



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

Theoretische Grundlagen und die Rolle parametrischer bzw. nicht-parametrischer Modellansätze in der Politikevaluation

Replik zu Anne Margarian, Johann Heinrich von Thünen-Institut, Braunschweig

Christian H.C.A. Henning und Jerzy Michalek

Christian-Albrechts-Universität Kiel

Der Kommentar von Frau Margarian zu unserem Beitrag „Ökonometrische Methoden der Politikevaluation: Meilenstein für eine sinnvolle Agrarpolitik der 2. Säule oder akademische Fingerübung?“, *Agrarwirtschaft* 3/4, Jahrgang 57 (2008) spiegelt das große Interesse an einer konsistenten Evaluation agrarpolitischer Maßnahmen gerade auch von Seiten der praktischen Politikevaluatoren, wie dem Johann von Thünen-Institut wider. Dieses Interesse an unserer Arbeit begrüßen wir sehr. Gleichzeitig zeigt der Kommentar von Frau Margarian aber auch klar auf, dass eine konsistente Politikevaluation ein komplexes methodisches Thema ist, das selbst von erfahrenen Evaluatoren nicht einfach erschlossen werden kann. Umso mehr freuen wir uns über die zusätzliche Gelegenheit, einige grundlegende Zusammenhänge hinsichtlich der theoretischen und methodischen Grundlagen einer konsistenten quantitativen Politikevaluation noch einmal verdeutlichen zu können.

Konkret wollen wir noch einmal explizit auf die Bedeutung von grundlegenden theoretischen ökonomischen Zusammenhängen im Rahmen der Politikevaluation und der Anwendung parametrischer und nicht-parametrischer Ansätze sowie auf die Unterscheidung von Methoden der Ex-post- und Ex-ante-Evaluation eingehen. Dies erscheint uns notwendig, da hier offensichtlich noch erhebliche Lücken zwischen dem Wissensstand der praktischen Evaluatoren und dem aktuellen Stand der internationalen Evaluationsforschung bestehen. Dies belegt einmal mehr, wie sinnvoll und wichtig politikorientierte EU-Forschungsprojekte wie „Advanced-Eval“ für eine effektive Vermittlung innovativer ökonomischer Methoden und Theorien sind.

Der zentrale Kritikpunkt des Kommentars von Frau Margarian stützt sich auf ihre Interpretation der mikro-ökonometrischen Evaluationsmethoden als „nicht theoriebasierte Evaluationsmethoden“. Frau Margarian geht in ihrem Kommentar explizit von der Annahme aus, dass die Anwendung eines nicht-parametrischen Ansatzes dem Paradigma einer nicht theoriebasierten Evaluation entspricht. Diese Annahme ist eindeutig falsch. Ein Paradebeispiel können hier Engelkurven bzw. Kostenfunktionskurven sein, die parametrisch oder aber nicht-parametrisch geschätzt werden können, wobei das theoretische Konstrukt der Engelkurve bzw. der Kostenfunktionskurven selbstverständlich vollkommen unabhängig von der Schätztechnik bleibt. Darüber hinaus ist jede Theorie, so auch die ökonomische Theorie, nicht auf explizit spezifizierte formale mathematische Modelle beschränkt, sondern theoretische Zusammenhänge können durchaus auch mit Hilfe von nicht voll spezifizierten funktionalen Zusammenhängen bzw. im Extremfall sogar allein mit Hilfe einfacher Wenn-Dann-Aussagen dargestellt werden (ESSER et al., 1977).

Um die Rolle parametrischer und nicht-parametrischer Ansätze sowie die Rolle der ökonomischen Theorie in dem konkreten Fall der Ex-post-Evaluierung noch einmal unmissverständlich darzulegen, werden wir im Folgenden einige ergänzende Ausführungen anführen.

