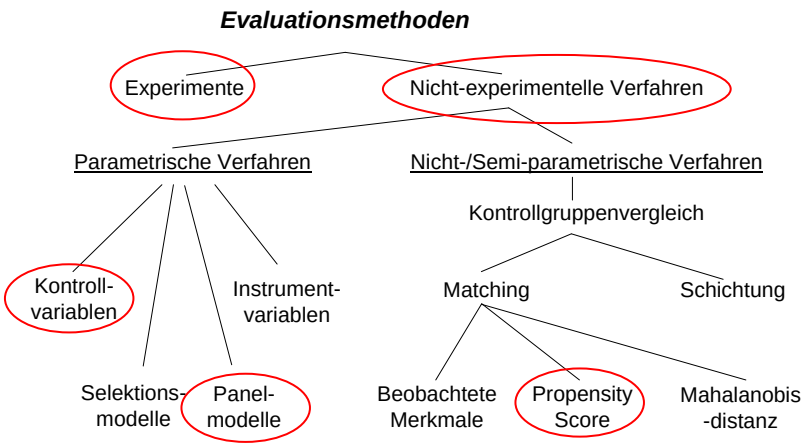
und Anwendung auf die Evaluation von Maßnahmen zur Entwicklung des ländlichen Raumes

Frühjahrsworkshop des Arbeitskreises Strukturpolitik der DeGEval
18./19. Juni 2009 in Wiesbaden

Überblick

1. Evaluationsmethoden im Überblick
2. Regression mit Kontrollvariablen
3. DiD im Panelmodell
4. Propensity Score Matching
5. Fazit

Evaluationsmethoden im Überblick



Experimente

- „Gold Standard“
- Teilnehmer und Nichtteilnehmer werden zufällig ausgewählt
- Zufällige Auswahl sichert (im Mittel) eine Gleichverteilung der (un-)beobachteten Merkmale (X)
- Potenzielles Nichtteilnahmeergebnis ist in beiden Gruppen gleich
- ***Kein Selektionseffekt !***

Durchschnittlicher Effekt auf die Teilnehmer = Ergebnis Teilnehmer – Ergebnis Nichtteilnehmer

- Pro: Teilnahme und andere Merkmale werden direkt kontrolliert
- Contra: Evaluator kann Bias verursachen, Verallgemeinerung der Ergebnisse fragwürdig

Nicht-experimentelle Verfahren - Parametrische vs. nicht-parametrische

	Parametrisch	Nicht-/Semi-parametrisch
Praktisches Handling	mittel - schwer	einfach
Datenbedarf	mittel	sehr hoch
Transparenz	mittel	hoch
Kommunizierbar	schwer (Koeffizienten)	einfach (Mittelwertvergleiche)
Annahmen		
- Verteilungen	hoch	keine / gering
- Programmeffekte	hoch (homogene Effekte)	keine (heterogene Effekte)

Regression mit Kontrollvariablen

Strukturelles Modell

i = Index Beobachtung

$$y_i = x_i \beta + d_i \delta + \varepsilon_i$$

$$Ergebnis_i = \underbrace{Alter_i \beta_1 + Quali \beta_2}_{\text{Kontrollvariablen}} + Programm_1 \delta + Fehler_i$$

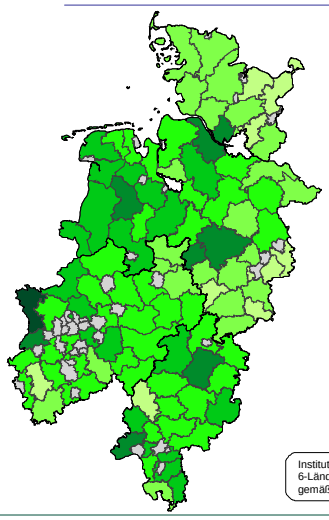
Effekt der Programmteilnahme

- Programmteilnahme kann binär oder stetig skaliert sein
- Auswahl Kontrollvariablen, funktionale Spezifikation sollte theoriebasiert erfolgen
- Pro: datenextensiv
- Contra: starke Annahmen, *selection on observables*

Beispiel Regression mit Kontrollvariablen – Untersuchung der Bestimmungsfaktoren für die räumliche Verteilung sektorbezogener Maßnahmen

Euro je ha LF

- <= 25
- 25 <= 50
- 50 <= 100
- 100 <= 150
- 150 <= 500
- > 500
- krfr. Städte



Öffentliche Ausgaben (2000 - 2006) Sektorbezogene Maßnahmen (AFP, VV, Beratung, Qualität etc.)

