

# **Methodische Ansätze zur Quantifizierung der Arbeitsplatzeffekte von Maßnahmen zur ländlichen Entwicklung**

**Dipl.-Ing. agr. Anne Margarian**

**Arbeitsberichte aus der vTI-Agrarökonomie**

**10/2010**

Frau Margarian ist wissenschaftliche Mitarbeiterin am Institut für Ländliche Räume des Johann Heinrich von Thünen-Instituts.

Adresse:        Institut für Ländliche Räume  
  
                  Johann Heinrich von Thünen-Institut (vTI),  
                  Bundesforschungsinstitut für Ländliche Räume, Wald und Fischerei  
  
                  Bundesallee 50  
  
                  D-38116 Braunschweig

Telefon:        (+49) (0)531 596 5511

E-Mail:         [anne.margarian@vti.bund.de](mailto:anne.margarian@vti.bund.de)

*Die Arbeitsberichte aus der vTI-Agrarökonomie* stellen vorläufige, nur eingeschränkt begutachtete Berichte über Arbeiten aus dem Institut für Betriebswirtschaft, dem Institut für Ländliche Räume und dem Institut für Marktanalyse und Agrarhandelspolitik des Johann Heinrich von Thünen-Instituts dar. Die in den Arbeitsberichten aus der vTI-Agrarökonomie geäußerten Meinungen spiegeln nicht notwendigerweise die der Institute wider. Kommentare sind erwünscht und sollten direkt an die Autoren gerichtet werden.

## Zusammenfassung

In der vorliegenden Arbeit werden mögliche methodische Ansätze zur Beurteilung von Arbeitsplatzeffekten von Fördermaßnahmen für den ländlichen Raum und insbesondere der Integrierten Ländlichen Entwicklung (ILE) vorgestellt und diskutiert. Aufgenommen wurden auch Methoden, deren Eignung für diese spezifischen Bewertungen verneint wird. Aus diesem Grund ist die vorliegende Arbeit nicht nur mit Blick auf die Bewertung der ILE zu lesen, sondern kann auch über mögliche Herangehensweisen an die Bewertung anderer Maßnahmen zur Entwicklung des ländlichen Raumes informieren. Die Diskussion der verschiedenen Methoden zeigt, dass es notwendig ist, die einer Wirkung zugrunde liegenden Zusammenhänge genau zu kennen, um die tatsächliche Wirkung einer Maßnahme auf den Arbeitsmarkt zu ermitteln. Es wird gezeigt, dass eine solche ausgereifte theoretische Basis für die ILE noch nicht existiert. Der Anspruch einer belastbaren Quantifizierung von Wirksamkeit und Effizienz kann daher zum jetzigen Zeitpunkt noch nicht erfüllt werden. Die Politikbewertung stellt in diesem Fall einen interaktiven Prozess der Wissensgenerierung dar.

**JEL:** C18, C54, R15, R58

**Schlüsselwörter:** Evaluation, Wirkungsanalyse, Politik für den ländlichen Raum, Arbeitsplatzeffekte

## Summary

The present work discusses possible methodological approaches for the assessment of employment effects of rural development measures. A special focus is on measures of the so called integrated rural development. Methods that are considered non-appropriate are considered nevertheless. Therefore, this paper may inform more generally about possible approaches towards the assessment of policy effects. The discussion of different methods clarifies the inevitable necessity of intimate causal knowledge of measures' intervention logics in order to reliably identify employment effects. Since such a sound theoretical basis does not exist yet for the measures of integrated rural development, the demand for a reliable quantification of the measures' effectiveness and efficiency cannot yet be met. In this case, policy evaluation has to be judged as an interactive process of knowledge generation.

**JEL:** C18, C54, R15, R58

**Keywords:** Evaluation, causal analysis, rural development measures, employment



## Inhaltsverzeichnis

<b>Zusammenfassung/Summary</b>	<b>i</b>
<b>1 Einleitung</b>	<b>1</b>
<b>2 Herausforderungen der Politikbewertung</b>	<b>3</b>
<b>3 Evaluation nach dem statistischen Vergleichsansatz</b>	<b>9</b>
3.1 Das Kontrollgruppendesign	9
3.1.1 Methodische Erläuterungen zum Kontrollgruppendesign	10
3.1.2 Beispielstudie zum Kontrollgruppendesign	12
3.1.3 Kritische Reflexion der Eignung des Kontrollgruppendesigns	14
3.1.4 Tabellarische Bewertung des Kontrollgruppenverfahrens	16
3.2 Das Propensity Score Matching (PSM)	16
3.2.1 Methodische Erläuterungen zum Propensity Score Matching	16
3.2.2 Beispielstudie zum Propensity Score Matching	17
3.2.3 Kritische Reflexion der Eignung des Propensity Score Matchings	19
3.2.4 Tabellarische Bewertung des PSM	22
3.3 Das Fixed Effects Panel Modell (FE-Panelmodell)	22
3.3.1 Methodische Erläuterungen zum FE-Panelmodell	23
3.3.2 Beispielstudie zum FE-Panelmodell	24
3.3.3 Kritische Reflexion der Eignung des FE-Panelmodells	25
3.3.4 Tabellarische Bewertung des FE-Panelmodells	26
<b>4 Theoriebasierte Evaluation</b>	<b>27</b>
4.1 Simulationsstudien	29
4.1.1 Die Regionale „Social Accounting Matrix“ (SAM)	29
4.1.1.1 Methodische Erläuterungen zur SAM	30
4.1.1.2 Beispielstudie zur Regionalen SAM	31
4.1.1.3 Kritische Reflexion der Eignung Regionaler SAM	33
4.1.1.4 Tabellarische Bewertung der RSAM-Modelle	35
4.1.2 Das ökonomische Input-Output-Modell (Ö-I-O-Modell)	36
4.1.2.1 Methodische Erläuterungen zum Ö-I-O-Modell	36
4.1.2.2 Beispielstudie zum Ö-I-O-Modell	37
4.1.2.3 Kritische Reflexion der Eignung des Ö-I-O-Modell	38
4.1.2.4 Tabellarische Bewertung der Ö-I-O-Modelle	40
4.2 Empirische Analysen	40
4.2.1 Die Partialanalyse	42
4.2.1.1 Methodische Erläuterungen zur Partialanalyse	43
4.2.1.2 Beispielstudie zur Partialanalyse	44

4.2.1.3	Kritische Reflexion der Eignung der Partialanalyse	47
4.2.1.4	Tabellarische Bewertung der Partialanalyse	49
<b>5</b>	<b>Die interventionslogikbasierte Evaluation</b>	<b>51</b>
5.1	Methodische Erläuterungen zur interventionslogikbasierten Evaluation	51
5.2	Beispielstudien zur interventionslogikbasierten Evaluation	53
5.3	Kritische Reflexion der Eignung der interventionslogikbasierten Evaluation	58
5.4	Tabellarische Bewertung der interventionslogikbasierten Evaluation	59
<b>6</b>	<b>Zusammenfassende Würdigung und Schlussfolgerungen</b>	<b>61</b>
	<b>Literaturverzeichnis</b>	<b>65</b>

## Abbildungsverzeichnis

Abbildung 4.1:	Variablen und Beziehungen in ASMACHERS Modell der Wirkungsanalyse regionalpolitischer Instrumente	46
Abbildung 5.1:	Beziehungen zwischen den Akteuren des Bolivianischen Innovationssystems mit Blick auf die Prioritätensetzung	54
Abbildung 5.2:	Beispielhafte Ursache-Wirkungszusammenhänge im Fallstudiendesign	57

## **Tabellenverzeichnis**

Tabelle 2.1:	Vergleichende Gegenüberstellung der Investitionsrechnung und der Kosten-Ertrags-Analyse	4
Tabelle 2.2:	Handlungsfelder der ILE zur Stärkung der Wirtschaftskraft und Schaffung von Arbeitsplätzen sowie Einkommensmöglichkeiten	5
Tabelle 3.1:	Einfaches DiD-Verfahren in schematischer Darstellung	10
Tabelle 3.2:	DiD-Schätzung im Panel-Regressions-Modell in schematischer Darstellung	23
Tabelle 4.1:	Beispiel einer SAM	30

## 1 Einleitung

Die vorliegende Studie wurde im Auftrag des Referates 525 des Bundesministeriums für Ernährung, Landwirtschaft und Verbraucherschutz (BMELV) erstellt. Hintergrund war der Wunsch, politische Programme, insbesondere Maßnahmen der „Integrierten Ländlichen Entwicklung“ (ILE), im Hinblick auf ihre Arbeitsplatzeffekte zu bewerten und so eine Vergleichbarkeit der Effektivität der entsprechenden Politiken mit anderen Interventionen zu ermöglichen. Die vorliegende Studie basiert auf einer Literaturrecherche zu Methoden, die in der Politikbewertung angewendet werden. Da die ILE noch ein relativ neues Politikkonzept darstellt, sind die entsprechenden Erfahrungen in diesem Bereich noch relativ gering. Aus diesem Grund wurden viele Methoden und Beispielstudien in die vorliegende Arbeit aufgenommen, die bisher nichts mit der ILE zu tun haben. Aufgenommen wurden auch Methoden, deren Eignung für diese spezifischen Bewertungen verneint wird. Aus diesem Grund ist die vorliegende Arbeit nicht nur mit Blick auf die Bewertung der ILE zu lesen, sondern kann auch über mögliche Herangehensweisen an die Bewertung anderer Maßnahmen zur Entwicklung des ländlichen Raumes informieren. Die Beschränkung auf die Bewertung der Arbeitsplatzeffekte entspricht der politischen Prioritätensetzung. Beim Vergleich von Fördermaßnahmen muss aber immer beachtet werden, dass mit diesen in der Regel mehrere Ziele in unterschiedlicher Gewichtung verfolgt werden.

Die Diskussion der verschiedenen Methoden zeigt, dass es notwendig ist, die einer Wirkung zugrunde liegenden Zusammenhänge genau zu kennen, um die tatsächliche Wirkung einer Maßnahme auf den Arbeitsmarkt zu ermitteln. In vergleichenden statistischen Ansätzen (Paradigma des Kontrafaktischen) ist die Kenntnis der Zusammenhänge notwendig, um die vollständige Kontrolle des Selektionseffektes glaubhaft zu machen und die den verwendeten Methoden zugrunde liegenden Annahmen zu verteidigen. Im theoriebasierten Ansatz müssen die relevanten funktionalen Zusammenhänge der unterstellten Modelle vor der Analyse explizit definiert und spezifiziert werden.

Programme wie die Integrierte Ländliche Entwicklung (ILE) und vergleichbare Maßnahmen der Regionalförderung dienen der Überwindung von Koordinationsproblemen und der Aktivierung brach liegender Potentiale. Sie basieren zwar auf neueren wissenschaftlichen Theorien zur Regionalentwicklung, in denen der Faktor Humankapital (z. B. durch Netzwerkeffekte und Wissens-Spillover) eine zentrale Rolle spielt. Eine wissenschaftliche Identifizierung notwendiger und hinreichender Bedingungen für die Aktivierung vorhandenen Humankapitals steht aber noch aus. Es existiert daher noch keine Grundlage für theoriebasierte Wirkungsmessungen. Die erwarteten indirekten Effekte und diskontinuierlichen Entwicklungen stellen auch die Anwendung vergleichender Analysen zur Ermittlung von Nettoeffekten in Frage. Für die Bewertung der ILE und vergleichbarer Maßnahmen eignen sich die entsprechenden Methoden daher nicht.

In diesem Fall einer noch unausgereiften wissenschaftlichen Basis stellt die wissenschaftliche Politikbegleitung einen interaktiven Prozess dar, in dem nicht nur die Politik von der Wissenschaft sondern auch die Wissenschaft von der Politik lernt. Sind beide Seiten offen für möglicherweise notwendige Änderungen alter (Denk-)Gewohnheiten, so kann die Effektivität der Politik so in einem iterativen Lernprozess weiter verbessert werden. Gleichzeitig ist die Erweiterung des Wissens um die Ursachen regionalwirtschaftlicher Dynamik durch die Grundlagenforschung eine Voraussetzung dafür, dass es zukünftig möglich werden kann, Nettoeffekte auf dem Arbeitsmarkt von Integrierten Fördeansätzen zu berechnen.

## 2 Herausforderungen der Politikbewertung

Im Zuge der steigenden Notwendigkeit, knapper werdende Mittel der öffentlichen Hand effizient und effektiv mit Blick auf die Verwirklichung gesellschaftlicher Ziele einzusetzen, steigt auch das Bedürfnis, die Wirkung der verwendeten Mittel zu messen und gegeneinander abzuwägen. Dieser Wunsch wurde auch mit Blick auf die Maßnahmen der „Zweiten Säule der EU-Agrarpolitik“ in den letzten Jahren immer deutlicher. Auch deshalb entstand in diesem Bereich, wie in vielen anderen, inzwischen ein umfassendes institutionalisiertes Evaluationswesen. Kritisch wurde mit Blick auf die Evaluationen in den letzten Jahren angemerkt, dass das verwendete methodische Handwerkszeug in der Regel nicht dazu geeignet erscheint, tatsächlich Nettoeffekte von Maßnahmen zu isolieren (siehe z. B. HENNING und MICHALEK, 2008; CASPARI und BARBU, 2008). In der vorliegenden Arbeit sollen, ausgehend von einer Literaturstudie, Methoden identifiziert werden, die dazu geeignet sind, die Arbeitsplatzeffekte der Maßnahmen für die Entwicklung des ländlichen Raumes zu messen. Ein besonderer Schwerpunkt soll dabei auf den Maßnahmen der ILE liegen.

Prinzipiell wäre die Kosten-Nutzen-Rechnung der geeignete Ansatz zur Abwägung unterschiedlicher Mittelverwendungen (KOESTER, 2001). Die folgende Tabelle 2.1 verdeutlicht das Konzept der Kosten-Nutzen-Rechnung in Abgrenzung zur unternehmerischen Investitionsrechnung. Charakteristisch ist vor allem die Notwendigkeit, die Wirkung auf *alle* direkt und indirekt Betroffenen zu erfassen. Das Entscheidungskriterium in der letzten Zeile der Tabelle macht deutlich, dass alle Erträge und Kosten in eine umfassende Bewertung einbezogen werden müssen. Dazu wäre es nötig, sowohl Kosten als auch Nutzen politischer Maßnahmen vollständig zu quantifizieren. Während für ein Unternehmen einfach die Maximierung des Ertragswertes als Zielsetzung unterstellt wird, wird nach Tabelle 2.1 in der Kosten-Nutzen-Analyse das Ziel der simultanen Maximierung des Wohlstandes aller Wirtschaftseinheiten unterstellt.

**Tabelle 2.1:** Vergleichende Gegenüberstellung der Investitionsrechnung und der Kosten-Ertrags-Analyse

	<b>Investitionsrechnung</b>	<b>Kosten-Nutzen-Analyse</b>
<b>Adressat der Analyse</b>	Unternehmensleitung	Politische Organe
<b>Objekt der Analyse</b>	Investitionen der Unternehmung	Politische Maßnahmen ohne Maßnahmen der Globalsteuerung
<b>Fragestellung der Analyse</b>	Lohnen sich bestimmte Investitionen (ex post und ex ante)	Lohnen sich bestimmte politische Maßnahmen (ex post und ex ante)
<b>Unterstellte Zielsetzung</b>	Maximierung des Ertragswertes der Unternehmung	Maximierung des Wohlstandes grundsätzlich aller Wirtschaftseinheiten einer Volkswirtschaft oder einer Region
<b>Einschränkende Nebenbedingungen</b>	Gesetzliche Vorschriften, Steuern, Liquidität etc.	Die Finanzverfassung, Budgetbeschränkungen etc.
<b>Nutzen oder Ertrag des zu beurteilenden Objekts</b>	Alle Einnahmen der Unternehmung, die aus den Investitionen entstehen	Alle Vorteile der Wirtschaftseinheiten, die direkt oder indirekt von der Maßnahme berührt werden
<b>Kosten des zu beurteilenden Objekts</b>	Alle Ausgaben der Unternehmung, die aus der Investition entstehen	Alle Nachteile der Wirtschaftseinheiten, die direkt oder indirekt von der Maßnahme berührt werden
<b>Bewertungsmaßstäbe zur Ermittlung der Kosten und Erträge</b>	Marktpreise der Güter und Dienste sowie der Produktionsfaktoren	Grundsätzlich Marktpreise (real); Schattenpreise, Alternativkosten oder andere Hilfsmaßstäbe, wenn kein Marktpreis vorhanden
<b>Zeitraum, der für die Analyse relevant ist</b>	Wirtschaftliche Lebensdauer des Investitionsobjektes	Unterschiedlich je nach Art der Maßnahmen und ihrer Wirkung auf die Wohlstandspositionen der Wirtschaftseinheiten
<b>Diskontierungssatz</b>	Kalkulatorischer Zinssatz der Unternehmung unter Berücksichtigung der Finanzierungskosten und des Risikos	Der von der Gesellschaft als richtig erachtete Diskontierungssatz
<b>Auswahl- und Entscheidungskriterien</b>	Interner Zinsfuß, Ertragswert, Annuitätenvergleich	Quotient aus Erträgen und Kosten

Quelle: Hedtke (1973).

Werden daher einzelne Kosten- oder Nutzelemente in einer solchen Analyse vernachlässigt, so kann das Ergebnis keine objektive Beurteilung der betreffenden Maßnahmen ermöglichen. Vergleicht man etwa die Arbeitsplatzeffekte einer Maßnahme, die primär dem Erhalt der kulturellen Vielfalt und nur sekundär dem Erhalt von Arbeitsplätzen dienen soll, mit denen einer rein arbeitsmarktpolitischen Maßnahme, so sagt das Ergebnis noch nichts über die gesellschaftliche Effektivität der beiden Maßnahmen im Vergleich aus. Dieses Problem einer multidimensionalen Zielstruktur betrifft Maßnahmen zur Ent-

wicklung des Ländlichen Raumes und insbesondere Maßnahmen der ILE besonders. Neben wirtschaftlichen Zielen bestehen hier Ziele der Daseinsvorsorge, der Bewahrung des kulturellen Erbes, des Umwelt- und des Landschaftsschutzes (s. Tabelle 2.2).

**Tabelle 2.2:** Handlungsfelder der ILE zur Stärkung der Wirtschaftskraft und Schaffung von Arbeitsplätzen sowie Einkommensmöglichkeiten

<b>Erhaltung und Schaffung von Arbeitsplätzen und Einkommensmöglichkeiten in ländlichen Räumen</b>	1.	Landwirtschaft als eine tragende Säule der ländlichen Räume stärken
	2.	Forstwirtschaft als Rohstofflieferant und zur Pflege des Waldes als prägendes Element der Kulturlandschaft unterstützen
	3.	Rahmenbedingungen für den Dienstleistungssektor als Wachstumsmarkt verbessern
	4.	Wertschöpfung im ländlichen Raum durch Ausbau der regionalen Wirtschaftskreisläufe erhöhen
	5.	Rahmenbedingungen für Unternehmensgründungen schaffen
<b>Allgemeine Unterstützung der wirtschaftlichen Entwicklung</b>	6.	Strategische Allianzen in stärkerer Zusammenarbeit mit der Wirtschaft aufbauen
	7.	Regionalbewusstsein stärken und Grundgedanken der Integrierten Ländlichen Entwicklung breit verankern
<b>Verbesserung der Daseinsvorsorge: Bauland, Infrastruktur, Wohn- und Arbeitsumfeld</b>	8.	Bauland bereitstellen und nachhaltige Raum- und Siedlungsstrukturentwicklung gewährleisten
	9.	Verkehrsinfrastruktur als einen grundlegenden Standortfaktor bedarfsgerecht ausbauen
	10.	Nahversorgung, soziale Dienste und kulturelle Angebote als notwendige Grundversorgung und attraktives Wohnumfeld sichern
	11.	Umwelt-, Erholungs- und Freizeitwert der Kulturlandschaft als weiche Standortfaktoren weiter entwickeln
	12.	Hochwasser- und Gewässerschutz als Standortvorteil im interregionalen Wettbewerb sichern

Quelle: Kötter und Thiemann (2006).

Der Vorteil, sich für die Beurteilung politischer Interventionen zunächst an einem besonders zentralen Ziel zu orientieren, besteht hingegen gerade in der Eindeutigkeit des relevanten Zielindikators, in unserem Falle der Zahl der geschaffenen Arbeitsplätze, der eine monetäre Bewertung der Nutzenseite zwecks Aggregation überflüssig macht. Vielmehr kann der Erfolg in diesem Fall durch die Relation „eingesetzte Mittel je Arbeitsplatz“ angegeben werden. Die Priorität des Ziels der Arbeitsplatzschaffung wird durch die Europäische Kommission unterstrichen: „The promotion of a high level of employment is a core objective of the Union. Employment plays a central role in raising of average living standards. It contributes to cohesion and social integration. It is an important measure of

Europe's success in using its economic and human resources fully and effectively" (EUROPÄISCHE KOMMISSION, 2000).

Die Schwierigkeit besteht nach dieser Vereinfachung vor allem darin, indirekte Kosten- und Nutzelemente zu identifizieren und zu quantifizieren. In der unternehmerischen Investitionsrechnung müssen allein die Kosten und die Erträge berücksichtigt werden, die dem Unternehmen selbst direkt zuzuschreiben sind. Für die Kosten-Nutzen-Analyse hingegen, die das Ziel der Maximierung einer gesamtgesellschaftlichen Wohlfahrt unterstellt, müssen auch die indirekten positiven und negativen Effekte, die eine Maßnahme verursacht, erfasst und bewertet werden (s. Tabelle 2.1). Um alle direkten und indirekten Effekte zu erfassen, muss bekannt sein, welche von ihnen der Maßnahme zuzuschreiben sind. Das ist entweder möglich, wenn alle Zusammenhänge theoretisch bekannt sind, oder aber, wenn die Situation bekannt ist, die sich ohne Intervention eingestellt hätte (die kontrafaktische Situation). Beides ist in der Regel nicht der Fall.

Die negativen indirekten Effekte einer Intervention zählen zu ihren Kosten. Um sie als solche aber auch zu bewerten, müssen die Preise aller betroffenen Faktoren bekannt sein. Führt etwa eine Investitionsförderung dazu, dass Kapital aus einem anderen Bereich abgezogen wird, so muss der Investitionsverlust an der anderen Stelle mit dem internen Zinssatz bewertet werden. Solche Zusammenhänge stellen in der Bewertung große Anforderungen an die Datenverfügbarkeit. Neben der Kenntnis aller theoretischen Zusammenhänge oder der Kenntnis der Entwicklung ohne Intervention besteht daher in der Verfügbarkeit relevanter Kontextinformationen eine weitere Bedingung für die quantitative Beurteilung politischer Maßnahmen.

Informationen sind insbesondere zu den sogenannten „weichen“ Erfolgsfaktoren oft nicht verfügbar. Gerade von diesen aber hängt der Erfolg der ILE ab: „Diese Faktoren, welche den Prinzipien von Regional Governance folgen, beruhen weitgehend auf der Freiwilligkeit der Akteure, der intersektoralen Kooperation und einer Einbindung von gesellschaftlichen Akteuren neben Politik und Verwaltung“ (BÖCHER und TRÄNKNER, 2008). Fließen etwa die Mittel der ILE insbesondere in die Regionen, in denen die entsprechenden Faktoren positiv ausgebildet sind und in denen die entsprechenden Prozesse daher sowieso teilweise oder in vollem Umfang abgelaufen wären, so werden die Prozesse möglicherweise dennoch der Förderung zugeschrieben, weil die tatsächlichen Erfolgsfaktoren nicht beobachtet wurden („Selektionsproblem“). Die ILE zielt darauf ab, das „lokale heuristische System“ der ländlichen Entwicklung fruchtbar zu machen: „It comprises such elements as: local economic, political and social actors; local development plans; social networks and kinship relations; local authorities, innovative individuals, development associations and partnerships as well as the development skills and experiences of these local actors“ (NEMES, 2005). Daraus folgt aber auch, dass die quantitative Erfassung konkreter Wirkungen einzelner Interventionen für die ILE besondere Herausforderungen birgt.

Darüber hinaus wurde von Wissenschaft und Politik in den letzten Jahren zunehmend anerkannt, dass es sich bei der Entwicklung ländlicher Räume nicht um einen linearen, monokausalen Trend handelt: „The term associated with this policy shift, namely integrated rural development reflects a new understanding of the problem. The word integrated stresses that the object is a complex, multidimensional one and consists of different interacting elements. This implies a need for proper assessment of these interactions, which requires an essentially dynamic approach” (KOSTOV und LINGARD, 2002). Dementsprechend ist es nicht klar, zu welchem Zeitpunkt welche Effekte zu erwarten sind. Es kann sein, dass in den ersten Jahren nach Anlauf eines Projektes noch kein monetärer Erfolg entsteht, während danach die Entwicklung schnell vorangeht. Die Wirkung einer Maßnahme kann zu verschiedenen Zeitpunkten und unter verschiedenen sozioökonomischen Rahmenbedingungen unterschiedlich, sogar gegensätzlich sein.