Das zentrale Ziel einer Ex-post-Evaluierung ist die quantitative Ermittlung des Effektes eines Politikprogramms auf spezielle Performance-Indikatoren Y , d.h. eine Ex-post-Evaluierung entspricht der Ermittlung der Veränderung eines (oder mehrerer) Performance-Indikatoren (-en), die ursächlich durch die Teilnahme ($D=1$) an dem Politikprogramm impliziert wurde:

$$(1) \quad \Delta Y = Y^1 - Y^0$$

Grundsätzlich lässt sich die Wirkung eines Politikprogramms auf der Grundlage entsprechender theoretischer Zusammenhänge erklären. Unterstellt man im Sinne von Tinbergens quantitativer Wirtschaftspolitik ein quantitatives Erklärungsmodell, ergibt sich formal:

$$(2) \quad F(Y, X, D) \equiv 0$$

Im allgemeinen Fall entspricht F einer Vektorfunktion, die einen funktionalen Zusammenhang zwischen den Performance-Indikatoren (Y) und der Programmteilnahme ($D=1$) bzw. Nichtteilnahme ($D=0$) sowie weiterer relevanter Faktoren (X) herstellt. Der in der Funktion F ausgedrückte Zusammenhang ist Gegenstand der ökonomischen Theorie. Dabei ist es ohne Frage ein Ziel der ökonomischen Forschung, möglichst viele ökonomische Phänomene möglichst detailliert zu erklären, d.h. den funktionalen Zusammenhang F sowie die relevanten Erklärungsfaktoren X möglichst exakt zu spezifizieren. Kennt man die relevanten Faktoren X und hat man die Funktion F spezifiziert, so ließe sich die Wirkung des Politikprogramms (ΔY) exakt bestimmen:

$$\Delta Y = Y^1 - Y^0$$

$$(3) \quad F(Y^1, X, D=1) = 0$$

$$F(Y^0, X, D=0) = 0$$

Eine quantitative Ex-post-Evaluierung ist also zweifelsohne auf der Grundlage von parametrischen Ansätzen möglich, solange relevante funktionale Zusammenhänge spezifiziert sind. Nun ist es aber leider eine offensichtliche Tatsache, dass die ökonomische Theorie für sehr viele empirische Phänomene noch keine unstrittigen und eindeutig spezifizierten funktionalen Zusammenhängen F liefert. Im Gegenteil; oft sind noch nicht einmal alle relevanten Erklärungsfaktoren unstrittig identifiziert bzw. es sind erst wenige Eigenschaften der Funktion F bekannt. Tatsächlich gibt es für die Mehrzahl realer wirtschaftlicher Phänomene immer noch konkurrierende theoretische Ansätze, d.h. formal gibt es unterschiedliche Funktionen F und F' , deren empirische Relevanz bei dem gegebenen Stand der Forschung nicht eindeutig bestimmt werden kann. Der letzte Punkt gilt selbst für vermeintlich einfache und bereits relativ gut erforschte wirtschaftliche Phänomene wie mikroökonomisches Nachfrage- bzw. Angebotsverhalten. Es gibt eine Reihe unterschiedlicher mikroökonomischer Ansätze zur

Ableitung der relevanten Verhaltensgleichungen, die formal jeweils einer anderen Spezifikation der Funktion F in Gl. (3) entsprechen und somit in der Regel zu einer anderen quantitativen Ex-post-Evaluierung des gleichen Politikprogramms bei gleicher empirischer Datengrundlage führen würden. Also selbst wenn wir unterstellen, dass ein Phänomen bereits sehr gut theoretisch erforscht ist, folgt, dass eine quantitative Ex-post-Evaluierung auf der Grundlage spezifizierter parametrischer Modellansätze keine eindeutigen Ergebnisse hervorbringt, solange letztere in der ökonomischen Theorie nicht eindeutig und unstrittig identifiziert worden sind. Besonders problematisch erscheint eine quantitative Ex-post-Evaluierung auf der Grundlage spezifizierter Modellansätze, wenn konkurrierende Modellansätze zu widersprüchlichen Ergebnissen führen.