Datengrundlage: Zahlstellendaten

Institut für Ländliche Räume des vTI
6-Länder-Ex-Post-Bewertung
gemäß VO (EG) Nr. 1257/1999

Evaluationsmethoden
Regression
Panelmodel
Matching
Fazit

Beispiel Regression mit Kontrollvariablen Ausgewählte Bestimmungsgründe der Mittelverteilung

Sektorbezogene Maßnahmen (AFP, VV, Beratung, Qualität etc.)

		SH	NI	NW	HE
Landwirtschaftliche Betriebe	Anzahl	n.s.	n.s.	-	n.s.
EMZ	Index		-	-	n.s.
Änderung Anzahl ldw. Betriebe	%	-	n.s.	n.s.	+
Anteil HE-Betriebe	%		+	+	n.s.
Anteil Gartenbau- / Dauerkulturbetriebe	%	+	+		
Besatzdichte Rind	GV/ha	+	n.s.	n.s.	n.s.
Besatzdichte Schwein	GV/ha		+		+
BWS Landwirtschaft	1000 Euro/Erwerbst.		n.s.	+	-
Herdengröße	n/Betrieb			-	
Kreise Viersen, Kleve	0 = nein, 1 = ja			+	
Anzahl Beobachtungen		11	37	31	20
R-Quadrat		0,98	0,76	0,79	0,76

n.s. = nicht signifikant

Andrea Pufahl

Evaluationsmethoden

Regression

Panelmodel

Matching

Fazit

Panelmodelle (Fixed-Effekt Modell)

i = Index Beobachtung
t = Index Zeitpunkt

Strukturelles Modell

$$y_{it} = x_{it}\beta + d_{it}\delta + \varepsilon_{it}$$

Transformiertes Modell (Differenz-in-Differenz)

Mittelwert über alle t

$$(y_{it} - \bar{y}_i) = (x_{it} - \bar{x}_i)\beta + (d_{it} - \bar{d}_i)\delta + \varepsilon_{it} - \bar{\varepsilon}_i$$

Effekt der Programmteilnahme

- Schätzung basiert auf jährlichen Abweichungen vom mehrjährigen Mittel
- Programmteilnahme: binär oder stetig skaliert und zeitvariabel
- Konstante Unterschiede kürzen sich heraus

- Pro: *selection on unobservables*
- Contra: mittlerer Datenbedarf (Panels), hohe Annahmen

Andrea Pufahl

Evaluationsmethoden

Regression

Panelmodel

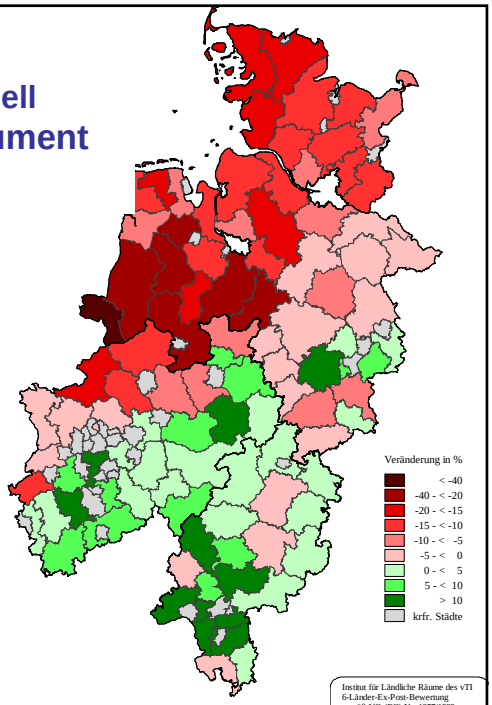
Matching

Fazit

Beispiel DiD-Panelmodell
Der EPLR – ein Instrument zum Grünlanderhalt?
Entwicklung (1999-2005)