In den vergangenen Jahren ist eine Vielzahl von Methoden entwickelt und weiterentwickelt worden, die der Ermittlung und Bewertung von erwünschten und unerwünschten Wirkungen politischer Interventionen dienen soll. Der Schwerpunkt mancher Methoden liegt in der Überwindung der Schwierigkeiten, die aus der theoretischen Unkenntnis von Zusammenhängen und Verhaltensregeln der Wirtschaftssubjekte resultieren. Sie versuchen, die Wirkungen durch den statistischen Vergleich mit dem approximierten Kontrafaktischen zu ermitteln. Andere Methoden hingegen gehen vom (unvollständigen) theoretischen Modell aus und simulieren durch eine möglichst detaillierte Rekonstruktion der bekannten Zusammenhänge die Entwicklung, um Wirkungen von Eingriffen ins System zu ermitteln. In beiden Gruppen gibt es verschiedene Arten, mit dem dritten Problem, den allgegenwärtigen Datenrestriktionen, umzugehen. Ein dritter Weg besteht in der vertieften Analyse von einzelnen Projekten im regionalen Kontext, wobei vor allem durch Befragungen zentrale Wirkungselemente ermittelt werden. Hier steht im Gegensatz zu den beiden anderen Methoden das Bestreben im Zentrum, den Wirkungsmechanismus vor Ort empirisch nachzuvollziehen und zu verstehen. Dabei kann die Interventionslogik einer Maßnahme direkt überprüft werden. Dieses Vorgehen wird gewählt, wo die Spezifität der Entwicklungen, die zu beurteilen sind, in Raum und Zeit hoch ist, und wo zentrale Informationen nicht sekundärstatistisch vorliegen. Im Folgenden werden zunächst die vorwiegend statistisch-empirischen Methoden, die dem Paradigma des Kontrafaktischen folgen, dann die vorwiegend modellhaft-simulierenden Methoden der theoriebasierten Evaluation und schließlich die qualitativ-empirischen Methoden der interventionslogikbasierten Analyse nacheinander diskutiert.

Nur wenige der präsentierten Methoden wurden tatsächlich für die Bewertung von Maßnahmen zur Unterstützung einer ILE eingesetzt. Dafür gibt es gute Gründe. Die entsprechenden Methoden werden dennoch präsentiert, um zum einen die spezifischen Schwierigkeiten der Analyse von Arbeitsplatzeffekten der ILE zu verdeutlichen und zum anderen auch Hinweise auf die Bewertung von Arbeitsplatzeffekten des weiteren Maßnahmenpektrums zur ländlichen Entwicklung zu geben. Die Methoden werden jeweils zunächst

knapp methodisch-technisch erläutert und dann beispielhaft anhand einer Studie vorgestellt, bevor eine Einordnung der Stärken und Schwächen der jeweiligen Methodik mit Blick auf die Maßnahmen zur Entwicklung des ländlichen Raumes insgesamt und auf die ILE im Besonderen erfolgt. Am Ende jeden Abschnittes werden zentrale Punkte, wie theoretische Annahmen, Sekundärdatenbedarf, Erhebungsaufwand, Anforderungen an Spezialkenntnisse und technischer Aufwand tabellarisch zusammenfassend dargestellt.

### 3 Evaluation nach dem statistischen Vergleichsansatz

In der Einleitung wurde angesprochen, dass das Ziel der statistisch-empirischen Methoden darin besteht, die kontrafaktische Situation, die sich ohne Intervention eingestellt hätte, real oder statistisch zu erzeugen, um im Vergleich der Entwicklungen mit und ohne Förderung eindeutige Aussagen zu Wirkungen treffen zu können. Der prototypische Fall besteht in der gezielten Planung der Intervention analog zur Planung, etwa eines Feldversuches in der Landwirtschaft, in dem Parzellen zur Behandlung zufällig ausgewählt werden und mit den nicht ausgewählten vor und nach der Behandlung in ihrem Ertrag verglichen werden. Dieses „Kontrollgruppendesign“ wird als erstes Beispiel angeführt. Ist die Durchführung eines solchen Experimentes nicht möglich, so werden verschiedene statistische Methoden verwendet, die die möglichst weitgehende Simulation des Kontrafaktischen anstreben.

Die in der Grundlagenliteratur zu den relevanten Methoden diskutierten, diesen zugrunde liegenden technischen Annahmen werden allerdings in der Anwendungsliteratur nur noch selten explizit beschrieben. Die getroffenen Annahmen aber bedürfen einer Untermauerung und Begründung. Während daher die vorgestellten Methoden entwickelt wurden, um das Problem des Theoriedefizits in der Wirkungsanalyse zu überwinden, kann dies nur in begrenztem Umfang gelingen: „[...] the data almost never speak for themselves, especially when there are missing data and missing counterfactual states. Econometric methods use economic theory to guide the construction of counterfactuals and to provide a discipline on empirical research in economics“ (HECKMANN, 2000). In den folgenden Kapiteln zum tatsächlichen Kontrollgruppendesign sowie zur hypothetischen Konstruktion des Kontrafaktischen in der Matching und zur Difference-in-Difference-Analyse wird es darum gehen, die Relevanz der versteckten Annahmen aufzuzeigen und zu diskutieren.

#### 3.1 Das Kontrollgruppendesign

Das Kontrollgruppendesign stellt das Ideal dar, dem man mit den anderen statistisch-empirischen Methoden möglichst nahe kommen möchte. Die folgenden Abschnitte werden zeigen, dass auch das „ideale“ Kontrollgruppendesign im sozialen Feldversuch im Vergleich zum „experimental design“ im Labor noch mit vielleicht überraschend vielfältigen und tiefgreifenden Problemen zu kämpfen hat. Die nachfolgenden Methoden können diese Probleme nicht beheben, sondern beruhen vielmehr auf weiteren nicht allgemeingültigen Annahmen. Der Weg kann daher nur darin bestehen, die im Folgenden zu beschreibenden Probleme zu kennen und bei der Planung und Durchführung der Analyse sowie bei der Interpretation der Ergebnisse zu berücksichtigen.

### 3.1.1 Methodische Erläuterungen zum Kontrollgruppendesign

Die Methode des Kontrollgruppendesigns ist nur im Rahmen einer im Voraus geplanten und in der Implementierung der zu untersuchenden Maßnahme bereits berücksichtigten Wirkungsanalyse möglich. Sie beruht auf der bewusst „zufälligen“ Auswahl geförderter Personen oder Regionen einerseits und nicht geförderter Personen oder Regionen, der Kontrollgruppe, andererseits bereits vor dem Hintergrund der späteren Analyse. Die Auswahl der zu Fördernden sollte entsprechend den Regeln der Stichprobenziehung erfolgen, um den zufälligen Charakter der Gruppen auch bei einer begrenzten Anzahl von Objekten in der Grundgesamtheit zu gewährleisten.<sup>1</sup>

Demnach sollten die Merkmale, die die Wirkung der analysierten Maßnahme nachweislich beeinflussen, zur Gruppenbildung herangezogen werden, um zu verhindern, dass die Gruppen sich zufällig in relevanten Merkmalen unterscheiden. Indem so Repräsentativität beider Gruppen für die Grundgesamtheit mit Blick auf die relevanten Merkmale hergestellt wird, kann die Präzision der Schätzung gesteigert werden. Auch hier gilt wieder, dass eine adäquate Schichtung der Gruppen Geförderter und Nicht-Geförderter nur anhand beobachtbarer Merkmale möglich ist. Eine zufällige Verzerrung der Ergebnisse aufgrund von Unterschieden, etwa in der Struktur sozialer Netzwerke in den Regionen, kann daher nicht ausgeschlossen werden. Diesem Problem wird begegnet, indem der Unterschied im Ausgangswert von den Unterschieden im Endwert, nach dem Interventionszeitraum, abgezogen wird (Difference-in-Difference (DiD) Methodik; Tabelle 3.1).

**Tabelle 3.1:** Einfaches DiD-Verfahren in schematischer Darstellung

	Interventions- gruppe	Vergleichs- gruppe	Unterschiede zwischen Gruppen
Ausgangssituation $t_0$	$I_0$	$V_0$	$I_0 - V_0$
Nach der Interventionsphase, $t_1$	$I_1$	$V_1$	$I_1 - V_1$
<b>Unterschiede in der Zeit</b>	$I_1 - I_0$	$V_1 - V_0$	$(I_1 - I_0) - (V_1 - V_0)$

Quelle: In Anlehnung an NONIE, Subgroup 1, (2007), S. 9.

Der „Unterschied im Unterschied“ findet sich in Tabelle 3.1 im unteren rechten, dunkel unterlegten Feld. Hat die zufällige Verzerrung einen im Zeitablauf konstanten Einfluss auf den Wirkungsindikator, so kürzt sie sich heraus.

<sup>1</sup> Die intuitiven Begründungen des folgenden Abschnittes finden sich detaillierter und technischer, aber kompakt, auch in BLOOM (2006).

Bewertet wird also nur der Unterschied in der Entwicklung Geförderter und Nicht-Geförderter, nicht der Absolutwert nach Intervention. Indem allerdings eine unbeobachtete unterschiedliche Anfangsstruktur auch unterschiedliche dynamische Effekte nach sich ziehen kann, bleibt dennoch ein Rest Unsicherheit bestehen. Es handelt sich aber jetzt nicht mehr um einen systematischen Selektionseffekt, sondern nur noch um einen zufälligen Stichprobenfehler, dessen Relevanz demnach deutlich geringer sein muss als im unvorbereiteten Vergleich Geförderter und Nicht-Geförderter: „Matching [...] enhances the precision of the experiment; the randomization again ensures that unknown influences on behavior are equalized across the treatment groups“ (BORUCH, 1997).

Darüber hinaus muss die Betrachtungsebene der Analyse richtig gewählt werden. Bei Arbeitsplatzmaßnahmen etwa betrifft der Erfolg des Einen möglicherweise den des Anderen negativ. Vergleicht man dann Geförderte mit Nicht-Geförderten, so wird die Erfolgsdifferenz gemessen, die aber allein aus dem Nachteil des anderen entstanden sein kann. Der gesellschaftliche Nutzen ist dann gleich null. Auch andersherum ist es denkbar, dass durch bestimmte Maßnahmen eines Projektes nicht nur die Projektteilnehmer eine positive Wirkung erfahren, sondern aufgrund von „Spillover-Effekten“ auch andere profitieren. Werden diese miteinander verglichen, so wird der positive Einfluss der Maßnahme unterschätzt. Eine Lösung dieses Problems kann darin bestehen, eine höhere Aggregationsebene in den zu vergleichenden Einheiten zu wählen, also statt einzelner Projekte etwa die Regionen, in denen Projekte gefördert wurden, mit Regionen ohne Projektförderung zu vergleichen. Ob die höhere Analyseebene das Problem löst, muss auf sachlogischer Ebene abgeschätzt werden. Geht es etwa um die Beurteilung der Förderung von Investitionen in großen Unternehmen, so ist es wahrscheinlich, dass die Standortwahl zugunsten einer Region zu Lasten einer anderen geht. In diesem Fall kann auch der direkte Vergleich auf Regionsebene irreführend sein.

Auf einer höheren Aggregationsebene wird es aufgrund der sinkenden Zahl der Vergleichsobjekte in der Grundgesamtheit zunehmend schwierig, „zufällige“ Vergleichsgruppen zu konstruieren (BLOOM, 2006). Auch die Anzahl der relevanten Merkmale, in denen sich die höher aggregierten Analyseeinheiten unterscheiden, nimmt potenziell zu. Insbesondere mit Blick auf regionale Einheiten muss festgehalten werden, dass die zentralen Bedingungen, die regionale Entwicklungsunterschiede begründen und nach denen daher geschichtet werden müsste, ja gerade nicht bekannt sind.

Eine andere Möglichkeit zur Lösung des Problems von Spillover- oder Verdrängungseffekten könnte darin bestehen, die Geförderten etwa so auszuwählen, dass sie in bestimmten Clustern konzentriert sind, was nur gerechtfertigt werden kann, wenn die räumliche Dimension für den Fördererfolg nachweislich keine Rolle spielt. Im Vergleich der Geförderten und Nicht-Geförderten in der Förderregion mit den Nicht-Geförderten in der Nicht-Förderregion lassen sich dann sowohl direkte als auch indirekte Wirkungen isolieren und miteinander verrechnen. Negative indirekte Effekte, die in Abzug gebracht wer-

den müssen, existieren demnach dann, wenn die Nicht-Geförderten in den Förderregionen im Schnitt eine ungünstigere Entwicklung durchgemacht haben als die Nicht-Geförderten in den Nicht-Förderregionen. Die im Folgenden vorgestellte Beispielstudie ist diesen Weg der Gruppenauswahl auf Regionsebene gegangen.

### 3.1.2 Beispielstudie zum Kontrollgruppendesign

Nach Kenntnis der Autorin gibt es innerhalb Europas keine politischen Maßnahmen, die nach dem Kontrollgruppendesign implementiert und bewertet worden sind. Vorgestellt wird daher eine Studie, die die Wirkung eines mexikanischen Programms zur Armutsbekämpfung durch die Förderung der Bildung von Humankapital untersucht (PROGRESA): „Because of the broader geographic nature of some of PROGRESA benefits, such as improvements in local schools and health facilities, and because it would be difficult to have both treatment and control groups in the same small locality, randomization is being implemented at the locality rather than at the household level” (BEHRMAN und TODD, 1999). Dieses Vorgehen der Studie, in dem auf die beschriebenen möglichen Probleme, auch des Kontrollgruppendesigns, reagiert wird, ist der Grund für ihre Auswahl an dieser Stelle. Die Vorbereitung der Studie und der Implementierung verlief folgendermaßen: Zunächst wurden 505 Orte ausgewählt, die an der Evaluierung teilnehmen würden. (1) Anhand von Wohlstandskriterien wurde dann innerhalb jeden Ortes bestimmt, welche Familien an dem Programm teilnehmen sollten. (2) Schließlich wurde zufällig bestimmt, in welchen 60 % der Orte das Programm implementiert würde. (3) In der Evaluation selbst werden die potenziellen und tatsächlichen Empfänger des Programms jeweils aus den verschiedenen Regionen verglichen.

BEHRMAN und TODD (1999) beschreiben auch Verzerrungen, die durch das Kontrollgruppendesign selbst auftreten können: So ist vorstellbar, dass sich Personen, die ohne Kontrollgruppendesign teilgenommen hätten, aufgrund der geringeren Wahrscheinlichkeit, genommen zu werden, nicht um eine Teilnahme bemühen. Es ist auch denkbar, dass die Mobilität durch die Eingrenzung der Förderung auf bestimmte Gebietskulissen beeinflusst wird. Auch wenn einige Teilnehmer aus dem Programm aussteigen, hat das Effekte, die in der Interpretation der Ergebnisse zumindest berücksichtigt werden müssen.

Die Autoren überprüfen die Vergleichbarkeit der Kontrollgruppen mit der Gruppe der Geförderten zunächst anhand der Verteilung einiger zentraler Kennzahlen. Dazu gehören zunächst Indikatoren der geographischen Lage; es wird also überprüft, ob die geförderten und die nicht geförderten Orte auf bestimmte geographische Einheiten gleich verteilt sind. Außerdem werden die Altersverteilung von Kindern in den beiden Gruppen, das Ausbildungsniveau, die Verteilung der medizinischen Versorgung und die Einkommensverteilung verglichen. BEHRMAN und TODD (1999) finden auch einige Hinweise auf eine nicht

zufällige ungleiche Verteilung der Ausprägungen zwischen den beiden Gruppen, die aber ihrer Einschätzung nach auf statistischen Effekten beruhen können.

Als Indikator für den Erfolg des Programms werden sowohl der Unterschied in der Situation nach der Intervention zwischen Förder- und Kontrollgruppe als auch der korrespondierende DiD-Wert verwendet (SKOUFIAS, 2005). Die Unterschiede wurden nicht im Mittelwertvergleich, sondern über multivariate Regressionen bestimmt. Die genaue Spezifizierung wird sehr genau in SKOUFIAS (2005) beschrieben. So kann der Einfluss beobachtbarer Charakteristiken der Region, des Haushalts und der einzelnen Person auf den Maßnahmenerfolg bestimmt werden. Die Regression wird so spezifiziert, dass die Logik der simultanen Überprüfung im Vorher-Nachher- und im Mit-Ohne-Vergleich im Modell nachvollzogen wird; sie beinhaltet daher auch einen erneuten Test der Hypothese fehlender Unterschiede in der Ausgangssituation zwischen den Gruppen der Geförderten und der Nicht-Geförderten. Darüber hinaus ermöglicht die Regression für die Gruppe der Nicht-Förderberechtigten die Ermittlung möglicherweise bestehender indirekter Effekte. In diesem Fall rechnen die Autoren in erster Linie mit positiven Nebeneffekten (SKOUFIAS, 2005). Eine weitere Möglichkeit besteht, wenn Daten für einen längeren Zeitraum nach der Intervention vorliegen, in der Analyse des zeitlichen Verlaufs des Fördereffektes.

SKOUFIAS (2005) weist weiterhin auf einen Aspekt von fundamentaler Bedeutung hin, der in der Beispielstudie aufgrund der klaren Identifizierung aller potenziellen Empfänger empirisch zugänglich ist: Der Erfolg einer Maßnahme kann einerseits an der Wirkung der tatsächlich erfolgten Förderung, aber andererseits auch an der Wirkung der intendierten Förderung gemessen werden. Nehmen einige der potenziellen Empfänger an der Maßnahme nicht teil, so wird deren Effektivität, gemessen an der intendierten Förderung, geringer. Wird hingegen nur die tatsächliche Förderung beurteilt, so sinkt die Effektivität der Förderung mit sinkender Teilnahmebereitschaft nicht unbedingt ab. Die Auswahl der Geförderten wird weiter kompliziert, wenn einzelne Empfänger diese Gruppe im Zeitablauf verlassen. Dieses Beispiel zeigt, wie wichtig es ist, die Fragestellung genau zu fassen, um die Methode und den Indikator zu wählen, die auch die gewünschten Antworten liefern können.

Um mögliche Unterschiede der Programmwirkung zwischen den Regionen und ihre Begründung durch unterschiedliche Implementierungsdesigns zu untersuchen, wird im Rahmen der Evaluierung von PROGRESA außerdem eine Prozessanalyse (vgl. Fußnote 1; MOSLEY und SOL, 2001) durchgeführt.

Auf die Ergebnisse der Studie wird hier nicht weiter eingegangen, da sie nicht im thematischen Zusammenhang zur Entwicklung des ländlichen Raumes in Deutschland steht. Die Literatur zur Bewertung des Programms PROGRESA wird aber aufgrund ihrer Verständlichkeit und der ausführlichen Diskussion aller methodischen Möglichkeiten und Probleme

me dem an weiterführenden Informationen Interessierten sehr empfohlen (<http://www.ifpri.org/search/publications?keys=PROGRESA>).

### 3.1.3 Kritische Reflexion der Eignung des Kontrollgruppendesigns

Das Kontrollgruppendesign erfordert eine integrierte Planung der Bewertung politischer Maßnahmen schon bei deren Implementierung. Es stellt damit hohe Anforderungen an die Administration und Konzeption der Förderung. Das ist oft nicht in Übereinstimmung zu bringen mit der Art, in der politische Entscheidungen in der Realität getroffen werden. Neben diesen praktischen Herausforderungen hängt die Praktikabilität des Verfahrens von sachlichen, politischen und auch ethischen Überlegungen ab.

So ist es etwa auch bei ausreichender Mittelverfügbarkeit in diesem Design nicht möglich, alle die auszuwählen, für die der höchste Bedarf identifiziert wurde. Vielmehr muss dafür gesorgt werden, dass sich diese Gruppe zu proportionalen Teilen in und außerhalb der Geförderten befindet. Damit sind ein entsprechendes Design und die entsprechende Auswahl an Geförderten politisch oft nicht vermittelbar und zum Teil auch sachlich/ethisch nicht vertretbar.

Die Schwierigkeiten, die aufgrund von Verdrängungs- und Spillover-Effekten auftreten können, wurden bereits in Kapitel 2.1.2 diskutiert. Je weniger die zu bewertenden Programme auf einzelne Empfänger ausgerichtet sind und je mehr ganze Regionalwirtschaften in ihrer Gesamtheit direkt beeinflusst werden sollen, umso mehr verstärken sich diese Probleme. Auch in der Diskussion der Bewertung in der Beispielstudie wurde insbesondere mit Blick auf die späteren Jahre der Laufzeit die Annahme problematisiert, dass Kontrollregionen unbeeinflusst von der Entwicklung in Förderregionen seien (SKOUFIAS, 2005).

BEHRMAN und TODD (1999) erwähnen noch ein weiteres Problem der Kontrollgruppenmethodik: „Another disadvantage of random assignment at the locality level is that we lose our ability to estimate locality specific treatment impacts, as within each locality all eligible persons do or do not receive the treatment.“ Hinter dieser Anmerkung steht das Problem, dass in Kontrollgruppendesigns immer nur der Durchschnittseffekt einer Maßnahme berechnet wird. Die Verteilung des Effektes zwischen den Teilnehmern und mehr noch die Ursachen heterogener Wirkungen einer Maßnahme werden nicht ermittelt. Der mittlere Effekt aber kann in seiner Aussage ein irreführendes Bild vermitteln, wenn z. B. im Extrem die Hälfte der Teilnehmer negativ und die andere Hälfte positiv von der Maßnahme beeinflusst werden und der mittlere Effekt gleich null ist. Um derartige Unterschiede im Rahmen der Methodik zu erfassen, müssen weiterführende Analysen mit Treatment- und Kontrollgruppe durchgeführt werden, deren Voraussetzung ist, dass Informatio-

nen zu den relevanten Unterscheidungsmerkmalen zu beiden Gruppen vorliegen. In der Beispielstudie wird gezeigt, wie hier flexibel und erfolgreich vorgegangen werden kann.

Vorraussetzung für einen erfolgreichen Einsatz der Methode ist also, dass die möglichen Verdrängungseffekte von vornherein ausgeschlossen oder auf die beschriebene Art herausgerechnet werden können und dass die politischen und administrativen Probleme überwunden werden, die sich aus dem anspruchsvollen Implementationsdesign ergeben. Sind dann auch noch die notwendigen sekundärstatistischen Daten für die Auswahl der Zufallsstrichproben vorhanden und ist die Grundgesamtheit groß und homogen genug, um die Konstruktion vergleichbarer Gruppen zu ermöglichen, so stellt das Kontrollgruppendesign eine Möglichkeit dar, trotz theoretischer Unsicherheiten relativ unumstrittene Ergebnisse etwa in der Quantifizierung von Arbeitsplätzen zu erhalten. Mit Blick auf die ILE gilt allerdings, dass die notwendige Ex-ante-Auswahl von Förderprojekten anhand exogen vorgegebener Kriterien durch den Mittelgeber dem Grundprinzip der Förderung wenigstens teilweise widerspricht. Eine Alternative könnte hier im sogenannten „Pipeline-Verfahren“ bestehen. Werden Projekte im Rahmen einer Maßnahme willkürlich (zufällig) zeitversetzt bewilligt, so können spätere Projekte als Kontrolle für frühere Projekte herangezogen werden (CASPARI und BARBU, 2008).

Die in der Beispielstudie angewendeten ergänzenden multivariaten Analysemethoden erhöhen den Informationsgehalt entsprechender Analysen, in denen oft nur die Nettowirkung als Ergebnis eines Mittelwertvergleichs ausgewiesen wird. Für eine Politikoptimierung aber sind weitere Informationen notwendig. SKOUFIAS (2005) betont, dass der ermittelte Effekt in der verwendeten Spezifizierung sowohl den Angebots- als auch den Nachfrageeffekt enthält. Um diese möglichen Effekte, die einerseits direkt durch den Finanzmitteltransfer und andererseits indirekt durch die nachfolgende Verbesserung der Versorgungslage entstehen, zu trennen, müssten hinreichende Kontrollvariable zur genauen Beschreibung der Versorgungslage vor und nach der Intervention existieren. Die Trennung erfolgt in der beschriebenen Studie aufgrund der Datenrestriktionen nicht. Das Bestreben drückt aber den Wunsch aus, alle erwünschten und unerwünschten, direkten und indirekten Effekte differenziert zu erfassen. Letztlich ermöglichen nur derart differenzierte Ergebnisse die Identifizierung kritischer Faktoren für die Planung einer effektiven Politikoptimierung.

### 3.1.4 Tabellarische Bewertung des Kontrollgruppenverfahrens

	1	2	3	4	5	6	7
<b>Technische Aspekte</b>	sehr gering				sehr hoch		
Zeitbedarf		■					
Anforderungen an technisches Know-How			■				
Anforderungen an Kontextwissen/Theoretisches Wissen			■				
Softwarebedarf		■					
Datenanforderungen				■			
Sonstige Kosten							■
<b>Inhaltliche Aspekte</b>	sehr gut möglich				nicht möglich		
Erfassung direkter Effekte	■						
Erfassung indirekter Effekte					■		
Direkte Modellierung der Interventionslogik							■
Erfassung diskontinuierlicher, auch zeitverzögerter Effekte			■				
Erfassung positiver externer Effekte					■		
Erfassung von Allokationseffekten					■		
Verallgemeinerbarkeit				■			
Beachtung des regionalen Kontextes/weiterer Wirkungsbedingungen		■					
	Pro				Contra		

## 3.2 Das Propensity Score Matching (PSM)

Das Propensity Score Matching wird im Anschluss an das Kontrollgruppendesign vorgestellt, weil hier das gleiche Prinzip aus der Ex-post-Perspektive angewendet wird. Der zentrale Unterschied besteht dann darin, dass die Auswahl der Empfänger nicht tatsächlich zufällig erfolgte. Zwar wird im Matching die Schichtung der Beobachtungseinheiten für den Vergleich in der Analyse vorgenommen, doch kann dies nur mit Blick auf beobachtbare Variable geschehen. Die faktische Randomisierung hingegen verspricht eine zufällige Verteilung auch der unbeobachtbaren Faktoren.