Es ist unstrittig, dass viele Phänomene der ländlichen Entwicklung, wie z.B. die exakte Abbildung der Wirkung von staatlichen Beratungsstellen auf die Innovationsdiffusion bzw. technischen „Spillover-Effekte, bis heute noch nicht eindeutig quantitativ modellierbar sind. Im Klartext: Für viele ländliche Entwicklungspolitiken ist der konkrete funktionale Zusammenhang $F(Y, X, D=1)$ nicht bekannt. Das heißt nicht, dass es überhaupt keine theoretischen Grundlagen gibt. Beispielsweise können ja bestimmte Zusammenhänge zwischen speziellen Faktoren (X) und Performance-Indikatoren durchaus theoretisch identifiziert worden sein, aber die exakten und kompletten funktionalen Zusammenhänge sind noch nicht bekannt, somit ist eine quantitative Evaluierung auf der Grundlage eines voll spezifizierten parametrischen Ansatzes nicht möglich.

In diesem Fall ist die Anwendung eines nicht-parametrischen Ansatzes, wie der vorgestellte Propensity-Score-Ansatz (PSM), eine vielversprechende Alternative, da auf der Grundlage dieses Ansatzes die Wirkung eines Politikprogramms auf spezielle Performance-Indikatoren quantitativ ermittelt werden kann, ohne eine spezielle Funktionsform oder gar Spezifikation zu unterstellen. Gerade dies ist der Vorteil dieser Methode, die selbstverständlich neben statistischer Theorie immer auch explizit auf ökonomischer Theorie aufbaut, da die entsprechenden Kovariaten X auf der Grundlage entsprechender ökonomischer Theorie identifiziert werden (müssen) (CALIENDO, 2006). Entsprechend wird PSM als eine innovative und sinnvolle Ex-post-Evaluierungsmethode in der Evaluationsforschung von führenden Ökonomen wie u.a. dem Nobelpreisträger Heckmann favorisiert.¹

¹ Der Vollständigkeit halber soll an dieser Stelle auch auf entsprechende parametrische Regressionsansätze hingewiesen werden, die entsprechende theoretische Zusammenhänge mit Hilfe von Verhaltensgleichungen in reduzierter Form, in der Regel werden lineare Funktionsformen angenommen, abbilden, die als lokale Approximation der unbekannteren tatsächlichen funktionalen Zusammenhänge $F(y, x)$ interpretiert werden. In der Regel wird in den parametrischen Regressionsansätzen die Teilnahme bzw. Nichtteilnahme an einem Politikprogramm mit Hilfe einer Dummy-Variable abgebildet. Die Vor- und Nachteile von PSM gegenüber diesen parametrischen Regressionsansätzen werden in der entsprechenden Literatur ebenfalls ausgiebig diskutiert (CALIENDO, 2006). Gemeinsam ist aber beiden Ansätzen, dass der Wirkungszusammenhang des Politikprogramms auf die relevanten Performance-Indikatoren nicht komplett spezifiziert wird.

Zusammenfassend können wir festhalten: (1) Mikro-ökonomische Evaluierungsmethoden basieren eindeutig auf der aktuellen ökonomischen wie auch statistischen und ökonomischen Theorie, allerdings erlauben diese eine quantitative Ex-post-Evaluierung ohne eine komplett spezifizierte Funktionsform für die jeweils relevanten ökonomischen Verhaltensgleichungen anzunehmen. (2) Ohne Frage reflektiert die Unkenntnis der konkreten Spezifikation der funktionalen Zusammenhänge (F) ein Theoriedefizit, was den Bedarf an zukünftigen Forschungsaktivitäten in diesem Bereich signalisiert. Allerdings erreichen nicht-parametrische Ex-post-Evaluierungsmethoden, wie das PSM, gerade vor dem Hintergrund dieses Theoriedefizits ihre eigentliche Bedeutung, da sie eine theoretisch fundierte quantitative Politikevaluation trotz des bestehenden Defizits ermöglichen. (3) Ließe sich die Funktion F eindeutig spezifizieren, so wäre eine Ex-post-Evaluierung im Grunde überflüssig.