	Grünland* / Grünlandanteil*
SH	**-13,4 % / -10,9 %**
NI	**-10,1 % / -9,4 %**
NRW	**1,4 % / 0,1 %**
HE	**8,2 % / 6,5 %**

*** Mittelwerte für Kreise, ohne kreisfreie Städte**



Andrea Pufahl

Beispiel DiD-Panelmodell
Einfluss ausgewählter Maßnahmen auf den Grünlanderhalt (2000-2005)

		Alle Länder	SH	NI	NW	HE
Bodennutzung und Agrarstruktur						
LF	ha	0,61 ***	1,22 ***	0,32	0,40 **	0,63 **
LF in HE-Betrieben	ha	0,02	-0,54 ***	-0,07	-0,05	0,02
Maisfläche	ha	-0,21	-0,30	-0,14	0,47 *	-0,34
Rapsfläche	ha	-0,005	0,42	-0,24	-0,08 **	0,49
Herdengröße Rind	GV/Betrieb	-72,84	22,80	159,90	-40,65	18,80
Rinderbestand	GV	0,33 ***	0,51 ***	0,54 ***	0,01	0,05
BWS Landwirtschaft (Vorjahr)	Euro	-0,03 *	-0,03	-0,08 ***	0,00	0,01
EPLR-Förderung						
Agrarinvestitionsförderung	1000 Euro	-0,37 *	-1,42 *	-0,53 **	-0,02	-0,02
Ausgleichszulage/-zahlung	1000 Euro	-0,01	-0,09	5,32	-0,37	-0,12
Agrarumweltmaßnahmen	1000 Euro	0,07	-2,64	0,65	0,17 **	-0,20
Beobachtungen	n	300	33	113	93	61
Anzahl Landkreise	n	101	11	38	31	21
R-Quadrat		0,64	0,96	0,80	0,68	0,79

Signifikanzniveau *** 1 %, ** 5 % und * 10 %.

Propensity Score Matching (1)

Intuitive Interpretation

- Matching = Gruppieren ähnlicher Teilnehmer & Nichtteilnehmer
- Propensity Score = Aggregiertes Maß beobachteter Merkmale

Definition Propensity Score

- Bedingte Wahrscheinlichkeit d. Teilnahme, bei gegebenen Merkmalen *X*
- Theoriebasiert: Auswahl relevanter Merkmale
- Gleicher Propensity Score = gleiche beobachtete Merkmale
- Matching auf Propensity Score "balanciert" Merkmalsunterschiede aus

- Pro: Kontrolle vieler Merkmale, gut kommunizierbar, wenig Annahmen
- Contra: datenintensiv, Annahmen tlw. nicht überprüfbar, *selection on observables*

Propensity Score Matching (2)

Vorher-Nachher-Vergleich

$$ATT = \underbrace{E\left(Y^1_{i,2000} \mid P_i = 1, p(\boldsymbol{X})\right)}_{\text{Ergebnis Teilnehmer}} - \underbrace{E\left(Y^0_{j,2000} \mid P_j = 0, p(\boldsymbol{X})\right)}_{\text{Ergebnis Nichtteilnehmer}}$$

Mit-Ohne/Vorher-Nachher-Vergleich (Differenz-in-Differenz)

$$ATT = E\left(\underbrace{Y^1_{i,2005} - Y^0_{i,2000}}_{\Delta \bar{Y}^1_i} \mid P_i = 1, p(\boldsymbol{X})\right) - E\left(\underbrace{Y^0_{j,2005} - Y^0_{j,2000}}_{\Delta \bar{Y}^0_j} \mid P_j = 0, p(\boldsymbol{X})\right)$$

$$ATT = \underbrace{\Delta \bar{Y}^1_i}_{\text{Ergebnisänderung Teilnehmer}} - \underbrace{\Delta \bar{Y}^0_j}_{\text{Ergebnisänderung Nichtteilnehmer}}$$

- Kontrolliert den Einfluss <u>unbeobachteter</u> Faktoren auf das Ergebnis nach der Auswahl vergleichbarer Teilnehmer & Nichtteilnehmer

Evaluationsmethoden
Regression
Panelmodel
Matching
Fazit

Beispiel Propensity Score Matching
Wirkung von AUM und AZ auf die Beschäftigung

Maßnahmen

- Agrarumweltmaßnahmen (z.B. Grünlandextensivierung etc.)
- Ausgleichszulage für benachteiligte Gebiete

Daten

- Buchführungsdaten ldw. Betriebe, 2000 bis 2005
- nicht repräsentativ
- Angaben zu Einkommen, Arbeitskräften etc.