### 3.2.1 Methodische Erläuterungen zum Propensity Score Matching

Das Ziel des Matchings besteht in der Identifizierung sehr ähnlicher Vergleichspaare. In herkömmlichen Verfahren kann Ähnlichkeit in Vergleichspaaren nur mit Blick auf einige wenige Eigenschaften, die durch kontinuierliche Variable beschrieben werden, hergestellt werden. Beim Matching wird die Wahrscheinlichkeit der Teilnahme an einem Programm mit Hilfe aller relevanten individuellen Charakteristiken in einem logistischen Regressi-

onsmodell<sup>2</sup> geschätzt. Anschließend werden die Paare mit den ähnlichsten Wahrscheinlichkeitswerten (Propensity Score) verglichen. Diese Paare ähneln sich dann nicht nur in der Wahrscheinlichkeit ihrer Teilnahme, sondern weisen gleichzeitig eine maximierte multidimensionale Ähnlichkeit in den relevanten Eigenschaften auf (ROSENBAUM und RUBIN, 1983).

Beim Matching selbst handelt es sich um ein nicht-parametrisches Verfahren, das deshalb auf wenigen Annahmen mit Blick auf die Verteilung der zugrunde liegenden Variablen beruht. Da der Propensity Score aber in einem logistischen Regressionsmodell bestimmt wurde, handelt es sich insgesamt um ein semi-parametrisches Verfahren. Der Propensity Score kann alternativ als Regressor in ein weiteres Schätzmodell eingebracht werden. Im parametrischen Schätzmodell wird allerdings zum Beispiel Linearität des Zusammenhangs unterstellt. Diese Annahme kann meist nicht sachlogisch untermauert werden. Das Problem kann abgemildert werden, indem der Propensity-Score entlang seines Wertebereiches in mehrere Gruppen unterteilt und als Klassifizierungsvariable in das Modell eingebracht wird.

### 3.2.2 Beispielstudie zum Propensity Score Matching

Notwendige Grundlage für die Anwendung des Matching-Verfahrens ist eine hinreichende Anzahl von Beobachtungen Geförderter und Nicht-Geförderter, die sich nicht systematisch in bestimmten relevanten Eigenschaften unterscheiden. Anders ausgedrückt muss es eine hinreichend große Schnittmenge Geförderter und Nicht-Geförderter geben, die sich in ihren relevanten Eigenschaften nicht unterscheiden. So ist das Matching-Verfahren nicht anwendbar, wenn z. B. systematisch alle Personen einer bestimmten sozialen Schicht gefördert werden, alle anderen hingegen nicht.

Aus diesen Gründen wird das Matching vor allem in der Arbeitsmarktforschung angewendet: Hier liegen in der Regel ausreichend viele Individualdaten zur Identifizierung einer hinreichend großen Anzahl vergleichbarer Paare vor. Um eine solche Studie handelt es sich auch bei der ausgewählten Beispieluntersuchung von EICHLER und LECHNER (2002). Sie analysieren die Effekte von Arbeitsbeschaffungsmaßnahmen (ABM) in Sachsen-Anhalt. Die von EICHLER und LECHNER zitierten Studien zum Arbeitsplatzeffekt aktiver Arbeitsmarktpolitik in verschiedenen europäischen Ländern zeigen ambivalente, zum

---

<sup>2</sup> Auf die Methodik der logistischen Regression wird an dieser Stelle nicht weiter eingegangen. Es handelt sich hierbei um eine sehr häufig verwendete Modellspezifizierung zur Schätzung von Ja-/Nein-Wahrscheinlichkeiten. Diese wird abgeleitet aus der geschätzten Chance (dem Wahrscheinlichkeitsverhältnis). Die Chance liegt zwischen null und unendlich und kann daher durch ein lineares Modell bestimmt werden. Informationen finden sich in allen weiterführenden Statistiklehrbüchern. Alternativ ist auch die Schätzung in einem Probit-Modell möglich.

Teil sogar negative Effekte auf. Diese Studien unterscheiden sich in den verwendeten Methodiken untereinander und von der der Autoren. Diese finden eine signifikante und substanzielle Verringerung der Arbeitslosigkeit durch ABM.

EICHLER und LECHNER verwenden einen unbalancierten Paneldatensatz (den Arbeitsmarktmonitor Sachsen-Anhalts) mit Daten für den Zeitraum von 1991 bis 1997. Sie schließen aufgrund von inhaltlichen Überlegungen Personen im Datensatz, die mit Beginn des Betrachtungszeitraums 1993 jünger sind als 22 Jahre und solche, die bei Beginn der ABM älter als 52 Jahre sind, aus der Analyse aus, da hier aufgrund des Berufseinstiegs oder einer möglichen sich anschließenden Frühverrentung abweichende Wirkungen nicht ausgeschlossen werden können. Die häufigste Dauer der ABM (der Median der Maßnahmendauer) beträgt zwölf Monate (ibid.). Die Gruppen der Teilnehmer und der Nicht-Teilnehmer im Datensatz unterscheiden sich in ihrem Arbeitslosigkeitsstatus, in ihrem Bildungsniveau und in ihrer Position ihrer letzten Anstellung. Aufgrund theoretischer Erwägungen stellen die Autoren fest, dass sie für eine klare Feststellung der Vergleichbarkeit von Teilnehmern und Nicht-Teilnehmern Informationen zur Schulbildung, zur Berufsausbildung und Berufserfahrung, zur vollständigen Arbeitsgeschichte jeder Person sowie zu weiteren sozioökonomischen Charakteristika haben müssten. Diese Informationen sind in den verfügbaren Daten nicht alle enthalten. Zur Schätzung der Teilnahmewahrscheinlichkeit und der Bestimmung des Propensity-Scores werden verfügbare Variable zu Geschlecht, Alter, Schulbildung, Universitätsabschluss und Berufsausbildung sowie zur Charakterisierung der regionalen Ebene herangezogen. Die Schätzung erfolgt hier in einem Probit- statt in einem Logit-Modell. Die regionale Arbeitslosigkeit steht nach den Schätzergebnissen in einem signifikanten Zusammenhang zur Teilnahmewahrscheinlichkeit, und auch eine Reihe regionaler Dummies, die Regionen mit bestimmten Industriestrukturen charakterisieren, sind signifikant.

Wie von HECKMAN et al. (1997) vorgeschlagen, kombinieren die Autoren, weil sie mit unbeobachteten Variablen rechnen, den Vergleichsgruppen- und den DiD-Ansatz. Weil ihr unbalancierter Paneldatensatz mit Beobachtungen vor, während und nach der Teilnahme mehr Möglichkeiten bietet als Datensätze mit nur zwei Beobachtungszeitpunkten (vorher – nachher), modifizieren sie allerdings den DiD-Ansatz. Da sie aufgrund unbeobachtbarer Einflussgrößen einen nicht korrigierten Selektionsbias nicht ausschließen können, treffen sie alternativ die Annahme eines in der Zeit stabilen Bias. Wird für Beobachtungen vor Teilnahme ein Effekt gemessen, so handelt es sich um den Bias, um den der gemessene Effekt nach Teilnahme korrigiert werden muss. Dieses Vorgehen ermöglicht es, dass nicht zu beiden Messzeitpunkten (vor und nach der Teilnahme) jeweils dieselben Beobachtungen verglichen werden müssen, vielmehr kann der Effekt für Vergleichspaare nach der Teilnahme geschätzt werden und ebenso der „Scheineffekt“ für Vergleichspaare vor der Teilnahme. Der tatsächliche Effekt wird bei Gültigkeit bestimmter Annahmen durch Saldierung ermittelt. Insbesondere darf kein tatsächlicher Effekt der Teilnahme bereits vor der Teilnahme existieren, wie es etwa denkbar wäre, wenn etwa zukünftige

Teilnehmer ihre Jobs aufgeben, um in den Genuss der Teilnahme zu kommen. Außerdem muss die Annahme einen konstanten Bias halten. Dadurch können im unbalancierten Panel zum einen mehr Beobachtungen genutzt werden und zum anderen können auch zeitvariable Kovariate zur Schätzung der Teilnahmewahrscheinlichkeit herangezogen werden. Das technische Vorgehen beim Matchen mit den zeitvariablen Faktoren wird genauer in LECHNER (1999) beschrieben.

EICHLER und LECHNER (2002) zeigen zunächst anhand des signifikanten, durch Matchen von Teilnehmern und Nicht-Teilnehmern *vor* Teilnahme ermittelten „Scheineffekts“, dass die unbeobachteten Variablen und das entsprechend unvollständige Schätzmodell zur Bestimmung des Propensity Scores zu einem Selektionsbias führen. Die Autoren zeigen außerdem, dass die Beachtung zeitvariabler Charakteristika zusätzlich zum Propensity Score beim Matching den Bias verringert, allerdings nicht eliminiert. Ohne die Beachtung des Selektionsbias, der durch den „Scheineffekt“ ermittelt wurde, hätte das Ergebnis in einer scheinbaren Erhöhung der Wahrscheinlichkeit der Arbeitslosigkeit durch die ABM bestanden. Indem aber der Selektionsbias korrigiert wird, dreht sich das Ergebnis um: Es wird gezeigt, dass die ABM die Wahrscheinlichkeit einer Arbeitslosigkeit nach der Maßnahme verringert. Dieses Ergebnis ist insbesondere für das Matching mit zeitvariablen Informationen deutlich.

### 3.2.3 Kritische Reflexion der Eignung des Propensity Score Matchings

Eine zentrale Annahme, die schon vom Urheber des Verfahrens Rubin intensiv diskutiert wurde, ist die „Stable Unit Treatment Value Assumption“ (SUTVA): “The structure would not be adequate when, for example, the response of unit  $i$  to treatment  $t$  depends on the treatment given to unit  $j$ , as could happen if they compete for resources” (ROSENBAUM und RUBIN, 1983). Dadurch, dass die Vergleichspersonen um eine gemeinsame Ressource konkurrieren, wird die Vergleichsperson durch den Erfolg der geförderten Person negativ in ihrer Entwicklung beeinflusst. Durch den Vergleich der beiden Personen erhält man nicht den Nettoeffekt der Intervention, sondern lediglich die Wirkung der Förderung auf die Entwicklung des Geförderten *unter den Bedingungen der Intervention*. Bei entsprechender Modifizierung der Fragestellung ist die Anwendung des Matchings so zwar technisch legitimierbar, oft wird dann aber nicht mehr die relevante politische Fragestellung beantwortet. Dieses Problem entspricht dem auch im Zusammenhang mit dem Kontrollgruppensdesign diskutierten Problem der Verdrängungs- und Spillover-Effekte. Wieder zeigt sich, dass eine klare Definition der Problem- und Fragestellung zentral für die entsprechende Durchführung einer Analyse ist.

Die Autoren der Beispielstudie stellen dies auch selber fest: „However, one should be careful in interpreting these results: Individual gains for the participants do not necessarily translate into benefits for the economy as a whole, because the costs of PEPs [Public

Employment Programs] and market interactions are ignored. The latter occur through possible negative impacts for competing private firms as well as through a possibly changing labour market supply structure during and after PEPs. No estimates of such general equilibrium effects of PEPs exist for East Germany, but STEINER, WOLF, EGELN, ALMUS, SCHRUMPF, and FELDOTTO (1998) find some hints that the positive effects found by earlier research for public training programs in East Germany do not correspond to similar effects on the macroeconomic level. Following from the above discussion, a cost-benefit analysis based on the results of this study and the expenditures on PEPs would not be valid. Furthermore, such an analysis has to take account of other goals and effects of PEP, which we have no information on”.

Zwischen 1991 und 1997 nahmen im Osten Deutschlands mehr als zwei Millionen Personen an ABM teil (EICHLER und LECHNER, 2002). Daraus könnten zwei Probleme entstehen: Zum einen können neben den Mikroeffekten, die direkt beim Teilnehmer entstehen, angebots- und nachfrageseitige Makroeffekte nicht ausgeschlossen werden. Das Matching würde diese positiven Multiplikatoreffekte nicht ermitteln, da alle Personen im Land potenziell davon profitieren. Zum anderen könnten die ABM reguläre Beschäftigungsverhältnisse verdrängen. Wie einleitend geschildert, führt das Matching dann zu einer deutlichen Überschätzung des Effekts der Maßnahme. Die Teilnehmer der ABM erhalten bis 1994 den Tariflohn, die Entlohnung wird in den darauffolgenden Jahren bis auf 90 % des Tarifes abgesenkt (ibid.). Demnach liegt das Einkommen der Teilnehmer in der Regel deutlich über der alternativen staatlichen Unterstützung, und ein positiver nachfrageseitiger Makroeffekt ist zu erwarten. Der angebotsseitige Makroeffekt ist möglicherweise begrenzt, da ein relativ großer Anteil der Beschäftigten im Bereich der Landschaftspflege, der Landwirtschaft oder auch der Administration beschäftigt ist (ibid.). Unzweifelhaft aber existiert ein positiver Wertschöpfungseffekt der Beschäftigung ansonsten arbeitsloser Personen, der mit dem Matching ebenso wenig wie der nachfrageseitige Effekt erfasst wird. Das größte potenzielle Problem aber ist der Verdrängungseffekt. Dieses wurde auch als solches erkannt und hat sich in der Gesetzesformulierung niedergeschlagen, in der die Förderung von ABM-Stellen, die andere Stellen de facto ersetzen, ausgeschlossen wird. Die Arbeit muss von öffentlichem Nutzen sein und ohne ABM-Stelle nicht ausgeführt werden. Während diese Bestimmungen ohne Zweifel einen abmildernden Effekt haben, kann doch bezweifelt werden, dass Verdrängungseffekte durch die administrative Beurteilung völlig verhindert werden können. Ihre Quantifizierung ist mit Hilfe des Matching aber nicht möglich.

Der zentrale Nachteil des PSM gegenüber dem Kontrollgruppendesign besteht darin, dass die Methode voraussetzt, dass alle relevanten Selektionskriterien, die in Zusammenhang zur Förderung stehen, bekannt sind, dass also ein allgemein akzeptiertes und überprüftes theoretisches Modell zur Entwicklung der zentralen Wirkungsindikatoren und der Wirkung der Intervention existiert. Darüber hinaus müssen Daten zu allen relevanten Faktoren existieren, nicht beobachtbare Faktoren führen sonst zu einer Verzerrung. Anders als

im Fall des Kontrollgruppenverfahrens kann jetzt nicht mehr davon ausgegangen werden, dass unbekannt oder vernachlässigte Einflussgrößen in der Regel durch die Randomisierung unschädlich gemacht werden. In der Beispielstudie wird versucht, dieses Problem zu umgehen, indem nicht der absolute Unterschied zwischen den Vergleichsgruppen geschätzt wird, sondern der Difference-in-Difference-Wert. Durch die verfeinerte Methodik der Beispielstudie wird dabei besonders deutlich, dass auch die Ausdifferenzierung nur zum erwünschten Ergebnis führt, wenn die Faktoren, die die Teilnahme bestimmen, selbst zeitkonstant sind. Ist das nicht der Fall, so müssen sie in der von den Autoren angewandten Methodik zum Matching herangezogen werden. Es wurde auch deutlich, dass für die Bestimmung eines validen Ergebnisses die Möglichkeit einer Wirkung der Teilnahme bereits vor Teilnahme ausgeschlossen werden muss.

Die Autoren der Beispielstudie verwenden für das Matching nicht nur individuelle sondern auch regionale Charakteristika. Ein methodisches Dilemma entsteht, wenn weder die Existenz von Verdrängungs- oder Spillover-Effekten noch unbeobachtbare relevante Einflüsse auf einer höheren, z. B. räumlichen Aggregationsebene ausgeschlossen werden können. Die Lösung des letzten Problems könnte z. B., wenn genügend Beobachtungen vorliegen, darin bestehen, nur Beobachtungen aus einer gemeinsamen räumlichen Einheit zu matchen (HECKMAN et al., 1997). Genau das aber sollte vor dem Hintergrund des ersten Problems vermieden werden. Insbesondere, wenn der Kontext die Entwicklung stark beeinflusst und möglicherweise ein nicht-linearer Verlauf vorliegt, reicht der DiD-Ansatz aber nicht mehr aus, um unbeobachtete Heterogenität zu kontrollieren: „These studies illustrate the risks involved in comparing the behaviour of individuals residing in different geographic areas and can be generalised to all situations where context matters, that is where people residing in different areas are subject to different environments, and these are likely to affect the selection and outcomes under study“ (ARPINO und MEALLI, 2008). Eine Möglichkeit, um das Problem unbeobachteter Heterogenität zwischen Gruppen von Analyseobjekten zu umgehen, ist die Spezifizierung des Propensity Scores in Mehr-Ebenen-Modellen (ARPINO und MEALLI, 2008). Hier existieren methodische Ansätze, die eine Kontrolle der räumlichen Dimension ermöglichen, ohne dass die Vergleichspaare sich im selben Cluster befinden müssen.

In der ILE wird nicht Personen- sondern Projektbezogen gefördert. Das Matchen nach Projekten wurde nach Wissen der Autorin bisher nie angestrebt. Voraussetzung wäre eine hinreichende Vergleichbarkeit der Projekte sowie die Verfügbarkeit relevanter Daten zu den Projekten. Sind diese Bedingungen nicht gegeben, so bliebe das Matchen nach Regionen. Neben der Schwierigkeit, vergleichbare Regionen aus der relativ geringen Grundgesamtheit zu identifizieren, besteht das weitere Problem in der geringen Bedeutung, die realistisch für die eher kleinen Projekte der ILE für die Regionalentwicklung insgesamt erwartet werden kann.

Ein weiteres großes Problem des Matching-Ansatzes besteht darin, dass nur die Beobachtungen verwendet werden, für die hinreichend ähnliche Vergleichsgruppen existieren. Es besteht daher die Gefahr, dass die Schlussfolgerungen auf dem Vergleich von Randgruppen, die nicht repräsentativ für die Gesamtpopulation sind, beruhen. Eine sinnvolle Erweiterung des beschriebenen Panelansatzes besteht in der Überprüfung möglicherweise heterogener Effekte einer Teilnahme für verschiedene Untergruppen. Wie in der Einleitung beschrieben, kann der reine Mittelwertvergleich zu wenig politikrelevanten oder sogar irreführenden Schlussfolgerungen führen, wenn die Heterogenität der Wirkungen, die im Mittelwert zusammengefasst sind, hoch ist.

### 3.2.4 Tabellarische Bewertung des PSM

Technische Aspekte	sehr gering	sehr hoch
	Zeitbedarf	
Anforderungen an technisches Know-How		
Anforderungen an Kontextwissen/Theoretisches Wissen		
Softwarebedarf		
Datenanforderungen		
Sonstige Kosten		
Inhaltliche Aspekte	sehr gut möglich	nicht möglich
	Erfassung direkter Effekte	
Erfassung indirekter Effekte		
Direkte Modellierung der Interventionslogik		
Erfassung diskontinuierlicher, auch zeitverzögerter Effekte		
Erfassung positiver externer Effekte		
Erfassung von Allokationseffekten		
Verallgemeinerbarkeit		
Beachtung des regionalen Kontextes/weiterer Wirkungsbedingungen		
	Pro	Contra

### 3.3 Das Fixed Effects Panel Modell (FE-Panelmodell)

Auch das Fixed-Effects-Panel-Modell schließt sich nahtlos an die bisherigen Überlegungen an, denn es stellt eine verfeinerte Variante des Difference-in-Difference-Schätzers dar. Im Gegensatz zum Matching, das zu den semiparametrischen Verfahren zählt, ist das FE-Panelmodell ein parametrisches Modell. Die Unterschiede zwischen den Beobachtungseinheiten werden also parametrisiert und so kontrolliert. Der „Trick“ der Methode liegt darin, dass alle Zeitkonstanten-Unterschiede sich durch die Differenzierung herauskürzen, sodass die Zahl der notwendigen Parameter eingeschränkt wird.

### 3.3.1 Methodische Erläuterungen zum FE-Panelmodell

In der DiD wird nicht der absolute Unterschied im Wirkungsindikator des Geförderten und der nicht geförderten Vergleichseinheit gemessen, sondern der Unterschied in der Entwicklung des Wirkungsindikators im Interventionszeitraum (vgl. Tabelle 3.2).

Die DiD-Schätzung im Regressionsmodell stellt eine Panelschätzung in folgender Form dar:

$$\text{Ergebnis} = a_0 * t_0 + b_0 * \text{Teilnehmer} * t_0 + a_1 * t_1 + b_1 * \text{Teilnehmer} * t_1 \quad (1)$$

Dabei beschreibt  $t$  den Zeitpunkt der Beobachtung und ist mit 1 für  $t_0$ <sup>3</sup> codiert und mit 2 für  $t_1$ . Die Ausgangssituation der Nichtteilnehmer mit Blick auf den Ergebnisindikator („Ergebnis“) zeigt sich im Achsenabschnitt  $a_0$ . „Teilnehmer“ drückt den Teilnahmezustand mit 1 für Teilnehmer und 0 für Nicht-Teilnehmer aus, sodass  $b_0$  den Unterschied in der Ausgangssituation von Teilnehmern zu Nicht-Teilnehmern aufzeigt. In gleicher Weise drückt  $a_1$  den Unterschied im Ergebnisindikator für Nicht-Teilnehmer zum Zeitpunkt 2 im Vergleich zum Zeitpunkt 1 aus und  $b_1$  den „Unterschied im Unterschied“ in der Zeit von Teilnehmern zu Nicht-Teilnehmern. Die folgende Tabelle 3.2 zeigt im Vergleich zu Tabelle 3.1 die prinzipielle Äquivalenz der beiden Verfahren der direkten Ermittlung des Unterschieds im Mittelwertvergleich und der Schätzung im Regressionsmodell.

**Tabelle 3.2:** DiD-Schätzung im Panel-Regressions-Modell in schematischer Darstellung

	Ausgangssituation, $t_0$	Nach der Interventionsphase, $t_1$	Unterschiede
<b>Interventionsgruppe</b>	$a_0 + b_0$	$a_0 + b_0 + a_1 + b_1$	$a_1 + b_1$
<b>Vergleichsgruppe</b>	$a_0$	$a_0 + a_1$	$a_1$
			$b_1$

Quelle: In Anlehnung an Buckley und Shang (2003).

Der Vorteil des Regressionsmodells besteht auch darin, dass es die Möglichkeit eröffnet, exogene Faktoren, die sich sowohl in der Zeit verändern als auch unterschiedliche Ausprägungen zwischen den Beobachtungseinheiten aufweisen, durch die Einfügung entsprechender Schätzkoeffizienten zu kontrollieren. Der Nachteil des parametrischen Verfah-

<sup>3</sup>  $t_0$  könnte also auch weggelassen werden, wodurch die Funktion von  $a_0$  als Achsenabschnitt unterstrichen würde. Die hier gewählte Formulierung dient vor allem der Verdeutlichung der Modellstruktur.

rens liegt in den stärkeren Verteilungsannahmen, die mit Blick auf die erklärenden Variablen getroffen werden müssen. Eine zentrale Annahme ist die der additiven Separabilität. Diese besagt mit Blick auf Gleichung (1), dass sich die Einflussgrößen in ihrer Wirkung additiv ergänzen.

### 3.3.2 Beispielstudie zum FE-Panelmodell

In einer Studie zu den Wirkungen der Regionalförderprogramme des EFRE von MAYERHOFER et al. (2008) (s. Kapitel 3.1.2.2) ergänzt eine Difference-in-Difference-Schätzung das zentrale Input-Output-Modell. Da dort die prinzipielle Äquivalenz der DiD-Ermittlung durch den Mittelwertvergleich sowie durch die Schätzung in der Panel-Regression demonstriert wird, ist die Studie für das Verständnis der Methodik sehr hilfreich. Darüber hinaus wird auch demonstriert, wie die Schätzung für unterschiedliche Subsample getrennt erfolgen und weitergehende Informationen aus dieser Differenzierung gewonnen werden können. Aus diesen Gründen wird die entsprechende Studie für den tiefer Interessierten empfohlen.

Ein großer Vorteil der Schätzung im FE-Panelmodell besteht aber in der Möglichkeit, nicht nur die Wirkung einer Teilnahme gegenüber einer Nicht-Teilnahme sondern auch etwa die Wirkung einer stärkeren Förderung im Vergleich zu einer geringeren Förderung festzustellen. Sie ermöglicht darüber hinaus die simultane Schätzung der Wirkung verschiedener Maßnahmen. Eine Studie, die die Methode aufgrund dieser Eigenschaften angewendet hat, ist PETRICK und ZIER (2009). Sie analysieren die Wirkung der Zahlungen der ersten und zweiten Säule der EU auf die Beschäftigungssituation in der Landwirtschaft in Brandenburg, Sachsen und Sachsen-Anhalt. Mit Blick auf die Zahlungen der ersten Säule wird von den Autoren zwischen flächen- und tierbezogenen Zahlungen differenziert. Von den Zahlungen der zweiten Säule wurden die Agrarinvestitionsförderung, die Förderung der Verarbeitung und Vermarktung und Maßnahmen zur Förderung ländlicher Räume entsprechend der Verordnung (EC) 1257/1999 analysiert. Darüber hinaus wurden Zahlungen der Agrarumweltmaßnahmen und der Ausgleichszahlungen für benachteiligte Gebiete aggregiert und berücksichtigt.