Neben diesem oben erläuterten fundamentalen Missverständnis enthält der Kommentar von Frau Margarian weitere falsche Aussagen hinsichtlich der Annahmen der PSM-Methode. Unter anderem ist es falsch, dass PSM lineare Beziehungen zwischen den Einflussgrößen unterstellt. Fakt ist, dass keine spezielle Funktionsform unterstellt wird. Ebenso ist es falsch, dass die PSM-Methode lediglich eine Evaluationsaussage in der Form „wirkt bzw. wirkt nicht“ zulässt. Tatsache ist, dass diese Methode eine quantitative Bestimmung des Effektes eines Politikprogramms auf unterschiedliche Performance-Indikatoren liefert. Weiterhin ist PSM gerade für komplexe Zusammenhänge auf der Aggregatebene, z.B. die Wirkung von Politikprogrammen auf regionale Entwicklungsindikatoren, eine hervorragend geeignete Methode zur Ex-post-Evaluierung, da diese komplexe dynamische Wechselwirkungen zwischen individuellen Wirtschaftseinheiten, die mit expliziten Modellen sehr schwer zu erfassen sind, implizit mit berücksichtigt. Selbstverständlich muss hierbei dann ein Matching über regionale Einheiten, z.B. Kreise, erfolgen. Entsprechende Verfahren wurden in der Literatur wie auch im Rahmen unseres EU-Projektes entwickelt und empirisch angewendet (VAN DE WALLE and CRATTY, 2002; LOKSHIN and YEMTSOV, 2005; MICHALEK, 2008).²

Ein weiteres Missverständnis, welches sich in dem Kommentar von Frau Margarian abzeichnet, bezieht sich auf die spezifische Unterscheidung von Ex-post- und Ex-Ante Politikevaluierung. Frau Margarian scheint in ihrem Kommentar die spezifischen Ansprüche an Evaluationsmethoden, welche sich für eine Ex-post- bzw. Ex-ante-Evaluierung ergeben, zu vermischen. Während Ex-post-Evaluierung zentral auf die quantitative Ermittlung des Effektes eines bereits implementierten Politikprogramms auf spezielle Performance-Indikatoren abzielt, das bereits aus einer unterstellten Interventionslogik abgeleitet worden ist, ist die Zielsetzung der Ex-ante-Evaluierung weiter gefasst. Insbesondere umfasst diese die explizite Ableitung der relevanten Interventionslogik eines Politikprogramms. Unser me-

² Es ist offensichtlich, dass ein komplett spezifizierter Ansatz, der alle dynamischen Interaktionseffekte in beliebigen ländlichen Regionen explizit quantitativ modelliert, nicht existiert bzw. nur unter extrem hohen Aufwendungen von Forschungsressourcen entwickelt werden könnte, so dass PSM gerade für eine routinemäßige Ex-post-Evaluierung eine sinnvolle und vielversprechende Methode darstellt.

thodischer Beitrag bezog sich eindeutig und ausschließlich auf PSM als Ex-post-Evaluierungsmethode, während entsprechende methodische Verfahren zur Ex-ante-Evaluation in dem Beitrag nicht explizit abgehandelt werden. Ohne Frage ist in der politischen Praxis sowohl die Ex-ante- wie auch eine Ex-post-Evaluierung von Politiken relevant, wobei sich die explizite Zielsetzung der Ex-post- im Vergleich zu einer Ex-ante-Evaluierung klar unterscheiden. Dies hat entsprechende methodische Implikationen. Es würde allerdings den Rahmen dieser Replik sprengen, diese Implikationen im Einzelnen abzuhandeln. Der interessierte Leser wird hier auf die entsprechende Literatur verwiesen. Dies sind u.a. MARSCHAK (1953), TODD and WOLPIN (2005) wie auch die Leitlinien zur Evaluation der EUROPÄISCHEN KOMMISSION (EC, 1999, 2001, 2006). Herausstellen möchten wir in diesem Zusammenhang aber, dass gerade vor dem Hintergrund der Tatsache, dass die ökonomische Theorie für viele Phänomene der ländlichen Entwicklung eben noch keine konsistenten, komplett spezifizierten ökonomischen Verhaltensmodelle bereitstellen kann und somit die jeweils einem Politikprogramm ex ante unterstellte Interventionslogik grundsätzlich unsicher ist, nicht-parametrischen Ex-post-Evaluationsmethoden eine besondere Schlüsselrolle zukommt, da diese eine von der unterstellten Interventionslogik unabhängige Überprüfung der Politikwirkung erlauben. Gerade eine solche quantitative Überprüfung von proklamierten Politikeffekten trägt nicht nur zu einem beschleunigten Erlernen effektiver wirtschaftspolitischer Maßnahmen (Policy-Learning) bei, sondern schafft auch eine entsprechende Transparenz hinsichtlich der Wirkung konkreter Politikmaßnahmen, die eine notwendige Voraussetzung für eine effektive Kontrolle administrativer bzw. politischer Handlungen durch die Gesellschaft darstellt.