Methode

- Propensity Score Matching (Caliper Matching)
- Differenz-in-Differenz-Vergleich (2000,2005)

Andrea Pufahl

Beispiel Propensity Score Matching
Wirkung von AUM auf die Beschäftigung
Average Treatment Effect on Treated (ATT), 2000-2005

Region		Agrarumweltmaßnahmen				
Kennwerte Beschäftigung	Einheit	Teilnehmer	Kontrollbetriebe	ATT	t-Wert	d
Hessen		n=68	n=68			
AK/Betrieb	AKE	-0,16	0,00	-0,17	-2,78 ***	-0,5
AK/100 ha LF	AKE	-0,42	0,02	-0,44	-3,14 ***	-0,5
Außerbetr. AK	AKE	0,03	0,04	-0,01	-0,28	0,0
Nordrhein-Westfalen		n=695	n=695			
AK/Betrieb	AKE	0,38	0,03	0,35	0,80	0,0
AK/100 ha LF	AKE	0,20	0,00	0,20	0,30	0,0
Außerbetr. AK	AKE	0,00	0,00	0,00	0,83	0,0
Niedersachsen		n=841	n=841			
AK/Betrieb	AKE	0,01	0,43	-0,41	0,91	0,0
AK/100 ha LF	AKE	-0,10	0,39	-0,49	-1,21	-0,1
Außerbetr. AK	AKE	0,00	0,01	-0,01	-0,55	0,0

Hinweise: AKE = Arbeitskrafteinheiten, LF = Landwirtschaftliche Fläche. Signifikanz der Mittelwerte (T-Test) auf einem Niveau von 1 % (***). Effektstärke (d): klein (< 0,2), mittel (>= 0,2 < 0,8), groß (>= 0,8).

Fazit

- Schätzung der *„counterfactual“* Situation
- Kontrolle beobachteter und unbeobachteter Faktoren
- Best: Differenz-in-Differenz-Schätzungen mit Paneldaten
- Second best: Querschnittsanalysen

- Datenverfügbarkeit determiniert meist zu verwendendes Verfahren
  - Stetige Programmteilnahme (Euro Förderung) > parametrisch
  - Binäre Programmteilnahme (0=nein, 1=ja) > nicht- und parametrisch
  - Wenig Beobachtungen (N < 400) > parametrisch
  - Viele Beobachtungen (N > 400) > nicht- und parametrisch
  - Verhältnis der Beobachtungen von Teilnehmern & Nichtteiln. ist gleich bzw. gering (1:1 … 1:4) > parametrisch

Literatur

- Rubin, D. (1974): Estimating Causal Effects of Treatments in Randominzed and Nonrandomized Studies. Journal of Educational Psychology 66: 688-701

Methoden der Wirkungsanalyse im Überblick

- Imbens, G. und Wooldridge, J. (2007): What's new in econometrics. NBER Summer Institute Course notes. (Skript, Präsentation, Video)
- Heckman, J. und Robb, R. (1985): Alternative Methods for Evaluating the Impact of Interventions. In: Heckman, J. und Singer, B. (Hrsg.): Longitudinal Analysis of Labor Market Data. New York (Cambridge University Press). S. 156-245.

Propensity Score Matching

- Heckman, J., Ichimura, H. und Todd, P. E. (1998): Matching as an econometric evaluation estimator. Review of Economic Studies 65, S. 261-294.
- Reinowski, E. (2006): Mikroökonomische Evaluation und das Selektionsproblem. Zeitschrift für Evaluation 2006, H. 2, S. 187-226. (deutsch)

Panelanalysen

- Wooldridge, J. M. (2002): Econometric Analysis of Cross Section and Panel Data. Cambridge (MIT Press).