Erklärt wird die Beschäftigung in Land- und Forstwirtschaft und Fischerei. Als zusätzliche Kontrollvariablen werden die Entwicklung der regionalen Bevölkerungsdichte sowie des durchschnittlichen Einkommens je Beschäftigten ins Modell aufgenommen. Die Ergebnisse eines FE-Modells, in dem zur Kontrolle der latenten Effekte regionale und zeitliche fixe Effekte beachtet werden, werden im Kontrast zu den Ergebnissen eines einfachen Pooled OLS-Modells ohne Kontrolle der latenten Einflüsse präsentiert. Während letzteres einen hoch signifikanten positiven Effekt für alle Maßnahmen auf die Beschäftigung in der Landwirtschaft aufzeigt, haben die Flächenzahlungen und die Zahlungen zur Unterstützung der Verarbeitung- und Vermarktung nach den Ergebnissen der FE-Modelle

einen signifikant negativen Effekt auf die Beschäftigungssituation. Positiv wirken nur noch die Ausgleichszahlungen für naturräumliche Benachteiligung.

### 3.3.3 Kritische Reflexion der Eignung des FE-Panelmodells

Diese Ergebnisse der Beispielstudie (PETRICK und ZIER, 2009) sind nicht unbedingt erwartet und intuitiv nachvollziehbar. Es muss auch gefragt werden, was genau sie messen. So werden die Ausgleichszahlungen in bestimmten Gebietskulissen gezahlt. Die FE-Schätzung vergleicht nun aber nicht die Entwicklung in den Regionen mit Förderung mit der Entwicklung ohne Förderung, sondern die Unterschiede innerhalb einer Region zwischen den Jahren in Abhängigkeit von der Höhe der jährlichen Zahlungen. Diese Schwankung sollte aber, solange keine Politikänderungen vorliegen, deutlich geringer sein als die Unterschiede zwischen den Regionen. Darüber hinaus müsste durch theoretische Überlegungen ausgeschlossen werden, dass die Schwankungen endogen sind, also selbst das Ergebnis eines strukturellen Wandels in der Landwirtschaft. Indem der ermittelte Effekt nur die Wirkung einer marginalen Schwankung darstellt, wird möglicherweise eine andere Frage beantwortet, als (implizit) gestellt wurde, denn was in den benachteiligten Regionen ohne Förderung geschehen würde bleibt fraglich, da eine entsprechende Referenz ohne Förderung nicht existiert.

Bei der impliziten Interpretation der Ergebnisse in der Studie hingegen, werden hier lineare Effekte der Förderung auf die Arbeitsplatzentwicklung unterstellt. Davon kann nicht automatisch ausgegangen werden, sodass für unterschiedliche Förderintensitäten getrennte Schätzungen durchgeführt werden müssten, um möglicherweise heterogene Effekte zu ermitteln. Weiß man von einzelnen beobachtbaren exogenen Einflussgrößen, dass sie einen nicht-linearen Effekt auf die Entwicklung haben, so muss, ebenso wie für den Fall unterschiedlicher Maßnahmenversionen, der DiD-Vergleich für die verschiedenen Untergruppen getrennt vorgenommen werden. Die stärkste denkbare Differenzierung zwischen Untergruppen erfolgt im Propensity Score Matching (s. letzten Abschnitt), in dem jeweils nur zwei Beobachtungseinheiten miteinander verglichen werden. Hier wird der Vorteil des nicht-parametrischen Matchings gegenüber dem parametrischen Schätzmodell deutlich, in dem additive Separabilität unterstellt werden muss, um eine gemeinsame Schätzung mit allen Beobachtungen simultan zu rechtfertigen.

Diese Argumentation gilt sowohl in der Querschnitts- als auch in der Zeitreihenbetrachtung. So kann eine marginale Veränderung der Förderhöhe in der Zeit einen völlig anderen Effekt haben, als wenn in der Vorperiode nicht und in der Folgeperiode deutlich gefördert wurde. Hier kommt es wieder darauf an, die Fragestellung zu schärfen und die Referenzsituation sorgfältig und inhaltlich begründet zu definieren. Beispielsweise ist es denkbar, dass die Existenz der Fördermaßnahmen der Beispielstudie in den vorausgegangenen Jahren den Strukturwandel bereits deutlich verlangsamt hat – gemessen wird aber hier nur der Effekt der marginalen Veränderung zwischen den Jahren. Über die Wirkung

etwa des Wegfalls einer der Maßnahmen auf die Arbeitsplatzentwicklung kann damit kaum etwas ausgesagt werden. Liegen hinreichend viele Beobachtungen in der Zeit vor, so können die Probleme, die sich aus Autokorrelationen in der Zeit ergeben, im Rahmen eines dynamischen Panelmodels mit entsprechender Kovarianzstruktur abgemildert werden.

Das Problem der fehlenden additiven Separabilität lässt sich am Beispiel unterschiedlicher regionaler Entwicklungspfade illustrieren. So kann die Kombination spezifischer Rahmenbedingungen nach den Theorien des endogenen Wachstums von Regionen zu jeweils spezifischen Wachstumspfaden führen. Zwar werden die zeitkonstanten Unterschiede in den Rahmenbedingungen und die Unterschiede in der Entwicklung, die jeder dieser Faktoren für sich genommen erzeugt, durch das Differenzieren „herausgekürzt“; wenn sich aber aus dem Zusammenspiel eine neue Qualität der Dynamik ergibt, so wird diese durch die Differenzierung nicht eliminiert.

Während mit Blick auf die ILE das größte Problem des Matchings in der Auswahl der geeigneten Vergleichsebene lag, besteht es hier in der Frage, was als Inputindikator gewählt werden sollte. Da die Wirkung der Maßnahme zentral vom Implementationsdesign und insbesondere von der Struktur der geförderten Netzwerke abhängt, erscheint es nicht angebracht, alleine die Fördermittel zu wählen. Da heterogene Fördereffekte erwartet werden, bestünde eine Möglichkeit darin, die Schätzung mit der Höhe der geflossenen Fördermittel als Indikator aber für Gruppen vergleichbarer Projekte getrennt durchzuführen. Die Gültigkeit auch der anderen diskutierten Annahmen ebenso wie das generelle Problem der Fristigkeit eintretender Wirkungen bleibt aber bestehen.

### 3.3.4 Tabellarische Bewertung des FE-Panelmodells

	1	2	3	4	5	6	7
<b>Technische Aspekte</b>	sehr gering			sehr hoch			
Zeitbedarf		■					
Anforderungen an technisches Know-How				■			
Anforderungen an Kontextwissen/Theoretisches Wissen		■					
Softwarebedarf				■			
Datenanforderungen				■			
Sonstige Kosten		■					
<b>Inhaltliche Aspekte</b>	sehr gut möglich			nicht möglich			
Erfassung direkter Effekte				■			
Erfassung indirekter Effekte			■				
Direkte Modellierung der Interventionslogik					■		
Erfassung diskontinuierlicher, auch zeitverzögerter Effekte				■			
Erfassung positiver externer Effekte					■		
Erfassung von Allokationseffekten					■		
Verallgemeinerbarkeit		■					
Beachtung des regionalen Kontextes/weiterer Wirkungsbedingungen		■					
	Pro			Contra			

## 4 Theoriebasierte Evaluation

Ein kontrollierter Versuch dient dem Ziel, Wirkungen in Bereichen zu identifizieren, in denen eine eindeutige theoretische Ableitung von Wirkungszusammenhängen auf theoretischer Basis (noch) nicht möglich ist. Kann kein kontrollierter Versuch durchgeführt werden, so müssen, um das Kontrafaktische dennoch zu konstruieren, in jedem Fall bestimmte Annahmen getroffen werden. Die Diskussion der Methoden in den bisherigen Abschnitten hat gezeigt, dass gerade in komplexen sozioökonomischen Systemen die notwendigen Annahmen oft nicht als gegeben angenommen werden können. Sie stellen letztlich Theorien dar, die durch die Hintertür eingeführt werden.

Im aktuellen Abschnitt geht es im Gegensatz dazu um Ansätze, in denen erwartete theoretische Zusammenhänge explizit modelliert werden. Wenn eine allgemein anerkannte Theorie oder auch nur ein allgemein bekannter empirischer Zusammenhang („stylised fact“) existieren, die den erwarteten Zusammenhang zwischen Intervention und Ergebnisindikator bestimmen, dann kann der erwartete funktionale Zusammenhang auch genutzt werden, um anhand empirischer Indikatoren in entsprechenden ökonometrischen Modellen zu überprüfen, ob die Intervention wie erwartet gewirkt hat. Es ist dann nicht nötig, jeden kausalen Zusammenhang der Intervention einzeln zu überprüfen, vielmehr stammen aus der Wissenschaft in Bereichen mit anerkannten Theoriegebäuden in der Regel Formeln in reduzierter Form, die den Netto-Zusammenhang zwischen verschiedenen Größen beschreiben.

In den Bereichen, in denen allgemein anerkannte Theorien zur Erklärung der relevanten Phänomene existieren, kann die Wirkung einer Intervention bereits aus der Theorie abgeleitet werden. Dieser Ansatz wird in den „modellhaft-simulierenden“ Methoden des folgenden Kapitels 4.1 verfolgt. Existieren derart anerkannte allgemeingültige Theorien, so wird sich in der Regel auch die Interventionslogik an diesen orientieren. Dieses ist bei den meisten makroökonomischen Interventionen zum Beispiel im Bereich der Geldpolitik der Fall. Werden entsprechende Interventionen überprüft und es zeigt sich, dass die erwarteten Effekte nicht eingetreten sind, so wird das zunächst den Rahmenbedingungen und den Umständen der Implementierung der Intervention zugeschrieben werden. Erst wenn sich entsprechende Ergebnisse häufen, wird die zugrunde gelegte Interventionslogik und damit die theoretische Grundlage der Intervention selbst hinterfragt werden.

Für die theoriebasierte Analyse ist es also zentral, dass der Zusammenhang zwischen Intervention und Wirkungsindikator funktional bekannt ist. Im Kapitel 2.1 wurde bereits deutlich, dass die ILE einer Interventionslogik folgt, die nicht den theoretischen Annahmen, die zum Beispiel einer Input-Output-Analyse zugrunde liegen, entspricht. Sie folgt neueren regionalökonomischen Theorieansätzen. Dabei muss auf der Suche nach einem „passenden“ Theorieansatz oder im Versuch der Wiederentdeckung des (implizit) zugrunde liegenden Theorieansatzes zunächst grundsätzlich festgestellt werden, dass die

ILE weniger darauf angelegt ist, die technische Effizienz in der existierenden Produktion zu erhöhen. Vielmehr wird implizit davon ausgegangen, dass Produktionsfaktoren aufgrund von Koordinationsproblemen brach liegen und nicht voll genutzt werden. In den Theorien zu Agglomerationseffekten in verschiedenen Regionen, die auf Myrdal zurückgehen, wird aber angenommen, dass sich, wenn der Markt größer wird, entweder die Produktion unter dem Einsatz der gleichen Faktoren ausdehnt, oder aber die Kosten bei gleicher Produktionsmenge abnehmen (CAPELLO, 2007). Neuere Ansätze erlauben es, die Entwicklung der Totalen Faktorproduktivität den drei zentralen Elementen einzeln zuzuschreiben: Technischem Fortschritt (eine Verschiebung der Grenzproduktionsfunktion, „Frontier“), zunehmende Effizienz (eine Bewegung der einzelnen Beobachtungen in Richtung der Frontier) und Kapitalakkumulation (Eine Bewegung auf der Frontier) (SALINAS-JIMENEZ, 2002). Koordinationsprobleme und Unterausnutzung von Faktoren spielen hier keine Rolle. Erst die auf Krugmann zurück gehenden Modelle der Neuen Ökonomischen Geographie führen die unterstellten positiven Skaleneffekte auf Unternehmensebene auf unvollkommene Märkte zurück (ibid.) und damit durchaus auf bestimmte Koordinationsprobleme.

Die Simulationsmethoden, die in aller Regel mit Modellen arbeiten, die noch nicht von der Möglichkeit positiver Skaleneffekte und den weitergehenden Erkenntnissen der Neuen Ökonomischen Geographie und ihrer Nachfolger beeinflusst sind, können daher kaum die erwarteten Effekte neuerer regionalpolitischer Ansätze abbilden. Sie werden im Folgenden dennoch diskutiert, um Möglichkeiten und Grenzen des Einsatzes solcher Modelle für die Analyse von Regionalpolitik weiter zu verdeutlichen.

Die Politik der ILE ist also Teil eines Paradigmenwechsels in der Regionalförderung. Dieser Wechsel der Politik stellt die andere Seite der Medaille eines entsprechenden Wechsels in der Regionalökonomie dar, in der Netzwerke, Cluster und „Knowledge-Spillover“ zentrale Rollen einnehmen. Allerdings ist die Ausbildung von Marktgleichgewichten in den entsprechenden Modellen nicht mehr eindeutig determiniert, sondern hängt vielmehr sensibel von der anfänglichen Ausstattung mit Arbeit und Kapital, aber auch von den Annahmen über die Agglomerationseffekte selbst ab (vgl. CAPELLO, 2007, S. 246 ff.). Insbesondere letztere unterscheiden sich vermutlich je nach regionaler Produktionsstruktur und sind vermutlich auch nicht-linear. Quantitative Informationen über die Wirkung von Wissens-Spillovern oder über Netzwerkeffekte unter verschiedenen Marktbedingungen gibt es kaum (vgl. hierzu TÖDTLING und TRIPPEL, 2005). Es gibt also (noch) keine Theorie in dem Sinne, dass die notwendigen und hinreichenden Bedingungen einer bestimmten Entwicklung in Gegenwart von Agglomerationseffekten identifiziert oder gar quantifiziert und in einen festen funktionalen Zusammenhang gebracht worden wären. Es existiert kein Gleichungssystem für dieses Entwicklungsparadigma, in das lediglich der Situation angepasste Parameter eingesetzt werden müssten, um dann durch den Vergleich verschiedener Simulationen die Wirkung zu isolieren.

In welchem Zusammenhang die durch die ILE angestrebte zunehmende Vernetzung von Akteuren zur Ausnutzung brach liegender Faktoren beiträgt, ist nicht allgemein bekannt. Ebenso wenig ist bekannt, in welchem Ausmaß und auf welchen Wegen die bessere Nutzung von Produktionsfaktoren (welcher?) zu einem Zuwachs an Arbeitsplätzen beiträgt. Eine Simulation von entsprechenden Politiken, wie es im Folgenden am Beispiel von Investitionsfördermaßnahmen der Regionalpolitik gezeigt wird, ist daher (noch) nicht möglich.

## 4.1 Simulationsstudien

Je umfassender das anerkannt wissenschaftliche Erklärungsmodell ist, das einer Politikbewertung zugrunde gelegt wird, umso weniger muss das Modell an die jeweiligen spezifischen Bedingungen der Intervention angepasst werden, weil die Parameter endogen durch als konstant anerkannte Gesetzmäßigkeiten (funktionale Zusammenhänge) bestimmt werden. Dabei ist es letztlich eine Ermessensfrage, in wiefern zum Beispiel Elastizitäten in der Produktionsfunktion als konstant angesehen werden, oder aber technische Unterschiede in Raum und Zeit mit in Betracht gezogen werden. Wird auch der technische Fortschritt modellendogen bestimmt, so werden die Elastizitäten modellendogen angepasst, ohne dass spezifische Parameter immer wieder neu geschätzt werden müssen.

Es wird deutlich, dass sich empirische und Simulationsstudien nicht scharf voneinander trennen lassen, sondern vielmehr ein Kontinuum darstellen. Dennoch wird die entsprechende Unterscheidung hier getroffen. Die vorzustellenden Simulationsstudien basieren auf der Anwendung allgemeiner Regionalmodelle, die alle wirtschaftlichen Zusammenhänge innerhalb einer Region und in einem Fall auch die Zusammenhänge zwischen verschiedenen Regionen gemeinsam betrachten. Im Gegensatz wird unter den empirischen Studien ein Partialmodell vorgestellt, in dem durch nur drei Gleichungen die Entwicklung eines Sektors der Region beschrieben wird. Die Entwicklung der Nachfrage nach Produkten wird dann im Gegensatz etwa zum SAM-Modell (Kapitel 4.1.1) nicht endogen bestimmt, sondern exogen geschätzt. Andersherum beruht auch für die allgemeinen Regionalmodelle die Simulation der Entwicklung der Nachfrage aus dem Ausland auf Modell-exogenen Schätzwerten.

### 4.1.1 Die Regionale „Social Accounting Matrix“ (SAM)

Diese Methodik wurde als Einstieg in die theoriebasierte Evaluation gewählt, weil sie die Bedeutung realer indirekter Effekte für die Beurteilung von Interventionen sehr plastisch aufzeigt. Gleichzeitig lässt sich aber an dieser im Prinzip einfachen Methodik die Problematik von theoretischen Annahmen, die die oft implizite Interventionslogik nicht adäquat widerspiegeln, gut diskutieren.

### 4.1.1.1 Methodische Erläuterungen zur SAM

Die SAM ist ein Bilanzierungswerkzeug, das die strukturellen Beziehungen zwischen allen relevanten Agenten einer Wirtschaft abbildet. In den Zellen der Matrix wird abgetragen, welcher Betrag vom jeweiligen Spaltenkonto an das entsprechende Zeilenkonto fließt. Die Summe der Einträge einer Zeile beschreibt das gesamte Einkommen, das dem entsprechenden Konto von allen anderen im betreffenden Zeitraum zugeflossen ist. Andersherum beschreibt die Summe der Einträge einer Spalte die gesamten Ausgaben, die von dem betreffenden Konto im entsprechenden Zeitraum an alle anderen Konten geflossen sind. Im Gleichgewicht ist das totale Bruttoeinkommen gleich den totalen Bruttoausgaben, d. h. die Zeilensummen sind gleich den korrespondierenden Spaltensummen (VOGEL, 1994). Eine zusammenfassende Beispielmatrix ist in Tabelle 4.1 dargestellt.

Die Matrix zeigt, dass in der SAM, anders als in der traditionellen Input-Output-Analyse, auch die Verteilungswirkungen von Gleichgewichtsänderungen erfasst werden. Unter den Faktoreinkommen befindet sich auch das Einkommen aus Arbeit, sodass mit der SAM eine direkte Bewertung von Arbeitsmarkteffekten ermöglicht wird.

**Tabelle 4.1:** Beispiel einer SAM

Konten:	PA	FE	INS	RdW	REG	KAP
<b>Produktionsaktivitäten</b>	A	0	C	Ex	G	I
<b>Faktoreinkommen</b>	V	0	0	0	0	0
<b>Institutionen</b>	0	Y	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	0
<b>Rest der Welt</b>	Im	0	0	0	0	0
<b>Regierung</b>	T <sub>1</sub>	0	T <sub>2</sub>	0	0	0
<b>Kapitalkonto</b>	0	0	S <sub>P</sub>	S <sub>F</sub>	S <sub>G</sub>	0

Quelle: Vogel (1994).

Anmerkungen: A: Matrix der Input-Output-Ströme

V: Verteilung der Einkommen aus Produktionsaktivitäten auf die Faktoren

Im: Zahlungen für Importe

T<sub>1</sub>: Produktionsabgaben an den Staat wie Steuern und Zölle

Y: Re-Allokation von Faktoreinkommen an Eigentümer, Haushalte und Unternehmen

C: Konsumausgaben dergleichen Institutionen

R<sub>1</sub>: Umverteilung zwischen Haushalten und Unternehmen

T<sub>2</sub>: Steuern und Abgaben von Haushalten und privaten und öffentlichen Unternehmen

S<sub>P</sub>: Sparen der verschiedenen privaten Institutionen

Ex: Verkauf an den Rest der Welt = Exporte

R<sub>2</sub>: Geldanweisungen aus dem Ausland

S<sub>F</sub>: Sparen aus dem Ausland

G: Güterkauf durch die Regierung

R<sub>3</sub>: Transfers von Einkommen an Haushalte und Firmen

S<sub>G</sub>: Sparen der Regierung

I: Investitionen in Kapitalgüter

Die drei makroökonomischen Gleichgewichtsmomente sind in der Matrix enthalten: die aggregierte Nachfrage ist gleich dem aggregierten Angebot; Die Summe der Investitionen sind gleich dem Spareinkommen und der gegenwärtige Kapitalzuwachs ist gleich dem Netto-Kapitalzufluss/-abfluss (VOGEL, 1994):

$$C + I + G + (Ex-Im)=V+T_1 \quad (1)$$

$$S_P + S_G + S_F = I \quad (2)$$

$$Im - (Ex + R_2) = S_F \quad (3)$$

Das SAM-Modell gehört daher zur Gruppe der komparativ-statischen Allgemeinen Gleichgewichts Modelle. Diese Modelle haben den Anspruch den „globalen ökonomischen Effekt“ und nicht nur die direkten Effekte politischer Interventionen zu quantifizieren (PSALTOPOULOS et al., 2004).

Auch für die SAM ist die Input-Output-Matrix zentral. Diese wird ausgehend von nationalen Input-Output-Tabellen geschätzt. Diese Schätzungen basieren in der Regel auf sogenannten beschäftigungsbasierten „lokalen Quotienten“, die dem Verhältnis des Anteils der Beschäftigten einer Industrie in einer Region zu dem Anteil der Beschäftigten der gleichen Industrie in der Gesamtwirtschaft entsprechen. Ist dieser Quotient größer als eins, so wird angenommen, dass in der Region mehr des betreffenden Gutes produziert wird als benötigt wird – die entsprechende „exportierende“ Industrie wird als sog. „Base-industry“ angesehen. Der nationale Multiplikator wird übernommen. Ist hingegen der Quotient kleiner als eins, so wird der nationale Multiplikator mit dem entsprechenden Quotienten multipliziert, um den regionalen Multiplikator zu bestimmen. Offensichtlich beruht ein solches Vorgehen auf einigen pragmatischen Annahmen. Die Alternative besteht darin, durch die Erhebung von Primärdaten in Umfragen bei Unternehmen Informationen zu den Multiplikatoren zu erhalten. Das entspricht einem kaum zu leistenden Aufwand. PSALTOPOULOS et al. (2004) haben einen Mittelweg gewählt und an zentralen oder besonders fragwürdigen Stellen der Input-Output-Matrix, die Grundlage ihrer SAM war, die Koeffizienten durch belastbarere Werte aus eigenen Erhebungen ebenso wie aus ergänzenden Sekundärstatistiken ersetzt. Auch die Koeffizienten der anderen Teile der SAM werden im Rückgriff auf vielfältige Sekundärstatistiken und eigene Erhebungen bestimmt.

#### 4.1.1.2 Beispielstudie zur Regionalen SAM

PSALTOPOULOS et al. (2004) haben für einige ausgewählte strukturschwache ländliche Regionen Europas die Wirkung der Europäischen Agrarstrukturpolitik beispielhaft und vergleichend mit Regionalen SAM analysiert. Neben anderen Wirkungen möchten die

Autoren auch die Wirkung der betreffenden Politikmaßnahmen auf die Beschäftigung in den ausgewählten Regionen untersuchen. Da unterstellt wurde, dass die Maßnahmen unterschiedliche Effekte in verschiedenen Regionen haben, wurden in den beteiligten drei Ländern Griechenland, Finnland und Schottland jeweils zwei abgelegene ländliche Regionen ausgewählt, von denen eine durch eine diversifizierte Ökonomie und die andere durch starke Fokussierung auf die Landwirtschaft gekennzeichnet waren. Die Studie baut auf der Förderung der Jahre 1989 bis 1993 auf. Neben strukturpolitischen Maßnahmen im engeren Sinne wurde auch die Wirkung der Mittel der ersten Säule der CAP als reine Transfers untersucht. Bei den geförderten Projekten der Strukturpolitik dominierten Investitions- und Infrastrukturprojekte in der Landwirtschaft. Insgesamt wurde die Projektdurchführung stark durch den öffentlichen Sektor dominiert: „[...] private sector responses were at low levels because of locally constrained entrepreneurial capacity“ (PSALTOPOULOS et al., 2004). LEADER-Maßnahmen waren bereits eingeschlossen.

Es gab deutliche Unterschiede in der Mittelverteilung zwischen den Regionen und innerhalb der Regionen. Insbesondere in einer der griechischen Regionen trugen die Mittel deutlich zur regionalen Nachfrage bei, während das insbesondere für eine der schottischen Regionen in sehr viel geringerem Ausmaß galt. Die vor allem auf der Landwirtschaft beruhenden Regionen erhielten vor allem Mittel aus der ersten Säule, von denen keine weiteren Kapazitätsanpassungseffekte für die weitere Wirtschaft ausgehen. Mittel aus der Strukturpolitik flossen eher an die diversifizierten Regionen und hier vielfach in die Verbesserung der Straßen- und Informations-Infrastruktur. Die Wirkungen der Ausgaben waren besonders hoch in der gemischten griechischen Region, in den anderen aber bedeutend geringer. Das gilt vor allem für die Beschäftigungseffekte.