Abschließend möchten wir noch einmal explizit darauf hinweisen, dass obwohl es nach unserer Einschätzung unstrittig ist, dass existierende mikro-ökonomische Evaluierungstechniken einen erheblichen Beitrag zu einer effektiven Ex-post-Politikevaluation leisten, sich selbstverständlich auch für diese Ansätze durchaus auch einige methodisch-theoretische Schwachstellen und Verbesserungspotentiale ergeben.

Beispielsweise wäre es auch im Rahmen einer Ex-post-Evaluation wünschenswert, nicht nur den Gesamteffekt eines Programms zu identifizieren, sondern darüber hinaus den konkreten Wirkungsmechanismus, d.h. die komplette Programmlogik im Sinne der Transformation der Programm-Inputs in entsprechende Programm-Outputs und daraus resultierende Programm-Outcomes, eindeutig zu erklären. Formal bedeutet dies, dass man die Wirkung eines Politikprogramms auf relevante Performance-Indikatoren in zwei Stufen zerlegen müsste, die Wirkung der Programm-Inputs (I) auf relevante Bestimmungsfaktoren der Performance X und die Wirkung veränderter Bestimmungsfaktoren auf die Performance: $\Delta Y = f(\Delta X(I))$. Auf diese Weise könnte man explizit verstehen, warum ein Programm nicht oder nicht in dem erwarteten Maße gewirkt hat. Die Ermittlung der kompletten Programmlogik kann mit Hilfe der diskutierten Matching-Verfahren wie auch parametrischen Regressionsverfahren in der Regel nicht geleistet werden. Hier besteht also ohne Frage noch Forschungsbedarf. Dies wird in der relevanten Evaluationsliteratur auch offen diskutiert, wobei in der aktuellen Literatur allerdings eine „program-

theory-based evaluation“ gegenüber der veralteten „social-science-theory-based evaluation“ favorisiert wird (Der Interessierte Leser sei hier auf CHEN (1990), BRICKMAYER and WEISS (2000), DONALDSON (2001, 2003, 2007), DONALDSON and LIPSEY (2006) verwiesen.). Entsprechend ist der Einsatz von Matching-Verfahren zur Ableitung der vollständigen Programm-Interventionslogik nur bedingt einsetzbar. Hier greift der Kommentar von Frau Margarian also durchaus einen relevanten Punkt auf. Klar widersprechen müssen wir allerdings ihrer Behauptung, dass eine Politikevaluation immer auch die Gültigkeit der dieser Evaluation zugrunde gelegten Theorie überprüft. Hier dreht sich die Autorin nun im Kreis: Eine fundierte Evaluation basiert immer auf entsprechenden theoretischen Annahmen, insofern ist es grundsätzlich unmöglich, simultan auf der Grundlage dieser Annahmen Politikeffekte zu ermitteln und diese Annahmen zu überprüfen (oder gar abzulehnen!). Wie WEISS (1997) schrieb, „The hope that evaluation will be able to say with any assurance that one theory fits the data seems destined for disappointment“. Gerade deshalb ist es für die breite Akzeptanz einer Evaluationsmethode von Vorteil, wenn diese auf theoretischen Annahmen basiert, die allgemein anerkannt sind. Hier ergeben sich wie bereits oben dargestellt in der Regel Vorteile für nicht-parametrische Matching-Verfahren, da diese im Vergleich zu parametrischen Verfahren von deutlich schwächeren theoretischen Annahmen ausgehen.