Die Wirkungen aus der Kapazitätsanpassung nach Infrastrukturprojekten waren ebenfalls für die griechischen Regionen am höchsten, für die diversifizierte aufgrund von Kapazitätserweiterungen im Tourismus-Bereich, für die landwirtschaftlich geprägte aufgrund von Investitionen in die Fischerei und in die Umstrukturierung von Bewirtschaftungsflächen. Die Autoren simulieren in den regionalen SAM auch die Wirkungen unterschiedlicher zukünftiger Politikszenerarien. Die Schlussfolgerungen hängen wesentlich davon ab, wie sich die Verteilung der regionalen Mittelströme unter verschiedenen Politikszenerarien ändern.

Insgesamt ergibt die Analyse, dass die Transfers der CAP geringere ökonomische Wirkungen haben als die strukturpolitischen Maßnahmen. Aus diesem Grunde waren auch die Wirkungen der eingesetzten Mittel in diversifizierten Regionen deutlich höher als in landwirtschaftlich geprägten.

### 4.1.1.3 Kritische Reflexion der Eignung Regionaler SAM

SAMs sind gut geeignet, um die Multiplikator- und Verteilungswirkungen von Einkommenstransfers und allgemeine Zahlungen, die in erster Linie auf Kapazitätserweiterungen und eine Ausdehnung der Zahl eingesetzter Arbeitskräfte abzielen, zu ermitteln. In Bezug auf die Analyse gezielter Strukturpolitiken bleiben sie aber oft unbefriedigend. So überraschen die Ergebnisse der Studie von PSALTOPOULOS et al. wenig. Es entsteht ein wenig der Eindruck, als würde schon eine Analyse der Mittelverteilung zwischen den Regionen und zwischen den Sektoren sowie dahingehend, ob es sich um strukturelle oder reine Transfermaßnahmen handelt, in Kenntnis der getroffenen Annahmen aber ohne aufwendige Datensammlung zu den gleichen Schlussfolgerungen führen. Es gibt in keinem Fall negative Auswirkungen einer Intervention zu beobachten, wie sie etwa bei strukturkonservierenden Eingriffen theoretisch durchaus möglich wären, wenn die Referenz in der alternativen Verwendung der eingesetzten Mittel besteht. Im Folgenden wird dargelegt, was die Hintergründe dieses skeptischen Fazits sind.

Die Datenanforderungen bei der Konstruktion Regionaler SAM sind sehr hoch (PSALTOPOULOS et al., 2004) und es ergeben sich in aller Regel Restriktionen in der Modellformulierung aus der Datenverfügbarkeit. Der Ausgangspunkt der Konstruktion Regionaler SAM ist in der Regel eine regionale Input-Output-Tabelle, deren Koeffizienten ausgehend von nationalen Input-Output-Tabellen geschätzt wurden. Der Trade-Off zwischen Belastbarkeit und Aufwand ist hier sehr deutlich.

Die Wirkung politischer Interventionen wird in einer Art Domino-Effekt gesehen. Je nachdem an welcher Stelle der Staat von außen Mittel an Konten der Region zahlt, ergeben sich unterschiedliche Multiplikatoreffekte, die zu weiterer Nachfrage, entsprechender Produktion und Einkommensgenerierung führen. Diese zugrunde gelegte Funktionsmodell der Wirtschaft hat zwei wichtige Implikationen mit Blick auf die Bewertung von ökonomischen Effekten von Interventionen: Zum einen werden mögliche Effekte, die sich aus dem Konzept der endogenen Entwicklung ergeben, wie positive Skaleneffekte, Agglomerationseffekte, positive Spillovers etc., und die mit hoher Wahrscheinlichkeit nicht linear sind, in den konstanten, zu einem bestimmten Zeitpunkt geschätzten Multiplikatoren nicht erfasst (1); zum anderen wird im Multiplikator-konzept implizit von der Existenz ungenutzter oder zu wenig genutzter Faktoren, wie etwa im Fall von Arbeitslosigkeit ausgegangen. Existieren diese Faktoren nicht (mehr) in dem Ausmaß in dem sie zur Realisierung des Multiplikators benötigt würden, so überschätzt die Simulation die positiven Wachstumseffekte (2). Das kann einerseits dadurch erklärt werden, dass der Multiplikator in der Gegenwart solcher Kapazitätsbegrenzungen abnimmt, und alternativ dadurch, dass dann auftretende Verdrängungseffekte oder Opportunitätskosten nicht beachtet werden.

Während Effekt 2 besonders die Eignung der SAM zur Analyse der Wirkung, etwa von Maßnahmen der Investitionsförderung, in Frage stellt, da in solchen Maßnahmen die

möglichen negativen Allokationseffekte von zentraler Bedeutung sind, ist Effekt 1 besonders kritisch für Maßnahmen wie die zur Integrierten Ländlichen Entwicklung, die ja gerade darauf abzielen, das endogene Wachstumspotenzial einer Region durch die Nutzbarmachung von Netzwerkeffekten und Humankapital zu erhöhen. Anders ausgedrückt zielen diese Maßnahmen weniger auf einen direkten Anstoß der Dominosteine durch einen exogenen Kapitalzuschuss an der richtigen Stelle ab, sondern vielmehr auf eine Erhöhung der Multiplikatoren an zentralen Stellen der SAM. Auch andere Maßnahmen zielen eher auf eine Änderung der Multiplikatoren ab. Die folgende Liste möglicher Ursachen für Änderungen in den Multiplikatoren über die Zeit (CONWAY, 1977) soll das noch einmal verdeutlichen:

- Technischer Wandel,
- Zu- oder abnehmende Skaleneffekte,
- Veränderungen im Produkt-Mix bestimmter Industrien,
- Entstehen neuer Industrien,
- Änderungen der relativen Preise wenn die Koeffizienten monetär fundiert sind,
- Substituierung von Inputs z. B. in Reaktion auf Preisänderungen oder technischen Wandel sowie die
- Änderung in Handelsmustern z. B. durch Substituierung von Importen und Verlagerung von Produktion über die Grenze.

Ein weiteres zentrales Problem der Input-Output- oder der SAM-Analyse besteht in der Wahl des geeigneten Zeithorizontes. Diese in der Einführung angesprochene Problem, dass die Wirkungszeitpunkte einer Intervention oft nicht bekannt sind, insbesondere wenn die Interventionslogik unklar bleibt, verstärkt die Schwierigkeit der Entscheidung für den Zeitraum, über den ein Multiplikator wirken soll. Wird dieser Zeitraum länger gewählt, so beinhaltet der empirisch bestimmte Multiplikator auch Effekte, die aus Investitionen und strukturellen Anpassungen stammen. Das kann zwar einerseits erwünscht sein, andererseits aber auch die logische Konsistenz des Modells unterminieren, in dem strukturelle Anpassungen im Prinzip nicht vorgesehen sind. Gerade im Bereich der ILE aber geht es ja darum, durch einen minimalen anfänglichen Stimulus und resultierende Netzwerkeffekte unter Umständen zu einem deutlich späteren Zeitpunkt auch strukturelle Änderungen der regionalen Ökonomie hervorzurufen. Diese wären in statischen Multiplikatoren, aber insbesondere in solchen, die für relativ kurze Zeiträume kreiert wurden, nicht enthalten.

Zusammenfassend betrachtet liegt der zentrale kritische Punkt der SAM und der Input-Output-Analyse in der rigiden Modellstruktur: Eine Erfassung spezifischer Interventionslogiken und der Intervention zugrunde liegender Annahmen ist kaum möglich: „Within the SAM analytical framework, impact analysis [...] generates changes in output, employment, etc. by applying multiplier and coefficient values to the injections of expen-

diture (treated as additional final demand) associated with structural policy.“ Das bedeutet, dass die Art und Weise in der das Geld im Detail ausgeschüttet wird keine Rolle spielt (MOSLEY und SOL, 2001). Ob es nun direkt an einen Produzenten fließt oder aber die Ausschüttung an komplexe Bedingungen wie etwa die Bildung lokaler Aktionsgruppen in der ILE geknüpft wird, ist für die Analyse in der SAM irrelevant; entsprechende Effekte können aber auch nicht erfasst werden<sup>4</sup>.

#### 4.1.1.4 Tabellarische Bewertung der RSAM-Modelle

	1	2	3	4	5	6	7
<b>Technische Aspekte</b>	sehr gering			sehr hoch			
Zeitbedarf							
Anforderungen an technisches Know-How							
Anforderungen an Kontextwissen/Theoretisches Wissen							
Softwarebedarf							
Datenanforderungen							
Sonstige Kosten							
<b>Inhaltliche Aspekte</b>	sehr gut möglich			nicht möglich			
Erfassung direkter Effekte							
Erfassung indirekter Effekte							
Direkte Modellierung der Interventionslogik							
Erfassung diskontinuierlicher, auch zeitverzögerter Effekte							
Erfassung positiver externer Effekte							
Erfassung von Allokationseffekten							
Verallgemeinerbarkeit							
Beachtung des regionalen Kontextes/weiterer Wirkungsbedingungen							
	Pro			Contra			

<sup>4</sup> In der Literatur wird manchmal unterschieden zwischen „Impact Evaluation“ und „Process Evaluation“. Letztere beschäftigt sich nach dieser Klassifizierung vor allem damit, wie die Implementierung einer Maßnahme ihre Wirkungen beeinflusst (MOSLEY und SOL, 2001). Mosley und Sol identifizieren Dezentralisierung, kooperative Netzwerke und Vertragslösungen als die großen Trends in der Politikimplementierung (ibid.). In der Analyse der Vorzüglichkeit bestimmter Implementierungsformen kann einerseits der direkte damit angestrebte Erfolg, wie etwa die Netzwerkentwicklung, gemessen werden oder aber die mittelbare Wirkung einer bestimmten Form der Implementierung wird im Erklärungsmodell kontrolliert und Wirkungsindikator ist doch wieder die Entwicklung der Arbeitsplätze. Der Vorteil der Analyse des unmittelbaren mit der spezifischen Implementierung angestrebten Erfolgs hat aber gerade den Vorteil, dass ein linearer und zeitlich wenig verzögerter Zusammenhang eher unterstellt werden kann.

## 4.1.2 Das ökonometrische Input-Output-Modell (Ö-I-O-Modell)

Die Methodik der ökonometrischen Input-Output-Modellierung schließt sich an die zuvor getroffenen Überlegungen an, weil sie es ermöglicht, einige der Schwächen teilweise zu überwinden, die im vorausgegangenen Kapitel mit Blick auf die Bewertung von Arbeitsplatzeffekten insbesondere der ILE mithilfe von Regionalen SAM angesprochen wurden.

### 4.1.2.1 Methodische Erläuterungen zum Ö-I-O-Modell

Gegenüber den SAM stellen die I-O-Modelle eine Vereinfachung dar, indem die Verteilung der Einkommen zwischen den „Spielern“ der Volkswirtschaft nicht explizit modelliert wird. Die Erweiterung in dem hier vorgestellten Modelltyp gegenüber der SAM des letzten Kapitels besteht im Wesentlichen in der Ergänzung der statischen Matrix durch „ökonometrisch geschätzte Zeitreihenmodelle, welche die aus der ökonomischen Theorie abgeleiteten Beziehungen zwischen verschiedenen Variablen (z. B. privater Konsumnachfrage und Haushaltseinkommen, Produktion und Beschäftigung etc.) empirisch quantifizieren und den dynamischen Veränderungen eines Wirtschaftssystems Rechnung tragen“ (MAYERHOFER et al., 2008). Damit wird der beklagte statische Charakter der klassischen I-O-Modelle teilweise überwunden. Die Schätzung des Nachfrageeffektes der Förderung erfolgt in diesem Modell allerdings nicht prinzipiell anders, als im vorausgegangenen Kapitel beschriebenen I-O-Modell.

Die zentrale Neuerung liegt zumindest in dem im Folgenden zu beschreibenden Modell in der Bestimmung des Preiseffektes auf Angebotsseite. Seine Berechnung beruht auf einer ökonometrisch geschätzten Kostenfunktion. Über die ebenfalls geschätzten Außenhandelselastizitäten führt ein geringerer Preis zu einem erhöhten Absatz auf dem Weltmarkt, bzw. zu verringerten Importen. Der Kapazitätseffekt beruht auf dem empirisch bestimmten Zusammenhang zwischen Kapitalstock und Output eines Sektors. Die zusätzliche Produktion wird annahmegemäß ebenfalls im Export abgesetzt (MAYERHOFER et al., 2008).

Ein weiterer Unterschied zum vorausgegangenen Beispiel besteht im Folgenden in der simultanen Modellierung aller Regionen eines Landes unter Nutzung empirisch bestimmter regionaler Im- und Exportströme zwischen den Regionen. Das hat den großen Vorteil, dass die regionale Verteilung der indirekten und induzierten Effekte in der gesamten Volkswirtschaft abgeschätzt werden kann.

### 4.1.2.2 Beispielstudie zum Ö-I-O-Modell

Bei der Beispielstudie handelt es sich ausdrücklich um eine Pilotstudie, die vor allem von methodischem Interesse ist, da die Analyse selbst sich noch zu vielen Datenrestriktionen gegenüber sah, um belastbare Ergebnisse liefern zu können (MAYERHOFER et al., 2008). Die Studie basiert auf der Anwendung eines Regionalen I-O-Modells für ganz Österreich, MultiReg, das zuvor für andere Fragestellungen in Kooperation zweier österreichischer Institutionen (WIFO und Institut für Regionalentwicklung und Technologiepolitik der Joanneum Research) entwickelt worden war: „MultiREG ist ein multiregionales, multi-sektorales ökonometrisches Input-Output-Modell und bildet die wirtschaftlichen Verflechtungen auf der Ebene von 32 Wirtschaftsbranchen bzw. Gütern und den neun Bundesländern ab. Es modelliert die sektoralen Produktions- und Konsumbeziehungen innerhalb eines Bundeslandes wie auch jene zwischen den Bundesländern (interregionaler Handel) und dem Ausland (internationaler Handel der Regionen). Damit werden auf Basis von theoretisch wie empirisch ermittelten Wirkungsketten die für einen Wirtschaftsraum typischen Kreislaufzusammenhänge zwischen Nachfrage, Produktion, Beschäftigung und Einkommen abgebildet“ (MAYERHOFER und FRITZ, 2009). Eine genaue Beschreibung dieses Modells findet sich in FRITZ et al. (2005).

Analysiert wurde mit dem Modell die Wirkung der EFRE-Fördermittel der Jahre 1995 bis 2007. Gemessen an den ausgezahlten Mitteln je Beschäftigten war die Förderintensität in den strukturschwachen ländlichen Regionen ungefähr drei mal höher als in den sachkapitalintensiven Regionen (MAYERHOFER und FRITZ, 2009). Die Gelder flossen überwiegend an Unternehmen in Industrie und Gewerbe und dienten hier vor allem der Finanzierung von Investitionen. Der weitaus größte Teil der restlichen Mittel floss an Tourismus und Freizeitwirtschaft, Forschung und Entwicklung sowie in die Bereiche Energie-, Umwelt- und Verkehrsinfrastruktur sowie Raumplanung und Sanierung (ibid.).

Mit Blick auf die nachfrageseitige Wirkung der Förderung berechnen die Autoren einen Multiplikator, der direkte, indirekte und induzierte Effekte umfasst. Induzierte Effekte entstehen aus den Beschäftigungs- und Einkommenseffekten der Förderung und bestehen über den Zeitraum der eigentlichen Förderung hinaus. Insgesamt beträgt der Multiplikator über den ganzen Zeitraum hinweg betrachtet 1,5, das heißt, 1 € an EFRE-Mitteln erhöht nachfrageseitig die nationale Bruttowertschöpfung um 1,50 €. Regional betrachtet entstehen, wie die Autoren zeigen, diese Wirkungen nicht unbedingt dort, wo die Mittel eingesetzt wurden. Die indirekten und induzierten Effekte entstehen vielmehr oft außerhalb der Förderregion, was naturgemäß umso eher der Fall ist, je kleiner die jeweilige Region ist. Andersherum erhält Wien nur einen sehr kleinen Teil der Fördermittel, profitiert anteils-

mäßig aber am stärksten von den Wirkungen.<sup>5</sup> Eine ähnliche Argumentation gilt auf sektoraler Ebene; so erhält das Baugewerbe mehr als ein Drittel der Mittel, sein Anteil am Gesamteffekt liegt aber nur etwa bei einem Fünftel.

Für Zahlungen für Unternehmensinvestitionen in Industrie und Gewerbe wurde auch der angebotsseitige Effekt berechnet. Das Investitionsvolumen wird dabei ohne Berücksichtigung von Abschreibungen in Relation zum Kapitalstock im Ausgangsjahr gesetzt. Für ursprünglich kleine Branchen in kleinen Regionen kommt so unter Umständen bei faktischen Neugründungen von Sektoren ein extrem hoher Quotient heraus. Ihrer Wirkung nach müssen Ersatz- und Erweiterungsinvestitionen unterschieden werden, die entweder nur einen Preiseffekt oder aber auch einen Kapazitätseffekt erwarten lassen (ibid.). Informationen zu dieser Unterscheidung liegen den Autoren aber nicht vor, sodass Ad-hoc-Annahmen getroffen werden müssen. Während nach den Ergebnissen die nachfrageseitigen Effekte in den ersten Jahren dominieren, gilt das Umgekehrte für die Folgejahre, in denen die Wirkung der Erhöhung des Kapitalstocks auf die Im- und Exporte dominiert (MAYERHOFER und FRITZ, 2009).

#### 4.1.2.3 Kritische Reflexion der Eignung des Ö-I-O-Modell

Die Autoren unterstreichen, dass sie keine Möglichkeit haben, das Ausmaß der Mitnahmeeffekte bei den geförderten Projekten zu bewerten. Sie arbeiten mit der Annahme, dass die Fördermittel, aber nur diese, zusätzlich eingesetzt wurden und erklären, andere Studien würden noch pessimistischere Einschätzungen nahelegen (MAYERHOFER und FRITZ, 2009). Die Autoren sprechen auch noch ein anderes, grundsätzliches Problem aller vergleichbaren Analysen an: Geht es um nationale Mittel, so bedeutet ihr Einsatz an einer Stelle bei gegebenem Budget, dass die staatlichen Mittel an anderer Stelle nicht eingesetzt werden können – es fallen also Opportunitätskosten an. MAYERHOFER et al. lösen das Problem bei den Nachfrageeffekten, indem sie nur die Gelder von der EU untersuchen, die Frage nach der Wirkung der Kofinanzierungsmittel bleibt damit aber unbeantwortet. Es wird aber unterstellt, dass kompensierend der öffentliche Konsum um die entsprechenden Mittel proportional auf Bundes- und Landesebene gesenkt wird (ibid.). Auch an anderen Stellen wurde notgedrungen mit weitgehenden Annahmen gearbeitet. An dieser Stelle sollen die Autoren selbst zitiert werden, die die Anforderungen formulieren, sollten die Ergebnisse der Methodik eine größere Belastbarkeit erreichen (MAYERHOFER und FRITZ, 2009):

---

<sup>5</sup> Die Ergebnisse des Simulationsmodells scheinen damit denen der von den Autoren ebenfalls durchgeführten Panel-Analyse im DiD-Design teilweise zu widersprechen, da im empirischen Schätzmodell die am stärksten geförderten Regionen auch am stärksten profitiert haben.

- „Auf Projektebene müsste zwischen Ersatz- und Erweiterungsinvestitionen unterschieden werden.
- Damit hängt die Identifikation des Investitionszwecks eng zusammen (Ausweitung der Produktionsmöglichkeit bestehender bzw. Ermöglichung neuer Produktlinien, Qualitätsverbesserung versus Rationalisierung, Erschließung neuer Märkte national und/oder international).
- Die Fördermittel wurden für Investitionen im Planungsstadium genehmigt. Für die Wirkungen ist wesentlich, ob sich die Planungsannahmen als realistisch erwiesen und die erhofften Absatzmöglichkeiten realisiert werden. Einer Analyse auf Makroebene wäre also eine Wirkungsanalyse auf Mikroebene voranzustellen.
- Einige Projekte wären ohne Förderung wahrscheinlich in anderen österreichischen Regionen verwirklicht worden. Abgesehen von den bei jedem Projekt zu erwartenden Mitnahmeeffekten werben Regionen mit Förderzusagen um Unternehmensstandorte und tragen damit zu einem regionalen Standortwettbewerb bei. Dies bewirkt eine innerösterreichische regionale Umverteilung von Investitionsmaßnahmen ohne unmittelbare Ausweitung des gesamten nationalen Kapitalstocks. Diese Problematik könnte nur im Rahmen einer Befragung der geförderten Unternehmen zufriedenstellend behandelt werden.“

Dies sind alles gravierende Schwierigkeiten, die nur unter Zuhilfenahme zusätzlicher Informationen von der Mikro-Ebene etwa aus Befragungen zu lösen sind. Sie stellen allerdings im Zusammenhang mit der Bewertung von Arbeitsplatzeffekten der ILE nicht das zentrale Problem dar. Dieses liegt darin begründet, dass die Methodik weiterhin keinerlei Aufschluss darüber gibt, ob Projekte ohne die Förderung auch initialisiert worden wären oder nicht. Es gilt darüber hinaus auch weiterhin, dass über die Vorzüglichkeit möglicher alternativer Projekte, die aufgrund der Förderung nicht realisiert wurden, und damit über die Kosten, die aus entgangenem Nutzen entstehen, nichts gesagt werden kann: „Diese privaten Mittel wären ja, auch wenn sie nicht in den modellierten Projekten verwendet worden wären, für andere Zwecke zur Verfügung gestanden, sei es innerbetrieblich oder über erhöhte Gewinnausschüttungen in Form privater Konsummöglichkeiten. Auch auf diese Weise hätten sie (positive) wirtschaftliche Auswirkungen mit sich gebracht, wenn auch nicht über dieselben Wirkungsketten. Speziell die regionale Verteilung des Mitteleinsatzes könnte dabei durchaus unterschiedlich gewesen sein“ (MAYERHOFER et al., 2008).

Ein weiterer Nachteil der vorgestellten I-O-Analyse im Vergleich zur SAM-Analyse des letzten Kapitels besteht insbesondere vor dem Hintergrund des Ziels der Berechnung von Arbeitsplatzeffekten darin, dass die Verteilung der Wirkungen zwischen Haushalten und Faktoren nicht erfasst wird. Die Bedeutung für den Faktor Arbeit wird also nicht klar. Dazu wäre es notwendig, zusätzliche Informationen zur Produktivitätsentwicklung in das Modell einzubringen. Das würde zwar das Problem der statischen Multiplikatoren aus dem SAM-Modell im vorausgegangenen Kapitel lösen, stellt aber seinerseits eine große Herausforderung dar.

Zusammenfassend muss daher festgestellt werden, dass das Zusammenspiel der Schätzung ökonomischer Verhaltensgleichung und modellhafter Simulation, das dem präsentierten Konzept im Kern zugrunde liegt, zwar theoretisch ein hohes Potenzial in sich birgt, wenn es um die Berechnung von Beschäftigungseffekten geht, dass die damit verbundenen Schwierigkeiten und Datenansprüche aber so hoch sind, dass bisher der Anspruch der zuverlässigen Berechnung makroökonomischer Effekte noch nicht befriedigend eingelöst werden konnte.

#### 4.1.2.4 Tabellarische Bewertung der Ö-I-O-Modelle

	1	2	3	4	5	6	7
<b>Technische Aspekte</b>	sehr gering			sehr hoch			
Zeitbedarf							■
Anforderungen an technisches Know-How							■
Anforderungen an Kontextwissen/Theoretisches Wissen							■
Softwarebedarf				■			
Datenanforderungen							■
Sonstige Kosten					■		
<b>Inhaltliche Aspekte</b>	sehr gut möglich			nicht möglich			
Erfassung direkter Effekte							■
Erfassung indirekter Effekte			■				
Direkte Modellierung der Interventionslogik				■			
Erfassung diskontinuierlicher, auch zeitverzögerter Effekte					■		
Erfassung positiver externer Effekte				■			
Erfassung von Allokationseffekten						■	
Verallgemeinerbarkeit					■		
Beachtung des regionalen Kontextes/weiterer Wirkungsbedingungen				■			
	Pro			Contra			

## 4.2 Empirische Analysen

Die im vorausgegangenen Kapitel vorgestellten Ansätze beruhen auf der Anwendung von sehr umfassenden allgemeinen Modellen, die schon aufgrund des großen Arbeitsaufwandes, der damit verbunden wäre, nicht bei jeder Anwendung wieder neu spezifiziert werden können. Es wird also angenommen, dass die zentralen Variablen modellendogen determiniert werden, während die exogenen Parameter weitgehend als konstante Daten behandelt werden. Davon abgegrenzt werden in diesem Abschnitt die empirischen theoriebasierten Methoden, die sich dadurch auszeichnen, dass zum einen alle Rahmendaten und Parameter bei jeder Analyse spezifisch ermittelt werden.

Die theoriebasierte empirische Analyse basiert auf wissenschaftlich untermauerten, als allgemeingültig anerkannten Zusammenhängen, nicht aber auf einem allgemeinen Theo-

riegebäude. Hier werden also die notwendigen Annahmen der empirischen Analyse, zum Beispiel Funktionsformen, nach dem theoretischen Vorwissen ausgewählt. Existiert ein solches Vorwissen, so kann es die Auswahl bestimmter Funktionsformen und kausaler Modelle leiten, ohne dass auch diese funktionalen Zusammenhänge selbst für jeden Einzelfall selbst wieder empirisch überprüft werden müssten. Das Problem besteht darin, dass viele spezifische Phänomene, die politischen Interventionen zugrunde liegen, zwar in der Praxis bekannt, aber wissenschaftlich noch nicht aufbereitet und theoretisch geklärt sind. In diesem Fall bleibt nur die Möglichkeit einer interventionslogikbasierten Analyse (Kapitel 4). Im Folgenden soll die Idee des Konzeptes an einem exemplarischen Fall dargestellt werden, um darzulegen in welcher Hinsicht weiterführende (Grundlagen-)Studien später einmal die Basis einer Bewertung von Arbeitsplatzeffekten der ILE bilden könnten.

Illustrativ wird dazu hier ein Beispiel aus dem ingenieurwissenschaftlichen Bereich herangezogen. Man stelle sich vor, eine Brücke sei vom Einsturz bedroht gewesen. Die Intervention der zuständigen Kommune bestand darin, die Brücke in ihrem Scheitel zu verstärken. Eine Möglichkeit der Analyse bestünde nun darin, eine andere, vergleichbare aber nicht gesicherte Brücke zu suchen und zu beobachten, wie beide unter den Bedingungen unterschiedlicher Belastungssituationen reagieren. Zeigen sich in der nicht gestützten Brücke Risse, in der gestützten hingegen nicht, oder stürzt die nicht gestützte im Gegensatz zur gesicherten in einem bestimmten Zeitraum gar zusammen, so wird die „Intervention“ der Verstärkung des Scheitels positiv beurteilt, ansonsten nicht. Abgesehen davon, dass dieses Beispiel einige der Probleme des kontrafaktischen Paradigmas gut verdeutlicht<sup>6</sup>, wird deutlich, dass uns ein solches Vorgehen in einem Umfeld, das (ingenieur-)wissenschaftlich bestens erforscht ist, nicht in den Sinn kommt. Vielmehr würde (und zwar streng genommen natürlich ex ante) ein Ingenieurbüro mit der Kalkulation der Wirkung der Brückenverstärkung beauftragt.<sup>7</sup>

---

<sup>6</sup> Wie lange muss gemessen werden? Wie hoch war die Wahrscheinlichkeit, dass eine der Brücken einstürzt, selbst wenn dieses nicht beobachtet wurde? Wie viele Vergleichsbeobachtungen müssen also gemacht werden, um ein belastbares Ergebnis zu erhalten? Wie werden die vergleichbaren Umstände erzeugt? Werden in einer Versuchsanordnung auf beiden Brücken die gleichen Mengen an LKW, Autos und Bussen in entsprechender zeitlicher Taktung zugelassen?

<sup>7</sup> Sind die Wirkungszusammenhänge den Experten so geläufig wie in diesem Beispiel, so würde der Bürgermeister nicht die entsprechende Intervention selbst kreieren, sondern seine Entscheidung ganz und gar auf der Empfehlung der Ingenieure basieren. Auch die Notwendigkeit einer Ex-post-Evaluierung würde in den meisten Fällen hier nicht gesehen werden. Der zentrale Unterschied liegt gerade in der inhärenten Unsicherheit sozioökonomischer Entwicklungen, die die Politik dazu zwingen, in einen Prozess des reziproken Lernens in Interaktion mit Wissenschaft und Politikanalyse einzutreten. (LAMBOUY und BOSCHMA, 2001). Die Wissenschaft liefert eben nicht die fertigen Rezepte zur Intervention und das auch nicht nach einem einmaligen Politikversuch. Vielmehr handelt es sich um ein iteratives Herantasten und schrittweises Lernen aus Erfahrung. Aus eben dem gleichen Grunde ist es vermutlich prinzipiell nicht möglich, eindeutige Politikeffekte zu ermitteln.

Das Ingenieurbüro würde, und das entspricht dem idealtypischen methodischen Vorgehen der theoriebasierten Wirkungsanalyse, zunächst die Rahmendaten der Brücke erheben. Dazu gehören zum Beispiel die maximale Schneelast, die maximale zu erwartende Windstärke, Festigkeit des Untergrundes etc. Die Ingenieure würden außerdem den Zustand der Brücke anhand der relevanten Parameter genau definieren: Bauart, Material, Stärke der Träger, Materialmüdigkeit etc. Entsprechend ist auch die theoriebasierte Bewertung ökonomischer Interventionen auf die Bereitstellung umfangreicher (sozioökonomischer) Rahmendaten zum Fördertatbestand und seinem Umfeld angewiesen, sie kann aber auf umfangreiches Datenmaterial zu nicht geförderten Vergleichsfällen verzichten.

In unserem Beispiel war die sogenannte Scheitelstärke  $D$  der Brücke durch Abnutzung gefährlich klein geworden. Das Ingenieurbüro erfragt die gewünschte Tragfähigkeit der Brücke als den Zielwert der Intervention. Es berechnet dann mit Hilfe der empirischen Formel (entspricht unserem „Stylised Fact“)  $D_S = 0.37 + 0.028 * L$  (BUSCH, 2005) die mindestens notwendige Dicke der Scheitelstärke  $D_S$  in Abhängigkeit von ihrer Spannweite  $L$ . Die erhobenen Rahmendaten gehen über empirisch oder physikalisch-theoretisch bestimmte Korrekturfaktoren nachträglich in die Berechnung ein, sodass die spezifische notwendige Scheitelstärke unter den gegebenen Bedingungen und vor dem Ziel der erwünschten Tragfähigkeit berechnet werden kann. Ist die Intervention, also die Brückenausbesserung, bereits eigenmächtig durch den Bürgermeister erfolgt, so kann das Ingenieurbüro im Umkehrschluss berechnen, ob die gewünschte Tragfähigkeit durch die Intervention gewährleistet wird. Das Ingenieurbüro kann auch kalkulieren, wie viel diese Verbesserung im Vergleich zu der von ihr berechneten Optimallösung gekostet hat und somit die Effizienz der „Intervention“ des Bürgermeisters angeben.

Ein Beispiel für eine solche Evaluation ist die Studie von BERGSCHMIDT et al. (2009, S. 43-57), zur Verringerung des Bodenabtrags durch Mulchsaat. Der erwartete Bodenabtrag mit und ohne Mulchsaat im Vergleich zur Situation bei Schwarzbrache in den verschiedenen Regionen wurde je nach dominierenden Fruchtarten nach einer empirisch abgeleiteten Formel aus der Literatur berechnet. Die relative und absolute Verringerung der Erosion durch die Mulchsaat konnte berechnet und den Kosten des Einsatzes gegenüber gestellt werden.

### 4.2.1 Die Partialanalyse

Leider liegen aus dem sozioökonomischen Bereich keine vergleichbaren Studien vor. Das liegt weniger an den methodischen Vorlieben der Analysten als vielmehr an den fehlenden konstanten Zusammenhängen, die eine Isolierung stabiler Gesetzmäßigkeiten im

funktionalen Sinne in den sozioökonomischen Wissenschaften erlauben würden.<sup>8</sup> So bleibt die mögliche Zukunft theoriebasierter empirischer Evaluationen sozioökonomisch begründeter Interventionen davon abhängig, in wie fern die ökonomische Grundlagenforschung belastbare empirische Gesetzmäßigkeiten und ihre begrenzenden Rahmenbedingungen ermitteln wird. Am ehesten existieren solche Theorien wieder im Bereich der Investitionsförderung. Die Zusammenhänge zwischen regionaler Faktorausstattung, Produktionstechnik und der Wirkung von Kapitalzuschüssen zu Investitionen sowie deren Rationalisierungs- und Outputeffekte und der abzuleitenden Wirkungen auf den Arbeitsmarkt sind theoretisch gut bekannt. Im vorausgegangenen Kapitel wurden sie in sehr aufwendigen und wenig spezifischen allgemeinen Regionalmodellen verwendet. Im folgenden Abschnitt wird der umgekehrte Weg eines kleinen Modells mit hoher Spezifität und daher größerer empirischer Ausrichtung verfolgt.

#### 4.2.1.1 Methodische Erläuterungen zur Partialanalyse

Das partialanalytische Modell besteht aus einem Gleichungssystem, das die Zusammenhänge mit Blick auf einen ausgewählten Markt beschreibt. Zusammenhänge zu anderen Märkten werden nicht integriert. Modelliert wird das Verhalten der Marktteilnehmer unter der Annahme konstanter Rahmenbedingungen.

Im partialanalytischen Modell werden die Zusammenhänge explizit modelliert, die als zentral für das zu untersuchende Problem erkannt wurden. In sofern ist die inhaltliche Auseinandersetzung mit der Intervention und den erwarteten Wirkungsmechanismen von noch größerer Bedeutung als beim allgemeinen Simulationsmodell. Dabei ermöglicht das partialanalytische Modell leichter eine Modellierung spezifischer Bedingungen als die großen allgemeinen Modelle. Ein Beispiel ist die Beachtung von Anpassungskosten in Investitionsentscheidungen, die sich aus der Kapitalstruktur, fehlender Flexibilität oder Teilbarkeit des Faktors Arbeit oder aber der Grenzen der Infrastruktur ergeben (OTT und SORETZ, 2007). OTT und SORETZ unterstellen zum Beispiel, dass die Vorzüglichkeit von Investitionen in Abhängigkeit von der Infrastrukturausstattung abhängt von der Relation eines (positiven) Produktivitätseffektes zum (positiven) Anpassungskosteneffekt zum (negativen) Rivalitätseffekt. Entsprechend werden Produktions-, Überfüllungs- und Anpassungskostenfunktionen spezifiziert, die es ermöglichen „die verschiedenen Einflusskanäle des öffentlichen Inputs innerhalb des Modells und ihre Auswirkungen auf die Mo-

---

<sup>8</sup> Interessant ist dabei am Rande, dass auch der Brückenbau ursprünglich sich nur an empirisch erkannten Gesetzmäßigkeiten orientierte, aus denen dann im ingenieurwissenschaftlichen Fortschritt nach und nach deterministische Regelmäßigkeiten abgeleitet wurden (BUSCH, 2005). Keineswegs wurden alle Einflussgrößen des Brückenbaus in ein einzelnes rein physikalisch begründetes Modell integrieren, mit dem jede mögliche Form beurteilt werden kann.

delldynamik zu identifizieren“ (OTT und SORETZ, 2007). Kern des Modells ist in diesem Fall die einzelbetriebliche Investitionsentscheidungsfunktion. In derartigen Modellen existieren viele Unbekannte. Ob ein solches Modell für die Politikbewertung in Frage kommt hängt vor allem davon ab, ob es möglich ist, diese empirisch zu bestimmen,

#### 4.2.1.2 Beispielstudie zur Partialanalyse

Die Beispielstudie zur Bewertung „Regionale[r] Strukturpolitik in der Bundesrepublik Deutschland“ stellt eine sehr ausführliche frühe Arbeit in dieser Richtung dar (ASMACHER, 1989). Es handelt sich um eine partialanalytische mikroökonomische Analyse, die in späteren Jahren als „Münsteraner Modell“ weitergeführt wurde. Als ein Motiv seiner Herangehensweise schreibt ASMACHER: „Weiterhin sollten die Modellbeziehungen aus fundierten theoretischen Vorstellungen über die möglichen Ursache-Wirkungs-Zusammenhänge zwischen Instrumenteneinsatz und Zielgrößen abgeleitet werden. Eine bloße (eher zufällige) Aneinanderreihung von plausibel erscheinenden Einflussfaktoren ist zu vermeiden. Das betrifft sowohl die Auswahl der für möglich gehaltenen erklärenden Variablen als auch deren funktionale Verknüpfung“ (ibid.).

Der Autor der Dissertation berechnet die Parameter des theoretischen Modells zur Nachfrage auf Faktormärkten, das seiner Wirkungsanalyse zugrunde liegt, im Rahmen seiner Analyse. Am Umfang dieser Arbeit wird bereits deutlich, dass dies ein Aufwand ist, der den Rahmen jeder Analyse zur Messung von Arbeitsplatzeffekten in der angewandten Forschung sprengt. Das beispielhafte Vorgehen ASMACHERS verdeutlicht aber auch viele Notwendigkeiten, die ebenso für Evaluationen anderer Art, insbesondere vor dem Hintergrund der notwendigen Einordnung der Ergebnisse angesichts der jeweiligen Schwächen verschiedener Methoden, gelten. Zu diesen Notwendigkeiten zählt in jedem Fall ASMACHERS erster Schritt, der in der expliziten Analyse der intendierten Wirkungsweise besteht. Die beiden Zielgrößen der Intervention sind das Einkommen und die Zahl der Arbeitsplätze. Beides soll durch eine Subventionierung des Kapitaleinsatzes in Investitionsförderungen gesteigert werden.

ASMACHER stellt klar, dass theoretisch nur ein indirekter und keineswegs eindeutiger Zusammenhang zwischen der Förderung des Kapitaleinsatzes in den wirtschaftlich schwachen Regionen und deren wirtschaftlicher Entwicklung besteht. Es ist nicht klar, ob die Förderung überhaupt neue Investitionen anstößt, welcher Zusammenhang zwischen dem Einsatz der Faktoren Arbeit und Kapital besteht und ob es für die mehr produzierten Güter ausreichend Nachfrage gibt. Theoretisch muss zwischen dem Substitutions- und dem Outputeffekt der Förderung unterschieden werden. Weil also zwischen der Nachfrage sowie dem Einkommen und dem Einsatz von Kapital sowie dem Einsatz von Arbeit ein Zusammenhang besteht, muss das Erklärungsmodell alle Faktoren simultan berücksichtigen. Das empirische Schätzmodell besteht dann aus einem simultanen Mehrgleichungssystem.

Weitere zu berücksichtigende relevante Faktoren, die die Zielgrößen beeinflussen, sind unter anderem: „die gesamtwirtschaftliche Kapazitätsauslastung, das Zinsniveau, die Arbeitskosten, Produkt- und Investitionsgüterpreise, Gewinnsteuern usw.“ (ASMACHER, 1989).

Dem theoretischen Modell zur Bestimmung der Faktornachfrage unter gegebenen Rahmenbedingungen beruht auf der Kostenminimierung der Unternehmen und einer unterstellten Produktionsfunktion vom Cobb-Douglas-Typ. Um das rein theoretisch begründete Modell der Realität anzupassen, müssen die Produktionselastizitäten der Faktoren empirisch bestimmt werden. Dazu müssen die Gleichungen, die das System aus Produktion, Arbeits- und Kapitaleinsatz und Nachfrage beschreiben, in eine schätzbare also lineare Form gebracht werden. Die Intervention wird in dem theoretischen Faktornachfragemodell berücksichtigt, indem die Kapitalkosten entsprechend der geltenden Fördersätze reduziert werden. Die Fördersumme ergibt sich damit als tatsächliche Inanspruchnahme der potentiell zur Verfügung stehenden Förderung endogen aus dem Modell. Wird nicht die gesamte Förderung in Anspruch genommen, so kann das bereits ein erster Faktor sein, der die angestrebte Wirkung reduziert<sup>9</sup>. Die regionalen Kapitalnutzungskosten werden außerdem durch die ökonomische Abschreibungsrate, den Zinssatz, den Steuersatz und den Investitionsgüterpreis (ibid.) beeinflusst, die daher als Rahmendaten in das Modell einfließen müssen.

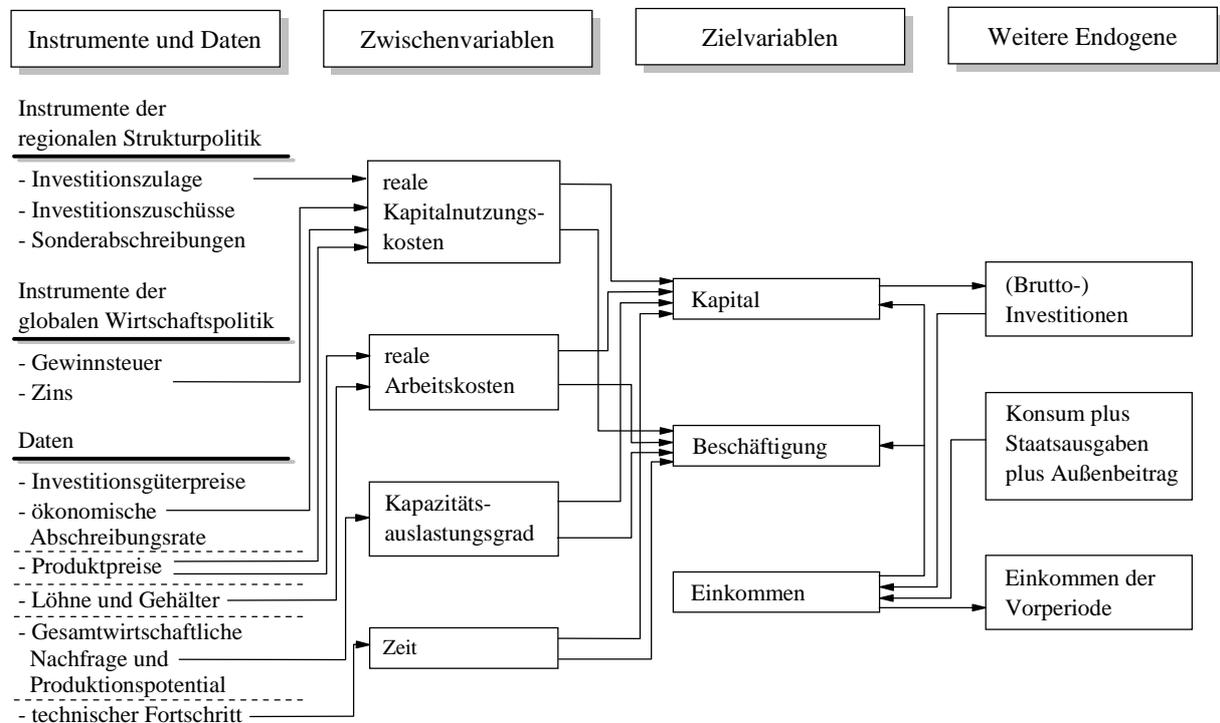
Dieses Faktornachfragesystem beruht auf der Annahme eines gegebenen Outputs. Es wird daher, um den Produktionseffekt zu erfassen, um zusätzliche Gleichungen erweitert, die die gesamten regionalen Wirkungszusammenhänge inklusive der Output- und Multiplikatoreffekte innerhalb der betrachteten Branche abbilden können. Zu diesem Zweck wird der Zusammenhang zwischen Kapitalstock und Bruttoinvestitionen funktional bestimmt. Das Modell wird ökonometrisch in seinen Parameterwerten bestimmt. Es ergibt sich ein spezifisches Modell der regionalen Faktornachfrage für die entsprechende Region. Schließlich wird eine Einkommensgleichung von der Verwendungsseite her definiert und eine Verhaltensfunktion, die die Bedeutung der Konsumausgaben am generierten Einkommen bestimmt. Insbesondere an dieser Stelle musste aufgrund fehlender regionalisierter Daten mit Ad-hoc-Annahmen gearbeitet werden, die allerdings empirisch überprüft wurden.

Abbildung 4.1 beschreibt schematisch die funktionalen Zusammenhänge zwischen Förderung, regionalen Bedingungen und Zielvariablen. Die Abbildung verdeutlicht die Notwendigkeit einer expliziten Interventionslogik für eine theoriebasierte Bewertung von Effekten.

---

<sup>9</sup> Es hängt von der genauen Fragestellung ab, ob man die Effektivität der Maßnahme am ausgezahlten Geld oder an den bereitgestellten Mitteln misst.

**Abbildung 4.1:** Variablen und Beziehungen in ASMACHERS Modell der Wirkungsanalyse regionalpolitischer Instrumente



Quelle: Asmacher (1989).

Nicht abbildbar ist die ebenfalls durch das Modell vorgegebene funktionale Form der Zusammenhänge, die nicht-linear ist. Anhand dieses Modells, in das die spezifischen regionalen Rahmendaten einfließen, werden die verschiedenen Szenarien mit und ohne Förderung simuliert: „Die Differenz zwischen den Ergebnissen stellt den isolierten Effekt der regionalpolitischen Maßnahme dar. Wir erhalten die zusätzliche Beschäftigung, den zusätzlichen Kapitalstock (Investitionen) und das zusätzliche Einkommen durch Regionalförderung“ (ASMACHER, 2005). Eine Schwachstelle des Vorgehens ASMACHERS, die der Autor selbst anspricht, besteht darin, dass nur Wirkungen auf die geförderte Industriebranche ermittelt werden. Die Quantifizierung der indirekten Effekte auf andere Branchen war nach den Angaben des Autors aufgrund fehlender Daten und fehlender Regionalmultiplikatoren nicht vorgenommen worden. Im Ergebnis zeigt sich in der kurzfristigen Wirkung ein Rückgang der Beschäftigung aufgrund des Substitutionseffektes des erhöhten Kapitaleinsatzes: Arbeit wird durch Kapital verdrängt. Schon nach kurzer Zeit überwiegt aber der Outputeffekt der Regionalförderung: Die Produktion wird gesteigert und ein positiver Beschäftigungseffekt stellt sich ein.

### 4.2.1.3 Kritische Reflexion der Eignung der Partialanalyse

Die beschriebene Analyse von ASMACHER (1989) war sehr aufwendig, wie bereits in der Einleitung herausgestellt. Das Modell wurde in den Folgejahren beständig weiter entwickelt und auch immer wieder auf die Bewertung der Regionalförderung angewandt. Dabei wird nicht ein und dasselbe Modell immer wieder verwendet, vielmehr bleibt nur die grundlegende Struktur des Mehrgleichungssystems mit Produktionsfunktion und Kostenminimierung und zusätzlichen Outputfunktionen erhalten. Ein neueres Beispiel der Anwendung des Konzepts findet sich in SCHALK und UNTIEDT (2000). Insbesondere wird deutlich, wie wissenschaftlicher und technischer Fortschritt hier zu einer Verbesserung in der Spezifizierung führen, indem zum Beispiel regionale Effizienzunterschiede zugelassen und neue Produktionsfunktionsformen erprobt werden.

Die Methodik wird von Niebuhr und LAMMERS (2002) diskutiert. Ihre Bewertung wird hier ausführlich wiedergegeben:

- „Ein zentrales Resultat der Wirkungsanalyse – der positive Beschäftigungseffekt der Investitionsförderung – ist abhängig von der gewählten Schätzmethode. Wird ein anderes Regressionsverfahren angewendet, stellt sich eine erhebliche Reduzierung der expansiven Beschäftigungswirkungen ein. Kommt statt des einstufigen ein zweistufiges Schätzverfahren zur Anwendung, reduziert sich die langfristige Elastizität der Kapitalnutzungskosten auf die Beschäftigung von -0,16 auf -0,03, d. h., bei einem alternativen methodischen Ansatz führt eine einprozentige Senkung der Kapitalnutzungskosten lediglich zu einem langfristigen Beschäftigungsanstieg von 0,03 % (WALZ, 1995, S. 303-306; DEITMER, 1993, S. 76 ff.).
- Der Ansatz beinhaltet keine modellendogene Untersuchung der Wirkungen, die von der Finanzierungsseite ausgehen (LAMMERS, 1990, S. 80 f.). Zwar wird basierend auf dem Konzept der Kapitalnutzungskosten die Frage diskutiert, inwieweit die positiven Effekte der Förderung durch gegenläufige Wirkungen der Finanzierung der Fördermaßnahmen konterkariert werden könnten. Das Ergebnis dieser Betrachtungen ist, dass weder eine allgemeine Steuererhöhung noch eine staatliche Kreditaufnahme die Kapitalnutzungskosten der privaten Investoren in einer Weise steigen lässt, dass sie ihre Investitionstätigkeit reduzieren (FRANZ und SCHALK, 1995, S. 286 ff.). Da dieses Ergebnis aber nicht innerhalb eines geschlossenen Modellansatzes abgeleitet und empirisch ermittelt wird, ist es wenig aussagekräftig. So bleiben z. B. die Effekte unberücksichtigt, die eine Steuererhöhung oder eine verstärkte staatliche Kreditaufnahme auf die gesamtwirtschaftliche und regionale Nachfrage haben.
- Die Analysen lassen die Arbeitsangebotsseite unberücksichtigt. Der Ansatz impliziert ein vollkommen elastisches Arbeitsangebot. Daher bleiben auch indirekte Faktoreffekte unberücksichtigt, die möglicherweise dem positiven direkten Outputeffekt entgegenwirken (WALZ, 1995, S. 303 f.; vgl. auch LAMMERS, 1989, S. 79).