Literatur

- BRICKMAYER, J. and C.H. WEISS (2000): Theory-based Evaluation in Practice. What Do We Learn?. In: Evaluation Review 24 (4): 407-431.
- CALIENDO, M. (2006): Microeconomic evaluation of labor market policies. Springer Verlag, Berlin.
- CHEN, H. (1990): Theory-driven evaluations. Sage, Newbury Park.
- CHEN, H. and P. ROSSI (1987): The theory-driven approach to validity. In: Evaluation and Program Planning 10 (1): 95-103.
- DONALDSON, S.I. (2001): Overcoming our negative reputation: Evaluation becomes known as a helping profession. In: American Journal of Evaluation 22 (3): 355-361.
- (2003): Theory-driven program evaluation in the new millennium. In: Donaldson, S.I. and M. Scriven (eds.): Evaluating social programs and problems: Visions for the new millennium. Erlbaum, Mahwah, NJ: 109-141.
- (2007): Program theory-driven evaluation science: Strategies and applications. Erlbaum, Mahwah, NJ.
- DONALDSON, S.I. and M.W. LIPSEY (2006): Roles for theory in contemporary evaluation practice: Developing practical knowledge. In: Shaw, I., J.C. Greene and M.M. Mark (eds.): The Handbook of Evaluation: Policies, Programs, and Practices. Sage, London: 56-75.
- ESSER, H., K. KLENOVITS und H. ZEHNPFENNIG (1977): Wissenschaftstheorie. Grundlagen und Analytische Wissenschaftstheorie. B.G. Teubner, Stuttgart.
- EC (2001): Ex ante evaluation: a practical guide for preparing proposals for expenditure programmes. In: http://www.ec.europa.eu/budget/library/documents/evaluation/guides/ex_ante_guide_2001_en.pdf.
- (1999): Evaluation of rural development programmes 2000-2006 supported from the European Agricultural Guidance and Guarantee Fund. Guidelines. Directorate General for Agriculture.
- (2007-2013): Handbook on common monitoring and evaluation framework. Guidance document, September 2006. Directorate General for Agriculture and Rural Development.

- LOKSHIN, M. and R. YEMTSOV (2005): Has Rural Infrastructure Rehabilitation in Georgia Helped the Poor? The World Bank Economic Review Advance Access originally published online on August 31, 2005. In: The World Bank Economic Review 19 (2): 311-333; doi:10.1093/wber/lhi007.
- MARSCHAK, J. (1953): Economic Measurements for Policy and Prediction. In: Hood, W. and T. Koopmans (eds.): Studies in Econometric Method. John Wiley, New York: 1-26.
- MICHALEK, J. (2008): Construction and measurement of rural development index. Application to evaluation of RD policies. Conceptual Report, University of Kiel.
- TODD, P.E. and K.I. WOLPIN (2005): Ex Ante Evaluation of Social Programs. University of Pennsylvania, June 1, 2005. In: http://www.homes.stat.unipd.it/erich/teaching/eval/papers/todd_wolpin05.pdf.
- VAN DE WALLE, D. and D. CRATTY (2002): Impact Evaluation of a Rural Road Rehabilitation Project. World Bank, Washington, D.C.
- WEISS, C.H. (1997): How can theory-based evaluation make greater headway? In: Evaluation Review 21 (4): 501-524.

Kontaktautor:

PROF. DR. CHRISTIAN HENNING
Department of Agricultural Economics,
Christian Albrechts University Kiel
Olshausenstr. 40-60, 24098 Kiel
Tel.: 04 31-880 44 53, Fax: 04 31-880 13 97
E-Mail: chenning@agric-econ.uni-kiel.de