- Analysiert werden im Rahmen der beschriebenen Untersuchungen lediglich die Effekte der direkten Förderung von Investitionen der gewerblichen Wirtschaft. Darüber hinaus unterstützt die Gemeinschaftsaufgabe aber auch Investitionen in die wirtschaftsnahe kommunale Infrastruktur und seit 1995 nicht-investive Maßnahmen. Zu den förderfähigen nicht-investiven Maßnahmen zählen Beratungsleistungen, betriebliche Schulungsmaßnahmen sowie in kleinen und mittleren Betrieben die Humankapitalbildung und anwendungsorientierte Forschung und Entwicklung (DEUTSCHER BUNDESTAG, 2000, S. 14 f.; KARL und KRÄMER-EIS, 1969, S. 19). [...] Auf eine erforderliche Ausweitung der Erfolgskontrolle auf die bisher vernachlässigten Komponenten der Regionalförderung weisen die Autoren selbst hin, indem sie die Bedeutung der Infrastruktur und des Humankapitals für den Abbau regionaler Einkommensdisparitäten besonders hervorheben (FRANZ/SCHALK, 1995, S. 294 ff.; WALZ, 1995, S. 304).“

Darüber hinaus wird von den Kritikern darauf hingewiesen, dass sich ein einmal für eine Region spezifiziertes Modell eben nicht auf andere Regionen übertragen und dort anwenden lässt. Das gleiche wird für die Übertragung auf andere Zeiträume zu gelten haben. Insgesamt können die angesprochenen Aspekte durch eine Erweiterung des theoretischen Rahmens mit genaueren theoretischen Spezifizierungen beachtet werden. Der Informationsbedarf steigt aber mit jeder dieser Anpassungen. NIEBUHR und LAMMERS (2002) kommen zu einer klaren Schlussfolgerung: „Ergebnisse auf der Basis des Münsteraner Modells zu den Wirkungen der Gemeinschaftsaufgabe stehen insgesamt gesehen auf unsicherem Grund. Sie sollten deshalb mit großer Vorsicht und in Kenntnis der genannten Kritikpunkte interpretiert und verwendet werden. Allerdings ist zu konstatieren, dass es bislang keine anderen, überzeugenderen empirischen Analysen gibt. Verlässlichere Aussagen zu den Effekten wären zu erzielen, wenn es gelänge, eine empirische Totalanalyse mit Hilfe eines allgemeinen Gleichgewichtsmodells durchzuführen. Insbesondere Datenrestriktionen auf regionaler Ebene stehen dem bislang entgegen“ (NIEBUHR und LAMMERS, 2002).

Ein heterogener Treatmenteffekt, je nachdem, wie im Einzelnen gefördert wird, kann ebenfalls nicht ermittelt und die besondere Wirkung der Förderung nur bestimmter Unternehmen daher auch nicht berechnet werden. Damit ist auch nicht klar, in wie fern das Ergebnis robust gegenüber Änderungen in der Unternehmensstruktur ist. Insgesamt hängt der Fördereffekt nach dieser Theorie ausschließlich vom Verhältnis des Substitutions- zum Outputeffekt ab. Wirkungen, die sich aufgrund der Theorien der Neuen ökonomischen Geographie erwarten ließen, in der mit positiven Skaleneffekten und positiven Agglomerationswirkungen operiert wird, werden ebenso wie negative Allokationseffekte nicht erfasst.

Während die Autoren der Studien selbst aus ihren Ergebnissen schlussfolgern, dass die Förderung mehr in Richtung der Effizienzsteigerung durch Humankapitalbildung, Innovationsförderung etc. ausgerichtet werden müsste, um etwa Konvergenzziele mit Blick auf die regionale Produktivität zu erreichen, wurden entsprechende Förderinstrumente mithilfe des

partialanalytischen Modells nicht analysiert. Der Grund dafür wird darin zu suchen sein, dass entsprechende Politiken sich in diesen Modellen nicht ohne weiteres darstellen lassen.

Mit Blick auf die ILE muss wieder konstatiert werden, dass eine solche Partialanalyse nicht bestimmen kann, welches endogene Potential durch die Förderung in der Region geweckt wird. Allerdings kann ein solches oder ähnliches Modell dazu dienen, die Arbeitsplatzeffekte zu kalkulieren, wenn der entsprechende „Shiftfaktor“, der etwa durch die Förderung von Netzwerken entsteht, vorab empirisch ermittelt wurde. Darum geht es im folgenden Kapitel.

### 4.2.1.4 Tabellarische Bewertung der Partialanalyse

	1	2	3	4	5	6	7
<b>Technische Aspekte</b>	sehr gering			sehr hoch			
Zeitbedarf						■	
Anforderungen an technisches Know-How						■	
Anforderungen an Kontextwissen/Theoretisches Wissen						■	
Softwarebedarf				■			
Datenanforderungen						■	
Sonstige Kosten				■			
<b>Inhaltliche Aspekte</b>	sehr gut möglich			nicht möglich			
Erfassung direkter Effekte		■					
Erfassung indirekter Effekte		■					
Direkte Modellierung der Interventionslogik		■					
Erfassung diskontinuierlicher, auch zeitverzögerter Effekte				■			
Erfassung positiver externer Effekte							■
Erfassung von Allokationseffekten					■		
Verallgemeinerbarkeit					■		
Beachtung des regionalen Kontextes/weiterer Wirkungsbedingungen		■					
	Pro			Contra			



## 5 Die interventionslogikbasierte Evaluation

Wenn in der Evaluationsliteratur von theoriebasierter Evaluation gesprochen wird, geht es aufgrund des Fehlens allgemeiner wissenschaftlicher Modelle zu den relevanten Zusammenhängen in der Regel um die spezifische, programmbezogene empirische theoriegeleitete Analyse: „It is an approach to evaluation that requires surfacing the assumptions on which the program is based in considerable detail: what activities are being conducted, what effect each particular activity will have, what the program does next, what the expected response is, what happens next, and so on, to the expected outcomes[...]. The evaluation then follows each step in the sequence to see whether the expected ministepts actually materialize” (BIRCKMAYER und WEISS, 2000).

Das Theoriedefizit ist auch im Bereich der Förderung des ländlichen Raumes evident. KOSTOV und LINGARD (2002) schreiben zum neuen Paradigma der ländlichen Entwicklung: „This shift represents a fundamental change in policy objectives and frameworks towards a more holistic approach to rurality and implies new tools of analysis. Conventional economic models are based on an instrumentalist methodology which links means to ends with little interest in the underlying structure. We argue here for a synergy approach to rural development. This approach incorporates both traditional network and institutional analysis and focuses on working mechanisms and processes rather than ends.” Die Evaluation müsste demnach zunächst analysieren, inwiefern die Unterstützung des Netzwerkmechanismus durch die Förderung gelingt und inwiefern dadurch eine Steigerung der regionalen Produktivität erreicht wird um erst in einem letzten Schritt abzuleiten, zu welchen Arbeitsplatzeffekten eine solche Entwicklung möglicherweise beiträgt. Indem eine solche Evaluation sich auf den Teil des Ursache-Wirkungs-Geflechts konzentriert, der als zentral für die Effektivität der Förderung betrachtet wird, handelt es sich auch hier um eine theoriegeleitete Evaluation, denn der zentrale Punkt muss identifiziert und in der Interpretation in seiner Bedeutung eingeordnet werden. Das geht nicht ohne explizite Darstellung der unterstellten theoretischen Zusammenhänge.

Abweichend von der üblichen Konvention wird hier zur Abgrenzung von den Ansätzen der vorausgegangenen Kapitel nicht von „theoriebasierter“, sondern von „interventionslogikbasierter“ Evaluation gesprochen.

### 5.1 Methodische Erläuterungen zur interventionslogikbasierten Evaluation

Wie bereits angesprochen stehen in den neuen Modellen des endogenen Wachstums der Regionalökonomie Wissens-Spillover und deren Bedeutung im Zentrum der Erklärung divergenter Entwicklungen. Die meisten Studien der neuen ökonomischen Wachstums-

theorie aber betrachten diese als „Black box“ und modellieren den Prozess der Wissensverbreitung nicht explizit (HENNING und SAGGAU, 2009). Da die Frage der Evaluation gerade ist, ob ein Effekt der Förderung auf die Wissensgenerierung in Netzwerken besteht, muss Licht in die Black box gebracht werden. Die Interaktion zwischen den Akteuren und der resultierende Wissensaustausch kann technisch am besten mit dem Netzwerk-konzept abgebildet werden (ibid.). Leider schreiben HENNING und SAGGAU (2009) auch: „However, economic research on networks is still in its infancy and studies on the impact of network structures on innovation diffusion and technological progress do not yet exist“.

Dementsprechend muss noch einmal betont werden, dass es sich bei dem neuen Politikan-satz der Integrierten Ländlichen Entwicklung und vergleichbaren anderen Regionalför-derprogrammen um experimentelle Politiken handelt. Die Wissenschaft kann hier nicht einseitig die Politik informieren, vielmehr ist der Wissensaustausch zwischen Politik und Wissenschaft wechselseitig und kumulativ. Dementsprechend wäre es falsch zu erwarten, dass Wirkungen unzweifelhaft identifiziert und quantifiziert werden können. Die Inter-ventionslogik dient der theoriebasierten Evaluation in solchen Fällen auch als spezifische Theorie, die es zu überprüfen gilt (vgl. MARGARIAN, 2008). Wieder werden in der Regel vor allem zentrale Elemente der Interventionslogik getestet und weniger die Zusammen-hänge, über die wenig Zweifel bestehen. Dieses Fehlen einer theoretischen und quantifi-zierten Verbindung ist typisch für Interventionen, die eher die „weichen“ Standortfak-toren einer Region ansprechen. Hier existieren zwar soziologische und ökonomische Theo-rien zu Netzwerkbildung und Humankapitalbildung, doch in wie fern diese durch be-stimmte Interventionen angestoßen werden können ist weitgehend unbekannt. Der Zu-sammenhang zwischen Intervention und Netzwerkbildung im spezifischen Fall eines In-terventionsszenarios sollte daher im Zentrum der Evaluation stehen.

Die Kunst einer solchen Interventionslogikbasierten Politikbewertung besteht weniger in der Anwendung einer bestimmten Methode. Aufgrund der fehlenden Sicherheit zu allge-meinen Zusammenhängen wird vielmehr an einzelnen Beispielen (Fallstudien) gearbeitet. Die Kunst einer solchen Studie besteht vor allem darin, die oft nur implizit dargelegte Interventionslogik explizit auszuarbeiten und die zentralen erwarteten Wirkungsmecha-nismen zu isolieren. Darüber hinaus ist die Beschaffung der oft sehr spezifischen Daten von zentraler Bedeutung für die Belastbarkeit der späteren Ergebnisse. Ein Großteil der Arbeit einer solchen Untersuchung besteht daher oft darin herauszufinden, wer die Wis-sensträger sind und welche Interessen sie möglicherweise verfolgen.

Statt einer einzelnen Beispielstudie werden im folgenden Absatz kurz etwas unterschied-liche vorwiegend qualitative Studien vorgestellt, die aber gemeinsam haben, dass sie sich an einer klaren, theoretisch begründeten Struktur orientieren.

## 5.2 Beispielstudien zur interventionslogikbasierten Evaluation

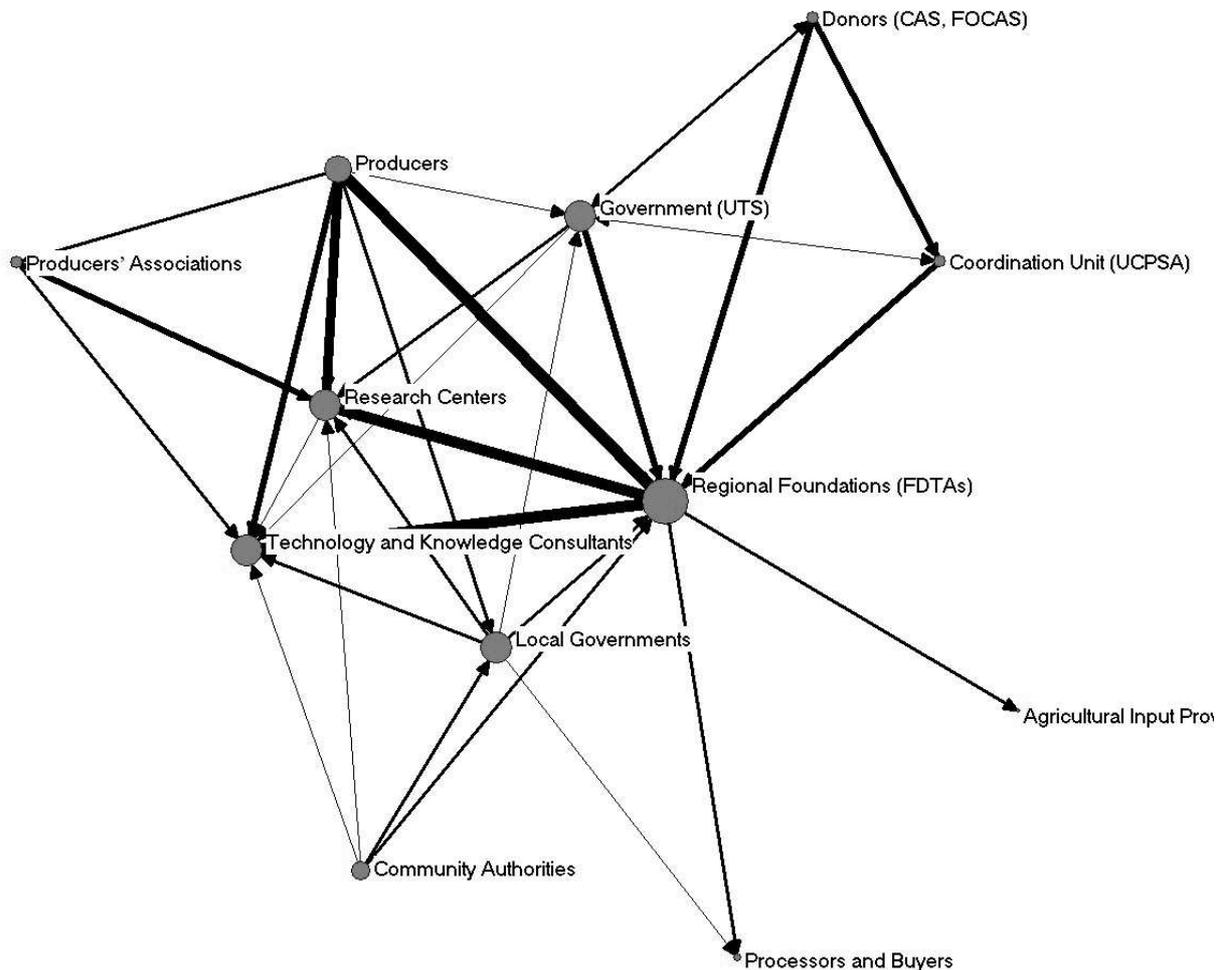
Alle Studien, die im Folgenden vorgestellt werden, sind in Datenerhebung und Datenaufbereitung sehr aufwendig und komplex. Aus diesem Grund werden sie in aller Regel nur als Fallstudien ohne den Anspruch von Repräsentativität durchgeführt. Schon deshalb kann von diesen und allen vergleichbaren Ansätzen nicht erwartet werden, dass sich mit ihnen die Nettoeffekte einer bestimmten Maßnahme berechnen ließen. Das war aber sowieso nicht mehr die Zielrichtung der hier zu diskutierenden Vorgehensweisen. Vielmehr geht es ja darum, die prinzipielle Art der Wirkung der ja auch in sich sehr heterogenen Projekte zu verstehen.

### Netzwerkanalysen

Netzwerkanalysen, die den formalen Anforderungen dieses neuen Wissenschaftsparadigmas entsprechen und zu Ergebnissen führen, die die existierenden formalisierten Modelle an die spezifische Untersuchungssituation anpassen können, sind nach Einschätzung der Experten dieses Bereiches sehr aufwendig. Als Beispiel einer pragmatischen Anwendung der Netzwerkkategorie im weiteren Sinne sei hier eine Studie zu „Innovation Systems Governance in Bolivia“ (HARTWICH et al., 2007) angeführt. In dieser Studie geht es um die Analyse der „Governance“ eines Innovationssystems, das der Verbreitung landwirtschaftlicher Technologien und anderer Innovationen dienen soll. Es werden Experten und zentrale Akteure befragt, um mithilfe dieser Informationen Matrizen zu konstruieren, die die Art und die Intensität der Interaktion zwischen den Agenten beschreiben. In der Art der Interaktion wurde zwischen Finanzierung und Prioritätensetzung unterschieden. Graphische Darstellungen der Parameter verdeutlichen dann, wer wen in welchem Ausmaß finanziert oder wer wessen Prioritäten folgt (Abbildung 5.1).

Die Ergebnisse sagen etwas darüber aus, wie wenig es gelungen ist, in diesem Innovationssystem die Initiative an dezentrale Stellen zu übertragen. Diese quantitativen Informationen werden in der Studie ergänzt durch qualitative Informationen aus den Befragungen, die die verschiedenen Dimensionen der „Governance“ betreffen: Partizipation, Transparenz und Offenheit, Verantwortlichkeit und Zurechenbarkeit, Konsensorientierung und Kohärenz sowie Strategische Vision. Das Beispiel ist alles in allem gut auf die ILE übertragbar, auch wenn es dort nicht in erster Linie um die Befähigung zur Innovation im engeren Sinne geht. Aber auch in der ILE kann man davon ausgehen, dass die Mobilisierung brach liegender Potenziale vor allem dann gelingt, wenn Akteure zu aktiven Kräften mit Initiative werden, die eher peripher oder dezentral im Netzwerk standen. Was dort an Projekten angeschoben wird hätte mit relativ hoher Wahrscheinlichkeit ohne die ILE nicht stattgefunden.

**Abbildung 5.1:** Beziehungen zwischen den Akteuren des Bolivianischen Innovations-systems mit Blick auf die Prioritätensetzung



Quelle: Hartwich et al. (2007).

### Funktionsorientierte Prozessanalyse

Andere systemorientierte Techniken stellen ebenfalls Möglichkeiten der Analyse dar. Zu diesen zählt die Funktionsorientierte Prozessanalyse, wie sie in HEKKERT et al. (2007) vorgeschlagen wurde. Die Autoren kombinieren eine Klassifizierung der Systemelemente nach Funktionen mit einer Prozessanalyse. Letztere kann auch mit dem Begriff des „story telling“ umschrieben werden, denn es handelt sich um eine chronologische Aufzählung von Ereignissen. Für Innovationssysteme haben die Autoren einen eigenen Satz zentraler Funktionen von Innovationssystemen identifiziert, die es ermöglichen sollen, die Entwicklung der Funktionssysteme im Zeitablauf zu erheben und abzubilden (ibid.):

- Unternehmerische Aktivitäten [Indikatoren: Anzahl von Neugründungen, Ausmaß Diversifizierungsaktivitäten bestehender Unternehmen, Anzahl der Experimente mit neuen Technologien (evtl. auch Organisationsformen, etc.),

- Wissenserweiterung (“learning by doing” and “learning by searching”), [Indikatoren: Anzahl F&E-Projekte, Patente, Investitionen und F&E (mit Blick auf die ILE wohl eher Fortbildungen, Trainingsmaßnahmen, Beratungen etc. )],
- Wissenstransfer in Netzwerken („learning by interacting“), [Anzahl von Treffen, Workshops, Darstellung der Netzwerkgröße und -intensität],
- Lenkung der Suche (z. B. durch verlässliche langfristige Politikziele), [Indikatoren: Zielsetzungen der Regierung, Anzahl der Veröffentlichungen, die entsprechende Projekte zum Thema haben und Erwartungen hervorrufen, Anzahl positiver und negativer öffentlicher Kommentare],
- Marktbildung (z. B. Schutz bestimmter Nischenmärkte bis zur vollen Marktreife der Innovation), [Indikatoren: Anzahl der entstandenen Nischenmärkte, besondere Steuererleichterungen (Förderungen?) für bestimmte Technologien (neue Aktivitäten?), neue Umweltstandards],
- Mobilisierung von Ressourcen [Indikatoren werden nicht vorgeschlagen, statt dessen direkte Befragung],
- Legitimierung/Überwindung von Widerständen [Indikatoren: Entwicklung von Interessengruppen und ihren Lobbyaktivitäten].

Wie an einigen Stellen bereits angemerkt, können diese Funktionen zur Analyse von Innovationspolitik durch entsprechendes Umdefinieren auf die regionale Betrachtungsebene und eine geringere Betonung der technischen Innovationen relativ leicht umgedeutet werden, um Politiken wie die ILE zu beschreiben. Die Funktionen sind untereinander nicht unabhängig, sondern können selbstverstärkende, aber auch selbstzerstörende Wirkungskreise bilden.

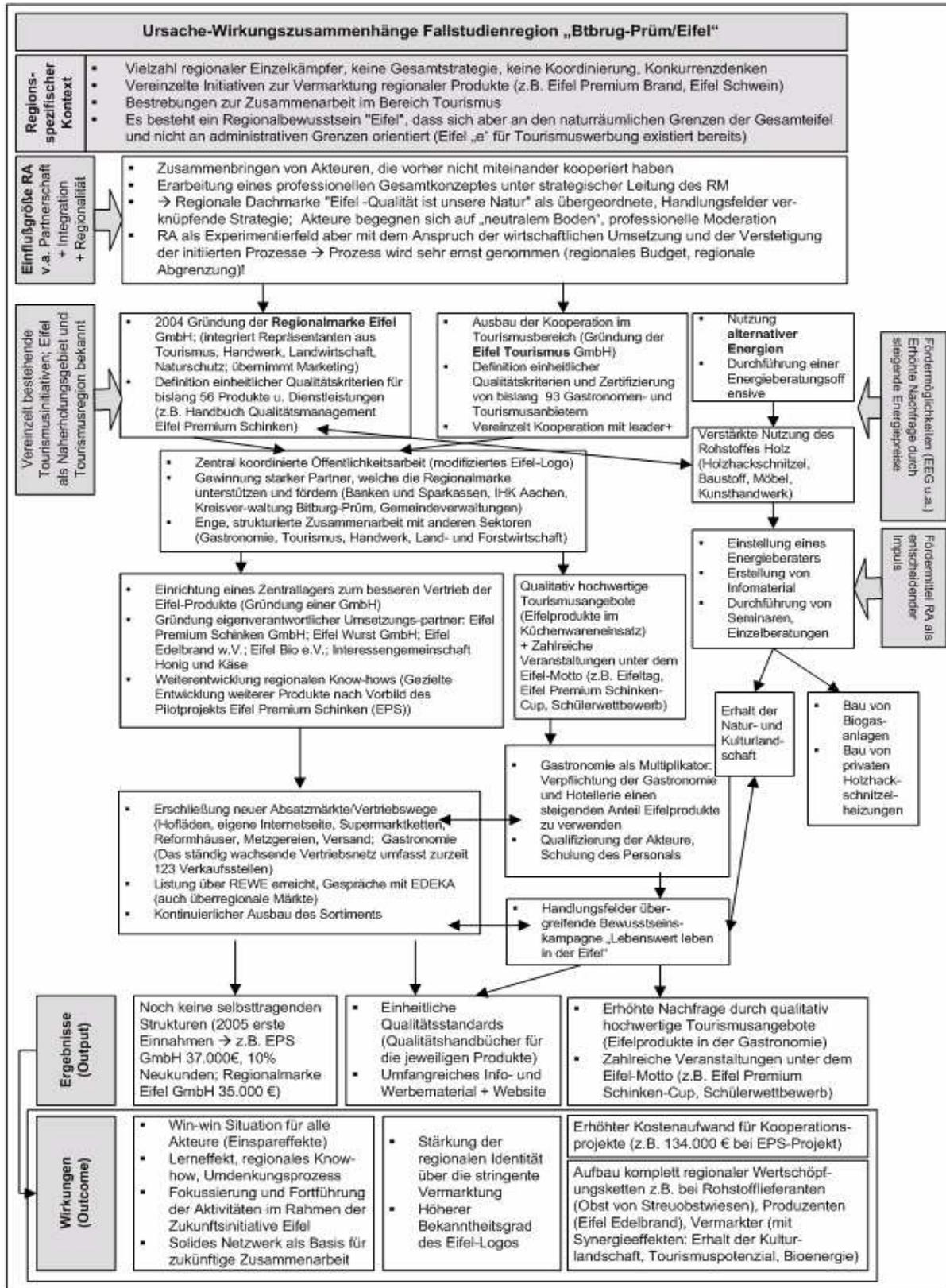
Um herauszufinden, welche Zusammenhänge in einem bestimmten Fall bestehen nutzen die Autoren die Prozessanalyse. Diese Analyse kann sogar in Echtzeit, also prozessbegleitend, durchgeführt werden. Es werden alle Ereignisse in Verbindung mit dem Projekt chronologisch festgehalten und einer der sieben Funktionen zugeordnet. Dann werden alle Ereignisse als positiv oder negativ kategorisiert. Danach können alle Ereignisse einer Funktion bildlich dargestellt werden (s. Abbildungen in NEGRO, 2007, Kapitel 5). Aus den sieben Abbildungen zu den sieben Funktionen ergibt sich ein schneller erster Eindruck zum Verlauf des Prozesses. Aus allem ergibt sich die Geschichte des Projektes. Im Vergleich der Ergebnisse verschiedener Studien kann dann nach verallgemeinerbaren Mustern gesucht werden (HEKKERT et al., 2007).

### **Umfassende Fallstudie**

In den bisherigen (Fall-)Studien wurde der Fokus der Untersuchung bewusst eng gefasst (Netzwerkanalyse) und alle weiteren Wirkungen blieben außen vor, oder der Untersuchungsgegenstand wurde stark schematisiert und damit auch aggregiert, um eine gewisse Vergleichbarkeit und letztlich Verallgemeinerbarkeit zu erreichen. Das ist in der letzten darzustellenden Herangehensweise nicht so. Hierbei handelt es sich tatsächlich um die Evaluation bestimmter Interventionen unter spezifischen Umständen ohne den Anspruch auf Verallgemeinerung. Zusammenhänge werden in großer Detailliertheit erfasst und in ihrer Relevanz bewertet, allerdings meist nicht quantifiziert. Das Beispiel für eine solche Fallstudie stammt aus einer Dissertation zum Regionen Aktiv Programm (WEIB, 2007).

Die Autorin beruft sich auf eine Methodik mit der Bezeichnung FACT-Analyse (Focused Assessment through Cause-effect Tracing). Im Prinzip handelt es sich dabei um das klassische Fallstudiendesign, in dem „qualitative und quantitative Daten verwendet werden, auf induktive und deduktive Vorgehensweisen (Aufstellung hypothetischer Wirkungsketten, Überprüfung und Vernetzung von Wirkungsketten) zurückgegriffen wird und zunächst breit erhobene Aspekte [...] sehr detailliert im Rahmen von Fallstudien analysiert und hinterfragt werden. Durch den Einsatz verschiedener Rückkopplungsschleifen und die Betrachtung eines Sachverhaltes aus verschiedenen Blickrichtungen liefert die Analyse fundierte Ergebnisse, die als Basis für die anschließende Bewertung dienen“ (WEIB, 2007, S. 68). Den Ansatz zeichnet insbesondere aus, dass die Wirkungsketten detailliert verfolgt werden. Das soll sicher stellen, dass zwischen tatsächlichen Wirkungen der Maßnahme und den Auswirkungen anderer Umstände klar unterschieden wird. Die Informationsquellen sind vielfältig und unterliegen keiner weiteren Qualifizierung. Die Ergebnisse für die Fallstudienregionen können in Flussdiagrammen zusammengefasst werden (Abbildung 5.2). In der Abbildung berücksichtigt sind nur die Wirkungen, die aufgrund der Ursache-Wirkungskette eindeutig der Förderung zugeschrieben werden konnten (WEIB, 2007).

Abbildung 5.2: Beispielhafte Ursache-Wirkungszusammenhänge im Fallstudienesign



Quelle: Weiß (2007).

### 5.3 Kritische Reflexion der Eignung der interventionslogikbasierten Evaluation

Die Darlegung einer an der Interventionslogik orientierten Erfolgs- (und natürlich auch Misserfolgs-)geschichte hat mit Recht sehr viel mehr Überzeugungskraft als auf einer stark aggregierten Ebene nachgewiesene „Netto“-effekte, deren zugehöriger Wirkungsmechanismus im Dunklen bleibt. Die entsprechenden Studien untersuchen die von den Intervenierenden unterstellten Wirkungszusammenhänge direkt. Dadurch ergeben sich wertvolle Informationen für die Politik, die über die reine Berechnung eines Wirkungsindikators weit hinaus gehen. Im Falle positiver Ergebnisse sollten diese Zusammenhänge eine glaubhafte Verteidigung entsprechender Politiken ermöglichen. Das Gleiche gilt andersherum für den Fall negativer Ergebnisse.

Unumgänglicher Schwachpunkt aller Fallstudien bleibt ihre fehlende Repräsentativität. Die gute oder die fehlende Wirkung kann immer auf möglicherweise unbeobachtbare Sondereinflüsse, wie zum Beispiel herausragende Begabungen oder intrinsische Motivation zentraler Akteure, zurückzuführen sein. Dieses Problem trifft den zuletzt behandelten allgemeinen Fallstudienansatz besonders hart. Aufgrund fehlender Abstraktion fehlt letztlich die Möglichkeit eines umfassenden Vergleichs. Dadurch aber bleibt offen, ob es bestimmte Entwicklungsmuster gibt, die mit der Förderung in engem Zusammenhang stehen. Solche wiederholten Muster aber machen die Zuschreibung von Entwicklungen zu Förderungen glaubhafter. Muster finden sich in den Graphen und Matrizen der Netzwerkanalyse und in den Verlaufsskizzen der Funktionalen Prozessanalyse. Die einfache Fallstudie bleibt letztlich immer auf der Ebene der Einzelfallbetrachtungen und der entsprechenden Angreifbarkeit stehen.

Von der Ermittlung der Nettoarbeitsplatzeffekte sind alle beschriebenen interventionslogikbasierten Methoden weit entfernt. Das wiederum ist der Spezifität der Interventionslogik (der zugrundeliegenden Theorie) geschuldet, die einen umfassenden funktionalen Ursache-Wirkungszusammenhang über mehrere Stufen hinweg nicht unterstützt. Wie bereits mehrfach angedeutet kann das auch nicht Ziel einer „experimentellen“ Politik wie der ILE sein.

## 5.4 Tabellarische Bewertung der interventionslogikbasierten Evaluation

	1	2	3	4	5	6	7
<b>Technische Aspekte</b>	sehr gering			sehr hoch			
Zeitbedarf						■	
Anforderungen an technisches Know-How				■			
Anforderungen an Kontextwissen/Theoretisches Wissen						■	
Softwarebedarf	■						
Datenanforderungen						■	
Sonstige Kosten							■
<b>Inhaltliche Aspekte</b>	sehr gut möglich			nicht möglich			
Erfassung direkter Effekte	■						
Erfassung indirekter Effekte			■				
Direkte Modellierung der Interventionslogik	■						
Erfassung diskontinuierlicher, auch zeitverzögerter Effekte			■				
Erfassung positiver externer Effekte			■				
Erfassung von Allokationseffekten				■			
Verallgemeinerbarkeit							■
Beachtung des regionalen Kontextes/weiterer Wirkungsbedingungen		■					
	Pro			Contra			



## 6 Zusammenfassende Würdigung und Schlussfolgerungen

Es gibt eine Vielzahl anderer Herangehensweisen an die Bewertung von Politikeffekten. Viele stellen Erweiterungen der Methoden des vergleichenden Ansatzes dar, die versuchen, deren Schwächen abzumildern. Dazu gehört die Wahl dynamischer Modelle oder die Beachtung räumlicher Korrelationen in den Fixed-Effects-Modellen. Die Anforderungen an das technische Wissen der Bewerter steigen mit jeder solcher Komplikation. Darüber hinaus steigt die Schwierigkeit in der Interpretation der ermittelten Zusammenhänge. In aller Regel steigt, wie die differenzierte Unterscheidung in der Matching-Studie nach direkten, Verdrängungs- und Spillover-Effekten deutlich macht, mit der weitergehenden Differenzierung des empirischen Modells meist auch der Datenbedarf. In der angewandten Forschung, die auf angemessenen Ressourceneinsatz achten muss, ergeben sich hier schnell Grenzen des Machbaren. Mit Blick auf die endogenen Entwicklungsprozesse, die die ILE anstoßen soll, werden auch weiter trotz aller technischen Raffinessen begründete Zweifel an der Angemessenheit der noch immer getroffenen Annahmen der vergleichenden Ansätze bestehen bleiben.

In Bezug auf die Fallstudien gibt es neben den Vorgestellten eine Vielzahl weiterer netzwerkanalytischer und systemtheoretischer Ansätze, die zum Teil auch kommerziell als Bewertungstools vermarktet wurden (z. B. der Vester-Ansatz unter <http://www.fredericvester.de/deu/sensitivitaetsmodell/>). Diese Ansätze stehen und fallen oft mit der verfügbaren Datenbasis bzw. den Erwartungen, die an die Informationen der involvierten Agenten gestellt werden. Mit Blick auf die ILE sollte ein Schwerpunkt einer Politikbewertung auf der Wirkung unterschiedlicher Interventionsdesigns liegen. Das geschieht in sogenannten Implementationsstudien (KONING und MOSLEY, 2001), auf die hier nicht weiter eingegangen wird. In solchen Studien geht es um eine empirische Unterstützung der Entscheidung für bestimmte Projektorganisationen.

Erst in letzter Zeit beginnt der tiefe Graben, der die Verfechter „harter“ quantitativer Methoden einerseits und die Anhänger der qualitativen Sozialforschung andererseits lange Zeit getrennt hat, an der einen oder anderen Stelle durch Pioniere überbrückt zu werden. In diesen mehrdimensionalen Arbeiten werden entweder die Ergebnisse quantitativen Studien im Nachhinein zum Beispiel durch Expertengespräche validiert, oder aber die Restriktionen und Annahmen von Erklärungsmodellen werden durch explorative Studien im Vorfeld gefestigt. Ein solcher „Methodenmix“ erscheint in Bereichen sinnvoll, in denen es keine gefestigten theoretischen Grundlagen für die Politikbewertung gibt. Allerdings ist es ein Irrtum zu glauben, dass man mit der Kombination verschiedener Methoden der Wahrheit in jedem Fall näher kommt, auch wenn der korrespondierende Begriff aus der Evaluationsforschung, „Triangulation“ (DENZIN, 1989), dies nahe legt. Einerseits bleibt die Gültigkeit der zugrundeliegenden Annahmen für den vergleichenden Ansatz immer zentral, auch wenn ihre Diskussion auf den ersten Blick wie eine akademische Fingerübung wirken mag. Gerade in Regressionen kann etwa die Beachtung eines einzi-

gen vorher ausgelassenen konditionalen Effektes zu einem Vorzeichenwechsel der Politikvariable führen – und zwar unabhängig davon, wie technisch ambitioniert das Schätzmodell schon vorher war.

Andererseits können auch qualitative Herangehensweisen und Befragungen irreführend sein. Dabei ist es seltener der Fall, dass bewusst aus einer bestimmten Interessenlage heraus falsch geantwortet wird. Vielmehr ist es auch den Befragten oft nicht möglich, sich in eine Situation ohne Förderung wirklich hineinzusetzen. Darüber hinaus gibt es das Phänomen, dass Menschen sich aus der Rückwärtsbetrachtung, wiederum unbewusst, Geschichten zu deren Erklärung zurechtlegen, die dann auch höchst plausibel erscheinen. Weiterhin gibt es das Problem der endogenen Präferenzen: Wenn es Personen nicht gelingt, die Umgebung ihren Wünschen anzupassen, passen sie ihre Wünsche der Umgebung an.

Da der Selektionseffekt der quantitativen Methoden und die Subjektivitätseffekte der qualitativen Methoden in die gleiche Richtung wirken können, ist auch ein Methodenmix keine Garantie für ein ausgewogenes Urteil.

Aus all den geschilderten Gründen gibt es auch für andere Regionalpolitiken keine unumstrittenen Methoden zur Ermittlung ihrer Wirkungen. Diese Situation spiegelt die Theorienvielfalt zur Regionalentwicklung wieder. Sie ergibt sich vor allem aus der Tatsache, dass es kaum Konstanten in den sozioökonomischen Wissenschaften gibt. Mit Blick auf die ILE sollte anerkannt werden, dass sie einer neuen Form von Politik angehört, die die andere Seite der Medaille jüngerer Entwicklungen in der Regionalökonomie ist. Bei den zugehörigen Interventionen handelt es sich daher immer ein wenig um „experimentelle“ Politik. Die Bewertung ist in diesem Fall keine Einbahnstraße: Die Wissenschaft informiert nicht einseitig die Politik, vielmehr informiert auch die Politik die Wissenschaft. Insbesondere die Politik muss um voranzukommen in einen iterativen Wissensaustausch eintreten. Dabei ist zumindest in diesem Feld die Suche nach *der* Evaluationsmethode ein zum Scheitern verurteilter Ansatz.

Nach eingehender Diskussion aller fundamentalen Möglichkeiten der Messung von Politikeffekten kann man sich sicher der folgenden Aussage anschließen: “To begin with, our starting point is that there is no single evaluation method amongst all the possible options that can serve as a methodological recipe applicable to each and every evaluation of these new regional policies. Therefore, there does not exist an ideal methodological design. Such a statement means implicitly that each situation requires a unique specific evaluation design. As some evaluators maintain, evaluation is more an art than a science: ‘The art of evaluation involves creating a design that is appropriate for a specific situation and particular action or policy-making context’ (PATTON, 1990)” (DIEZ und ESTEBAN, 2000).

Mit Blick auf die Bewertung der ILE muss Folgendes festgehalten werden:

- (1) Eine Ermittlung der Netto-Arbeitsplatzeffekte entsprechender Maßnahmen ist heute noch nicht möglich.
- (2) Dennoch kann durch eine wissenschaftliche Politikbegleitung und durch weitere Grundlagenforschung zur Entwicklung der Wirtschaft im ländlichen Raum auch heute schon zur laufenden Politikoptimierung beigetragen werden. Wissenschaft und Politik könnten dabei wechselseitig voneinander profitieren.
- (3) Anhand von vertieften Fallstudien lässt sich für den Einzelfall die Wirksamkeit der Maßnahme entlang ihrer Interventionslogik qualitativ überprüfen. Zwar lassen sich die Ergebnisse nicht verallgemeinern, doch sind entsprechende qualitativ hochwertige Analysen für den Rezipienten meist nachvollziehbar und überzeugend.
- (4) Darüber hinaus wird durch grundlegende Studien zur Initialisierung von Netzwerken, zur Bedeutung von Sozialkapital für wirtschaftliche Entwicklung und zum Problem ungenutzter Ressourcen im ländlichen Raum heute die Grundlage für eine theoriebasierte Politikplanung und -bewertung in der Zukunft gelegt.

Empfohlen wird daher die wissenschaftliche Begleitforschung zur Integrierten Ländlichen Entwicklung aufrecht zu erhalten und ggf. methodisch zu optimieren. Der Fokus sollte weniger auf die Erzeugung von bestimmten Wirkungsindikatoren als auf die Beleuchtung von Ursache-Wirkungszusammenhängen gelegt werden. Darüber hinaus wird noch mal unterstrichen, dass erst die Grundlagenforschung zur Entwicklung ländlicher Räume die Basis für mögliche spätere quantitative Wirkungsanalysen legen wird.



## Literaturverzeichnis

- ARPINO B, MEALLI F (2008) The specification of the propensity score in multilevel observational studies. Dondena Working Paper No. 6, Carlo F. Dondena Centre for Research on Social Dynamics, Milan
- ASMACHER C (1989) Regionale Strukturpolitik in der Bundesrepublik Deutschland: Wirkungsweise und zielkonforme Gestaltung. In: Beiträge zum Siedlungs- und Wohnungswesen und zur Raumplanung, Band 129. Institut für Siedlungs- und Wohnungswesen, Münster
- BERGSCHMIDT A, DIRKSMEYER W, EBERS H, FITSCHEN-LISCHEWSKI A, FORSTNER B, MARGARIAN A, HEUER J (2008) Ex-post-Bewertung des Agrarinvestitionsförderungsprogramms (AFP) im Förderzeitraum 2000 bis 2006. Materialband. Johann Heinrich von Thünen-Institut, Institut für Betriebswirtschaft, Braunschweig. HTML: <http://www.vti.bund.de/de/institute/bw/publikationen/>
- BEHRMAN JR, TODD PE (1999) Randomness in the experimental samples of PROGRESA (education, health, and nutrition program). International Food Policy Research Institute, Washington, D.C.  
[http://www.ifpri.org/sites/default/files/publications/behrmantodd\\_random.pdf](http://www.ifpri.org/sites/default/files/publications/behrmantodd_random.pdf)
- BIRCKMAYER JD, WEISS CH (2000) Theory-Based Evaluation in Practice: What Do We Learn? Evaluation Review 24, S. 407-431
- BLOOM HS (2006) The Core Analytics of Randomized Experiments for Social Research. In: ALASUUTARI P, BICKMAN L, BRANNEN J (eds.): The SAGE Handbook of Social Research Methods. Sage Publications: Thousand Oaks
- BÖCHER M, TRÄNKNER S (2008) Erfolgsfaktoren integrierter ländlicher Entwicklung. In: BÖCHER M, KROTT M, TRÄNKNER S (Hrsg.) Regional Governance und integrierte ländliche Entwicklung: Ergebnisse der Begleitforschung zum Modell- und Demonstrationsvorhaben „Regionen Aktiv“. VS Verlag für Sozialwissenschaften: Wiesbaden.
- BORUCH RF (1997) Randomized Experiments for Planning and Evaluation. Sage Publications: Thousand Oaks
- BUCKLEY J, SHANG Y. (2003) Estimating policy and program effects with observational data: the “differences-in-differences” estimator. In: Practical Assessment, Research & Evaluation, 8(24). Retrieved April 13, 2009 from <http://PAREonline.net/getvn.asp?v=8&n=24>
- BUSCH P (2005) Tragfähigkeit, Tragsicherheit und Tragreserven von Bogenbrücken. 5. Dresdner Brückenbausymposium. HTML: [http://www.tu-dresden.de/biwitb/mbau/download/05bbs\\_11.pdf](http://www.tu-dresden.de/biwitb/mbau/download/05bbs_11.pdf)
- CAPELLO R (2007) Regional Economics. New York: Routledge

- CASPARI A, BARBU R (2008) Wirkungsevaluierungen: Zum Stand der internationalen Diskussion und dessen Relevanz für Evaluierungen der deutschen Entwicklungszusammenarbeit. Evaluation Working Papers. Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung
- CONWAY RS (1977) The Stability of Regional Input-Output Multipliers. *Environment & Planning A*, vol. 9, pp. 197-214
- DENZIN NK (1989) *The Research Act: A Theoretical Introduction to Sociological Methods*. Englewood Cliffs, NJ: Prentice Hall
- DIEZ MA, ESTEBAN MS (2000) The evaluation of regional innovation and cluster policies: looking for new approaches. Paper presented to the session: Decentralisation and Evaluation. Fourth EES Conference. Lausanne. October 12-14, 2000
- EICHLER M, LECHNER M (2002) An Evaluation of Public Employment Programmes in the East German State of Sachsen-Anhalt. *Labour Economics* 9 (2), S. 143-186
- EUROPÄISCHE KOMMISSION (2000) Communication from the Commission: Community policies in support of employment. Com (2000) 78 final
- FRITZ O, STREICHER G, ZAKARIAS G (2005) MultiREG – Ein multiregionales, multisektorales Prognose- und Analysemodell für Österreich. WIFO Monatsbericht 08/2005, Wien
- HARTWICH F, ALEXAKI A, BAPTISTA R (2007) Innovation Systems Governance in Bolivia: Lessons for Agricultural Innovation Policies. IFPRI Discussion Paper 00732. International Food Policy Research Institute, Washington
- HEDTKE R (1973) Möglichkeiten und Grenzen einer Anwendung der Kosten-Ertrags-Analyse in der Strukturpolitik. Berlin: Duncker & Humblot
- HEKKERT MP, SUURS RAA, NEGRO SO, KUHLMANN S, SMITS REHM (2007) Functions of innovation systems: A new approach for analysing technological change. *Technological Forecasting & Social Change* 74, S. 413-432
- HECKMAN JJ, ICHIMURA H, TODD P (1997) Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme, *Review of Economic Studies*, 64(4), S. 605-654
- HENNING CHCA, MICHALEK J (2008) Ökonometrische Methoden der Politikevaluation: Meilenstein für eine sinnvolle Agrarpolitik der 2. Säule oder akademische Fingerübung? *Agrarwirtschaft* 57 (3/4), S. 232-243
- HENNING CH, SAGGAU V (2009) Information networks and knowledge spillovers: simulations in an agent-based model framework. In: SALVADORI N (Hrsg.). *Institutional and Social Dynamics of Growth and Distribution*. Edward Elgar: Cheltenham, UK, Northampton, MA, USA. S. 253-289

- KARL H, KRÄMER-EIS H (1969) Entwicklung der regionalen Wirtschaftspolitik in Deutschland, in: VON EBERSTEIN HH, KARL H (Hrsg.), Handbuch der regionalen Wirtschaftsförderung, 4. Aufl., Köln, Beitrag A II
- KOESTER U (2001) Europäische Agrarpolitik: Ein Spannungsfeld divergierender Interessen. In: OHR R, THEURL T (Hrsg.): Kompendium Europäische Wirtschaftspolitik. Verlag Vahlen, München, S. 309-362.
- KÖTTER T, THIEMANN KH (2006) Der Beitrag Integrierter Ländlicher Entwicklung zur Schaffung von Arbeitsplätzen unter besonderer Berücksichtigung der demographischen Entwicklung in Deutschland. Ländlicher Raum auf Roter Liste, SH 1
- KONING J DE, MOSLEY H (Hrsg.) (2001) Labour Market Policy and Unemployment. Impact and Process Evaluation in Selected European Countries. Edward Elgar: Cheltenham, UK
- KOSTOV P, LINGARD J (2002) Integrated rural development - do we need a new approach? In: ARZENI A, ESPOSTI R, SOTTE F (Hrsg.). European Policy Experience With Rural Development, Kiel: Wissenschaftsverlag Vauk KG.
- LAMBOOY JG, BOSCHMA RA (2001) Evolutionary economics and regional policy. In: Annals of Regional Science, 35, 113-131
- LAMMERS K, NIEBUHR A (2002) Erfolgskontrolle in der deutschen Regionalpolitik: Überblick und Bewertung. HWWA-Report 214. Hamburgisches Welt-Wirtschafts-Archiv (HWWA)
- LECHNER M (1999) Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany after Unification. Journal of Business & Economic Statistics 17
- MARGARIAN A (2008) Diskussionsbeitrag zu: CHRISTIAN
- HCA, MICHALEK H und J: Ökonometrische Methoden der Politikevaluation : Meilenstein für eine sinnvolle Agrarpolitik der 2. Säule oder akademische Fingerübung? In: Agrarwirtschaft 57 (7), S. 346-347
- MAYERHOFER P, FRITZ O (2009) Erste Analyse der Wirkungen der EU-Regionalpolitik in Österreich. WIFO Monatsberichte 6/2009, Wien
- MAYERHOFER P, FRITZ O, HIERLÄNDER R, STREICHER G (2008) Quantitative Effekte der EU-Regionalförderung in Österreich. Eine Pilotstudie. Österreichisches Institut für Wirtschaftsförderung, Wien
- MOSLEY H, SOL E (2001) Process evaluation of active labour market policies and trends in implementation regimes. In: KONING J DE, MOSLEY H (Hrsg.) Labour Market Policy and Unemployment. Impact and Process Evaluation in Selected European Countries. Edward Elgar: Northampton, MA, S. 163-178
- NEGRO SO (2007) Dynamics of Technological Innovation Systems: The case of biomass energy. Doctoral thesis Utrecht University.  
URL: <http://igitur-archive.library.uu.nl/dissertations/2007-0219-200257/>

- NEMES G (2005) Integrated Rural Development: The Concept and its Operation. Discussion Papers MT–DP. 2005/6. Institute of Economics, Hungarian Academy of Sciences, Budapest
- NONIE, Subgroup 1 (2007) NONIE: Impact Evaluation Guidance Section 2: Experimental and Quasi-Experimental Approaches to Impact Evaluation.  
URL: <http://www.worldbank.org/ieg/nonie/docs/subgroup1.doc>
- OTT I, SORETZ S (2006): Infrastruktur als Investitionsdeterminante von KMU. Universität Hannover, Diskussionspapier Nr. 329
- PATTON MQ (1990) Qualitative Evaluation and Research Methods. SAGE Publications.
- PETRICK M, ZIER P (2009) Employment impacts of the Common Agricultural Policy in Eastern Germany – A regional panel data approach. Contributed Paper prepared for presentation at the International Association of Agricultural Economists Conference, Beijing, China, August 16-22, 2009
- PSALTOPOULOS D, THOMSON KJ, EFSTRATOGLOU S, KOLA J, DAOULI A (2004) Regional social accounting matrices for structural policy analysis in lagging EU rural regions. *European Review of Agricultural Economics* 31(2), S. 149-178
- ROSENBAUM PR, RUBIN DB (1983) The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70, S. 41–55
- SALINAS-JIMENEZ MM (2003) Technological change, efficiency gains and capital accumulation in labour productivity growth and convergence: an application to the Spanish regions, in: *Applied Economics*, 35, 1839-1851
- SCHALK HJ, UNTIEDT G (2000) Regional Investment Incentives in Germany: Impacts on Factor Demand and Growth. *Annals of Regional Science*, 34, S. 173-195
- SKOUFIAS E (2005) PROGRESA and Its Impacts on the Welfare of Rural Households in Mexico. Research Report 139, International Food Policy Research Institute. Washington, D.C.
- TEMEL T, JANSSEN W, KARIMOV F (2003) Systems analysis by graph theoretical techniques: assessment of the agricultural innovation system of Azerbaijan. *Agricultural Systems* 77 (2), S. 91-116
- TÖDTLING F, TRIPPL M (2005) One size fits all? Towards a differentiated regional innovation policy approach. *Research Policy* 34, S. 1203-1219
- VOGEL SJ (1994) Structural Changes in Agriculture: Production Linkages and Agricultural Demand-led Industrialisation. *Oxford Economic Papers* 46, S. 136-156
- WEIB K (2007) Der Mehrwert des regionalen Wettbewerbsansatzes gegenüber dem herkömmlichen Förderansatz – Das Beispiel Regionen Aktiv. Promotionsvorhaben zur Erlangung des Doktorgrades (Dr. rer. pol.). Fakultät Raumplanung, Universität Dortmund.

- 
- WOLF K (2002) Analyse regionaler Beschäftigungsentwicklung mit einem ökonometrischen Analogon zu Shift-Share-Techniken. In: KLEINHENZ G (Hrsg.): IAB-Kompendium Arbeitsmarkt- und Berufsforschung. Beiträge zur Arbeitsmarkt- und Berufsforschung, BeitrAB 250, S. 325-333